Essays in Public Finance and Information Economics

By

André Medeiros Sztutman

Submitted to the Department of Economics
in partial fulfillment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY IN ECONOMICS

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2023

Authored by: André Medeiros Sztutman
Department of Economics
May 12, 2023

Certified by: Robert M. Townsend
Elizabeth and James Killian Professor of Economics
Thesis supervisor

Certified by: Iván Werning
Robert M. Solow Professor of Economics
Thesis supervisor

Accepted by: Abhijit Banerjee
Ford International Professor of Economics
Chairman, Departmental Committee on Graduate Studies

Essays in Public Finance and Information Economics

By

Andre Medeiros Sztutman

Submitted to the Department of Economics on May 12, 2023
in partial fulfillment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY IN ECONOMICS

ABSTRACT

This thesis comprises three chapters on public finance and information economics. The first focuses on the interaction of imperfect information in labor markets and the tradeoffs the government faces when setting non-linear taxes. The second focuses on the role of heterogeneity in elasticities in affecting those tradeoffs. The third, focuses on the imperfect information in financial markets, and on how to design disclosure rules to increase the size of gains from trade in lending markets when these markets are adversely selected.

The first chapter asks how optimal taxes are affected by reputation building and imperfect information in labor markets. To answer that question, I build a model of labor markets with incomplete and asymmetric information where job histories play a crucial role in transmitting information about workers' productivity, which allows us to better understand the efficiency and distributive consequences of imperfect monitoring and screening in labor markets, and the tradeoffs the government faces when setting taxes. Optimal taxes are described by generalized versions of standard redistributive and corrective taxation formulas, which depend crucially on labor wedges: the marginal contribution to output relative to the increases in lifetime earnings that result from supplying one extra unit of labor at each period. Using data from the Health and Retirement Study, I find that the corrective component of taxes is likely to be large, especially at the top of the income distribution.

The second chapter (joint with John Sturm) asks how income taxes should account for heterogeneity in elasticities of taxable income. We address this question with a test that passes if and only if there exists a weighted utilitarian planner for whom taxes are locally optimal. Our test incorporates standard sufficient statistics and a novel ingredient: the variance of elasticities conditional on income. Theoretically, we show that the test fails when these variances are sufficiently high. Empirically, we find they are indeed large in a panel of US tax returns. We thereby conclude, without taking a stance on redistributive preferences, that there are welfare-improving tax reforms.

The increasing availability of data in credit markets may appear to make adverse selection concerns less relevant. However, when there is adverse selection, more information does not necessarily increase welfare. The third chapter (joint with Robert M. Townsend and Nicole Immorlica) provides tools for making better use of the data that is collected from potential borrowers, formulating and solving the optimal disclosure problem of an intermediary with commitment that seeks to maximize the probability of successful transactions, weighted by the size of the gains of these transactions. We show that any optimal disclosure policy needs to satisfy some simple conditions in terms of local sufficient statistics. These conditions relate prices to the price elasticities of the expected value of the loans for the investors. Empirically, we apply our method to data from the Townsend Thai Project – a long panel dataset with rich information on credit histories, balance sheets, and income statements – to evaluate whether it can help develop rural credit markets in Thailand, finding economically meaningful gains from adopting limited information disclosure policies.

JEL Classification: H2, D8, J2, I3, G2.

Thesis supervisor:	Robert M. Townsend
Title:	Elizabeth and James Killian Professor of Economics
Thesis supervisor:	Iván Werning
Title:	Robert M. Solow Professor of Economics

Acknowledgements

I am honored to have this opportunity to express my sincere gratitude and appreciation to the many individuals who have supported me throughout my Ph.D. journey.

First and foremost, I would like to express my heartfelt gratitude to my esteemed advisors, Rob Townsend and Iván Werning, for their invaluable guidance, unwavering support, and insightful feedback throughout my Ph.D. program. Their expertise, dedication, and encouragement have been instrumental in shaping my research and academic development.

I also want to express my gratitude to Jon Gruber and Jim Poterba for their valuable feedback and guidance in my academic pursuits in the field of public finance. I am also grateful to Nicole Immorlica for her support in explorations on the intersection between economics and computer science.

My gratitude extends to the brilliant professors I had the privilege of interacting with at MIT. In particular, I would like to thank Alberto Abadie, Daron Acemoglu, Arnaud Costinot, Dave Donaldson, and Amy Finkelstein for their constructive feedback and valuable insights throughout my Ph.D. program.

I would like to thank the many people with whom I have had the privilege of working, especially my dear friend John Sturm, and also Benjamin Grossman, Yingju Ma, Andre Mendes, Andre Schlieker, and Bryant Xia. To all my colleagues and dear friends, including Oğuzhan Çelebi, Alden Chen, Jonathan Cohen, Brandon Enriquez, Joel Flynn, Adam Harris, Giacomo Lanzani, Anh Nguyen, Kramer Quist, Garima Sharma, Rahul Singh, Diana Sverdlin Lisker, as well as Lucas Finamor, Ana Carolina Diniz, Rodolfo Rigato, and Camila Caron Rigato, thank you for your unwavering support, motivation, and encouragement throughout this journey.

I would also like to thank Grace Koo, Clyo Castle, Fundação Estudar, and Accenture Industry Convergence Initiative for their financial support and mentoring programs.

I would like to extend my thanks to the MIT economics department staff, including Deborah Jamiol, Kim Scantlebury, Gary King, Julia Martyn-Shah, and Shannon May; the MIT Sailing Pavilion staff; and the MIT Medical staff, in special Maureen Rezendes.

To my beloved family, Haroldo and Isabel, and my brother Felipe, thank you for your endless love, support, and encouragement that have been an invaluable source of strength for me throughout my academic journey.

Finally, to my wife, Michaela Blanc, thank you for your love, patience, and unwavering support throughout this journey. Your steadfast love and support have been the anchor that has kept me grounded and focused. I am deeply grateful for your love and support.

Table of Contents

List of Figures

List of Tables

Chapter 1

Dynamic Job Market Signaling and Optimal Taxation

1.1 Introduction

It takes time for a highly productive worker to be recognized as such in labor markets. For a worker with little experience, it is hard for employers to infer their abilities from their short resumes. However, as workers advance in their careers, increasing the length of their resumes, they make information about their abilities available to potential employers. In line with this idea, there is evidence that measures of ability that are not observed by firms become increasingly more predictive of workers' salaries as they progress in their careers.[1]

These informational imperfections drive workers to exert effort throughout their careers. In the beginning, foreseeing higher salaries in the future, they work hard even though their current salaries are small. As they accumulate experience, employers become able to recognize who the more productive workers are and start paying them differentially, allowing workers to reap the benefits of their past effort. Those dynamic career concerns may not in general induce workers to exert the efficient amount of effort.[2]

Besides shaping the incentives to exert effort, these informational imperfections affect the income distribution. Whenever employers cannot differentiate among workers of distinct productivities at some point in their careers, those workers may receive the same wages. These informational imperfections, thus, may make the higher productivity workers implicitly subsidize the lower productivity workers. Because these informational asymmetries affect incentives and the distribution of income, it is natural to ask what can be done to counterbalance those effects, and how the presence of these dynamic informational asymmetries affects the efficiency and redistribution tradeoffs the government faces when setting taxes.

This paper proposes a simple model of career concerns, building on an otherwise standard dynamic model of labor supply and demand, that addresses the connection between information transmission in labor markets, taxes, and inequality. The model builds on the signaling logic from Spence (1973), shifting the focus from one-time investment decisions on education to the dynamics of job market experience accumulation and the effort decisions throughout the lifetime of a worker. Concretely, the model is built around two ideas.

[1] For example, Armed Forces Qualification Test (AFQT) exam scores become increasingly more predictive of workers' salaries (Farber and Gibbons, 1996; Altonji and Pierret, 2001; Lange, 2007).

[2] Indeed, Holmström (1999) shows that dynamic career concerns introduce labor supply distortions and may only partially replace the incentives that pay-for-performance contracts could provide for workers to exert effort.

First, it is easy for the firms to see and pay for the execution of clearly specified tasks, or deliverables, but it is much harder for them to assess the individual contribution of each worker to the firm's total output, or the value of the firm. Second, firms assess the productivity of workers by looking at the resumes, which parsimoniously summarize the history of deliverables the workers have produced so far in their careers.[3]

In equilibrium, working hard has dual benefits: it generates larger payments today, and it also establishes workers' reputations, signaling their productivity to employers. The logic behind this is the following: employers do not know the productivities of the workers they are hiring, but they can see how many tasks they have completed. Under the assumption that for those who are productive it is less costly to complete deliverables, the firms can infer that those who have longer resumes are also more productive. This generates a "rat race," and pushes workers to exert additional effort to signal to employers their productivities. While doing that, they build up their resumes, and progressively earn more. There is a positive return to experience even without human capital accumulation, which is entirely due to signaling.

The simple structure of the model will also allow us to derive comparative statics and to incorporate and understand the dual role of nonlinear taxes in correcting for the labor market imperfections and redistributing income. This understanding is brought by a series of results that i) delineate what the optimal tax base is, ii) write optimal tax formulas in terms of simple sufficient statistics that can be estimated and compared to standard optimal tax formulas, which would ignore the role of information imperfections, and iii) derive comparative statics on taxes and welfare from changes in the degree of information asymmetries.

Optimal lifetime income taxes are described by a generalization of standard Mirrleesian nonlinear taxation formulas. This generalization accounts for the labor market imperfections, and introduces a Pigouvian component to the standard formulas. Intuitively, optimal taxes can be thought of in two steps. First, they correct for the informational friction at the margin where it matters by making sure that, for any extra unit of effort, the marginal benefits to workers is equal to their marginal product of labor. Then, simultaneously, redistributive taxes are imposed on top of these taxes following standard Mirrleesian formulas, where the costs and benefits of redistribution are expressed in terms of the shape of the income distribution, compensated and income elasticities, and welfare weights.[4]

Surprisingly, in a simple benchmark case, to correct for the career concerns distortion is not necessary to keep track of the ratio of marginal productivities over salaries throughout the lifetime of the worker.[5] It is enough to look at this ratio only around the time of retirement. Moreover, under the assumption that the willingness to provide deliverables and the unobserved productivities are positively correlated, and holding constant the usual sufficient statistics for optimal taxes (elasticities and the distribution of income), taking career concerns into account pushes towards higher marginal taxes. The reason behind this is the following: employers do not know who are the employees retiring at the time of contracting, so those who are retiring are offered the same wage as those who have the same experience but are more productive and are willing to work even more. By the assumption that the willingness to complete deliverables and the unobserved productivities are positively correlated, those who are retiring are the least productive in the pool of workers with the same resume. This explains the wedge between their salaries and marginal product. This wedge,

[3]The model is a counterpart of the Arrow (1962) learning-by-doing model, where instead of affecting human capital, the history of completion of deliverables affect, in equilibrium, employers' perceptions of the worker productivity. This connection is especially clear in the special case where the resume is defined as the total experience, or the cumulative sum of deliverables a worker has produced. For this reason, we can call it a "signaling-by-doing model."

[4]Moreover, these generalized formulas apply independently of the model of competition with imperfect information or of the source of the friction that makes workers to receive salaries that are not their marginal products, provided some simple conditions are satisfied, pretax earnings are an invertible function of the supply of deliverables, and firms make zero profits.

[5]Nor it is sufficient: another simple benchmark case shows that salaries can be equal to marginal productivities at all points in the career of a worker, but the timing of labor supply decisions can be severely distorted.

more precisely one minus the wedge, can be interpreted as the negative externality of the "rat race". That is, holding the remuneration schedule of firms constant, if a worker would work more and earn one more dollar, they would generate less than one dollar of output for the firm. This difference can be thought of as a negative production externality, and can be corrected with Pigouvian taxes. For these reasons, taking into account these dynamic informational imperfections pushes towards higher marginal taxes over lifetime income.[6]

Recent technological changes are reshaping how workers are monitored and screened.[7] This paper shows that the elimination of informational frictions is a force that: i) pushes lower marginal corrective taxes, ii) pushes towards higher marginal redistributive taxes, iii) has a generally ambiguous net effect on taxes, but in sensible simple cases the corrective component dominates, and iv) it is a force that hurts society from the point of view of a redistributive planner. On one hand, the corrective taxes diminish because, as the strength of the "rat race" diminishes, it takes lower taxes to correct for the fact that workers work too much over their lifetimes. On the other hand, because better information benefits disproportionally more the more productive workers, the shape of the income distribution may change and push towards a higher redistributive component of taxes. Relatedly, whenever redistribution towards the poorer is valued by society, society is hurt when there is less information asymmetry because it becomes easier for high productivity workers to separate themselves from lower productivity workers, undoing part of the implicit cross subsidies that happen in labor markets from those high productivity workers towards the lower productivity workers.

From an empirical point of view, to calibrate the optimal taxation formulas, a key statistic that needs to be estimated is the ratio of the marginal productivity of workers at retirement over their salaries, as a function of their lifetime income. In principle, this labor wedge could be inferred from the shape of the lifetime income distribution and from the signaling component of the return to experience. Although there are estimates for the shape of the income distribution and the growth rate of salaries across the income distribution (for example, Guvenen et al. (2021)), decomposing this growth rate into its signaling and human capital accumulation is a considerable challenge. There are not estimates for each of those in regards to the return to experience, but a growing literature has decomposed the return to schooling into its different components. A simple calibration exercise combining estimates of the growth rate of salaries across the income distribution from Guvenen et al. (2021) and assuming that the signaling share of the return to experience is equal to the signaling share of the return to schooling from Aryal et al. (2019), indicates that the Pigouvian component of taxes could be as high as 25% for top earners, and on average would be around 6%.

To get more direct estimates of the signaling component of the return to experience and the wedge between salaries and marginal productivities at retirement, this paper develops an empirical strategy relying on tax changes as a source of exogenous variation in wages and using data from the Health and Retirement Study survey. This empirical strategy adapts results from the literature that has quantified the degree of adverse selection in insurance markets by leveraging exogenous variation in prices (as in Einav et al. (2010); Einav and Finkelstein (2011); Cabral et al. (2022)). In the context of imperfect information and career

[6]Optimal taxes do not depend only on labor wedges but also on elasticities and the shape of the income distribution. This paper shows that from the point of view of optimal taxation formulas, similarly to results in Scheuer and Werning (2017) common elasticity estimates are biased downwards. The reason for it is that increases in marginal retention induce the marginal types – who are the less productive types – to work more, and therefore reduce pretax salaries, making the effective change in post-tax wages smaller. Thus, the estimated elasticities of taxable income are lower in magnitude than the elasticities that enter optimal taxation formulas, which keep pretax salaries fixed. This is a countervailing force that makes accounting for career concerns push less towards higher marginal taxes.

[7]These changes appear in the form of increasing availability of data and new tools to analyze it (Chalfin et al., 2016; Autor, 2019; Acemoglu et al., 2020; Bales and Stone, 2020), and changes in task composition of jobs from the automation of routine tasks (Autor et al., 2003), and the advent of "new work" (Autor, 2019).

concerns in labor markets, the key idea is that with a source of exogenous variation in wages for a specific labor contract, one can non-parametrically trace the shape of the labor demand curve by looking at average productivities as a function of salaries. However, inferring productivities in labor markets is a difficult endeavor. To circumvent that challenge two complementary approaches are adopted.

The first strategy is to assume that, although there are informational assymetries, labor markets are competitive and, thus, we can treat hourly salaries as the average marginal productivity of workers. Under this assumption by observing individual salary changes before and after a tax change, and observing the number of people who retired in response to the tax change, we can infer the productivity of those who were originally almost indifferent between retiring or not. Applying this strategy to data from the Health and Retirement Study survey shows that, for an average worker, the Pigouvian component of taxes is of the order of 5%, while for high earners it ranges from 10% to as high as 60%. Since optimal taxes can be thought of as the product of a Pigouvian and a Mirrleesian component, the fact that the Pigouvian component is so high implies that the redistributive component of taxes is potentially quite small or close to zero. In other words, the current tax system is significantly less redistributive than otherwise might be thought, if imperfect information was not taken into account.

The second strategy leverages the rich set of questions asked in the Health and Retirement Study and allows for more direct tests for the mechanism highlighted in this paper. In particular, the Health and Retirement Study includes data on cognitive scores, assessed at each interview from questions involving counting, naming, and vocabulary tasks. We find that the pool of individuals who keep working after a tax increase has better cognitive scores (as measured before the tax change), in line with the idea that the changes in pretax salaries induced by tax changes are due to selection, or changes in the composition of the pool of individuals who are still working. This effect is also larger at the top of income distribution, in line with the idea that those informational imperfections are more pronounced for high-paying occupations.

To add realism to an otherwise stylized model, several extensions to the basic model are presented, including on-the-job learning, richer heterogeneity in elasticities of labor supply, and richer signal structures. This paper shows that, in these extensions, the key insights from the generalized optimal taxation formulas and their empirical implications still hold with minor caveats. When there is human capital accumulation on the form of on-the-job learning, the return to experience features both a signaling and human capital accumulation component; however the same optimal lifetime income tax formula applies. When there is rich heterogeneity in elasticities, the average labor wedge at each lifetime income level should be weighted by those lifetime income elasticities. When resumes include richer exogenous signals which cannot be observed by the government, the benefits and costs in optimal tax formulas should be weighted by the sensitivity of post-tax salaries to tax changes. When the signal the firm sees is a richer function of the history of deliverables, there may be additional distortions to be corrected, but the same optimal lifetime income tax formula applies.

Related Literature

This paper is related and contributes to several strands of the literature, including a public finance optimal taxation literature, the empirical and the theoretical literature on imperfect information in labor markets, and an empirical literature on dynamic labor supply choices.

First, the taxation results in this paper build on the optimal taxation literature that goes back to the seminal contributions of Mirrlees (1971); Diamond (1998); Saez (2001), and more precisely contribute to a growing literature on optimal taxation with richer models of labor markets (Hariton and Piaser, 2007;

Rothschild and Scheuer, 2013; Stantcheva, 2014, 2017; Bastani et al., 2015; Ales et al., 2015; Scheuer and Werning, 2016; Ales and Sleet, 2016; Scheuer and Werning, 2017; da Costa and Maestri, 2019; Costinot and Werning, 2018; Craig, 2020; Hummel, 2021; Guerreiro et al., 2022). Methodologically perhaps the closest papers are Scheuer and Werning (2017) and Scheuer and Werning (2016), who show that standard optimal taxation formulas apply quite generally, including a broad range of models where wages are endogenous. Relative to these papers, this paper adds further generality to optimal taxation formulas, enriching them to cover situations where labor market frictions introduce labor market distortions. That is, the generalized taxation formulas in this paper hold when there are additional labor market inefficiencies, independently of their nature. The results on welfare in this paper speak to welfare theorems for economies with informational frictions from Prescott and Townsend (1984) and generalize results from Stantcheva (2014). In the latter, welfare comparisons are drawn between economies with "double adverse selection" – in the form of non-linear screening as in Miyazaki (1977) – to economies where firms know the productivity of workers. This paper extends the comparison to more general and arbitrary frictions, as well as to other intermediary levels of informational frictions.

Second, this paper contributes to the literature on imperfect information in labor markets, which goes back as far as the seminal contributions of Spence (1973), to the models of the "rat race" as in Akerlof (1976) and Miyazaki (1977), and in Stantcheva (2014), and the dynamic career concerns model of Holmström (1999), extended and further analyzed more recently by Bonatti and Hörner (2017), Cisternas (2018), and Hörner and Lambert (2021). Relative to this literature, this paper provides a new model of dynamic signaling combining elements from both of these classes of models. The assumption in this paper, that firms see and pay for the execution of clearly specified tasks, or deliverables, borrows from static competitive screening models of the labor markets as in Miyazaki (1977), and bypasses a key limitation from the canonical career concerns setup in Holmström (1999), where firms can see the individual contribution that each worker makes to the firms' profits but cannot pay for performance. Conversely, the assumption that firms assess the productivity of workers by looking at resumes borrows from Holmström (1999), and bypasses a key limitation from the static setup in Miyazaki (1977), where firms learn the productivity of the workers through a one-time interaction, and resumes play no role in transmitting information. Furthermore, the idea that firms see a simple public signal builds on the motivational rating setup from Hörner and Lambert (2021), and bear resemblance to aggregation and linearity results from Holmstrom and Milgrom (1987).

Third, a related, but more empirically focused literature, has looked at how firms learn about the productivity of workers and whether there are information asymmetries in labor markets, including Jovanovic (1979); Farber and Gibbons (1996); Acemoglu and Pischke (1998, 1999); Altonji and Pierret (2001); Lange (2007); Kahn and Lange (2014); Cella et al. (2017); Aryal et al. (2019). Relative to this literature, this paper provides new evidence for the importance of informational asymmetries in labor markets, in particular for workers in later stages of their careers. This complements the evidence from Kahn and Lange (2014) who found that firms have substantial uncertainty over the productivities of older workers. This paper also contributes to the literature on technological changes and their impacts on labor markets (Autor et al., 2003; Brynjolfsson and Mitchell, 2017; Brynjolfsson et al., 2018; Autor, 2019; Acemoglu et al., 2020; Acemoglu, 2021; Autor et al., 2022), delineating key welfare and normative implications from changes in the technologies for monitoring and screening workers.

There is a large literature that has looked at dynamic labor supply decisions, human capital accumulation and on-the-job learning (Heckman, 1976; Eckstein and Wolpin, 1989; Shaw, 1989; Altuğ and Miller, 1998; Keane and Wolpin, 2001; Imai and Keane, 2004; Keane, 2011; Altonji et al., 2013). One key insight from that literature is that workers, when making their labor supply decisions today, would consider the impact of those decisions today on their future salaries, and that in this context the workers' opportunity cost of time may not be equal to their current wages. The same effect is present in this paper, where the dependence of future salaries on current effort decisions are alternatively explained by signaling effects.

The remainder of the paper is structured as follows. Section 1.2 presents the dynamic signaling model. Section 1.3 presents two simple benchmark cases, and discusses a range of possibilities that the dynamic signaling model can acommodate. Section 1.4 discusses positive properties that hold beyond these benchmark cases. Section 1.5 introduces taxes and presents the main normative results. Section 1.6 discusses the existing empirical evidence, the empirical strategy, and the empirical results. Section 1.7 discusses extensions, including human capital accumulation, heterogeneous elasticities, and richer signal structures. Section 1.8 concludes.

1.2 Signaling-by-doing model

This paper adapts a standard dynamic model of labor supply, demand and taxation, by adding the constraint that firms have limited information over the workers' productivities and can only contract based on a subset of the observed activities workers perform. The general setup encompasses a standard neoclassical life-cycle labor supply problem, and with minor modifications can encompass several models of screening and imperfect information in labor markets as special cases. These special cases include a) models where individual contributions to output are observed and workers get paid a fixed salary independent of the realized output, as in Holmström (1999), as well as b) models where hours are observed but output is not, and workers are screened through the total amount of hours or observable effort they commit to offer, as in Miyazaki (1977) or in Akerlof (1976). For exposition, we focus on one specialized version of the general model that allows to derive clear comparative statics and simple optimal tax formulas, while retaining the essential economic assumptions that describe dynamic job market signaling. That is, employers will see a worker's resume – how much in terms of deliverables a worker has provided so far in their career – and will pay workers for the execution of these deliverables. Workers, aware of how their resumes will be read, will choose their labor supply balancing costs and benefits in terms of current and future wages. Section 1.7 will consider several departures from this simple model, including richer signal structures and richer heterogeneity in preferences.

1.2.1 Preferences and Technology

The household block of the model consists of a continuum overlapping generations of workers who live for a continuum of periods going from zero to one. These workers have arbitrary preferences over labor and consumption flows, and are forward-looking: they understand that their labor supply choices can affect the information firms will have about them in the future. The production block of the model is described by competitive firms with linear production functions.

More formally, workers are indexed by their types θ, which determines their productivity and their preferences, and there are different cohorts of workers. Each worker lives for a continuum of periods that goes from zero to one, where zero corresponds to the time the worker is born and one corresponds to the time their life ends. At each period, workers of all ages coexist. They supply labor ($\tilde{h}(\cdot)$) and consume ($\tilde{c}(\cdot)$) at each period. That is, $\tilde{c}(\cdot)$ denotes the flow of consumption function $\tilde{c} : [0, 1] \mapsto \mathbb{R}^+$, and $\tilde{h}(\cdot)$ denotes the flow of labor supply function $\tilde{h} : [0,1] \mapsto \mathbb{R}^+$. An individual of type θ has a productivity $v(\theta) > 0$, and production is linear, that is the flow of production is equal to the product of productivities and the labor supply $\tilde{h}$. Preferences are denoted $U(\tilde{c}(\cdot), \tilde{h}(\cdot), \theta)$,[8] where $\tilde{c}(\cdot)$ is the time-flow of consumption and $\tilde{h}(\cdot)$ is the time-flow of labor supply.

The worker problem is standard: they maximize a utility functional, subject to a lifetime budget constraint, where flows in the future, at age a, are discounted at the rate $q(a)$.[9] However, salaries w depend on the information the firms have about the worker $I(\tilde{h}(\cdot), a, \theta)$, which will be specified below, but more generally could be a function of the flow of labor supply across all the periods $\tilde{h}(\cdot)$, the type θ and age a of the worker. This captures the possibility that the workers may want to change their labor supply to influence their future salaries, by changing the employers' perceptions of their productivities.

$$V = \max_{\tilde{c}(\cdot), \tilde{h}(\cdot)} U(\tilde{c}(\cdot), \tilde{h}(\cdot), \theta) \ st. \ \int_0^1 q(a)\Big(\tilde{c}(a) - w(I(\tilde{h}(\cdot), a, \theta))\tilde{h}(a) - T(\tilde{y}(\cdot), a)\Big)da \leq 0 \tag{1.1}$$

Notice that a standard life cycle labor supply problem features as a special case of this, where wages are equal to the productivity of the worker, and this productivity would be independent of their labor supply decisions. Notice as well that human capital accumulation of the form of on-the-job learning would generate analogous concerns for the worker, with labor supply decisions affecting future wages through real increases in their productivities.

Workers pay taxes $T(\tilde{y}(\cdot), a)$ on their income flows $\tilde{y}$, which are the product of their wages w and the flow of labor supply $\tilde{h}$. Those taxes can be used to shape incentives and redistribute income. They will be discussed in more detail in Section 1.5.

1.2.2 Contracts and Information

Firms are constrained to offer infinitesimal contracts, which is a contract for one unit of the labor supply. Firms are unable to commit to long-term contracts; they cannot promise that they will not try to renegotiate labor contracts once more information becomes available. This makes the firm problem essentially static. Finally, there is free entry and exit, therefore firms make zero profits.

Firms do not observe types or productivities, but instead, they observe a signal of their past experience. In the simplest case, we can think of that signal as how much labor a worker has supplied so far, which is denoted by $I(\tilde{h}(\cdot), a, \theta) = h(a) = \int_0^a \tilde{h}(\tilde{a})q(\tilde{a})d\tilde{a}$, or the length of the resume when the worker has age a. More generally, as covered in Sections 1.5.2 and 1.7, $I(\tilde{h}(\cdot), a, \theta)$ can accommodate richer signal specifications, including more general functions of past experience $I(\tilde{h}(\cdot), a, \theta) = \int_0^a \phi(\tilde{a}, a)\tilde{h}(\tilde{a})d\tilde{a}$ (with $\phi(\tilde{a}, a) > 0$, continuous in both arguments), or exogenous signals $I(\tilde{h}(\cdot), a, \theta) = (h(a), z(\theta))$. The more general idea is that the resume is an imperfect measure of the past history of the completion of deliverables and their timing.

This detailed history, if it was perfectly observed by firms and if workers' types are single-dimensional, in most sensible cases, would contain enough information to allow firms to infer the workers' type almost instantaneously. While explicitly introducing stochastic noise on preferences or on the information could

[8] Although it is not necessary and it is a especially restrictive in the context of continuous time choices, assuming time-separability of preferences can help us understand the trade-offs workers face when deciding their labor supply flows. For an example with time-separability, see Section 1.3.2.

[9] These discount rates are assumed to be exogenously given, i.e. there is a technology for transferring resources across periods at rates $q(\cdot)$, and that these rates are such that budget and resource constraints are well defined; that is, the present value of resources in the economy is finite. While the first assumption is not essential, the fact that the present value of resources is finite is important to guarantee that the economy is dynamically efficient.

equivalently capture this imperfect inference problem, this formulation avoids dealing with technical difficulties arising from multidimensional screening problems, where in this case both the type space and the space of goods would be high-dimensional. This formulation can also flexibly accommodate other concerns that would lie outside the immediate scope of the model, such as information becoming harder to retrieve for work experiences that are further in the past, or firms assigning disproportional weights to some experiences because of limited attention or because they have the wrong model of how the economy works and mistakenly attribute increases in productivity to human capital accumulation instead of signaling. Special cases of this formulation include observing the total experience $\int_0^a q(\tilde{a})\tilde{h}(\tilde{a})d\tilde{a}$, and the pace of experience accumulation $\int_0^a \frac{\tilde{h}(\tilde{a})}{a}d\tilde{a}$, which are discussed in more detail in Section 1.3.

An important point to emphasize is that we should interpret what has been referred to as labor supply $\tilde{h}(\cdot)$ not necessarily as hours or effort, but as what will be referred to in this paper as deliverables. That is, $\tilde{h}(\cdot)$ represents the specific piece of the information on which the firms can condition their contracts. This will allows us to consider the possibility that changes in technology make these deliverables a better or worse measure of output. In the extreme case where $\tilde{h}(\cdot)$ is the flow of output, there is no information asymmetry problem between the firms and workers, and workers will get paid for their marginal contributions to the value of the firm output.[10] Firms, moreover, do not observe output, or cannot individually assess the contribution of each worker to output, or the firm's profits. They can, however, infer the expected productivity given the signal of experience, either because they have hired many employees with the same experience and have seen how much output these employees have generated on average or because they know the economy-wide distribution of productivities and labor supply.[11]

Figure 1.1 represents diagrammatically the flows of production, payments, and information. Workers complete deliverables for the firms, and the firm sees the completion of the deliverables and the workers' resume. Each worker's resume keeps track of their deliverables completion history, adding those deliverables up. Payments are based on what the firm observes, that is, resumes and the execution of the deliverable. Output and profits are realized, but the firm cannot individually assess the contribution of each unit of deliverable to the firm output.

1.2.3 Taxes

This model features what has been called "double adverse selection" (Stantcheva, 2014). Both firms and the government do not observe workers' types. We assume that firms observe the workers' resumes, while the government observes histories of earnings $\tilde{y}$.[12] This assumption captures the idea that every year, tax payers send their tax returns to the the tax authority, and the tax authority keeps track of these tax returns. The government faces a budget constraint and can save and borrow at the same exogenous discount rates the workers face. Notice that because the workers face a lifetime budget constraint, the timing of tax payments

[10]To match this setup to canonical models of imperfect information in labor markets notice that, in the Miyazaki (1977) model the only difference would be that the information set of firms would consist of not how much a worker has supplied so far $h(a)$, but how much the worker would supply over the lifetime $h(1)$. In the Holmström (1999) model, labor supply $\tilde{h}(a)$ would be a two-dimensional vector of effort and an indicator function for whether the individual decides to work at any period. Workers would get paid only by the second component of $\tilde{h}(a)$. The information set of firms would consist of a stochastic function of the history of effort and the type of the agent.

[11]The fact that firms observe hours but not individual output can also be thought of through the lens of team production (Alchian and Demsetz, 1972). Production needs to take place inside a firm, that aggregates the work of multiple workers. For example, the firm production function could be $\text{F}=\prod_i^n \mathbb{1}(h_i > 0) \cdot \sum_i^n h_i \cdot v_i$, that is, production inside the firm is linear but it needs multiple (n) workers to be present. In the limit of a large number of workers (n), the firm would assess the productivity of workers only from their average productivity.

[12]In Section 1.5, we discuss the extent that this assumption can be relaxed, and conditions under which optimal taxes take even simpler form.

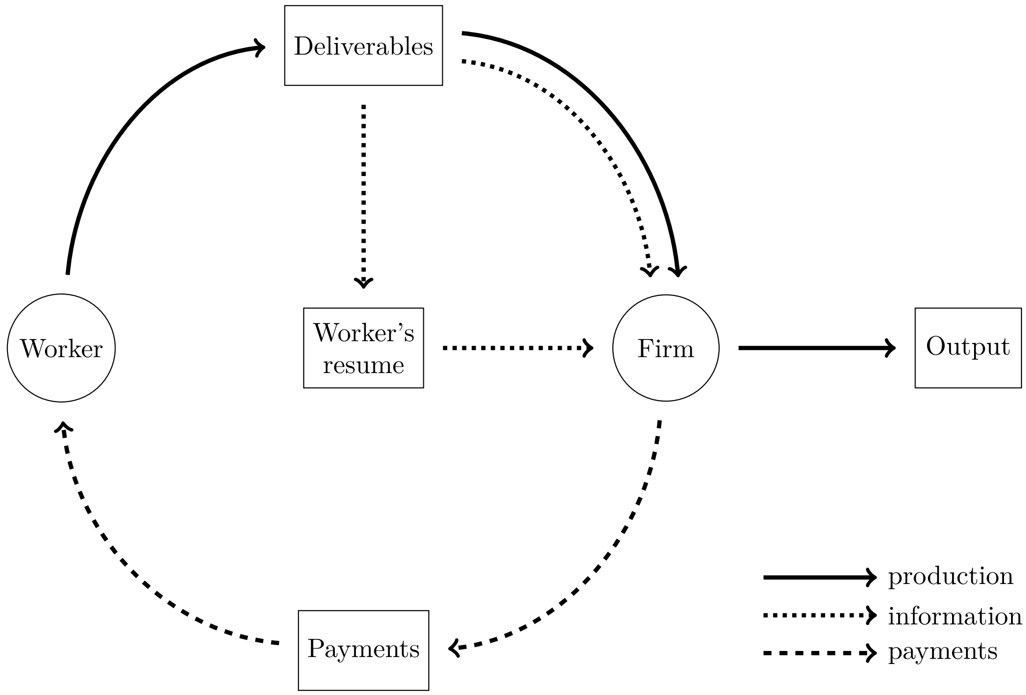


Figure 1.1: Flow of production, information, and payments

and transfers to and from the government is not pinned down by the model.

1.2.4 Equilibrium Definition

Equilibrium is described by workers choosing optimal consumption and labor supply flows taking salaries and taxes as given, and anticipating the effect that their labor supply decisions have on their future salaries and taxes, as stated in 1.1, while firms simultaneously set salaries according to the zero profit condition 1.2, presented below, paying each worker for their expected productivity conditional on their resumes. This zero profit condition is justified by competition among firms to enter the market and hire labor.

Definition. For each marginal unit of labor, the firms' free entry condition is described by an Akerlof (1970) lemons condition,[13] where salaries are equal to productivity of the workers with the same resume:[14]

$$w(I(\tilde{h}(\cdot), a, \theta)) = \mathbb{E}[v(\theta)|I(\tilde{h}(\cdot), a, \theta)] \tag{1.2}$$

Equation 1.2 lies at the core of the model. It simply states that at any period people with equivalent resumes will be paid equally, and that firms will on average break-even.

Equilibrium, thus, is described by workers choosing optimal consumption and labor supply flows taking salaries as given, and anticipating the effect that their labor supply decisions have on their future salaries,

[13] This condition perhaps would more precisely be named in this context, an Akerlof "peach condition", because since it is assumed that those who are more willing to work are those who are more productive, there is advantageous selection instead of adverse selection.

[14] In this definition, it is assumed that I and $\tilde{h}(\cdot)$ are continuously distributed, so that the expectation is the same if it is conditioned on the workers who accept the contract.

as stated in 1.1, while firms simultaneously are setting salaries according to the condition 1.2, paying each worker for their expected productivity conditional on their resumes.

1.3 Two Polar Cases

In this section, two polar cases illustrate a range of possibilities this simple setup can accommodate. These two benchmark cases, and additional parametric examples, are illustrative of the patterns of wages and the return to experience, as well as of the type of distortions that dynamic labor market signaling can generate.

In the first case, payments per unit of the deliverable will be, for every worker, increasing with experience and almost at all times different than their marginal product of labor. In spite of salaries being higher at advanced stages of the career of the workers, there will be no intertemporal distortions. A dynamic rat race will generate high-powered incentives for the workers to exert effort throughout their careers.

In the second case, payments per unit of the deliverable will be constant throughout the career of the workers, and equal to their marginal product of labor. In spite of pre-tax salaries being constant, corrective taxes will be necessary to avoid large intertemporal distortions, and excessively strong incentives to work hard at earlier stages of the career of workers. A dynamic rat rate will generate high-powered incentives to exert effort, but those incentives will be declining over time.

1.3.1 First Case: The Length of the Resume

The first benchmark case, or example of the informational structure we just presented, is one where the worker's resume is summarized by the cumulative discounted sum of what the worker has produced or its "length," that is, $I(\tilde{h}(\cdot), a, \theta) = \int_0^a q(\tilde{a})\tilde{h}(\tilde{a})d\tilde{a}$. This case has features that are closely analogous to the Arrow (1962) learning-by-doing model, where productivity is a function of the cumulative use of a factor. Similarly in this case, firms' perceptions about each worker's productivity are also a function of their cumulative labor supply.

This case is natural if most or all of the heterogeneity in labor supply is at the extensive margin, and if, for firms, it is hard to observe anything other than how many years or hours of experience a worker has. In fact, in many occupations, it is hard for firms to observe and verify much more than the start and end date of previous positions, or the number of hours someone has been working for the current position.

This case is also natural if firms believe that there is a lot of human capital accumulation, so that the model they may have in mind is one where there is on-the-job learning, in the form of the Arrow (1962) learning-by-doing model.[15] In that vein, firms could mistake the return to experience as human capital accumulation instead of changes in the composition of the pool of workers. That is, the firm could hire many workers with different lengths of resumes, and observe that the pool of workers with longer resumes is, on average, more productive. Without knowing why the pool is more productive, it may attribute that fact to human capital accumulation instead of signaling. Firms could think the change in the composition of workers who happen to achieve a higher degree of experience is a real return to human capital when it is not, and still post the same salaries and make zero profits. Workers as well may not know whether they will become more productive by working more, or whether they would just signal to employers they are more productive. From the workers' perspective, to make their labor supply choices, it only matters how experience accumulation and labor supply decisions today will likely impact their future salaries.

[15] A more general version of the model that has both ingredients is presented and discussed in Section 1.7.

Under the assumption that resumes are summarized by their "length," salaries take a particularly simple form, and satisfy a modified Akerlof lemons condition, as stated in Lemma 1.

Lemma 1. *If the information of firms is defined as the "length-of-the-resume," that is,* $I(\tilde{h}(\cdot), a, \theta) = \int_0^a q(\tilde{a})\tilde{h}(\tilde{a})d\tilde{a}$*, then the Akerlof lemons condition (Equation 1.2) is equivalent to:*

$$w(h) = \mathbb{E}[v(\theta)|h(\theta) \geq h] \tag{1.3}$$

Proof. See Appendix Section 1.9.1. □

Where $h(\theta)$ denotes the length of the resume of type θ at the end of their career, that is $h(\theta) \equiv h(1, \theta)$. Lemma 1 states that salaries as a function of the length of resume h are the average productivity of all the types who eventually reach a point in their careers where their resumes are longer than h. This is the case, because at any period, for every type who eventually reaches a longer than h resume, there is someone, potentially from a different generation, who has a resume of length h today.[16] However, if workers' labor supply decisions are heterogeneous only at the extensive margin (differing in how long they stay in the labor force) or if the deliverables were defined as working additional years, then the set of workers that would be pooled together would consist only of workers of the same cohort.[17]

Given equation 1.3, which describes salaries as a function of the cumulative labor supplied, it is useful to note that lifetime income can be written quite simply. We can change variables to express the worker lifetime income as $y(h) \equiv y(h(1)) = \int_0^1 w(h(a))\tilde{h}(a)q(a)da = \int_0^h w(z)dz$, that is we can express lifetime income as instead of integrating over time, as integrating over increases in the length of the resume (dz).That is, how much more the worker would receive for providing one extra unit of the deliverable (in present value) over their lifetime, is the same at any point in the career of the worker and it is equal to the payment for the last unit of the deliverable, that is $y'(h) = \mathbb{E}[v(\theta)|h(\theta) \geq h]$. This equation also helps simplify the analysis behind this dynamic model of career concerns, making it as simple as a static model.

In the economically-sensible case where it is less costly for the more productive people to supply the deliverables, salaries increase with experience. Workers at the beginning of their careers are willing to work more, relative to the myopic trade-off between current salaries and current effort, because they expect higher future salaries as an outcome of building up their experience, signaling to employers that they have higher productivities. The property that exerting effort increases future salaries, by affecting employers' perceptions of the ability of the worker, will be shared by a large range of informational structures, as described in more detail in Section 1.4.

At the beginning of their working life, for the first job they can get, all individuals face the same wage. Employers cannot distinguish between workers who have no experience at all. As workers advance in their careers, completing tasks and increasing their lifetime supply of labor, the length of the resume of hard-working individuals works to separate them from the other workers that execute fewer tasks and have shorter resumes. These more productive workers initiate their careers subsidizing the less productive, but at each new task they execute, some less productive workers are left behind with a shorter resume. For this reason, the remuneration that the more productive workers receive for the execution of tasks becomes progressively higher.

[16] The same argument could be applied to any specification of I of the form $I(\tilde{h}(\cdot), a, \theta) = \int_0^a \phi(\tilde{a})\tilde{h}(\tilde{a})d\tilde{a}$.

[17] Given this structure of labor demand and supply, a natural question in this setup is whether there exists an equilibrium, as defined in Section 1.2.4. Proposition 11 in the Appendix Section 1.9.3, shows that it does exist under standard assumptions. Namely, these assumptions state that the distribution of types, productivities and marginal rates of substitution are continuous, and that marginal rates of substitution are smooth as a function of lifetime consumption and lifetime income, with bounded derivatives.

The size of the return to experience depends on how many people are being left behind by these more productive workers as they advance in their careers, and on how much more productive they are relative to those with shorter resumes. That is, the return to experience depends on the joint distribution of preferences and productivities.

Fortunately, there are simple and sensible assumptions on preferences and heterogeneity that allow us to derive equally simple expressions for the return to experience and the shape of the income distribution, as well as to analyze how they would change in response to information becoming more symmetric. Towards this goal, let's assume that preferences over lifetime labor supply and lifetime consumption are such that there is a constant elasticity of lifetime labor supply and those preferences are quasilinear in lifetime consumption.[18]

$$U(c,h,\theta) = c - \left(\frac{h}{b(\theta)}\right)^{1+\frac{1}{\epsilon}}\left(1+\frac{1}{\epsilon}\right)^{-1}$$

where $b(\theta) = \theta^{1-\delta}$,$v(\theta) = \theta^{\delta}$, $\theta \sim$ Pareto with shape parameter $\alpha > 1$, and $0 \leq \delta < 1$. In this example, δ governs the amount of information asymmetry: a higher δ means more heterogeneity comes from unobserved productivities ($v(\theta)$) instead of observable productivities ($b(\theta)$). Under this formulation $b(\theta)$ should be thought of as how many deliverables a worker can provide per unit of effort l, that is, $h = l \cdot b(\theta)$. $v(\theta)$ should be interpreted as how much output the worker generates per unit of the deliverable, that is, $y = h \cdot v(\theta)$. The total productivity as a function of effort then is just the product of $v(\theta)$ and $b(\theta)$, and it is equal to θ. Using the equilibrium definition and the convenient properties of Pareto distributions, we can guess and verify that salaries are also a power function, so wages and experience follow a log-linear relationship:

$$log(w) = \gamma \cdot log(h) + \kappa$$

where $\gamma = \frac{\delta}{1-\delta+\epsilon}$.[19] There is a constant proportional return to experience that is entirely driven by selection or employers learning about the types who are willing to take the jobs they are offering. This return to experience is larger when heterogeneity comes mostly from unobserved productivities, that is when δ is higher, as there is more heterogeneity to be screened out by experience. When the elasticity parameter ϵ is low, the return to experience is also larger: in this case, a higher experience is really indicating that the worker is more productive. When the elasticity is low, the difference between how a marginal increase in labor supply hurts the less productive relative to the more productive workers becomes larger.

Because higher types are more productive and are the workers who are willing to work longer, wages increase over time, and at the time of retirement, each worker would be facing a higher salary than their productivity. As noted earlier, whenever the resume is summarized by its "length", the lifetime benefits of increasing the lifetime labor supply do not depend on when the worker completes these extra units of the deliverable. Therefore, workers face high-powered incentives to work not only at the time of retirement, but at all moments in their careers. This implies that Pigouvian component of taxes as defined in Section 1.5 will be positive, and will correct for that distortion. This also will imply that dynamic job market signaling, in this case, will not introduce intertemporal distortions, and the optimal tax base will be the workers' lifetime earnings.

As will be discussed further in Section 1.5 the marginal Pigouvian component of taxes can be thought

[18] Note that given the result from Proposition 1, and assuming exogenous discount rates, it is without loss to specify preferences in terms of lifetime labor supply and lifetime consumption. This means that behind this preference specification there could be either heterogeneity only at the extensive margin (in which case, workers of different cohort would not be pooled together), or richer heterogeinity in preferences (in which case, workers of different cohorts potentially would be pooled together).

[19] For the algebra behind this example, see Appendix Section 1.9.2.

of as guaranteeing that workers at the margin are paid for their marginal product, i.e. post- Pigouvian tax payments for an additional increase in lifetime labor supply marginal unit of labor ($r(h)$) should be such that $r(h) = v(\theta)$. We can translate this condition as a restriction on marginal retention on earnings ($r(y)$), which state that $r(y) = \frac{v(\theta(y))}{E[v(\theta)|y(\theta)\geq y]}$. In this case, under the parametric assumptions above, $r(y) = \frac{\alpha-\delta}{\alpha}$.

Thus, these Pigouvian taxes will also be increasing in the amount of information asymmetry δ, and, given the parametric assumptions in this example, they will be constant as a function of lifetime income. In this case, the Pigouvian corrections can be implemented with linear taxes. This makes the implementation particularly simple, because since these taxes are linear, they do not need to be a function of lifetime income per se and can be implemented with history-independent annual income taxes.[20]

To summarize, whenever the public signal the firms can obtain from workers is the "lenghth-of-their-resumes," payments per unit of the deliverable will be, for every worker, increasing with experience, and almost at all times different than their marginal product of labor. In spite of salaries being higher at advanced stages of the career of the workers, there will be no intertemporal distortions. A dynamic rat race will generate high-powered incentives for the workers to exert effort throughout their careers. Pigouvian taxes can be used to correct for these high-powered incentives as will be discussed in Section 1.5, taking lifetime income as an optimal tax base. In a simple parametric example, the return to experience will take a familiar log-linear form, and marginal Pigouvian taxes will be constant, implying that they can be implemented with history-independent annual taxes.

1.3.2 Second Case: The Pace of the Resume

This section presents the second benchmark case, where the resume is summarized by "pace of the resume": the ratio of the total amount of deliverables a worker has produced and the age of the worker, i.e. $I(\tilde{h}(\cdot), a, \theta) = \int_0^a \frac{\tilde{h}(\tilde{a})}{a} d\tilde{a}$. This case is particularly natural if all the heterogeneity in preferences are at the intensive margin, and employers are aware of the relationship between the heterogeneity in productivities and the willingness to supply deliverables.

In this example payments per unit of the deliverable will be constant thoughout the career of the worker, and equal to their marginal product of labor. However, the corrective component of taxes will be positive and will correct for intertemporal distortions and the incentives to work too much when young.

We assume that workers have additively separable preferences over time, and preferences are heterogenous only on the intensive margin. That is, every worker starts and ends their careers at the same age, but each of them may be willing to supply deliverables at different rates. In other words, we can think of all the heterogeneity as coming from how each worker evaluates different paces of work, and not from how long they would like to stay in the labor market.

Under that assumption, any optimal allocation has each worker consuming and supplying labor at constant rates over their lifetimes. As we will see in more detail in Section 1.5, the government can implement those allocations using age-dependent and history-dependent taxes.

Moreover, whenever there is separation, more productive types will supply labor at higher rates, and therefore will have stronger resumes. Because they supply labor at constant rates, each worker will also reveal a constant "pace-of-the-resume," and each type, being associated uniquely with a certain "pace" will receive a different pre-tax salary, which will be equal to their marginal productivity. Therefore, for each

[20]Linearity is one property that guarantees that lifetime income taxes can be implemented with history-independent annual taxes. Besides linearity, whenever there is heterogeneity only at the extensive margin, lifetime income taxes can be implemented with history-independent non-linear annual taxes.

worker, salaries will be equal to their marginal productivity almost at every point in their careers.[21]

The fact that pre-tax salaries are equal to the productivity of the workers does not mean, however, that there are no distortions that taxes should take care of. In fact, taxes should correct for potentially large intertemporal distortions that arise from the dynamic reputation building effects. Whenever young workers consider increasing their labor supply, they receive not only larger payments now, they also improve their resumes in all future periods, and increase the remuneration they would receive for the future completion of deliverables. The role of Pigouvian taxes in this case is to counterbalance this effect, exactly cancelling the dynamic reputation building benefits, and making sure that the lifetime benefits the workers would receive by increasing their labor supply are equal to only their current salaries which in turn are already equal to their marginal productivities.

The size of the reputation building effects, and the return to experience depend on the joint distribution of productivities and willingness to provide the deliverables. Fortunately, there are simple parametric assumptions that allows us to derive simple expressions for the signaling return to experience and for the corrective component of taxes. In particular, for illustrative purposes, let us assume that preferences take the form:

$$U = \int_0^1 e^{-\rho a}\left(\tilde{c}(a) - \frac{\tilde{h}(a)^{1+\frac{1}{\epsilon}}}{1+\frac{1}{\epsilon}} b(\theta)^{-(1+\frac{1}{\epsilon})}\right)da,$$

where $b(\theta)$ is the cost of a worker of type θ to work more and provide more deliverables per unity of time. We denote productivities by $v(\theta)$, and assume production is linear in the flow of labor supply. Productivities and preferences are parameterized so that $v(\theta) = \theta^{\delta}$, $b(\theta) = \theta^{1-\delta}$, with $0 < \delta < 1$.

Any optimal allocation has a constant labor supply over time. Moreover, for concreteness let's focus on the allocation where the redistributive component of taxes is set to zero, so that for each worker, the optimal flow of labor supply satisfies the following first order condition:

$$\tilde{h}(\theta)^{\frac{1}{\epsilon}} b(\theta)^{-(1+\frac{1}{\epsilon})} = v(\theta)$$

Denoting a particular level of a resume pace as h_α, using the condition above we can conclude that pre-tax salaries as a function of the pace would satisfy $w(h_\alpha) = \mathbb{E}[v(\theta)|\frac{\int_0^a \tilde{h}(\tilde{a})d\tilde{a}}{a} = h_a] = h_a^{\frac{\delta}{1+\epsilon-\delta}}$. Notice these are exactly the productivity of the workers for whom it is optimal to supply a certain pace h_α. That is, salaries at each period are equal to the marginal productivity of the workers. As discussed above, this does not imply that there should be no corrective taxes; in fact the relevant labor wedge is not the static difference between productivities and current salaries, but the difference of productivities and the sum of current salaries and the increases in future lifetime earnings from reputation builiding effects.

In fact, for a worker who has supplied labor at the pace h_α up until the age $\bar{a}$, increasing their labor supply today and then reverting back to their constant flow increases their lifetime earnings by:

$$\frac{dy}{d\tilde{h}(\bar{a})}(h_a, \bar{a}) = e^{-\rho\bar{a}} h_a^{\frac{\delta}{1+\epsilon-\delta}} + \left(\frac{\delta}{1+\epsilon-\delta}\right)\int_{\bar{a}}^1 \frac{e^{-\rho a}}{a} da \cdot \left(h_a\right)^{\frac{\delta}{1+\epsilon-\delta}-1}$$

where the first term is the current salary, and the second term are the reputation effects, that is the increase of future payments for the completion of future deliverables (holding the future flow at the constant pace h_α). Because the first term is exactly the productivity of the workers, corrective taxes as a function of labor supply (τ) should be exactly equal to the second term, that is:

[21] The exception being a zero measure moment when they have just entered the workforce and have empty resumes.

$$\tau(h_a, \bar{a}) = \left(\frac{\delta}{1+\epsilon-\delta}\right) \int_{\bar{a}}^{1} \frac{e^{-\rho a}}{a} da \cdot \left(h_a\right)^{\frac{\delta}{1+\epsilon-\delta}-1} \tag{1.4}$$

where h_a can be inferred from the history of earnings. In this simple case, those Pigouvian taxes can be implemented as relatively simple age- and history-dependent taxes. For workers of age a, if their lifetime earnings has been increasing at a ratio that is consistent with a constant labor supply flow equal to h_α, then the marginal taxes on labor they should face should be equal to those in equation 1.4.[22] Those Pigouvian taxes correct not for the difference between current payment and productivities (which in this example is zero!), but for the difference between the lifetime gains from a marginal increase in effort today and the productivities. Those extra lifetime gains come from the positive reputation effects of exerting higher effort, which result in higher salaries in the future.

These large intertemporal corrective taxes are suggestive that, without them, younger workers would exert a lot more effort and work longer hours. This prediction is in line with the evidence from Landers et al. (1996) who show that, in law firms, hours decrease with tenure and associates work too many hours relative to their desires and relative to partners. This prediction however is at odds with the commonly found inverted U-shaped pattern for hours over the life-cycle: younger and less experienced workers work fewer hours, hours are mostly constant for workers between the ages of 25 to 55, and they quickly fall for older workers (Card, 1991; Kaplan, 2012).

To summarize, whenever the public signal the firms can obtain from workers is the "pace-of-their-resumes," payments per unit of the deliverable will be, for every worker, constant throughout their lifetime, and almost at all times equal to their marginal product of labor. In spite of salaries being constant, there will be large intertemporal distortions. A dynamic rat race will generate high-powered incentives for the workers to exert effort throughout their careers, but those incentives will be stronger for younger workers and will decline as workers approach their retirement age. Pigouvian taxes can be used to correct for these high-powered incentives and intertemporal distortions as will be discussed in Section 1.5.

1.4 General Positive Properties

The examples presented in the previous section were stylized, and relied on particular assumptions on the information structure and on parametric assumptions on preferences and productivities. This section shows properties of salaries and the return to experience that hold under a broader set of assumptions. We focus on the case where the resume is summarized by a signal of the form $I(\tilde{h}(\cdot), a, \theta) = \int_0^a \tilde{h}(\tilde{a})\phi(\tilde{a}, a)d\tilde{a}$, with $\phi(\tilde{a}, a) > 0$, bounded and continuous. In contrast to the previous cases, where the signal was referred to the "length-of-the-resume" and the "pace-of-the-resume", we say that in this more general case the firms receive a public signal which can be called the "strength-of-the-resume" of the candidate.

Together with the overlapping generation structure of the model, this formulation can be thought of as a parsimonious way of introducing noise in how employers assess what has been done and when. While doing that, it implies that employers may pool together workers of different generations, whenever their past experiences are assessed as equivalent. It allows firms to see and put different weights on past experience, and those weights may depend on the current age of the employee as well as the age the employees had when they completed each task.

[22] Notice that Lemma 5 guarantees that quite generally in this setup the government can infer the labor supply flows from the earning flows.

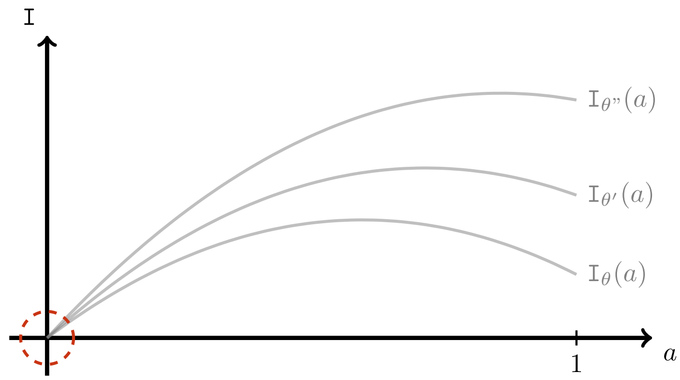


Figure 1.2: Career of Workers: Initial Salaries

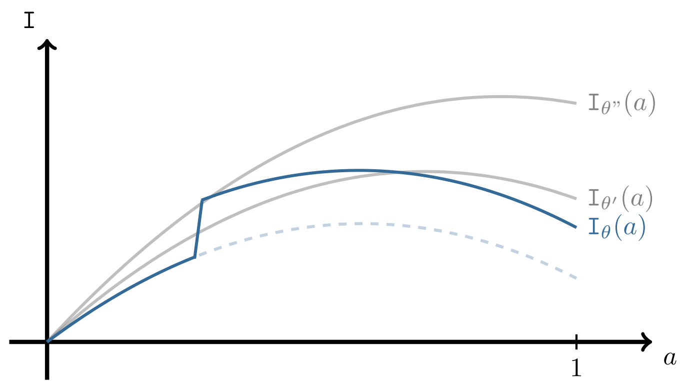


Figure 1.3: Career of Workers: Signaling Return to Experience

First, for any preferences and distribution of productivities workers, every worker starts their career with nothing to show in their resumes, and therefore the strenght of their resumes is zero. Everyone has the same initial salary, which is equal to the average productivity of all workers. In particular, workers with above-average productivity start their career earning less than their marginal product, as the average productivity is everyone's initial salary. All workers have to climb the same career ladder and before they can show anything in their resume they are all indistinguishable for employers. This is illustrated in Figure 1.2, where the trajectories of the "strength-of-the-resume" of different hypothetical workers are plotted as a function of their age, and the pool of workers who just started their careers is highlighted in red.

Second, at any point, assuming more productive types are more willing to provide the deliverables at all periods, exerting more effort increases future salaries. This is the case because whenever workers decide to exert more effort and improve their resumes, they will show to employers resumes that are more similar to the resumes of the more productive types. This is illustrated in Figure 1.3, where in blue workers of a certain type decide to exert more effort at a certain age, and as a result they make their resume stronger in all future periods. This is a general property that will be satisfied in other information structures as discussed in Section 1.5.2.

Third, workers at the peak of their careers (when they reach the highest value of the index $\int_0^a \tilde{h}(\tilde{a})\phi(\tilde{a},a)d\tilde{a}$) will earn more than their marginal productivity because they will be pooled together only with the types

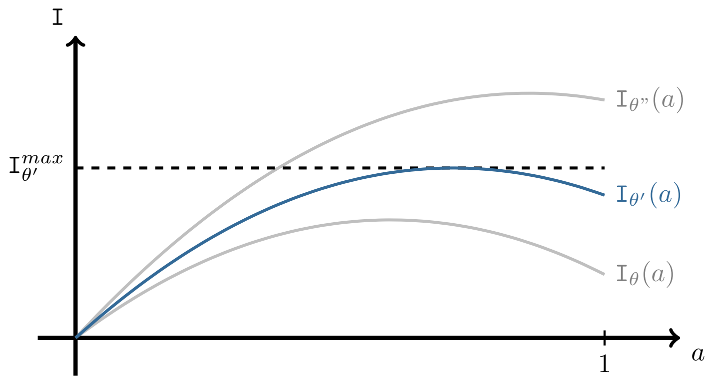


Figure 1.4: Career of Workers: Salaries at the Peak

that are willing to provide higher amounts of deliverables across all periods. This is illustrated in Figure 1.4. In Figure 1.4, a dashed horizontal line shows the pool of workers who at some point in their careers have a resume of a certain "strenght." As illustrated there, whenever this line is tangent to the highest "strength-of-the-resume" a certain type of worker achieves, that this worker will be pooled together only with people who reach even higher levels of that "strenght" and who are more productive than this worker (under the assumption that those who are more productive are more willing to supply deliverables).

These three observations are summarized in Proposition 1, below.

Proposition 1. *For any distribution of preferences and productivities, and for any ϕ, where $\mathtt{I} = \int_0^a \phi(\tilde{a}, a)\tilde{h}(\tilde{a})d\tilde{a}$, with $\phi > 0$ and $\tilde{h} \geq 0$ bounded and continuous, and provided that the workers who are more productive are more willing to supply deliverables:*

1. *All workers start with an empty resume and earn the same initial salaries;*

2. *There is a positive signaling return to experience;*

3. *At the peak of the career of each worker, salaries are higher than marginal productivity.*

Relative to Holmström (1999), this model imposes a different assumption on the contract space. In Holmström (1999) firms cannot pay for performance, firms do not observe effort but observe output, and overworking in a short period of time is a way for workers to trick employers into believing they are more productive. Here, firms must pay for the observable component of effort, the deliverables, and the fulfillment of these deliverables add to the resumes, which is a publicly available signal to employers. At the cost of other simplifying assumptions, the fact that firms in the model can condition payments on the performance measures that are available to them – as in fact many of them do – adds realism to the way that career concerns are modeled. Interestingly, employer learning under this assumption can be seen as a problem in itself instead of a side effect of an insurance and incentives tradeoff.

This matters not only from a positive perspective, but also from a normative point of view. Workers enter a "rat race" as in Akerlof (1976) and Holmström (1999), but interestingly, this "rat race," in the case where resumes are defined as the cumulative discounted sum of deliverables, does not generate an intertemporal distortion, but rather, generates a lifetime labor supply distortion. Workers work too much, but the timing of their labor supply decisions is not distorted. For building a resume, in that simple information structure, there is no advantage of concentrating efforts in a particular period of their lives. The discounted sum of payments a worker receives from increasing the length of the resume by one unit is the same independently of when the worker decides to increase it, and at the same time the contribution to the total output of the economy from increasing the length of the resume by one unit is also the same independently of when the worker decides to increase it. When the resume is alternatively summarized by its "pace," then there are large intertemporal distortions, but differently than Holmström (1999), the incentives to exert effort are never too low, and they approach undistorted levels as workers approach the retirement age and the reputation motive to exert effort disappears.

This also illustrates that the difference between current salaries and productivities does not in itself represent a distortion. Rather, distortions come from differences between productivities and the lifetime gains from exerting effort, where those lifetime gains are the sum of two components: current payments and increases in future payments from reputational effects. In this particular case, salaries are initially low but increase with experience, while the size of the distortion is constant over the lifetime of the worker. In Section 1.3.2, a simple example – where the resume instead of being defined as the cumulative sum of deliverables is defined as the pace of production of deliverables over time (i.e. $I(\tilde{h}(\cdot), a, \theta) = \int_0^a \frac{\tilde{h}(\tilde{a})}{a} d\tilde{a}$) – illustrates that the opposite case is also a possibility. Salaries can be exactly equal to the productivities of the workers at all times and, at the same time, there could be large intertemporal distortions coming from reputational effects and how effort today affects future payments.

1.5 Taxation

In this section, we focus on the normative implications of the model and how dynamic signaling effects should affect optimal taxation. We first focus on the case where the public signal that the firms see is defined as the "length-of-the-resume", or the discounted cumulative sum of what the worker has produced, and then we extend the results to more general information structures.

As in Mirrlees (1971), the government does not observe the workers' types, and because firms also do not observe the types, this model features what has been called "double adverse selection" (Stantcheva, 2014). We are going to place special focus on the case where the government sets taxes on lifetime income (y) to maximize a welfarist functional of utilities. This focus is motivated by two sets of reasons. On the practical side, many taxes and transfers actually condition on lifetime income: most notably, the key determinant of US social security benefits is a measure of the average earnings over the lifetime of a worker. Furthermore, for many taxes the timing of earnings can be manipulated: the realization of capital gains can be delayed indefinitely to avoid taxes, and there is flexibility in reporting the timing of income coming from C-corporations. On the theoretical side, an extension of the Atkinson and Stiglitz (1976) result shows that it is still optimal to tax lifetime income even with "double adverse selection."

Proposition 2. *(AS extension): If preferences take the form $U(C(\tilde{c}(\cdot)), H(\tilde{h}(\cdot)), \theta)$, then it is optimal to tax lifetime income, even with "double adverse selection."*

Proof. See Appendix Section 1.9.4. □

In this Proposition $C(\tilde{c}(\cdot))$ and $H(\tilde{h}(\cdot))$ are common aggregators of flows of consumption and labor supply, respectively. The assumption on preferences says that if preferences across households are homogenous over the timing of consumption and labor supply decisions, that is, if preferences can be written as a function of common indexes (C, H) that aggregate the flows of consumption and labor supply, and these indexes are the same across households, then it is optimal to use only lifetime income as the tax base. It means that dynamic job market signaling, under the baseline assumptions here in this paper, does not introduces intertemporal distortions and under the benchmark case where preferences over timing of consumption and labor flows are homogeneous across households, there would be no reason to introduce further distortions on intertemporal decisions. This resembles a special case in Holmström (1999), where there are also no intertemporal distortions, namely when there is no discounting and productivities follow a random walk process. More generally, in Holmström (1999), reputation concerns in general induce workers to exert relatively more effort earlier in their careers, and progressively exert less effort as their reputations are consolidated.

This result echoes tax smoothing results as in Werning (2007a), and the tax smoothing ideas that go as far back as Vickrey (1947), and adds another reason to use income averaging rules, as they are present for example in the calculation of social security benefits. Importantly, although the income base should be the lifetime income of a worker, those taxes do not need to be raised only at the end of a worker's life, and can be raised annually, as long as taxes each year depend on the history of earnings of each worker up to the current date.

Moreover, the result is general in an important way. It essentially says that post-tax lifetime earnings, when it is possible, should be a function of lifetime labor supply . In the simple model we just presented, pretax lifetime earnings are a function of lifetime labor supply, and thus taxes on lifetime income are enough to guarantee that post-tax lifetime earnings are a function of lifetime labor supply. But in extensions of the model where pretax lifetime earnings would cease to be a function of lifetime labor supply (for example if the signal the firm sees about a worker is a different function of the flows of labor supply instead of the discounted sum of these flows), then the optimal policy would entail taxes meant to undo these intertemporal distortions, and meant to make post-tax lifetime earnings a function of discounted lifetime labor supply.

1.5.1 Optimal Taxation Formulas

In this section, we derive necessary conditions for optimal taxes in terms of sufficient statistics, as in Saez (2001). We assume that the government maximizes a welfarist functional of worker utilities, $W(V(R;\theta))$.[23] The government solves:

$$\underset{R(y)}{max}\ \mathbb{E}[W(V(R;\theta))]\ s.t.\ \mathbb{E}[y(h(R;\theta),R) - R(y(h(R;\theta))] \geq 0$$

and subject to the constraint that pretax wages are determined by the differential equation $y'(h) = \mathbb{E}[v(\theta)|h(R;\theta) \geq h]$ with initial condition $y(0) = 0$.

Solving directly for $R(y)$ is complicated; changing taxes at an income level y has cascading effects on salaries of everyone earning y or more, by shifting the composition of workers. However, the following proposition allows us to simplify the problem by allowing the planner to keep salaries fixed when taxes change. In other words, we can frame the problem as the planner solving directly for post-tax salaries, effectively ignoring how pretax salaries are set.

Lemma 2. *Without loss, we can solve directly for $\tilde{R}(h) = R(y(h))$, and then find $y(h)$, and $R(y)$. That is, the planner can solve the simpler problem:*

[23]This formulation is quite general, and can be converted to Pareto efficiency tests as in Werning (2007b), by picking linear functionals of the form $W(V(R;\theta)) = \lambda(\theta) \cdot V(R;\theta)$, making it the dual of a revenue maximization problem subject to a minimum utility requirement for each type.

$$\underset{\tilde{R}(h)}{max}\ \mathbb{E}[W(V(\tilde{R};\theta))]\ s.t.\ \mathbb{E}[v(\theta)h(\tilde{R};\theta) - \tilde{R}(h(\tilde{R};\theta))] \geq 0$$

Proof. See Appendix Section 1.9.5. □

This result applies more generally to other models of labor market frictions provided that firms make zero profits,[24] and, given any allocation, $y(h)$ is well defined and invertible. An important example where these conditions hold is the Azevedo and Gottlieb (2017) model of competition with adverse selection. More generally, there could be other production externalities, imperfect competition generating compressed wages, or monopolistic screening by a single firm. This feature allows us to conceptually separate what are the relevant externalities coming from the information frictions that the planner would like to correct from what are just regular transfers and innocuous price adjustments. Importantly, the same idea will be applied in Section 1.5.2, to more general information structures, allowing us to separate issues of efficiency from issues of redistribution quite generally.[25]

Proposition 3. *(For single-dimensional θ) If a tax schedule is optimal, then it satisfies the following optimal tax formula:*

$$\left(\frac{\chi(y) - r(y)}{r(y)}\right)\epsilon_r^c(y)g(y)y = \int_y^\infty \left(1 - \lambda(\tilde{y})\right)g(\tilde{y})d\tilde{y} + \int_y^\infty \left(\frac{\chi(\tilde{y}) - r(\tilde{y})}{r(\tilde{y})}\right)\eta_I(\tilde{y})g(\tilde{y})d\tilde{y} \tag{1.5}$$

Proof. See Appendix Section 1.9.6.

where $r(y)$ denotes the marginal retention function; $\epsilon_r^c(y) = \frac{dy}{dr_y}\frac{r(y)}{y}$ are the local compensated elasticities; $g(y)$ denotes the density of y; $\lambda(y) \equiv \frac{W'(V)U_c}{\mu}$ is the marginal value the planner places on transfer to a worker earning y; $\eta_I(y) = \frac{dy}{dI}r(y)$ is the income elasticity; and $\chi(y) \equiv \frac{v(y)}{y'(h(y))}$, and $v(y)$ is the productivity of the worker with lifetime income y. □

This equation is almost the same as the standard first-order condition that appears in Saez (2001) and Werning (2007b), with one additional ingredient: $\chi(y)$. The equation says that if a tax schedule is optimal, then three sorts of effects should balance each other: compensated effects, mechanical and welfare effects, and income effects. When a planner considers increasing marginal retention over a small region (holding post-tax salaries fixed everywhere else), there are compensated effects (left-hand side of the equation) coming from the fact that people would work more. These effects are proportional to the densities and compensated elasticities, and they ease the feasibility constraint not by $1 - r(y)$ as in the standard Mirrleesian model, but by $\chi(y) - r(y)$, that is, a worker who earns an extra dollar in their lifetime generates $\chi(y)$ units of output per dollar earned (and retains $r(y)$). The mechanical and welfare effects (first term on the right-hand side of the equation) are the same as before: the planner is giving one dollar to those who earn at least y, and this has mechanical costs of one dollar per person and welfare effects that are weighted by the marginal value of a dollar that the planner attributes to a transfer to the each of these people. Finally, there are income effects that affect everyone who earns at least y. These people are induced to work less (if income effects are negative), and, as in the compensated effects, they damage the feasibility constraint by $\chi(y) - r(y)$.

The same equation can alternatively be written in two blocks: one block that translates the redistributive motive, and that looks exactly like the standard Mirrleesian formula,[26] and a second block that defines

[24] Or, more generally, there is full taxation of profits, or, still, profits are uniformly shared between workers, as in Scheuer and Werning (2017).

[25] Innocuous in the sense of the first and second welfare theorems, where changes in fundamentals could result in changes in relative prices which no planner that has access to individual-specific transfers would not like to undo.

[26] As it appears for example in Scheuer and Werning (2017), and Saez (2001).

retention as the product of two components, the Mirrleesian, and a Pigouvian component that corrects for the production externality:

$$\left(\frac{1-r_m(y)}{r_m(y)}\right)\epsilon_r^c(y)g(y)y = \int_y^\infty \left(1-\lambda(\tilde{y})\right)g(\tilde{y})d\tilde{y} + \int_y^\infty \left(\frac{1-r_m(\tilde{y})}{r_m(\tilde{y})}\right)\eta_I(\tilde{y})g(\tilde{y})d\tilde{y}$$

$$r(y) = r_m(y)\cdot\chi(y)$$

This decomposition allows us to understand what exactly are the externalities that career concerns make workers impose on each other. The key insight is that, from the point of view of individual workers, their individual actions do not affect the whole remuneration schedule of other workers, but they do not get paid their marginal products. When increasing their lifetime labor supply so that they get paid one extra dollar, they are contributing to the economy not one dollar but $\chi(y)$ dollars. In other words, it is as if they would produce one dollar but generate negative production externalities of the size of $1-\chi(y)$ dollars.

It is interesting to compare how (a) the economy without any information asymmetries, (b) the economy with information asymmetries only between the government and the workers, as in the standard Mirrlees model, and (c) the economy with information asymmetries between the firms, workers and government compare in terms of their utility possibilities frontiers. In (b), when there are no information asymmetries in the labor markets, the planner could set taxes to zero, and then by the first welfare theorem, we know this is a first best allocation. Hence, there is a common point in the first and second best utility possibilities frontiers, i.e., in the utility possibilities frontiers of (a) and (b). Now, what is perhaps surprising is that there is a common point in the utility possibilities frontiers of (a) and (c), that is, between the third best utility possibilities frontier and the first best utility possibilities frontier, which is achieved by setting the Mirrleesian component of taxes to zero, so that $r(y) = \chi(y)$, as shown in Proposition 14, in the Appendix Section 1.9.8. This further justifies the decomposition above as one between a Mirrleesian and a Pigouvian component.

Employers, when facing workers with the same resume, do not know who are the workers who will retire and will not extend their resumes further. By the assumption that those who are more willing to extend their resumes are those who are more productive in the unobservable dimension, the workers who are retiring have the smallest productivity among those with the same resume. Therefore, we have that $\chi(y) \leq 1$, workers at retirement get paid more than their marginal products. From the tax formula above, holding estimates of elasticities and densities of the income distribution constant, we can see that taking into account these career concerns unambiguously pushes towards higher marginal taxes at every income level.

These optimal taxes do not depend only on labor wedges but also on elasticities and the shape of the income distribution. Appendix Section 1.9.7 shows that from the point of view of optimal taxation formulas, similarly to results in Scheuer and Werning (2017) common elasticity estimates are biased downwards. The reason for it is that increases in marginal retention induce the marginal types – who are the less productive types – to work more, and therefore reduce pretax salaries, making the effective change in post-tax wages smaller. Thus, the estimated elasticities of taxable income are lower in magnitude than the elasticities that enter optimal taxation formulas, which keep pretax salaries fixed. This is a countervailing force that pushes towards lower instead of higher marginal taxes.

These formulas can also be read as Pareto efficiency tests, as in Werning (2007b). If a tax schedule is Pareto efficient, then there are weights $\lambda(y) \geq 0$ such that, given the current tax rates, the income distribution, and the estimated elasticities and labor wedges $\chi(y)$, the formula above holds. Relative to the standard Pareto efficiency test, the inclusion of $\chi(y)$, holding the other estimated statistics fixed, makes the formula easier to be satisfied, that is, higher marginal tax rates can be rationalized.

1.5.2 Taxes Under Richer Signal Structures

So far in this Section, we have assumed that firms summarize a resume by looking at its length, which we have defined as the discounted cumulative sum of deliverables the worker has completed. This section relaxes this assumption and considers a richer set of signal structures under the same framework.

Conditioning on the "Strength of the Resume"

In this subsection, we will allow for a broad class of signals, that take the form of $h_\phi(\tilde{h}(\tilde{a})_0^1, a) = \int_0^a \phi(\tilde{a}, a)\tilde{h}(\tilde{a})d\tilde{a}$, with $\phi(\tilde{a}, a) > 0$, continuous in a and $\tilde{a}$.[27] The general idea behind this formulation is that the signal the firm sees can be thought of as the "strength of the resume". This measure is an imperfect signal of the past history of deliverables of the worker, but completing more deliverables always makes the resume stronger.

This formulation can capture in a reduced form different possibilities: firms may be interested in the pace at which the worker has produced deliverables, and thus may use $\phi(\tilde{a}, a) = 1/a$. The firms may want to look at the total experience, as a proxy of human capital, and use $\phi(\tilde{a}, a) = q(\tilde{a})$. It may become hard to verify experiences in the distant past, so that $\phi(\tilde{a}, a) < q(\tilde{a})$. The firms may have all those concerns at the same time, as long as, when put together, they can be summarized by the idea that firms would evaluate the experience through the lens of an index of the form $h_\phi(\tilde{h}(\tilde{a})_0^1, a) = \int_0^a \phi(\tilde{a}, a)\tilde{h}(\tilde{a})d\tilde{a}$. This formulation bypasses the need to introduce explicitly all those elements, and to postulate explicit stochastic processes for shocks for each of those elements and considerations. Not introducing those shocks directly makes the analysis of optimal tax systems tractable, and avoids technical issues arriving from multidimensional screening problems, as well as from failures of the homogeneity assumption on preferences over the timing of consumption and labor supply flows.[28]

Despite this apparent complexity, optimal taxation formulas will take a simple structure. We will follow the same logic from Proposition 2, and assume preferences are homogeneous over the timing of labor supply and consumption as in Proposition 2, and that more productive types are more willing to provide the deliverables.[29] Under these assumptions, it will be shown that, as in the previous section, taxes can be described as the composition of i) corrective taxes that guarantee that the workers benefits from each marginal increase in labor supply is equal to the their contribution to the output of the firm, and ii) redistributive taxes that are described by the same Mirrleesian optimal taxation formulas.

To show that, we first set aside the issue of implementation, and characterize the set of optimal incentive compatible allocations. Those allocations need to satisfy two properties, more formally stated in Lemma 3, in Appendix Section 1.9.11. These properties state that, ii) intertemporally, labor supply decisions will not be distorted, and ii) an analogous equation to the standard Mirrleesian formula should hold in terms of the lifetime labor disutility index H.

[27] This assumption guarantees that expectations of productivities are well defined, and further it will be shown that those expectations are increasing in the completion of tasks $\tilde{h}$. But, more generally, we could consider any informational structure with these properties. Later in this section, the full history of labor supply decisions will be assumed to be observed, and there will be richer heterogeneity in preferences to guarantee that those expectations are well-defined.

[28] One may worry that those weights should be endogenous to the tax system. However, as it will be shown, the same optimal taxation formulas will hold even if those weights were to depend on the tax system.

[29] Which in this case follows from additionally assuming that $\frac{d^2 H(\tilde{h}(\cdot))}{d\tilde{h}(a')d\tilde{h}(a)} < 0$.

Then, towards the goal of implementing these optimal allocations with a system of taxes and transfers, we establish that, quite generally, there is a positive return to experience – formally stated in Lemma 4 (in Appendix Section 1.9.11). In other words, exerting more effort is a way to signal to employers that you are a more productive worker, and thus, will impact future salaries positively. Intuitively, the signals considered in this section are increasing in the effort decisions, and under the assumption that those who are willing to exert more effort are those who are more productive, employers can infer that higher signals translate into higher expected productivities. Besides being of interest in itself, this result is used to show that the planner can infer labor supply decisions from earnings histories, as formally stated in Lemma 5 (in Appendix Section 1.9.11).

This result allows us to decentralize the incentive compatible and efficient allocations, analogous to the way that Lemma 2 was used to derive the optimal lifetime income taxation formulas from Proposition 3. That is, we can think of the planner as solving for the optimal allocation as in Lemma 3, or equivalently, solving for post-tax wages $R(\tilde{h}(\cdot)_0^1)$, and then implementing this post-tax retention function with history-dependent earnings taxes.

Finally, we summarize these taxation results in Proposition 4, below, which states that the tax system should be such that i) history-dependent taxes (R_p) should be used to correct for labor wedges, and ii) after correcting for these distortions, lifetime income redistributive taxes should be imposed on top of these taxes, according to standard redistributive formulas.

Proposition 4. *If $R(\tilde{y}(\cdot)_0^1)$ is optimal, then, there exists R_m, R_p with $R(\tilde{y}(\cdot)_0^1) = R_m(R_p(\tilde{y}(\cdot)_0^1)$, such that R_m and R_p satisfy the following conditions:*

1. *Intertemporal, Pigouvian: for any $\bar{a}$,$\underline{a}$, and $H(\tilde{h}(a)_0^1) = H$, switching the timing of labor supply decisions and holding fixed lifetime labor supply and H should leave lifetime earnings unaffected:*

$$\int_{\bar{a}}^1 \frac{dR_p}{d\tilde{y}(\tilde{h}(\tilde{a})_0^a)} \frac{d\tilde{y}(\tilde{h}(\tilde{a})_0^a)}{d\tilde{h}(\bar{a})q(\bar{a})} \frac{d\tilde{h}(\bar{a})q(\bar{a})}{dH} q(a)da = \int_{\underline{a}}^1 \frac{dR_p}{d\tilde{y}(\tilde{h}(\tilde{a})_0^a)} \frac{d\tilde{y}(\tilde{h}(\tilde{a})_0^a)}{d\tilde{h}(\underline{a})q(\underline{a})} \frac{d\tilde{h}(\underline{a})q(\underline{a})}{dH} q(a)da$$

2. *Lifetime, Pigouvian: increasing H^* should increase lifetime earnings proportionally to the increase in output:*

$$\tilde{R}_p'(H) = \int_0^1 \frac{dR_p(\tilde{y}(\tilde{h}(\tilde{a})_0^a; H^*)_0^1)}{d\tilde{y}(\tilde{h}(\tilde{a})_0^a; H^*)} \frac{d\tilde{y}(\tilde{h}(\tilde{a})_0^a; H^*)}{d\tilde{h}(a)} \frac{d\tilde{h}(a)}{dH} da = v(H) \int_0^1 q(a) \frac{d\tilde{h}(a)}{dH} da$$

3. *Lifetime, redistributive: Define the retention that workers face as $R_m(R_p(\tilde{y}(\cdot)))$, and $r_m = R_m'(R_p)$. After correcting for distortions, then R_m should satisfy standard Mirrleesian formulas:*

$$\left(\frac{1 - r_m(R_p)}{r_m(R_p)}\right) g(R_p) R_p \epsilon_{\tilde{r}}^c(R_p) = \int_{R_p}^{\infty} g(\tilde{R}_p) \left(1 - \lambda(\tilde{R}_p)\right) d\tilde{R}_p + \int_{R_p}^{\infty} \left(\frac{1 - r_m(\tilde{R}_p)}{r_m(\tilde{R}_p)}\right) g(\tilde{R}_p) \eta_I(\tilde{R}_p) d\tilde{R}_p,$$

Notice that in this Proposition we describe taxes in "layers", and there is a choice in describing taxes as one layer of Pigouvian taxes and another layer of Mirrleesian taxes. An alternative formulation could have one layer of intertemporal taxes, and another double layer of Pigouvian and Mirrleesian taxes. The first would keep lifetime income unchanged as a function of H, and the second would feature an analogous lifetime Pigouvian component and a redistributive component as in Proposition 3. The following remark presents an alternative simple way of describing the Pigouvian component of taxes.

Remark 1. We can define R_p to be such that:

$$v(H)q(a) = \frac{dR_p(\tilde{y}(\cdot)_0^1)}{d\tilde{y}(a)} w(\tilde{h}(\cdot)_0^a) + \int_a^1 \frac{dR_p(\tilde{y}(\cdot)_0^1)}{d\tilde{y}(\tilde{a})} \frac{dw(\tilde{h}(\cdot)_0^{\tilde{a}})}{d\tilde{h}(a)} \tilde{h}(\tilde{a}) d\tilde{a},$$

which, in the case where there are no intertemporal distortions in pre-tax salaries, simplifies to:

$$\frac{dR_p(\tilde{y}(\cdot)_0^1)}{d\tilde{y}(a)} \frac{1}{q(a)} = \frac{v(H)q(a)}{q(a)w(\tilde{h}(\cdot)_0^a) + \int_a^1 \frac{dw(\tilde{h}(\cdot)_0^{\tilde{a}})}{d\tilde{h}(a)} \tilde{h}(\tilde{a}) q(\tilde{a}) d\tilde{a}},$$

where $v(H)$ is the marginal productivity of the type that supplies the level H of labor, and where for ease of notation the dependence on $\tilde{y}(\cdot)_0^1$ is omitted. That is, the formula should be read as a function of earnings flows $\tilde{y}(\cdot)$, through the inverse operator $\tilde{h}(\tilde{y}(\cdot)_0^1)_0^1$.

Notice that correcting for intertemporal distortions is a significantly more complicated endeavor: taxes should be history-dependent, and depend on how much a change in labor supply today translates into higher lifetime earnings, not only through its impact on current earnings, but additionally through its indirect impact on future salaries. For workers of different earnings histories, the tax rate they would face on the next dollar at a given period would depend on their future earnings and past earnings (which can be mapped to their labor supply choices), and how an increase of one unit of their labor supply today would impact current and future earnings, with the latter capturing the private benefits from the signaling effects from that increase in labor supply.

Notice that if there is a tax system in place that already corrects for intertemporal distortions R_p, then Proposition 4 tells us that we are back to the simpler case covered in Proposition 1. Moreover, as a corollary of Proposition 4, we know that if the information structure does not introduce intertemporal distortions, then the simpler optimal taxation formulas 1.5 hold.

Corollary 1. *If the information structure is such that no intertemporal distortions are generated, then if taxes are optimal they satisfy condition 1.5.*

Conditioning on the Full History of Deliverables

The second case of more general endogenous information structures we consider is the case in which firms observe the full history of deliverables, and the detailed timing of the execution of those. That is, for a worker of age a, the firm will see the history of labor supply decisions from the time the worker was born up to their current age a, which is denoted $\tilde{h}(\cdot)_0^a$.[30]

We assume in this section that preferences now take the form $\mathrm{U}(C, \tilde{h}(\cdot), \theta)$, and that θ is high-dimensional. To talk meaningfully about expectations of productivities conditional on some history $\tilde{h}(\cdot)_0^a$, we need some workers to be willing to provide this rich set of histories, which would not be possible for preferences of the form $U(C, H(\tilde{h}(\cdot)), \theta)$, or with a low dimensional type space. To keep this section simple, we additionally assume there is a single consumption good C.[31] We retain the assumption that more productive types are more willing to provide the deliverables.[32]

Under those assumptions, we will show that optimal taxes can also be thought of as the composition of i) corrective taxes that guarantee that, for each additional unit of labor, workers receive lifetime benefits

[30] In this case, workers of different ages are never pooled together, as any history that ends up at age a is different than a history that ends at age a', when $a \neq a'$.

[31] This is without loss, since discount rates $q(\cdot)$ are assumed to be exogenous.

[32] Which in this case can be stated as: if $v(\theta) > v(\theta')$ then for any $\tilde{h}(\cdot)_0^1$, C, $MRS_{C,\tilde{h}(a)}(C, \tilde{h}(\cdot), \theta) < MRS_{C,\tilde{h}(a)}(C, \tilde{h}(\cdot), \theta')$, where $MRS_{C,\tilde{h}(a)}(C, \tilde{h}(\cdot)_0^1, \theta) = -\frac{U_{\tilde{h}(a)}(C, \tilde{h}(\cdot), \theta)}{U_C(C, \tilde{h}(\cdot), \theta)}$.

equivalent to their contribution to output, and ii) redistributive taxes, which, in this case, do not necessarily take the simple Mirrleesian form of the previous section.

Towards showing that, analogously to the previous section, first we establish that if the planner could choose the allocation, while being restricted to the set of incentive compatible allocations, any optimal allocation would lie at the frontier of production possibilities set (Lemma 6 in the Appendix Section 1.5.2). Then, we show that the planner can use Pigouvian taxes to achieve the frontier of the production efficient set of allocations. This will be the case because the planner will be able to infer labor supply choices from the history of earnings, that is, sequences of $\tilde{h}(\cdot)_0^1$ will map to sequences of $\tilde{y}(\cdot)_0^1$ one-to-one, as in Lemma 2 and Lemma 5, from the previous sections.

To show that the planner can infer labor supply choices from the history of earnings, an important intermediate result is that there is a positive signaling return to experience. This result is interesting on its own, and it shows that the career concerns logic carries through this more general environment (see Lemma 7 in the Appendix Section 1.5.2, which is the analogous counterpart to Lemma 4) . This signaling return to experience is driven by the fact that higher productivity types are those who are willing to provide labor supply paths with a larger number of deliverables. That property is used to show that the planner can infer labor supply decisions from earnings histories (see Lemma 8 in the Appendix Section 1.5.2), which in turn is crucial for the ability of the planner to implement allocations at the frontier of production possibilities set. That is, analogously to Lemma 7 and Lemma 4, from the previous Sections, Lemma 8 in this Section allows the planner to decentralize feasible, incentive compatible allocations that lie at the efficient possibilities frontier of the economy with the use of taxes on earnings. In doing that, Pigouvian taxes play an important role, as stated in the following Proposition.

Proposition 5. *The planner can guarantee that the allocation would lie at the frontier of the production possibilities set by using Pigouvian taxes.*

These Pigouvian taxes take the same general form as in the previous section's Remark 1. Thus, although this economy may look quite complicated, the same principles of tax design can be applied. There is a caveat though. Because we have unrestricted preferences, and multiple goods, now the design of optimal redistributive taxes, after correcting for the Pigouvian distortions, is more complicated, and without further normative assumptions, we cannot point to lifetime income taxation as the preferred form of redistribution. From an implementation perspective, analogously to the case analyzed in the previous Section, the structure of taxes can be thought of as a composition of two layers of taxes. First, Pigouvian taxes correct for the distortions. Second, on top of these taxes, another layer of taxes is imposed to take care of redistribution.[33]

1.5.3 Changes in Information Processing Technologies

As new technologies and richer datasets begin to affect the workplace (Chalfin et al., 2016; Autor, 2019; Acemoglu et al., 2020; Bales and Stone, 2020), an important question to consider is the impact of these technological changes in on the workers' incentives to exert effort, on the distribution of income, and on the tradeoffs the government faces when setting taxes. Those changes should affect how workers get paid and how resumes are read, arguably making information imperfections less pronounced. On the other hand, the rising automation of routine tasks (Autor et al., 2003), and the advent of "new work" (Autor, 2019), that is, novel jobs reflecting changes in technology and preferences, have contributed to changes in the task composition

[33] Multidimensional heterogeneity can also make the second layer interact with the first layer, while retaining the general structure of a Mirrleesian formulas. For an example, see Section 1.9.7.

of jobs. As non-routine cognitive tasks become more prevalent, it may become harder to monitor and assess the productivity of workers. As a result, these imperfections in the information transmission process may become more salient. In this section, we look at the consequences of changes in fundamentals that affect information asymmetries in labor markets.

We define information asymmetry problems in labor markets to decrease if deliverables become a better measure of product, affecting how firms pay workers and how firms read resumes. More precisely, we adopt the following definition.

Definition. Let preferences be $U(c,h,\theta) = \tilde{U}\left(c, h/b(\theta)\right)$. If new productivities and tastes are such that $v^n(\theta) = v(\theta)/\Delta(\theta)$ and $b^n(\theta) = b(\theta) \cdot \Delta(\theta)$, $\Delta > 0$ and increasing in θ, order preserving (meaning that if $\theta > \theta'$, then $v(\theta) > v(\theta')$, $v^n(\theta) > v^n(\theta')$, $MRS^{\theta}_{c,h} > MRS^{\theta'}_{c,h}$, and $MRS^{n,\theta}_{c,h} > MRS^{n,\theta'}_{c,h}$), then information asymmetries in labor markets decrease.[34]

Under this definition, decreasing information asymmetries keeps the first-best utility possibilities frontier unchanged. It also increases the ratio $v(\theta)/\mathbb{E}[v(\tilde{\theta})|\tilde{\theta} > \theta]$ for every type of worker, and preserves the relative ranking of types in terms of productivities and marginal rates of substitution.

A perhaps surprising result is that information asymmetries in labor markets increase welfare in this setup, as stated in Proposition 6. This result implies that those frictions are, in a particular way, good for redistribution, and attenuate inequality. That is, if the way in which the government evaluates inequality is expressed in how it sets marginal tax rates, then an increase in information asymmetries in labor markets increases welfare, or, in other words, improves the distribution of outcomes from the point-of-view of a redistributive planner. Information asymmetries help with redistribution, because they make it harder for high productivity workers to separate themselves from low productivity workers. Moreover, because under the assumptions of Lemma 2, (or the more general Lemma 5) the optimal post-tax wages do not depend on the specifics of the model of asymmetric information, the impact on welfare is the same across different models, provided that taxes are set optimally and following the same preferences for redistribution.

In other words, this result does not depend on the particular assumption we imposed on the information structure that determines how salaries are set by firms. It relies on the planner being able to solve for the allocation directly, which is possible whenever there is a one-to-one mapping between earnings and labor supply decisions. This property is satisfied by different information structures as shown in Section 1.5.2. Further, the result also holds across different models of career concerns, provided they satisfy the invertibility conditions behind Lemma 2 or the more general Lemma 5, that allow for labor supply decisions of workers to be perfectly inferred from their history of earnings. In fact, this result is a generalization of Stantcheva (2014) Proposition 13, which compares welfare in an economy under a Miyazaki-Wilson-Spence (MWS) model of the labor market relative to a Mirrleesian economy.[35]

Proposition 6. *If the original tax schedule is optimal and the planner has strong enough redistributive preferences, and leisure is a normal good, then decreasing information asymmetries in labor markets decreases welfare.*

Proof. See Appendix Section 1.9.9. □

[34] Notice that under the assumptions behind Proposition 2, any optimal allocation makes post-tax wages a function of lifetime labor supply. Thus, in this case, it is without loss to specify preferences as a function of lifetime labor supply and lifetime consumption.

[35] Another similar setup where these conditions are satisfied is the Azevedo and Gottlieb (2017) model of competitive screening.

The idea behind this result is that if the planner has redistributive tastes, then it sets positive marginal redistributive taxes to transfer resources from higher types to lower types, an the incentive compatibility constraints bind downwards (Seade, 1982; Werning, 2000). Whenever the degree of informational asymmetry in labor markets decreases, these downward incentive compatibility constraints become tighter, as it becomes less costly for the high types to imitate the low types. With less information asymmetry in labor markets, the high productivity workers can use more of their previously unobserved productivities to imitate the deliverable production of the lower productivity workers.

The impact on taxes is more subtle. When resumes are defined as the cumulative discounted sum of deliverables, the corrective component unambiguously falls for each type, because $\chi(\theta) = \frac{v(\theta)}{\mathbb{E}[v(\tilde{\theta})|\tilde{\theta}\geq\theta]}$ increases proportionally more in the numerator relative to the denominator. In that vein, we can think of an decrease in information asymmetries in labor markets as a force towards lower taxes. To understand how the Mirrleesian component and total marginal taxes would be affected is useful to write optimal taxes in terms of types θ, as in the following proposition.

Proposition 7. *Optimal taxes as a function of types θ must satisfy the following equations:*

$$r_y(\theta) = r_m(\theta) \cdot \chi(\theta)$$

$$\frac{1-r_m(\theta)}{r_m(\theta)} f(\theta)\left(-\frac{\partial \log MRS}{\partial \theta}\right)^{-1} = \int_\theta^\infty (1-\hat{\lambda}(\tilde{\theta}))f(\tilde{\theta})d\tilde{\theta} + \int_\theta^\infty \left(\frac{1-r_m(\tilde{\theta})}{r_m(\tilde{\theta})}\right)\eta(\tilde{\theta})f(\tilde{\theta})d\theta$$

Proof. See Appendix Section 1.9.10. □

Changes in the degree of informational asymmetry affect at least two key ingredients in the formula above χ and $\frac{\partial \log MRS}{\partial \theta}$, and each of them, respectively, affects directly the Pigouvian and the Mirrleesian component of taxes. They also operate very differently. The Pigouvian component of taxes for a type θ is affected by the changes in the unobservable component of the productivity of everyone that ends up with a lifetime income higher than $y(\theta)$. As the type of labor these people supply become more easily measurable, type θ receives smaller implicit subsidies from higher productivity workers, and $\chi(\theta)$ increases, approaching one. The Mirrleesian component on the other hand depends on how type θ is more or less willing to provide more deliverables relative to their local neighboring types. Thus, for any given θ, one can imagine a decrease in informational asymmetries that can have impacts on the Pigouvian or Mirrleesian component of taxes of arbitrarily different magnitudes. That is, the net effect on total marginal taxes, taking into account both the Pigouvian and Mirrleesian components, is in general ambiguous.

A simple example shows that we may expect the Pigouvian component to dominate over the Mirrleesian component in certain circumstances. Assuming, there are no income effects, and that preferences take the simple form as in the example from Section 1.3.1, and types are Pareto distributed, we can solve for optimal marginal retention as a function of types as:

$$\chi = \frac{\alpha-\delta}{\alpha}, \quad \frac{1-r_m}{r_m} = \frac{1-\bar{\lambda}}{\frac{\epsilon}{1+\epsilon}\frac{\alpha}{1-\delta}} \quad \Longrightarrow r = r_m \cdot \chi = \frac{\frac{\epsilon}{1+\epsilon}\frac{\alpha-\delta}{1-\delta}}{(1-\bar{\lambda}) + \frac{\epsilon}{1+\epsilon}\frac{\alpha}{1-\delta}}$$

where $\bar{\lambda}(\theta) = \mathbb{E}[\hat{\lambda}(\tilde{\theta})|\tilde{\theta} \geq \theta]$ denotes the marginal value of a one dollar transfer to the types above θ, which is assumed to be constant in θ starting from some level $\hat{\theta}$. χ is decreasing in the degree of informational asymmetry, thus corrective taxes increase as informational asymmetries increase. r_m is increasing in the degree of informational asymmetry, and therefore the redistributive component of taxes decrease. Provided that $\frac{\alpha}{1+\epsilon} \geq 1$ (which incidentally guarantees that output is finite), holding fixed this marginal value of transfers $\bar{\lambda}$, marginal retention r decreases with the degree of informational asymmetry δ, that is, marginal taxes increase as the degree of informational asymmetry increases. In this case, thus, the Pigouvian component dominates over the Mirrleesian component, calling for higher marginal taxes as the degree of informational asymmetry increases.

1.6 Empirical Evidence

The previous section has offered an understanding of the tradeoffs a government faces between incentives and redistribution through the lens of simple sufficient statistics. In this sense, a key new element that needs to be estimated is the ratio of marginal productivities over salaries for the last unit of labor that workers supply.[36]

1.6.1 Data

To evaluate the magnitude of the signaling return to experience and the Pigouvian component of taxes we use Health and Retirement Study data. The survey is a representative sample of the US population older than 50 and a biannual panel covering the period from 1992-2018. It follows around 20,000 workers and it is rich on covariates, including job histories, hours worked, education, cognition, measures of lifetime income, and geographic, industry, and occupation variables.

There were several state and federal tax reforms over the period, which we are going to explore as a source of exogenous variation in wages. The federal tax reforms include the Omnibus Budget Reconciliation Act of 1993, which affected mostly top income earners; the Economic Growth and Tax Relief Reconciliation Act of 2001, which affected those at the bottom and at top of the income distribution; the Jobs and Growth Tax Relief Reconciliation Act of 2003, which affected tax rates for middle and top income earners; the American Recovery and Reinvestment Act of 2009, with tax changes across different parts of the income distribution; and the American Taxpayer Relief Act of 2012, which changed tax rates at the top of the income distribution.

There were also several state tax reforms. These were widely dispersed across the US, as shown in Figure A1, although they were more prevalent in some states such as California, Connecticut, Delaware and Idaho, and were almost completely absent in Alaska, Florida, Nebraska, North Dakota, South Dakota, Tennessee, Texas, Washington, and Wyoming. A more detailed description of state tax reforms is presented in the Appendix Sections 1.10.2 and 1.10.5.

To explore the variation induced by those changes in marginal rates, we construct simulated changes in marginal tax rates at initial incomes using the NBER tax simulator. We assign a wide range of income, consumption and demographic variables from HRS to the 32 inputs in the NBER tax simulator, adapting and extending to our period of analysis (1992-2018) the procedure developed by Pantoja et al. (2018).[37]

1.6.2 Empirical Strategy and Results

The key statistic we would like to estimate is the ratio of the marginal productivity of workers at retirement over their salaries, as a function of their lifetime income. Our approach relies on tax changes as

[36] More generally, we would like to estimate the difference in marginal productivities and the sum of current salaries and all future earnings increases that result from supplying one unit of labor, at each point in the career of the workers, as discussed in Section 1.7. While estimating marginal productivities over salaries is a considerably hard challenge, estimating the latter is an even harder challenge. Direct attempts at estimating the latter are left for future work.

[37] A more detailed description of the mapping between the variables is presented in Appendix Section 1.10.1.

a source of exogenous variation in wages, to quantify the degree of informational asymmetry. The key idea is that, looking at the labor markets around the time of retirement, the change in average productivities induced by the exogenous variation in wages is informative of the marginal productivity of those who are almost indifferent between working more or retiring.

This approach builds on the literature that has quantified the degree of adverse selection in markets described by simple Akerlof (1970) lemons conditions, such as Einav et al. (2010). In the context of health insurance markets they analyze, the key argument is that with a source of exogenous variation in prices, one can non-parametrically trace the shape of the cost curve for the insurance contract by looking at the average cost as a function of prices. The average cost, for a given price, can be inferred from data on insurance claims. Translated into our context, with a source of exogenous variation in wages (for a specific labor contract), one can non-parametrically trace the shape of the labor demand curve by looking at average productivities as a function of salaries.

However, inferring productivities in labor markets is a more difficult endeavor than inferring costs in insurance markets. In insurance markets, detailed data on insurance claims can be used to compute reasonably precise measures of expected costs for the insurance contracts. In labor markets, most often direct data on productivities is not available. To circumvent this challenge, two complementary approaches are adopted. The first assumes that labor markets are competitive and therefore leverages the observation that wages would be equal to the average productivity of workers with the same resume, as in the model presented in this paper. The second takes advantage of the rich set of covariates available from the Health and Retirement Study and looks at cognitive measures as proxies for productivities.

Salary Changes and Retirement Decisions

At the general level, to infer the degree of information asymmetry from salary changes, consider an increase in marginal taxes for those who are near retirement and how it affects a pool of workers with equivalent resumes and the same original salaries. This tax change may induce some of the individuals in this pool to actively retire. The pool of workers who remain in the workforce in the next period is different, because it no longer contains those who were almost indifferent between retiring or not. If those who were almost indifferent were also less productive in the unobservable dimension, as in the model, then wages should increase. Moreover, by observing wages before and after the change in marginal rates, we can find the average productivity of those who are close to being indifferent between retiring or not. That is, we can decompose wages before and after as a weighted sum of the productivites of those who are almost indifferent and those who are not, that is:

$$\mathbb{E}[v_{before}] = \mathbb{E}[v_{after}](1 - s_{mg}) + v_{mg}s_{mg},$$

where, $\mathbb{E}[v_{before}]$ and $\mathbb{E}[v_{after}]$ denotes the average productivity (or salaries) of the workers with a common resume, respectively before and after the tax increase; s_{mg} is the share of people who are induced to retire by the increase in marginal taxes. This simple relationship can be written in terms of labor wedges (χ), elasticities of salaries (ϵ_r^w) and semi-elasticities of labor market participation (η_r^p). That is:

$$\chi = \frac{v_{mg}}{\mathbb{E}[v_{after}]} = 1 - \frac{\frac{\mathbb{E}[v_{after}] - \mathbb{E}[v_{before}]}{\mathbb{E}[v_{after}]}}{s_{mg}} \implies \chi = 1 + \frac{\epsilon_r^w}{\eta_r^p}$$

Thus, from observing salaries before and after, and the share of people who retire as a result of a small tax increase, one can infer the productivity of the people who are marginally indifferent between retiring or not. Labor wedges χ, thus, can be inferred from elasticities of salaries (ϵ_r^w) and semi-elasticities of labor market participation (η_r^p). Data from the Health and Retirement Study is particularly well suited for recovering these elasticities, as it includes carefully calculated measures of lifetime income, the timing of retirement decisions, salaries, hours of work, and detailed income measures allowing us to approximate the marginal tax rates faced by workers using the NBER taxsim model.

While this explanation has focused on salary levels, and a simple before and after comparison, in practice we will be looking at salary changes at the individual level, and we will be pooling together different tax reforms, across different state, time periods, and parts of the income distribution, while at the same time controlling for year-fixed effects, initial hourly wages, and marital status. The key identification assumption for the results that rely on changes in salaries is that future productivity changes, and elasticities of participation, are independent of each other and of tax changes conditional on the set of controls.[38]

The Pigouvian externality ($\chi - 1$) can also be interpreted as the coefficient of a two sample instrumental variable regression of changes in log salaries on changes in participation, where changes in taxes are the instrument for changes in participation. Intuitively, the coefficient on that regression tells how salaries change when the marginal worker is forced to stay in the labor force, and thus is informative of their productivities.

To obtain estimates for the elasticities of wages, we regress changes in log hourly salaries on simulated changes in the log of marginal retention rates, including different sets of controls X_{it}, as in equation 1.6. These controls aim to capture i) the possibility that changes in the tax schedule and in hourly wages may both respond to business cycles fluctuations (thus the inclusion of year-fixed effects), ii) the possibility that tax changes may have targeted different income groups and wages may evolve differently for those different groups (thus the inclusion of log hourly wages, and other non-linear functions of hourly wages), and iii) similarly, tax changes may have targeted differentially people of different marital status, for whom wages may evolve differentially as well (thus the inclusion of marital status indicator variables).

$$\Delta \log w_{it} = \epsilon^w \Delta \log r_{it} + \gamma' X_{it} + u_{it} \tag{1.6}$$

The results of this set of regressions are presented in Figure 1.11 and Table 1.3 in the Appendix Sections 1.10.4 and 1.10.5. For the main specification, which includes year fixed effects, marital status dummies, and a 10-piece linear spline on hourly wages as controls, the estimated elasticity of wages is -0.16, with a standard deviation of 0.1, implying that a 1% increase in marginal retention between years 0 and 2 causes a 0.16% decrease in salaries between years 0 and 4. The effects are stronger at the 6-year horizon, with a point estimate of -0.27, and revert back at the 8-year horizon to -0.14, when estimates also get noisier. This is in line with the idea that marginal tax increases push the people who were almost indifferent between retiring or not into retirement, and those people are less productive than the average worker although they were receiving the same salaries. Salaries then would increase as employers realize that there was a change in the productivity composition of the pool of workers still on the labor force.

It is a common wisdom that workers experience most of their salary changes when they change jobs. To more precisely capture the effects of changes in marginal tax rates on the wages of workers, we also consider the effects of restricting the sample to only those who switch jobs over the relevant time period. These results are presented in Figure 1.12 and Table 1.4 in the Appendix Section 1.10.4. These elasticities are higher in

[38] This assumption becomes weaker when we estimate heterogeneous elasticities by different groups as in 1.6.2, as then only within groups the elasticities of participation need to be independent of future productivity changes.

magnitude, in line with the idea that wages are more flexible when workers switch jobs. The results from the main specification (including the full set of controls and a 10-piece spline on hourly wages) imply a 1% increase in marginal retention causes wages to fall by 0.34% over the 4-year horizon, and by 0.43% over the 6-year and 8-year horizons.

To obtain estimates for the semi-elasticities of participation, analogously, we regress changes in participation[39] on simulated changes in the log of marginal retention rates, including different sets of controls, as in equation 1.7. Again, these controls aim to capture the possibility that changes in the tax schedule and in labor market participation may both respond to business cycles fluctuations (thus the inclusion of year-fixed effects), the possibility that tax changes may have targeted different income groups and labor market participation may evolve differently for those different groups (thus the inclusion of log hourly wages, and other non-linear functions of hourly wages), and similarly, tax changes may have targeted differentially people of different marital status, for whom labor market participation may evolve differentially as well (thus the inclusion of marital status indicator variables).

$$\Delta p_{it} = \eta_r^p \Delta \log r_{it} + \gamma' X_{it} + u_{it} \tag{1.7}$$

The estimated semi-elasticities of participation when including the full set of controls and the 10-piece spline on hourly wages are of the order of 0.10 at the 2-year horizon, 0.01 at the 4-year horizon and 0.03 at the 6-year horizon, implying that at the 4-year horizon, a one percent increase in marginal retention causes a 1 percentage point decrease in the probability of a worker getting out of the labor force. The relatively larger effects at the shorter horizons when compared to the elasticity of wages is consistent with the idea that, after a tax increase, first some of the workers drop out of the labor market, and then, as the employers learn that the pool of the remaining workers is more productive, wages gradually increase as time passes.

Our estimates are in line with a literature that has found that the implied tax rates on labor income from public pension rules have large disincentive effects on work (Gruber and Wise, 1998; Coile and Gruber, 2007). On the other hand, they are higher than the substitution elasticities inferred from variation in Social Security benefits from the 1977 Social Security Act that created the so called "Notch generation" Gelber et al. (2016). While our confidence intervals are relatively wide, the difference can also be explained by the fact that the substitution incentives of their reform, and of Social Security benefits more generally, are relatively more opaque (Blinder et al., 1980) than the substitution incentives from income tax reforms.

Taking the ratio of the estimated coefficients $\frac{\epsilon_r^w}{\eta_r^p}$ (while multiplying η_r^p by one hundred so numerator and denominator are in the right units) we obtain estimates for the magnitude of the labor market informational externality $(1 - \chi)$. These results are presented in Tables 1.6 and 1.7. For the main specification, which includes the full set of controls and the 10-piece spline on hourly wages in both the participation and wages regressions, and looks at the effects over a 4-year horizon, the estimated negative informational externality is around 0.16. In other words, workers are paid around 16% more than their marginal productivity for the last unit of labor they supply.

Cognitive Measures: Inspecting the Mechanism

There could be other stories that explain why wages increase and participation falls when there are tax increases. Most notably, firms' labor demand may be partially elastic. In order to provide further evidence that the pool of workers is playing a role in salary changes, we look at cognitive measures collected by the

[39] That is, changes in the indicator variable that is equal to one whenever the individual is working.

Health and Retirement Study. The average cognitive measure of the individuals working is lower after a tax increase, even when that measure is taken before the tax change.

Mental status scores in the Rand harmonized longitudinal files from the Health and Retirement Survey are computed as the sum of vocabulary, naming, and counting scores from the HRS. Those scores are the sum of correct answers from questions ranging from "Who are the current president and vice-president of the United States?" to "How much is 100 minus 7? How much is that minus 7? [...]." The detailed construction of this variable is presented in Appendix 1.10.3. Those measures can be seen as a proxy for ability, similarly to how Armed Forces Qualification Test (AFQT) scores in the National Longitudinal Survey of Youth is often used as a measure of ability (Farber and Gibbons, 1996; Altonji and Pierret, 2001; Lange, 2007; Craig, 2020). While it has the disadvantage of being less detailed, it has the advantage of being assessed repeatedly for each respondent, at every survey year, and for that reason may be a more accurate measure of ability if ability is not constant but evolves dynamically over time.

Looking at the HRS total mental status scores as a proxy of ability, we regress those scores as measured two years before the baseline year, on changes in marginal retention and a set of control variables (X_{it}, including year fixed effects, marital status and flexible functions of hourly wages), restricting the sample to those who are working in the baseline year and four years in the future. The coefficient η_r^m on regression 1.8 translates how different are the average mental status scores of those who are working after changes in marginal retention.

$$scores_{it} = \eta_r^m \Delta \log r_{it} + \gamma' X_{it} + u_{it} \tag{1.8}$$

The results for different sets of controls are presented in Table 1.10 in the Appendix Section 1.10.4. Fixing the set of controls, the results for different time horizons are presented in Figure 1.14 in the Appendix Section 1.10.5. For a horizon of four years after the baseline year, under the most stringent specification, a one percent increase in marginal retention between the baseline year a two years ahead a change in the composition of the pool of workers such that average mental status scores (as measured before the tax change) decrease by 1.3 points (out of 15), with a standard deviation of 0.5, conditional on hourly salaries. This effect is in line with the mechanism emphasized in this paper, where salaries change as a response to changes in the productivity composition of the workers who are are willing to supply the deliverables given the current incentives, and where salaries and taxes work as screening devices both for the government and for the firms.

Heterogeneity Across Income Levels

The argument made on Section 1.6.2 is built on observing salaries and cognitive measures for the workers who have the same equivalent resumes, and thus would face the same remuneration for their next unit of labor they would supply. However, the elasticities of wages and the semi-elasticities of participation may be heterogeneous across different resumes and labor contracts.

To address that heterogeneity, we estimate equations 1.6 and 1.7 locally as a function of hourly salaries,[40] using local polynomial methods. That is, for different dependent variables dep_{it}, regression equations as in 1.9, where y denotes an hourly wage level. In those regressions, observations are weighted by their distance from the hourly wages level y where the equation is being evaluated using the Epanechnikov kernel, and an optimal bandwidth selected using a leave-one-out cross validation procedure. An additional cross-term

[40] Even within remuneration levels, there could be heterogeneity in elasticities and labor market wedges, in which case the relevant optimal taxation formulas also call for the estimation of correlation among those, as in the tax formula 1.10. For evidence of further heterogeneity across education groups and occupations, see the Appendix Section 1.10.4.

$\Delta \log r_{it}(y_{it} - y)$ is included to improve on the bias-variance tradeoffs, as explained in more detail in Fan and Gijbels (1996). Optimal bandwidths are selected with the leave-one-out cross validation procedure proposed by Racine (1993). Bootstrap confidence intervals are generated using the basic bootstrap method described in Chapter 5 of Davison and Hinkley (1997).

$$dep_{it} = \epsilon_r(y)\Delta \log r_{it} + \beta(y)\Delta \log r_{it}(y_{it} - y) + \gamma(y)'X_{it} + u_{it} \tag{1.9}$$

The local results are presented in Figures 1.15 to 1.17 in the Appendix Section 1.10.5. They show that elasticities of wages are higher in magnitude for high-earners, which also have lower participation semi-elasticities. Moreover, semi-elasticities of mental status scores are also higher in magnitude for high-earners. This is in line with the idea that informational imperfections are a larger issue for high-earning occupations and jobs. The point estimate for the elasticities of wages suggest that at the top of the distribution of hourly salaries the labor wedges could be very high, but also are imprecisely estimated. The lower bound on the confidence interval at the 90th percentile would rule out a value lower the 0.5, implying that those workers could be paid more than twice their marginal productivities. However, these values come from a combination of high elasticities of hourly wages and low participation elasticities, which approach zero, raising concerns about the validity of the bootstrap confidence intervals. The general message however, is that we should expect the labor wedge χ to be decreasing in income, and it is reasonable to expect values ranging from 0.9 to less than 0.5, where 0.9 is the estimated χ for the upper third of the income distribution.

1.6.3 Comparison to Existing Evidence

In this section, we compare our estimates to the available evidence on the time patterns of salaries and signaling on labor markets. We show that a back-of-the envelope calculation using the existing evidence from the literature would result in a similar magnitude for the Pigouvian component of taxes we found in the previous section.

There is documented evidence that workers experience large growth rates of salaries as a function of experience. In fact, Guvenen et al. (2021) have documented that the top 1% earners have a very steep growth rate of salaries, of around 2700% over a 30-year period, or 11.3% per year (and approximately 3% per year on average across workers). While part of this pattern may be thought as the result of human capital accumulation, and another part of it may be thought of as the result of pure luck, it is reasonable to expect that at least another part of it is due to selection and the career concerns logic we have uncovered. In fact, Guvenen et al. (2021) argue that no empirically plausible model of stochastic productivities could explain the large growth rates of salaries observed at the top. Moreover, there is evidence that signaling and learning are important to explain the dynamics of salaries and tenure in some occupations. For example, Cella et al. (2017) have shown that the relationship between the volatility of stock returns and the tenure of CEOs of large US firms is consistent with the idea that the market gradually learns about the CEO ability throughout the years of the CEO tenure.

For the purpose of this paper, the key question is what share of the growth rate of salaries is due to signaling and how can it be translated into an estimate of the corrective component of taxes. However, there is no direct estimate in the literature that can readily be used to answer these questions. To get a sense of what would be reasonable magnitudes for the signaling component of the growth rate of salaries, we can look at the evidence on the return to schooling and the role of signaling in that context. Out of the return to schooling, recent work has concluded that on average 30% can be attributed to signaling and 70%

to human capital accumulation (Aryal et al., 2019). Combining the growth rate of salaries for top earners from Guvenen et al. (2021) with the signaling fraction of the return to schooling from Aryal et al. (2019), we can guess that the return to experience due to signaling, for top earners, may be of the order of 3.4% (30% of 11.3%) per year at 10 years of experience. To translate those numbers to a magnitude for the corrective component of taxes, using the free entry condition 1.3, we can show that the return to experience is related to the Pigouvian component of taxes $(1-\chi)$, and the shape of the lifetime income distribution by the formula $\gamma = \frac{\alpha_y(1-\chi)}{1+\alpha_y(1-\chi)}$. This implies that, for the values above, the Pigouvian component of taxes at the top that could be as high as 25%, and on average around 6%.

1.7 Richer Type Space, Signal Structure and On-the-job Learning

To add realism to an otherwise stylized model, several extensions to the basic model are presented. These extensions include human capital accumulation, richer heterogeneity in elasticities, and richer signal structures. Key insights from the generalized optimal taxation formulas and their empirical implications will hold with some caveats in those extensions.

The first extension allows human capital accumulation in the form of learning-by-doing as in Arrow (1962), and relaxes the assumption that workers have constant productivities over their lifetimes, which is an evidently implausible assumption. It is reasonable to expect that at least some part of the return to experience observed in the data is due to increases in productivity due to on-the-job learning, or training efforts. However, the extended model shows that this assumption is to some extent innocuous. Although it complicates the relationship between the observed return to experience, the degree of information asymmetry in the market and the rate of human capital accumulation, the same optimal tax formula applies when on-the-job learning is costless.

The second extension allows for richer heterogeneity in elasticities. This extension is motivated by the empirical evidence that there is substantial heterogeneity in how people respond to taxes, even within tax brackets (Eissa and Liebman, 1996; Gruber and Saez, 2002; Blau and Kahn, 2007; Vere, 2011; Sturm and Sztutman, 2021). The key modification that multidimensional types introduce is that now, at a given income level, it matters not only how much a worker produces per unit of pretax income, but how these are correlated with the elasticities. Intuitively, an increase in post-tax salaries at a given income will affect people of different elasticities differently, and production will increase proportionally to the product of elasticities and productivities, and therefore will increase more if the elasticities and the unobserved productivities are positively correlated.

The third extension allows firms to see additional signals the government does not see. For example, from the point of view of the employers it may be clear that some workers are on different career tracks, and that those workers can hardly change that. But for the government, it may be hard to distinguish them, or it may be hard to codify those distinctions into the tax system in a way that cannot be manipulated. In this case, optimal taxes are described by a weighted version of the basic taxation formula, where the weights are given by the sensitivity of the different post-tax retention functions to changes in marginal taxes. This modification of the optimal tax formulas follows from the fact that there are multiple career tracks, there are multiple pretax salary functions, while there is a single nonlinear taxation instrument the government can use. Considering a variation in the income retention schedule and tracking how this variation affects the post-tax salaries of different careers in response to it results in the weighted version of the basic optimal tax formula. The effects of changes in marginal taxes on post-tax salaries are attenuated by changes in pre-tax salaries whenever there are information asymmetries. For this reason taking into account these different career tracks may attenuate the magnitude of the corrective component of taxes.

The fourth extension, similarly considers the possibility that other functions of the detailed timing of the completion of tasks may be observed by firms, while the government could see the history of earnings. This possibility makes the signaling return to experience more involved, and introduces the possibility that without taxes there may be intertemporal distortions. Under common preferences over the timing of labor and consumption decisions, the optimal tax system can be written in a way where taxes that depend on the history of earnings would correct for this distortion, and on top of these taxes, optimal lifetime income taxes that are described by the same optimal lifetime income taxation formulas.

1.7.1 On-the-job Learning

The assumption that workers have a constant productivity over their lifetimes is, of course, extreme. It is reasonable to expect that at least some part of the return to experience observed in the data is due to increases in productivity due to on-the-job learning, or training efforts. A simple way to accommodate these concerns is to allow for productivities to depend not only on the types of workers but also on experience itself, that is $v = v(\theta, h)$. Indeed, one of the key reasons why employers may focus on experience as a signal for the productivity of workers is exactly because accumulating experience may directly increase the productivity of workers.

Turning back to the example from Section 1.3.1, a simple way to enrich the setup is to assume that that productivities increase proportionally with experience, that is, productivities are $v(\theta, h) = \tilde{v}(\theta)h^{\beta}$. Then, because wages are the expectation of productivities conditional on h, $w(h) = \mathbb{E}[v(\tilde{\theta}, h)|\tilde{\theta} \geq \theta(h), h] = h^{\beta}\mathbb{E}[\tilde{v}(\tilde{\theta})|\tilde{\theta} \geq \theta(h)]$, and the return to experience would have two components, a signaling and human capital accumulation component. Salaries still satisfy a log-linear relationship, with a coefficient that is the sum of the signaling $\frac{\delta}{1-\delta+\epsilon}$ and the human capital component β:

$$log(w) = \left(\frac{\delta}{1-\delta+\epsilon} + \beta\right) \cdot log(h) + \left(\frac{(1-\delta)(1+\epsilon)}{(1-\delta)(1+\epsilon)-\delta\epsilon}\right) \cdot log\left(\frac{\alpha}{\alpha-\delta}\right)$$

More generally, while human capital accumulation very much changes the meaning of the return to experience, the next proposition, focusing on the simpler case where the resume is defined as the cumulative discounted sum of deliverables, shows that the necessary condition for the optimality of taxes from 1.5 is unchanged.

Proposition 8. *Suppose productivity depends on experience and the unobserved types of workers. Then we can write the planner's problem as:*

$$\max_{\tilde{R}(h)} \mathbb{E}[W(V(\tilde{R};\theta))] \; s.t. \; \mathbb{E}\left[\int_0^{h(\tilde{R},\theta)} (v(\theta,\tilde{h}) - \tilde{r}_h(\tilde{h}))d\tilde{h} - I\right] \geq 0$$

And a necessary condition for a tax schedule to be optimal is given by the following formula (which is analogous to 1.5):

$$\left(\frac{\chi(y)-r(y)}{r(y)}\right)\epsilon_r^c(y)g(y)y = \int_y^{\infty}\left(1-\lambda(\tilde{y})\right)g(\tilde{y})d\tilde{y} + \int_y^{\infty}\left(\frac{\chi(\tilde{y})-r(\tilde{y})}{r(\tilde{y})}\right)\eta_I(\tilde{y})g(\tilde{y})d\tilde{y},$$

where $\chi(y) = \frac{v(y)}{y'(h(y))}$, and $v(y) = v(\theta(y), h(\theta(y)))$ is the productivity at retirement of the worker with lifetime income y. While the formula looks exactly the same, there is a subtle but important distinction. We cannot use estimates of productivity of workers far from retirement to infer their productivity at retirement, as these two quantities can be significantly different. Second, we cannot infer the degree of degree of informational asymmetry by only looking at the return to experience. For these reasons, the empirical strategies aimed at estimating $\chi(y)$ applied in the previous section are designed to be robust to these concerns, and to aim precisely at disentangling the return to experience coming from human capital accumulation from the return to experience coming from employers learning about productivities through job histories.

1.7.2 Heterogeneity in Elasticities

The main version of the model presented features a single-dimension of heterogeneity. Types of different productivities and willingness to provide the deliverables sort themselves into different lifetime income levels, and within lifetime income levels there is no heterogeneity.[41] However, it is reasonable to expect that productivities, and elasticities might be heterogenous within lifetime income levels. Although the standard Mirrleesian first order condition is basically unchanged when agents have heterogeneous elasticities (Scheuer and Werning, 2016; Jacquet and Lehmann, 2021; Bierbrauer et al., 2020; Sturm and Sztutman, 2021), when there is imperfect information this heterogeneity creates a subtle interaction between the Pigouvian and Mirrleesian component of taxes in a way that is reminiscent of Diamond (1973). Intuitively, when taxes increase at a given bracket, different workers respond differently, while simultaneously facing different labor wedges. Thus, how the correlation of elasticities and labor wedges matter for the total effect of the imperfect information externality, or the total impact on the resource constraint of the economy. This is formally presented in the following proposition, where a more general version of the tax formula 1.5 is presented, focusing on the simpler case where the resume is defined as the cumulative discounted sum of deliverables.

Proposition 9. *If a tax schedule is optimal then it needs to satisfy the following relationship:*

$$\mathbb{E}\left[\left(\frac{\chi(y) - r(y)}{r(y)}\right)\epsilon_r^c(y)\right]g(y)y = \int_y^\infty g(\tilde{y})\left(1 - \mathbb{E}[\lambda(\tilde{y})]\right)d\tilde{y} + \int_y^\infty \mathbb{E}\left[\left(\frac{\chi(\tilde{y}) - r(\tilde{y})}{r(\tilde{y})}\right)\eta_I(\tilde{y})\right]g(\tilde{y})d\tilde{y}, \quad (1.10)$$

where $\chi(y) \equiv v(y,\theta)/y'(h(y))$, that is, how much more product is generated per unit of wages for someone who is currently earning y, which conditional on the income level can still depend on the type θ.

Proof. See Appendix Section 1.9.13. □

If there are no income effects, we can write: $r(y) = r_m(y) \cdot r_d(y)$, where $r_d(y) = \frac{\mathbb{E}[\chi(y)\epsilon_r^c(y)]}{E[\epsilon_r^c(y)]}$, and $\left(\frac{1-r_m(y)}{r_m(y)}\right)\mathbb{E}[\epsilon_r^c(y)]g(y)y = \int_y^\infty g(\tilde{y})\left(1-\mathbb{E}[\lambda(\tilde{y})]\right)d\tilde{y}$. In this case, the formula looks very similar to Diamond (1973) equation 10 in the context of a model of Pigouvian taxation with limited instruments, linear utilities in income and separable in externalities. As in that paper, the elasticity weights how much the externality matters. However, here there are no further multiplier effects: what we call externalities in this model – the fact that workers appropriate more than their marginal products when they work more – does not impact the labor supply of others except to the extent that salaries change.

[41] Indeed, Sturm and Sztutman (2021), among others, have shown that there is substantial heterogeneity in elasticities within (annual) income levels.

1.7.3 Richer Signal Structure, Exogenous Signals

The baseline version of the model features a very simple signal structure. This simplicity allows us to clearly understand how incentives and the distribution of income interact when there are career concerns. However, it leaves out important elements of real labor markets, such as richer signals that firms can extract from workers. For example, from the point-of-view of firms it may be very clear that some workers are on different career tracks, and that those workers can hardly change that. For the government, it may be harder to distinguish these workers, and even harder to create taxes that are specific to each career track. This possibility can be accommodated by introducing exogenous signals that the only firms and not the government sees.

If firms see additional exogenous signals the government does not see, then we cannot apply proposition 2. The reason behind it is that there are multiple pretax salary functions, a different career path for each signal realization. However, we can still consider a variation a in the income retention schedule and track how this variation affects the salaries of different careers change in response to it. The following proposition uses this idea to derive comparable tax formulas in this more complex environment, but focusing on the simpler case where the resumes are defined as the cumulative discounted sum of deliverables.

Proposition 10. *If a tax schedule is optimal then it needs to satisfy the following relationship:*

$$\mathbb{E}_z\left[\int_y^\infty \left(\frac{\chi(\tilde{y},z)-r_{\tilde{y}}}{r_{\tilde{y}}}\right)\epsilon^c_{r_{\tilde{y}}}(\tilde{y},z)g(\tilde{y}|z)\tilde{y}\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{y}\right]=\mathbb{E}_z\left[\int_y^\infty\int_{\tilde{y}}^\infty\left(1-\lambda(\tilde{\tilde{y}};z)\right)g(\tilde{\tilde{y}}|z)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{\tilde{y}}d\tilde{y}\right]$$
$$+\mathbb{E}_z\left[\int_y^\infty\int_{\tilde{y}}^\infty\left(\frac{\chi(\tilde{\tilde{y}},z)-r_{\tilde{\tilde{y}}}}{r_{\tilde{\tilde{y}}}}\right)\eta^h_I(\tilde{\tilde{y}})g(\tilde{\tilde{y}}|z)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{\tilde{y}}d\tilde{y}\right]$$

where $\chi(\tilde{y},z)=\frac{v(h(\tilde{y},z);z)}{y'(h(\tilde{y},z);z)}$, *and* $\frac{dr_{h(\tilde{y};z)}}{dr_y}$ *is the response of post-tax salaries of someone who initially earn income* y, *and who gets the signals* z.

Proof. See Appendix Section 1.9.14. □

In this case the optimal policy described by a weighted version of the standard first order condition, where weights are given by how much post tax wages change when income taxes change at different career paths. Taking into account this heterogeneity across career tracks may attenuate the size of the corrective component of taxes. For example, if there are no income effects, then $\frac{dr_{h(\tilde{y};z)}}{dr_y}=0$ for $\tilde{y}\neq y$, and the formula above reduces to:

$$\mathbb{E}_z\left[\left(\frac{\chi(y,z)-r_y}{r_y}\right)\epsilon^c_{r_y}(y,z)g(y|z)y\frac{dr_{h(y;z)}}{dr_y}\right]=\mathbb{E}_z\left[\frac{dr_{h(y;z)}}{dr_y}\int_y^\infty\left(1-\lambda(\tilde{y};z)\right)g(\tilde{y}|z)\cdot d\tilde{y}\right]$$

For any signal realization$z$, the weights $\frac{dr_{h(y;z)}}{dr_y}$ are equal to one if there is no information asymmetry and are less than one if there is any informational asymmetry.[42] In this sense, relative to the basic Equation 1.5, holding fixed the other sufficient statistics, taking into account this source of heterogeneity would call for relatively lower marginal taxes, attenuating the intensity with which informational asymmetries push towards higher marginal taxes.

[42]For more details on how changes in salaries and the degree of informational asymmetry $\chi$are related see Section 1.6.2.

1.8 Conclusion

While incomplete information is a key feature of labor markets, standard benchmark models of taxation often ignore it. In this paper, we developed a simple model that allows job histories and resumes to play this informational role, with firms using them to predict productivities and forward looking individuals making labor supply decisions that anticipate the impact of these decisions today on future wages. In this model, the interest from firms on learning the productivity of workers arises even when firms are allowed to pay-for-performance and both workers and firms are risk neutral.

Moreover, we incorporated optimal taxation in this model, deriving generalized Mirrleesian formulas that apply not only to this particular model of imperfect information in the labor market, but more generally to models with labor market frictions, as long as the mechanism that explains how firms set wages satisfies some simple conditions. Furthermore, the main insights from the optimal taxation formulas hold under several extensions of the basic model, including richer signal structures, human capital accumulation, and multidimensional heterogeneity. These generalized formulas, of independent interest, can be applied to other setups, such as health insurance and financial markets, where taxes may need to play a dual role: correct for informational frictions or other sources of externalities, and redistribute between different types of workers.

These formulas decompose optimal taxes into two components: a redistributive component, and a corrective component. While a large non-linear income taxation literature has explored and estimated the statistics that appear in the redistributive component of taxes, there is limited work estimating the second, especially in the context of dynamic imperfect information in labor markets. Using data from the Health and Retirement Study survey this paper has shown that for an average worker, the corrective component of taxes is of the order of 5%, while for high earners it ranges from 10% to as high as 60%. This result has implications the redistributive effects of the tax system, and is consistent with the view that the current tax system may be less redistributive than it would have been thought, if imperfect information was not taken into account.

Bibliography

(2018). Rand hrs longitudinal file 2018 (v1) public use dataset. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA.

Acemoglu, D. (2021). Harms of ai. Technical report, National Bureau of Economic Research.

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2020). Ai and jobs: Evidence from online vacancies. Technical report, National Bureau of Economic Research.

Acemoglu, D. and Pischke, J.-S. (1998). Why do firms train? theory and evidence. *The Quarterly journal of economics*, 113(1):79–119.

Acemoglu, D. and Pischke, J.-S. (1999). The structure of wages and investment in general training. *Journal of political economy*, 107(3):539–572.

Akerlof, G. (1976). The economics of caste and of the rat race and other woeful tales. *The Quarterly Journal of Economics*, pages 599–617.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.

Alchian, A. A. and Demsetz, H. (1972). Production, information costs, and economic organization. *The American economic review*, 62(5):777–795.

Ales, L., Kurnaz, M., and Sleet, C. (2015). Technical change, wage inequality, and taxes. *American Economic Review*, 105(10):3061–3101.

Ales, L. and Sleet, C. (2016). Taxing top ceo incomes. *American Economic Review*, 106(11):3331–66.

Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.

Altonji, J. G., Smith Jr, A. A., and Vidangos, I. (2013). Modeling earnings dynamics. *Econometrica*, 81(4):1395–1454.

Altuğ, S. and Miller, R. A. (1998). The effect of work experience on female wages and labour supply. *The Review of Economic Studies*, 65(1):45–85.

Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, 29(3):155–173.

Aryal, G., Bhuller, M., and Lange, F. (2019). Signaling and employer learning with instruments. Technical report, National Bureau of Economic Research.

Atkinson, A. B. and Stiglitz, J. E. (1976). The design of tax structure: direct versus indirect taxation. *Journal of Public Economics*, 6(1-2):55–75.

Autor, D. (2019). Work of the past, work of the future. Technical report, National Bureau of Economic Research.

Autor, D., Chin, C., Salomons, A. M., and Seegmiller, B. (2022). New frontiers: The origins and content of new work, 1940–2018. Technical report, National Bureau of Economic Research.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4):1279–1333.

Azevedo, E. M. and Gottlieb, D. (2017). Perfect competition in markets with adverse selection. *Econometrica*, 85(1):67–105.

Bales, R. A. and Stone, K. V. (2020). The invisible web at work: artificial intelligence and electronic surveillance in the workplace. *Berkeley J. Emp. & Lab. L.*, 41:1.

Bastani, S., Blumkin, T., and Micheletto, L. (2015). Optimal wage redistribution in the presence of adverse selection in the labor market. *Journal of Public Economics*, 131:41–57.

Bierbrauer, F., Boyer, P., and Hansen, E. (2020). Pareto-improving tax reforms and the earned income tax credit.

Blau, F. D. and Kahn, L. M. (2007). Changes in the labor supply behavior of married women: 1980–2000. *Journal of Labor economics*, 25(3):393–438.

Blinder, A. S., Gordon, R. H., and Wise, D. E. (1980). Reconsidering the work disincentive effects of social security. *National tax journal*, 33(4):431–442.

Bonatti, A. and Hörner, J. (2017). Career concerns with exponential learning. *Theoretical Economics*, 12(1):425–475.

Brynjolfsson, E. and Mitchell, T. (2017). What can machine learning do? workforce implications. *Science*, 358(6370):1530–1534.

Brynjolfsson, E., Mitchell, T., and Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? In *AEA papers and proceedings*, volume 108, pages 43–47.

Cabral, M., Cui, C., and Dworsky, M. (2022). The demand for insurance and rationale for a mandate: Evidence from workers' compensation insurance. *American Economic Review*, 112(5):1621–68.

Card, D. (1991). Intertemporal labor supply: An assessment.

Cella, C., Ellul, A., and Gupta, N. (2017). Learning through a smokescreen: earnings management and ceo compensation over tenure. *Riksbank Research Paper Series Forthcoming, Kelley School of Business Research Paper*, (15-18):15–02.

Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J., and Mullainathan, S. (2016). Productivity and selection of human capital with machine learning. *American Economic Review*, 106(5):124–27.

Cisternas, G. (2018). Career concerns and the nature of skills. *American Economic Journal: Microeconomics*, 10(2):152–89.

Coile, C. and Gruber, J. (2007). Future social security entitlements and the retirement decision. *The review of Economics and Statistics*, 89(2):234–246.

Costinot, A. and Werning, I. (2018). Robots, trade, and luddism: A sufficient statistic approach to optimal technology regulation. Technical report, National Bureau of Economic Research.

Craig, A. C. (2020). Optimal income taxation with spillovers from employer learning.

da Costa, C. E. and Maestri, L. J. (2019). Optimal mirrleesian taxation in non-competitive labor markets. *Economic Theory*, 68(4):845–886.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their application.* Number 1. Cambridge university press.

Diamond, P. A. (1973). Consumption externalities and imperfect corrective pricing. *The Bell Journal of Economics and Management Science*, pages 526–538.

Diamond, P. A. (1998). Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. *American Economic Review*, pages 83–95.

Diamond, P. A. and Mirrlees, J. A. (1971). Optimal taxation and public production i: Production efficiency. *The American economic review*, 61(1):8–27.

Eckstein, Z. and Wolpin, K. I. (1989). Dynamic labour force participation of married women and endogenous work experience. *The Review of Economic Studies*, 56(3):375–390.

Einav, L. and Finkelstein, A. (2011). Selection in insurance markets: Theory and empirics in pictures. *Journal of Economic perspectives*, 25(1):115–38.

Einav, L., Finkelstein, A., and Cullen, M. R. (2010). Estimating welfare in insurance markets using variation in prices. *The Quarterly Journal of Economics*, 125(3):877–921.

Eissa, N. and Liebman, J. B. (1996). Labor supply response to the earned income tax credit. *The quarterly journal of economics*, 111(2):605–637.

Fan, J. and Gijbels, I. (1996). Local polynomial modelling and its applications.

Farber, H. S. and Gibbons, R. (1996). Learning and wage dynamics. *The Quarterly Journal of Economics*, 111(4):1007–1047.

Gelber, A. M., Isen, A., and Song, J. (2016). The effect of pension income on elderly earnings: Evidence from social security and full population data. *NBER Working paper*.

Gruber, J. and Saez, E. (2002). The elasticity of taxable income: evidence and implications. *Journal of public Economics*, 84(1):1–32.

Gruber, J. and Wise, D. (1998). Social security and retirement: An international comparison. *The American Economic Review*, 88(2):158–163.

Guerreiro, J., Rebelo, S., and Teles, P. (2022). Should robots be taxed? *The Review of Economic Studies*, 89(1):279–311.

Guvenen, F., Karahan, F., Ozkan, S., and Song, J. (2021). What do data on millions of us workers reveal about life-cycle earnings risk? *Econometrica*, 89(5):2303–2339.

Hariton, C. and Piaser, G. (2007). When redistribution leads to regressive taxation. *Journal of Public Economic Theory*, 9(4):589–606.

Health and Retirement Study (2018). Rand hrs longitudinal file 2018 (v1) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI.

Heckman, J. J. (1976). A life-cycle model of earnings, learning, and consumption. *Journal of political economy*, 84(4, Part 2):S9–S44.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1):169–182.

Holmstrom, B. and Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, pages 303–328.

Hörner, J. and Lambert, N. S. (2021). Motivational ratings. *The Review of Economic Studies*, 88(4):1892–1935.

Hummel, A. J. (2021). Monopsony power, income taxation and welfare.

Imai, S. and Keane, M. P. (2004). Intertemporal labor supply and human capital accumulation. *International Economic Review*, 45(2):601–641.

Jacquet, L. and Lehmann, E. (2021). Optimal income taxation with composition effects. *Journal of the European Economic Association*, 19(2):1299–1341.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of political economy*, 87(5, Part 1):972–990.

Kahn, L. B. and Lange, F. (2014). Employer learning, productivity, and the earnings distribution: Evidence from performance measures. *The Review of Economic Studies*, 81(4):1575–1613.

Kaplan, G. (2012). Inequality and the life cycle. *Quantitative Economics*, 3(3):471–525.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature*, 49(4):961–1075.

Keane, M. P. and Wolpin, K. I. (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review*, 42(4):1051–1103.

Landers, R. M., Rebitzer, J. B., and Taylor, L. J. (1996). Rat race redux: Adverse selection in the determination of work hours in law firms. *The American Economic Review*, pages 329–348.

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics*, 25(1):1–35.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208.

Miyazaki, H. (1977). The rat race and internal labor markets. *The Bell Journal of Economics*, pages 394–418.

Pantoja, P., Hayes, J., Hurd, M., Mallett, J., Martin, C., Meijer, E., Rohwedder, S., and St.Clair, P. (2018). Rand hrs tax calculations 2014 (v2) documentation. *RAND Center for the Study of Aging, Santa Monica, CA.*

Prescott, E. C. and Townsend, R. M. (1984). Pareto optima and competitive equilibria with adverse selection and moral hazard. *Econometrica: journal of the econometric society*, pages 21–45.

Racine, J. (1993). An efficient cross-validation algorithm for window width selection for nonparametric kernel regression. *Communications in Statistics-Simulation and Computation*, 22(4):1107–1114.

Rothschild, C. and Scheuer, F. (2013). Redistributive taxation in the roy model. *The Quarterly Journal of Economics*, 128(2):623–668.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The Review of Economic Studies*, 68(1):205–229.

Scheuer, F. and Werning, I. (2016). Mirrlees meets diamond-mirrlees. Technical report, National Bureau of Economic Research.

Scheuer, F. and Werning, I. (2017). The taxation of superstars. *The Quarterly Journal of Economics*, 132(1):211–270.

Seade, J. (1982). On the sign of the optimum marginal income tax. *The Review of Economic Studies*, 49(4):637–643.

Shaw, K. L. (1989). Life-cycle labor supply with human capital accumulation. *International Economic Review*, pages 431–456.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, pages 355–374.

Spence, M. (1978). Product differentiation and performance in insurance markets. *Journal of Public Economics*, 10(3):427–447.

Stantcheva, S. (2014). Optimal income taxation with adverse selection in the labour market. *Review of Economic Studies*, 81(3):1296–1329.

Stantcheva, S. (2017). Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy*, 125(6):1931–1990.

Sturm, J. and Sztutman, A. (2021). Income taxation and heterogeneity. *Available at SSRN 3820880*.

Vere, J. P. (2011). Social security and elderly labor supply: Evidence from the health and retirement study. *Labour Economics*, 18(5):676–686.

Vickrey, W. (1947). *Agenda for Progressive Taxation*. Ronald Press Co., New York.

Werning, I. (2000). An elementary proof of positive optimal marginal taxes. Technical report, mimeo MIT.

Werning, I. (2007a). Optimal fiscal policy with redistribution. *The Quarterly Journal of Economics*, 122(3):925–967.

Werning, I. (2007b). Pareto efficient income taxation. Technical report, mimeo, MIT.

Wilson, C. (1977). A model of insurance markets with incomplete information. *Journal of Economic theory*, 16(2):167–207.

1.9 Appendix

1.9.1 Salaries with Overlapping Generations (Proof of Lemma 1)

Lemma. *Given the overlapping generations structure of the model, equation 1.2 is equivalent to:*

$$w(h(s)) = \mathbb{E}[v(\theta)|h(\theta) \geq h(s)] \tag{1.11}$$

Proof. First notice that for every type who supplies $h(\theta) > h(s)$ over their lifetime, there is someone (potentially from a different cohort) who now has the experience $h(s)$. Thus $\mathbb{E}[v(\theta)|h(\theta,s) = h(s)] = \mathbb{E}[v(\theta)|h(\theta) \geq h(s)]$. Now suppose a firm sets salaries that are not equal to the average productivity, so that equation above is violated for some $h(s)$. Then, a firm that offers a contract conditional on the experience level $h(s)$ is either making losses or positive profits. In the first case, the firm would be better off by not offering the contract and in the second case a firm could enter the market offering an infinitesimally lower price and making strictly positive profits. □

1.9.2 Algebra Behind Example in Section 1.3.1

Salaries are given by the expectation of productivity of those who provide at least h. In terms of types salaries are given by $w(h(\theta)) = \mathbb{E}[v(\tilde{\theta})|\tilde{\theta} \geq \theta] = \frac{\alpha}{\alpha-\delta}\theta^{\delta}$. The first order conditions of the worker imply that:

$$w(h) = h^{1/\epsilon} b(\theta)^{-(1+1/\epsilon)}$$

Guess and verifying that the wage is a power function ,i e, $w(h) = kh^{\gamma}$, for some k and γ. We conclude that:

$$w(h) = kh^{\gamma} = \frac{\alpha}{\alpha-\delta} k^{\frac{-\delta\epsilon}{(1+\epsilon)(1-\delta)}} h^{\frac{1-\epsilon\gamma}{(1+\epsilon)(1-\delta)}}.$$

And therefore

$$k = \left(\frac{\alpha}{\alpha-\delta}\right)^{\frac{(1-\delta)(1+\epsilon)}{(1-\delta)(1+\epsilon)-\delta\epsilon}}$$

$$\gamma = \frac{\delta}{1-\delta+\epsilon}$$

1.9.3 Equilibrium Existence Without Taxes

Proposition 11. *Assuming $v(\theta)$ smoothly increasing in types, MRS smoothly decreasing in types, consumption and labor, and a continuous distribution of types, there exists an equilibrium.*

Proof. Consider the direct mechanism that offers the allocation $c(\theta), h(\theta)$, where $c(\theta)$ and $h(\theta)$ is the solution to the following system of equations.

$$\frac{c'(\theta)}{h'(\theta)} - MRS(c(\theta), h(\theta), \theta) = 0$$

$$\mathbb{E}[v(\tilde{\theta})|\tilde{\theta} \geq \theta] = \frac{c'(\theta)}{h'(\theta)}$$

$$\int (c(\theta) - v(\theta)h(\theta)))f(\theta)d\theta = 0$$

This allocation satisfies IC, feasibility. The first is equation is the local IC. $c(\theta)$ and $h(\theta)$ are increasing in θ because $\mathbb{E}[v(\tilde{\theta})|\tilde{\theta} \geq \theta]$ is increasing in θ and MRS is decreasing in θ, and increasing in c and h. Because of single-crossing the local IC and monotonicity together impliy that incentive constraints are globally satisfied . The third equation is the feasibility constraint. In this case, the feasibility constraint is automatically satisified given the first two equations, as workers get paid what they produce in expectation, firms are offering salaries $\mathbb{E}[v(\tilde{\theta})|\tilde{\theta} \geq \theta]$ and making zero profits. Notice as well that $\int^{\theta} v(\tilde{\theta})h'(\tilde{\theta})d\tilde{\theta} < v(\theta)\int_0^{\theta} h'(\tilde{\theta})d\tilde{\theta} \leq v(\theta)h(\theta)$, thus $c(\underline{\theta}) > 0$, that is, higher types workers in this allocation subsidize lower type workers and consumption is positive everywhere. The existence of a solution is guaranteed by the Picard-Lindelöf theorem, given the regularity assumptions on MRS , v and the distribution of types. □

1.9.4 Atkinson-Stiglitz with "Double Adverse Selection" (Proof of Proposition 2)

Proof. Any tax system generates a common budget set B that determines which pairs (C, H) are feasible. As a first step in the proof, we are going to show that we can replicate this budget set for the workers and save resources, without imposing taxes on timed consumption or labor flows. To do that, we generate a new tax system where, for each pair (C, H) the new pre tax income is $e(H)$ and post tax income is $e(C)$, where these are defined below:

$e(C) = \underset{\tilde{c}(\cdot)}{min} \int q(t) \cdot \tilde{c}(t)dt \ st. \ C(\tilde{c}(\cdot)) \geq C$ and $e(H) = \underset{\tilde{h}(\cdot)}{max} \int q(t) \cdot w(h()) \cdot \tilde{h}(t)dt \ st. \ H(\tilde{h}(\cdot)) \leq H$.

Under this new tax system, the worker problem can be written in three parts:

$\underset{C,H}{max}\, U(C, H, \theta) \ st. \ (C, H) \in B$ and $e(C) = \underset{\tilde{c}(\cdot)}{min} \int q(t) \cdot \tilde{c}(t)dt \ st. \ C(\tilde{c}(\cdot)) \geq C$ and $e(H) = \underset{\tilde{h}(\cdot)}{max} \int q(t) \cdot w(h(t)) \cdot \tilde{h}(t)dt \ st. \ H(\tilde{h}(\cdot)) \leq H$

Notice that $e(H)$ is the maximum pre-tax income that can be generated by generating at most the disutility H. Moreover, it depends only on lifetime labor supply and not on the timing of these labor supply decisions, because $\int q(t) \cdot w(h(t)) \cdot \tilde{h}(t) = W(h)$. Thus, it is the maximum lifetime labor that generates at most the disutility H. Because production depends only on lifetime labor supply (an is increasing in lifetime labor supply), it is also the maximum production that generates at most the disutility H.

Further, because $e(C)$ is the smallest amount of resources that achieves the subutility level C, and $e(H)$ is the maximum production that can be generated by generating at most the disutility H, whenever choices change, there are more resources than required to obtain the same allocation. These extra resources can be used to increase all the consumption possibilities $(C + \Delta(H), H)$ by some small amount $\Delta(H)$, chosen in such way that everyone still prefers their originally labor aggregate choice H.

We have assumed that the discount rates $q(\cdot)$ were exogenously given, and in this case we have shown that there is a lifetime income taxation system that is weakly better than any other tax system, as claimed in the Proposition. An analogous argument extends the result to arbitrary endogenous discount rates $q(\cdot)$, that is, discount rates that arise from competitive firms doing the intertemporal allocation of resources. Indeed, the result is implied by a production efficiency argument as in (Diamond and Mirrlees, 1971). Moreover, we assume throughout the paper that the economy is dynamically efficient in the sense that the present value of the output exists and it is finite.

□

1.9.5 Invertibility Condition (Proof of Lemma 2)

Lemma. *Without loss, we can solve directly for $\tilde{R}(h) = R(y(h))$, and then find $y(h)$, and $R(y)$. That is, the planner can solve the simpler problem:*

$$\max_{\tilde{R}(h)} \mathbb{E}[W(V(\tilde{R};\theta))] \ s.t. \ \mathbb{E}[v(\theta)h(\tilde{R};\theta) - \tilde{R}(h(\tilde{R};\theta))] \geq 0$$

Proof. Because $y'(h)$ is a well-defined function of the allocation and is always positive, $y(h)$ always exists and it is strictly increasing. Thus, there exists an inverse function $h^{-1}(y)$. Therefore we can define $R(y)$ so that $R(y(h)) = \tilde{R}(h)$. Thus, we found the income tax schedule and equilibrium salaries that prevail in the economy where the planner solved directly for the retention function $\tilde{R}(h)$. □

1.9.6 Optimal Taxes with Single Dimensional Heterogeneity (Proof of Proposition 3)

Proof. The optimal tax schedule solves the following problem:

$$\begin{aligned} \max_{\tilde{R}} \ & E[\lambda(\theta)V(\tilde{R},\theta)] \ s.t. \ E[v(\theta)(h(\theta)) - \tilde{R}(h(\theta))] \geq 0 \\ & \text{where } E[v(\theta)|MRS^{\theta}_{c,h} \leq \tilde{R}'(h)] = y'(h) \\ & \tilde{R}(h) = R(y(h)) \end{aligned}$$

Considering a small variation on marginal retention rates as a function of effort at a given level of effort:

$$E\left[\lambda(\theta)\frac{dV(\tilde{r},I,\theta)}{d\tilde{r}_h}\right] = -\mu E\left[v(\theta)\frac{dh(\theta)}{d\tilde{r}_h} - \tilde{r}(h)\frac{dh(\theta)}{d\tilde{r}_h} - 1(h(\theta) \geq h)\right]$$

$$\begin{aligned} E\left[\frac{\lambda(\theta)}{\mu}\frac{dV(\tilde{r},I,\theta)}{dI}1(h(\theta) \geq \theta)\right] = -E\Bigg[&(v(\theta) - \tilde{r}(h(\theta)))\frac{dh^c(\theta)}{d\tilde{r}_h}1(h(\theta) = h) \\ &- (v(\theta) - \tilde{r}(h(\theta))\frac{dh(\theta)}{dI}1(h(\theta) \geq h) - 1(h(\theta) \geq h)\Bigg] \end{aligned}$$

$$\Big(v(h) - \tilde{r}(h)\Big)f(h)\frac{dh^c(\theta)}{d\tilde{r}_h} = \int_h^\infty f(\tilde{h})\Big(1 - \lambda(\tilde{h})\Big)d\tilde{h} + \int_h^\infty \Big(v(\tilde{h}) - \tilde{r}(\tilde{h})\Big)f(\tilde{h})\frac{d\tilde{h}}{dI}d\tilde{h}$$

Now two steps: writing it in terms of elasticities and converting to a formula in terms of earnings. First:

$$\left(\frac{v(h) - \tilde{r}(h)}{\tilde{r}(h)}\right)f(h)h\frac{dh^c(\theta)}{d\tilde{r}_h}\frac{\tilde{r}(h)}{h} = \int_h^\infty f(\tilde{h})\Big(1 - \lambda(\tilde{h})\Big)d\tilde{h} + \int_h^\infty \left(\frac{v(\tilde{h}) - \tilde{r}(\tilde{h})}{\tilde{R}(\tilde{h})}\right)f(\tilde{h})\frac{d\tilde{h}}{dI}\tilde{R}(\tilde{h})d\tilde{h}$$

$$\left(\frac{v(h) - \tilde{r}(h)}{\tilde{r}(h)}\right)f(h)h\epsilon^c_{\tilde{r}}(h) = \int_h^\infty f(\tilde{h})\Big(1 - \lambda(\tilde{h})\Big)d\tilde{h} + \int_h^\infty \left(\frac{v(\tilde{h}) - \tilde{r}(\tilde{h})}{\tilde{R}(\tilde{h})}\right)f(\tilde{h})\eta_I(\tilde{h})d\tilde{h}$$

Then, using that:

$$\begin{aligned} \tilde{r}(h) &= r(y(h))y'(h) \\ \frac{dy(h)}{dr} &= y'(h)\frac{dh}{d\tilde{r}}\frac{d\tilde{r}}{dr} = y'(h)^2\frac{dh}{d\tilde{r}} \ \text{(micro elasticities)} \\ f(h) &= g(y(h))y'(h) \\ \frac{dy(h)}{dI} &= y'(h)\frac{dh}{dI} \end{aligned}$$

$$dh = \frac{1}{y'(h)} dy$$

$$\left(\frac{v(h)/y'(h) - r(y(h))}{r(y(h))}\right) g(y(h)) y(h) \epsilon_r^c(y(h)) = \int_h^\infty f(\tilde{h}) \left(1 - \lambda(\tilde{h})\right) d\tilde{h} + \int_h^\infty \left(\frac{v(\tilde{h})/y'(\tilde{h}) - r(y(\tilde{h}))}{R(y(\tilde{h}))}\right) f(\tilde{h}) \eta_I(y(\tilde{h})) d\tilde{h}$$

$$\left(\frac{v(y)/y'(h(y)) - r(y)}{r(y)}\right) g(y) y \epsilon_r^c(y) = \int_y^\infty g(\tilde{y}) \left(1 - \lambda(\tilde{y})\right) d\tilde{y} + \int_y^\infty \left(\frac{v(\tilde{y})/y'(h(\tilde{y})) - r(\tilde{y})}{R(\tilde{y})}\right) g(\tilde{y}) \eta_I(\tilde{y}) d\tilde{y}$$

□

1.9.7 Elasticities

The optimal tax formulas presented in Section 1.5 feature compensated and income elasticities. In the formulas, because we are keeping the pretax salaries fixed, these elasticities can be called "micro elasticities". A number of empirical studies use aggregate variation in taxes (such as state-tax variations, or kinks in the tax schedule) to estimate elasticities of taxable income. These statistics are better described as "macro elasticities" – how income changes when the tax schedule changes for everyone, potentially affecting pretax salaries. That is, these elasticities are inferred from the observation of how incomes from people living in different states react to state tax reforms, or how income in different years reacted differently to federal tax reforms (Gruber and Saez, 2002). When measuring the change in individual taxable income, they conflate the change in behavior that responds directly to the changes in marginal taxes and the change in behavior that responds to changes in wages that are induced by these economy-wide tax reforms. To extract the micro elasticities from the estimated macro elasticities it is necessary to rescale them up to account for the endogenous changes in wages due to imperfect information in labor markets. Increases in marginal retention induce the marginal types – who are the less productive types – to work more, and therefore reduce pretax salaries, making the effective change in post-tax wages smaller. Thus, the estimated elasticities of taxable income are lower in magnitude than the micro elasticities that keep the pretax salaries fixed. Proposition 12 relates steady state compensated "micro elasticities" to "macro elasticities", in the case where resumes are defined as the cumulative discounted sum of deliverables a worker has produced.

Proposition 12. *Compensated "micro elasticities" and "macro elasticities" are related by the following formula:*

$$\epsilon_{r_y}^{y,m}(y) = \frac{\epsilon_{r_y}^{y,M}(y)}{1 - \alpha(y) \epsilon_{r_y}^{y,M}(y) \left(1 - \chi(y)\right)}$$

"Macro elasticities" are lower than "micro elasticities". A locally flat increase in marginal retention at pretax income level y makes workers who were just indifferent at that region increase their labor supply, lowering the average productivity of the workers at y, and therefore lowering pretax wages, and attenuating the original increase in labor supply. How much salaries decrease is proportional to the density of people at y relative to the mass of people above y (from which the shape parameter $\alpha(y) = \frac{g(y)y}{1-G(Y)}$ shows up in the formula), to how much people are changing their income ($\epsilon_{r_y}^{y,M}(y)$), and to how far their productivities are from their salaries (hence $1-\chi(y)$ in the formula). The "micro elasticity" can be thought of as renormalizing the original elasticity by the effective change in retention, coming both from the mechanical change induced by the reform and from the endogenous change in salaries.

These are “own elasticities”: they tell how income changes for someone who initially is earning y as a response to a change in marginal rates at the same income level. Changes in marginal retention in other income levels affect salaries in multiple other income levels: the pool of workers from each income level is shifted, generating further compensated and uncompensated changes. For this reason, the expression for income elasticities is also more involved. A change in the intercept of the tax schedule generates not only further income effects but also further compensated effects. Proposition 13 relates the change in pretax salaries to changes in income.

Proposition 13. *Income “macro” elasticities and changes in pretax salaries and income “micro” elasticities are related by the following formula:*

$$\epsilon_I^{y'(h(y))} = -\alpha(y)\epsilon_I^{y,M}(y)\left(1-\chi(y)\right)$$

$$\eta(y) = \frac{\epsilon_I^{y,M}(y)\left(1+\epsilon_{r_h}^{y,M}(y)\alpha(y)\left(1-\chi(y)\right)\right)}{1-\int_0^y \alpha(\tilde{y})\epsilon_I^{\tilde{y},M}(\tilde{y})\left(1-\chi(\tilde{y})\right)\frac{y'(h(\tilde{y}))}{R(\tilde{y})}d\tilde{y}}$$

There are two effects playing a role in the relationship between income “micro” and “macro” elasticities. First, assuming that income elasticities are negative, an income transfer to all workers induces them to work less, and those who work less are the least productive, pushing towards higher salaries. Now because salaries increase, a compensated effect is increasing the labor supply of workers, making the income “macro” elasticity higher (or smaller in absolute value) than the income “micro” elasticity. Second, because salaries are increasing not only at a given income level y, but everywhere below (and above) it, the income transfer is effectively higher than what the “macro” elasticity accounts for. Thus, this is a force making the “micro” elasticity lower (higher in absolute value) than the “macro” elasticity. Thus, whether the “macro” income elasticity or the “micro” is larger depends on which effect dominates the other.

It should be noticed that these elasticities and the relationship between them hold in the steady state. Moreover, they are long-term elasticities – they look at how lifetime income changes as a response to tax reforms. As long-term elasticities, the empirical evidence on the magnitude of these is relatively scarcer compared to elasticities over shorter horizons.

Proofs of Propositions 12 and 13

Proposition: compensated “micro elasticities” and “macro elasticities” are related by the following formula:

$$\epsilon_{r_y}^{y,m} = \frac{\epsilon_{r_y}^{y,M}}{1-\alpha(y)\epsilon_{r_y}^{y,M}\left(1-\chi(y)\right)}$$

Proof. because

$$\frac{dlogr_h}{dlogr_y} = \frac{dlogy'}{dlogr_y} + 1$$

$$\epsilon_{r_y}^{y,M} = \epsilon_{r_y}^{y,m}\left(1+\epsilon_{r_y}^{y'(h(y))}\right)$$

and thus

$$\epsilon_{r_y}^{y'(h(y))} = -\alpha(y)\epsilon_{r_y}^{y,m}\left(1+\epsilon_{r_y}^{y'(h(y))}\right)\left(1-\chi(y)\right)$$

$$\implies \epsilon_{r_y}^{y'(h(y))} = \frac{-\alpha(y)\epsilon_{r_y}^{y,m}\left(1-\chi(y)\right)}{1+\alpha(y)\epsilon_{r_y}^{y,m}\left(1-\chi(y)\right)}$$

$$\epsilon_{r_y}^{y,m} = \frac{\epsilon_{r_y}^{y,M}}{\left(1+\epsilon_{r_y}^{y'(h(y))}\right)} = \epsilon_{r_y}^{y,M}\left(1+\alpha(y)\epsilon_{r_y}^{y,m}\left(1-\chi(y)\right)\right)$$

$$\implies \epsilon_{r_y}^{y,m} = \frac{\epsilon_{r_y}^{y,M}}{1-\alpha(y)\epsilon_{r_y}^{y,M}\left(1-\chi(y)\right)}$$

□

Proposition. *Income "macro" elasticities and changes in pretax salaries and income "micro" elasticities are related by the following formula:*

$$\epsilon_I^{y'(h(y))} = -\alpha(y)\epsilon_I^{y,M}(y)\left(1-\chi(y)\right)$$

$$\eta(y) = \frac{\epsilon_I^{y,M}(y)\left(1+\epsilon_{r_h}^{y,m}(y)\alpha(y)\left(1-\chi(y)\right)\right)}{1-\int_0^y \alpha(\tilde{y})\epsilon_I^{\tilde{y},M}(\tilde{y})\left(1-\chi(\tilde{y})\right)\frac{y'(h(\tilde{y}))}{R(\tilde{y})}d\tilde{y}}$$

Proof. because

$$dz = \frac{dz}{dr_h}\cdot dr_h + \frac{dz}{dI_f}\cdot dI_f$$

$$\frac{dz}{dI_o}\frac{I_o}{z} = \frac{dz}{dr_h}\cdot\frac{r_h}{z}\cdot\frac{dr_h}{dI_o}\cdot\frac{I_o}{r_h} + \frac{dz}{dI_f}\cdot\frac{I_o}{z}\cdot\frac{dI_f}{dI_o}$$

$$\epsilon_I^{y,M} = \epsilon_{r_h}^{y,m}\epsilon_I^{y'} + \eta^y\frac{dI_f}{dI_o}$$

$$\frac{dI_f^y}{dI_o} = 1 + \int_0^y \frac{dr_h}{dI_o}d\tilde{y} = 1 + \int_0^y \frac{I_o}{y'(h(\tilde{y}))}\frac{dy'(h(\tilde{y}))}{dI_o}\frac{y'(h(\tilde{y}))}{I_o}d\tilde{y} = 1 + \int_0^y \epsilon_I^{y'}\frac{y'(h(\tilde{y}))}{I_o}d\tilde{y}$$

$$\epsilon_I^{y,M} = -\epsilon_{r_h}^{y,m}\alpha(y)\epsilon_I^{y,M}(y)\left(1-\chi(y)\right) + \eta(y)\left(1-\int_0^y \alpha(\tilde{y})\epsilon_I^{\tilde{y},M}(\tilde{y})\left(1-\chi(\tilde{y})\right)\frac{y'(h(\tilde{y}))}{R(\tilde{y})}d\tilde{y}\right)$$

$$\eta(y) = \frac{\epsilon_I^{y,M}(y)\left(1+\epsilon_{r_h}^{y,m}(y)\alpha(y)\left(1-\chi(y)\right)\right)}{1-\int_0^y \alpha(\tilde{y})\epsilon_I^{\tilde{y},M}(\tilde{y})\left(1-\chi(\tilde{y})\right)\frac{y'(h(\tilde{y}))}{R(\tilde{y})}d\tilde{y}}$$

□

1.9.8 First Best with Pigouvian Taxes

Proposition 14. *Assuming $v(\theta)$ smoothly increasing in types, MRS smoothly decreasing in types, and a continuous distribution of types, a first best allocation can be achieved, with $r(y) = \chi(y)$.*

Proof. Consider the direct mechanism that offers the allocation $c(\theta), h(\theta)$, where $c(\theta)$ and $h(\theta)$ is the solution to the following system of equations.

$$\frac{c'(\theta)}{h'(\theta)} - MRS(c(\theta), h(\theta), \theta) = 0$$

$$v(\theta) = \frac{c'(\theta)}{h'(\theta)}$$

$$\int (c(\theta) - v(\theta)h(\theta)))f(\theta)d\theta = 0$$

This allocation satisfies IC, feasibility. The first is equation is the local IC. $c(\theta)$ and $h(\theta)$ are increasing in θ because $v(\theta)$ is increasing in θ and MRS is decreasing in θ, and increasing in c and h. Because of single-crossing the local IC together with monotonicity implies global IC's are satisfied . The third equation is the feasibility constraint. Notice as well that $\int^{\theta} v(\tilde{\theta})h'(\tilde{\theta})d\tilde{\theta} < v(\theta)\int_0^{\theta} h'(\tilde{\theta})d\tilde{\theta} \leq v(\theta)h(\theta)$, thus $c(\underline{\theta}) > 0$, that is, higher types workers in this allocation subsidize lower type workers and we do not need to worry about consumption being negative. Further, notice that in this allocation $r(y) = \chi(y)$. The existence of a solution to these differential equations is guaranteed by the Picard-Lindelöf theorem, given the regularity assumptions on MRS , v and the distribution of types.

Finally, we show that it is a first best allocation: it corresponds to an allocation where workers face linear budgets and get paid their marginal products while receiving transfers $I(\theta) = c(\theta) - w(\theta)h(\theta)$. □

1.9.9 Welfare (Proof of Proposition 6)

Proposition. *If the original tax schedule is optimal and the planner has strongly enough redistributive preferences, and labor is a normal good, then decreasing information asymmetries in labor markets decreases welfare.*

Proof. We can set the planner's problem as maximizing a welfare function of workers utility, subject to incentive compatibility constraints and a feasibility constraint.

$$\max_{l(\theta), c(\theta)} \int W(u(c(\theta), l(\theta)))f(\theta)d\theta$$

$$s.t.\ u\Big(c(\theta), l(\theta)\Big) \geq u\Big(c(\theta'), l(\theta')\frac{b(\theta')}{b(\theta)}\Big)\ \forall \theta, \theta'\ \text{[IC's]}$$

$$\int (c(\theta) - v(\theta)b(\theta)l(\theta)f(\theta)d\theta \leq 0\ \text{[Feasibility]}$$

The incentive compatibility constraints translate the idea that under the results from Lemma 2 (or the more general Lemma 4) the planner can solve for the allocation in terms of consumption and deliverables $h(\theta)$ subject to workers not being willing to misreport their types and trade their allocation $(c(\theta), h(\theta))$ to another allocation $(c(\theta'), h(\theta'))$, and that the amount of deliverables is equal to the product of effort $l(\theta)$ and observable component of productivities $b(\theta)$, i.e. $l(\theta) \cdot b(\theta) = h(\theta)$.

In our context the results in Seade (1982); Werning (2000), imply that out of the set of IC's, only the local downward are binding, provided that $W'(U)U_c$ is positive and decreasing, and that leisure is a normal good (they also further imply $v - r$ is positive,. For more details, see Appendix Section 1.9.10 and the argument in Werning (2000)). Now, notice that those downward incentive compatibility constraints become tighter whenever informational asymmetries decrease, while the feasibility constraint is unchanged. Thus, welfare decreases. The result does not depend on the particulars of the information structure and besides holding in the MWS case (as in Stantcheva (2014)), aad in other models of competition with imperfect information such as the competitive screening model in Azevedo and Gottlieb (2017), it also holds in the case where resumes are defined as the length of the resume, also holds under the more general information structure of Section 1.5.2, where take the general form $I(\tilde{h}(\cdot)_0^a, a) = \int_0^a \phi(\tilde{a}, a)\tilde{h}(\tilde{a})d\tilde{a}$. □

1.9.10 Formula in Terms of Types (Proof of Proposition 7)

Proposition. *Optimal taxes as a function of types θ must satisfy the following equations:*

$$r_y(\theta) = r_m(\theta) \cdot \chi(\theta)$$

$$\frac{1 - r_m(\theta)}{r_m(\theta)} f(\theta)\left(-\frac{\partial \log MRS}{\partial \theta}\right)^{-1} = \int_\theta^\infty (1 - \hat{\lambda}(\theta))f(\theta)d\theta + \int_\theta^\infty \left(\frac{1 - r_m(\theta)}{r_m(\theta)}\right)\eta(\theta)f(\theta)d\theta$$

Proof. Let's set up the planners problem as:

$$\begin{aligned} \max_{u,h} & \int \lambda(\theta)u(\theta)f(\theta)d\theta \\ s.t.\ & u'(\theta) = U_\theta(e(u(\theta), h(\theta), \theta), h(\theta), \theta) \\ & \int (e(u(\theta), h(\theta), \theta) - v(\theta)h(\theta))f(\theta)d\theta \leq 0 \end{aligned}$$

Lagrangian

$$\max_{u,h} \int \lambda\ u\ f + \mu(u' - U_\theta) - \kappa(e - vh)f\ d\theta$$

Integrate by parts

$$\max_{u,h} \int \lambda\ u\ f - \mu' u - \mu U_\theta - \kappa(e - vh)f\ d\theta + u\mu|_{\underline{\theta}}^{\overline{\theta}}$$

FOC's for $u(\theta)$,and $h(\theta)$

$$\lambda f - \mu' - \mu U_{\theta,c} e_u - \kappa e_u f = 0$$

$$\mu(U_{\theta,c} e_h + U_{\theta,h}) - k e_h f + k v f = 0$$

Replacing $e_u = U_c^{-1}$ on the first equation.

$$\lambda U_c f - \mu' U_c - \mu U_{\theta,c} - f\kappa = 0$$

Define $\hat{\mu} = \mu U_c/\kappa$ and $\hat{\lambda} = \lambda U_c/\kappa$, plus some rearrangement (as in Scheuer and Werning (2017))

$$\hat{\lambda} f - \hat{\mu}' - \hat{\mu} MRS_c h' = f$$

Rearrange second line and take a derivative

$$\hat{\mu}\frac{U_{\theta,c}MRS + U_{\theta,h}}{U_c MRS} = f\ \frac{v - MRS}{MRS}$$

$$-\hat{\mu}\frac{\partial \log MRS}{\partial \theta} = f\ \frac{v - MRS}{MRS}$$

We arrived at two equations:

$$\hat{\lambda} f - \hat{\mu}' - \hat{\mu} MRS_c h' = f$$

$$-\hat{\mu}\frac{\partial \log MRS}{\partial \theta} = f\ \frac{v - r}{r}$$

Differentiating the second equation:

$$-\hat{\mu} = \frac{f\ \frac{v-r}{r}}{\frac{\partial \log MRS}{\partial \theta}}$$

$$-\hat{\mu}' = \frac{f'\ \frac{v-r}{r} + f\ \frac{d}{d\theta}(\frac{v-r}{r})}{\frac{\partial \log MRS}{\partial \theta}} - \frac{\frac{\partial^2 \log MRS}{\partial \theta^2} f\ \frac{v-r}{r}}{\frac{\partial \log MRS}{\partial \theta}^2}$$

Plugging back in the first equation:

$$\hat{\lambda} f + \frac{f'\ \frac{v-r}{r} + f\ \frac{d}{d\theta}(\frac{v-r}{r})}{\frac{\partial \log MRS}{\partial \theta}} - \frac{\frac{\partial^2 \log MRS}{\partial \theta^2} f\ \frac{v-r}{r}}{\frac{\partial \log MRS}{\partial \theta}^2} + \frac{f\ \frac{v-r}{r}}{\frac{\partial \log MRS}{\partial \theta}} MRS_c h' = f$$

$$\hat{\lambda} - 1 + \frac{f'/f\ \frac{v-r}{r} + \frac{d}{d\theta}(\frac{v-r}{r})}{\frac{\partial \log MRS}{\partial \theta}} - \frac{\frac{\partial^2 \log MRS}{\partial \theta^2} \frac{v-r}{r}}{\frac{\partial \log MRS}{\partial \theta}^2} + \frac{\frac{v-r}{r}}{\frac{\partial \log MRS}{\partial \theta}} MRS_c h' = 0$$

$$\frac{v-r}{r}\left(\frac{\partial \log MRS}{\partial \theta}\right)^{-1}\left(\frac{f'}{f} - \frac{\frac{\partial^2 \log MRS}{\partial \theta^2}}{\frac{\partial \log MRS}{\partial \theta}} + MRS_c h' + \frac{d \log}{d\theta}(\frac{v-r}{r})\right) = 1 - \hat{\lambda}$$

$$\frac{v-r}{r}\left(f'\left(\frac{\partial \log MRS}{\partial \theta}\right)^{-1} - f\frac{\frac{\partial^2 \log MRS}{\partial \theta^2}}{\frac{\partial \log MRS}{\partial \theta}^2}\right) + \\ f\left(\frac{\partial \log MRS}{\partial \theta}\right)^{-1}\frac{d}{d\theta}\left(\frac{v-r}{r}\right) = (1-\hat{\lambda})f - \left(\frac{v-r}{r}\right)f\left(\frac{\partial \log MRS}{\partial \theta}\right)^{-1} MRS_c h'$$

Integrate both sides with respect to θ

$$\frac{v-r}{r} f\left(\frac{\partial \log MRS}{\partial \theta}\right)^{-1}\Bigg|_{\theta}^{\infty} = \int_{\theta}^{\infty}(1-\hat{\lambda})f d\theta - \int_{\theta}^{\infty}\left(\frac{v-r}{r}\right)\frac{MRS_c h'}{\frac{\partial \log MRS}{\partial \theta}} f d\theta$$

Notice that: $\eta(\theta) = \frac{dh}{dI} = \frac{MRS_c h'}{\frac{\partial \log MRS}{\partial \theta}}$,

Using that $\lim_{\theta \to \infty} \frac{v-r}{r} f \left(\frac{\partial \log MRS}{\partial \theta} \right)^{-1} = 0$:

$$\frac{v-r}{r} f \left(-\frac{\partial \log MRS}{\partial \theta} \right)^{-1} = \int_\theta^\infty (1-\hat{\lambda}) f d\theta + \int_\theta^\infty \left(\frac{v-r}{r} \right) \eta f d\theta$$

Finally, notice that $\frac{v-r}{r} = \frac{\chi - r_y}{r_y}$

$$\frac{\chi - r_y}{r_y} f \left(-\frac{\partial \log MRS}{\partial \theta} \right)^{-1} = \int_\theta^\infty (1-\hat{\lambda}) f d\theta + \int_\theta^\infty \left(\frac{\chi - r_y}{r_y} \right) \eta f d\theta$$

□

1.9.11 Conditioning on the "Strength of the Resume" (Proof of Proposition 4)

In this section, we assume that signals take the form of $h_\phi(\tilde{h}(\tilde{a})_0^1, a) = \int_0^a \phi(\tilde{a}, a)\tilde{h}(\tilde{a})d\tilde{a}$, with $\phi(\tilde{a}, a) > 0$, continuous in a and $\tilde{a}$.

Using the logic from Proposition 2 and the assumption of homogeneous preferences over the timing of labor supply and consumption as in Proposition 2, we will show that optimal taxation formulas will still take a simple structure.

We first set aside the issue of implementation, and assume and characterize the set of optimal incentive compatible allocations. Those allocations need to satisfy two properties, as stated in Lemma 3.

Lemma 3. *Any optimal and incentive compatible allocation satisfies two properties:*
Efficient timing: For any $H(\theta)$, $\tilde{h}(\cdot;\theta)_0^1 = argmax \int_0^1 q(a)\tilde{h}(a)da$ st. $H(\tilde{h}(\cdot)) = H$
Lifetime optimality:

$$\left(\frac{F_H(\theta, H) - \tilde{r}(H)}{\tilde{r}(H)} \right) f(H) H \epsilon_{\tilde{r}}^c(H) = \int_H^\infty f(\tilde{H}) \left(1 - \lambda(\tilde{H}) \right) d\tilde{H} + \int_H^\infty \left(\frac{F_H(\theta, \tilde{H}) - \tilde{r}(\tilde{H})}{\tilde{r}(\tilde{H})} \right) f(\tilde{H}) \eta_I(\tilde{H}) d\tilde{H},$$

where $F(\theta, H) = \max_{\tilde{h}} v(\theta) \int q(a)\tilde{h}(a)da$ s.t. $H(\tilde{h}(\cdot)) = H$, and $\tilde{r}(H) = \tilde{R}'(H)$.

Proof. Efficient timing is an implication of production efficiency theorem of Diamond and Mirrlees (1971). Lifetime optimality follows from standard variational argument over a retention schedule $R(H)$, which maps choices of H to an assigned consumption $C = R(H)$, where we assume that no bunching takes place at the optimal assignment. □

Lemma 4. *Assuming $\frac{d^2 H(\tilde{h}(\cdot)_0^1)}{d\tilde{h}(a) d\tilde{h}(a')} < 0$, $MRS(C, H, \theta)$ decreasing in θ, $I(\tilde{h}(\cdot)_0^a, a) = \int_0^a \phi(a, \tilde{a})\tilde{h}(\tilde{a})$, salaries are increasing in I.*

Proof. Because $MRS(C, H, \theta)$ is decreasing in θ, for any $R(H)$ strictly increasing – which is a property of the optimal allocation described in Lemma 3 – higher types θ pick higher levels of H. Since $\frac{d^2 H(\tilde{h}(\cdot)_0^1)}{d\tilde{h}(a) d\tilde{h}(a')} < 0$, higher levels of H are followed by higher levels of each $\tilde{h}(a)$. Therefore, conditional on age, higher types will pick strictly higher indexes, and thus for any $I > I'$ and $a > 0$, $\mathbb{E}[v(\theta)|I, a] > \mathbb{E}[v(\theta)|I', a]$. Therefore, $\mathbb{E}[v(\theta)|I] > \mathbb{E}[v(\theta)|I']$, and salaries are increasing in I. □

Lemma 5. *No two different continuous sequences $\tilde{h}(a)_0^1$ and $\tilde{h}'(a)_0^1$ map into the same $\tilde{y}(a)_0^1$.*

Proof. Remember that $\tilde{y}(\tilde{h}(a), \tilde{h}(\tilde{a})_0^a) = \tilde{h}(a) \cdot w(\tilde{h}(\tilde{a})_0^a) = \tilde{h}(a) \cdot \mathbb{E}[v(\theta)|I(\tilde{h}(\tilde{a})_0^a)]$. Consider the first non zero measure interval where $\tilde{h}(a)_0^1$ and $\tilde{h}'(a)_0^1$ differ, and without loss consider a ball $(\underline{a}, \bar{a})$ where $\tilde{h}(a) > \tilde{h}'(a)$. If salaries were the same for both sequences, $\tilde{y}(\tilde{h}(\tilde{a})_0^{\bar{a}} > \tilde{y}(\tilde{h}'(\tilde{a})_0^{\bar{a}}$, and the proof would be complete. However, by Lemma 4 salaries are increasing as a function of $\tilde{h}(a)$, thus indeed we have that $\tilde{y}(\tilde{h}(\tilde{a})_0^{\bar{a}} > \tilde{y}(\tilde{h}'(\tilde{a})_0^{\bar{a}}$. □

Lemma 5 allows us to decentralize any allocation with the properties from Lemma 3.

Proposition. *If $R(\tilde{y}(\cdot)_0^1)$ is optimal, then, there exists R_m, R_p with $R(\tilde{y}(\cdot)_0^1) = R_m(R_p(\tilde{y}(\cdot)_0^1)$. Such that R_m and R_p satisfy the following conditions:*

1. *Intertemporal, Pigouvian: for any $\bar{a}$,$\underline{a}$, and $H(\tilde{h}(a)_0^1) = H$, switching the timing of labor supply decisions and holding fixed H^*, should leave lifetime earnings unaffected:*

$$\tilde{R}_p(H(\tilde{h}(a)_0^1) = R_p(\tilde{y}(\tilde{h}(\tilde{a})_0^a)_0^1)$$

$$\int_{\bar{a}}^1 \frac{dR_p}{d\tilde{y}(\tilde{h}(\tilde{a})_0^a)} \frac{d\tilde{y}(\tilde{h}(\tilde{a})_0^a)}{d\tilde{h}(\bar{a})q(\bar{a})} \frac{d\tilde{h}(\bar{a})q(\bar{a})}{dH} q(a)da = \int_{\underline{a}}^1 \frac{dR_p}{d\tilde{y}(\tilde{h}(\tilde{a})_0^a)} \frac{d\tilde{y}(\tilde{h}(\tilde{a})_0^a)}{d\tilde{h}(\underline{a})q(\underline{a})} \frac{d\tilde{h}(\underline{a})q(\underline{a})}{dH} q(a)da$$

   Notice that this only describes what happens when labor supply changes across time, holding H fixed. So leaves both retention as a function of the index H, $\tilde{R}_p(H)$, and retention as a function of the timing of earning $R_p(\tilde{y}(\cdot)_0^1)$, partially defined.

2. *Lifetime, Pigouvian: increasing H^* should increase lifetime earnings proportionally to the increase in output:*

$$\tilde{R}(H) = R_p(\tilde{y}(\tilde{h}(\tilde{a})_0^a; H^*)_0^1)$$

$$\tilde{R}'_p(H) = \int_0^1 \frac{dR_p(\tilde{y}(\tilde{h}(\tilde{a})_0^a; H^*)_0^1)}{d\tilde{y}(\tilde{h}(\tilde{a})_0^a; H^*)} \frac{d\tilde{y}(\tilde{h}(\tilde{a})_0^a; H^*)}{d\tilde{h}(a)} \frac{d\tilde{h}(a)}{dH} da = v(H) \int_0^1 q(a) \frac{d\tilde{h}(a)}{dH} da$$

3. *Lifetime, redistributive: Define the retention that workers face as $R_m(R_p(\tilde{y}(\cdot)))$, and $r_m = R'_m(R_p)$. After correcting for distortions, then R_m should satisfy standard Mirrleesian formulas:*

$$\left(\frac{1 - r_m(R_p)}{r_m(R_p)}\right) g(R_p) R_p \epsilon_{\tilde{r}}^c(R_p) = \int_{R_p}^\infty g(\tilde{R}_p) \left(1 - \lambda(\tilde{R}_p)\right) d\tilde{R}_p + \int_{R_p}^\infty \left(\frac{1 - r_m(\tilde{R}_p)}{r_m(\tilde{R}_p)}\right) g(\tilde{R}_p) \eta_I(\tilde{R}_p) d\tilde{R}_p,$$

Proposition 1.9.11 states that the tax system should be such that i) history dependent taxes (R_p) should be used to correct for labor wedges, and ii) after correcting for these distortions, lifetime income redistributive taxes should be imposed on top these taxes, according to standard redistributive formulas.

Remark. We can define R_p to be such that:

$$v(H)q(a) = \frac{dR_p(\tilde{y}(\cdot)_0^1)}{d\tilde{y}(a)} w(\tilde{h}(\cdot)_0^a) + \int_a^1 \frac{dR_p(\tilde{y}(\cdot)_0^1)}{d\tilde{y}(\tilde{a})} \frac{dw(\tilde{h}(\cdot)_0^{\tilde{a}})}{d\tilde{h}(a)} \tilde{h}(\tilde{a}) d\tilde{a},$$

which, in the case where there are no intertemporal distortions in pre-tax salaries, simplifies to:

$$\frac{dR_p(\tilde{y}(\cdot)_0^1)}{d\tilde{y}(a)} \frac{1}{q(a)} = \frac{v(H)q(a)}{q(a)w(\tilde{h}(\cdot)_0^a) + \int_a^1 \frac{dw(\tilde{h}(\cdot)_0^{\tilde{a}})}{d\tilde{h}(a)} \tilde{h}(\tilde{a})q(\tilde{a})d\tilde{a}},$$

where $v(H)$ is the marginal productivity of the type that supplies the level H of labor, and where for ease of notation the dependence on $\tilde{y}(\cdot)_0^1$ is omitted. That is, the formula should be read as a function of earnings flows $\tilde{y}(\cdot)$, through the inverse operator $\tilde{h}(\tilde{y}(\cdot)_0^1)_0^1$.

Proof. We have established that we can find the optimal post-tax wages $\tilde{R}(\tilde{h}(\cdot))$,by solving the following problem:

$$\underset{\tilde{R}(\cdot)}{max}\ \mathbb{E}[W(V(\tilde{R},\theta))] \text{ s.t. } \mathbb{E}\left[\int_0^1 v(\theta)q(a)\tilde{h}(a)da - \tilde{R}(\tilde{h}(\cdot))\right] = 0$$

That is, we solve for the optimal post-tax wages as if there was a first layer of taxes guaranteeing that the lifetime gains from increasing the labor supply are equal to the contribution to output, ie:

$$\frac{\partial \tilde{R}_p(\tilde{h}(\cdot))}{\partial \tilde{h}(a)} = v(\theta)q(a)$$

In terms of earnings, we have that:

$$R_p(\tilde{y}(\tilde{h}(\cdot))) = \tilde{R}_p(\tilde{h}(\cdot))$$

Therefore,

$$\frac{\partial \tilde{R}_p(\tilde{h}(\cdot))}{\partial \tilde{h}(a)} = \int_0^1 \frac{\partial R_p(\tilde{y}(\cdot))}{\partial \tilde{y}(\tilde{a})}\frac{\partial \tilde{y}(\tilde{a})}{\partial \tilde{h}(a)}d\tilde{a}$$

and

$$v(H)q(a) = \frac{dR_p(\tilde{y}(\cdot)_0^1)}{d\tilde{y}(a)}w(\tilde{h}(\cdot)_0^a) + \int_a^1 \frac{dR_p(\tilde{y}(\cdot)_0^1)}{d\tilde{y}(\tilde{a})}\frac{dw(\tilde{h}(\cdot)_0^{\tilde{a}})}{d\tilde{h}(a)}\tilde{h}(\tilde{a})d\tilde{a},$$

To arrive at the the special case with no intertemporal distortions, notice that having no (pre-tax) intertemporal distortions implies that $\frac{dR_p(\tilde{y}(\cdot)_0^1)}{d\tilde{y}(\tilde{a})}q(\tilde{a}) = \frac{dR_p(\tilde{y}(\cdot)_0^1)}{d\tilde{y}(a)}q(a)$ for every $\tilde{a}$ and a. □

1.9.12 Conditioning on the Full History of Completion of Deliverables (Proof of Proposition 5)

We assume in this section that preferences now take the form $\mathrm{U}(C,\tilde{h}(\cdot),\theta)$ and that θ is high-dimensional, in the sense that, given a retention function $R(\tilde{h}(\cdot))$, strictly increasing in $\tilde{h}(a)$ for any a, for any continous $\tilde{h}(\cdot)_0^1$, there exists a type θ, for which supplying $\tilde{h}(\cdot)_0^1$ is optimal. We restrict household choices to the set of continuous $\tilde{h}(\cdot)$. We retain the assumption more productive types are more willing to provide the deliverables, which is stated as: for any set of salaries $w(\cdot)$, for any two continous flows $\tilde{h}_A(\cdot)_0^1 \geq \tilde{h}_B(\cdot)_0^1$, if there is an non-zero measure interval L such that for $\tilde{a} \in L$, $\tilde{h}_A(\tilde{a}) > \tilde{h}_B(\tilde{a})$, then for all $\bar{a} \geq \sup L$ the set of types who, given $w(\cdot)$, supply a labor supply flow which coincides with $\tilde{h}_A(\cdot)_0^{\bar{a}}$ is on average more productive than the set of types who supply $\tilde{h}_B(\cdot)_0^{\bar{a}}$.

Lemma 6. *If the planner could choose the allocation, while being restricted to set the of incentive compatible allocations, any optimal allocation would lie at the frontier of production possibilities set.*

Proof. Follows from analogous arguments from the production efficiency theorem of Diamond and Mirrlees (1971). That is, consider the problem

$$\underset{C_\theta,\tilde{h}_\theta(\cdot)}{max}\ \mathbb{E}[W(\mathrm{U}(C_\theta,\tilde{h}_\theta(\cdot),\theta))]\ st\ \mathbb{E}[v(\theta)\int_0^1 q(a)\tilde{h}_\theta(a)da - C_\theta] \geq 0$$

$$U(C_\theta,\tilde{h}_\theta(\cdot),\theta) \geq U(C_{\theta'},\tilde{h}_{\theta'}(\cdot),\theta)\ \forall\theta,\theta'$$

By the taxation principle, we can incorporate the incentive compatibility constraints into the indirect utility function of the workers and solve the equivalent problem:

$$\underset{R(\tilde{h}_\theta(\cdot))}{max}\ \mathbb{E}[W(\mathrm{V}(R,\theta))]\ st\ \mathbb{E}[v(\theta)\int_0^1 q(a)\tilde{h}_\theta(a;R)da - R(\tilde{h}_\theta(\cdot;R))] \geq 0 \tag{1.12}$$

Now, towards a contradiction, suppose production takes place at the interior of the production possibility frontier. Then, we can increase $R(\tilde{h}_\theta(\cdot))$ uniformly (as lowering the price of the consumption good). Because the indirect utility function is increasing in R, everyone would be better off, and welfare would be higher. This is feasible, because under the assumption that labor supply decisions are continuous in a uniform increase in R, there is a small enough increase in $R(\tilde{h}_\theta(\cdot))$ that keeps the allocation inside the production possibilities set.

□

The second result is that the planner can use Pigouvian taxes to achieve the frontier of the production efficient set of allocations, because sequences of $\tilde{h}(\cdot)_0^1$ will map to sequences of $\tilde{y}(\cdot)_0^1$ one-to-one, as in Lemma 2 and Lemma 5.

Lemma 7. *There is a positive return to experience. That is, in any optimal allocation, salaries* $w(\tilde{h}(\cdot)_0^a) = \mathbb{E}[v(\theta)|\tilde{h}(\cdot)_0^a]$ *are increasing in labor supply choices* $\tilde{h}(\tilde{a})$, *where* $\tilde{a} \leq a$.

Proof. This is an immediate consequence of the assumption that the more productive types are more willing to provide the deliverables, and that the type space is rich enough so that for any path $\tilde{h}(\cdot)$, expectations are well-defined. □

This Lemma, analogously to 4, establishes that there is positive return to experience.

Lemma 8. *The planner can infer labor supply choices from earnings. That is, no two continuous* $\tilde{h}(\cdot)_0^1$ *map to the same* $\tilde{y}(\cdot)_0^1$.

Proof. Remember that $\tilde{y}(\tilde{h}(a),\tilde{h}(\tilde{a})_0^a) = \tilde{h}(a)\cdot w(\tilde{h}(\tilde{a})_0^a) = \tilde{h}(a)\cdot\mathbb{E}[v(\theta)|\tilde{h}(\tilde{a})_0^a]$. Consider the first non zero measure interval where $\tilde{h}(a)_0^1$ and $\tilde{h}'(a)_0^1$ differ, and without loss consider a ball $(\underline{a},\bar{a})$ where $\tilde{h}(a) > \tilde{h}'(a)$. If salaries were the same for both sequences, $\tilde{y}(\tilde{h}(\tilde{a})_0^{\bar{a}} > \tilde{y}(\tilde{h}'(\tilde{a})_0^{\bar{a}}$, and the proof would be complete. However, by the previous Lemma salaries are increasing as a function of $\tilde{h}(a)$, thus indeed we have that $\tilde{y}(\tilde{h}(\tilde{a})_0^{\bar{a}} > \tilde{y}(\tilde{h}'(\tilde{a})_0^{\bar{a}}$. □

Those results imply Pigouvian taxes should play an important role, as stated in the following Proposition.

Proposition. *The planner can guarantee that the allocation would lie at the frontier of the production possibilities set by using Pigouvian taxes.*

Proof. Consider the problem 1.12. This formulation can be thought of as solving for the redistributive wage schedule after Pigouvian taxes have been imposed, so that pre-tax salaries of the workers would have been equal to their productivities (or the average productivity of the workers with the same labor supply history $\tilde{h}(\cdot)$, if at the solution there are multiple types sharing the same history). The solution of this problem results in a wage schedule, $R(\tilde{h}(\cdot))$, which is as a function of labor supply decisions. This wage schedule, by the previous Lemma, can be written as a function of the history of earnings $\tilde{y}(\cdot)$, $R(\tilde{y}(\cdot))$, so it can be implemented with history dependent earnings taxes. □

These Pigouvian taxes take the same general form as in the previous section. Thus, although this economy may look quite complicated, the same principles of tax design can be applied. There is a caveat though.

Because we have unrestricted preferences, and multiple goods, now the design of optimal redistributive taxes, after correcting for the Pigouvian distortions is more complicated, and without further normative assumptions, we cannot point to lifetime income taxation as the preferred form of redistribution.

1.9.13 Heterogeneity in Elasticities (Proof of Proposition 9)

Proposition. *If a tax schedule is optimal then it needs to satisfy the following relationship:*

$$\mathbb{E}\left[\left(\frac{\chi(y)-r(y)}{r(y)}\right)\epsilon_r^c(y)\right]g(y)y = \int_y^\infty g(\tilde{y})\left(1-\mathbb{E}[\lambda(\tilde{y})]\right)d\tilde{y} + \int_y^\infty \mathbb{E}\left[\left(\frac{\chi(\tilde{y})-r(\tilde{y})}{r(\tilde{y})}\right)\eta_I(\tilde{y})\right]g(\tilde{y})d\tilde{y}$$

where $\chi(y) \equiv v(y,\theta)/y'(h(y))$

Proof. The proof is analogous to the single dimensional heterogeneity case, except that the expectation on the outside cannot be dropped. □

1.9.14 Firms See Additional Signals (Proof of Proposition 10)

Proof.

$$\max_R E[\lambda(\theta)V(\tilde{R},\theta)] \; s.t. \; E[v(\theta)(h(\theta)) - \tilde{R}(h(\theta))] \geq 0$$
$$\text{where } E[v(\theta)|MRS^\theta_{c,h} \leq \tilde{R}'(h)] = y'(h)$$
$$\tilde{R}(h) = R(y(h))$$

Considering a small variation on marginal retention rates as a function of earnings, and noticing that this variation translates into an $\frac{dr_h}{dr_y}$ variation in the marginal retention as a function of effort:

$$E\left[\lambda(\theta)\frac{dV(r,\theta)}{dr_h}\cdot\frac{dr_h}{dr_y}\right] = -\mu E\left[\theta\frac{dh(\theta)}{dr_h}\cdot\frac{dr_h}{dr_y} - \frac{dh(\theta)}{dr_h}\cdot\frac{dr_h}{dr_y}r_h(h(\theta),z(\theta)) - 1(y(h(\theta),z(\theta)) \geq y)\cdot\frac{dr_h}{dr_y}\right]$$

$$E\left[\int_y^\infty \lambda(\theta)\frac{dV(r,\theta)}{dr_{h(\tilde{y};z)}}\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{y}\right] =$$
$$-\mu E\left[\int_y^\infty v(\theta)\frac{dh(\theta)}{dr_{h(\tilde{y};z)}}\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y} - \frac{dh(\theta)}{dr_{h(\tilde{y};z)}}\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}r_h(h(\theta),z(\theta)) \right.$$
$$\left. -1(y(h(\theta),z(\theta)) \geq y)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{y}\right]$$

$$E\left[\int_y^\infty \frac{\lambda(\theta)}{\mu}\frac{dV(r,\theta)}{dI}1(h(\theta) \geq h(\tilde{y};z))\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{y}\right] =$$
$$-E\left[\int_y^\infty \left(v(\theta) - r_h(h(\theta),z(\theta))\right)\left(\frac{dh^c(\theta)}{dr_{h(\tilde{y};z)}}1(h(\theta) = h(\tilde{y};z)) - \frac{dh(\theta)}{dI}1(h(\theta) \geq h(\tilde{y};z))\right)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}\right.$$
$$\left. -1(y(h(\theta),z(\theta)) \geq y)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{y}\right]$$

$$\mathbb{E}_z\left[\int_y^\infty \int_{h(\tilde{y};z)}^\infty \lambda(\tilde{h};z)f(\tilde{h}|z)\cdot \frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{h}d\tilde{y}\right] =$$

$$-E\left[\int_y^\infty \left(v(\theta)-r_h(h(\theta),z(\theta))\right)\left(\frac{dh^c(\theta)}{dr_{h(\tilde{y};z)}}1(h(\theta)=h(\tilde{y};z))-\frac{dh(\theta)}{dI}1(h(\theta)\geq h(\tilde{y};z))\right)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}\right.$$
$$\left.-1(y(h(\theta),z(\theta))\geq y)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{y}\right]$$

$$\mathbb{E}_z\left[\int_y^\infty \int_{h(\tilde{y};z)}^\infty \lambda(\tilde{h};z)f(\tilde{h}|z)\cdot \frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{h}d\tilde{y}\right] =$$
$$-\mathbb{E}_z\left[\int_y^\infty \left(v(\theta)-r_h(h(\theta),z(\theta))\right)\frac{dh^c(\theta)}{dr_{h(\tilde{y};z)}}f(h(\tilde{y};z)|z)\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{y}\right]$$
$$+\mathbb{E}_z\left[\int_y^\infty \int_{h(\tilde{y};z)}^\infty \left(\left(v(\theta)-r_h(h(\theta),z(\theta))\right)\frac{dh(\theta)}{dI}f(\tilde{h}|z)\right)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{h}d\tilde{y}\right]$$
$$+\mathbb{E}_z\left[\int_y^\infty \int_{h(\tilde{y};z)}^\infty f(\tilde{h}|z)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{h}d\tilde{y}\right]$$

$$\mathbb{E}_z\left[\int_y^\infty \left(v(\theta)-r_h(h(\theta),z(\theta))\right)\frac{dh^c(\theta)}{dr_{h(\tilde{y};z)}}f(h(\tilde{y};z)|z)\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{y}\right] =$$
$$\mathbb{E}_z\left[\int_y^\infty \int_{h(\tilde{y};z)}^\infty \left(1-\lambda(\tilde{h};z)\right)f(\tilde{h}|z)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{h}d\tilde{y}\right]$$
$$+\mathbb{E}_z\left[\int_y^\infty \int_{h(\tilde{y};z)}^\infty \left(\left(v(\theta)-r_h(h(\theta),z(\theta))\right)\frac{dh(\theta)}{dI}f(\tilde{h}|z)\right)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{h}d\tilde{y}\right]$$

$$\mathbb{E}_z\left[\int_y^\infty \left(\frac{v(h(\tilde{y},z);z)-r_h(h(\tilde{y};z);z)}{r_h(h(\tilde{y};z);z)}\right)\epsilon^c_{r_h}(h(\tilde{y};z);z)f(h(\tilde{y};z)|z)h(\tilde{y};z)\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{y}\right] =$$
$$\mathbb{E}_z\left[\int_y^\infty \int_{h(\tilde{y};z)}^\infty \left(1-\lambda(\tilde{h};z)\right)f(\tilde{h}|z)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{h}d\tilde{y}\right]$$
$$+\mathbb{E}_z\left[\int_y^\infty \int_{h(\tilde{y};z)}^\infty \left(\frac{v(\tilde{h};z)-r_h(\tilde{h};z)}{R_h(\tilde{h};z)}\right)\eta^h_I(\tilde{h};z))f(\tilde{h}|z)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{h}d\tilde{y}\right]$$

$$\mathbb{E}_z\left[\int_y^\infty \left(\frac{\chi(\tilde{y},z)-r_{\tilde{y}}}{r_{\tilde{y}}}\right)\epsilon^c_{r_{\tilde{y}}}(\tilde{y},z)g(\tilde{y},z)\tilde{y}\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{y}\right] = \mathbb{E}_z\left[\int_y^\infty \int_{\tilde{y}}^\infty \left(1-\lambda(\tilde{\tilde{y}};z)\right)g(\tilde{\tilde{y}}|z)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{\tilde{y}}d\tilde{y}\right]$$
$$+\mathbb{E}_z\left[\int_y^\infty \int_{\tilde{y}}^\infty \left(\frac{\chi(\tilde{\tilde{y}},z)-r_{\tilde{\tilde{y}}}}{R(\tilde{\tilde{y}})}\right)\eta^h_I(\tilde{\tilde{y}})g(\tilde{\tilde{y}}|z)\cdot\frac{dr_{h(\tilde{y};z)}}{dr_y}d\tilde{\tilde{y}}d\tilde{y}\right]$$

□

1.10 Empirical Appendix

1.10.1 Computing Simulated Marginal Rates and Their Changes

Whenever possible, to match available HRS variables to TAXSIM32 variables, we follow the same treatment of input variables to TAXSIM as outlined in RAND's 2014 HRS tax calculations (Pantoja et al., 2018). Discrepancies are recorded in footnotes under the corresponding HRS variables.

Table 1.1: Correspondence of HRS and taxsim32 Variables

	HRS Variable(s) Used	taxsim32 variable
1	N/A	taxsimid
2	[43]RwIWENDY	year
3	RwSTATE	state
4	[44]RwMSTAT	mstat
5	RwAGEY_B	page
6	SwAGEY_B	sage
7	[45]We use the dependents variable when available and impute values when needed.	depx
8	N/A	dep13
9	N/A	dep17
10	N/A	dep18
11	RwIEARN	pwages
12	SwIEARN	swages
13	[46]HwIDIVIN	dividends
14	N/A (but this variable is not available in RAND's version of taxsim)	intrec
15	N/A	stcg
16	N/A	ltcg
17	[47]HwRNTIN, HwIOTHI1, HwIOTHI2, HwITRSIN, HwIBNDIN, HwIBUSIN, HwICHKIN, HwIDIN, HwIALMNY, HwICDIN, HwICHKIN, H1ISAV1, H2ISAV2, HwIBUSIN, HwILUYR1-HwILUYR3	otherprop
18	N/A	nonprop
19	RwIPENA, SwIPENA	pensions
20	[48]RwISSDI, RwISSI, SwISSDI, SwISSI, Medicare part b) coverage variable from FAT Files	gssi
21	RwIUNEM, SwIUNEM	ui
22	HwISSI, HwIFOOD, HwIWELF, RwIWCMP, RwIVET	transfers
23	[49]Dollar amount of rent paid variable from FAT Files	rent paid
24	[50]Dollar amount of real estate tax paid variable from FAT Files	proptax
25	[51]RwOOPMD, SwOOPMD	otheritem
26	N/A	childcare
27	[52]HwAMORT, Dollar amount of donations variable from FAT Files, RwOOPMD, SwOOPMD	mortgage
28	N/A	scorp
29	N/A	pbusinc
30	N/A	pprofinc
31	N/A	sbusinc
32	N/A	sprofinc

1.10.2 State Variation in Taxes

Figures 1.5 to 1.10 illustrate in which states have had the largest and most frequent changes in real marginal rates, for both single and joint tax returns. To observe the changes in state income tax codes, we obtain income percentile cutoffs by converting all nominal incomes reported in the HRS dataset to 2021 dollars, using the PCE index. Then, we construct a pseudo dataset with these cutoffs in all 50 states plus Washington DC from 1992 to 2018. We use the NBER tax simulator (taxsim32) to simulate marginal income tax rates for these constructed individuals. Finally, we increment year by one, inflate accordingly using the PCE index, obtain a new set of marginal income tax rates, and take the difference between the two rates to find the policy change at each income level.

1.10.3 Mental Status Scores

The mental status summary sums the scores for serial 7's (RwSER7, 0-5), backwards counting from 20 (RwBWC2, 0-2), and object (RwCACT, RwSCIS; 0-2 total), date (RwDY, RwMO, RwYR, RwDW; 0-4 total), and President/Vice-President (RwPRES, RwVP; 0-2 total) naming tasks. The resulting range is 0-15. Since these items were not included in Waves 1 and 2H, there is no mental summary score for these waves, and the Wave 2A summary is called R2AMSTOT to indicate that it is limited to the AHEAD cohort in Wave 2.

Those questions are presented in the table below.

Table 1.2: Mental Status Scores - Questions

Variable	Content	Score
RwCACT	"What do you call the kind of prickly plant that grows in the desert?"	0-1
RwSCIS	"What do you usually use to cut paper?"	0-1
RwSCIS and RwVP	whether the Respondent was able to correctly name the current president and vice-president of the United States, respectively.	0-2
RwSER7	Number of correct subtractions in the serial 7s test. This test asks the individual to subtract 7 from the prior number, beginning with 100 for five trials. Correct subtractions are based on the prior number given, so that even if one subtraction is incorrect subsequent trials are evaluated on the given (perhaps wrong) answer.	0-5
RwBWC20 and RwBWC86	whether the Respondent was able to successfully count backwards for 10 continuous numbers from 20 and 86, respectively. Two points are given if successful on the first try, one if successful on the second, and zero if not successful on either try.	0-2
RwDY, RwMO, RwYR, and RwDW	whether the Respondent was able to report today's date correctly, including the day of month, month, year, and day of week, respectively.	0-4

1.10.4 Tables

Table 1.3: Elasticities of wages

	(1)	(2)	(3)	(4)	(5)	(6)
ϵ^w	-0.13	-0.061	-0.066	-0.092	-0.16	-0.16
	(0.10)	(0.10)	(0.10)	(0.099)	(0.099)	(0.099)
year f.e.	no	yes	yes	yes	yes	yes
marital status	no	no	yes	yes	yes	yes
hourly wages	no	no	no	linear	c. spline	l. spline
observations	39179	39179	39179	39179	39179	39179

Notes. Robust standard errors in parentheses. Elasticities ϵ^w are computed from linear regressions of changes in log hourly wages over four years on changes in log marginal retention rates over two years (evaluated at the base year income data). Each column includes different sets of controls: year fixed effects, marital status dummies, and hourly wages. Column (4) includes log hourly wages. Column (5) includes a 5 piece cubic spline of log hourly wages. Column (6) includes a 10-piece linear spline of log hourly wages.

Table 1.4: Elasticities of wages for job switchers

	(1)	(2)	(3)	(4)	(5)	(6)
ϵ^w	-0.31	-0.19	-0.19	-0.28	-0.32	-0.34
	(0.18)	(0.19)	(0.19)	(0.18)	(0.18)	(0.18)
year f.e.	no	yes	yes	yes	yes	yes
marital status	no	no	yes	yes	yes	yes
hourly wages	no	no	no	linear	c. spline	l. spline
observations	13958	13958	13958	13958	13958	13958

Notes. Robust standard errors in parentheses. Elasticities ϵ^w are computed from linear regressions of changes in log hourly wages over four years on changes in log marginal retention rates over two years (evaluated at the base year income data). The sample is restricted to those who switch jobs at least once between the baseline year and four years later. Each column includes different sets of controls: year fixed effects, marital status dummies, and hourly wages. Column (4) includes log hourly wages. Column (5) includes a 5 piece cubic spline of log hourly wages. Column (6) includes a 10-piece linear spline of log hourly wages.

Table 1.5: Participation semi-elasticities

	(1)	(2)	(3)	(4)	(5)	(6)
η^p	-0.039	0.063	0.044	0.037	0.013	0.010
	(0.056)	(0.058)	(0.056)	(0.062)	(0.062)	(0.062)
year f.e.	no	yes	yes	yes	yes	yes
marital status	no	no	yes	yes	yes	yes
hourly wages	no	no	no	linear	c. spline	l. spline
observations	72526	72526	72526	61526	61526	61526

Notes. Robust standard errors in parentheses. Semi-elasticities η^p are computed from linear regressions of changes of participation over four years on changes in log marginal retention rates over two years (evaluated at the base year income data). Each column includes different sets of controls: year fixed effects, marital status dummies, and hourly wages. Column (4) includes log hourly wages. Column (5) includes a 5 piece cubic spline of log hourly wages. Column (6) includes a 10-piece linear spline of log hourly wages.

Table 1.6: "Rat race" externality estimates (4y/4y) $-(1-\chi)$

	(1)	(2)	(3)
ratio	0.034	-0.014	-0.015
	(-0.0095, 2.42)	(-2.36, 0.027)	(-2.69, 0.024)
year f.e.	no	yes	yes
marital status	no	no	yes
hourly wages	no	no	no
observations	85706	85706	85706

Table 1.7: "Rat race" externality estimates (4y/4y) $-(1-\chi)$ (cont.)

	(1)	(2)	(3)
ratio	-0.025	-0.12	-0.16
	(-47.4, 0.0071)	(-9.03, -0.044)	(-31.3, -0.076)
year f.e.	yes	yes	yes
marital status	yes	yes	yes
hourly wages	linear	c. spline	l. spline
observations	85706	85706	85706

Notes. Bootstrapped bias-corrected confidence intervals in parentheses (with 2000 bootstrap replications). Estimates for externality ($1-\chi$, or one minus the labor wedge) are obtained from dividing the elasticity of wages by the participation semi-elasticity multiplied by one hundred. Elasticities of wages are computed from linear regressions of changes in log hourly wages over the next four years on changes in log marginal retention rates over two years (evaluated at the base year income data). Participation semi elasticities are computed from regressing changes in participation over the next four years on changes in log marginal retention rates over two years (evaluated at the base year income data). Each column includes different sets of controls: year fixed effects, marital status dummies, and hourly wages. In the bottom table, column (1) includes log hourly wages. Column (2) includes a 5 piece cubic spline of log hourly wages. Column (3) includes a 10-piece linear spline of log hourly wages.

Table 1.8: "Rat race" externality estimates (4y/2y) $-(1-\chi)$

	(1)	(2)	(3)
ratio	-0.017	-0.0045	-0.0049
	(-0.22, 0.019)	(-0.029, 0.013)	(-0.030, 0.013)
year f.e.	no	yes	yes
marital status	no	no	yes
hourly wages	no	no	no
observations	85706	85706	85706

Table 1.9: "Rat race" externality estimates (4y/2y) $-(1-\chi)$ (cont.)

	(1)	(2)	(3)
ratio	-0.0072	-0.015	-0.015
	(-0.043, 0.0086)	(-0.12, 0.0045)	(-0.13, 0.0053)
year f.e.	yes	yes	yes
marital status	yes	yes	yes
hourly wages	linear	c. spline	l. spline
observations	85706	85706	85706

Notes. Bootstrapped bias-corrected confidence intervals in parentheses (with 2000 bootstrap replications). Estimates for externality ($1-\chi$, or one minus the labor wedge) are obtained from dividing the elasticity of wages by the participation semi-elasticity multiplied by one hundred. Elasticities of wages are computed from linear regressions of changes in log hourly wages over the next four years on changes in log marginal retention rates over two years (evaluated at the base year income data). Participation semi elasticities are computed from regressing changes in participation over the next two years on changes in log marginal retention rates over two years (evaluated at the base year income data). Each column includes different sets of controls: year fixed effects, marital status dummies, and hourly wages. In the bottom table, column (1) includes log hourly wages. Column (2) includes a 5 piece cubic spline of log hourly wages. Column (3) includes a 10-piece linear spline of log hourly wages.

Table 1.10: Mental status scores

	(1)	(2)	(3)	(4)	(5)	(6)
η^m	0.14	-0.82	-1.08	-1.20	-1.30	-1.31
	(0.42)	(0.51)	(0.51)	(0.54)	(0.54)	(0.54)
year f.e.	no	yes	yes	yes	yes	yes
marital status	no	no	yes	yes	yes	yes
hourly wages	no	no	no	linear	c. spline	l. spline
observations	16027	13187	13187	11030	11030	11030

Notes. Robust standard errors in parentheses. Semi-elasticities η^m are computed from linear regressions of mental status scores two years in the past on changes in log marginal retention rates over two years (evaluated at the base year income data), restricting the sample for those who are working in the baseline year and four years ahead. Each column includes different sets of controls: year fixed effects, marital status dummies, and hourly wages. Column (4) includes log hourly wages. Column (5) includes a 5 piece cubic spline of log hourly wages. Column (6) includes a 10-piece linear spline of log hourly wages.

Table 1.11: Elasticities of wages

	-8y	-6y	-4y	-2y	+2y	+4y	+6y	+8y
ϵ^w	0.12	-0.0089	-0.014	0.055	0.036	-0.16	-0.27	-0.14
	(0.11)	(0.075)	(0.062)	(0.047)	(0.076)	(0.099)	(0.12)	(0.15)
N	17898	25425	34685	45922	53931	39179	29253	21119

Notes. Robust standard errors in parentheses. Elasticities ϵ^w are computed from linear regressions of changes in log hourly wages over k-years on changes in log marginal retention rates over two years (evaluated at the base year income data). All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Table 1.12: Elasticities of wages for job switchers

	-8y	-6y	-4y	-2y	+2y	+4y	+6y	+8y
ϵ^w	0.36	0.19	0.023	0.30	0.23	-0.34	-0.43	-0.43
	(0.16)	(0.13)	(0.12)	(0.12)	(0.19)	(0.18)	(0.18)	(0.21)
N	10040	11950	12375	9430	10996	13958	13703	11737

Notes. Robust standard errors in parentheses. Elasticities ϵ^w are computed from linear regressions of changes of log hourly wages over k-years on changes in log marginal retention rates over two years (evaluated at the base year income data). For each column, the sample is restricted to those who switch jobs at least once between the baseline year and k years ahead. All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Table 1.13: Participation semi-elasticities

	-8y	-6y	-4y	-2y	+2y	+4y	+6y	+8y
η^p	-0.039	-0.021	-0.048	-0.039	0.10	0.010	0.032	0.15
	(0.026)	(0.019)	(0.017)	(0.015)	(0.047)	(0.062)	(0.068)	(0.075)
N	20522	29053	39482	51895	72301	61526	53964	46435

Notes. Robust standard errors in parentheses. Semi-elasticities η^p are computed from linear regressions of changes of hours wages over k-years years on changes in log marginal retention rates over two years (evaluated at the base year income data). All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Table 1.14: Mental status scores

	-8y	-6y	-4y	-2y	+2y	+4y	+6y	+8y
η^m	0.32	-0.034	-0.56	0.65	-0.77	-1.31	-1.19	-0.61
	(0.55)	(0.61)	(0.45)	(0.38)	(0.45)	(0.54)	(0.57)	(0.66)
N	4362	5784	9566	13055	14705	11030	8264	5005

Notes. Robust standard errors in parentheses. For columns (5) to (8), semi-elasticities η^p are computed from linear regressions of memory scores two years in the past over years on changes in log marginal retention rates over two years (evaluated at the base year income data) , restricting the sample for those who are working in the baseline year and k-years ahead. For columns (1) to (4), semi-elasticities η^p are computed from linear regressions of memory scores (2+|k|) years in the past over years on changes in log marginal retention rates over two years (evaluated at the base year income data) , restricting the sample for those who are working in the baseline year and |k|-years before. All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Table 1.15: Income percentiles

	wg. 4y	wg (j.s.) 4y	partic. 2y	mental st. 2y
bottom third	-0.14	-0.40	0.087	-0.13
	(0.17)	(0.25)	(0.082)	(0.81)
middle third	-0.064	-0.37	0.19	-1.09
	(0.13)	(0.26)	(0.078)	(0.82)
upper third	-0.29	-0.22	0.035	-1.18
	(0.20)	(0.44)	(0.073)	(0.62)
Observations	39179	13958	72301	14705

Notes. Robust standard errors in parentheses. Elasticities and semi-elasticities are computed from linear regressions of different dependent variables on changes in log marginal retention rates over two years (evaluated at the base year income data). Each column corresponds to a different dependent variable: the first column reports results where the dependent variable are changes in log hourly wages over the next four years; the second column additionally restricts the sample to respondents who switched jobs over the same period. The third column reports results for changes in labor force participation over two years. The fourth column has a dependent variable mental status scores two years in the past, and restricts the sample to those who are working in the baseline year and two years ahead. All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Table 1.16: Occupations

	wg.	wg (j.s.)	partic.	mental st.
Managerial, sales, clerical	0.14	-0.10	0.10	0.48
	(0.17)	(0.29)	(0.087)	(0.70)
Professional	-0.43	-0.45	0.11	0.063
	(0.26)	(0.54)	(0.12)	(0.83)
Other services	0.21	0.17	-0.18	-4.69
	(0.23)	(0.42)	(0.14)	(1.95)
Farming, forestry, mechanics, construction	-0.038	-1.22	0.050	-2.60
	(0.43)	(0.79)	(0.16)	(1.70)
Operators	-0.49	-0.83	0.090	-2.59
	(0.22)	(0.38)	(0.13)	(1.38)
Observations	39179	13958	72301	14705

Notes. Robust standard errors in parentheses. Elasticities and semi-elasticities are computed from linear regressions of different dependent variables on changes in log marginal retention rates over two years (evaluated at the base year income data). Each column corresponds to a different dependent variable: the first column reports results where the dependent variable are changes in log hourly wages over the next four years; the second column additionally restricts the sample to respondents who switched jobs over the same period. The third column reports results for changes in labor force participation over two years. The fourth column has a dependent variable mental status scores two years in the past, and restricts the sample to those who are working in the baseline year and two years ahead. All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Table 1.17: Education levels

	wg.	wg (j.s.)	partic.	mental st.
high school or less	-0.27	-0.28	0.018	-2.38
	(0.12)	(0.23)	(0.067)	(0.76)
more than high school	-0.054	-0.39	0.18	0.43
	(0.15)	(0.26)	(0.062)	(0.52)
Observations	39179	13958	72301	14705

Notes. Robust standard errors in parentheses. Elasticities and semi-elasticities are computed from linear regressions of different dependent variables on changes in log marginal retention rates over two years (evaluated at the base year income data). Each column corresponds to a different dependent variable: the first column reports results where the dependent variable are changes in log hourly wages over the next four years; the second column additionally restricts the sample to respondents who switched jobs over the same period. The third column reports results for changes in labor force participation over two years. The fourth column has a dependent variable mental status scores two years in the past, and restricts the sample to those who are working in the baseline year and two years ahead. All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

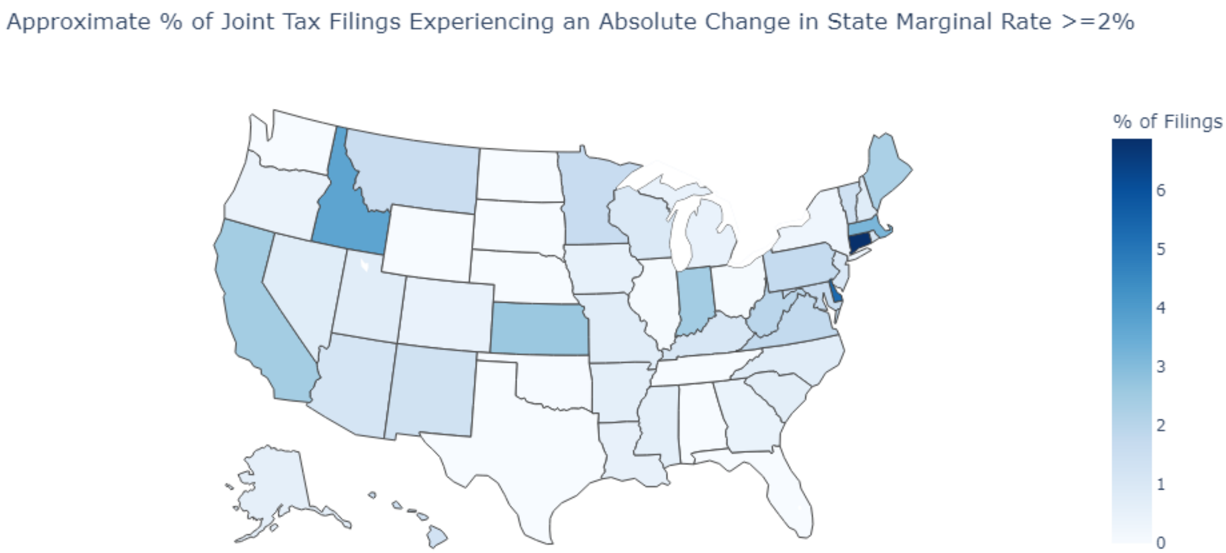
Table 1.18: Income percentiles and education

	wg.	wg (j.s.)	partic.	mental st.
bottom third, highschool or less	-0.28	-0.24	-0.020	-1.78
	(0.19)	(0.27)	(0.099)	(1.14)
middle third, highschool or less	-0.21	-0.33	0.10	-3.21
	(0.17)	(0.38)	(0.10)	(1.22)
upper third, highschool or less	-0.37	-0.30	-0.042	-2.50
	(0.37)	(0.88)	(0.16)	(1.63)
bottom third, more than highschool	0.15	-0.74	0.31	2.04
	(0.35)	(0.50)	(0.14)	(1.13)
middle third, more than highschool	0.093	-0.40	0.30	0.84
	(0.21)	(0.33)	(0.11)	(1.07)
upper third, more than highschool	-0.26	-0.20	0.061	-0.77
	(0.23)	(0.50)	(0.081)	(0.64)
Observations	39179	13958	72301	14705

Notes. Robust standard errors in parentheses. Elasticities and semi-elasticities are computed from linear regressions of different dependent variables on changes in log marginal retention rates over two years (evaluated at the base year income data). Each column corresponds to a different dependent variable: the first column reports results where the dependent variable are changes in log hourly wages over the next four years; the second column additionally restricts the sample to respondents who switched jobs over the same period. The third column reports results for changes in labor force participation over two years. The fourth column has a dependent variable mental status scores two years in the past, and restricts the sample to those who are working in the baseline year and two years ahead. All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

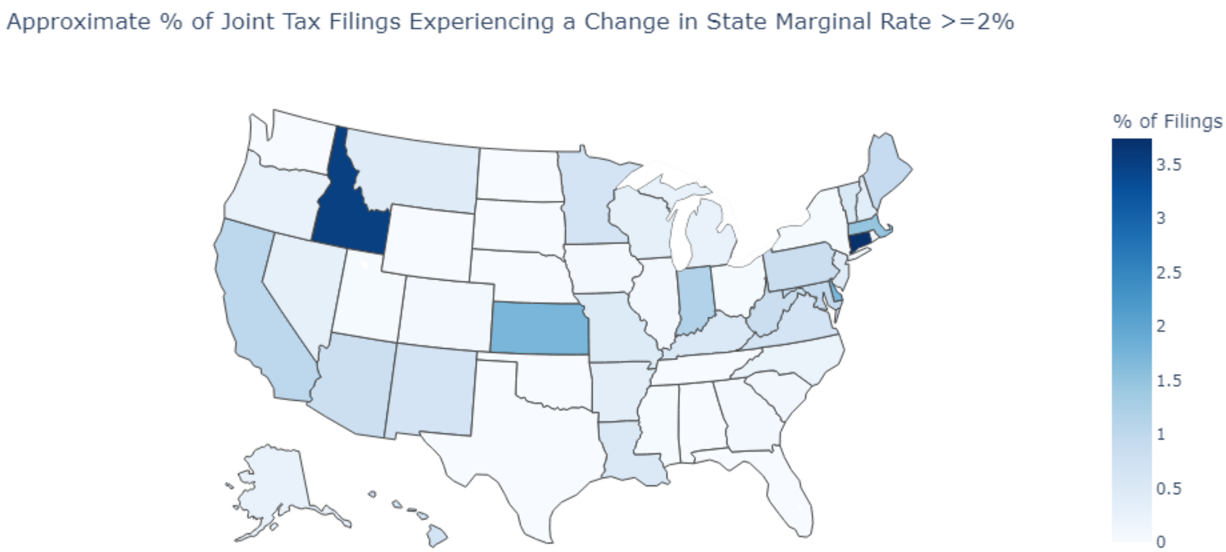
1.10.5 Figures

Figure 1.5: State variation in real marginal tax rates



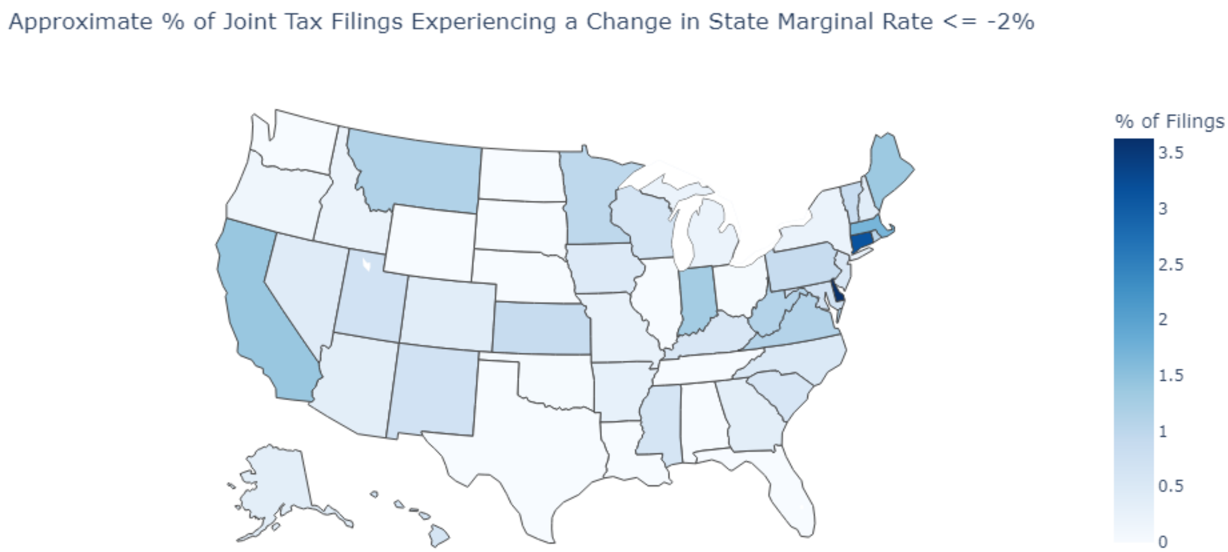
Notes. Percentages were calculated using the NBER tax simulator and the publicly available HRS RAND longitudinal file, imputing each state distribution of income from the national distribution of income. Shades are proportional to the percentage of jointly-filing taxpayers experiencing an absolute marginal tax change larger than 2%.

Figure 1.6: State variation in real marginal tax rates



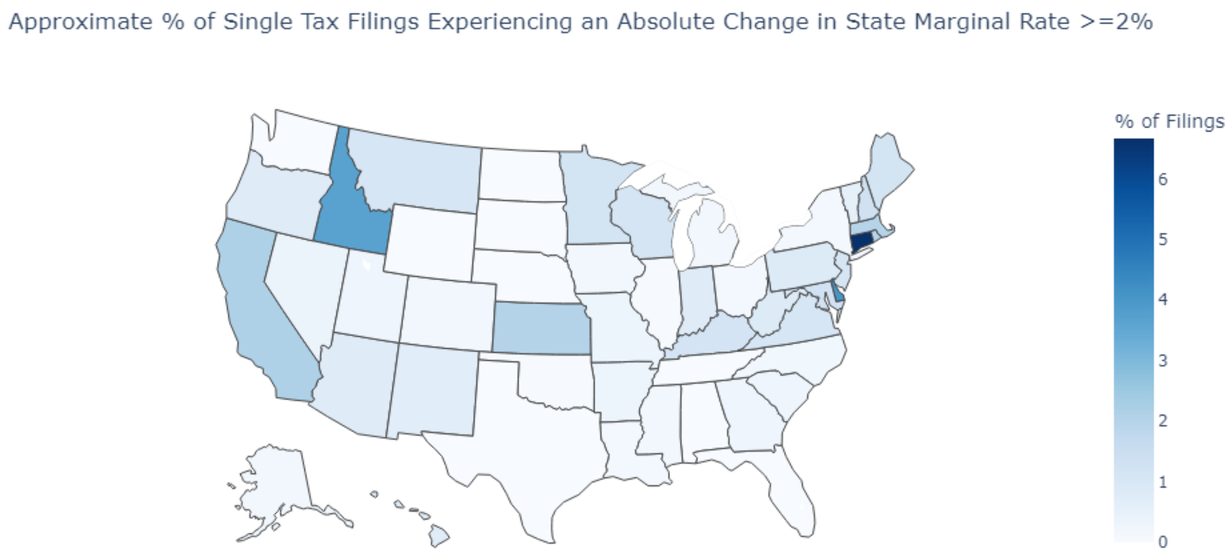
Notes. Percentages were calculated using the NBER tax simulator and the publicly available HRS RAND longitudinal file, imputing each state distribution of income from the national distribution of income. Shades are proportional to the percentage of jointly-filing taxpayers experiencing a marginal tax change larger than 2%.

Figure 1.7: State variation in real marginal tax rates



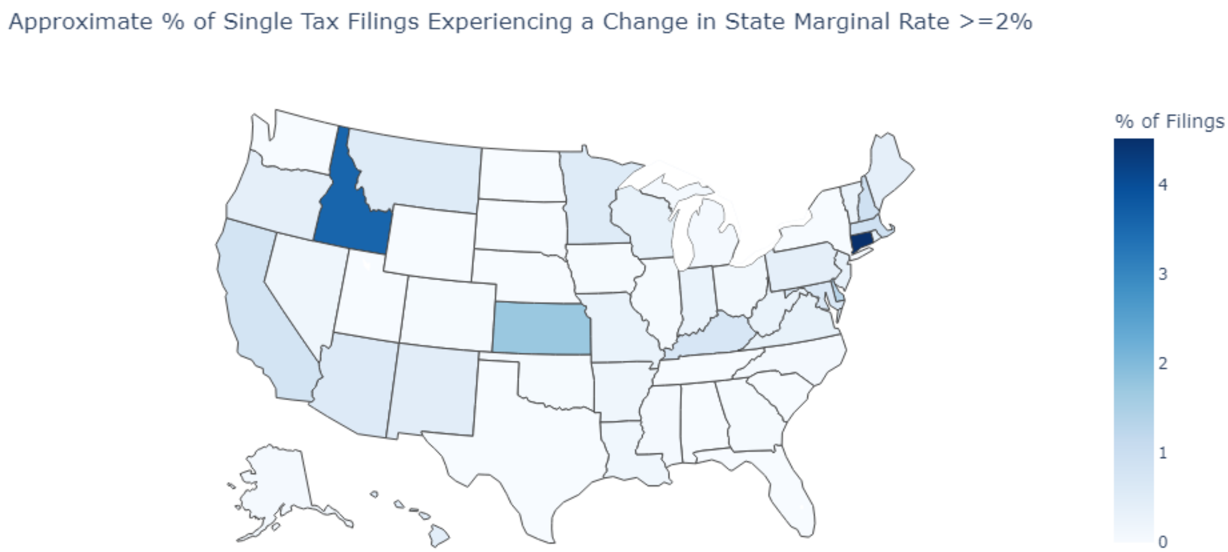
Notes. Percentages were calculated using the NBER tax simulator and the publicly available HRS RAND longitudinal file, imputing each state distribution of income from the national distribution of income. Shades are proportional to the percentage of taxpayers experiencing a marginal tax decrease larger than 2%.

Figure 1.8: State variation in real marginal tax rates



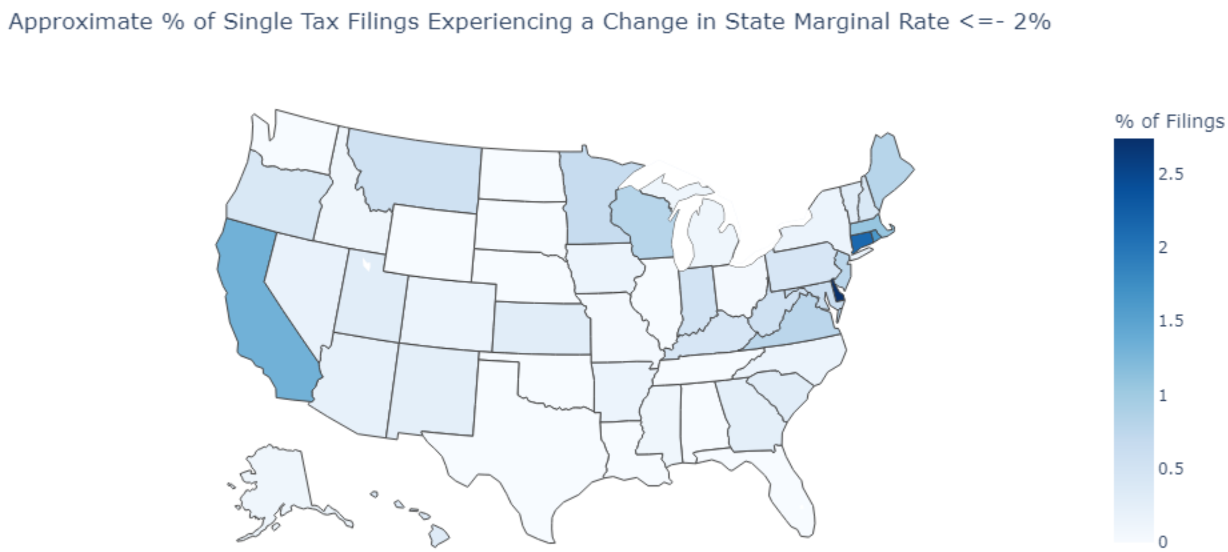
Notes. Percentages were calculated using the NBER tax simulator and the publicly available HRS RAND longitudinal file, imputing each state distribution of income from the national distribution of income. Shades are proportional to the percentage of single taxpayers experiencing an absolute marginal tax change larger than 2%.

Figure 1.9: State variation in real marginal tax rates



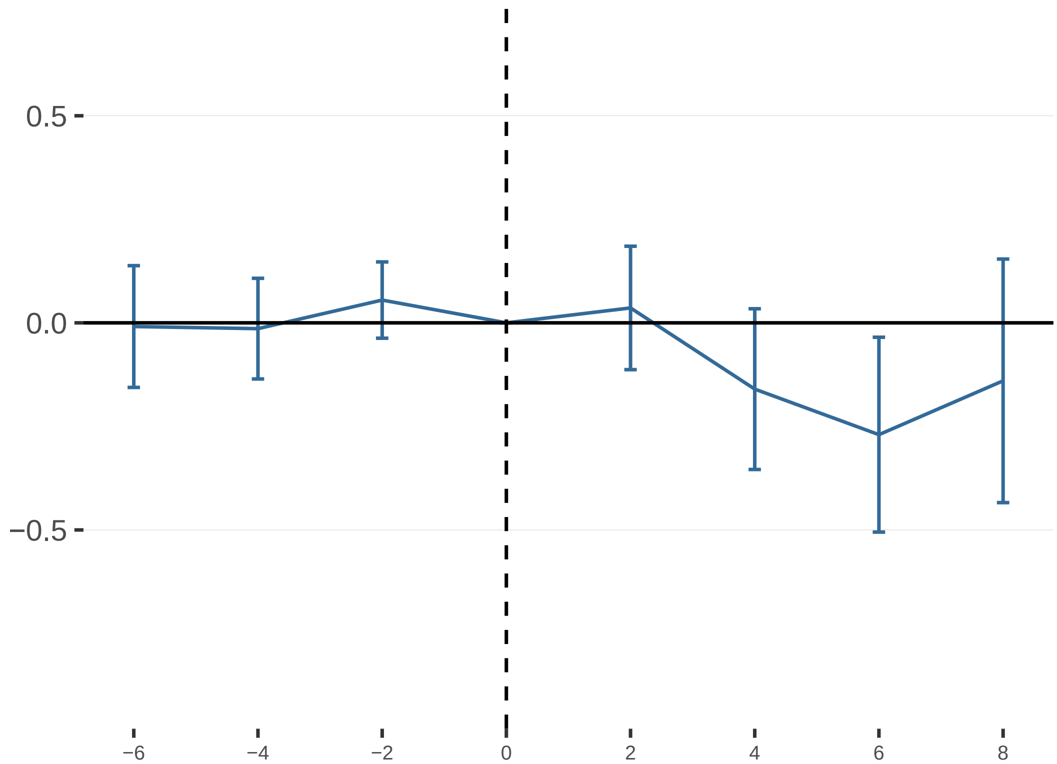
Notes. Percentages were calculated using the NBER tax simulator and the publicly available HRS RAND longitudinal file, imputing each state distribution of income from the national distribution of income. Shades are proportional to the percentage of single taxpayers experiencing a simulated marginal tax increase larger than 2%.

Figure 1.10: State variation in real marginal tax rates



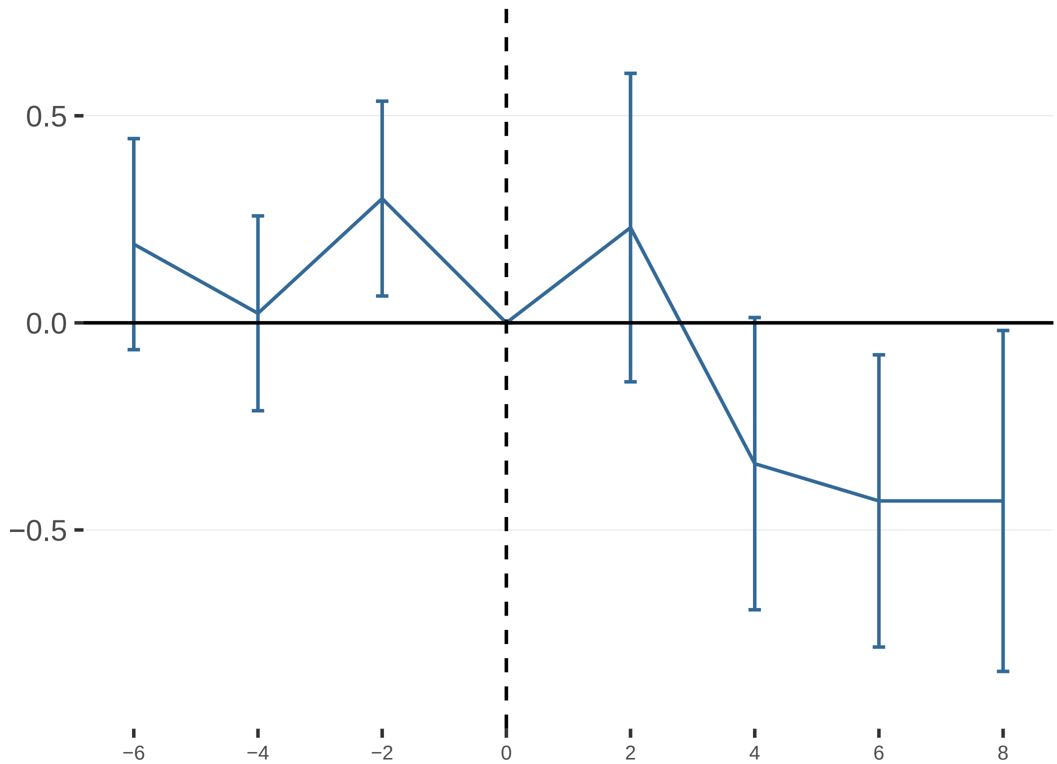
Notes. Percentages were calculated using the NBER tax simulator and the publicly available HRS RAND longitudinal file, imputing each state distribution of income from the national distribution of income. Shades are proportional to the percentage of single taxpayers experiencing a simulated marginal tax decrease larger than 2%.

Figure 1.11: Elasticity of wages over different horizons



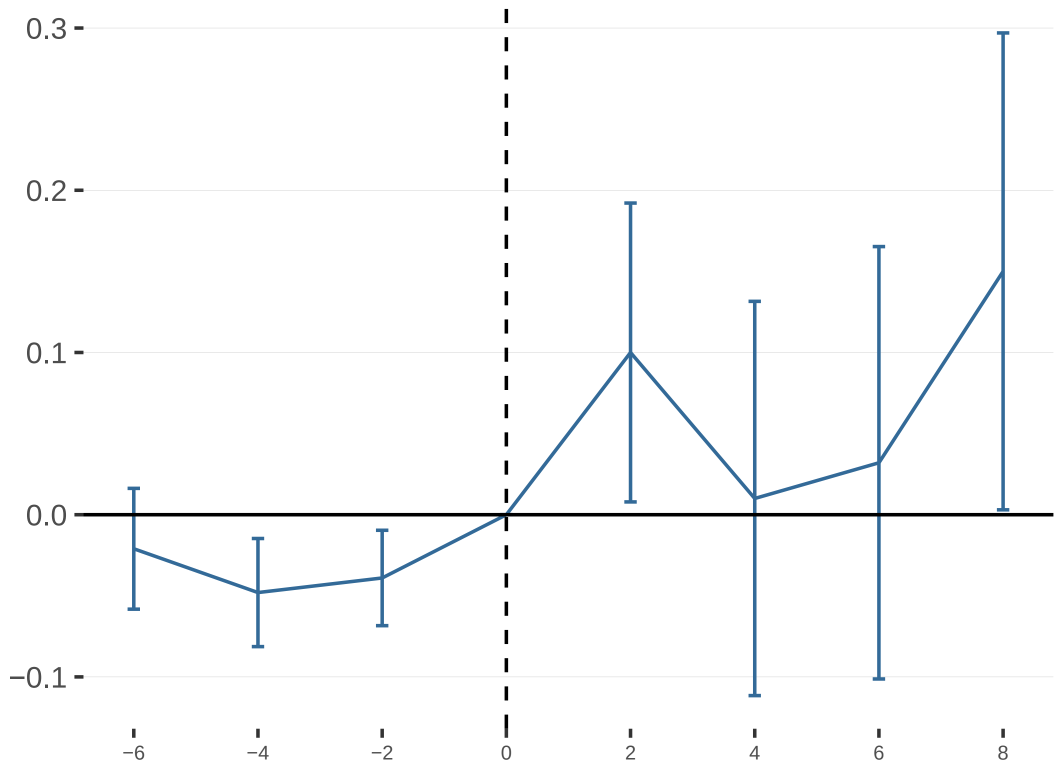
Notes. 95% confidence intervals. Elasticities are computed from linear regressions of changes in log hourly wages over k-years on changes in log marginal retention rates over two years (evaluated at the base year income data). All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Figure 1.12: Elasticity of wages for job switchers over different horizons



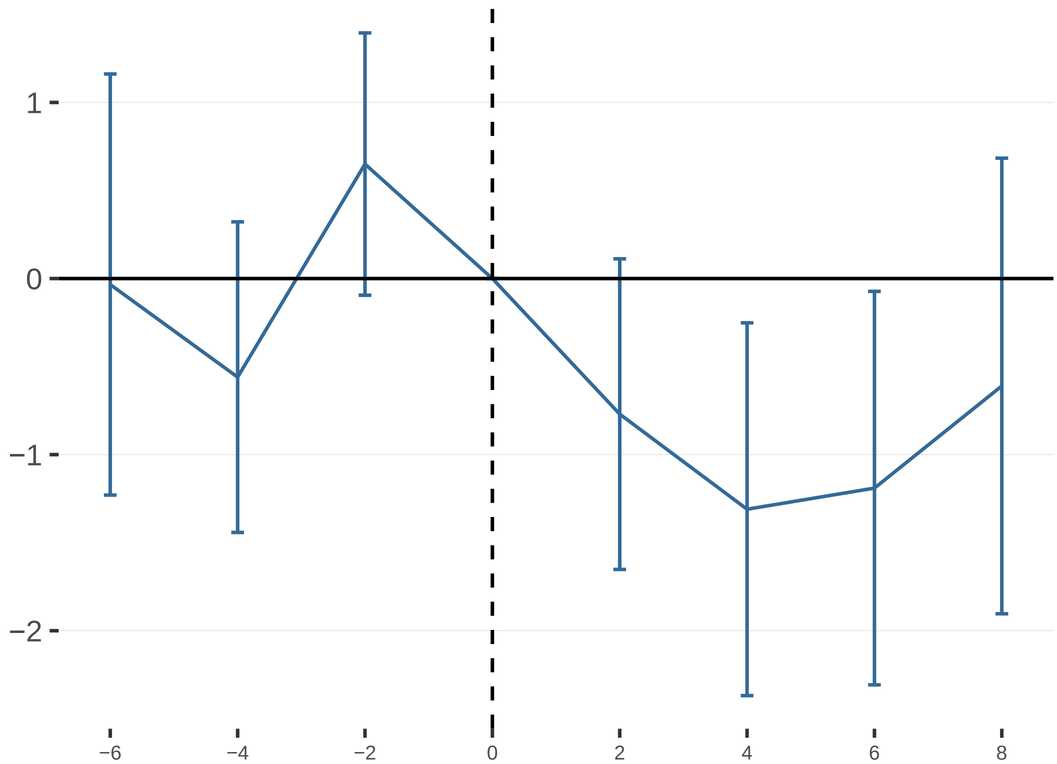
Notes. 95% confidence intervals. Elasticities are computed from linear regressions of changes of log hourly wages over k-years on changes in log marginal retention rates over two years (evaluated at the base year income data). For each column, the sample is restricted to those who switch jobs at least once between the baseline year and k years ahead. All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Figure 1.13: Semi-elasticities of participation over different horizons



Notes. 95% confidence intervals. Semi-elasticities are computed from linear regressions of changes of hours wages over k-years years on changes in log marginal retention rates over two years (evaluated at the base year income data). All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Figure 1.14: Semi-elasticities of mental status scores over different horizons

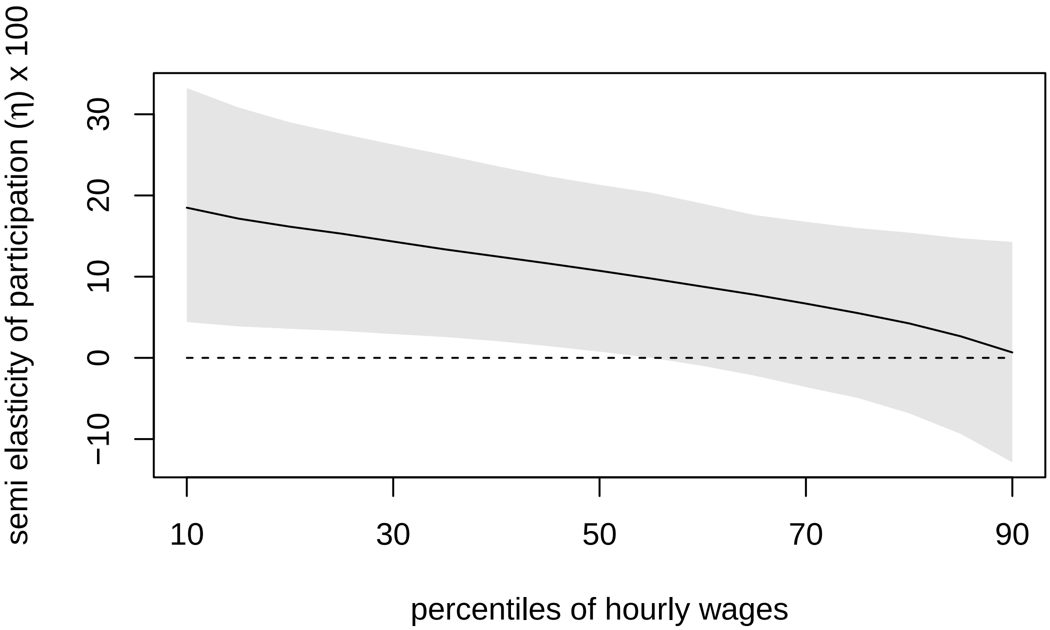


Notes. 95% confidence intervals. For the periods (k) between 2 and 8, semi-elasticities are computed from linear regressions of memory scores two years in the past over years on changes in log marginal retention rates over two years (evaluated at the base year income data), restricting the sample for those who are working in the baseline year and k-years ahead. Similarly, for the periods (k) between -6 to -2, semi-elasticities are computed from linear regressions of memory scores (2+|k|) years in the past over years on changes in log marginal retention rates over two years (evaluated at the base year income data), restricting the sample for those who are working in the baseline year and |k|-years before. All specifications include year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Figure 1.15: Elasticity of wages over different hourly wages percentiles

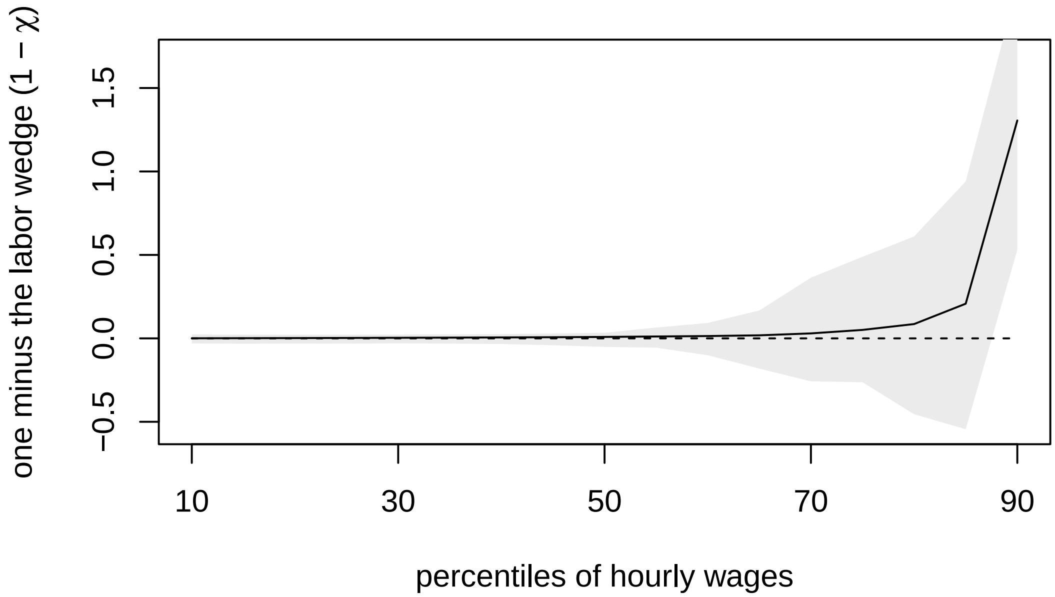
Notes. 95% bootstrap confidence intervals. Elasticities are computed from local linear regressions of changes in log hourly wages over four years on changes in log marginal retention rates over two years (evaluated at the base year income data), including year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Figure 1.16: Semi-elasticities of participation over different hourly wages percentiles



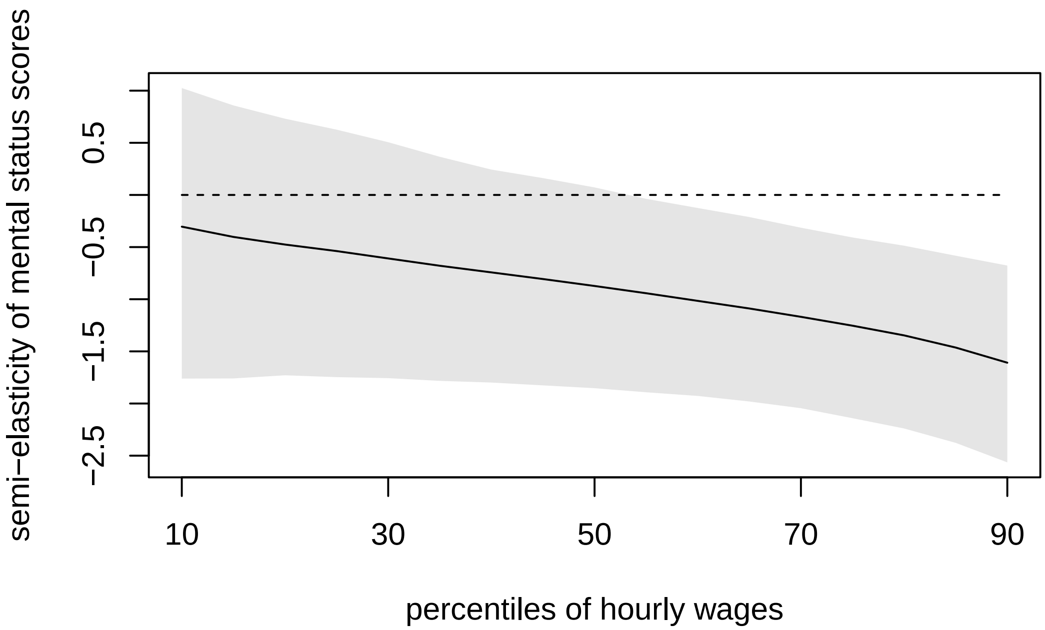
Notes. 95% bootstrap confidence intervals. Semi-elasticities are computed from linear regressions of changes of hours wages over two years on changes in log marginal retention rates over two years (evaluated at the base year income data), including year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Figure 1.17: Estimates for one minus the labor wedge for different hourly wages percentiles



Notes. 95% bootstrap confidence intervals. Estimates are computed from the ratio of elasticities of wages and semi-elasticities of participation. Elasticities of participation are computed from local linear regressions of changes in log hourly wages over four years on changes in log marginal retention rates over two years (evaluated at the base year income data). Semi-elasticities are computed from linear regressions of changes of hours wages over two years on changes in log marginal retention rates over two years (also evaluated at the base year income data). Each regression includes year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Figure 1.18: Semi-elasticities of mental status scores over different hourly wages percentiles



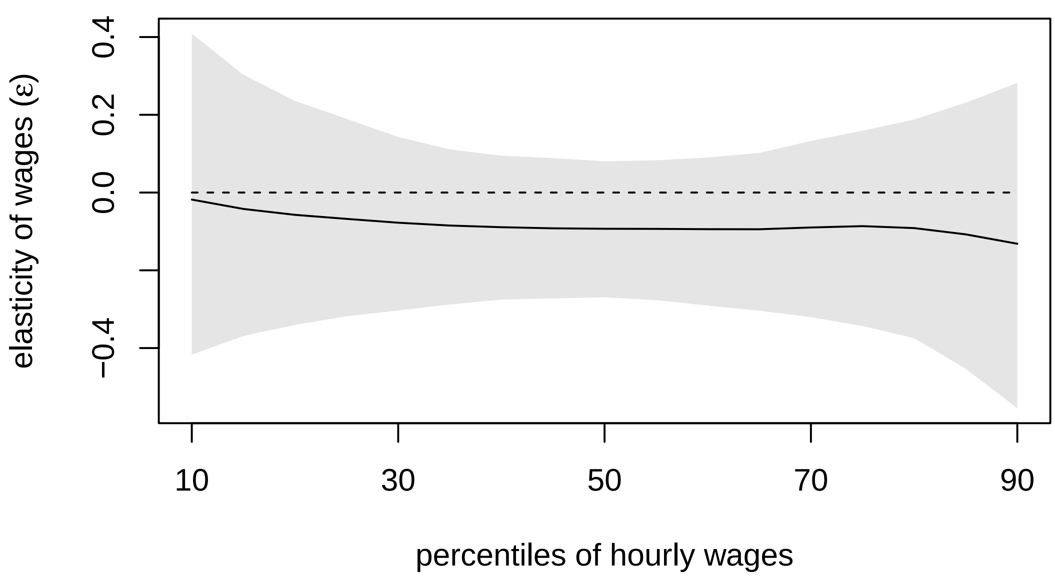
Notes. 95% bootstrap confidence intervals. Semi-elasticities are computed from linear regressions of memory scores 2 years in the past over years on changes in log marginal retention rates over two years (evaluated at the base year income data), restricting the sample to those who are working in the baseline year and 2 years ahead. Each regression includes year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

1.10.6 Robustness Checks

Bandwidth Choice

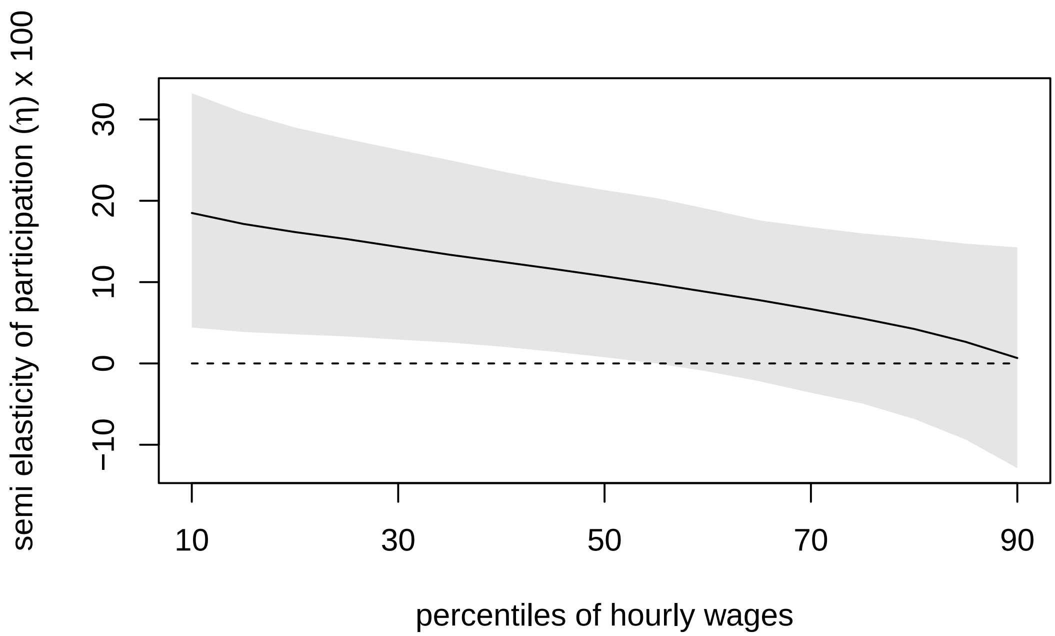
The figures 1.14 to 1.18 show non-parametric regressions with an optimal bandwidth selected by a computationally tractable leave-one-out cross-validation criterium (Racine, 1993). This section shows that results are not driven by the choice of bandwidth, presenting results for when the bandwidths are set to half the size of the optimized bandwidths from the previous section.

Figure 1.19: Elasticity of wages over different hourly wages percentiles



Notes. 95% bootstrap confidence intervals. Elasticities are computed from local linear regressions of changes in log hourly wages over four years on changes in log marginal retention rates over two years (evaluated at the base year income data), including year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Figure 1.20: Semi-elasticities of participation over different hourly wages percentiles

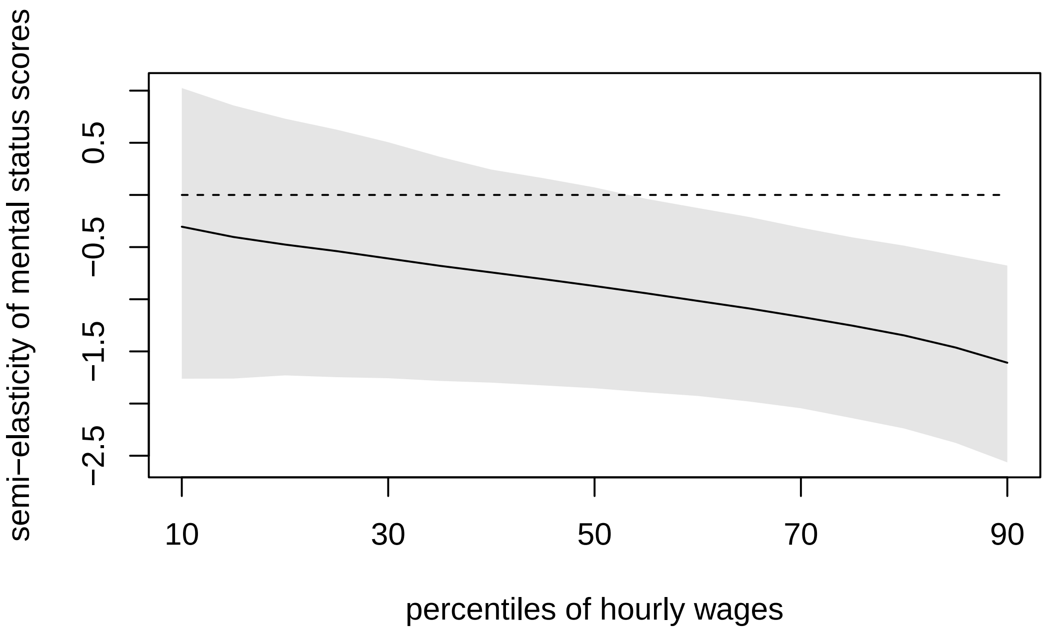


Notes. 95% bootstrap confidence intervals. Semi-elasticities are computed from linear regressions of changes of hours wages over two years on changes in log marginal retention rates over two years (evaluated at the base year income data), including year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Figure 1.21: Estimates for one minus the labor wedge for different hourly wages percentiles

Notes. 95% bootstrap confidence intervals. Estimates are computed from the ratio of elasticities of wages and semi-elasticities of participation. Elasticities of participation are computed from local linear regressions of changes in log hourly wages over four years on changes in log marginal retention rates over two years (evaluated at the base year income data). Semi-elasticities are computed from linear regressions of changes of hours wages over two years on changes in log marginal retention rates over two years (also evaluated at the base year income data). Each regression includes year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Figure 1.22: Semi-elasticities of mental status scores over different hourly wages percentiles



Notes. 95% bootstrap confidence intervals. Semi-elasticities are computed from linear regressions of memory scores 2 years in the past over years on changes in log marginal retention rates over two years (evaluated at the base year income data), restricting the sample to those who are working in the baseline year and 2 years ahead. Each regression includes year fixed effects, marital status dummies, and a 10-piece linear spline of log hourly wages.

Set of Controls

In the main part of the paper we followed the empirical literature on the elasticties of taxable income in our choice of the set of controls, which include year-fixed effects, marital status, and initial hourly wages. In this section, we report results when this set also includes age dummies, which are often included in the set of control variables when estimating the effects of accruals and retirement benefits on retirement decisions. The results are only slighlty changed, suggesting that initial set of control variables was rich enough, and there is not much of a correlation between the individual changes in marginal tax rates and age dummies after netting out the effects of year dummies, marital status, and initial hourly wages.

Table 1.19: Elasticities of wages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ϵ^w	-0.13	-0.061	-0.066	-0.069	-0.100	-0.16	-0.16
	(0.10)	(0.10)	(0.10)	(0.10)	(0.099)	(0.099)	(0.099)
year f.e.	no	yes	yes	yes	yes	yes	yes
marital status	no	no	yes	yes	yes	yes	yes
age	no	no	no	yes	yes	yes	yes
hourly wages	no	no	no	no	linear	c. spline	l. spline
observations	39179	39179	39179	39179	39179	39179	39179

Notes. Robust standard errors in parentheses. Elasticities ϵ^w are computed from linear regressions of changes in log hourly wages over four years on changes in log marginal retention rates over two years (evaluated at the base year income data). Each column includes different sets of controls: year fixed effects, marital status dummies, age, and hourly wages. Column (4) includes log hourly wages. Column (5) includes a 5 piece cubic spline of log hourly wages. Column (6) includes a 10 piece linear spline of log hourly wages.

Table 1.20: Elasticities of wages for job switchers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ϵ^w	-0.31	-0.19	-0.19	-0.20	-0.29	-0.32	-0.34
	(0.18)	(0.19)	(0.19)	(0.19)	(0.18)	(0.18)	(0.18)
year f.e.	no	yes	yes	yes	yes	yes	yes
marital status	no	no	yes	yes	yes	yes	yes
age	no	no	no	yes	yes	yes	yes
hourly wages	no	no	no	no	linear	c. spline	l. spline
observations	13958	13958	13958	13958	13958	13958	13958

Notes. Robust standard errors in parentheses. Elasticities ϵ^w are computed from linear regressions of changes in log hourly wages over four years on changes in log marginal retention rates over two years (evaluated at the base year income data). The sample is restricted to those who switch jobs at least once between the baseline year and four years later. Each column includes different sets of controls: year fixed effects, marital status dummies, age, and hourly wages. Column (4) includes log hourly wages. Column (5) includes a 5 piece cubic spline of log hourly wages. Column (6) includes a 10 piece linear spline of log hourly wages. *Notes.* Robust standard errors in parentheses. Elasticities ϵ^w are computed from linear regressions of changes in log

Table 1.21: Elasticities of wages

	-8y	-6y	-4y	-2y	+2y	+4y	+6y	+8y
ϵ^w	0.12	-0.0075	-0.016	0.054	0.035	-0.16	-0.27	-0.16
	(0.11)	(0.075)	(0.062)	(0.047)	(0.076)	(0.099)	(0.12)	(0.15)
N	17898	25425	34685	45922	53931	39179	29253	21119

hourly wages over k-years on changes in log marginal retention rates over two years (evaluated at the base year income data). All specifications include year fixed effects, marital status dummies, age dummies, and a 10 piece linear spline of log hourly wages.

Table 1.22: Elasticities of wages for job switchers

	-8y	-6y	-4y	-2y	+2y	+4y	+6y	+8y
ϵ^w	0.36	0.19	0.037	0.30	0.23	-0.34	-0.40	-0.43
	(0.16)	(0.13)	(0.12)	(0.12)	(0.19)	(0.18)	(0.18)	(0.21)
N	10040	11950	12375	9430	10996	13958	13703	11737

Notes. Robust standard errors in parentheses. Elasticities ϵ^w are computed from linear regressions of changes of log hourly wages over k-years on changes in log marginal retention rates over two years (evaluated at the base year income data). For each column, the sample is restricted to those who switch jobs at least once between the baseline year and k years ahead. All specifications include year fixed effects, marital status dummies, age dummies, and a 10 piece linear spline of log hourly wages.

Notes. Robust standard errors in parentheses. Semi-elasticities η^p are computed from linear regressions of changes of hours

Table 1.23: Participation semi-elasticities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
η^p	-0.039	0.063	0.044	-0.020	-0.013	-0.018	-0.021
	(0.056)	(0.058)	(0.056)	(0.054)	(0.060)	(0.060)	(0.060)
year f.e.	no	yes	yes	yes	yes	yes	yes
marital status	no	no	yes	yes	yes	yes	yes
age	no	no	no	yes	yes	yes	yes
hourly wages	no	no	no	no	linear	c. spline	l. spline
observations	72526	72526	72526	72526	61526	61526	61526

wages over four years on changes in log marginal retention rates over two years (evaluated at the base year income data). Each column includes different sets of controls: year fixed effects, marital status dummies, age dummies, and hourly wages. Column (4) includes log hourly wages. Column (5) includes a 5 piece cubic spline of log hourly wages. Column (6) includes a 10 piece linear spline of log hourly wages. *Notes.* Robust standard errors in parentheses. Semi-elasticities η^p are computed from linear

Table 1.24: Participation semi-elasticities

	-8y	-6y	-4y	-2y	+2y	+4y	+6y	+8y
η^p	-0.039	-0.021	-0.049	-0.040	0.087	-0.021	-0.011	0.11
	(0.026)	(0.019)	(0.017)	(0.015)	(0.046)	(0.060)	(0.066)	(0.072)
N	20522	29053	39482	51895	72301	61526	53964	46435

regressions of changes of hours wages over k-years years on changes in log marginal retention rates over two years (evaluated at the base year income data). All specifications include year fixed effects, marital status dummies, age dummies, and a 10 piece linear spline of log hourly wages. *Notes.* Robust standard errors in parentheses. Semi-elasticities η^m are computed

Table 1.25: Mental status scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
η^m	0.14	-0.82	-1.08	-0.98	-1.02	-1.12	-1.12
	(0.42)	(0.51)	(0.51)	(0.51)	(0.55)	(0.54)	(0.54)
year f.e.	no	yes	yes	yes	yes	yes	yes
marital status	no	no	yes	yes	yes	yes	yes
age	no	no	no	yes	yes	yes	yes
hourly wages	no	no	no	no	linear	c. spline	l. spline
observations	16027	13187	13187	13187	11030	11030	11030

from linear regressions of mental status scores two years in the past on changes in log marginal retention rates over two years (evaluated at the base year income data), restricting the sample for those who are working in the baseline year and four years ahead. Each column includes different sets of controls: year fixed effects, marital status dummies, age dummies, and hourly wages. Column (4) includes log hourly wages. Column (5) includes a 5 piece cubic spline of log hourly wages. Column (6) includes a 10 piece linear spline of log hourly wages. *Notes.* Robust standard errors in parentheses. For columns (5) to (8),

Table 1.26: Mental status scores

	-8y	-6y	-4y	-2y	+2y	+4y	+6y	+8y
η^m	0.32	0.0097	-0.57	0.67	-0.65	-1.12	-1.02	-0.40
	(0.55)	(0.60)	(0.45)	(0.38)	(0.45)	(0.54)	(0.57)	(0.66)
N	4362	5784	9566	13055	14705	11030	8264	5005

semi-elasticities η^p are computed from linear regressions of mental status scores two years in the past over years on changes in log marginal retention rates over two years (evaluated at the base year income data) , restricting the sample for those who are working in the baseline year and k-years ahead. For columns (1) to (4), semi-elasticities η^m are computed from linear regressions of mental status scores (2+k) years in the past over years on changes in log marginal retention rates over two years (evaluated at the base year income data) , restricting the sample for those who are working in the baseline year and k-years before. All specifications include year fixed effects, marital status dummies, age dummies, and a 10 piece linear spline of log hourly wages.

Chapter 2

Income Taxation with Elasticity Heterogeneity

Joint with John Sturm

2.1 Introduction

It is uncontroversial that different people respond to incentives differently. Tax responses are no exception. Indeed, heterogeneity in tax responsiveness is a consistent finding of the empirical literature on the elasticity of taxable income (ETI). In part, this reflects that some households—such as secondary earners and those near retirement—simply have more elastic labor supply (Blau and Kahn, 2007; Eissa and Liebman, 1996; Vere, 2011). It also reflects that some households pay more attention to taxes, are more adept at taking advantage of itemizations, or are more able to avoid taxes all together (Taubinsky and Rees-Jones, 2018; Gruber and Saez, 2002; Kopczuk, 2005).

Does heterogeneity in ETIs matter for income taxation? There is good reason to think it may be unimportant: To first-order, the effects of any tax change on everything a planner values—i.e. tax revenue and the distribution of household welfare—can be computed simply from the *mean* elasticities at each income level (Werning, 2007; Jacquet and Lehmann, 2015; Scheuer and Werning, 2018). But beyond the first-order approach, it is less obvious what role heterogeneity may play.

This paper explores the simple but overlooked idea that—unlike the first-order condition explored in the literature—the planner's *second-order condition* depends intimately on the structure of household heterogeneity. The second-order condition captures how the effects of an infinitesimal variation in taxes *change* when one repeats that variation for a second time. One key source of such changes is that, due to each household's response to the first variation, it experiences the second tax variation at a new income level. Notably, if households with the same initial income respond to the first variation differently and the tax change differs across incomes, then the second variation can even have different effects on households who faced the same taxes ex-ante.

The advantage of this differential taxation is that—all else equal—the planner prefers to increase taxes on households with lower ETI, whose smaller responses result in smaller tax revenue losses. While the planner cannot explicitly condition taxes on ETI, we show that whenever there is enough heterogeneity in ETI, she can construct a particular tax variation that allows her to set a tax system *as if* she could directly condition on the elasticities of taxable income, via a sorting mechanism. Namely, the variation causes low-ETI households to move to incomes where marginal taxes increase and high-ETI households to move to incomes where marginal taxes decrease. Repeating the same variation for a second time, therefore, *disproportionately targets* marginal tax increases toward those who respond to them less. In other words, the planner "sorts" households on ETI and then "extorts" them by raising taxes disproportionately more on the unresponsive.

Our main theoretical results encapsulate these ideas in a simple, planner-agnostic "rationalizability test" for the tax-schedule, formulated in terms of locally observable sufficient statistics. Our test passes if and only if *there exists a social planner* within a broad class for whom the tax schedule is optimal among all nearby schedules. While our test—by virtue of remaining agnostic about planner preferences—is relatively permissive, it is at the same time stricter than the familiar test for constrained Pareto efficiency. This is because we impose some mild restrictions on planner preferences: We assume that the social planner maximizes a weighted sum of household utilities, where the weight on any household may be endogenous so long as (a) it changes smoothly in that household's utility and (b) the average weight at each income is continuous in income. These assumptions are satisfied by most welfarist objectives studied in the optimal taxation literature. By contrast, a "Pareto planner"—who seeks to bring each household's utility above a threshold—changes households' welfare weights discontinuously at those thresholds. We show that this subtle distinction has important implications, giving rise to a further requirement that taxes need to satisfy, specifically, a new second-order condition.

Interestingly, we show that this condition is not only necessary but—when combined with a standard first-order condition—also locally sufficient for tax rationalizability. This is surprising because this condition is constructed by looking at only a small subset of possible deviations from a given tax schedule. However, we demonstrate that all other tax deviations have significant welfare effects on a finite mass of households, which implies that they are sub-optimal for a planner who is sufficiently biased toward the status quo. Our test therefore provides a complete characterization of local tax rationalizability.

This analysis benefits in several ways from our approach of "testing" tax schedules by asking whether some preferences rationalize them instead of fixing preferences and "solving" for the optimal schedule. First, we can avoid embedding strong normative judgments into the positive analysis of tax systems. Second, we need not extrapolate locally observed sufficient statistics to non-local tax schedules where one does not know their true values. Third, in our particular setting, the version of the second-order condition that guarantees the existence of preferences that rationalize the tax schedule is dramatically simpler than the second-order condition for any specific utilitarian objective function.

In contrast to the existing literature (e.g. the first-order Pareto efficiency test of Werning (2007)), our second-order test depends not only on the mean ETI at each income level, but also on its variance, reflecting the scope for a planner to "sort and extort" households by elasticity. In light of the well-documented empirical variation in ETIs, this raises the question: "Is the heterogeneity in ETIs large enough that real world tax schedules can be improved?" The second part of the paper seeks to answer this question by evaluating our rationalizability test in a panel of US tax returns from 1979 to 1990. Doing so requires us to move beyond the existing literature on ETI estimation, which does not study the variance of elasticities conditional on income.

We estimate ETI variances with a number of empirical strategies: First, we provide a lower bound on these variances by estimating mean elasticities and group size as nonparametric functions of income for two different groups of taxpayers—those who claim a high number of itemizations and those who claim a low number of itemizations—and computing the variance of elasticities across those two groups at each income level. Second, we take a less conservative structural approach, postulating and estimating a flexible distribution for elasticities, and then using it to obtain non-parametric point estimates for the variance of elasticities at each income level. Third, we validate these explicit estimates by providing reduced-form evidence for the sorting mechanism that underlies failures of the second-order condition. Specifically, we document that income-conditional ETIs change across years as predicted by the differential reshuffling of high- and low-elasticity households in response to observed tax changes.

Our preferred estimates have stark implications for income taxation. Although the first-order test passes, we find that our novel second-order test fails in every year of our sample. This implies it is impossible to rationalize the income tax schedule within the set of social planners we consider. Said differently, any social planner willing to make at least some minimal welfare trade-offs between households would prefer a different tax schedule. In this sense, a "free lunch" is available through tax reform. A conservative quantification exercise suggests that either raising or lowering top taxes by 20 percentage points results in yearly welfare gains equivalent to approximately $3000 per top earner.

Related Literature Theoretically, we contribute to a long literature on the design of non-linear income taxation schedules when labor is supplied on the intensive margin and production is linear. Recent work by Werning (2007), Jacquet and Lehmann (2015), and Scheuer and Werning (2018) has shown that the so-called "ABC formula" of Diamond (1998)—a necessary first-order condition for optimality—extends to the case of multi-dimensional heterogeneity if one simply uses the average elasticities at each income level.[1] We complement this result by providing an additional necessary condition—a second-order condition—which, when combined with the first-order condition, is also sufficient. To our knowledge, we are the first to explicitly study the planner's second-order condition with respect to non-linear income taxes.[2]

The approach of "testing" a tax schedule rather than "solving" for optimal taxes is most similar to that of Werning (2007); Hosseini and Shourideh (2019), and Bierbrauer et al. (2020). However, our notion of "rationalizability" departs slightly from the Pareto efficiency criterion these papers study by requiring the planner's utilitarian objective to be smooth with respect to each household's utility, as well as continuous across households of similar income levels. This small distinction leads to large differences in our conclusions. Namely, the main result of Bierbrauer et al. (2020) implies that, in our setting, the first-order condition of Werning (2007) is not only necessary but also locally sufficient for taxes to be Pareto efficient, even when one allows for multi-dimensional heterogeneity. By contrast, we show that a distinct second-order condition is necessary—and together with the first-order condition, locally sufficient—in order for taxes to be rationalizable by a planner who values household utility in a smooth way. While our result does not apply to a Pareto planner—since her desire to make transfers to a given household changes discontinuously when that household's utility crosses a threshold—it does apply to arbitrarily accurate smooth approximations of any Pareto objective. Under any such smooth approximation, the planner becomes willing to tolerate a small welfare loss to any household if it is accompanied by a large enough gain for other households. Our results therefore illustrate a sense in which the Pareto efficiency criterion is knife-edge, falling apart once one allows arbitrarily small trade-offs to be made across households.

Closely related to our paper is Jacquet and Lehmann (2020), who highlight shifts in the elasticity-composition of the income distribution as a source of endogeneity in the average elasticity at any income. They provide analytical cases and numerical examples where accounting for heterogeneity leads to a larger

[1] Also see a much earlier informal derivation by Saez (2001).

[2] Werning (2007) shows that when households vary along one dimension and their preferences have a certain functional form, the planner's problem is concave. This implies that the second-order condition *would hold*, were one to compute it.

gap between optimal and actual income taxes at high incomes.[3] We also emphasize these compositional shifts, but take a different approach. First, by testing existing tax schedules rather than solving for new optimal schedules, we emphasize a novel and distinct aspect of elasticity heterogeneity—that it can be so strong to prevent existing tax schedules from being rationalized by any (smooth) social preferences. Second, our results accommodate arbitrary multidimensional heterogeneity in preferences. The generality of our analytical results allows us to decompose the contributions of many empirically-relevant quantities—such as the shape of the income distribution and the curvature of the tax schedule—in mediating compositional effects. This generality also allows us to connect our test to the data without restrictive assumptions on the joint distribution of elasticities and productivities.

Our findings also complement other existing theoretical results in public finance and mechanism design. First, we share with the random taxation literature an emphasis on the non-convex nature of certain taxation problems (Stiglitz, 1982b,a; Hines Jr and Keen, 2021). While this literature shows that stochastic tax schedules can be welfare-enhancing in some circumstances,[4] we demonstrate that under similar—and testable—conditions, targeted deterministic deviations can provide an alternative way to improve outcomes. Second, our interest in targeting households according to their elasticities relates to a strand of the literature that proposes using multi-dimensional policy instruments to improve efficiency in the presence of multi-dimensional heterogeneity. Concretely, Moser and Olea de Souza e Silva (2019) argue that income-contingent retirement savings policies can be used to more efficiently screen on productivity when households differ in terms of both productivity and present bias. We show how a social planner can, in fact, use a single-dimensional instrument to target households "as if" they could condition on untaxed characteristics.[5]

On the empirical side, we build on the approach of Gruber and Saez (2002) to estimate elasticities of taxable income (ETIs). We use the same NBER panel of tax returns as these authors and a similar identification strategy, identifying ETIs from households' changes in income following changes in taxes. Our work extends beyond Gruber and Saez (2002) in two ways. First, we estimate elasticities *conditional on income level* nonparametrically using local polynomial methods. Second, we place a greater emphasis on the *variance* of ETI across households. Of course, our interest in ETI heterogeneity has significant precedent in the literature: Gruber and Saez (2002) and Kopczuk (2005) show that itemizers are more elastic, and related work on labor supply elasticities suggests that ETIs may be higher for second-earners, single mothers, and those near retirement (Blau and Kahn, 2007; Eissa and Liebman, 1996; Vere, 2011). More recent work by Kumar and Liang (2020) emphasizes heterogeneity in average ETIs *between income brackets* as a key source of variation across ETI estimates in the literature.[6] What distinguishes our approach to the variation in ETIs is an interest in a new, model-implied, sufficient statistic: the variance in ETI *conditional on income level*. We employ a number of identification strategies tailored to this statistic. Among these strategies is a structural procedure that relies on flexible parametric assumptions on the distribution of elasticities

[3] In a similar vein, Lockwood et al. (2020) discuss how accounting for a planner's uncertainty over households' taxable income elasticities affects optimal taxes. Although different in its motivation, we view Lockwood et al. (2020) as closely related to Jacquet and Lehmann (2020) since—when a planner faces an expected budget constraint—uncertainty over states of the world and heterogeneity across households are interchangeable for the purposes of income taxation.

[4] This possibility appears more generally in other mechanism design problems, see for example Maskin and Riley (1984b,a).

[5] More generally, our work relates to the multidimensional screening and nonlinear pricing literatures. Whereas the multidimensional screening literature has emphasized direct mechanisms and revelation principle (e.g. Rochet and Stole (2003)), we apply the taxation principle and focus on the indirect mechanism (Hammond, 1979; Rochet, 1985; Guesnerie, 1998). This approach allows us to sidestep challenging technical issues and derive new conditions for optimality. Our approach can easily be applied to non-linear pricing problems, where our second-order condition would translate to a condition for the local convexity of revenues.

[6] Although we do not focus on inattention to tax changes, our estimation is related in spirit to Taubinsky and Rees-Jones (2018), who estimate heterogeneity in responses to sales taxes.

and on the linearity of tax responses. We combine this identification strategy with methods from the recent literature on discrete-approximation algorithms in econometrics (Bonhomme and Manresa, 2015; Bonhomme et al., 2017; Lewis et al., 2019).

Outline: The paper is organized as follows: Section 2.2 presents a simple example that introduces the key ideas of the paper. Section 2.3 lays out the formal model. Section 2.4 presents and discusses our theoretical test for rationalizability. Section 2.5 presents our empirical estimates and evaluates our test using US tax return. Section 2.6 presents a simple quantification of the welfare losses from un-rationalizable tax policy. Section 2.7 discusses our findings and Section 2.8 concludes.

2.2 Motivating example

Before proceeding to the general model, we present a simple example that introduces the key ideas featured in our later formal results.[7]

A unit measure μ of households $h \in \mathcal{H}$ supply labor and consume a good—produced one-for-one with labor—in a static economy. Households face a non-linear tax schedule T. Each household h's preferences are additively separable between consumption and labor, feature a constant elasticity of labor dis-utility, and have no income effects:

$$V^h(T) \quad \equiv \quad \max_z \quad z - T(z) \quad - \quad \frac{z^{1+\frac{1}{\beta^h}}}{1+\frac{1}{\beta^h}} \Big/ (\theta^h)^{\frac{1}{\beta^h}}, \tag{2.1}$$

where $\theta^h > 0$ and $\beta^h > 0$ are productivity and elasticity parameters. We let $z^h(T)$ denote the maximizer. Within elasticity groups, productivity is Pareto with a common shape, i.e. $\theta^h \mid \beta^h \sim \text{Pareto}(\alpha > 1)$.

Now, consider the problem of a planner reforming a constant top tax rate $\bar{\tau}$ that applies to all income earned above income level $\bar{z}$.[8] Let $T_{\bar{\tau}}$ denote the entire schedule as a function of $\bar{\tau}$. We assume the planner places a constant weight $\widetilde{\lambda}$ on transfers to each household in the top bracket, relative to tax revenue, and therefore chooses $\bar{\tau}$ to maximize a weighted sum of welfare and tax revenue in the top bracket, or

$$\mathcal{L}(\bar{\tau}) \quad \equiv \quad \widetilde{\lambda} \cdot W_{\text{top}}(\bar{\tau}) \; + \; \text{Rev}_{\text{top}}(\bar{\tau}) \quad = \quad \mathbb{E}_h \left[\widetilde{\lambda} \cdot V^h(T_{\bar{\tau}}) \; + \; \bar{\tau} \cdot z^h(T_{\bar{\tau}}) \;\middle|\; z^h(T_{\bar{\tau}_0}) \geq \bar{z} \right], \tag{2.2}$$

Our assumptions imply that, within each elasticity group, the average post-reform income of households with pre-reform incomes in the top bracket is a power function of the top tax rate: $\mathbb{E}_h \left[z^h(T_{\bar{\tau}}) \;\middle|\; z^h(T_{\bar{\tau}_0}) \geq \bar{z}, \; \beta^h = \beta \right] \propto (1-\bar{\tau})^{\alpha\beta}$.

A natural starting place for this analysis is the standard first-order condition for the revenue effects of a small increase in top taxes around its initial level $\bar{\tau}_0$.

$$\left.\frac{d}{d\bar{\tau}}\right|_{\bar{\tau}=\bar{\tau}_0} \mathcal{L}(\bar{\tau}) \quad \propto \quad \underbrace{-\widetilde{\lambda}}_{\text{Welfare effect}} \quad + \quad \underbrace{1}_{\text{Mechanical effect}} \quad - \quad \underbrace{\frac{\bar{\tau}_0}{1-\bar{\tau}_0} \, \alpha \, \mathbb{E}_{\text{top}}[\beta]}_{\text{Behavioral effect}}, \tag{2.3}$$

where $\mathbb{E}_{\text{top}}[\cdot]$ weights each elasticity proportionally to the total top-bracket earnings of households with that elasticity. The welfare effect captures welfare losses that result—by the envelope theorem—from the

[7]Appendix 2.9.2.1 walks through the algebra, which is straightforward.

[8]We also assume that (a) the planner is considering a tax *increase*—to ensure no responses by earners outside of the top bracket—and (b) the tax schedule is convex—which ensures intensive margin responses.

increased tax burden at households' initial incomes. The mechanical effect captures revenue gains as taxes increase at households' initial incomes, whereas the behavioral effect captures negative fiscal externalities as households reduce their incomes in response to higher taxes.

We later show that in the data, the sum of the two revenue effects is always positive on net: raising taxes increases tax revenues, i.e. we are on the correct side of the "Laffer curve." Importantly, this implies that there exists a welfare weight $\widetilde{\lambda} > 0$ that rationalizes the planner's first-order condition. To verify that we are not a local minimum, or—as we will say throughout—in a "Laffer valley," we check the second-order condition:

$$0 \quad \geq \quad \frac{d^2}{d\bar{\tau}^2}\bigg|_{\bar{\tau}=\bar{\tau}_0} \mathcal{L}(\bar{\tau}) \quad \geq \quad \frac{d^2}{d\bar{\tau}^2}\bigg|_{\bar{\tau}=\bar{\tau}_0} \mathrm{Rev}_{\mathrm{top}}(\bar{\tau}) \quad \propto \quad -(2-\bar{\tau}_0)\,\mathbb{E}_{\mathrm{top}}[\beta] \;+\; \bar{\tau}_0\;\alpha\;\mathbb{E}_{\mathrm{top}}[\beta^2] \tag{2.4}$$

Interestingly, this second-order test is sensitive to not only the mean elasticity of top earners, but also its variance $\mathbb{V}\mathrm{ar}_{\mathrm{top}}[\beta] = \mathbb{E}_{\mathrm{top}}[\beta^2] - \mathbb{E}_{\mathrm{top}}[\beta]^2$. And *we are in a "Laffer valley" (i.e. (2.4) fails) if—fixing the mean elasticity—elasticities vary enough.* Moreover, this conclusion is normatively neutral, as the test does not depend on the welfare weight $\widetilde{\lambda}$.[9]

What economic mechanism underlies the planner ability to improve taxes when faced with heterogeneous elasticities? This is best illustrated by conceptualizing the second-order effects of a tax increase as the difference between the effects of two successive, infinitesimal tax increases. How does the first tax change shape the effects of the second? Here, the key idea is that *the behavioral effects of the first tax change determine the income level at which each household experiences the second tax change*, and therefore determine the marginal tax change it faces. For low-elasticity households—who by definition do not respond much to the first tax change—this has little bite. However, many high elasticity households exit the top bracket in response to the first tax change, and so are not affected by the second one. Due to this differential sorting, *the first tax change decreases the effective elasticity of top earners; the planner's second increase in taxes exploits this.*

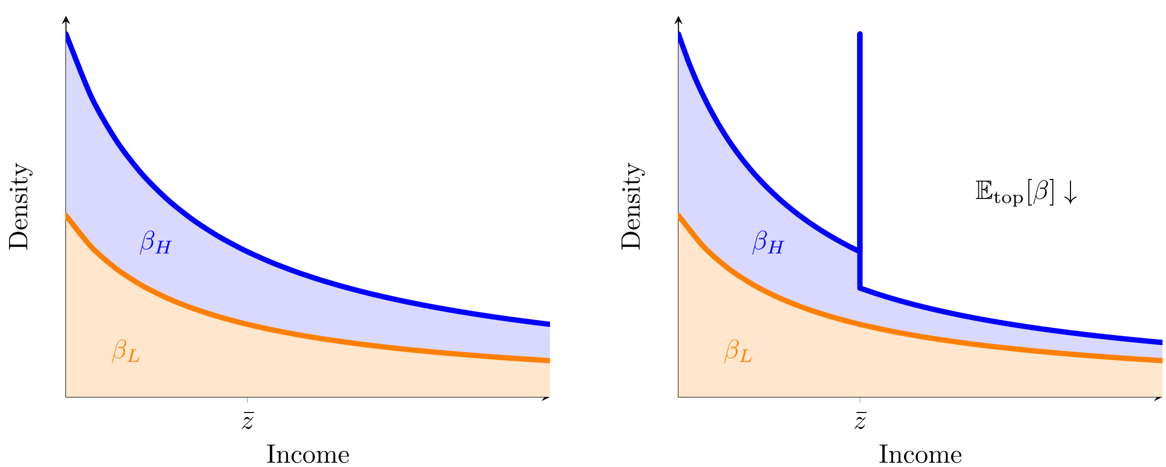


Figure 2.1: Income density of high- (blue) and low-elasticity (orange) households before (left panel) and after (right panel) taxes are increased. The compositional shift induced by an increase in taxes reduces the aggregate elasticity of top earners. (The vertical line in the right panel is bunching at a kink.)

[9]The second inequality of (2.4) requires that second-order welfare effects are weakly positive. For the purpose of this example, this relies on our assumption that the planner's desire to make transfers to each top-earner is constant, not changing as their consumption or income. However, our later results show that this assumption is unnecessary when the planner uses more finely-tuned tax variations.

Is this theoretical consideration of any practical importance? Consider a back-of-the-envelope calibration: Suppose top marginal tax rates $\bar{\tau}_0$ are 50%, the Pareto shape parameter α is 2.5, and the average elasticity of the top earners is 0.3, consistent with the literature. Then the first-order condition holds if the planner places a weight $\widetilde{\lambda} = 0.25$ on transfers to top earners. The second order test (2.4) holds—so that taxes locally maximize rather than minimize revenue—if and only if $\mathbb{V}\mathrm{ar}_{\mathrm{top}}[\beta] \leq 0.27$. This is roughly the variance if, for example, three quarters of households have elasticity zero and one quarter have elasticity ≈ 1.2. This value for the variance is only slightly above a conservative lower bound we estimate in Section 2.5 based on differences in elasticities between low- and high-itemizers. This suggests that the planner's second-order condition may well be violated in practice.

The rest of the paper develops the ideas illustrated above in order to (a) accommodate general planner and household preferences, (b) provide an efficiency test that applies to the whole tax schedule and is both necessary and sufficient, and (c) formulate this test in terms of sufficient statistics that we then evaluate in the data. The same key ideas—elasticity variance in the planner's second-order condition and the "sort and extort" motive—continue to play central roles in a general environment with non-linear income taxation.

2.3 Model

We study a standard, static Mirrlees model of income taxation, but allow for arbitrary preference heterogeneity. After laying out the model, we describe a number of mild regularity conditions used for our main results.

2.3.1 Environment

Time is static. There is a single consumption good and a single labor factor. Production is linear, and we normalize the price and the wage to one.

A unit measure μ of households $h \in \mathcal{H}$ supply labor z and consume subject to a non-linear income retention schedule to maximize utility u^h. For any income retention schedule $\widetilde{R} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, we define

$$z^h(\widetilde{R}) \equiv \arg\max_{z \geq 0} u^h(\widetilde{R}(z), z) \tag{2.5}$$

if the arg max exists, and let $V^h(\widetilde{R})$ denote the associated max.

We take as a primitive a particular income retention schedule

$$R : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \tag{2.6}$$

that maps each household's pre-tax income to its post-tax income, i.e. consumption. We denote labor supply, consumption, and utility at R by z_0^h, c_0^h, and V_0^h, respectively. Throughout, we will use "income retention schedule" and "tax schedule" interchangeably.

Throughout the paper, we study tax changes $\Delta : \mathbb{R}_{\geq 0} \to \mathbb{R}$ relative to R. Our main results focus in particular on very "small"—i.e. local—tax changes. Put simply, we call a tax change small if its level and derivatives are uniformly bounded by those of R, with a small bound. More formally, we define a norm on the space of tax changes[10] by

[10] Our regularity assumptions ensure the derivatives used in this definition exist. In Appendix 2.9.1.1 we define a Banach space of tax deviations on which $||\cdot||$ is defined and is a norm in the formal sense.

$$||\Delta||= \sup\left\{B\in\mathbb{R}_{>0}\;\middle|\; \forall z\in\mathbb{R}_{>0},\; |\Delta(z)|\leq B|R(z)| \text{ and } |\Delta'(z)|, \left|\frac{d\Delta'(z)}{d\log z}\right|, \left|\frac{d^2\Delta'(z)}{d\log z^2}\right| \leq B|R'(z)|\right\}. \tag{2.7}$$

We denote by $\mathbf{0}$ the zero tax change and by $B_\delta(\mathbf{0})$ the open ball of radius δ around $\mathbf{0}$.

2.3.2 Class of social planners

Throughout the paper, we study the problem of hypothetical social planners who assess welfare according to "generalized utilitarian" criteria and face tax revenue constraints.

Definition 1. *We call a level of government expenditure G and a collection of welfare-weighting functions $(w^h)_{h\in\mathcal{H}}$ with $w^h : \text{Im}(u^h) \to \mathbb{R}\cup\{-\infty\}$ a* ***standard social objective*** *if there exists $\delta > 0$ such that, for all $h \in \mathcal{H}$, w^h satisfies the following conditions on the domain $V^h(R + B_\delta(\mathbf{0}))$:*

- *w^h is finite,*
- *w^h is twice-continuously differentiable,*
- *w^h is weakly increasing, and strictly increasing for a positive measure of $h \in \mathcal{H}$;*

and $w^h \circ V^h(\widetilde{R})$ and its first two Frechet derivatives in $\widetilde{R}$ are h-measurable within $R + B_\delta(\mathbf{0})$.[11]

Our main results rely on the notion that a social objective may "rationalize" the tax schedule R by providing a metric according to which it is optimal among all other (nearby) tax schedules.

Definition 2. *A standard social objective $((w^h)_{h\in\mathcal{H}}, G)$* ***locally rationalizes*** *R if there exists $\delta > 0$ such that*

$$R \quad\in\quad \underset{\widetilde{R}\in R+B_\delta(\mathbf{0})}{\arg\max} \int w^h \circ V^h(\widetilde{R}) \qquad s.t. \qquad \int \left[z^h(\widetilde{R}) - \widetilde{R}(z^h(\widetilde{R}))\right] d\mu \geq G. \tag{2.8}$$

The regularity conditions we impose for our formal results ensure that aggregate welfare and tax revenue always exist locally to R, so (2.8) is well-posed.

By working within the class of standard social objectives, we restrict our focus in a few important ways. First, the importance of any set of households to welfare is proportional to their measure. Second, welfare is additively separable across households; this rules out certain explicit forms of fairness concerns. Third, welfare is increasing in each household's utility. Finally, welfare is sufficiently smooth in each household's utility.

The class of standard social objectives nests most, but not all planners considered in the income taxation literature. One notable exception is a "Pareto planner" who prioritizes getting each household's utility above a (household-specific) "target utility." A Pareto planner's marginal desire to allocate additional utility to any particular household is discontinuous at its target utility for that household, whereas a planner with a standard social objective values this additional utility in a continuous way. Another exception is a "Rawlsian planner" who maximizes the utility of the worst-off household. Although such a planner lacks a standard social objective due to non-separability, one may think of the Rawlsian motive as a limit case of standard social objectives with increasingly concave weighting functions w^h.

[11] These derivatives exist by w^h's differentiability, Assumption 2, and (for small enough δ) Lemma 2; see Appendix 2.9.1.

2.3.3 Regularity conditions

Whereas many results in the optimal income taxation literature rely on specific function form assumptions, our theoretical results instead hold under much weaker regularity conditions. For readability, we summarize the main content of these assumptions below and state them formally in Appendix 2.9.1.

Our regularity conditions serve four main purposes. First they ensure that welfare and tax revenue are well-defined and well-behaved (i.e. twice-continuously differentiable) in a neighborhood around the initial tax schedule R. The main conditions we impose to this end are that the initial tax schedule R is sufficiently smooth and has bounded elasticities, that—in a similar spirit—household utility is sufficiently smooth and generates income and compensated elasticities of labor supply that are bounded across households.

Second, our assumptions ensure that welfare and tax revenue can be expressed as integrals over household income z, facilitating integration by parts. To this end we assume the existence of sufficiently smooth functions for the income density $g(z)$ and income-conditional elasticity moments (e.g. the mean compensated elasticity of labor supply conditional on income z).[12]

Third—and of somewhat more qualitative interest—we assume that, local to the initial tax schedule R, all labor is supplied on the intensive margin. In doing so, we abstract away from interesting extensive margin decisions, such as labor force participation and migration. We view the extensive margin of labor supply as an interesting area for future work.

Fourth, we adopt a slightly stronger notion of a standard social objective, which we call a **standard, regular social objective**. This ensures the existence of income-conditional moments for several welfare statistics—such as the planner's desire to transfer income to households at income z—and imposes that these moments are sufficiently smooth in income. Our results use the assumption that average income-condition welfare weights are continuous in income in order to argue that a small decrease in the welfare of households earning z can be fully compensated (from the perspective of aggregate welfare) by a proportional increase in the welfare of $z + dz$ earners.

2.4 Rationalizability test for tax schedules

In this section, we present our main theoretical results. We provide a set of simple conditions on locally observable sufficient statistics—i.e. a "test"—that holds if and only if the tax schedule is locally rationalizable. Our test augments the first-order test of Werning (2007) with a novel, second-order condition.

After showing these conditions are necessary for rationalizability, we offer two interpretations: First, we illustrate variational arguments about the first- and second-order changes in welfare and tax revenue that tax change. Second, we explain the economic mechanisms at work, drawing particular attention to a novel "sort and extort" motive for taxation with multi-dimensional heterogeneity. Finally, we show that the conditions are also locally sufficient for rationalizability.

2.4.1 Necessary conditions

In order to state our first result, we will now introduce notation used throughout the paper. For all $z \in \operatorname{supp} g$, we denote by $\eta(z)$, $\varepsilon(z)$, and $\alpha(z)$ the average income and compensated elasticities[13] and the

[12]While the necessity of our rationalizability test only relies on weak assumptions of this form, the sufficiency of our test additionally requires some mild conditions on the relationship between taxes, incomes, and conditional elasticity moments in the limits of zero and infinite incomes.

[13]We define these elasticities more formally in Appendix 2.9.1.2.

local shape parameter of the income distribution at income z.[14]

$$\eta(z) \equiv \mathbb{E}[\eta^h(R)|z_0^h = z] \qquad \varepsilon(z) \equiv \mathbb{E}[\varepsilon^h(R)|z_0^h = z], \qquad \alpha(z) \equiv -\frac{d\log(zg(z))}{d\log z} \tag{2.9}$$

Income elasticities capture behavioral responses to changes in the *level* of income retention, whereas compensated elasticities capture behavioral responses to changes in the *slope* of income retention. In addition to these familiar statistics, our results also depend on two less well-studied statistics. These are the income-conditional average compensated elasticity *squared* and an income-conditional average compensated *super*-elasticity $\varepsilon^{+h}(R)$ that captures how households' compensated elasticity change due to their labor supply responses to marginal tax changes.[15] We use the notation

$$\varepsilon^2(z) \equiv \mathbb{E}[\varepsilon^h(R)^2|z_0^h = z], \qquad \varepsilon^+(z) \equiv \mathbb{E}[\varepsilon^{+h}(R)|z_0^h = z]. \tag{2.10}$$

Our first main result provides necessary conditions that constrain the relationships between these parameters and taxes at any locally rationalizable tax schedule. This test consists of a first order condition, which is standard in the literature, and a complementary second order condition, which is novel.

Theorem 1. *Suppose R is locally rationalized by some standard, regular social objective. Then for all $z \in \operatorname{supp} g$,*

$$0 \;\geq\; -1 \;+\; (1-R'(z))z\left[\frac{\eta(z)}{R(z)} \;+\; \left(\alpha(z) - \frac{d\log}{d\log z}\left(\frac{1-R'(z)}{R'(z)}\right) - \frac{d\log\varepsilon(z)}{d\log z}\right)\frac{\varepsilon(z)}{R'(z)z}\right]. \tag{ABC}$$

Moreover if (ABC) *does not everywhere hold with equality,*[16] *then for all $z \in \operatorname{supp} g$,*

$$0 \;\geq\; -\left(1+R'(z)\right)z\varepsilon(z) \;+\; \left(1-R'(z)\right)z\left[\alpha(z)\varepsilon^2(z) \;-\; \frac{d\varepsilon^2(z)}{d\log z} \;+\; \varepsilon^+(z)\right]. \tag{DEFG}$$

Proof. See Appendix 2.9.3.1. □

We label the first-order condition "ABC" in analogy to Diamond (1998)'s well-known "ABC formula", because they both derive from the planner's first-order condition. More precisely, (ABC) can be understood as a differentiated version of Diamond's formula. Analogously, we dub our second-order test "DEFG."

The first-order condition (ABC) and the second-order condition (DEFG) reflect two very different senses in which R is robust to changes in taxes. First, (ABC) says that the planner cannot lower taxes locally to z without reducing tax revenue, i.e. we are on the correct side of the Laffer curve. If this condition fails then—since it is possible to lower taxes at the same time as raising tax revenue—taxes are *Pareto inefficient* (Werning, 2007). By contrast, a failure of the second-order test does not imply Pareto inefficiency but rather a slightly weaker condition: *non-rationalizability* within the class of planners we consider. The distinction is that, whereas the planner responds to a failure of (ABC) by only lowering taxes, the planner responds to a failure of (DEFG) at z with *a small decrease in taxes just above z and an equal and opposite increase in taxes just below z*.[17] Although this makes some households just below z worse off, we show that their

[14]This is related to the well-known concept of the income distribution's "Pareto tail." The local shape parameter is a local version of this concept that coincides with the Pareto shape parameter if g is the density of a Pareto distribution.

[15]See Appendix 2.9.1.4.1 for a more formal definition.

[16](ABC) does holds everywhere with equality if and only if R is a stationary point for tax revenue; this is generically false.

[17]The reason we use this particular deviation is that a planner's second-order condition is only informative in directions in which their Lagrangian is constant to first order (other directions are either infeasible or welfare-reducing). Our assumptions guarantee that the Lagrangian is constant to first order under this deviation because they guarantee that the marginal effects of tax changes around some income level—both on household behavior and on welfare—are continuous both in (a) what that

decrease in welfare can be made arbitrarily small relative to the planner's revenue gains. This willingness to make welfare trade-offs—so long as the downside is small enough—is a key feature that distinguishes our notion of rationalizability from Pareto efficiency.

Another important contrast between the standard first-order condition and our new second-order condition is that the former only depends on *average* elasticities within each income level, whereas the latter depends on a higher moment of the distribution of elasticities. Critically, in order to evaluate (DEFG), one must—since $\varepsilon^2(z) = \varepsilon(z)^2 + \mathbb{V}\text{ar}[\varepsilon^h(R)|z_0^h = z]$— consider the *variance* of compensated elasticities within each tax bracket. For this reason, the inclusion of multi-dimensional heterogeneity can have profound effects on the planner's problem. Namely, when the income density is falling quickly enough ($\alpha(z) > 0$)—as at high income levels—the planner's second-order condition fails whenever the variance in elasticities is sufficiently large. In this case, it is not possible to rationalize the tax schedule R within the broad class of social objectives we consider, even though the first-order (ABC) test known to the existing literature may pass.

2.4.2 Variational interpretation

In this section, we explain the first- and second-order conditions of Theorem 1 by illustrating the variations in taxes that they reflect. Intuitively, if R is rationalizable—i.e. optimal for some social planner—then, for any small variation in taxes, that planner's Lagrangian must be constant to first order in that variation and concave to second order.

The necessary conditions of Theorem 1 reflect these first- and second- order requirements for a *particular* type of tax change: a smooth increase in retention, narrowly concentrated between $z - dz$ and $z + dz$.[18] By "narrow," we mean that the amount this variation changes retention at its peak is much greater than the range of incomes over which it changes retention. Crucially, a typical household affected by a narrow tax change experiences a much greater change in marginal taxes than in the level of taxes. This simplifies our analysis by allowing us to ignore (a) curvature in the planner's desire to transfer income to any given household and (b) income effects, which become small compared to compensated effects.[19] Although our necessary conditions focus on narrow tax variations, we later show all other tax changes are undesirable for some planner, so that these necessary conditions are also sufficient for rationalizability.

To begin the analysis, consider the first-order effects of a tax change concentrated at some income z, as in Figure 2.2. These effects can be decomposed, à la Saez (2001), into welfare effects, mechanical revenue effects, and behavioral revenue effects. Welfare effects simply reflect that—since we have weakly raised retention everywhere—households are better off as a result of the tax change. Mechanical revenue effects capture the losses in tax revenue that occur because, ignoring any change in household labor supply behavior, taxes are lower at each income level. Finally, behavioral revenue effects reflect that tax-change-induced changes in household behavior—through both income and compensate effects—have fiscal externalities on the planner whenever marginal taxes are non-zero.

Since the welfare effects are positive, revenue effects must be negative at any rationalizable schedule (otherwise the tax schedule can be improved). The RHS of (ABC) captures these revenue effects: The -1

income level is and (b) the sign of the tax changes.

[18]As discussed below Theorem 1, our second-order condition couples this narrow increase in retention with a nearby, narrow decrease in retention, which ensures that the planner's Lagrangian is constant to first order. Below, we keep things simple by discussing the second-order effects of only the increase in retention. This is without loss because although the decrease in retention has *opposite first-order effects*, it has *identical second-order effects*.

[19]Formally, we do not "ignore" these effects. Rather, our proofs explicitly model them but show that, for sufficiently "narrow" tax variations, they are dominated by other considerations.

corresponds to mechanical revenue losses and the remaining terms correspond to behavioral effects, which can increase or decrease revenue.

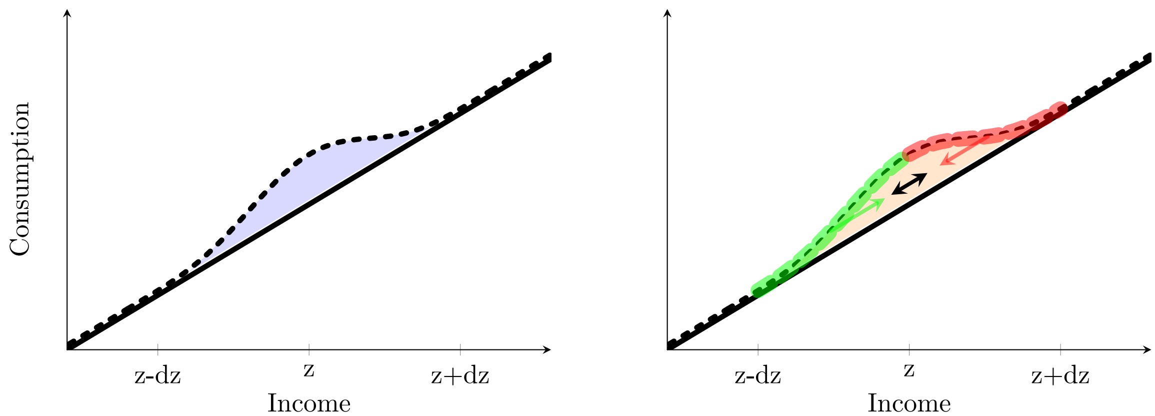


Figure 2.2: Left panel: An initial income retention schedule, before (solid) and after (dashed) a small variation in taxes. The shaded blue area represents the size of the mechanical effect and (if appropriately weighted) the welfare effect. Right panel: Behavioral effects due to a the same variation in taxes. Behavioral income effects (black arrow) act on households whose level of income is affected (shaded orange area), and can have positive or negative sign. Behavioral compensated effects increase the incomes of households whose marginal taxes decrease (green), and vice-versa (red).

We now consider the *second-order effects* of the same variation in taxes we have already studied to first order. The second derivative of the planner's Lagrangian captures how its first derivative *changes* as the same variation is done more or less. In other words, we ask, "how do the effects of an infinitesimal variation in taxes differ when we do it for a second time?" If R is optimal for some planner, then these second-order effects must be negative on net, i.e. the Lagrangian must be locally concave. In order to compute the Lagrangian's local curvature, we now walk through how each component of its first derivative with respect to the narrow variation studied above—i.e. the welfare, mechanical, and welfare effects—change when we repeat that variation for a second time.

To begin, consider how the welfare effect changes when our tax variation is repeated for a second time. For a narrow tax variation, the only second-order effect on welfare is as follows: When taxes change for a second time, the amount of post-tax income transferred to each household may differ the amount transferred during the first tax change, since—due to its behavioral response to the first tax change—that household's income is different than it was during the first tax change. Since these first-order behavioral effects shift all households with incomes in $[z-dz, z+dz]$ toward z, where the change in retention is higher, the second tax variation transfers more post-tax income to each household than did the first; see Figure 2.3. For this reason, the second-order effect on welfare is positive. Rather than quantifying the precise size of this convexity, our second-order test captures its sign with an inequality. This expresses that—because the change in welfare effects is always positive—the change in (mechanical and behavioral) revenue effects must be negative in order for the planner's Lagrangian to be concave on the whole. The test therefore simply states that tax revenues must be concave with respect to narrow tax variations.

We now turn to these revenue effects. We start with our tax variation's mechanical effects, asking how they change when the variation is repeated for a second time. Conveniently, this change is simply the flip side of the second-order welfare effect considered above: When taxes change for a second time, each household receives a greater transfer than it did the first time, since—due to its behavioral response to the first tax change—that household's income has moved toward z, where the variation increases retention the most. This effect contributes a term of $-R'(z)z\varepsilon(z)$ to (DEFG).

Just as the second tax variation's mechanical effect changes due to households' behavioral response to the first variation, so too does the second tax variation's behavioral effect change due to the mechanical effects of the first variation. Intuitively, when taxes change for a second time, each household's behavioral effect on revenue is smaller, since it moves to a new income at which—due to the first variation—retention is higher, i.e. taxes are lower. This effect contributes a second term of $-R'(z)z\varepsilon(z)$ to (DEFG), a manifestation of Hessian symmetry

The tax variation's second-order behavioral revenue effects are also shaped by its first-order effects on household behavior. These "behavioral-affecting-behavioral" effects can be divided into three groups: (a) one effect proportional to elasticities, (b) one key effect proportional to elasticity-squared, and (c) two small "correction" terms. The elasticity-proportional effect captures the following mechanism, which leads to less revenue-advantageous changes in behavior: Where the first variation has increased marginal retention, households respond less to the second tax change since the change in log marginal retention is smaller,[20] and vice-versa. This dampens positive labor supply responses and amplifies negative labor supply responses, contributing a factor of $-(1-R'(z))z\varepsilon(z)$.

We now turn our attention to a key effect of interest, one proportional to the average square of elasticities at z. The dependence of this effect on a higher moment of the elasticity distribution means that it—unlike the previous effects discussed—can have a qualitatively different effect when elasticities vary conditional on income. The effect is as follows: Because households adjust their incomes in response to the first tax change, they experience the second tax change at a different income level, where it may be larger or smaller. Moreover, how many households relocate to each different income level—and so how many households experience greater or smaller increases in taxes during the second tax change—depends on how many households start at each income level, i.e. on the local shape of the income distribution. For example, suppose the income density is sharply decreasing locally to z. In this case, the amount of income earned by households who—due to their behavioral responses to the first tax change—face larger marginal tax decreases or smaller marginal tax increases during the second tax change than the first is greater than the income earned by households who face smaller tax decreases or larger tax increases during the second tax change. This is because the former set of households are those inhabiting $[z-dz, z-\frac{dz}{2}]$ and $[z, z+\frac{dz}{2}]$, which contains more mass than the complementary range of incomes when the income density is sharply decreasing; see, e.g. Figure 2.3. Since the larger first group responds with more advantageous behavioral effects to the second tax change than the first—whereas the second group does the opposite but is smaller—the second tax change has a more positive behavioral effect on revenue than the first. If, on the other hand, the income density is increasing around z, the effect flips. In total, this effect accounts for the term $(1-R'(z))z\alpha(z)\varepsilon^2(z)$ in (DEFG). $\alpha(z)$ reflects the shape of the income density locally; $\varepsilon^2(z)$ reflects that households' elasticities determine not only how much they relocate different income levels (and so different tax changes) but also how strongly they respond to those changes.

Finally, two additional "corrections" to these second-order behavioral effects account for the final two terms of (DEFG). We emphasize these terms less because—to the extent we can estimate them—they are quantitatively unimportant, and because they are zero in many parameterizations used in the literature. The first adjusts the effect described in the previous paragraph to account for the possibility that the distribution of elasticities conditional on income may vary with income locally to z. The second captures the fact that, due to changes in the curvature of household preferences and/or the tax schedule, households may have different elasticities during the second tax change than during the first. To the extent we can estimate these

[20]This is true to the extent that household elasticities remain constant; see later discussion.

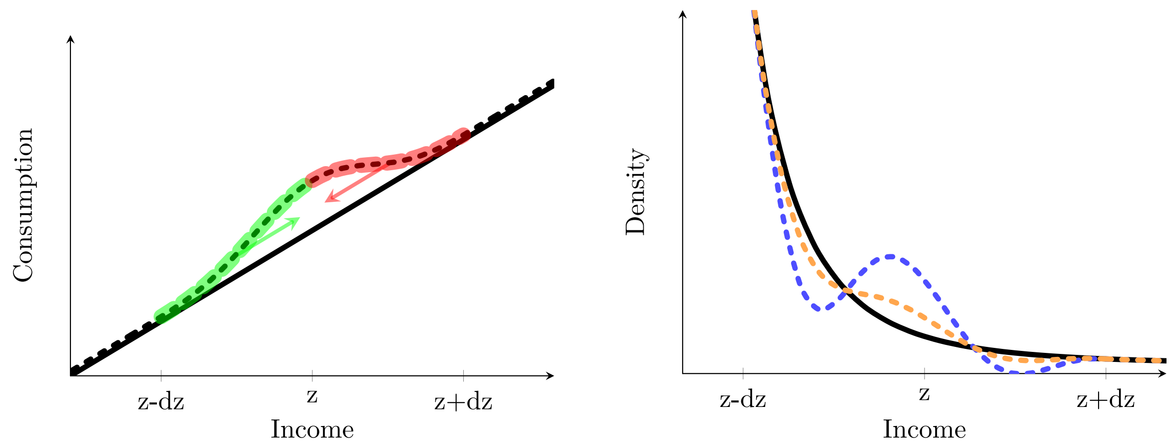


Figure 2.3: Left panel: Changes in behavior through compensated effects following the first tax change. Right panel: Initial income density (black) and income density following the first tax change for high- (blue) and low-elasticity (orange) households. Each household adjusts its income toward z, but the change in the income density depends on the shape of the initial income density and the distribution of elasticities.

effects, they appear to be small (see Section 2.5.3.2). However, we do not estimate the component of these "super-elasticities" stemming from changes in curvature of household preferences; this is an interesting area for future work.

2.4.3 Exploring the economic mechanism

While the analysis of Section (2.4.2) explains the various terms of our second-order test (DEFG) mathematically, we would also like to develop a more intuitive and economic understanding of it. Indeed, whether the test can fail in any reasonable description of the economy is an important question, considering that Werning (2007) shows the first-order condition (ABC) is both necessary and sufficient for global Pareto efficiency in a certain one-dimensional setting.

In this section, we show under what conditions (DEFG) fails and, when it does, provide an intuitive explanation of how the planner can improve the tax schedule. Our main focus is on the way that the test fails when—due to multi-dimensional household heterogeneity—elasticities vary widely within income levels. Here, we emphasize a novel "sort and extort" procedure through which the planner can increase tax revenue by using a first tax change to shift the elasticity-composition of the income distribution and a second tax change to exploit the separation of high- and low-elasticity households with differential taxes. Finally, we also explain how the test can fail even when households differ along single dimension of heterogeneity (i.e. productivity).

2.4.3.1 Multi-dimensional heterogeneity: "sort and extort"

The most natural and empirically relevant case in which (DEFG) fails is when there is significant variation in ETIs within income levels. Since one dimension of heterogeneity in household preferences is required to rationalize income differences, within-income differences are only possible with multi-dimensional heterogeneity. These within-income differences in household ETI point to an important limitation faced by the social planner. In general, a social planner might like to tax different households with the same income differently, either because she prefers some households to others or because she anticipates that they may respond to taxes differently. For example, a revenue-maximizing planner levies high taxes on low-ETI households and low taxes on high-ETI households, all else equal. However, when she cannot condition taxes explicitly on these elasticities, the planner is forced set a single tax schedule that—at each income level z—must balance her desire to tax different z-earners differently. In this sense, the planner's ability to *target* taxes is limited by the fact that her tax instrument is lower-dimensional than the space of households she taxes with it.

This constraint—that the planner must set a single tax schedule for households who vary within income levels—is at the core of failures of the second-order condition. To see why, consider the simple case where households can be partitioned into a finite set of groups $i \in I$, each satisfying our regularity conditions in isolation. We show in Appendix 2.9.3.3 that in this case the *aggregate* first- and second-order tests are simply the averages of the *within-group* first- and second-order tests. That is, if we let $\Pi^i_{ABC}(z)$ and $\Pi^i_{DEFG}(z)$ represent the right-hand sides of (ABC) and (DEFG) but when expectations are taken only over households within each group i, then the our test can be expressed, for each $z \in \operatorname{supp} g$ as:

$$\mathbb{E}\left[\Pi^{i(h)}_{ABC}(z)\middle| z^h_0 = z\right] \;\leq\; 0 \qquad \text{and} \qquad \mathbb{E}\left[\Pi^{i(h)}_{DEFG}(z)\middle| z^h_0 = z\right] \;\leq\; 0. \tag{2.11}$$

Moreover, if preferences within each group i are concave and additively separable over consumption and labor and vary only in a labor dis-utility shifter, then one may show that[21]

$$\Pi^i_{ABC}(z) \;\leq\; 0 \qquad \Longrightarrow \qquad \Pi^i_{DEFG}(z) \;\leq\; 0. \tag{2.12}$$

Together, these two observations imply that the second-order test can only fail if the first-order test fails for at least one group. In other words, all failures of (DEFG) can be attributed to the fact that for at least one group, R is on the wrong side of the Laffer curve.

In a world with group-specific taxes, the planner could address the fact that one group is on the wrong side of the Laffer curve by simply lowering its taxes. However, this option is not available to the planner we consider: Any desirable decrease in taxes for one group requires a (potentially) undesirable decrease in taxes for another group. Indeed, if group-specific failures of the first-order condition are weak enough, it is possible for the second-order condition to hold still and for taxes to be rationalizable. In this context, the key insight of Theorem 1 is as follows: *If group-specific failures of first-order conditions are so severe that the aggregate second-order condition fails, then the planner can improve taxes with a special, "as-if" group-specific tax change.*

The rest of this section explains how the tax variation underlying (DEFG)—the one explored more mechanically in Section 2.4.2—approximates group-specific taxation. A first step is to recall that since this tax variation is arbitrarily "narrow", it has arbitrarily small welfare effects and arbitrarily small revenue effects through income elasticities. This implies that all of its impacts operate through compensated elasticity effects.[22] In particular, the sense in which our tax variation approximates group-specific taxation is by effectively lowering taxes more for high-elasticity households and raising taxes more for low-elasticity households.

Our tax variation accomplishes this "as-if" elasticity-dependent taxation with a two-step change in taxes, captured by the first and second derivatives of the planner's Lagrangian. To first-order, changing taxes causes households to change their labor supply behavior as depicted in Figure 2.3. Crucially, low-elasticity

[21] See Appendix 2.9.3.3. This fact is consistent with Werning (2007)'s observation that the dual of the planner's problem is convex when there is one such group i.

[22] These compensated elasticity effects are—as illustrated for the first order variation at the beginning of Section 2.4.2—moderated by the local shape parameter $\alpha(z)$. Although our test accounts for cross-group differences in this shape, we keep this discussion simple by assuming that the shape parameter is constant across groups i.

households end up with roughly the same income density after an infinitesimal change in taxes, whereas high-elasticity households end up with a very different density, one which depends both on their initial density and the shape of the tax change itself. Most importantly, we show that when the initial density is declining steeply, high-elasticity households are disproportionately drawn toward incomes in two regions—one where the variation decreases taxes by more and one where it increases taxes by less. That is, the variation's first-order effect is to "sort" households differentially, by elasticity. The second derivative of the planner's Lagrangian asks how the first-order effects change when the variation is performed a second time.[23] Critically, these effects improve as the variation is performed a second time, because they now operate on households who have been sorted in just the right way to generate positive behavioral effects on revenue: High elasticity-households are disproportionately represented where taxes decrease and low-elasticity households are disproportionately represented where taxes increase. That is, the planner exploits the sorting of households by "extorting" only the least-responsive households with high taxes, *as if* she could tax them differentially.

Quantitatively, this sort-and-extort mechanism accounts for a factor of $(1-R'(z))z\alpha(z)\,\mathbb{V}\text{ar}[\varepsilon^h(R)^2|z^h(R)=z]$ in our second-order test, (DEFG). We show in Appendix 2.9.3.3 that in the simplest natural case—when household preferences are linear in consumption and CES in labor supply, and when marginal taxes and the mean and variance of income-conditional elasticities are constant local to z—this term is the only factor that can cause (DEFG) to fail when (ABC) holds. That is, if the first-order test holds at some income z where marginal taxes and the income shape parameter $\alpha(z)$ are positive, then the second-order test fails there if and only if—holding average elasticities at z constant—there is enough heterogeneity in these elasticities. In other words, the second-order test binds if and only if there is enough scope for a sort-and-extort tax reform.

2.4.3.2 One-dimensional heterogeneity: "shift and exploit"

Although the most economically interesting and empirically plausible failures of (DEFG) derive from within-income household heterogeneity, failures are also possible in models where households are homogeneous conditional on income. The main idea behind this possibility is similar to the "sort and extort" mechanism explored in the previous section: The planner can change elasticities at each income level with a first tax variation and then exploit this change by repeating the variation. The difference relative to the heterogeneity case is that elasticity changes come not through sorting but rather through "shifts" – either in the elasticity of individual households or in the identity of which households (each having fixed elasticities) are represented at each income level.

Notably, it is impossible for a planner to "shift and exploit" households when utility satisfies the functional form studied by Werning (2007).[24] This functional form implies that all variation in compensated elasticities across income levels is due to differences in income and consumption levels at which households' preferences are evaluated, rather than differences in the curvature of their preferences at a given level of income and consumption. It therefore prevents the planner from using a small tax variation to shift local elasticities, since—for any income level z—the households who each z after a tax change have the same post-tax change elasticity as the household who previous earned z had before the tax change. In Appendix 2.9.2.2, we provide explicit examples showing how—outside of this class of preferences—(DEFG) can fail with one-dimensional preference heterogeneity.

[23] Because the variation is narrow, these effects can be simplified to compensated effects on revenue.

[24] This functional form nests additive CES preferences, which is perhaps the most commonly analyzed example in the literature.

2.4.4 Sufficient conditions

So far, we have provided a novel set of necessary conditions that must be satisfied by any locally rationalizable tax schedule. However, we derived the second-order condition of our test using only very "narrow" variations in taxes. So tax schedules which pass the (DEFG) test might still allow for other sorts of improvements. Our second main result says that this is not the case: The necessary conditions of Theorem 1 are also sufficient for local rationalizability.

This sufficiency result relies on a few mild assumptions beyond those required for Theorem 1. Aside from the additional regularity conditions in Assumption 6 and the condition—stated in the theorem itself—that some households at each income level have small enough elasticities, we also require that (ABC) and (DEFG) hold in a slightly stronger sense. Finally, our sufficiency result uses a slightly weaker notion of local optimality than the necessity result.[25]

We can now present Theorem 2, which shows that our characterization of locally rationalizable tax schedules in Theorem 1 is tight: If (ABC) or (DEFG) fail, then the schedule sub-optimal for all planners; if they pass then the schedule is optimal for some planner.

Theorem 2. *Suppose that*

- (ABC) *and* (DEFG) *hold by amounts* $\Pi_{ABC}(z)$ *and* $\Pi_{DEFG}(z)$ *that satisfy:*
  - *For all* $z \in \operatorname{supp} g$, $\Pi_{ABC}(z) > 0$.
  - *There exist* $\bar{b}_c, \bar{b}_z > 0$ *such that* $\Pi_{DEFG}(z) \geq \bar{b}_c R(z) + \bar{b}_z z$ *for all* $z \in \operatorname{supp} g$.
  - *There exists* M *such that* $z|\Pi'_{ABC}(z)| \leq M|\Pi_{ABC}(z)|$ *for all* $z \in \operatorname{supp} g$.
- *For some sufficiently small* $\epsilon > 0$, $\mathbb{P}[\varepsilon^h(R) \leq \epsilon | z_0^h = z] > 0$ *for all* $z \in \operatorname{supp} g$.[26]

Then there exists a standard, regular social objective that—for any $M > 0$*—locally rationalizes* R *within the sub-space of deviations*

$$\mathbf{\Delta}^*_M \equiv \left\{ \Delta \in \mathbf{\Delta} \;\middle|\; ||\Delta|| \leq \; M \left[\int g(z)\,(R(z)+z) \left(\left(\frac{\Delta(z)}{R(z)}\right)^2 + \left(\frac{\Delta'(z)}{R'(z)}\right)^2 \right) dz \right]^{\frac{1}{2}} \right\}. \tag{2.13}$$

Proof. See Appendix 2.9.3.2. □

Considering that (DEFG) reflects only the robustness of R to very "narrow" deviations in taxes, this is a surprising result. Indeed, given a fixed social objective, we know that the planner must consider the robustness of R to all possible deviations. Why do we not need additional conditions in order to rule out these other second-order deviations?

The main insight is as follows: *For an arbitrary planner*, one must indeed verify additional second-order conditions corresponding to variations in taxes not spanned by "narrow" deviations. However, *there exists* a planner for whom all other deviations are undesirable. Intuitively, any non-narrow deviation has an appreciable effect on the level of income retention, instead of just its slope. While changes in the level

[25] Even without this weaker notion, there exists—provided that R satisfies the test—a welfare function that, for any $\Delta \in \mathbf{\Delta}$, prefers R to $R + \epsilon\Delta$ for all $\epsilon > 0$ below some $\epsilon_\Delta > 0$. Our weaker notion of local optimality slightly restricts $\mathbf{\Delta}$ so as to ensure ϵ_Δ is uniform across all remaining tax changes Δ.

[26] More formally, we assume there exists a conditional expectation function $p_{\leq}(z;\epsilon) \equiv \mathbb{E}[\mathbb{1}_{\varepsilon^h(R)\leq\epsilon} | z_0^h = z]$ such that for all $z \in \operatorname{supp} g$, $p_{\leq}(z;\epsilon) > 0$. A conditional probability function exists because z_0^h is measurable by Assumption 3, because—since $\varepsilon^h(R)$ is measurable (see the second-to-last step in the proof of Lemma 2)—the indicator $\mathbb{1}_{\varepsilon^h(R)\leq\epsilon}$ is measurable, and because this indicator is also therefore integrable by dominated convergence.

of retention have several effects, one of them corresponds the curvature of the planner's re-distributional preferences: As she transfers more money to any one household, her desire to do so may increase or decrease. By taking the planner's preferences to be sufficiently risk-averse around each household's initial utility level, we can always make negative the second-order effects of tax variations that significantly vary the level of retention. Since (DEFG) ensures all variations that do *not* significantly vary the level of retention have negative second-order effects, this is sufficient to ensure local optimality.

While our focus on rationalizability facilitates this clean characterization, we also provide results relevant to the approach that fixes a welfare function and solves for optimal taxes. Specifically, Lemmas 6 and 7 in Appendix 2.9.3.1 provide the planner's full first- and second-order conditions.[27]

2.5 Empirical test of rationalizability

We now set out to empirically evaluate the (ABC) and (DEFG) tests. To do so, we estimate each of the sufficient statistics that comprise them in the NBER tax return datasets, which include both a panel of household tax returns and larger repeated cross sections within each year. After evaluating the test, we check the robustness of our estimates using direct evidence that tax changes can indeed "sort" households by elasticities.

To set the stage, recall our second-order (DEFG) test for "Laffer valleys." In the simplest case where (a) the distribution of elasticities is locally constant in z, (b) taxes are locally linear in z, and (c) labor supply preferences are additive CES, the test simplifies to

$$(1 + R'(z))\varepsilon(z) \leq (1 - R'(z))\alpha(z)\left(\varepsilon(z)^2 + \mathbb{V}\text{ar}[\varepsilon^h(R)|z_0^h = z]\right). \tag{2.14}$$

This expression serves two expository purposes. First, it clarifies what are our main statistics of interest: The shape of the tax schedule ($R'(z)$), the shape of the income distribution ($\alpha(z)$), income-conditional mean compensated elasticities of taxable income ($\varepsilon(z)$), and the income-conditional variance of these elasticities. Of these statistics, income-conditional variance is the most novel to our analysis and poses the greatest estimation challenges.

Second, (2.14) facilitates simple, back-of-the-envelope calculations. In the US, top marginal taxes are roughly $R'(z) \approx 0.5$, the income distribution features a Pareto tail with shape $\alpha(z) \approx 2.5$, and existing elasticity estimates—while they vary widely—are in the vicinity of $\varepsilon(z) \approx 0.3$ (Gruber and Saez, 2002; Saez et al., 2012). So, the second-order condition fails if and only if the variance of elasticities of top earners is ≥ 0.27. Indeed, this number is close to the lower bound on variance that is implied by the difference of elasticity between itemizers and non-itemizers estimated by Gruber and Saez (2002): these groups are roughly evenly sized and their mean elasticities differ by about one, implying the variance of elasticities across only the two groups is about one quarter. We will estimate that within-income variances that are significantly higher, implying that taxes are in a Laffer valley.

2.5.1 Data

We use the NBER panel and repeated cross-section of tax returns from 1979 to 1990. This sample period includes major tax reforms such as the Economic Recovery Tax Act 1981, which decreased marginal rates in

[27] While we provide these conditions in full, they (a) require the estimation of many statistics in order to be evaluated and (b) do not lend themselves to simple verification of the second-order condition, even if all of the required sufficient statistics were known.

3 years from 1982 to 1984; the Tax Reform Act 1986, which decreased the number of brackets and reduced the top marginal rate to 28%; the 1987 EITC expansion; and some state level tax reforms.

The data include limited demographic information, individual-specific federal and state income tax schedules,[28] and various measures of income. To construct a consistent measure of taxable income for the whole period, we closely follow Gruber and Saez (2002). The measure we use includes wage, business and capital income, and subtracts exemptions, standard and itemized deductions. Within the panel data, we—following the procedure that Gruber and Saez (2002) use to compute medium run elasticities—compute income changes over 3-year windows. We compute marginal tax changes (at initial income) over the same windows, and we drop individual-years with changes in marital status, initial (pre-deductions) income lower than $10 thousand dollars (in 1990 terms).

The panel sample consists of a random selection of four digit endings of social security numbers. There is purely random attrition; in some years a random subset of the sample social security numbers are excluded. The repeated cross-section sample is larger but not a panel; it over-represents higher-income individuals, which we account for with sample weights. Table 2.1 provides summary statistics.

Table 2.1: Summary statistics for panel data and repeated cross sections (CS)

	Panel: Mean	SD	CS: Mean	SD
Taxable income (1990 dollars)	29,489	48,501	286,603	1,222,002
Single dummy	0.30	0.46	0.22	0.42
Marginal tax rate (state+federal)	28.37	9.33	35.87	14.07
Change in rate at initial income (pp)	-1.86	4.71		
Log change in income	0.03	0.85		
Number of observations	59,199		1,380,590	
Max # obs in a year	10,717 (1987)		203,448 (1979)	
Min # obs in a year	4,448 (1983)		76,134 (1983)	

While this data set may not offer the cleanest imaginable identification—tax changes are not randomly assigned to households at the individual level—we believe it is appropriate for our exercise. For one, it contains tax changes that affect people throughout the income distribution, which allows us to estimate ETIs at each income level. Second, the data includes substantial tax variation at both the state and federal level, and as such workers' behavioral responses to these changes can be thought of as typical for real tax reforms. Finally, it contains tax changes of different sizes, and one of our strategies for estimating ETI variance leverages this variation in treatment size.

2.5.2 Estimation of sufficient statistics

In order to evaluate our rationalizability test, we require estimates of the shape of income tax schedule, the shape of the income distribution, and several moments of households compensated, income, and super-ETIs conditional on income. Of these, the last is by far the most ambitious relative to the existing literature, which typically estimates only mean uncompensated within coarse income bins.

To make this task possible, we impose three structural assumptions:

1. Households have no income elasticity.

2. Households have CES disutility of labor supply.

[28] These are computed using the NBER TAXSIM program, which calculates liabilities under US Federal and State income tax laws from individual data.

3. Households respond fully to changes in taxes within three years.

These assumptions allow us to focus on estimating the income-conditional moments of medium-run compensated elasticities.[29] While these assumptions are strong, they are common in the literature and we feel that—despite them—our estimates provide a plausible assessment of our rationalizability test. Still, our results should all be interpreted with significant caution.

2.5.2.1 Mean elasticities

We begin by estimating mean ETIs by income level. Our empirical strategy closely mirrors that of Gruber and Saez (2002), except that—instead of estimating a single ETI—we estimate ETIs locally in the space of year-demeaned log income $\tilde{z}$. Concretely, we estimate the following local regression:

$$\Delta \log z_t^h \;=\; a(\tilde{z}) \;+\; \varepsilon(\tilde{z}) \cdot \Delta \log R_t'(z_{t-1}^h) \;+\; c(\tilde{z}) \cdot mrs_{t-1}^h \;+\; d_{t-1} \;+\; \xi_t^h \tag{2.15}$$

where Δ represents time differences in t holding $t-1$ fixed, so that $\Delta \log R_t'(z_{t-1}^h) = \log R_t'(z_{t-1}^h) - \log R_{t-1}'(z_{t-1}^h)$ is the change in marginal retention at h's initial income. The year-demeaned-log-income-specific constant $a(\tilde{z})$ controls for differences in typical income changes in different brackets so that our estimates are not biased by, for example, mean reversion. We also control for a marital status dummy and time fixed effects.

Methodologically, we implement this regression by first differencing out by year fixed-effects and then running local-polynomial regressions in year-demeaned-log-income space with a constant bandwidth and using the Epanechnikov kernel. We optimally select this bandwidth using a leave-one-out cross validation procedure.[30] We compute confidence intervals by using the basic bootstrap method described in Chapter 5 of Davison and Hinkley (1997).

Under the assumption—which we maintain throughout—that, conditional on our controls, changes in marginal tax rates *within a (demeaned) income level* are as good as randomly assigned, this regression identifies the average ETI of households with each year-demeaned log-income $\tilde{z}$. This is exactly the statistic required by our theory.[31] Given the presence of year fixed effects, our identifying variation consists of (a) within-year, within-income variation in tax changes across individuals (especially those living in different states) and (b) across-year variation in the relative tax rates between different income levels.[32] Year fixed effects control for the fact that the government may adjust taxes in anticipation of changes in aggregate economic conditions, so long as these tax changes are uniform across tax brackets.

Figure 2.4 shows our estimates of ETI by income level.[33] Our estimates are consistent with—though somewhat on the lower end of—estimates in the literature. In line with Gruber and Saez (2002) and Kumar

[29] They imply that compensated elasticities are equal to uncompensated elasticities and that there households have no super-elasticities, except through changes in the curvature of the tax schedule.

[30] For more details, see Appendix 2.9.5.

[31] Our estimation strategy differs from some earlier work, such as that of Gruber and Saez (2002), which identifies a hypothetical elasticity concept based on locally linear taxes. Our elasticity concept is inclusive of changes in income caused by "knock-on" changes in marginal taxes as a household adjusts its income.

[32] Decomposing these sources of variation helps to clarify in what cases our identifying assumptions are violated. For example, variation through (a) violates the identifying assumptions if incomes decline in a state in the same year that state has a tax change, but for unrelated—and yet not statistically independent—reasons. Variation through (b) violates the identifying assumptions if incomes at some income level decline differentially (relative to those at other income levels) in the same year that taxes change differentially across those brackets, but for unrelated—and yet not statistically independent—reasons.

[33] As discussed above, we estimate ETI locally in year-demeaned-log-income space. Figure 2.4 converts these estimates to income space in 1990 by combining log income levels that are the same distance from average log incomes in their respective years. Appendix Figure 2.12 shows the original estimates in year-demeaned-log-income space.

and Liang (2020), we find somewhat higher elasticities at the bottom and the top of the income distribution.

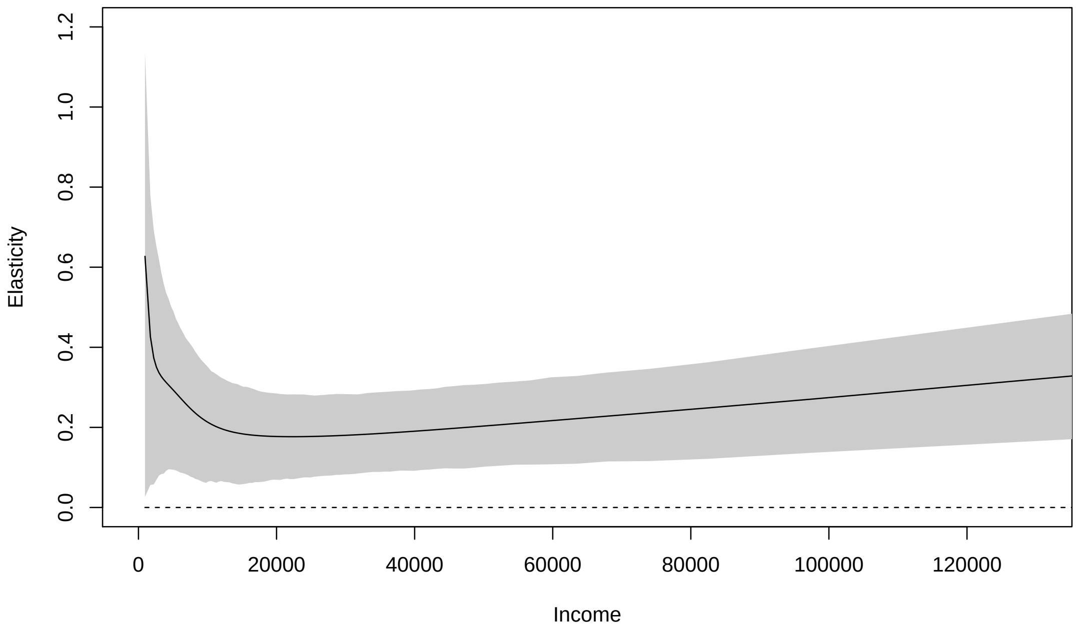


Figure 2.4: Mean ETI by income level in 1990 USD, and 95% confidence bands.

2.5.2.2 Variance of elasticities

Our second object of interest is the variance of ETI conditional on income level. While there is some precedent for studying differences in ETIs across groups, our approach varies in two main ways. First, the existing literature does not emphasize within-income variation and in some cases explicitly focuses on across-income variation (Eissa and Liebman, 1996; Gruber and Saez, 2002; Kopczuk, 2005; Kumar and Liang, 2020). By contrast, our test explicitly calls for within-income variance estimates. Second, the literature has focused on differences in elasticity conditional on observed characteristics. However—insofar as our test does not apply when taxes can condition on observed heterogeneity—we are particularly interested in unobserved heterogeneity (see Section 2.4.3.1). This focus brings with it additional identification challenges, which we discuss.

We pursue two main strategies: (a) a conditioning on observables approach based on differences across itemization status and (b) a structural estimation approach that relies on the linearity of tax responses.

Conditioning on observables: itemization status

Motivated by work that has documented higher elasticities for households who itemize deductions, we use the number of itemizations to split our sample into heavy and low itemizers (Gruber and Saez, 2002; Kopczuk, 2005). While we could simply compare household who do and do not itemize (at all), this would be problematic in our setting because almost all high-income households do itemize. So that we have variation in itemization status at each income level, we categorize households as low or high itemizers depending on whether they have below or above the mean level of itemizations *at their income level*.[34] We define

[34] Concretely, we compute the mean number of itemizations by income bracket (*nitem*) non-parametrically by estimating the

itemization status in year before the tax changes we consider, so that it does not implicitly control for changes in income.

After classifying individuals into these two groups, we estimate the local following regression at each year-demeaned income level $\tilde{z}$:

$$\begin{aligned} \Delta \log z_t^h \;=\; a(\tilde{z}) \;+\; \varepsilon_{L,t}(\tilde{z}) \cdot \Delta \log R_t'(z_{t-1}^h) \;+\; \delta(\tilde{z}) \cdot HI_{t-1}^h \cdot \Delta \log R_t'(z_{t-1}^h) \\ +\; c(\tilde{z}) \cdot mrs_{t-1}^h \;+\; d_{t-1} \;+\; \xi_t^h \end{aligned} \tag{2.16}$$

where HI_{t-1}^h is an indicator for high itemizers. This differs from our estimation of unconditional mean elasticities in two ways. First, we interact tax changes with a high income dummy in order to estimate the difference in elasticity by itemization status. Second, by allowing $\varepsilon_{L,t}(\tilde{z})$ to differ across years, we ensure that $\delta(\tilde{z})$ measures only *within-year* differences in elasticities by itemization status, despite the fact that during our sample period, a broadening of the tax base during our sample period may have both reduced elasticities and reduced itemizations (Kopczuk, 2005).[35] We estimate this regression locally in year-demeaned log-income space with the same methodology described in Section 2.5.2.1. To compute confidence intervals, we bootstrap both steps of the procedure just described, holding fixed the choices for the bandwidths, and use the basic bootstrap confidence interval.

The left panel of Figure 2.5 shows our estimates of differences in elasticity between heavy and light itemizers.[36] Next, we combine these estimates with estimates of the shares of heavy-itemizers by income bracket, in order to compute an implied lower bound on the variance of elasticities by income brackets.[37] Again, we compute confidence intervals by bootstrapping the whole procedure while holding bandwidth choices fixed. The right panel of Figure 2.5 shows the implied variance as a function of income. At high incomes, the lower bound on variance we estimate is already close to the level required to violate the back-of-the-envelope second-order condition discussed above.

This approach generates a conservative but robust estimate of the variance in ETI by income level. It is conservative because it leverages only differences in elasticities across two observable groups, ignoring observed and unobserved variation within these groups. At the same time, it is robust to the concern that (due to super-elasticities or tax salience) households' tax responses may be non-linear in the size of tax changes. Such non-linearities could cause us to mistake variances in the size of tax changes for variance in ETI at a given tax change; this is not an issue when estimating means, so long as each group faces similarly-sized tax changes.

Structural estimation

In order to estimate unobserved heterogeneity in elasticities, we complement the approach above with a second, more structural approach.

The key idea underlying this approach that if (a) each household's elasticities do not depend on the size

---

following local regression in the space of year-demeaned log income

$$nitem_{t-1}^h = a(\tilde{z}) + b(\tilde{z}) \log z_t^h + \xi_{t-1}^h.$$

Here, $a(\tilde{z})$ is our actual estimate, but we include the linear term as recommended by Fan and Gijbels (1996).

[35] Had we not allowed $\varepsilon_{L,t}(\tilde{z})$ to depend on year, we may have picked up *between-year* differences related to that there may be relatively more itemizers in years where both itemizers and non-itemizers have high elasticities—in which case the difference between average elasticities of itemizers and non-itemizers across years could exceed the average difference within years—or vice-versa.

[36] Appendix Figure 2.13 show the analogous plots of Figure 2.5 in the space of year-demeaned log income, where we perform the actual estimation before mapping to 1990.

[37] We estimate these shares by the same local-polynomial approach used for our other regressions.

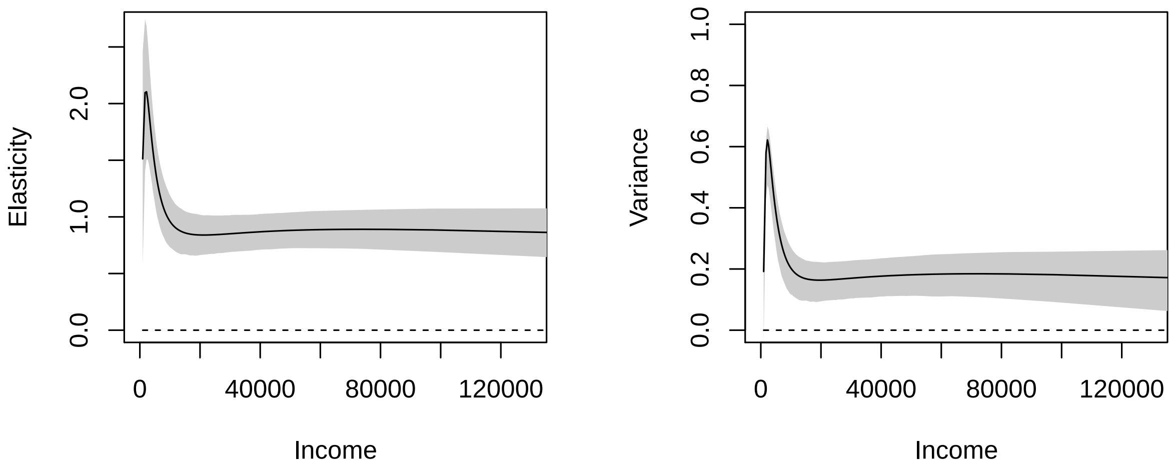


Figure 2.5: Left panel: Difference in elasticities between heavy and light itemizers, by income level in 1990. Right panel: Implied lower bound on variance in elasticity, by income level in 1990. In gray: 95% confidence bands.

of the tax changes they face and (b) one has access to continuous variation in the size of tax changes, then all moments of the distribution of elasticities—and in particular its variance—are non-parametrically identified. Intuitively, linearity implies that as the treatment size increases, the variance of treatment responses can be identified out of the change in the dispersion of the outcomes. This is because the heterogeneity in treatment effects should progressively magnify the dispersion of outcomes as the size of the treatment increases.[38]

While linearity is, admittedly, a strong assumption, it allows us to remain very flexible in modelling the structure of ETI heterogeneity. In particular, we assume that the distribution of elasticities has finite support, but allow this support to have many points, each taking arbitrary values. This flexibility allows us to approximate many different distributions, while discreteness ensures that our resulting estimates of ETI variance cannot be driven by a fat tail.

More explicitly, we assume that

$$\widehat{\Delta \log z_t^h} \quad = \quad a_{g_t^h} \quad + \quad \varepsilon_{g_t^h} \cdot \widehat{\Delta \log R_t'(z_{t-1}^h)} \tag{2.17}$$

where g_t^h takes a finite number of values, and where $\widehat{\Delta \log z_t^h}$ and $\widehat{\Delta \log R_t'(z_{t-1}^h)}$ are changes in log income and change in log taxes at initial income, respectively, after partialling out by marital status, a 10-piece linear spline in log taxable income, and year dummies. We estimate this regression by running a k-means algorithm that minimizes the mean square error of (2.17) by assigning each (h, t) pair to a group, and iteratively estimating $\{a_g, \varepsilon_g\}$ within groups.[3940]

The left panel of (2.6) displays our estimated distribution of ETIs. The estimated distribution has a very large mass around zero and some mass at higher values, up to a maximum ETI of 16. While this number is large, we believe it is plausible as an intensive margin elasticity at low incomes or as a proxy for extensive margin labor supply or tax avoidance decisions at any income. For robustness, we also present versions of

[38] For a proof, see Appendix 2.9.5.1.

[39] The k-means algorithm dates back to Sebestyen (1962) and MacQueen et al. (1967), and has recently been applied by Bonhomme and Manresa (2015); Bonhomme et al. (2017); Lewis et al. (2019).

[40] We do this while fixing the number of groups k, which we then select with a Bayesian information criterion subject to a maximum possible number of groups $k_{max} = 100$ as in Bonhomme and Manresa (2015).

our main results in which ETIs are capped at 5.

Given assignments $\hat{\varepsilon}_t^h$ of ETIs to each individual at each date t (i.e. between t and $t+1$), we compute the variance in ETIs by income level using local polynomial regressions.[41] In order to ensure that we capture only *within-year* variation in elasticities, we estimate the following regression locally in the space of year-demeaned log incomes $\tilde{z}$:

$$\left(\hat{\varepsilon}_t^h - \hat{m}_t(\tilde{z})\right)^2 = a(\tilde{z}) + b(\tilde{z}) \log z_t^h + \xi_{t-1}^h \tag{2.18}$$

where $\tilde{z}_t^h$ is year-demeaned log income, and $\hat{m}_t(\tilde{z}_t^h)$ is the year and income specific mean of elasticities, also estimated using first order local polynomials.

One important concern about this procedure is that it may be prone to small sample bias, as we estimate a large number of parameters. To address potential small sample bias and to obtain confidence intervals for our estimates of variance, we bootstrap the entire procedure, holding fixed the number of groups k and all bandwidths, and then subtract from our point estimates the bootstrap estimator of the bias.

The right panel of Figure 2.6 shows our structural estimates of income-conditional variance in ETI.[42] Notably, we estimate significantly higher variances than with the lower bound approach of the prior section. Appendix Figure 2.14 shows our structural estimates when ETIs are capped at 5.

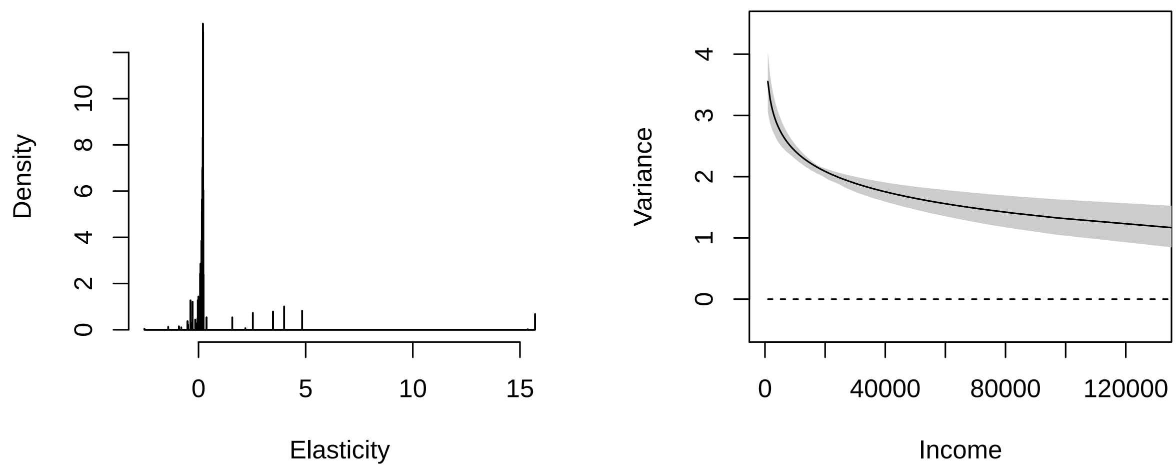


Figure 2.6: Left panel: Distribution of ETIs identified by k-means procedure. Right panel: Implied variance in elasticities by income level. In gray: 95% confidence bands.

2.5.2.3 Estimation of tax schedule, income density

Our rationalizability test also relies on estimates of the shape of the tax schedule and the income distribution. We estimate these moments in the NBER cross sectional files, which contain more observations and so allow for more precise estimation.

To compute $\alpha(z)$ in each year, we estimate a smooth functional form for the CDF of taxable income $G_t(z)$ in each year t with local polynomials regressions to third degree in log taxable income. Here we follow Fan and Gijbels (1996), who recommend including terms up to one order above the derivative of interest. We then translate our estimated coefficients for the first and second derivatives of $G_t(z)$ in log income into

[41] We follow an analogous procedure to compute the third moment of ETIs.

[42] Here again we have combined incomes across years that share a common year-demeaned log income.

an estimate of $\alpha_t(z)$.[43] We obtain confidence intervals by bootstrapping the whole procedure while holding the estimated optimal bandwidth constant. Our estimates are broadly consistent across years and similar to other recent work such as those of Hendren (2020). The left panel of Figure 2.7 shows our estimates in 1990.

We take a similar approach to estimate marginal tax rates and each of its derivatives required by our test. In these cases we estimate local-polynomial regressions of order d, where d is the nearest odd integer greater than the derivative of interest (in log income space). Again, we boostrap confidence intervals holding optimal bandwidths fixed. The right panel of Figure 2.7 shows our estimates in 1990.[44]

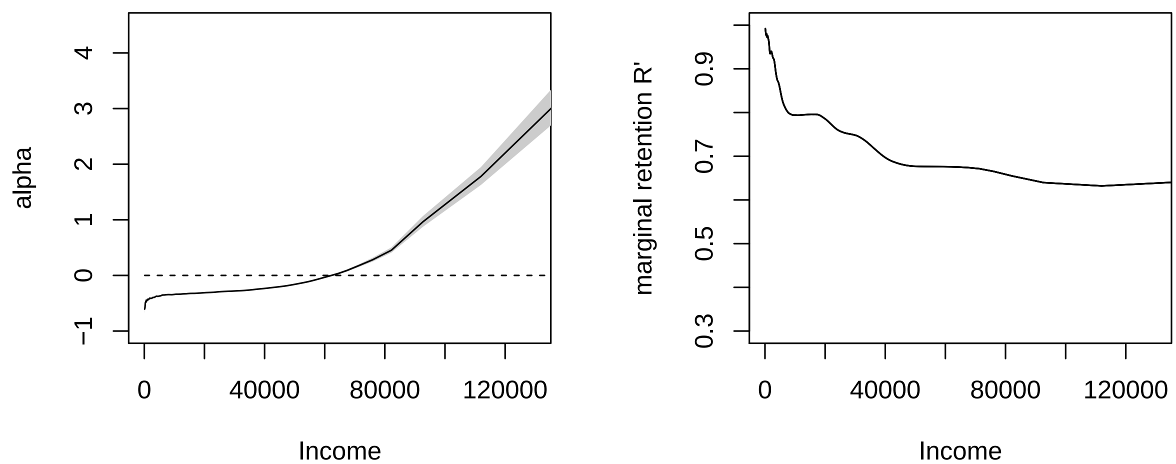


Figure 2.7: Left panel: Local shape parameter of the 1990 income distribution, by income level. Right panel: Average marginal taxes in 1990, by income level. 95% confidence bands.

By construction, our estimates of the tax schedule do not feature any kinks. This reflects that we run local regressions with a finite bandwidth, which smooths out the schedule. To the extent that kinks affect our rationalizability test, we therefore ignore their effects. However, we think that a smoothed tax schedule may be a realistic interpretation of the way that households *perceive* a kinked tax schedule, or a proxy for their inability to perfectly adjust their incomes, and so do not dwell on this issue (Rees-Jones and Taubinsky, 2020).[45]

2.5.3 Evaluation of test

Having computed each of the elements in the (ABC) and (DEFG) formulas, we can evaluate whether the tax schedule is rationalizable.

[43] Namely our estimate of $\alpha_t(z)$ is equal to the negative of our local estimate of $\frac{d^2 G_t(z)}{d \log z^2}$ divided by our local estimate of $\frac{dG_t(z)}{d \log z}$.

[44] We show all years together in Appendix Figure 2.16 to 2.26.

[45] To ensure our test is robust to this treatment of kinks, we also estimate our second-order test with an alternative methodology that excludes the term proportional to tax curvature. This has negligible effects, see Appendix Figure 2.39.

2.5.3.1 First-order test

We start with the first-order test (ABC), which recall assesses whether taxes are above the top of the Laffer curve at each income level z. Concretely, we compute—for each year t in our data—the statistic

$$\widehat{ABC}_t(z) \;=\; -1 \;+\; \frac{1-R_t'(z)}{R_t'(z)}\left(\alpha_t(z) - \frac{d\log}{d\log z}\left(\frac{1-R_t'(z)}{R_t'(z)}\right) - \frac{d\log\varepsilon(z)}{d\log z}\right)\varepsilon(z) \tag{2.19}$$

where recall R_t and α_t are estimated separately in each year. We compute confidence intervals for $\widehat{ABC}_t(z)$ by combining bootstrap replications from the two different data sets we use—the panel and the cross-sectional file for year t—assuming that observations in separate datasets are drawn independently from each other.

Our estimates of $\widehat{ABC}_t(z)$ are broadly consistent across years and consistently negative, implying that taxes are below the top of the Laffer curve. The left panel of Figure 2.8 shows this for one representative year; Appendix Figures 2.27 to 2.37 show our estimates for all years.

2.5.3.2 Second-order test

We now evaluate our new, second-order test (DEFG), which recall assesses whether taxes are in a Laffer valley at each income level z. Concretely, we compute—for each year t in our data—the statistic[46]

$$\widehat{DEFG}_t(z) = -\left(1+R_t'(z)\right)\varepsilon(z) + \left(1-R_t'(z)\right)\left[\alpha_t(z)\varepsilon^2(z) - \frac{d\varepsilon^2(z)}{d\log z} + \frac{d^2\log R_t'(z)}{d\log z^2}\varepsilon^3(z)\right] \tag{2.20}$$

where the elasticities are derived from our structural estimates. We compute confidence intervals in the same way described for $\widehat{ABC}_t(z)$.

Our estimates of $\widehat{DEFG}_t(z)$ are broadly consistent across years. Strikingly, they fail at incomes above around $90,000$ in 1990 USD. The right panel of Figure 2.8 shows this for one representative year; Appendix Figures 2.27 to 2.37 show our estimates for all years. In Appendix Figures 2.38 to 2.40, we compare $\widehat{DEFG}_t(z)$ with and without including the last two terms of (2.20); this has almost no effect, implying that our estimates are not driven by the third moment of elasticities or steep changes in the second moment.

An alternative way to visualize (DEFG)—as well as to compare the implications of our different estimates of variance—is to reframe the second-order test in terms of the minimum level of variance at which it fails. We can then compare this level with our estimates of variance, at each income level. Figure 2.9 illustrates this comparison.[47] Blue and orange shaded regions represent the levels of ETI variance at which the (DEFG) test passes and fails, respectively, at each income. On top of this background, we superimpose our estimates of variance based on the two strategies in Section (2.5.2.2). The test easily fails under our structural estimates of variance, while our lower bounds on variance based on itemization status come about half of the way to violating the test. Not shown below (see Appendix Figure 2.14) are our structural estimates of variance when elasticities are capped at 5; these violate the test on point estimates but cannot reject that the test passes.

[46] This is equal to the RHS of (DEFG) when, as we have assumed, income effects are zero and labor supply preferences are additive-CES; see Appendix 2.9.1.4.1 for details regarding the super-elasticity term.

[47] It does so under the assumptions that, at each z, (a) the tax schedule is locally linear and (b) the variance of ETIs is locally constant. This allows the test to be framed as simple "variance test" without considering the other roles played by second and third moments of ETIs. Figure 2.39 shows that this simplification is without significant loss.

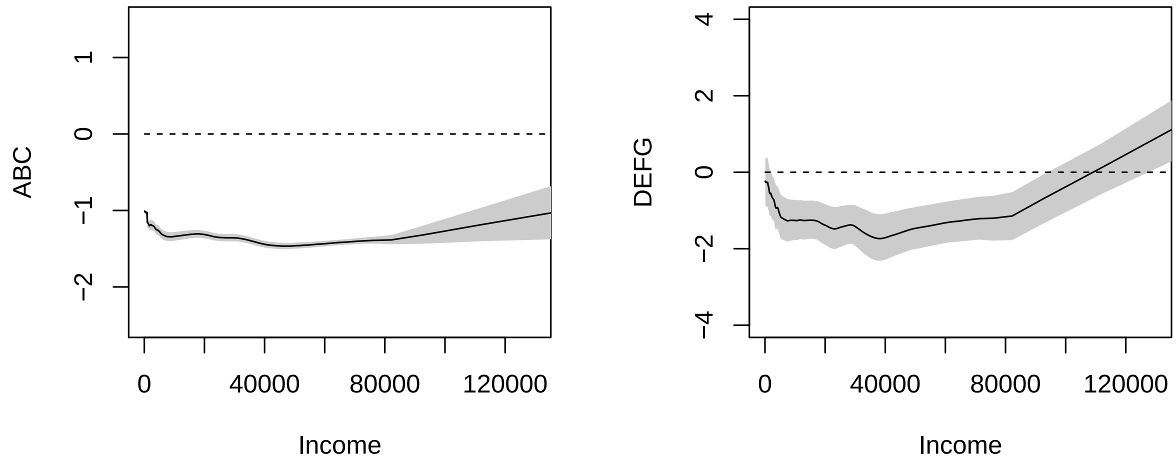


Figure 2.8: Left panel: ABC test evaluated in 1990. The test passes at all incomes. Right panel: DEFG test evaluated in 1990. The test fails at high incomes. 90% confidence bands.

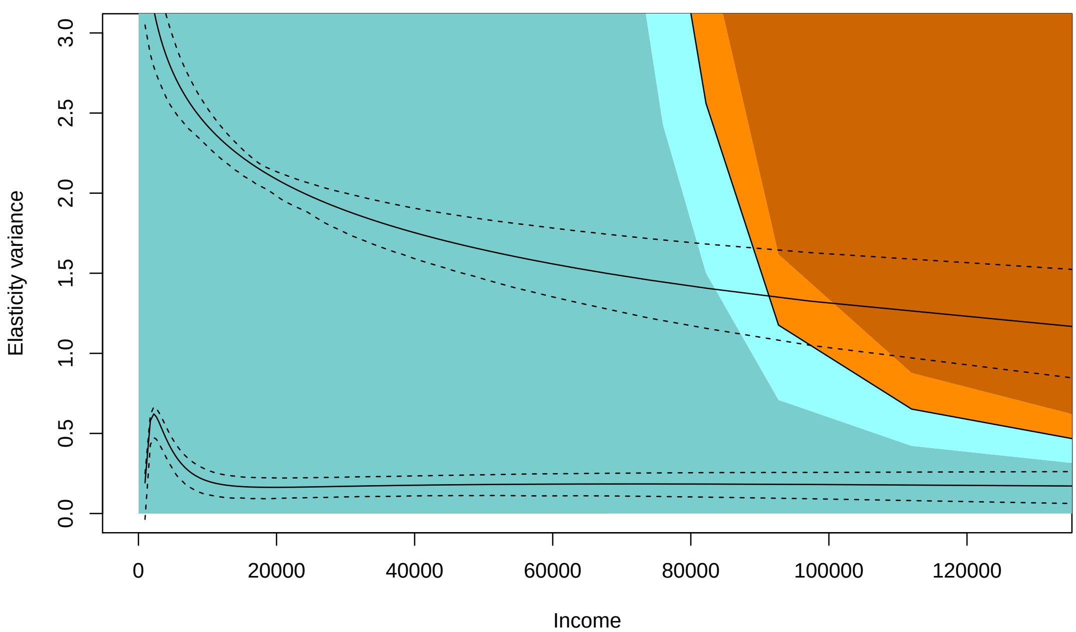


Figure 2.9: 1990. Background: ETI variance consistent (blue) and inconsistent (orange) with (DEFG). Foreground: Estimates of ETI variance. Lighter colors and dashed lines colors are 95% confidence bands.

2.5.4 Inspecting the mechanism

A number of potential confounding factors could prevent the heterogeneity in elasticities we have estimate from translating into the type of elasticity-sorting that our theory predicts. For example, perhaps our estimates mistake differences across households in the *timing* of tax responses for differences in the *size* of tax responses.[48] Another possibility is that super-elasticities—which we do not estimate—could work to attenuate the level of elasticity sorting, similar to the discussion in Section 2.4.3.2. Alternatively, heterogeneity

[48]Our conditioning-on-observables estimates are robust to this concern, but our structural estimates are not.

in elasticities across income levels may reflect differences in the institutional arrangements of employment in different jobs, which may not respond to tax changes.

To address these concerns, we now attempt to directly validate the mechanism that underlies our theory, i.e. the idea that changes in taxes cause changes in the average elasticity within each income level. The main idea behind our strategy is as follows: In regions of the tax schedule where the income density is strongly decreasing ($\alpha(z) > 0$), a locally flat increase in marginal retention should increase the average elasticity conditional on income. This is because—when there is significant heterogeneity in elasticities and when all elasticity groups have proportional densities—the high elasticity types that end up at each income level z come from much lower incomes, where the density is higher. This pattern flips where $\alpha(z) < 0$, so that an increase in marginal retention should lower the average elasticity at z.

More concretely, one can show that that for small tax changes,

$$\Delta\varepsilon(z) \quad \approx \quad \alpha(z) \;\; \mathbb{V}\mathrm{ar}[\varepsilon^h(R)|z_0^h = z] \;\; \Delta R'(z). \tag{2.21}$$

In other words, there is a positive interaction between the ETI $\varepsilon(z)$ at a given income level and the product of the Pareto shape $\alpha(z)$ and the level of marginal retention $R'(z)$, and the size of this interaction term is equal to the local variance of ETIs. In principle, we could therefore identify ETI variance at each income level by including in the mean ETI regression (2.15) an interaction between tax changes and the product of $\alpha(z) \cdot R'(z)$. In practice, this approach has two limitations, and these limitations motivate our actual strategy. First, it may be under-powered, so we pool across income levels. Second, one source of variation in $R'(z)$—tax differences across states (rather than years)—may be correlated with elasticity differences for reasons other than our mechanism and therefore bias this regression. Namely, one should expect that states with more income-elastic or cross-state-migration-elastic populations (at any income z) respond by levying lower marginal taxes. We therefore focus on the component of variation in taxes that comes solely from time variation.

Our preferred specification is as follows:

$$\begin{aligned} \Delta \log\left(z_t^h\right) &= \varepsilon \cdot \Delta \log R_t'(z_{t-1}^h) \\ &\quad + \gamma \cdot \Delta \log R_t'(z_{t-1}^h) \cdot (\overline{\log R'}_{t-1}(z_{t-1}^h) - \overline{\overline{\log R'}}(z_{t-1}^h)) \cdot \alpha_{t-1}(z_{t-1}^h) \\ &\quad + c \cdot mrs_{t-1}^h + d_{t-1} + f(z_{t-1}^h) + \xi_t^h \end{aligned} \tag{2.22}$$

where $\overline{\log R'}_{t-1}(z_{t-1}^h)$ is the average log marginal retention rate at z_{t-1}^h in year $t-1$, $\overline{\overline{\log R'}}(z_t^h)$ is the average log marginal retention rate at z_t^h across years, and where $f(z_{t-1}^h)$ is a control for the level of income (either a single linear term or ten piece splines). As discussed above, one may interpret the regression coefficient on this interaction term as an average of within-income ETI variances, taken across incomes. Secondarily, we also estimate (2.22) using only cross-state variation—in which case the cross-state within-year average shape $\alpha_{t-1}(z_{t-1}^h)$ and average marginal taxes $\overline{\log R'}_{t-1}(z_{t-1}^h)$ are replaced by the analogous within-state across-year average shape and average marginal taxes. Finally, we also estimate (2.22) using both cross-state and cross-year variation, using the analogous within-state and within-year average shapes and average marginal taxes.

Table 2.2 presents our results. In most cases we estimate positive but small and statistically insignificant implied variances, which we interpret as moderately supportive of our theory. The small size of our estimates using year × state variation is likely explained by attenuation bias stemming from noise in our estimates of

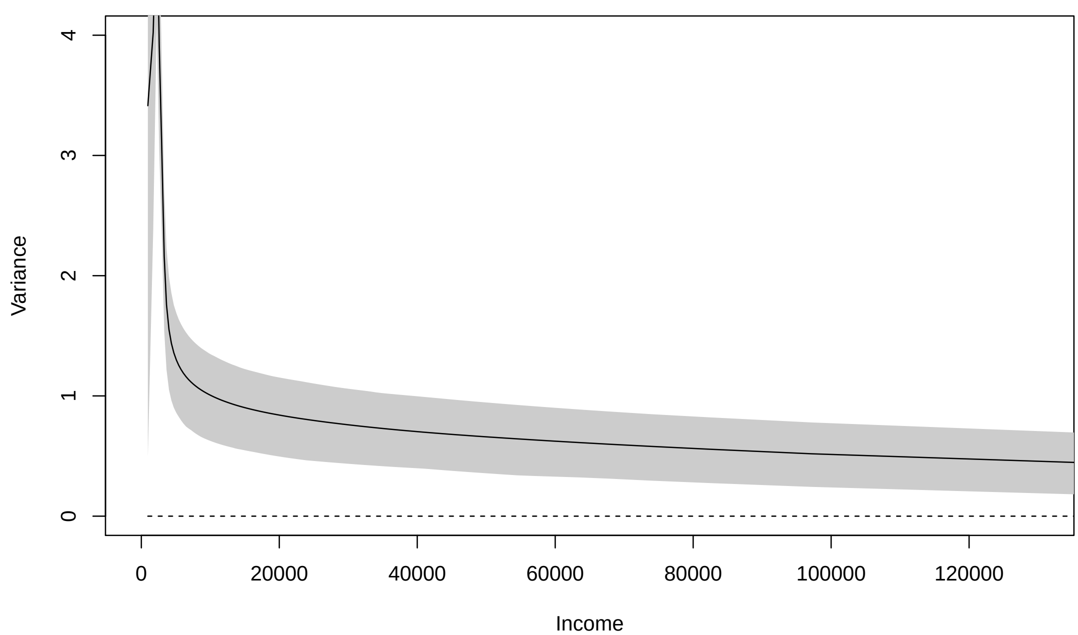


Figure 2.10: Income-conditional variance implied by elasticity differences across low- and high-tax years in 1990, 95% confidence bands.

local αs.

Table 2.2: Inspecting the mechanism

	year		state		year state	
elasticity	0.44	0.21	0.49	0.21	0.48	0.22
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
variance	1.05	0.11	0.38	0.10	-0.05	0.22
	(0.19)	(0.12)	(0.29)	(0.18)	(0.32)	(0.31)
income control	linear	splines	linear	splines	linear	splines

Notes. Estimates for means and variances of elasticities using Equation 2.22. Bootstrapped standard errors in parentheses. The first two columns report results where the variance of elasticities is inferred from variation in elasticities across years. In the third and fourth column, it is inferred from variation across states, and in the fifth and sixth it is inferred from variation across years and states.

Finally, we extract an implicit measure of variances by income bracket by running the regression above non-parametrically, allowing $\gamma(z), \epsilon(z)$ and $c(z)$ to depend on demeaned income. Figure 2.10 shows that, in the year-variation-only case, we estimate implicit variances similar to those in Section 2.5.2.2. In the other specifications, our estimates are very noisy and close to zero; see Appendix Figure 2.15. As we believe the year-only variation is better identified, we conclude that this exercise lends moderate support to our theory overall.

2.6 Quantitative application

So far, we have presented evidence that the taxes in our sample can—from the perspective of any planner in a broad class—be improved. But how large are these potential welfare gains, and what sort of tax changes achieve them? In this section, we speak to these questions using a simple calibrated model.

Model: Motivated by the fact that our efficiency test fails only at high incomes, we remain agnostic to the behavior of households with incomes below \$100,000 and to the planner's preferences regarding their welfare. For households with incomes above \$100,000, we impose both positive and normative assumptions. Positively, we assume the following: First, consistent with the year 1990, top incomes are distributed as Pareto($\alpha_{\text{top}} = 2.5$), are subject to an initial top marginal tax rate of $\tau_{\text{top}} = 35\%$. Second, each top-earning household h has a constant compensated elasticity of labor supply β^h and no income elasticity. Third, β^h is distributed independently of productivity according to a Gamma distribution with mean 0.3 and variance 1.2.[49]

Normatively, we must take a stance on how the planner values changes in household welfare relative to government revenue in order to quantify the gains from tax reform. To do so, we assume the following: First, the planner places a common and constant value λ_{top} on transfers to each top-earner, relative to government revenue.[50] Second, in setting the initial tax schedule, the planner has "followed the literature" by checking the first- but not the second-order condition of the tax schedule. Together with our positive assumptions, this allows us to back out the welfare weight:

$$\lambda_{\text{top}} \quad = \quad 1 \; - \; \frac{1 - R'_{\text{top}}}{R'_{\text{top}}}\, \alpha_{\text{top}}\, \varepsilon_{\text{top}} \quad \approx \quad 0.6. \tag{2.23}$$

Simple tax reforms: To motivate the particular tax changes we consider, we return to the planner's second-order condition. In our specialized model, this takes a particularly uncomplicated form for any tax change Δ that only affects top earners:[51]

$$\frac{d^2}{d\epsilon^2}\text{ Welfare}(R + \epsilon\Delta) \;=\; \int \underbrace{\left[-(1 + R_{\text{top}})\varepsilon_{\text{top}} + (1 - R_{\text{top}})\alpha_{\text{top}}\varepsilon^2_{\text{top}}\right]}_{\text{LHS of (DEFG)}} \left(\frac{\Delta'(z)}{R'_{\text{top}}}\right)^2 zg(z)dz. \tag{2.24}$$

One may easily verify that the "LHS of (DEFG)" term is positive under our positive assumptions, implying that the planner's second-order condition fails, i.e. taxes are in the Laffer valley.

This expression says that a change in taxes on top earners improves welfare to the extent it changes the marginal taxes they face. In particular, there is no need to restrict to very "narrow" tax changes—as was done in the main text—in order to exploit the failure of the second-order condition; narrow tax changes are just a technical device used to handle income effects and non-constant welfare weights, both of which are absent here. Even simple tax changes, such as raising or lowering marginal taxes on all top earners, can "sort and extort," as illustrated in the motivating example of Section 2.2.

This in mind, we consider two very simple tax reforms: A 20 percentage point increase and a 20 percentage point decrease in the top marginal tax rate.

[49] We view this as a sensible approximation of our empirical estimates of elasticity heterogeneity.

[50] Here, we implicitly assume that the planner does not face a hard budget constraint, but instead has a constant marginal value of public funds.

[51] This follows from Lemmas 6 and 7; see Appendix 2.9.3.1.

Welfare gains and discussion:

Consistent with the observation that taxes are in a Laffer valley, we document welfare gains from either raising or lowering taxes. Concretely, the planner's gains from either raising or lowering top taxes by 20pp are equivalent to the social value of transferring a little more than \$3000 to each household in the top bracket or $\lambda_{\text{top}} \cdot$ \$3000 $\approx$ \$2000 to the planner for each household in the top bracket.

The estimate above is a lower bound on welfare gains for several reasons. First, we take all taxes below \$100,000 as given, thus ignoring the potential gains from simultaneously changing them. Second, we restrict taxes above \$100,000 to be linear, ignoring the additional gains from using non-linear tax changes. Third, we limit ourselves to 20 percentage point changes in the top tax rate, ignoring the gains from using larger tax changes. Fourth—and somewhat more subtly—we ignore any behavioral effects on households with initial incomes outside of the top bracket. This results in an underestimate since the only case in which such households respond to top rate changes is when they jump into the top bracket, which has positive fiscal externalities.

One topic not addressed above is the shape of the *optimal* tax reform, and in particular whether it is similar to the simple reforms we consider. For instance, one may worry that a schedule with many "squiggles" would be more effective in sorting households with different elasticities and therefore optimal. While we do not answer this question in general, Appendix 2.9.2.3 presents a suggestive exercise in which, starting from a sub-optimal tax schedule, we update taxes in the direction suggested by the planner's first-order condition until converging to a new local optimum. In a simple case with two elasticity types, we find that taxes converge to a "two-part" schedule that smoothly transitions between a high tax rate used on low incomes and a low tax rate used on high incomes—rather than a "squiggly" schedule.

2.7 Discussion

2.7.1 Relation to inverse optimality literature

Overall, we interpret the empirical violation of our rationalizability test as new and surprising observation about the (in)efficiency of US income taxes. The existing literature—typified by the first-order (ABC) test—has tended to interpret the US income tax schedule as Pareto efficient and reflective of a particular set of social preferences that places somewhat greater weight on transfers to lower-income households (Bargain et al., 2014). Our interpretation differs in two ways: First, the violation of our second-order test implies that taxes are, for any social planner in the broad class we consider, inefficient. Second, not only are taxes inefficient, but also—because they are un-rationalizable,—it is misguided to interpret taxes as reflecting any set of social preferences. In this sense, our findings are critical of the "inverse optimum" literature which attempts to infer a planner's distributional preferences from the tax schedule.

At the same time, our results are consistent with a more recent interpretation of the inverse optimum approach which shows that first-order-condition-implied "as if" welfare weights can be used to value the distributional impacts of policies in a *welfare-function-independent* way (Hendren, 2020; Hendren and Sprung-Keyser, 2020). Specifically, Hendren (2020) shows that if the weighted incidence of a policy change is positive, then—if it is accompanied by an appropriate change in income taxes—it can be used to create a Pareto improvement.[52] While this result does not rely on our second-order condition, its interpretation is somewhat different when the second-order condition fails. Namely, the compensatory adjustments do lead to Pareto

[52] Although this result is derived in a one-dimensional model, a version of it carries over to our setting if policy changes have homogeneous impacts on households within each income level.

improvements, but also operate strictly within the frontier, since any planner could improve welfare through further tax changes.

2.7.2 Robustness of ETI variance estimates

We have presented a variety of estimates for variances of elasticities, some more conservative and others requiring stronger assumptions.

Perhaps the most robust are the implicit estimates of variance presented in Section 2.5.4, because they directly test the sorting method that underlies our theory. As we have discussed, these estimates are robust to concerns that super-elasticities may undo the sorting effects of ETI variance or that heterogeneity in elasticities across income levels reflect differences in the institutional arrangements that do not respond to tax changes. However, they may be confounded by other changes in the tax system over the decade. In particular, a broadening of the base may have reduced individual elasticities—for example by restricting tax avoidance—at the same time as marginal taxes decreased (Kopczuk, 2005). This force counterbalances the effects of the sorting mechanism we highlight at the top, and reinforces it at the bottom. It would therefore bias our estimates of ETI variance downward at the top and upward at the bottom.

Our lower-bound estimates of variance based on itemization status are robust to a different set of concerns. For one, they are not subject to the criticism above about base broadening. Also, as we discuss in Section 2.5.2.2, they do not risk mistaking non-linearities in tax responses—i.e. heterogeneity in ETIs across sizes of tax changes—with heterogeneity in ETIs for tax changes of a given size.

By contrast, our structural estimates rely on stronger assumptions, particularly the linearity of household tax responses and the discreteness of the elasticity distribution. Of course, they also have the advantage of flexibly estimating unobserved heterogeneity. One potential concern to which they are robust is that household preferences may follow a dynamic process through time, so that the same individual who is very elastic or very productive in one year is less so during the next. So long as the *distribution* of types in the population is constant, this reshuffling has no effect whatsoever on our results.

Taken together, our estimates present a consistent and robust picture of the variance of ETIs by income level. In particular, our estimates suggest that the (DEFG) is very likely to fail for high incomes, even if a few of the possible identification issues discussed above are active.

A final potential concern is that our theory only models intensive labor supply decisions, whereas actual ETIs may involve extensive margin decisions. For example, households may join or leave the labor force, find a second job, migrate between states, or hire a tax accountant. To the extent these decisions cannot be repeated, they stand against the idea that households have a stable elasticity according to which a planner can sort them. Of course, many extensive margin decisions, such as moving between jobs or states, can be repeated, opening up the possibility for the planner to sort individuals based on sequences of these decisions. Labor force participation may pose a problem for our estimates at low incomes, but insofar as those decisions are less prevalent at the top of the income distribution, they should not affect our conclusion that the test fails at high incomes.

2.8 Conclusion

We take a second-order approach to the classical non-linear income taxation problem. Far from a technical detail, the second-order condition introduces a new qualitative idea for income taxation: Taxes must not only be below the top of the Laffer curve, but also must not lie in a “Laffer valley”.

Our theoretical results shed light on the relationship between the Laffer valley and household heterogeneity. As our discussion emphasizes, heterogeneity in household elasticities within income levels provides a tax reform motive for planners who are constrained to use a single income tax schedule. By changing taxes once, the planner can (in an imperfect way) sort high- and low-elasticity households into different parts of the income distribution; by changing taxes again, she can exploit this separation, as if she had access to elasticity-dependent taxes. We capture this insight in a simple test for the local rationalizability of the tax schedule in terms of locally estimable sufficient statistics.

Our empirical results take this novel test to the data in order to assess whether actual US tax schedules from 1979 to 1990 were rationalizable by any planner. We extend the approaches of existing empirical work to estimate ETIs by income level and to estimate not only ETI means but also ETI variance—a key statistic for our theory. Strikingly, our estimates reject the rationalizability of the tax schedule in every year of our sample. Said differently, any planner in the class we consider would prefer a different tax schedule; there is a free lunch available through tax reform. A conservative quantification exercise suggests that either raising or lowering top taxes by 20 percentage points results in yearly welfare gains equivalent to approximately $3000 per top earner.

Bibliography

Aliprantis, C. and Border, K. (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer-Verlag Berlin Heidelberg.

Bakker, L. (2018). Uniform convergence and differentiation.

Bargain, O., Dolls, M., Neumann, D., Peichl, A., and Siegloch, S. (2014). Tax-benefit revealed social preferences in europe and the us. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, (113/114):257–289.

Bierbrauer, F., Boyer, P., and Hansen, E. (2020). Pareto-improving tax reforms and the earned income tax credit.

Billingsley, P. (2008). *Probability and measure*. John Wiley & Sons.

Blau, F. D. and Kahn, L. M. (2007). Changes in the labor supply behavior of married women: 1980–2000. *Journal of Labor economics*, 25(3):393–438.

Bonhomme, S., Lamadon, T., and Manresa, E. (2017). Discretizing unobserved heterogeneity. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2019-16).

Bonhomme, S. and Manresa, E. (2015). Grouped patterns of heterogeneity in panel data. *Econometrica*, 83(3):1147–1184.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their application*. Number 1. Cambridge university press.

Diamond, P. A. (1998). Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. *American Economic Review*, pages 83–95.

Eissa, N. and Liebman, J. B. (1996). Labor supply response to the earned income tax credit. *The quarterly journal of economics*, 111(2):605–637.

Fan, J. and Gijbels, I. (1996). *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*, volume 66. CRC Press.

Folland, G. B. (1999). *Real analysis: modern techniques and their applications*, volume 40. John Wiley & Sons.

Gruber, J. and Saez, E. (2002). The elasticity of taxable income: evidence and implications. *Journal of public Economics*, 84(1):1–32.

Guesnerie, R. (1998). *A contribution to the pure theory of taxation*. Number 25. Cambridge University Press.

Hammond, P. J. (1979). Straightforward individual incentive compatibility in large economies. *The Review of Economic Studies*, 46(2):263–282.

Hendren, N. (2020). Measuring economic efficiency using inverse-optimum weights. *Journal of Public Economics*, 187:104198.

Hendren, N. and Sprung-Keyser, B. (2020). A unified welfare analysis of government policies. *The Quarterly Journal of Economics*, 135(3):1209–1318.

Himmelberg, C. (1975). Measurable relations. *Fundamenta Mathematicae*, (87):53—72.

Hines Jr, J. R. and Keen, M. J. (2021). Certain effects of random taxes. *Journal of Public Economics*, 203:104412.

Hosseini, R. and Shourideh, A. (2019). Retirement financing: An optimal reform approach. *Econometrica*, 87(4):1205–1265.

Jacquet, L. and Lehmann, E. (2015). Optimal income taxation when skills and behavioral elasticites are heterogeneous.

Jacquet, L. and Lehmann, E. (2020). Optimal income taxation with composition effects. *Journal of the European Economic Association*.

Kammar, O. (2016). A note on frechet diffrentiation under lebesgue integrals.

Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics*, 89(11-12):2093–2119.

Kumar, A. and Liang, C.-Y. (2020). Estimating taxable income responses with elasticity heterogeneity. *Journal of Public Economics*, 188:104209.

Lewis, D. J., Melcangi, D., and Pilossoph, L. (2019). Latent heterogeneity in the marginal propensity to consume. *FRB of New York Staff Report*, (902).

Li, Q. and Racine, J. S. (2007). *Nonparametric econometrics: theory and practice*. Princeton University Press.

Lockwood, B., Sial, A. Y., and Weinzierl, M. C. (2020). Designing, not checking, for policy robustness: An example with optimal taxation. Technical report, National Bureau of Economic Research.

MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA.

Martinsson, G. (2006). The implicit and inverse function theorems.

Maskin, E. and Riley, J. (1984a). Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196.

Maskin, E. and Riley, J. (1984b). Optimal auctions with risk averse buyers. *Econometrica: Journal of the Econometric Society*, pages 1473–1518.

Maurer, H. and Zowe, J. (1979). First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Mathematical programming*, 16(1):98–110.

Moser, C. and Olea de Souza e Silva, P. (2019). Optimal paternalistic savings policies. *Columbia Business School Research Paper*, (17-51).

Racine, J. (1993). An efficient cross-validation algorithm for window width selection for nonparametric kernel regression. *Communications in Statistics-Simulation and Computation*, 22(4):1107–1114.

Rees-Jones, A. and Taubinsky, D. (2020). Measuring “schmeduling”. *The Review of Economic Studies*, 87(5):2399–2438.

Rochet, J.-C. (1985). The taxation principle and multi-time hamilton-jacobi equations. *Journal of Mathematical Economics*, 14(2):113–128.

Rochet, J.-C. and Stole, L. A. (2003). The economics of multidimensional screening. *Econometric Society Monographs*, 35:150–197.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The review of economic studies*, 68(1):205–229.

Saez, E., Slemrod, J., and Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of economic literature*, 50(1):3–50.

Scheuer, F. and Werning, I. (2018). Mirrlees meets diamond-mirrlees: Simplifying nonlinear income taxation.

Sebestyen, G. S. (1962). Decision-making processes in pattern recognition.

Shorack, G. R. (2000). *Probability for statisticians.* Number 04; QA273, S4. Springer.

Stiglitz, J. (1982a). Utilitarianism and horizontal equity: The case for random taxation. *Journal of Public Economics*, 18(1):1–33.

Stiglitz, J. E. (1982b). Self-selection and pareto efficient taxation. *Journal of public economics*, 17(2):213–240.

Taubinsky, D. and Rees-Jones, A. (2018). Attention variation and welfare: theory and evidence from a tax salience experiment. *The Review of Economic Studies*, 85(4):2462–2496.

Vere, J. P. (2011). Social security and elderly labor supply: Evidence from the health and retirement study. *Labour Economics*, 18(5):676–686.

Werning, I. (2007). Pareto efficient income taxation. Technical report, mimeo, MIT.

2.9 Appendix

2.9.1 Formalized statements from the main text

Below, we provide formal statements of several definitions and assumptions stated loosely in the main text. These are organized into statements about taxes and tax changes, conditions on household and aggregate labor supply, social objective definitions, and supporting concepts for the rationalizability test. Within these subsections, we also provide basic technical Lemmas that illustrate the roles of several of the assumptions and provide a foundation used in the proofs of our main results.

2.9.1.1 Taxes and tax deviations

We begin with a basic regularity condition on taxes.

Assumption 1. *R is continuous on $\mathbb{R}_{\geq 0}$ and three-times continuously differentiable on $\mathbb{R}_{>0}$, and there exists $B^R > 0$ such that for all $z \in \mathbb{R}_{>0}$,*

$$\left|\frac{dR(z)}{d\log z}\right| \leq B^R|R(z)|, \qquad \text{and} \qquad \left|\frac{dR'(z)}{d\log z}\right|, \left|\frac{d^2R'(z)}{d\log z^2}\right| \leq B^R|R'(z)|. \tag{2.25}$$

Next, we define a space of feasible tax changes $\boldsymbol{\Delta}$ by

$$\boldsymbol{\Delta} = \left\{\Delta : \mathbb{R}_{\geq 0} \to \mathbb{R} \;\middle|\; \begin{array}{c} \Delta \text{ continuous, } \Delta \text{ three-times continuously differentiable on } \mathbb{R}_{>0}, \text{ and } \exists B \in \mathbb{R} : \\ \forall z \in \mathbb{R}_{>0},\ |\Delta(z)| \leq B|R(z)| \text{ and } |\Delta'(z)|, \left|\frac{d\Delta'(z)}{d\log z}\right|, \left|\frac{d^2\Delta'(z)}{d\log z^2}\right| \leq B|R'(z)| \end{array}\right\}. \tag{2.26}$$

This space is well-defined under Assumption 1. Note that the function $||\cdot||$ defined in (2.7) is well-defined on $\mathbf{\Delta}$. The following Lemma establishes that not only is $||\cdot||$ a norm on $\mathbf{\Delta}$, but also $(\mathbf{\Delta}, ||\cdot||)$ is a Banach space. We later leverage this fact in order to apply existing results on optimization in Banach spaces.

Lemma 1. $(\Delta, ||\cdot||)$ *is a Banach space.*

Proof. See Appendix 2.9.6.1. □

Throughout the paper, we will consider many functions of the form $f : R + \mathbf{\Delta} \to \mathbb{R}$, where $R + \mathbf{\Delta} \equiv \{R + \Delta \mid \Delta \in \mathbf{\Delta}\}$. As any such function may alternatively be understood as a function $\widetilde{f}(\Delta) = f(R + \Delta)$ on $\mathbf{\Delta}$, we will WLOG refer to such functions f as being Frechet in Δ when the corresponding $\widetilde{f}$ is, and with derivatives equal to those of $\widetilde{f}$. Notationally, we denote the Frechet derivative of any function $f : R+\mathbf{\Delta} \to \mathbb{R}$ evaluated at a point $\widetilde{R}$ by $Df(\widetilde{R})$, if it exists. For any $\Delta \in \mathbf{\Delta}$, we let $D_\Delta f(\widetilde{R})$ denote $Df(\widetilde{R})(\Delta)$, i.e. the first Frechet derivative of f at $\widetilde{R}$ in direction Δ with magnitude $||\Delta||$ (and similarly for higher derivatives).

2.9.1.2 Labor supply regularity conditions

Our first assumption on labor supply is a basic regularity condition on household preferences, satisfied by typical functional forms used in the literature.

Assumption 2. *Household utility is given by a function* $u : \mathcal{H} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R} \cup \{-\infty\}$. *On the restricted domain* $\mathcal{H} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0}$, *utility* $u^h(c, z)$ *is finite, has three continuous derivatives in* (c, z) *which are measurable*[53] *in* (h, c, z) *and satisfy* $u^h_c(c, z) > 0$ *and* $u^h_z(c, z) < 0$.

Our next assumption is of more qualitative importance. It is a set of three conditions which together guarantee that—locally to R—each household supplies labor purely on the intensive margin. In words, they are as follows: First, each household h's problem has a unique solution at R. Second, h's labor supply preferences—given the tax schedule—have some positive level of concavity locally, i.e. within a ball of radius ϵ^h around $\log z^h_0$. Third h has a strong enough preference for supplying labor at z^h_0 relative to any level outside the local neighborhood of concavity. Importantly for later results, the level of concavity and the relative consumption preference are both uniform across households. Intuitively, these conditions hold when households have sufficiently concave preferences relative to the curvature of the tax schedule.

Assumption 3. *There exist* $(\epsilon^h)_{h\in\mathcal{H}}, \eta, \bar{c} > 0$ *with* ϵ^h *h-measurable, such that for all* $h \in \mathcal{H}$

- *The problem* $\max_{z\in\mathbb{R}_{\geq 0}} u^h(R(z), z)$ *has a unique, strictly positive, and h-measurable solution* z^h_0 *at* $\widetilde{R} = R$,
- *there exists a function* $v^h(\tilde{z}) : B_{\epsilon^h}(\log z^h_0) \to \mathbb{R}$ *such that, for all* $z \in e^{B_{\epsilon^h}(\log z^h_0)}$,

$$u^h\left(R(z)e^{v^h(\log z)}, z\right) = u^h\left(c^h_0, z^h_0\right) \quad \text{and} \quad v^{h\prime\prime}(\log z) \geq \eta,^{54} \tag{2.27}$$

- *and for all* $z \notin e^{B_{\epsilon^h}(\log z^h_0)}$,

$$u^h\left(e^{\bar{c}}R\left(z\right), z\right) \leq u^h\left(e^{-\bar{c}}c^h_0, z^h_0\right). \tag{2.28}$$

[53] Throughout the paper, we interpret $\mathbb{R}$ as a measure space with respect to the linear Borel σ-algebra and the Lebesgue measure. The only spaces other than (subsets of) $\mathbb{R}$ and $\mathcal{H}$ on which we refer to measurability are products thereof; we interpret all such spaces as σ-algebras with respect to the canonical product σ-algebra.

[54] Here, $e^{B_{\epsilon^h}(\log z^h_0)} \equiv \left\{z \in \mathbb{R}_{>0} \mid \log z \in B_{\epsilon^h}(\log z^h_0)\right\}$. Also note that any such function v^h is unique since consumption utility is strictly increasing. Also $v^h(z)$ is twice continuously differentiable since u^h and R are and $u^h_c > 0$, by the implicit function theorem.

The following Lemma establishes that, indeed, the assumptions stated so far guarantee each household's labor supply—locally to the initial tax schedule—is well-defined and unique, purely intensive, and well-behaved.

Lemma 2. *There exists $\delta > 0$ such that $z^h(\widetilde{R})$ and $c^h(\widetilde{R})$ are well-defined, $\mathcal{H}$-measurable, and strictly positive on $\mathcal{H} \times (R + B_\delta(\mathbf{0}))$, and on this domain have two continuous and $\mathcal{H}$-measurable Frechet derivatives in $\widetilde{R}$.*

Proof. See Appendix 2.9.6.2 □

In order to state the next assumption, we first introduce the concepts of labor supply compensated and income elasticities and "super-elasticities". The compensated (income) elasticity of labor supply for a household describe how labor supply changes due to changes in local marginal (the level of) taxes fixing the local level of (marginal) taxes:

$$\eta^h(\widetilde{R}) = \frac{1}{\widetilde{R}(z^h(\widetilde{R}))} \left. \frac{d \log z^h(\widetilde{R}^h(\cdot;\epsilon_0,\epsilon_1))}{d\epsilon_0} \right|_{\substack{\epsilon_0=0\\ \epsilon_1=0}}, \qquad \varepsilon^h(\widetilde{R}) = \frac{1}{\widetilde{R}'(z^h(\widetilde{R}))} \left. \frac{d \log z^h(\widetilde{R}^h(\cdot;\epsilon_0,\epsilon_1))}{d\epsilon_1} \right|_{\substack{\epsilon_0=0\\ \epsilon_1=0}} \tag{2.29}$$

where $\widetilde{R}^h(z;\epsilon_0,\epsilon_1) = \widetilde{R}(z) + \epsilon_0 + (z - z^h(\widetilde{R}))\epsilon_1$. The super-elasticities—denoted by $\eta^h_{+0}(\widetilde{R})$, $\eta^h_{+1}(\widetilde{R})$, $\varepsilon^h_{+0}(\widetilde{R})$, and $\varepsilon^h_{+1}(\widetilde{R})$—are defined as the *indirect* change in these elasticities caused by the change in the curvature of preferences that (for non-CES preferences) results when labor supply and income respond to tax changes. The super-elasticities denoted with "+0" correspond to changes in elasticities induced by changes in the level of taxes, whereas those denoted with "+1" correspond to changes induced by changes in the slope of taxes.[55]

Assumption 4. *There exists $\delta, M > 0$ such that:*

- *At R, pre- and post-tax income, z^h_0 and c^h_0, are $\mathcal{H}$-integrable.*
- *If $\widetilde{R} \in R + B_\delta(\mathbf{0})$, then for all $h \in \mathcal{H}$*

$$|\eta^h(\widetilde{R})|,\ |\varepsilon^h(\widetilde{R})|,\ |\varepsilon^h(\widetilde{R})^{-1}|,\ |\eta^h_{+1}(\widetilde{R})|,\ |\eta^h_{+0}(\widetilde{R})|,\ |\varepsilon^h_{+1}(\widetilde{R})|,\ |\varepsilon^h_{+0}(\widetilde{R})| \leq\ M. \tag{2.30}$$

The following Lemma establishes that, under these integrability assumptions and elasticity bounds, not only individual but also aggregate labor supply and consumption are defined and well-behaved locally to R.

Lemma 3. *There exists $\delta > 0$ such that $z^h(\widetilde{R})$, $c^h(\widetilde{R})$, and their first two Frechet derivatives in $\widetilde{R}$ are bounded (as linear maps) across all $\widetilde{R} \in R + B_\delta(\mathbf{0})$ by linear combinations of z^h_0 and c^h_0 and, in particular, are $\mathcal{H}$-integrable.*[56]

Proof. See Appendix 2.9.6.3. □

Next, we impose conditions that make it possible to express aggregate tax revenue in terms of an integral over income levels, which allows us to integrate by parts.

Assumption 5. *The distribution of initial pre-tax income, z^h_0, admits a twice-continuously differentiable density g on $\mathbb{R}_{\geq 0}$; and for each of the following elasticity variables x^h_n, a conditional expectation*[57][58] $x(z) \equiv$

[55]We present complete and formal definitions of elasticities and super-elasticities in Appendix 2.9.6.2.

[56]Moreover the first two Frechet derivatives of $\log z^h(\widetilde{R})$ and $\log c^h(\widetilde{R})$ are bounded uniformly across $h \in \mathcal{H}$, $\widetilde{R} \in R + B_\delta(\mathbf{0})$; this fact is useful in later proofs.

[57]We say "a" rather than "the", as conditional expectation is unique only up to measure zero sets.

[58]Each such conditional expectation function $\mathbb{E}[x^h_n|z^h_0]$ exists since z^h_0 is measurable by Assumption 4 and x^h_n is integrable, as it is measurable by arguments in the second-to-last section of the proof of Lemma 2 and bounded by Assumption 4 (then apply dominated convergence).

$\mathbb{E}[x_n^h|z_0^h = z]$ *is n-times continuously differentiable*[59] *on* $\operatorname{supp} g$. *This holds for* $x_2^h = \varepsilon^h(R)$, $\eta^h(R)\varepsilon^h(R)$, ; $x_1^h = \eta^h(R)$, $\varepsilon^h(R)$, $\eta^h(R)^2$, $\varepsilon^h(R)^2$, $\eta^h(R)^2\varepsilon^h(R)$, $\eta^h(R)\varepsilon^h(R)^2$, $\eta_{+1}^h(R)$; *and* $x_0^h = \eta^h(R)^3$, $\varepsilon^h(R)^3$, $\eta_{+0}^h(R)$, $\varepsilon_{+0}^h(R)$, $\varepsilon_{+1}^h(R)$.

Theorem 2, but not Theorem 1, relies on the following, additional regularity condition. It allows us to integrate by parts not only in a local region of the income schedule but over all incomes from zero to infinity.[60]

Assumption 6. *The following, additional regularity conditions hold:*

1. $\lim_{z\to L} z^2 g(z) = 0$ *and* $\lim_{z\to L} zR(z)g(z) = 0$ *for* $L = 0, \infty$.[61]

2. *For each of the following variables* x_n^h *and for any* $\epsilon > 0$, *a conditional expectation*[62] $x_{\leq}(z;\epsilon) \equiv \mathbb{E}[x_n^h(R) \mid z^h(R) = z,\ \varepsilon^h(R) < \epsilon]$ *is n-times continuously differentiable on* $\operatorname{supp} g$. *This holds for* $x_1^h = \eta^h(R)$ *and* $x_0^h = \varepsilon^h(R)$, $\frac{\eta^h(R)^2}{\varepsilon^h(R)}$.

3. *Each of the following are bounded in magnitude across all* $z \in \operatorname{supp} g$:

$$\begin{gathered}\left(\frac{d\log R(z)}{d\log z}\right)^{-1}, \quad \alpha(z), \quad z\alpha'(z), \quad z\eta'(z), \quad z\varepsilon'(z), \quad z(\eta^2)'(z), \\ z(\eta\varepsilon)'(z), \quad z^2(\eta\varepsilon)''(z), \quad z(\varepsilon^2)'(z), \quad z(\eta^2\varepsilon)'(z), \quad z(\eta\varepsilon^2)'(z), \quad z\eta_{+1}'(z), \quad z\eta_{\leq}'(z;\epsilon)\end{gathered} \tag{2.31}$$

2.9.1.3 Social objective definitions

The basic structure imposed by the definition of a standard social objective is enough to guarantee that each individual's contribution to aggregate welfare is locally well-behaved.

Lemma 4. *Suppose that* $((w^h)_{h\in\mathcal{H}}, G)$ *is a standard social objective. Then there exists* $\delta > 0$ *such that* $w^h \circ V^h(\widetilde{R})$ *is finite and* $\mathcal{H}$*-measurable on* $\mathcal{H} \times (R + B_\delta(\mathbf{0}))$, *and on this domain has two continuous and* $\mathcal{H}$*-measurable Frechet derivatives in* $\widetilde{R}$.

Proof. See Appendix 2.9.6.4. □

Toward studying aggregate welfare, we now introduce additional structure to the problem of a planner.

Definition 3. *A standard social objective* $((w^h)_{h\in\mathcal{H}}, G)$ *is* ***regular*** *if the following hold:*

1. *For some* $\delta > 0$ *and integrable functions* $b_0, b_1, b_2 : \mathcal{H} \to \mathbb{R}$ *such that for all* $\widetilde{R} \in R + B_\delta(\mathbf{0})$,

$$\begin{aligned} \left|(w^h \circ u^h)(c^h(\widetilde{R}), z^h(\widetilde{R}))\right| &\leq b_0(h), \\ \left|c^h(\widetilde{R})(w^h \circ u^h)_c(c^h(\widetilde{R}), z^h(\widetilde{R}))\right| &\leq b_1(h), \\ \text{and} \quad \left|c^h(\widetilde{R})^2(w^h \circ u^h)_{cc}(c^h(\widetilde{R}), z^h(\widetilde{R}))\right| &\leq b_2(h). \end{aligned} \tag{2.32}$$

[59] By 0-times continuously differentiable, we mean continuous.

[60] Of these conditions in Assumption 6, we view all but one innocuous. Specifically, in any tax schedule with a finite marginal taxes at 0 and a positive intercept—i.e. $R(0) > 0$—the ratio of average to marginal retention diverges in the limit as $z \to 0$. Since we have already assumed that all households supply positive income, this condition—which note only need hold on the support of h—can be made less objectionable by simply assuming there is a minimum earned income across all households. However, this issue may warrant further attention in extensions of our work to models of labor supply on the extensive margin.

[61] Note that these limits are implied by the existence of aggregate income and tax revenue so long as $g(z)$ and $R(z)$ are not too "squiggly."

[62] Such a conditional expectation function exists since z_0^h is measurable (by Assumption 3); since $\eta^h(R)$, $\varepsilon^h(R)$, and $\frac{\eta^h(R)^2}{\varepsilon^h(R)}$ are measurable (by the second-to-last step of the proof of Lemma 2); since this implies $\mathbb{1}_{\varepsilon^h(R)\leq\epsilon}$ is measurable; and since by Assumption 4 the elasticities are bounded and so integrable.

2. *For each of the following variables x_n^h, a conditional expectation*[63] $\mathbb{E}[R(z_0^h)x_n^h|z_0^h = z]$ *is n-times continuously differentiable on* $\operatorname{supp} g$. *This holds for $x_0^h = \lambda^h(R) \equiv (w^h \circ u^h)_c(c_0^h, z_0^h)$, $\lambda^h(R)\varepsilon^h(R)$, $\lambda^h(R)\frac{\eta^h(R)^2}{\varepsilon^h(R)}$, $(\lambda\gamma)^h(R) \equiv (w^h \circ u^h)_{cc}(c_0^h, z_0^h)c_0^h$; and $x_1^h = \lambda^h(R)\eta^h(R)$.*

For each variable x_n^h referred to in the previous definition, we—for any $z \in \operatorname{supp} g$—denote by $x(z)$ the ratio[64]

$$x(z) \equiv \frac{\mathbb{E}[R(z_0^h)x_n^h|z_0^h = z]}{R(z)}. \tag{2.33}$$

for the remainder of the paper. Note that by Assumption 1, $x(z)$ is n-times continuously differentiable when the corresponding social objective is regular.

Under the additional structure imposed by a social objective's regularity, not only individual contributions to welfare, but also aggregate welfare is defined and well-behaved locally to R.

Lemma 5. *Suppose $((w^h)_{h\in H}, G)$ is a standard, locally regular social objective. Then there exists $\delta > 0$ such that $w^h \circ V^h(\widetilde{R})$ and its first two Frechet derivatives in $\widetilde{R}$ are bounded (as linear maps) across all $\widetilde{R} \in R + B_\delta(\mathbf{0})$ by linear combinations of the functions $b_n(h)$ from Definition 3 and, in particular, are $\mathcal{H}$-integrable.*

Proof. See Appendix 2.9.6.5. □

2.9.1.4 Other supporting concepts

2.9.1.4.1 Super-elasticity concepts

Fix any $\delta > 0$ small enough that Lemma 2 applies. In the proof of Lemma 2, we establish that, at any $\widetilde{R} \in B_\delta(R)$, each household h's compensated elasticity $\varepsilon^h(\widetilde{R})$ satisfies

$$\varepsilon^h(\widetilde{R}) = \widehat{\varepsilon}^h(z^h(\widetilde{R}), \widetilde{R}), \qquad \text{where} \qquad \widehat{\varepsilon}^h(z, \widehat{R}) \equiv \frac{1}{\frac{d\log}{d\log z}M^h(\widehat{R}(z), z) - \frac{d\log \widehat{R}'(z)}{d\log z}} \tag{2.34}$$

where $M^h(c, z) \equiv -\frac{u_z^h(c,z)}{u_c^h(c,z)}$ is h's elasticity of substitution between consumption and leisure. One may use the expression $\widehat{\varepsilon}^h(z, \widehat{R})$ in order to decompose changes in h's elasticity with respect to taxes into two components: First, elasticity changes through the change in the tax schedule at a fixed labor supply. Second, h's elasticity changes through h's change in labor supply at a fixed tax schedule. One may divide the latter changes, in turn, into changes in elasticity stemming from income and compensated effects on labor supply. It is the latter that defines $\varepsilon^+(\widetilde{R})$. Formally,

$$\varepsilon^{+h}(\widetilde{R}) = \left[\frac{d}{d\log z}\widehat{\varepsilon}^h(z, \widetilde{R})\right]\varepsilon^h(\widetilde{R}). \tag{2.35}$$

We show in the proof of Lemma 2 that the changes in elasticity contained in $\varepsilon^{+h}(\widetilde{R})$ reflect two basic channels: The change in elasticity due to changes in the curvature of preferences as taxes change, and the

[63] To see that conditional expectations of $R(z_0^h)x_n^h$ conditional on z_0^h (which is measurable by Assumption 3) exist, we argue that they are measurable and bounded by integrable functions; then dominated convergence implies they are integrable. Each $R(z_0^h)x_n^h$ is measurable since w^h and u^h are (locally) twice-continuously differentiable by the definition of a standard social objective and Assumption 2, since c_0^h and z_0^h are measurable by Asumption 4, and since elasticities are measurable by arguments in the second-to-last section of the proof of Lemma 2. Each $R(z_0^h)x_n^h$ is—by Assumption 4—bounded by a constant times one of the integrable functions $b_n(h)$ from the definition of local regularity.

[64] This is a valid definition since for any $z \in \operatorname{supp} z \not\ni 0$, $R(z) > 0$ (see the proof of Lemma 2).

change in elasticity due to changes in the curvature of the tax schedule as taxes change. That is,

$$\varepsilon^+(z) \equiv \varepsilon_{+1}(z) \ + \ \frac{d^2 \log R(z)}{d \log z^2}(\eta\varepsilon^2)(z) \ + \ \frac{d^2 \log R'(z)}{d \log z^2}\varepsilon^3(z) \tag{2.36}$$

The former are zero when preferences are (locally) additively CES, whereas the latter are zero when the tax schedule is (locally) CES.

2.9.2 Additional Discussion

2.9.2.1 Algebra for motivating example

Below, we briefly walk through the very straightforward algebra behind the example presented in Section 2.2.

Setup:

A unit measure μ of households $h \in \mathcal{H}$ supply labor and consume in a static economy subject to a tax schedule T. The problem of each household h is

$$V^h(T_{\bar{\tau}}) \quad = \quad \max_z \quad z - T(z) \quad - \quad \frac{z^{1+\frac{1}{\beta^h}}}{1+\frac{1}{\beta^h}} \Big/ (\theta^h)^{\frac{1}{\beta^h}} \tag{2.37}$$

We denote by $z^h(T_{\bar{\tau}}$ the maximizer of the household's problem. Conditional on elasticity, productivity is distributed Pareto; $\theta^h \mid \beta^h \sim \text{Pareto}(\alpha > 1)$.

We assume the tax schedule T is convex and initially imposes a constant top rate $\bar{\tau}_0$ on all incomes above some level $\bar{z}$.

Household labor supply

To begin, we characterize the labor supply problem (2.37) of each household h. We break this analysis into two cases. First, consider a household h for whom $\bar{z} \geq \theta^h(1-\bar{\tau}_0)^{\beta^h}$. Whenever the top tax rate is any $\bar{\tau} \geq \bar{\tau}_0$, h must have an income z weakly below $\bar{z}$ since otherwise decreasing z increases utility at a rate

$$-1 + \bar{\tau} + \left(z/\theta^h\right)^{\frac{1}{\beta^h}} \quad \geq \quad -1 + \bar{\tau}_0 + \left(z/\theta^h\right)^{\frac{1}{\beta^h}} \quad > \quad 0. \tag{2.38}$$

Moreover, note that the labor supply of any such household h is unaffected by increases in the top tax rate, since h already prefers some income $z < \bar{z}$.

Second, consider a household h for whom $\bar{z} < \theta^h(1-\bar{\tau}_0)^{\beta^h}$. At any top tax rate $\bar{\tau}$, h must have an income z weakly above $\bar{z}$ since otherwise—by T's convexity—increasing z increases utility at a rate weakly greater than

$$1 - \bar{\tau}_0 - \left(z/\theta^h\right)^{\frac{1}{\beta^h}} \quad > \quad 0. \tag{2.39}$$

An analogous argument implies that h has income z^h strictly above $\bar{z}$ at $\bar{\tau} > \bar{\tau}_0$ if and only if $\bar{z} < \theta^h(1-\bar{\tau})^{\beta^h}$. When this latter inequality holds, the local differentiability of taxes and preferences implies the first order condition:

$$1 - \bar{\tau} = \left(z^h/\theta^h\right)$$

In summary, household labor supply is given as a function of preferences and the top tax rate $\bar{\tau} \geq \bar{\tau}_0$ as

$$z^h(T_{\bar{\tau}}) \quad = \quad \begin{cases} \max\left[\theta^h(1-\bar{\tau})^{\beta^h}, \bar{z}\right] & \text{if} \quad \theta^h(1-\bar{\tau}_0)^{\beta^h} > \bar{z} \\ z^h(T_{\bar{\tau}_0}) \leq \bar{z} & \text{otherwise.} \end{cases} \tag{2.41}$$

Top incomes in each elasticity group:

We now compute, for any top tax rate $\bar{\tau} \geq \bar{\tau}_0$ the total income $Z_{\text{top}}(\tau|\beta)$ earned above $\bar{z}$ by households of elasticity β. We define this amount to be normalized by the size of the group, i.e. the number of households with elasticity β.

$$\begin{aligned} Z_{\text{top}}(\bar{\tau}|\beta) &= \int \max\left[\theta(1-\bar{\tau})^{\beta} - \bar{z}, 0\right] \ \text{density}(\theta|\beta) \ d\theta \\ &= \int \max\left[\theta(1-\bar{\tau})^{\beta} - \bar{z}, 0\right] \ \alpha\theta^{-1-\alpha} \ d\theta \\ &= \alpha \int_{\bar{z}/(1-\bar{\tau})^{\beta}}^{\infty} \left(\theta(1-\bar{\tau})^{\beta} - \bar{z}\right) \ \theta^{-1-\alpha} \ d\theta \\ &= \alpha \left[(1-\bar{\tau})^{\beta} \int_{\bar{z}/(1-\bar{\tau})^{\beta}}^{\infty} \theta^{-\alpha} d\theta - \bar{z} \int_{\bar{z}/(1-\bar{\tau})^{\beta}}^{\infty} \theta^{-1-\alpha} d\theta\right] \\ &= \alpha \left[(1-\bar{\tau})^{\beta} \frac{1}{\alpha-1} \left(\bar{z}/(1-\bar{\tau})^{\beta}\right)^{1-\alpha} - \bar{z}\frac{1}{\alpha} \left(\bar{z}/(1-\bar{\tau})^{\beta}\right)^{-\alpha}\right] \\ &= \alpha \ \bar{z}^{1-\alpha} \ (1-\bar{\tau})^{\alpha\beta} \left[\frac{1}{\alpha-1} - \frac{1}{\alpha}\right] \\ &= \underbrace{\frac{\bar{z}^{1-\alpha}}{\alpha-1}}_{\equiv k} (1-\bar{\tau})^{\alpha\beta} \end{aligned} \tag{2.42}$$

Tax revenue and its derivatives:

Total tax revenue earned in the top bracket is simply the top tax rate times the income earned by each elasticity-group above $\bar{z}$, or

$$\text{Rev}_{\text{top}}(\bar{\tau}) \quad = \quad \mathbb{E}_{\beta}\left[\ \bar{\tau} \ \cdot \ Z_{\text{top}}(\bar{\tau}|\beta)\ \right] \quad = \quad k \cdot \mathbb{E}_{\beta}\left[\ \bar{\tau} \ (1-\bar{\tau})^{\alpha\beta}\ \right], \tag{2.43}$$

where $\mathbb{E}_{\beta}[\cdot]$ an expectation over β values according to their prevalence in the population.

Given the simple functional form for income earned in the top tax bracket, it is easy to compute derivatives of tax revenue as taxes increase:

$$
\begin{aligned}
\text{Rev}'_{\text{top}}(\bar{\tau}) &= \mathbb{E}_\beta\left[\ k\ (1-\bar{\tau})^{\alpha\beta}\ \right] \quad - \quad \bar{\tau}\ \mathbb{E}_\beta\left[\ \alpha\beta\ k\ (1-\bar{\tau})^{\alpha\beta-1}\ \right] \\
&= \mathbb{E}_\beta\left[\ Z_{\text{top}}(\bar{\tau}|\beta)\ \right] \quad - \quad \frac{\bar{\tau}}{1-\bar{\tau}}\ \alpha\ \mathbb{E}_\beta\left[\ \beta\ Z_{\text{top}}(\bar{\tau}|\beta)\ \right] \\
\text{Rev}''_{\text{top}}(\bar{\tau}) &= -\mathbb{E}_\beta\left[\ \alpha\beta\ k\ (1-\bar{\tau})^{\alpha\beta-1}\ \right] \quad - \quad \mathbb{E}_\beta\left[\ \alpha\beta\ k\ (1-\bar{\tau})^{\alpha\beta-1}\ \right] \\
&\quad + \quad \bar{\tau}\ \mathbb{E}_\beta\left[\ \alpha\beta\ (\alpha\beta-1)\ k\ (1-\bar{\tau})^{\alpha\beta-2}\ \right] \\
&= -\frac{2\alpha}{1-\bar{\tau}}\mathbb{E}_\beta\left[\ \beta\ Z_{\text{top}}(\bar{\tau}|\beta)\ \right] \quad - \quad \frac{\bar{\tau}\alpha}{(1-\bar{\tau})^2}\mathbb{E}_\beta\left[\ \beta\ Z_{\text{top}}(\bar{\tau}|\beta)\ \right] \\
&\quad + \quad \frac{\bar{\tau}\alpha^2}{(1-\bar{\tau})^2}\ \mathbb{E}_\beta\left[\ \beta^2\ Z_{\text{top}}(\bar{\tau}|\beta)\ \right] \\
&= \frac{\alpha\mathbb{E}_\beta\left[Z_{\text{top}}(\bar{\tau}|\beta)\right]}{(1-\bar{\tau})^2}\left(-(2-\bar{\tau})\mathbb{E}_{\text{top}}\left[\beta\right] + \tau\alpha\ \mathbb{E}_{\text{top}}\left[\beta^2\right]\right)
\end{aligned}
\tag{2.44}
$$

where $\mathbb{E}_{\text{top}}[\cdot]$ is an expectation over elasticity groups that weights each proportionally to the share of income earned in the top bracket by households with that elasticity.

Welfare and its derivatives:

Finally, we compute the welfare of each top-earner and its derivatives with respect to $\bar{\tau}$.

Plugging in our expression for incomes $z^h(T_{\bar{\tau}})$ into the household utility function, we obtain that for all h with $z^h(T_{\bar{\tau}_0}) > \bar{z}$ and $\bar{\tau}$ near enough to $\bar{\tau}_0$,

$$
\begin{aligned}
V^h(T_{\bar{\tau}}) &\equiv \max_z \quad \underbrace{z - T(z)}_{=\bar{\tau}\bar{z}-T(\bar{z})+(1-\bar{\tau})z} \quad - \quad \frac{z^{1+\frac{1}{\beta^h}}}{1+\frac{1}{\beta^h}} \Big/ (\theta^h)^{\frac{1}{\beta^h}} \\
&= \bar{\tau}\bar{z} \quad - \quad T(\bar{z}) \quad - \quad \frac{\theta^h(1-\bar{\tau})^{1+\beta^h}}{1+\beta^h}.
\end{aligned}
\tag{2.45}
$$

We may therefore differentiate:

$$
\begin{aligned}
\left.\frac{d}{d\bar{\tau}}\right|_{\bar{\tau}=\bar{\tau}_0} V^h(T_{\bar{\tau}}) &= \bar{z} \quad - \quad \theta^h(1-\bar{\tau}_0)^{\beta^h} \quad = \quad -(z^h(T_{\bar{\tau}_0}) - \bar{z}) \\
\left.\frac{d^2}{d\bar{\tau}^2}\right|_{\bar{\tau}=\bar{\tau}_0} V^h(T_{\bar{\tau}}) &= \beta^h\theta^h(1-\bar{\tau}_0)^{\beta^h-1} > 0
\end{aligned}
\tag{2.46}
$$

Of course, the utilities of households with initial incomes below $\bar{z}$ are not affected by increases in $\bar{\tau}$.

We conclude that

$$
\left.\frac{d}{d\bar{\tau}}\right|_{\bar{\tau}=\bar{\tau}_0} \widetilde{\lambda}\cdot W_{\text{top}}(\bar{\tau}) \quad = \quad -\widetilde{\lambda}\cdot\mathbb{E}_\beta[Z_{\text{top}}(\bar{\tau}|\beta)] \qquad \text{and} \qquad \left.\frac{d^2}{d\bar{\tau}^2}\right|_{\bar{\tau}=\bar{\tau}_0} \widetilde{\lambda}\cdot W_{\text{top}}(\bar{\tau}) \geq 0, \tag{2.47}
$$

as we have used in the main text.

2.9.2.2 One-dimensional heterogeneity: "shift and exploit"

Werning (2007) shows that if households differ only along a single dimension θ^h and, for some common concave function v_c and convex function v_z, both increasing, have utility

$$u^h(c,z) \quad = \quad v_c(c) \quad - \quad v_z(z) \;/\; \theta^h, \tag{2.48}$$

then the problem of Pareto planner is globally convex. A similar result holds in our setting: The problem of any planner with a standard objective so long as the weighting functions w^h are concave. In particular, the first-order condition (ABC) implies the second-order condition (DEFG).[65]

A natural question is whether this convexity is an essential feature of one-dimensional settings or a consequence of the function form (2.48). To answer this question, we consider the planner's second-order condition in the *general* one-dimensional case where there is a unique type $h(z)$ who earns each income $z \in \operatorname{supp} g$. Proposition 1 in Appendix 2.9.3.3 leverages the following two special properties of (2.48) in order to show that—for these preferences—(ABC) implies (DEFG):

- The marginal rate of substitution between consumption and leisure is weakly increasing in consumption and decreasing in leisure.[66]

- All of the variation in compensated elasticity across income levels is due to differences in the income and consumption levels at which households' preferences are evaluated, rather than differences in the curvature of their preferences at a given level of income and consumption.

The first property is quite weak and satisfied by the standard functional forms in the literature.[67] By contrast, the second property is a much more "special" feature. The following two examples illustrate how the planner may improve taxes when this knife-edge assumption fails by a significant enough amount.

Example 1. *Taxes are linear, i.e. $R(z) = rz$ for $r \in (0,1)$. A one-dimensional continuum of households h have additive-CES preferences with idiosyncratic elasticities $\beta(\theta^h)$:*

$$u^h(c,z) \quad = \quad c \quad + \quad \frac{z^{1+1/\beta(\theta^h)}}{1+1/\beta(\theta^h)} \Big/ \theta^h \tag{2.49}$$

Finally, suppose $\beta(\theta)$ is very sharply decreasing around some type $\theta^{h(z^*)}$.

Example 2. *Taxes are linear, i.e. $R(z) = rz$ for $r \in (0,1)$. A one-dimensional continuum of households h have preferences:*

$$u^h(c,z) \quad = \log(c) \quad - \quad v\left(z/\theta^h\right) \tag{2.50}$$

for some increasing and concave function v with variable elasticity. It is easy to verify that this implies each household supplies labor $z^h = \theta^h n^*$ for some common n^* and has a (common) compensated elasticity

$$\varepsilon^h(R) = \frac{1}{1 + \frac{d \log v'(n^*)}{d \log n^*}}. \tag{2.51}$$

Finally, suppose $\frac{d \log v'(n)}{d \log n}$ is very sharply decreasing locally to n^*.

[65]See Proposition 1.

[66]This is different than what is commonly referred to as "decreasing marginal rates of substitution", which applies to shifts along an indifference curve. The condition we study here is slightly stronger than the convexity of preferences.

[67]In the cases where it does fail—such as when households have certain non-convex preferences—there seems to be little reason to hope the planner's problem should be convex, anyways.

We now explain why—in either example—any planner can increase the value of her Lagrangian using a narrow variation in taxes around z^*, as in Figure 2.2. To first order, this variation changes the elasticity at each income level. In Example 1, this change occurs because each household adjusts its income, shifting which elasticities are present at each income level. This effect is strongest where the change in taxes is the most steeply increasing (decreasing), where households with much higher (lower) elasticities are drawn in from the left (right). In Example 2, this change occurs because each household adjusts its elasticity. Namely, those whose labor supplies increase (decrease) also experience increased (decreased) elasticities, due to the changing curvature of $v'(\cdot)$. In either example, the effect of the shift in elasticities is that the same tax variation has much better behavioral effects if performed a second time: Elasticities are particularly high where marginal taxes increase the most and particularly low where marginal taxes decrease the most. More succinctly, the planner can *shift* elasticities with a first tax change and then *exploit* this shift with a second. If the shift in elasticities is dramatic enough, this effect causes the two deviations to improve the planner's Lagrangian on net.

2.9.2.3 Simulation exercise

We now describe in detail the simulation exercise alluded to in Section 2.6.

Concretely, each household h has constant compensated elasticities β^h and no income elasticity. Within every income level 80% of households have $\beta^h = 0.01$ and 20% have $\beta^h = 2$. We roughly approximate the distribution of income in our data by assuming that productivity is distributed according to a generalized Pareto distribution with location 0, scale e^9, and shape 1/2 (this implies a Pareto tail of shape 2). Marginal taxes are initially 50% at every income level. The planner places a constant but not common welfare weight on each household, is indifferent between households with the same initial income, and ensures that her first-order condition holds at the initial tax schedule. Under this calibration, the planner's first-order condition holds exactly but the second-order condition fails for incomes above $\approx \$90,000$.

Next, we perturb marginal taxes slightly, lowering them by 1% at every income level. After computing how households respond to this change, we recompute the planner's first-order condition, which no longer holds. We then move taxes slightly in the direction in which the first-order condition fails. This does not simply push taxes back to where they began, as—at high incomes—taxes were initially at a local minimum. We iterate this procedure until it converges.

Figure 2.11 shows how marginal retention evolves between the initial tax schedule and the final schedule to which our procedure converges. Notably, taxes converge to a new, much lower marginal rate at high incomes—where the second-order condition initially failed—but are barely changed at lower incomes—where the second order condition initially held. Intuitively, an increase in marginal retention sorts high-elasticity households into higher income levels, rationalizing further decreases in marginal taxes at the top and so on until eventually taxes are high enough.

2.9.3 Proofs of main results

This section contains proofs of our main results. These proofs focus on the main conceptual steps and relegate many supporting details to Appendix 2.9.6.

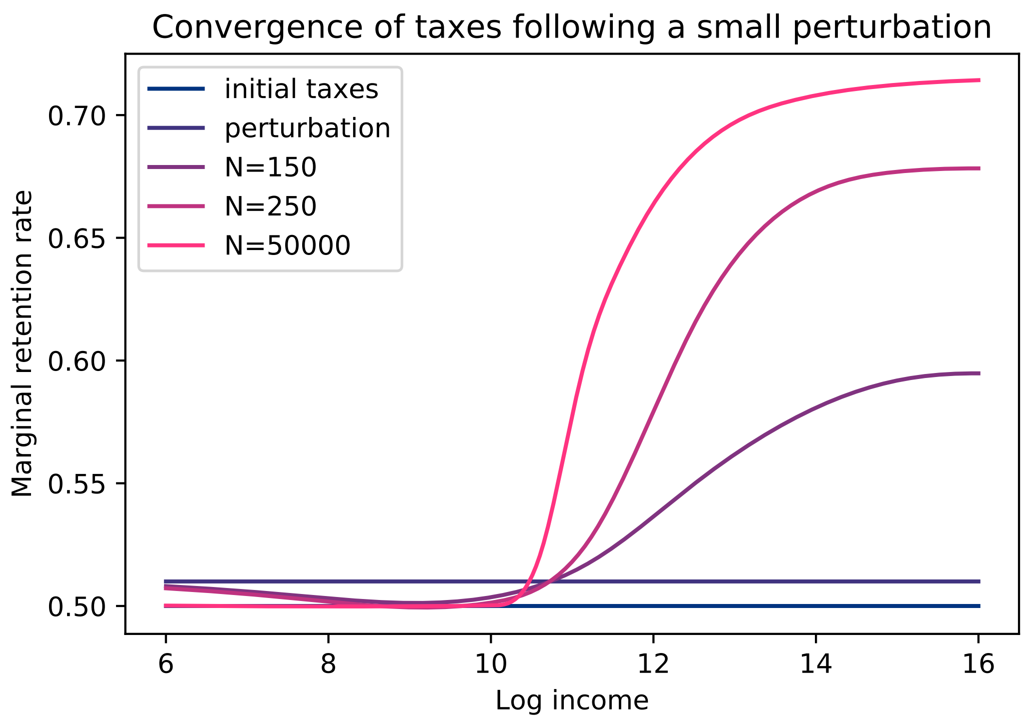


Figure 2.11: Schedule of marginal retention rates following a small perturbation away from initial taxes in the Laffer valley.

2.9.3.1 Proof of Theorem 1

The proof has three main steps. First, Lemma 6 computes the first- and second-order derivatives of aggregate tax revenue. Second, Lemma 7 does the same for aggregate welfare. Third, we use these derivatives to study the planner's first- and second-order necessary conditions for optimality of the tax schedule.

Lemma 6. *Take $\Delta \in \boldsymbol{\Delta}$ and suppose Δ is non-zero only on some interval $[\underline{z}, \overline{z}] \subset \operatorname{supp} g$. Then*

$$\begin{aligned}
D_\Delta \int \left(z^h(R) - R(z^h(R))\right) d\mu &= \int_{\operatorname{supp} g} g(z)\psi(z)\Delta(z)dz \\
D^2_{\Delta\Delta} \int \left(z^h(R) - R(z^h(R))\right) d\mu &= \int_{\operatorname{supp} g} g(z)\left[\Psi_0(z)\left(\frac{\Delta(z)}{R(z)}\right)^2 + \Psi_1(z)\left(\frac{\Delta'(z)}{R'(z)}\right)^2\right] dz \\
\text{where} \quad \psi(z) &\equiv \frac{1-R'(z)}{R'(z)}\left(\frac{d\log R(z)}{d\log z}\eta(z) + \left(\alpha(z) - \frac{d\log}{d\log z}\left(\frac{1-R'(z)}{R'(z)}\right) - \frac{d\log\varepsilon(z)}{d\log z}\right)\varepsilon(z)\right) - 1 \\
\Psi_1(z) &\equiv -z\left(1+R'(z)\right)\varepsilon(z) + z\left(1-R'(z)\right)\left[\left(\alpha(z) - \frac{d\log\varepsilon^2(z)}{d\log z}\right)\varepsilon^2(z) + \varepsilon^+(z)\right]
\end{aligned} \tag{2.52}$$

and $\psi(z)$, $\Psi_0(z)$, and $\Psi_1(z)$ are continuous functions of z on $\operatorname{supp} g$. ($\Psi_0(z)$ is defined in the proof.)

Proof. Let $f^h(\widetilde{R}) \equiv z^h(\widetilde{R}) - \widetilde{R}(z^h(\widetilde{R}))$ denote the tax revenue earned from each household $h \in \mathcal{H}$. In Appendix 2.9.6.5.2 we establish that for some $\delta > 0$, aggregate tax revenue $\int f^h(\widetilde{R})d\mu$ is defined and has two continuous Frechet derivatives at all $\widetilde{R} \in R + B_\delta(\mathbf{0})$. Moreover, these derivatives satisfy

$$D^n \int f^h(\widetilde{R})d\mu = \int D^n f^h(\widetilde{R})d\mu. \tag{2.53}$$

for $n = 1, 2$.

We now proceed to compute these derivatives.

First derivative of tax revenue

Fix any $\Delta \in \mathbf{\Delta}$ satisfying $\Delta(z) = 0$ for all z outside of some interval $[\underline{z}, \overline{z}] \subset \operatorname{supp} g$. To compute the first derivative of tax revenue, we combine (2.53) with the expression (2.126) for $D_\Delta z^h(\widetilde{R})$ in the proof of Lemma 2:[68]

$$\begin{aligned} D_\Delta \int f^h(\widetilde{R}) d\mu &= \int \left[(1 - R'(z_0^h)) z_0^h \left(\eta^h(R) \frac{\Delta(z_0^h)}{R(z_0^h)} + \varepsilon^h(R) \frac{\Delta'(z_0^h)}{R'(z_0^h)} \right) - \Delta(z_0^h) \right] d\mu \\ &= \int\limits_{\operatorname{supp} g} g(z) \left[(1 - R'(z)) z \left(\eta(z) \frac{\Delta(z)}{R(z)} + \varepsilon(z) \frac{\Delta'(z)}{R'(z)} \right) - \Delta(z) \right] dz \end{aligned} \tag{2.54}$$

Finally, we integrate by parts in order to convert the term proportional to $\Delta'(z)$ into a term proportional to $\Delta(z)$. Here, we use that (a) $\Delta(z) = 0$ outside of $[\underline{z}, \overline{z}] \subset \operatorname{supp} g$, (b) by $R'(z)$'s continuity (from Assumption 1) and the argument in the proof of Lemma 2 that $R'(z) > 0$ at all $z > 0$, $R'(z)$ is bounded above zero on $[\underline{z}, \overline{z}]$, (c) by Assumptions 1 and 5 and the definition of $\mathbf{\Delta}$, $g(z) \frac{1 - R'(z)}{R'(z)} \varepsilon(z) z$ and $\Delta(z)$ are continuously differentiable.

$$\begin{aligned} &\int\limits_{\underline{z}}^{\overline{z}} g(z) \frac{1 - R'(z)}{R'(z)} \varepsilon(z) z \Delta'(z) dz \\ = \quad & - \int\limits_{\underline{z}}^{\overline{z}} g(z) \frac{1 - R'(z)}{R'(z)} \varepsilon(z) \left(-\alpha(z) + \frac{d \log}{d \log z} \left(\frac{1 - R'(z)}{R'(z)} \right) + \frac{d \log \varepsilon(z)}{d \log z} \right) \Delta(z) dz. \end{aligned} \tag{2.55}$$

Subtituting this in gives us the expression $\psi(z)$ in the statement of the lemma.

Finally, the continuity of $\psi(z)$ on $\operatorname{supp} g$ follows from Assumptions 1 and 4 and the fact, noted above, that $R'(z) > 0$ for all $z \in \operatorname{supp} g \not\ni 0$.

Second derivative of tax revenue

To begin, note that simplifying the expression (2.140) for $D^2_{\Delta\Delta} \widetilde{R}(z^h(\widetilde{R}))$ in the proof of Lemma 3 and combining it with $D^2_{\Delta\Delta} z^h(R) = z^h[(D_\Delta \log z^h(R))^2 + D^2_{\Delta\Delta} \log z^h(R)]$ gives us

$$\begin{aligned} D^2_{\Delta\Delta} f^h(R) = z^h \Bigg[& \left(1 - R^{h\prime}\right) \left(\left(D_\Delta \log z^h(R)\right)^2 + D^2_{\Delta\Delta} \log z^h(R) \right) - \\ & R^{h\prime} \left(2 \left(D_\Delta \log z^h(R)\right) \frac{\Delta^{h\prime}}{R^{h\prime}} + \frac{d \log R^{h\prime}}{d \log z} \left(D_\Delta \log z^h(R)\right)^2 \right) \Bigg] \end{aligned} \tag{2.56}$$

Substituting in for $D_\Delta \log z^h(R)$ and $D_{\Delta\Delta} \log z^h(R)$ using the expressions (2.126) and (2.134) in the proof of Lemma 2, employing (2.53), and finally changing variables to integrate over income rather than

[68] Appendix 2.9.6.5.3 walks through the measure-theoretic steps used below to move between the first and second line; we use similar steps without explicit reference for the rest of the proofs.

households, we obtain

$$D^2_{\Delta\Delta}\int f^h(R)d\mu = \int_{\text{supp}\, g} g(z)\left[A(z)\left(\frac{\Delta(z)}{R(z)}\right)^2 + B(z)\frac{\Delta(z)}{R(z)}\frac{\Delta'(z)}{R'(z)} + C(z)\left(\frac{\Delta'(z)}{R'(z)}\right)^2 + D(z)\frac{\Delta(z)}{R(z)}\frac{\Delta''(z)z}{R'(z)} + E(z)\frac{\Delta'(z)}{R'(z)}\frac{\Delta''(z)z}{R'(z)}\right]dz,$$

where

$$\begin{aligned}
A(z) &\equiv z\left(1-R'(z)\right)\left[-\eta(z) + \left(1 - 2\frac{d\log R(z)}{d\log z}\eta^2(z)\right) + \frac{d}{d\log z}\left(\frac{d\log R(z)}{d\log z}\right)\eta^3(z)\right. \\
&\qquad \left. + \frac{d}{d\log z}\left(\frac{d\log R'(z)}{d\log z}\right)(\eta^2\varepsilon)(z) + \eta_{+0}(z)\right] - zR'(z)\left[\frac{d\log R'(z)}{d\log z}\eta^2(z)\right] \\
B(z) &\equiv 2z\left(1-R'(z)\right)\left[\frac{d\log R(z)}{d\log z}\eta^2(z) + \left(1 - \frac{d\log R(z)}{d\log z} - \frac{d\log R'(z)}{d\log z}\right)(\eta\varepsilon)(z)\right. \\
&\qquad \left. + \frac{d}{d\log z}\left(\frac{d\log R(z)}{d\log z}\right)(\eta^2\varepsilon)(z) + \frac{d}{d\log z}\left(\frac{d\log R'(z)}{d\log z}\right)(\eta\varepsilon^2)(z) + \eta_{+1}(z)\right] - 2zR'(z)\left[\eta(z) + \frac{d\log R'(z)}{d\log z}(\eta\varepsilon)(z)\right] \\
C(z) &\equiv z\left(1-R'(z)\right)\left[-\varepsilon(z) + 2\frac{d\log R(z)}{d\log z}(\eta\varepsilon)(z) + \left(1 - 2\frac{d\log R'(z)}{d\log z}\right)\varepsilon^2(z)\right. \\
&\qquad \left. + \frac{d}{d\log z}\left(\frac{d\log R(z)}{d\log z}\right)(\eta\varepsilon^2)(z) + \frac{d}{d\log z}\left(\frac{d\log R'(z)}{d\log z}\right)\varepsilon^3(z) + \varepsilon_{+1}(z)\right] - zR'(z)\left[2\varepsilon(z) + \frac{d\log R'(z)}{d\log z}\varepsilon^2(z)\right] \\
D(z) &\equiv 2z\left(1-R'(z)\right)(\eta\varepsilon)(z) \\
E(z) &\equiv 2z\left(1-R'(z)\right)\varepsilon^2(z)
\end{aligned} \tag{2.57}$$

Our assumptions guarantee that $A(z)$ and $C(z)$ are continuous, $B(z)$ and $E(z)$ are continuously differentiable, and $D(z)$ is twice-continuously differentiable in z on $\text{supp}\, g$;[69] and moreover that each additive term of (2.57) is integrable in isolation.[70]

In order to reach the expression in the statement of the Lemma, we integrate by parts:[71]

$$\begin{aligned}
\int_{\text{supp}\, g} g(z)B(z)\frac{\Delta(z)}{R(z)}\frac{\Delta'(z)}{R'(z)}dz &= -\int_{\text{supp}\, g}\frac{d}{dz}\left(\frac{g(z)B(z)}{R(z)R'(z)}\right)\frac{1}{2}\Delta(z)^2dz \\
\int_{\text{supp}\, g} g(z)E(z)\frac{\Delta'(z)}{R'(z)}\frac{\Delta''(z)z}{R'(z)}dz &= -\int_{\text{supp}\, g}\frac{d}{dz}\left(\frac{g(z)zE(z)}{R'(z)^2}\right)\frac{1}{2}\Delta'(z)^2dz \\
\int_{\text{supp}\, g} g(z)D(z)\frac{\Delta(z)}{R(z)}\frac{\Delta''(z)z}{R'(z)}dz &= -\int_{\text{supp}\, g}\frac{d}{dz}\left(\frac{g(z)zD(z)}{R(z)R'(z)}\right)\Delta(z)\Delta'(z)dz - \int_{\text{supp}\, g}\frac{g(z)zD(z)}{R(z)R'(z)}\Delta'(z)^2dz \\
&= \int_{\text{supp}\, g}\frac{d^2}{dz^2}\left(\frac{g(z)zD(z)}{R(z)R'(z)}\right)\frac{1}{2}\Delta(z)^2dz - \int_{\text{supp}\, g}\frac{g(z)zD(z)}{R(z)R'(z)}\Delta'(z)^2dz
\end{aligned} \tag{2.58}$$

[69] This follows from Assumptions 1 and 5 and the facts that (a) as shown in the proof of Lemma 2, $R(z), R'(z) > 0$ at all $z > 0$, and (b) by Assumption 3, $0 \notin \text{supp}\, g$.

[70] Integrability is immediate from the continuity discussed above, the continuity and positivity of $R(z)$ and $R'(z)$, the continuity of $\Delta(z)$ implied by the definition of $\boldsymbol{\Delta}$, and the fact that $\Delta(z)$ is zero outside of $[\underline{z}, \overline{z}]$.

[71] The validity of each integration by parts follows from that (a) since Δ is zero outside of $(\underline{z}, \overline{z})$, we may restrict each integral to that interval (b) $\Delta(z)$ is zero at the endpoints of the interval, and (c) since $[\underline{z}, \overline{z}] \subset \text{supp}\, g$, similar continuity arguments to those above ensure each term of each integrand is continuously differentiable as needed.

We conclude that

$$
\begin{aligned}
D^2_{\Delta\Delta}\int\left(z^h(R)-R(z^h(R))\right)d\mu &= \int\limits_{\operatorname{supp} g} g(z)\left[\Psi_0(z)\left(\frac{\Delta(z)}{R(z)}\right)^2+\Psi_1(z)\left(\frac{\Delta'(z)}{R'(z)}\right)^2\right]dz \\
\text{where}\quad \Psi_0(z) &\equiv A(z)-\frac{R(z)^2}{2g(z)}\frac{d}{dz}\left(\frac{g(z)B(z)}{R(z)R'(z)}\right)+\frac{R(z)^2}{2g(z)}\frac{d^2}{dz^2}\left(\frac{g(z)zD(z)}{R(z)R'(z)}\right) \\
\Psi_1(z) &\equiv C(z)-\frac{d\log R(z)}{d\log z}D(z)-\frac{R'(z)^2}{2g(z)}\frac{d}{dz}\left(\frac{g(z)zE(z)}{R'(z)^2}\right)
\end{aligned}
\tag{2.59}
$$

The continuity of Ψ_0 and Ψ_1 on $\operatorname{supp} g$ follows from our earlier observations about $A(z),...,E(z)$, Assumptions 1 and 5, and the fact that $R(z),R'(z)>0$ on $\operatorname{supp} g$ (see the proof of Lemma 2.

To complete the proof, it remains to simplify the expression for $\Psi_1(z)$. Here, the main step is to compute the $E(z)$ term. Since by definition $g(z)\neq 0$ for $z\in\operatorname{supp} g$, we have

$$
\begin{aligned}
-\frac{R'(z)^2}{2}\frac{d}{dz}\left[\frac{g(z)zE(z)}{R'(z)^2}\right] &= -g(z)\frac{1}{2}\left[-\alpha(z)E(z)+zE'(z)-2\frac{d\log R'(z)}{d\log z}E(z)\right] \\
&= g(z)\left[z(1-R'(z))\varepsilon^2(z)\left(\alpha(z)+2\frac{d\log R'(z)}{d\log z}-1-\frac{d\log\varepsilon^2(z)}{d\log z}\right)+zR'(z)\frac{d\log R'(z)}{d\log z}\varepsilon^2(z)\right]
\end{aligned}
\tag{2.60}
$$

Substituting this expression, the definitions of $C(z)$ and $D(z)$, and the definition $\varepsilon^+(z)\equiv\frac{d\left(\frac{d\log R(z)}{d\log z}\right)}{d\log z}(\eta\varepsilon^2)(z)+\frac{d\left(\frac{d\log R'(z)}{d\log z}\right)}{d\log z}\varepsilon^3(z)+\varepsilon_{+1}(z)$ (see Appendix 2.9.1.4.1) into the definition of $\Psi_1(z)$ and cancelling terms gives the expression in the statement of the lemma.

□

Lemma 7. *Let $((w^h)_{h\in\mathcal{H}},G)$ be a standard, regular social objective. Take $\Delta\in\boldsymbol{\Delta}$ and suppose Δ is non-zero only on some interval $[\underline{z},\overline{z}]\subset\operatorname{supp} g$. Then*

$$
\begin{aligned}
D_\Delta\int w^h\circ V^h(R)d\mu &= \int\limits_{\operatorname{supp} g} g(z)\lambda(z)\Delta(z)dz, \\
D^2_{\Delta\Delta}\int w^h\circ V^h(R)d\mu &= \int\limits_{\operatorname{supp} g} g(z)\left[\Phi_0(z)\left(\frac{\Delta(z)}{R(z)}\right)^2+\Phi_1(z)\left(\frac{\Delta'(z)}{R'(z)}\right)^2\right]dz \\
\textit{where}\quad \Phi_0(z) &\equiv R(z)\left[(\lambda\gamma)(z)+\frac{d\log R(z)}{d\log z}\left(\lambda\frac{\eta^2}{\varepsilon}\right)(z)+\frac{1}{2}\left(\alpha(z)+\frac{d\log R(z)}{d\log z}\right)(\lambda\eta)(z)-\frac{1}{2}z(\lambda\eta)'(z)\right] \\
\Phi_1(z) &\equiv R(z)\frac{d\log R(z)}{d\log z}(\lambda\varepsilon)(z)
\end{aligned}
\tag{2.61}
$$

where $\Phi_0(z)$ and $\Phi_1(z)$ are continuous functions of z on $\operatorname{supp} g$.

Proof. Let $f^h(\widetilde{R})\equiv w^h\circ V(\widetilde{R})$. In Appendix 2.9.6.5.2 we establish that for some $\delta>0$, aggregate welfare $\int f^h(\widetilde{R})d\mu$ is defined and has two continuous Frechet derivatives at all $\widetilde{R}\in R+B_\delta(\mathbf{0})$. Moreover, these derivatives satisfy

$$
D^n\int f^h(\widetilde{R})d\mu = \int D^n f^h(\widetilde{R})d\mu. \tag{2.62}
$$

for $n=1,2$.

First derivative of welfare

Fix any $\Delta\in\boldsymbol{\Delta}$ satisfying $\Delta(z)=0$ for all z outside of some interval $[\underline{z},\overline{z}]\subset\operatorname{supp} g$. To compute the first derivative of tax revenue, we combine (2.62) with the expression (2.141) for $D_\Delta w^h\circ V^h(\widetilde{R})$ in the proof of

Lemma 5:

$$D_\Delta \int f^h(R) d\mu \;=\; \int R(z_0^h) \underbrace{(w^h \circ u^h)_c \left(c_0^h, z_0^h\right)}_{=\lambda^h(R)} \frac{\Delta(z_0^h)}{R(z_0^h)} d\mu \;=\; \int\limits_{\operatorname{supp} g} g(z)\lambda(z)\Delta(z) dz \tag{2.63}$$

where $\lambda(z)$ is the average marginal value of transfers to z-earners relative to tax revenue, as defined in and below Definition 3.

Second derivative of welfare

To begin, we simplify the expression (2.142) for $D^2_{\Delta\Delta} w^h \circ V^h(R)$ in the proof of Lemma 5:

$$\begin{aligned} D^2_{\Delta\Delta} f^h(R) = &(w^h \circ u^h)_{cc} \left(c_0^h, z_0^h\right) (c_0^h)^2 \left(\frac{\Delta(z_0^h)}{R(z_0^h)}\right)^2 \\ &+ (w^h \circ u^h)_c \left(c_0^h, z_0^h\right) c_0^h \frac{d \log R(z_0^h)}{d \log z} \frac{1}{\varepsilon^h(R)} \left(D_\Delta \log z^h(R)\right)^2 . \end{aligned} \tag{2.64}$$

Substituting in for $D_\Delta \log z^h(R)$ using the expression (2.126) in the proof of Lemma 2, employing (2.62), and finally changing variables to integrate over income rather than households, we obtain

$$\begin{aligned} D^2_{\Delta\Delta} \int f^h(R) d\mu &= \int\limits_{\operatorname{supp} g} g(z) \left[A(z) \left(\frac{\Delta(z)}{R(z)}\right)^2 + B(z) \frac{\Delta(z)}{R(z)} \frac{\Delta'(z)}{R'(z)} + C(z) \left(\frac{\Delta'(z)}{R'(z)}\right)^2\right] dz. \\ \text{where} \quad A(z) &\equiv R(z) \left((\lambda\gamma)(z) + \frac{d \log R(z)}{d \log z} \left(\lambda \frac{\eta^2}{\varepsilon}\right)(z)\right) \\ B(z) &\equiv R(z)(\lambda\eta)(z) \frac{d \log R(z)}{d \log z} \\ C(z) &\equiv R(z) \frac{d \log R(z)}{d \log z} (\lambda\varepsilon)(z) \end{aligned} \tag{2.65}$$

where the various terms $(\lambda x)(z)$ are as defined in and below Definition 3. Our assumptions guarantee that $A(z)$ and $C(z)$ are continuous and $B(z)$ is continuously differentiable in z on $\operatorname{supp} g$;[72] and moreover that each additive term of (2.65) is integrable in isolation.[73]

In order to reach the expression in the statement of the Lemma, we integrate by parts:[74]

$$\int\limits_{\operatorname{supp} g} g(z) B(z) \frac{\Delta(z)}{R(z)} \frac{\Delta'(z)}{R'(z)} dz \quad = \quad - \int\limits_{\operatorname{supp} g} \frac{d}{dz} \left(\frac{g(z)B(z)}{R(z)R'(z)}\right) \frac{1}{2} \Delta(z)^2 dz \tag{2.66}$$

We conclude that

$$\begin{aligned} D^2_{\Delta\Delta} \int \left(z^h(R) - R(z^h(R))\right) d\mu \quad &= \quad \int\limits_{\operatorname{supp} g} g(z) \left[\Phi_0(z) \left(\frac{\Delta(z)}{R(z)}\right)^2 + \Phi_1(z) \left(\frac{\Delta'(z)}{R'(z)}\right)^2\right] dz \\ \text{where} \qquad \Phi_0(z) \equiv A(z) - \frac{R(z)^2}{2g(z)} \frac{d}{dz} &\left(\frac{g(z)B(z)}{R(z)R'(z)}\right) \qquad \text{and} \qquad \Phi_1(z) \equiv C(z) \end{aligned} \tag{2.67}$$

The continuity of Φ_0 and Φ_1 on $\operatorname{supp} g$ follows from our earlier observations about $A(z), B(z), C(z)$, Assumptions 1 and 5, Definition 3, and the fact that $R(z), R'(z) > 0$ on $\operatorname{supp} g$ (see the proof of Lemma 2. To

[72] This follows from Assumptions 1 and 5, Definition 3 and the facts that (a) as shown in the proof of Lemma 2, $R(z), R'(z) > 0$ at all $z > 0$, and (b) by Assumption 3, $0 \notin \operatorname{supp} g$.

[73] Integrability is immediate from the continuity discussed above, the continuity and positivity of $R(z)$ and $R'(z)$, the continuity of $\Delta(z)$ implied by the definition of $\boldsymbol{\Delta}$, and the fact that $\Delta(z)$ is zero outside of $[\underline{z}, \overline{z}]$.

[74] The validity of integration by parts follows from that (a) since Δ is zero outside of $(\underline{z}, \overline{z})$, we may restrict each integral to that interval (b) $\Delta(z)$ is zero at the endpoints of the interval, and (c) since $[\underline{z}, \overline{z}] \subset \operatorname{supp} g$, similar continuity arguments to those above ensure each term of each integrand is continuously differentiable as needed.

complete the proof, it remains to simplify the expression for $\Psi_0(z)$ by expanding the $B(z)$ term. Since, by definition $g(z) \neq 0$ for $z \in \operatorname{supp} g$, we have

$$
\begin{aligned}
-\frac{R(z)^2}{2g(z)}\frac{d}{dz}\left(\frac{g(z)B(z)}{R(z)R'(z)}\right) &= -\frac{1}{g(z)}\frac{R(z)^2}{2z}\frac{d}{d\log z}\left[\frac{g(z)z(\lambda\eta)(z)}{R(z)}\right] \\
&= -\frac{1}{g(z)}\frac{R(z)^2}{2z}\left[-g(z)z\alpha(z)\frac{(\lambda\eta)(z)}{R(z)} + g(z)z\frac{z(\lambda\eta)'(z)}{R(z)} - g(z)\frac{z(\lambda\eta)(z)}{R(z)}\frac{d\log R(z)}{d\log z}\right] \\
&= \frac{R(z)}{2}\left[\left(\alpha(z) + \frac{d\log R(z)}{d\log z}\right)(\lambda\eta)(z) - z(\lambda\eta)'(z)\right]
\end{aligned}
\tag{2.68}
$$

Simplifying this expression and substituting it into the definition of $\Phi_0(z)$ gives the expression in the statement of the Lemma. □

We now turn to the third step of the proof of Theorem 1. Here, the main idea is to use the derivatives computed in Lemmas 6 and 7 to study the first- and second-order necessary conditions that must hold for the planner who rationalizes R.

More concretely, we may—since by assumption R is locally rationalized by a standard, regular social objective—take $((w^h)_{h\in\mathcal{H}}, G)$ to be such a social objective. Since R is locally rationalized by the social objective, $\mathbf{0}$ must solve:

$$
\begin{aligned}
&\mathbf{0} \in \underset{\Delta\in\boldsymbol{\Delta}}{\arg\max}\ F(\Delta) \quad \text{s.t.} \quad H(\Delta) \in \mathbb{R}_{\geq 0} \\
&\text{where } F(\Delta) \equiv \begin{cases} \int w^h \circ u^h\left((R+\Delta)(z^h(R+\Delta)), z^h(R+\Delta)\right) d\mu, & \text{if } \Delta \in B_\delta(\mathbf{0}) \\ F(\mathbf{0}) - 1 & \text{if } \Delta \notin B_\delta(\mathbf{0}) \end{cases} \\
&H(\Delta) \equiv \begin{cases} \int \left[z^h(R+\Delta) - (R+\Delta)(z^h(R+\Delta))\right] d\mu - G, & \text{if } \Delta \in B_\delta(\mathbf{0}) \\ 0, & \text{if } \Delta \notin B_\delta(\mathbf{0}) \end{cases}
\end{aligned}
\tag{2.69}
$$

where $\delta > 0$ is small enough that F and H are well-defined and within $B_\delta(\mathbf{0})$ have well-defined and continuous first and second Frechet derivatives (see Appendix 2.9.6.5.2).

In Appendix 2.9.6.5.4 we show that the optimization problem (2.69) satisfies the conditions required to apply standard results from optimization theory on Banach spaces. In particular, the fact that $\mathbf{0}$ solves (2.69) implies: If $DH(\mathbf{0}) \neq 0$, then there exists $\kappa \in \mathbb{R}_{\geq 0}$ such that

- A first-order condition holds: $DF(\mathbf{0}) + \kappa DH(\mathbf{0}) = 0$
- A second-order condition holds: for all non-zero $\Delta \in \boldsymbol{\Delta}$ satisfying $D_\Delta H(\mathbf{0}) = 0$, $D^2_{\Delta,\Delta}F(\mathbf{0}) + \kappa D^2_{\Delta,\Delta}H(\mathbf{0}) \leq 0$.

Since the case where $DH(\mathbf{0}) = 0$—i.e. tax revenue is, to first-order, invariant to tax changes—is unlikely to apply in practice, we relegate it to Appendix 2.9.6.5.5. We show that in this case (ABC) holds with equality for all $z \in \operatorname{supp} g$, so the theorem holds. The remainder of the proof focuses on the complementary case where $DH(\mathbf{0}) \neq 0$.

First-order condition

Recall that there exists $\kappa \geq 0$ for which $DF(\mathbf{0}) + \kappa DH(\mathbf{0}) = 0$. As we have assumed $DH(\mathbf{0}) \neq 0$, this implies there exists $\Delta \in \boldsymbol{\Delta}$ with either $D_\Delta F(\mathbf{0}) \neq 0$, so we must have $\kappa > 0$. Since, by the definition of a standard social objective, $\lambda(z) \geq 0$ for all z, Lemma 6 implies $D_\Delta F(\mathbf{0}) \geq 0$ for all $\Delta \in \boldsymbol{\Delta}$ satisfying $\Delta(z) \geq 0$ for all $z \in \operatorname{supp} g$.

Putting together these observations with the expression for the first derivative of revenue in Lemma 6, we have that for all $\Delta \in \boldsymbol{\Delta}$ satisfying $\Delta(z) \geq 0$ for all $z \in \operatorname{supp} g$,

$$D_{\Delta} H(\mathbf{0}) \quad = \quad \int_{z \in \operatorname{supp} g} g(z)\psi(z)\Delta(z)dz \quad \geq \quad 0, \tag{2.70}$$

where $\psi(z)$ is as in Lemma 6 and recall $\psi(z)$ is continuous on $\operatorname{supp} g$.

We conclude that $\psi(z) \geq 0$ for all $z \in \operatorname{supp} g$, i.e. the first part of the theorem holds. Otherwise, the continuity of $g(z)$ (from Assumption 5 and $\psi(z)$ imply there exists an interval $[\underline{z}, \overline{z} > \underline{z}] \subset \operatorname{supp} g$ so that i.e. $g(z)\psi(z) < 0$ at all $z \in [\underline{z}, \overline{z}]$. The result then follows from considering any weakly positive function $\Delta(z)$ that is strictly positive on a non-zero-measure sub-interval of $[\underline{z}, \overline{z}]$, zero outside of $[\underline{z}, \overline{z}]$, and is contained in $\boldsymbol{\Delta}$.[75]

Second-order condition

We now argue that the second-order condition stated above implies $\Psi_1(z) \leq 0$ for all $z \in \operatorname{supp} g$—where Ψ_1 is as defined in Lemma 6—as claimed in the statement of the theorem. It suffices to show $\Phi_1(z) + \kappa\Psi_1(z) \leq 0$ for all $z \in \operatorname{supp} g$—where Ψ_0 is as defined in Lemma 7—since $\kappa > 0$ and since $\Phi_1(z) \geq 0$ because (a) $R(z), R'(z) > 0$ (see the proof of 2) and (b) by Assumptions 2, the definition of a standard social objective, and the observation in the proof of 2 that $\varepsilon^h(R) > 0$, we have $\varepsilon^h(R), \lambda^h(R) \geq 0$ for all $h \in \mathcal{H}$.

To see this, suppose not, i.e. $\Phi_1(z^*) + \kappa\Psi_1(z^*) > 0$ at some $z^* \in \operatorname{supp} g$. By the continuity established in Lemma 6 and 7 as well as that of $g(z)$ from Assumption 5, there then exists an interval $[\underline{z}, \overline{z} > \underline{z}] \subset \operatorname{supp} g$ on which $g(z)(\Phi_1(z) + \kappa\Psi_1(z))$ is bounded above zero and $g(z)(\Phi_0(z) + \kappa\Psi_0(z))$ and $g(z)\psi(z)$ are bounded. Letting $k > 0$ be some number for which $\min_{z \in [\underline{z}, \overline{z}]} g(z)(\Phi_0(z) + \kappa\Psi_0(z)) > -k \min_{z \in [\underline{z}, \overline{z}]} g(z)\left(\Phi_1(z) + \kappa\Psi_1(z)\right)$, we now consider a tax changes of the form $\Delta(\cdot; \hat{z}, r, k)$ that is in $\boldsymbol{\Delta}$, is zero outside of the interval $B_r(\hat{z})$, and satisfies $\int_{\hat{z}-r}^{\hat{z}+r} \Delta'(z; \hat{z}, r, k)^2 dz > k \int_{\hat{z}-r}^{\hat{z}+r} \Delta(z; \hat{z}, r, k)^2 dz$. Here the idea is to take $\Delta(\cdot; \hat{z}, r, k)$ to be a sufficiently narrow "bump" centered at z; we provide an explicit example in Appendix 2.9.6.5.7.

To complete the proof, consider $\Delta \in \boldsymbol{\Delta}$ defined by $\Delta(z) \equiv \alpha_1\Delta_-(z) + \alpha_2\Delta_+(z)$ for $\Delta_-(z) \equiv \Delta(\cdot; \frac{3}{4}\underline{z} + \frac{1}{4}\overline{z}, \frac{1}{4}(\overline{z} - \underline{z}), k)$, $\Delta_+(z) \equiv \Delta(\cdot; \frac{1}{4}\underline{z} + \frac{3}{4}\overline{z}, \frac{1}{4}(\overline{z} - \underline{z}), k)$, and some constants $\alpha_-, \alpha_+ \in \mathbb{R}$.[76] By choosing α_- and α_+ appropriately—and without setting both to zero—we can use ensure that

$$D_{\Delta} H(\mathbf{0}) \quad = \quad \alpha_- \int_{\operatorname{supp} g} g(z)\psi(z)\Delta_-(z)dz \quad + \quad \alpha_+ \int_{\operatorname{supp} g} g(z)\psi(z)\Delta_+(z)dz \quad = \quad 0. \tag{2.71}$$

The planner's second order condition then implies $D^2_{\Delta\Delta} F(\mathbf{0}) + \kappa D^2_{\Delta\Delta} H(\mathbf{0}) \leq 0$. However, applying Lemmas

[75] We give an example in Appendix 2.9.6.5.6.

[76] Any such tax change is in $\boldsymbol{\Delta}$ since by Lemma 1 is a Banach space.

6 and 7 to our construction implies the opposite, a contradiction:

$$
\begin{aligned}
&D^2_{\Delta\Delta}F(\mathbf{0}) + \kappa D^2_{\Delta\Delta}H(\mathbf{0}) \\
&= \alpha_-^2 \int\limits_{\text{supp } g} g(z)\left[(\Phi_0(z)+\kappa\Psi_0(z))\Delta_-(z)^2 + (\Phi_1(z)+\kappa\Psi_1(z))\Delta'_-(z)^2\right]dz \\
&+ \alpha_+^2 \int\limits_{\text{supp } g} g(z)\left[(\Phi_0(z)+\kappa\Psi_0(z))\Delta_+(z)^2 + (\Phi_1(z)+\kappa\Psi_1(z))\Delta'_+(z)^2\right]dz \\
&\geq \alpha_-^2 \underbrace{\left(\min_{z\in[\underline{z},\overline{z}]} g(z)(\Phi_0(z)+\kappa\Psi_0(z))\right)}_{>-k\cdot\min_{z\in[\underline{z},\overline{z}]} g(z)(\Phi_1(z)+\kappa\Psi_1(z))} \underbrace{\int\limits_{\text{supp } g} \Delta_-(z)^2 dz}_{<\frac{1}{k}\int_{\text{supp } g}\Delta'_-(z)^2 dz} + \left(\min_{z\in[\underline{z},\overline{z}]} g(z)(\Phi_1(z)+\kappa\Psi_1(z))\right) \int\limits_{\text{supp } g} \Delta'_-(z)^2 dz \\
&+ \alpha_+^2 \underbrace{\left(\min_{z\in[\underline{z},\overline{z}]} g(z)(\Phi_0(z)+\kappa\Psi_0(z))\right)}_{>-k\cdot\min_{z\in[\underline{z},\overline{z}]} g(z)(\Phi_1(z)+\kappa\Psi_1(z))} \underbrace{\int\limits_{\text{supp } g} \Delta_+(z)^2 dz}_{<\frac{1}{k}\int_{\text{supp } g}\Delta'_+(z)^2 dz} + \left(\min_{z\in[\underline{z},\overline{z}]} g(z)(\Phi_1(z)+\kappa\Psi_1(z))\right) \int\limits_{\text{supp } g} \Delta'_+(z)^2 dz \\
&> \quad 0.
\end{aligned}
\tag{2.72}
$$

Above, the first equality uses that $\Delta_-(z)$ and $\Delta_+(z)$ are never non-zero at any common z. The final inequality uses that α_- or α_+ is non-zero and that $\int_{\text{supp } g}\Delta'_-(z)^2, \int_{\text{supp } g}\Delta +_- (z)^2 > 0$.

Thus, by computing the first and second derivatives of revenues and welfare, combining those expressions, and applying them to "narrow" deviations, we have shown that equations (ABC) and (DEFG) are necessary conditions for the optimality of a tax schedule.

2.9.3.2 Proof of Theorem 2

The proof has three main parts. The first main part is to infer from the tax schedule a profile of marginal welfare weights that rationalize its first order condition and a profile of marginal welfare weight curvatures consistent with its second condition. The second main part is to construct a standard, regular social objective that generates the desired marginal welfare weights and curvatures and to demonstrate its various properties. The third main part studies the planner's Lagrangian given the social objective of the previous part, and in particular makes a Lagrangian sufficiency argument for local optimality.

Part 1: Inferring / assigning welfare weights and curvatures

Implied marginal welfare weights

Our first step in constructing a welfare function will be to infer an implied average marginal welfare weight at each income level and then assign it to particular households within the income level. In a later step of the proof, we will connect these inferred marginal welfare weights to an exact welfare function.

To begin, define $\widehat{\lambda}(\cdot) : \mathbb{R}_{\geq 0} \to \mathbb{R}$ and $\widehat{\lambda}^{\cdot}$ by

$$
\widehat{\lambda}(z) = \begin{cases} \Pi_{ABC}(z), & \text{if } z \in \operatorname{supp} g \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad \widehat{\lambda}^h = \begin{cases} \widehat{\lambda}(z_0^h)\frac{\mathbb{1}_{\varepsilon^h(R)\leq\epsilon}}{p_{\leq}(z_0^h;\epsilon)}, & \text{if } z \in \operatorname{supp} g \\ 0, & \text{otherwise} \end{cases} \tag{2.73}
$$

where $\epsilon > 0$ is any value $< \epsilon = \frac{1}{B^R}\min\left[\frac{\overline{b}_c}{\overline{\lambda}_c}, \frac{\overline{b}_z}{\overline{\lambda}_z}\right]$, where $\overline{b}_c > 0$ and $\overline{b}_z > 0$ are the constants in the statement of the proof,[77] and where $p_{\leq}(z_0^h;\epsilon) \equiv \mathbb{P}[\varepsilon^h(R) \leq \epsilon | z_0^h = z]$ is as defined Footnote 26 in the statement of the Theorem.

By the definition of $\Pi_{ABC}(z)$ and the assumptions in the statement of the theorem, we have $\widehat{\lambda}(z) > 0$ for all $z \in \operatorname{supp} g$. In Appendices 2.9.6.6.1 and 2.9.6.6.2, we establish several other properties of $\widehat{\lambda}(z)$ and $\widehat{\lambda}^h$ related to measurability, continuity, integrability, and conditional expectations. For the purpose of the main argument of the proof, only two are essential: We show that (a) $\widehat{\lambda}(z)$ is continuously differentiable on $\operatorname{supp} g$ and (b) letting $\eta_{\leq}(z;\epsilon) \equiv \mathbb{E}[\eta^h(R) \mid z^h(R) = z,\ \varepsilon^h(R) < \epsilon]$, $\varepsilon_{\leq}(z;\epsilon) \equiv \mathbb{E}[\varepsilon^h(R) \mid z^h(R) = z,\ \varepsilon^h(R) < \epsilon]$, and $(\frac{\eta^2}{\varepsilon})_{\leq}(z;\epsilon) \equiv \mathbb{E}[\frac{\eta^h(R)^2}{\varepsilon^h(R)} \mid z^h(R) = z,\ \varepsilon^h(R) < \epsilon]$ be defined as in Assumption 6, we have that for $x^h = 1, \eta^h(R), \varepsilon^h(R)$, and $\frac{\eta^h(R)^2}{\varepsilon^h(R)}$, $R(z_0^h)\widehat{\lambda}(z_0^h)x_{\leq}(z_0^h;\epsilon)$ is a conditional expectation for $R(z_0^h)\widehat{\lambda}^h x^h$ conditional on z_0^h.

Implied sufficient welfare curvature

We now proceed to use $\hat{\lambda}^h$ and our observations from the previous step of the proof in order define a convenient curvature-of-welfare variable $\hat{\gamma}^h$.

Recalling $\Psi_0(z)$ from Lemma 6, we begin by defining, for all $z \in \mathbb{R}_{\geq 0}$,

$$(\widehat{\lambda\gamma})(z) = \begin{cases} -1 - \frac{z}{R(z)} - \frac{\Psi_0(z)}{R(z)} - \frac{d\log R(z)}{d\log z}\widehat{\lambda}(z)\left(\frac{\eta^2}{\varepsilon}\right)_{\leq}(z;\epsilon) \\ -\frac{1}{2}\left(\alpha(z) + \frac{d\log R(z)}{d\log z}\right)\widehat{\lambda}(z)\eta_{\leq}(z;\epsilon) + \frac{1}{2}\frac{d}{d\log z}\left(\widehat{\lambda}(z)\eta_{\leq}(z;\epsilon)\right), & \text{if } z \in \operatorname{supp} g \\ 0, & \text{otherwise.} \end{cases} \tag{2.74}$$

Next, we define a household-level version of this curvature, constructed so as to (a) aggregate up to the income conditional mean $\widehat{\gamma}(z)$ and (b) allocation all weight to households with low elasticities, as with $\widehat{\lambda}^h$.

$$\widehat{\gamma}^h = \begin{cases} \frac{(\widehat{\lambda\gamma})(z_0^h)}{\widehat{\lambda}^h p_{\leq}(z_0^h;\epsilon)}, & \text{if } z \in \operatorname{supp} g \text{ and } \varepsilon^h(R) \leq \epsilon \\ 0, & \text{otherwise.} \end{cases} \tag{2.75}$$

In Appendix 2.9.6.6.3, we establish several properties of $(\widehat{\lambda\gamma})(z)$ and $\widehat{\gamma}^h$ related to measurability, continuity, integrability, and conditional expectations. For the purpose of the main argument of the proof, only one is essential: We show that $R(z_0^h)(\widehat{\lambda\gamma})(z_0^h)$ is a conditional expectation for $R(z_0^h)\widehat{\lambda}^h\widehat{\gamma}^h$ given z_0^h.

Part 2: Definition and properties of a social objective

At this point, we are finally ready to define a social objective using our inferred marginal welfare weights and curvatures $\widehat{\lambda}^h$ and $\widehat{\gamma}^h$. To do so, take $\delta > 0$ small enough that Assumption 4 applies, $z^h(\widetilde{R})$ and $c^h(\widetilde{R})$ exist and have two continuous and integrable Frechet derivatives on $\widetilde{R} \in R + B_\delta(\mathbf{0})$ (see Lemmas 2 and 3), and $\hat{c}^h(u) \equiv u^h(\cdot, z_0^h)^{-1}(u)$ is defined for all $u \in V^h(R + B_\delta(\mathbf{0}))$ (see Appendix 2.9.6.6.4). Then define

[77] Our motivation for this particular choice of ϵ will become evident in the final step of the proof.

$((w^h)_{h\in\mathcal{H}}, G)$ by:

$$w^h : \operatorname{Im}(u^h) \to \mathbb{R} \cup \{-\infty\}$$
$$u \mapsto \begin{cases} \widehat{\lambda}^h \displaystyle\int_{[V_0^h, u]} \frac{e^{\Phi\left(\widehat{\gamma}^h\left(\log \hat{c}^h(\tilde{u}) - \log c_0^h\right)\right)}}{u_c^h(\hat{c}^h(\tilde{u}), z_0^h)} d\tilde{u}, & \text{if } u \in V^h(R + B_\delta(\mathbf{0})) \\ 0, & \text{else} \end{cases} \tag{2.76}$$

$$G = \int \left(z_0^h - c_0^h\right) d\mu$$

where $\Phi(\cdot) \equiv \sqrt{2\pi}\left(\Phi_0(\cdot) - \frac{1}{2}\right)$ and Φ_0 is the standard normal CDF.[78]

In Appendix 2.9.6.6.5, we establish two important technical properties of this social objective: We show that w^h is well-defined (the fact that G is well-defined is immediate from Assumption 4) and $((w^h)_{h\in\mathcal{H}}, G)$ is a *standard, regular social objective* in the sense of Definitions 1 and 3.

The key to w^h's construction is that it ensures that, for all $h \in \mathcal{H}$ with $z_0^h \in \operatorname{supp} g$,

$$\lambda^h(R) = (w^h \circ u^h)_c(c_0^h, z_0^h) = \widehat{\lambda}^h \quad \text{and} \quad (\lambda\gamma)^h(R) = c_0^h\ (w^h \circ u^h)_{cc}(c_0^h, z_0^h) = \widehat{\lambda}^h \widehat{\gamma}^h. \tag{2.77}$$

The former of these properties is immediate from the chain rule and the fundamental theorem of calculus (here, we use the continuity of the integrand, from Assumption 2 and our characterization of $c^h(\tilde{u})$ in Appendix 2.9.6.6.4.) The latter is somewhat more complicated; we relegate its proof to Appendix 2.9.6.6.5 (See "Proof of regularity, part 2"). These facts allow us to relate conditional expectations of $\lambda^h(R)$ and $(\lambda\gamma)^h(R)$ to $\widehat{\lambda}(z_0^h)$ and $(\widehat{\lambda\gamma})(z_0^h)$: We show in Appendix 2.9.6.6.5 that expectations of $R(z_0^h)\lambda^h(R)$, $R(z_0^h)\lambda^h(R)\eta^h(R)$, $R(z_0^h)\lambda^h(R)\varepsilon^h(R)$, $R(z_0^h)\lambda^h(R)\frac{\eta^h(R)^2}{\varepsilon^h(R)}$, $R(z_0^h)(\lambda\gamma)^h(R)$ conditional on income z_0^h exist and are equal to $R(z_0^h)\widehat{\lambda}(z_0^h)$, $R(z_0^h)\widehat{\lambda}(z_0^h)\eta_{\le}(z_0^h;\epsilon)$, $R(z_0^h)\widehat{\lambda}(z_0^h)\varepsilon_{\le}(z_0^h;\epsilon)$ and $R(z_0^h)\widehat{\lambda}(z_0^h)\left(\frac{\eta^2}{\varepsilon}\right)_{\le}(z_0^h;\epsilon)$, and $R(z_0^h)(\widehat{\lambda\gamma})(z_0^h)$ respectively.

Part 3: Lagrangian sufficiency

Derivatives of planner's Lagrangian

Define $\mathcal{L} : R + B_\delta(\mathbf{0}) \to \mathbb{R}$ by $\mathcal{L}(\widetilde{R}) \equiv \int w^h \circ V^h(\widetilde{R}) d\mu + \int \left(z^h(\widetilde{R}) - c^h(\widetilde{R})\right) d\mu - G$. Toward the eventual goal of making a Lagrangian sufficiency argument, we now consider and compute the derivatives of $\mathcal{L}(\widetilde{R})$. By Lemmas 3 and 5, we may take δ small enough that, by the linearity of differentiation, $\mathcal{L}(\widetilde{R})$ is twice-continuously Frechet-differentiable on $R + B_\delta(\mathbf{0})$. In Appendix 2.9.6.6.6 we show that—under Assumption 6—strengthened versions of Lemmas 6 and 7 hold for the objective defined above, so that *for all* $\Delta \in \mathbf{\Delta}$,

$$D_\Delta \mathcal{L}(R) = \int_{\operatorname{supp} g} g(z)\left[\lambda(z) + \psi(z)\right]\Delta(z) dz$$
$$D^2_{\Delta\Delta}\mathcal{L}(R) = \int_{\operatorname{supp} g} g(z)\left[(\Phi_0(z) + \Psi_0(z))\left(\frac{\Delta(z)}{R(z)}\right)^2 + (\Phi_1(z) + \Psi_1(z))\left(\frac{\Delta'(z)}{R'(z)}\right)^2\right] dz, \tag{2.78}$$

[78] The only properties of Φ we will use are that Φ is infinitely continuously differentiable, $\Phi(0) = 0$, $\Phi'(0) = 1$, and some scalar $\overline{\Phi}$ bounds $|\Phi(x)|$ and $|\Phi'(x)|$ across all $x \in \mathbb{R}$.

where $\psi(z), \Phi_0(z), \Psi_0(z), \Phi_1(z), \Psi_1(z)$ are as defined in Lemmas 6 and 7; recall these functions are continuous on $z \in \operatorname{supp} g$.

This expression for the Lagrangian's first derivative has a straightforward implication: By (2.73) and (2.77), $\lambda(z) + \psi(z) = 0$, so $D\mathcal{L}(R) = 0$. Turning now to the Lagrangian's second derivative, we claim there exists a scalar $b > 0$ such that for all $z \in \operatorname{supp} g$,

$$\Phi_0(z) + \Psi_0(z) \leq -b(R(z) + z) \qquad \text{and} \qquad \Phi_1(z) + \Psi_1(z) \leq -b(R(z) + z). \tag{2.79}$$

To see this, first note that (a) by the definitions of $(\widehat{\lambda\gamma})(z)$ (see (2.74)) and $\Phi_0(z)$, and (b) since—as $\lambda^h(R) = \widehat{\lambda}^h$ and $(\lambda\gamma)^h(R) = \widehat{\lambda}^h\widehat{\gamma}^h$ for measure one of households—the conditional expectations discussed above remain valid if $\widehat{\lambda}^h$ and $\widehat{\lambda}^h\widehat{\gamma}^h$ are replaced by $\lambda^h(R)$ and $(\lambda\gamma)^h(R)$, respectively, we have that for all $z \in \operatorname{supp} g$,

$$\begin{aligned} \Phi_0(z) &= R(z)\left[(\widehat{\lambda\gamma})(z) + \frac{d\log R(z)}{d\log z}\widehat{\lambda}(z)\left(\frac{\eta^2}{\varepsilon}\right)(z;\epsilon) + \frac{1}{2}\left(\alpha(z) + \frac{d\log R(z)}{d\log z}\right)\widehat{\lambda}(z)\eta(z;\epsilon) - \frac{1}{2}\frac{d\left(\widehat{\lambda}(z)\eta(z;\epsilon)\right)}{d\log z}\right] \\ \implies \quad \Phi_0(z) + \Psi_0(z) &= R(z)\left[-1 - \frac{z}{R(z)} - \frac{\Psi_0(z)}{R(z)}\right] + \Psi_0(z) = -R(z) - z. \end{aligned} \tag{2.80}$$

So $\Phi_0(z) + \Psi_0(z) \leq -b(R(z) + z)$ for any $b \in (0, 1)$. Moreover, by the definitions of $\Phi_1(z)$ and $\Psi_1(z)$, we have, for all $z \in \operatorname{supp} g$,

$$\Phi_1(z) + \Psi_1(z) = R(z)\frac{d\log R(z)}{d\log z}\underbrace{(\lambda\varepsilon)(z)}_{=\widehat{\lambda}(z)\varepsilon_{\leq}(z;\epsilon)} - \Pi_{DEFG}(z) \leq B^R\epsilon(\overline{\lambda}_c R(z) + \overline{\lambda}_z z) - \bar{b}_c R(z) - \bar{b}_z z. \tag{2.81}$$

Above, the second inequality uses Assumption 1, the definitions of $\overline{\lambda}_c$ and $\overline{\lambda}_z$, and the bounding assumption on $\Pi_{DEFG}(z)$ in the statement of the theorem. Recalling our definition of $\epsilon < \frac{1}{B^R}\min\left[\frac{\bar{b}_c}{\overline{\lambda}_c}, \frac{\bar{b}_z}{\overline{\lambda}_z}\right]$, we have shown that there exists $b > 0$ such that $\Phi_1(z) + \Psi(z) \leq -b(R(z) + z)$.

We summarize our observations about $\mathcal{L}$'s derivatives as follows: $\mathcal{L}$ is twice-continuously differentiable on $R + B_\delta(\mathbf{0})$ and there exists $b > 0$ such that for all $\Delta \in \mathbf{\Delta}$,

$$\begin{aligned} &D_\Delta\mathcal{L}(R) = 0 \qquad \text{and} \qquad D^2_{\Delta\Delta}\mathcal{L}(R) \leq -b||\Delta||_*^2 \\ &\text{where} \quad ||\Delta||_* \equiv \left[\int g(z)(z + R(z))\left(\left(\frac{\Delta(z)}{R(z)}\right)^2 + \left(\frac{\Delta'(z)}{R'(z)}\right)^2\right)dz\right]^{\frac{1}{2}}. \end{aligned} \tag{2.82}$$

[79]

Lagrangian sufficiency on a restricted domain

To finish the proof, it suffices to show that, given any $M > 0$, there exists sufficiently small $\tilde{\delta} > 0$ so that $\mathcal{L}(\widetilde{R})$ is maximized within $R + B_{\tilde{\delta}}(\mathbf{0}) \cap \mathbf{\Delta}^*_M$ by $\widetilde{R} = R$, where $\mathbf{\Delta}^*_M = \{\Delta \in \mathbf{\Delta} \mid ||\Delta|| \leq M||\Delta||_*\}$ is as defined in the theorem statement. A Lagrangian sufficiency argument then completes the proof.[80] The reason we restrict the domain to $\mathbf{\Delta}^*_M$ is that this domain rules out sequences of variations which become much larger in the sense of $||\cdot||$ (the relevant norm for Taylor's theorem) than in the sense of $||\cdot||_*$ (the relevant sense for our bounds on the Lagrangian's second derivative).

To see this, first note that for any $R + \Delta \in R + B_\delta(\mathbf{0})$, Taylor's theorem applied along the line between

[80]Put simply, Lagrangian sufficiency says that if there exists a Lagrange multiplier—in our case 1—so that a feasible point—in our case R—maximizes the Lagrangian among all points in some set that includes all feasible points, then that point solves the constrained optimization problem.

R and $\widetilde{R}$ (allowed by $\mathcal{L}$'s twice-continuous differentiability), we have

$$\begin{aligned}\mathcal{L}(R+\Delta) &= \mathcal{L}(R) \;+\; D_\Delta \mathcal{L}(R) \;+\; \frac{1}{2} D^2_{\Delta\Delta}\mathcal{L}(R+\alpha_\Delta \Delta) \\ &= \mathcal{L}(R) \;+\; D_\Delta \mathcal{L}(R) \;+\; \frac{1}{2} D^2_{\Delta\Delta}\mathcal{L}(R) + \frac{1}{2}\left(D^2_{\Delta\Delta}\mathcal{L}(R+\alpha_\Delta\Delta) \;-\; D^2_{\Delta,\Delta}\mathcal{L}(R)\right)\end{aligned} \tag{2.83}$$

for some $\alpha_\Delta \in [0,1]$. We next argue that $\frac{D^2_{\Delta\Delta}\mathcal{L}(R+\alpha_\Delta\Delta)-D^2_{\Delta\Delta}\mathcal{L}(R)}{||\Delta||^2}$ converges to 0 uniformly across Δ and α_Δ for all sequences $\Delta \to 0$. To see this, note that by $\mathcal{L}$'s twice-continuous differentiability, we have that for all $\hat{\epsilon} > 0$, there exists $\hat{\delta} > 0$ such that for all $\check{\Delta} \in B_{\hat{\delta}}(\mathbf{0})$ and all non-zero $\widetilde{\Delta}, \widehat{\Delta} \in \mathbf{\Delta}$,

$$\frac{\left|D^2\mathcal{L}_{\widetilde{\Delta},\widehat{\Delta}}(R+\check{\Delta}) - D^2\mathcal{L}_{\widetilde{\Delta},\widehat{\Delta}}(R)\right|}{||\widetilde{\Delta}||||\widehat{\Delta}||} < \hat{\epsilon}. \tag{2.84}$$

The desired conclusion is implied by taking $\check{\Delta} = \alpha_\Delta \Delta$ and $\widetilde{\Delta} = \widehat{\Delta} = \Delta$. In particular, this implies that the last term in (2.83) is $o(||\Delta||^2)$ in the sense that

$$\frac{\frac{1}{2}\left(D^2_{\Delta\Delta}\mathcal{L}(R+\alpha_\Delta\Delta) \;-\; D^2_{\Delta\Delta}\mathcal{L}(R)\right)}{||\Delta||^2} \to 0 \qquad \text{as} \qquad ||\Delta|| \to 0. \tag{2.85}$$

Combining these observations with the previous step, we may, for any $R+\Delta \in R + B_\delta(\mathbf{0})$, write

$$\begin{aligned}\mathcal{L}(R+\Delta) \;&=\; \mathcal{L}(R) \;+\; D_\Delta\mathcal{L}(R) \;+\; \frac{1}{2}D^2_{\Delta\Delta}\mathcal{L}(R) \;+\; o(||\Delta||^2) \\ &\leq\; \mathcal{L}(R) \;-\; b\frac{||\Delta||^2_*}{2} \;+\; o(||\Delta||^2) \\ \implies \quad \mathcal{L}(R+\Delta) - \left(\mathcal{L}(R) \;-\; \frac{b}{2}\frac{||\Delta||^2_*}{2}\right) \;&\leq\; -\frac{b}{2}\frac{||\Delta||^2_*}{2} \;+\; o(||\Delta||^2)\end{aligned} \tag{2.86}$$

Finally, we claim that—for any $M > 0$—there exists $\tilde{\delta}$ such that for all $R+\Delta \in R + B_{\tilde{\delta}}(\mathbf{0}) \cap \mathbf{\Delta}^*_M$, $\mathcal{L}(R+\Delta) \leq \mathcal{L}(R) - \frac{b}{2}\frac{||\Delta||^2_*}{2}$; note this implies $\mathcal{L}(R+\Delta) \leq \mathcal{L}(R)$.

To see this, suppose otherwise, i.e. there exists a sequence $\Delta_n \to 0$, with each $\Delta_n \in B_\delta(\mathbf{0}) \cap \mathbf{\Delta}^*_M$, such that for all n, $\mathcal{L}(R+\Delta_n) > \mathcal{L}(R) - \frac{b}{2}\frac{||\Delta_n||^2_*}{2}$. By (2.86), this implies

$$0 < -\frac{b}{2}\frac{||\Delta_n||^2_*}{2} + o(||\Delta||^2) \leq -\frac{b}{2\sqrt{M}}\frac{||\Delta_n||^2}{2} + o(||\Delta_n||^2). \tag{2.87}$$

where the second inequality is by the definition of $\mathbf{\Delta}^*_M$. For small enough n, the RHS is strictly negative, a contradiction.

This guarantees that—within $\mathbf{\Delta}^*_M$—the Lagrangian of the planner we have constructed is locally maximized at R. By Lagrangian sufficiency, this completes the proof.

2.9.3.3 Proof of claims from "sort and extort" section (2.4.3.1)

Suppose that each household h belongs to one of finitely many groups $i^h \in I$. We denote income-conditional statistics $x^h | z^h_0 = z, i^h = i$ within each group by $x(z;i)$ and so on. To begin, we will make use of the following simple Lemma:

Lemma 2.9.1. *For any differentiable function $f(z, i)$,*

$$\left(\alpha(z) - \frac{d}{d\log z}\right)\mathbb{E}[f(z;i^h) \mid z_0^h = z] = \mathbb{E}\left[\left(\alpha(z;i^h) - \frac{d}{d\log z}\right) f(z;i^h) \;\middle|\; z_0^h = z\right] \tag{2.88}$$

Proof. Recalling that $\alpha(z) = -\frac{d\log zg(z)}{d\log z}$ and multiplying through by $-h(z)$, note that

$$\begin{aligned}
& - g(z)\left(\alpha(z) - \frac{d}{d\log z}\right)\mathbb{E}[f(z;i^h) \mid z_0^h = z] \\
&= g(z)\frac{d\log}{d\log z}\left(g(z)z\right)\mathbb{E}[f(z;i^h)|z_0^h = z] + g(z)\frac{d}{d\log z}\mathbb{E}[f(z;i^h)|z_0^h = z] \\
&= \frac{d}{dz}\left(g(z)z\right)\mathbb{E}[f(z;i^h)|z_0^h = z] + g(z)\frac{d}{d\log z}\mathbb{E}[f(z;i^h)|z_0^h = z] \\
&= \frac{d}{dz}\left(g(z)z\mathbb{E}[f(z;i^h)|z_0^h = z]\right) \\
&= \frac{d}{dz}\left(\sum_{i\in I} g(z;i)zf(z;i)\right) \\
&= -\sum_{i\in I} g(z;i)\left(\alpha(z,i) - \frac{d}{d\log z}\right) f(z;i) \\
&= -g(z)\mathbb{E}\left[\left(\alpha(z;i^h) - \frac{d}{d\log z}\right) f(z;i^h) \;\middle|\; z_0^h = z\right]
\end{aligned} \tag{2.89}$$

□

Applying this Lemma to DEFG and substituting for $\varepsilon^+(z)$ implies that DEFG can be written as:

$$\begin{aligned}
0 \;\geq\; & -\left(1 + R'(z)\right)\mathbb{E}[\varepsilon(z;i^h) \mid z_0^h = z] \\
& + \left(1 - R'(z)\right)\left(\mathbb{E}\left[\left(\alpha(z;i^h) - \frac{d}{d\log z}\right)\varepsilon^2(z;i^h) \;\middle|\; z_0^h = z\right] + \mathbb{E}\left[\varepsilon(z;i^h)\frac{d}{d\log z}\widehat{\varepsilon}(z;i^h) \;\middle|\; z_0^h = z\right]\right) \\
= & -\mathbb{E}\left[\left(1 + R'(z)\right)\varepsilon(z;i^h) + \left(1 - R'(z)\right)\left(\left(\alpha(z;i^h) - \frac{d}{d\log z}\right)\varepsilon^2(z;i^h) + \varepsilon(z;i^h)\frac{d}{d\log z}\widehat{\varepsilon}(z;i^h)\right) \;\middle|\; z_0^h = z\right] \\
= & \sum_{i\in I}\mathbb{P}\left[i^h = i|z_0^h = z\right]\Pi^i_{DEFG}(z)
\end{aligned} \tag{2.90}$$

Next, we assume that within each group i, preferences satisfy the function form studied in Werning (2007):

$$u^h(c, z) \quad = \quad v_c^i(c) \quad - \quad v_z^i(z) \;/\; \theta^h, \tag{2.91}$$

where v_c^i and v_z^i are increasing and v_c^i is concave and v_z^i is convex.

Proposition 1. *If preferences with a group $i \in I$ satisfy* (2.91)*, then* $\Pi^i_{ABC}(z) \implies \Pi^i_{DEFG}(z)$

Proof. Starting from (DEFG), we divide out by a factor of $z\varepsilon(z;i)$. Letting $h(z,i)$ be the type (if one exists) in group i who supplies labor z, we this gives us

$$0 \geq -\left(1 + R'(z)\right) + \left(1 - R'(z)\right)\left[\alpha(z)\varepsilon(z;i) - \frac{d\varepsilon(z;i)}{d\log z} + \frac{d}{d\log\tilde{z}}\Big|_{\tilde{z}=z}\widehat{\varepsilon}^{h(z,i)}(\tilde{z}, \widetilde{R}) - \frac{d\varepsilon(z;i)}{d\log z}\right] \tag{2.92}$$

where here the fact that the last two terms cancel is an easily-verifiable property of (2.91).

Next we rearrange the first-order condition (ABC) to get

$$(1-R'(z))\left(\alpha(z;i)\varepsilon(z;i)-\varepsilon'(z;i)z\right)=R'(z)\left(1+\Pi^i_{ABC}(z)\right)-(1-R'(z))\frac{d\log R(z)}{d\log z}\eta(z;i)-\frac{d\log R'(z)}{d\log z}\varepsilon(z;i) \tag{2.93}$$

Combining the two previous equations—the former from the second-order condition and the latter from the first-order condition—(DEFG) for group i is equivalent to:

$$0\geq R'(z)\Pi^i_{ABC}(z)-1-(1-R'(z))\frac{d\log R(z)}{d\log z}\eta(z;i)-\frac{d\log R'(z)}{d\log z}\varepsilon(z;i) \tag{2.94}$$

In order to simplify the LHS of this expression, we note that—from the definition of income and compensated elasticities in (2.125)—

$$\begin{aligned}
&1+(1-R'(z_0^h))\frac{d\log R(z_0^h)}{d\log z_0^h}\eta^h(R)+\frac{d\log R'(z_0^h)}{d\log z_0^h}\varepsilon^h(R)\\
&=\frac{\cancel{\frac{d\log M^h(c_0^h,z_0^h)}{d\log c_0^h}\frac{d\log R(z_0^h)}{d\log z_0^h}}+\frac{d\log M^h(c_0^h,z_0^h)}{d\log z_0^h}-\cancel{\frac{d\log R'(z_0^h)}{d\log z_0^h}}-(\cancel{1}-R'(z_0^h))\frac{d\log R(z_0^h)}{d\log z_0^h}\frac{d\log M^h(c_0^h,z_0^h)}{d\log c_0^h}+\cancel{\frac{d\log R'(z_0^h)}{d\log z_0^h}}}{\frac{d\log M^h(c_0^h,z_0^h)}{d\log c_0^h}\frac{d\log R(z_0^h)}{d\log z_0^h}+\frac{d\log M^h(c_0^h,z_0^h)}{d\log z_0^h}-\frac{d\log R'(z_0^h)}{d\log z_0^h}}\\
&=\varepsilon^h(R)\left(\frac{d\log M^h(c_0^h,z_0^h)}{d\log z_0^h}+R'(z)\frac{d\log R(z)}{d\log z}\frac{d\log M^h(c_0^h,z_0^h)}{d\log c_0^h}\right)
\end{aligned} \tag{2.95}$$

So (DEFG) is equivalent to

$$\Pi^i_{ABC}(z)\leq\frac{1}{R'(z)}\varepsilon(z;i)\left(\frac{\partial\log M(R(z),z;i)}{\partial z}+R'(z)\frac{d\log R(z)}{d\log z}\frac{\partial\log M(R(z),z;i)}{\partial c}\right). \tag{2.96}$$

Since $R'(z)>0$ by the proof of Lemma 2, since compensated elasticities are positive, and since the concavity and convexity of v^i_c and v^i_z, respectively, imply $M^h(c,z)$ is increasing in consumption and labor for all h, the RHS is positive. (ABC) for group $i\in I$ therefore implies (DEFG) for group i.

□

Proposition 2. *Suppose each household h has preferences of the form*

$$u^h(c,z) \quad=\quad c \quad-\quad \frac{(z/\theta^h)^{1+1/\beta^h}}{1+1/\beta^h} \tag{2.97}$$

and, local to some income z, $R(z)$ is linear in z and the mean and variance of income-conditional elasticities are constant in z. Then

$$\Pi_{DEFG}(z) \;=\; z\varepsilon(z)\left(R'(z)\Pi_{ABC}(z)-1\right) \;+\; (1-R'(z))z\alpha(z)\mathbb{V}ar[\varepsilon^h(R)^2|z^h(R)=z]. \tag{2.98}$$

Proof. This functional form for preferences implies that there are no income effects and no preference super-elasticities. This fact and the assumptions that $R''(z)=\varepsilon'(z)=\varepsilon^{2\prime}(z)=0$ imply that the first- and second-order tests can be written, respectively, as

$$\begin{aligned}
\Pi_{ABC}(z) &= -1 \;+\; (1-R'(z))\alpha(z)\frac{\varepsilon(z)}{R'(z)}\\
\Pi_{DEFG}(z) &= -\left(1+R'(z)\right)z\varepsilon(z) \;+\; \left(1-R'(z)\right)z\alpha(z)\varepsilon^2(z).
\end{aligned} \tag{2.99}$$

Decomposing $\varepsilon^2(z) = \varepsilon(z)^2 + \mathbb{V}\text{ar}[\varepsilon^h(R)^2|z^h(R) = z]$ then implies that

$$\Pi_{DEFG}(z) - R'(z)z\varepsilon(z)\Pi_{ABC}(z) = -z\varepsilon(z) + (1 - R'(z))\, z\alpha(z)\mathbb{V}\text{ar}[\varepsilon^h(R)^2|z^h(R) = z]. \quad (2.100)$$

□

2.9.4 Empirical Appendix

2.9.5 Selection of bandwidths for local regressions

For picking our bandwidths, we minimize the leave-one-one cross validations criteria, which is the average of squared leave-one-out residuals, that is

$$LOOCV = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{g}_{-i}(x_i))^2, \quad (2.101)$$

where $\hat{g}_{-i}(x_i))$ is the predicted value for y_i using the estimated model $\hat{g}_{-i}$ that does use the observation i, but evaluated at the covariate values x_i. The average of these residuals is an estimate for the asymptotic mean integrated square error (AMISE) of model with bandwith h. By minimizing it, we pick the model with lowest estimated AMISE (Li and Racine, 2007).

For regressions with a large number of observations, the procedure above is computationally demanding. To speed up computations we use the procedure described in Racine (1993), where the leave-one-out cross validation is computed in subsamples and then is scaled down, using the fact that the optimal bandwith should be proportional to $c\sigma_x n^{-1/5}$, where c is a constant that does not depend on the number of observations n, and σ_x is the standard deviation of x.

2.9.5.1 Non-parametric identification of elasticity moments

In this section we show that we can recover all moments of the distribution of elasticities from moments in the data. We assume that behavioral responses of each worker to tax changes are linear, that is, each of them has constant elasticities.

Under this assumption, the change in income and the change in marginal retention are related through an equation of the form:

$$y_t^h = a_t^h + b_t^h x_t^h$$

We now show that we can recover all moments of the joint distribution of (a_t^h, b_t^h).

We now raise the structural equation above to the $n - power$ and algebraically manipulate it into a regression equation:

$$(y_t^h)^n = (a_t^h + b_t^h x_t^h)^n$$

$$(y_t^h)^n = \sum_{k=0}^{n}\binom{n}{k}(a_t^h)^{n-k}(b_t^h x_t^h)^k$$

$$(y_t^h)^n = \sum_{k=0}^{n} \binom{n}{k} \mathbb{E}[(a_t^h)^{n-k}(b_t^h)^k](x_t^h)^k + \sum_{k=0}^{n} \binom{n}{k} \left((a_t^h)^{n-k}(b_t^h) - \mathbb{E}[(a_t^h)^{n-k}(b_t^h)^k]\right)(x_t^h)^k$$

Now, notice that, assuming x_t^h is randomly assigned, for any k and k',

$$\mathbb{E}\left[(x_t^h)^{k'k}\left((a_t^h)^{n-k}(b_t^h) - \mathbb{E}[(a_t^h)^{n-k}(b_t^h)^k]\right)\right] = 0.$$

Therefore, the equation above is a regression equation, where the coefficients are moments of the distribution of a_t^h and b_t^h. Crucially, note than all moments of the joint distribution of (a_t^h, b_t^h) take the form $\mathbb{E}[(a_t^h)^{n-k}(a_t^h)^k]$ for some n and k. So we can recover it from some such regression, assuming that there is enough variation in x (which prevents collinearity).

2.9.5.2 Derivation of mechanism regression equation

In this section, we show that the regression equation 2.22 recovers the variance of elasticities within brackets.

We start from the following expression, based on (2.21):

$$\mathbb{E}[\varepsilon^h|z, \log R_t'] \approx \mathbb{E}[\varepsilon^h|z, \overline{\log R'}] + \alpha_t(z)\mathbb{V}\mathrm{ar}[\varepsilon^h|z](\log R_t'(z) - \overline{\log R'}(z)).$$

Writing the elasticity ε^h as the sum of its expectation and an expectational error ξ^h, we use this expression to substitute in for elasticities in the definition of the elasticity elasticity as the approximate change in income that results from an exogenous change in marginal taxes at the initial income level:

$$\begin{aligned}\Delta \log z_t^h &\approx \varepsilon^h \Delta \log R_t'(z_t^h) + a_t^h \\ &\approx \left[\mathbb{E}[\varepsilon^{\tilde{h}}|z^h, \overline{\log R'}] + \alpha(z^h)\mathbb{V}\mathrm{ar}[\varepsilon^{\tilde{h}}|z^h]\left(\log R_t'(z_t^h) - \overline{\log R'}(z_t^h)\right) + \xi_t^h\right] \Delta \log R_t'(z_t^h) + a_t^h\end{aligned}$$

where here the expectation and variances are over the household $\tilde{h}$, whose income we condition on being equal to z^h.

Finally, we add and subtract terms to arrive at a regression equation:

$$\begin{aligned}\Delta \log z_t^h \approx \quad & \mathbb{E}[\mathbb{E}[\varepsilon^{\tilde{h}}|z^{\hat{h}}, \overline{\log R'}]] \cdot \Delta \log R'(z^h) \\ + \quad & \mathbb{E}[\mathbb{V}\mathrm{ar}[\varepsilon^{\tilde{h}}|z^{\hat{h}}]] \cdot \alpha_t(z_t^h)\left(\log R_t'(z_t^h) - \overline{\log R'}(z_t^h)\right)\Delta \log R'(z_t^h) \\ + \quad & \left(\mathbb{V}\mathrm{ar}[\varepsilon^{\tilde{h}}|z^h] - \mathbb{E}[\mathbb{V}\mathrm{ar}[\varepsilon^{\tilde{h}}|z^{\hat{h}}]]\right)\alpha_t(z_t^h)\left(\log R_t'(z_t^h) - \overline{\log R'}(z_t^h)\right)\Delta \log R_t'(z_t^h) \\ + \quad & \left(\mathbb{E}[\varepsilon^{\tilde{h}}|z^h, \overline{\log R'}] - \mathbb{E}[\mathbb{E}[\varepsilon^{\tilde{h}}|z^{\hat{h}}, \overline{\log R'}]]\right)\Delta \log R_t'(z_t^h) \\ + \quad & \xi_t^h \Delta \log R_t'(z_t^h) \quad + \quad a_t^h,\end{aligned}$$

where all nested expectations are over $\tilde{h}$ for the inner expectation and $\hat{h}$ for the outer expectation.

This is a regression equation, and it is identified under the assumption that tax changes are randomly assigned. The coefficient on $\alpha_t(z_t^h)(\log R_t'(z_t^h) - \overline{\log} R'(z_t^h))\Delta \log R_t'(z_t^h)$ recovers the average variance measure $\mathbb{E}[\mathbb{V}\mathrm{ar}[\varepsilon^{\tilde{h}}|z^{\hat{h}}]]$.

2.9.6 Omitted Proofs

This appendix contains proofs of supporting results used in the main theorems, as well as various technical details.

2.9.6.1 Proof of Lemma 1

This proof shows that $(\mathbf{\Delta}, ||\cdot||)$ is a Banach space.

$\mathbf{\Delta}$ is a real vector space by standard arguments (the main property to consider is closure under addition). Moreover, $||\cdot||$ induces a norm on $\mathbf{\Delta}$; this is easy to check:

- If $||\Delta||= 0$, then by definition, $\Delta(z) = 0$ for all $z > 0$. Continuity then implies $\Delta(z) = 0$ for all $z \geq 0$. In the other direction, if $\Delta(z) = 0$, then $||\Delta||= 0$ by definition.
- For any $a \in \mathbb{R}$, $||a\Delta||= |a|\cdot||\Delta||$.
- To show the triangle inequality, take $\Delta, \widetilde{\Delta} \in \mathbf{\Delta}$. Let B^{Δ} and $B^{\widetilde{\Delta}}$ be bounds for which

$$\begin{aligned} &\forall z \in \mathbb{R}_{>0},\ |\Delta(z)|\leq B^{\Delta}|R(z)| \text{ and } |\Delta'(z)|, \left|\frac{d\Delta'(z)}{d\log z}\right|, \left|\frac{d^2\Delta'(z)}{d\log z^2}\right| \leq B^{\Delta}|R'(z)| \\ &\text{and} \quad |\widetilde{\Delta}(z)|\leq B^{\widetilde{\Delta}}|R(z)| \text{ and } |\widetilde{\Delta}'(z)|, \left|\frac{d\widetilde{\Delta}'(z)}{d\log z}\right|, \left|\frac{d^2\widetilde{\Delta}'(z)}{d\log z^2}\right| \leq B^{\widetilde{\Delta}}|R'(z)| \end{aligned} \tag{2.102}$$

  By the linearity of differentiation, we have that

$$\begin{aligned} &\forall z \in \mathbb{R}_{>0},\ |\Delta(z) + \widetilde{\Delta}(z)|\leq (B^{\Delta} + B^{\widetilde{\Delta}})|R(z)| \\ &\text{and} \quad |\Delta'(z) + \widetilde{\Delta}'(z)|, \left|\frac{d\left(\Delta'(z) + \widetilde{\Delta}'(z)\right)}{d\log z}\right|, \left|\frac{d^2\left(\Delta'(z) + \widetilde{\Delta}'(z)\right)}{d\log z^2}\right| \leq (B^{\Delta} + B^{\widetilde{\Delta}})|R'(z)| \end{aligned} \tag{2.103}$$

  which implies $\Delta + \widetilde{\Delta} \in \mathbf{\Delta}$.

In order to argue that $(\mathbf{\Delta}, ||\cdot||)$ is a real Banach space, it remains to verify that it is complete. To this end, consider a sequence $(\Delta_n)_{n=0}^{\infty}$ of functions in $\mathbf{\Delta}$. Moreover, suppose the sequence is Cauchy with respect to $||\cdot||$. We will show the sequence converges to some limit contained in $\mathbf{\Delta}$.

To begin, note that since Δ_n is Cauchy, so are—for each $z \in \mathbb{R}_{>0}$—$\Delta_n(z), \Delta'_n(z)$, $\frac{d\Delta'_n(z)}{d\log z}$, and $\frac{d^2\Delta'_n(z)}{d\log z^2}$. Each therefore converges pointwise to some functions $\Delta^0(z), \Delta^1(z), \Delta^2(z), \Delta^3(z) : \mathbb{R}_{>0} \to \mathbb{R}$. Note that the sequence $\Delta_n(0)$ is also Cauchy: Otherwise the continuity of the Δ_ns implies that for some $\epsilon > 0$, there exist arbitrarily large n and m as well as $z \in (0, 1)$ for which $\Delta_n(z) - \Delta_m(z) > \epsilon$. Since $R(z)$ is, by continuity, bounded on $[0, 1]$, this would violate that Δ_n is Cauchy. So $\Delta^0(z)$ is defined on $\mathbb{R}_{\geq 0}$.

We next establish two facts about these functions $\Delta^k(z)$: First, for each $k = 0, 1, 2, 3$ there exists $B_k \in \mathbb{R}$ such that

$$\forall z \in \mathbb{R}_{>0}, \quad |\Delta^k(z)|\leq B_k|R^{(\max[1,k])}(z)|. \tag{2.104}$$

To see this, suppose not. In this case, the bounds $B_n = ||\Delta||_n$ associated with each Δ_n must diverge. But it is easy to see this violates that the sequence is Cauchy. Second, we note that each $\Delta^k(z)$ function is continuous in z. This follows from the fact that, for any $z \in \mathbb{R}_{>0}$, $|R^{(\max[1,k])}(z)|$ achieves some max on $[z/2, 2z]$ by continuity (from Assumption 1). The definition of $||\cdot||$ then implies that whichever of $\Delta_n(\cdot), \Delta'_n(\cdot)$, $\frac{d\Delta'_n(\cdot)}{d\log z}$, and $\frac{d^2\Delta'_n(\cdot)}{d\log z^2}$ converges to $\Delta^k(\cdot)$ on $[z/2, 2z]$ does so uniformly. Since each Δ_n is three-times continuous differentiable and uniform convergence preserves continuity, this implies Δ^k is continuous on $[z/2, 2z]$; varying z, we get that Δ^k is continuous on $\mathbb{R}_{>0}$.

In the case of $k = 0$, both of these observations on $\Delta^{k=0}$ extend to $\mathbb{R}_{\geq 0}$. First, the fact that $\Delta_n \to \Delta^0$ uniformly on $(0, 1]$, the fact that Δ_n is continuous on [0,1], and the triangle inequality imply that $\Delta_n \to \Delta^0$ uniformly on $[0, 1]$. So Δ^0 inherits Δ_n's continuity on $\mathbb{R}_{\geq 0}$. Applying this continuity and that of $R(z)$ to the fact that for all $z > 0$, $|\Delta^0(z)| \leq B_k |R(z)|$ implies this is also true at $z = 0$.

Finally, we claim that for all $z \in \mathbb{R}_{>0}$, $\Delta^{0\prime}(z)$ exists and equals $\Delta^1(z)$, $\frac{d\Delta^1(z)}{d\log z}$ exists and equals $\Delta^2(z)$, and $\frac{d\Delta^2(z)}{d\log z}(z)$ exists and equals $\Delta^3(z)$. To see this, recall we have already argued that for all $m \in \mathbb{N}$, $\Delta_n \to \Delta^0, \Delta_n' \to \Delta^1, \frac{d\Delta_n'(\cdot)}{d\log z} \to \Delta^2$, and $\frac{d^2\Delta_n'(\cdot)}{d\log z^2} \to \Delta^3$ uniformly on $[\frac{1}{m}, m]$. The lemma stated below therefore implies that each of the derivatives described above exists and coincides in the desired way on $[\frac{1}{m}, m]$. Applying this argument for each $m \in \mathbb{N}$ gives us the desired claim on $\mathbb{R}_{>0}$.

Lemma 2.9.2. *Let $f_n : [a, b] \to \mathbb{R}$ be a sequence of differentiable functions whose derivatives f_n' are continuous. If $f_n \to f$ uniformly and $f_n' \to g$ uniformly, then f is differentiable and $f' = g$.*

Proof. This is a standard fact in real analysis. See, e.g., Bakker (2018). ☐

At this point we have shown that the sequence (Δ_n) converges to a function Δ^0 that is continuous on $\mathbb{R}_{\geq 0}$ and that has first, second, and third derivatives $\mathbb{R}_{>0}$ (this follows from the previous step because $z > 0$ on $\mathbb{R}_{>0}$). We have shown that these derivatives are continuous on $\mathbb{R}_{>0}$ (similarly). Finally, we have shown the existence of some $B = \max_{k=0,1,2,3} B_k$ so that $|\Delta^0(z)| \leq B|R(z)|$, $|\Delta^{0\prime}(z)| \leq B|R'(z)|$, $\left|\frac{d\Delta^{0\prime}(z)}{d\log z}\right| \leq B|R'(z)|$, and $\left|\frac{d^2\Delta^{0\prime}(z)}{d\log z^2}\right| \leq B|R'(z)|$. Together, these observations imply $\Delta^0 \in \boldsymbol{\Delta}$.

2.9.6.2 Proof of Lemma 2

We complete the proof in seven steps. First, we establish a convenient fact we will use throughout; second, we show the existence of a unique labor supply function within $R + B_{\tilde{\epsilon}}(\mathbf{0})$ for some $\tilde{\epsilon}$ (common across $h \in \mathcal{H}$); and, third and fourth, we show this labor supply function is twice continuously differentiable and measurable. Fifth and sixth, we provide explicit expressions for the first and second derivatives of labor supply; and, seventh, we show these as well are measurable. Eighth, we show that $\widetilde{R}(z^h(\widetilde{R}))$ inherits the relevant properties from $z^h(\widetilde{R})$.

Positivity of tax schedule

Fix any $h \in \mathcal{H}$. We will show that an implication of the fact that $z_0^h > 0$ is that for all $z > 0$, $R(z) > 0$ and $R'(z) > 0$. Under Assumption 3, this implies that for all $h \in \mathcal{H}$, $R(z_0^h), R'(z_0^h) > 0$.

We begin with the R case and then proceed to the R' case. To see the former, suppose not. Then since R is continuous, there exists some highest $z < z_0^h$ for which $R(z) = 0$. But then for all $z' \in (z, z_0^h)$, we have

$$\begin{aligned}
\log R(z') &= \int_{\log z_0^h}^{\log z'} \underbrace{\frac{R'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})}}_{\in[-B^R, B^R]} d\tilde{z} \in [-B^R|\log z_0^h - \log z'|, B^R|\log z_0^h - \log z'|] \\
\implies R(z') &\in \left[e^{-B^R|\log z_0^h - \log z'|}, e^{B^R|\log z_0^h - \log z'|}\right] \cdot R(z_0^h)
\end{aligned} \tag{2.105}$$

where the underbrace is by Assumption 1. Since $R(z') \geq e^{-B^R|\log z_0^h - \log z|} R(z_0^h) > 0$ for all $z' > z$, $R(z) > 0$ by continuity, a contradiction.

We may apply a similar argument to R's first derivative, leveraging the observation that, since $z_0^h > 0$ and u^h and R are differentiable, h's labor supply satisfies the first-order condition:

$$\underbrace{u_c^h(c_0^h, z_0^h)}_{>0} R'(z_0^h) + \underbrace{u_z^h(c_0^h, z_0^h)}_{<0} = 0 \tag{2.106}$$

where the underbraces are by Assumption 2 and the fact that $z_0^h, c_0^h > 0$. This implies $R'(z_0^h) > 0$. The same argument as above applied to R' instead of R the implies that $R'(z) > 0$ at all $z > 0$.

Existence and uniqueness of labor supply function

We now claim that, for any $\Delta \in B_{\hat{c}(\bar{c})}(\mathbf{0})$, h's problem (2.5) at $R + \Delta$ has a solution within $e^{B_{\epsilon^h}(\log z_0^h)}$. To see this, first recall that by assumption, for any $z \notin e^{B_{\epsilon^h}(z^h(R))}$,

$$u^h\left(R\left(z\right)e^{\bar{c}},\ z\right) \leq u^h\left(R(z_0^h)e^{-\bar{c}},\ z_0^h\right) \tag{2.107}$$

Second, we claim that for $||\Delta|| \leq \hat{c}(\bar{c}) \equiv \frac{\min(1,\bar{c})}{2}$, we have $R(z) + \Delta(z) \in [R(z)e^{-\bar{c}}, R(z)e^{\bar{c}}]$. To see this, note that $||\Delta|| < \hat{c}(\bar{c}) \leq \bar{c} < e^{\bar{c}} - 1$ implies $R(z) + \Delta(z) \leq R(z)e^{\bar{c}}$ for all $z \in \mathbb{R}_{\geq 0}$. Moreover it is easy to verify that for all $\bar{c} > 0$, $1 - \frac{\min(1,\bar{c})}{2} < e^{-\bar{c}}$, implying that for $||\Delta|| < \hat{c}(\bar{c})$, $R(z) + \Delta(z) \geq R(z)e^{-\bar{c}}$ with strict inequality wherever $|R(z)| > 0$. Combining these observations and using that utility is strictly increasing in consumption, we have for any $z \notin e^{B_\delta(\log z_0^h)}$ that

$$\begin{aligned} u^h\left(R(z) + \Delta(z), z\right) &\leq u^h\left(R\left(z\right)e^{\bar{c}},\ z\right) \\ &\leq u^h\left(R(z_0^h)e^{-\bar{c}},\ z_0^h\right) < u^h\left(R(z_0^h) + \Delta(z_0^h),\ z_0^h\right) \end{aligned} \tag{2.108}$$

The final, strict, inequality uses the fact that $R(z_0^h) > 0$, as established above.

Since we have shown that h strictly prefers z_0^h to any $z \notin e^{\log B_{\epsilon^h}(\log z_0^h)}$ when facing the schedule $R + \Delta$, h's problem (2.5) can for any $\Delta \in B_{\bar{c}(\bar{c})}(\mathbf{0})$ be rewritten as

$$\max_{z \in e^{\overline{B}_{\epsilon^h}(\log z^h(R))}} u^h\left((R + \Delta)(z), z\right). \tag{2.109}$$

Since the objective is continuous and the domain is compact, (2.109) has a solution, and in particular (by the comparison-to-z_0^h argument above) one within $e^{B_{\epsilon^h}(\log z_0^h)}$.

We now argue that, for any $\Delta \in B_{\tilde{\epsilon}}(\mathbf{0})$, this problem has a *unique* solution, where $\tilde{\epsilon} \equiv \min(\hat{c}(\bar{c}), b(\eta, R))$ and $b(\eta, R) \equiv \min\left(\frac{1}{2}, \frac{\eta}{40\max(1, B^R)^2}\right)$. To see this, consider the compensating variation function $v^h(\cdot)$ of Assumption 3. Since—by Assumption 2—u^h is three-times continuously differentiable and $u_c^h > 0$, and—by Assumption 1—R is three-times continuously differentiable, the implicit function theorem implies that v is three-times continuously differentiable within $B_{\epsilon^h}(\log z_0^h)$. Note that, for any $\Delta \in \mathbf{\Delta}$ with $||\Delta|| < 1$ and $z \in e^{B_{\epsilon^h}(\log z_0^h)}$, (2.27) implies that

$$\begin{aligned} u^h\Big(\left(R(z) + \Delta(z)\right)e^{v^h(\log z) - \widetilde{\Delta}(\log z)}, z\Big) &= u^h\left(R(z_0^h), z_0^h\right), \\ \text{where } \widetilde{\Delta}(\tilde{z}) &\equiv \log\left(1 + \frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})}\right). \end{aligned} \tag{2.110}$$

Note that $\widetilde{\Delta}(\log z)$ is three-times continuously differentiable in $B_{\epsilon^h}(\log z_0^h)$ since $R(z) > 0$ and Δ and R are three-times continuously differentiable on $\mathbb{R}_{>0}$ by Assumption 1 and the definition of $\mathbf{\Delta}$.

Combining (2.110) with (2.109) implies that, for $\Delta \in B_{\hat{c}(\bar{c})<\frac{1}{2}}(\mathbf{0})$, z solves h's problem at $R+\Delta$ if and only if

$$\log(z) \in \underset{\tilde{z}\in B_{\epsilon^h}(\log z_0^h)}{\arg\min}\ v^h(\tilde{z}) - \widetilde{\Delta}(\tilde{z}). \tag{2.111}$$

To show that h's problem has a unique solution at $R+\Delta$, it therefore suffices to show that the objective $v^h(\cdot) - \widetilde{\Delta}(\cdot)$ is strictly convex on $B_{\epsilon^h}(\log z_0^h)$. Since we know by assumption that $v^{h\prime\prime}(\cdot) \geq \eta$, it suffices to show that $|\widetilde{\Delta}''(\cdot)| < \eta$. To this end, we compute the derivatives of $\widetilde{\Delta}$ below.

$$\begin{aligned}
\widetilde{\Delta}'(\tilde{z}) &= \frac{\frac{\Delta'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})} - \frac{R'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})}\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})}}{1+\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})}}\\
\widetilde{\Delta}''(\tilde{z}) &= \frac{\frac{\Delta''(e^{\tilde{z}})e^{2\tilde{z}}}{R(e^{\tilde{z}})} + \frac{\Delta'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})} - \frac{\Delta'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})}\frac{R'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})} - \frac{R''(e^{\tilde{z}})e^{2\tilde{z}}}{R(e^{\tilde{z}})}\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})} - \frac{R'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})}\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})} + 2\frac{R'(e^{\tilde{z}})^2e^{2\tilde{z}}}{R(e^{\tilde{z}})^2}\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})} - \frac{R'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})}\frac{\Delta'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})}}{1+\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})}}\\
&\quad - \left(\frac{\frac{\Delta'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})} - \frac{R'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})}\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})}}{1+\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})}}\right)^2
\end{aligned} \tag{2.112}$$

It remains to show that these derivatives are bounded for small enough tax deviations Δ. To this end, recall first that by Assumption 1, there exists $B^R \geq 0$ such that for all $z \in \mathbb{R}_{>0}$, $|R'(z)|z \leq B^R|R(z)|$ and $|R''(z)|z \leq B^R|R'(z)|$. Moreover, recall that for any constant $b>0$, if $||\Delta|| < b$, then $|\Delta(z)| \leq b|R(z)|$, with strict inequality if $|R(z)| > 0$, and $|\Delta'(z)| \leq b|R'(z)|$ and $|\Delta'(z)| \leq b|zR''(z)|$, with strict inequality if $|R'(z)| > 0$. Combining these observations, we may note that for any $\Delta \in B_b(\mathbf{0})$, $z \in e^{B_{\epsilon^h}(\log z_0^h)}$,

$$\begin{gathered}
\left|\frac{\Delta(z)}{R(z)}\right| < b,\ \left|\frac{\Delta''(z)z^2}{R(z)}\right| \leq bB^R,\ \left|\frac{\Delta'(z)z}{R(z)}\right| \leq bB^R,\ \left|\frac{\Delta'(z)z}{R(z)}\frac{R'(z)z}{R(z)}\right| \leq b(B^R)^2,\\
\left|\frac{R''(e^{\tilde{z}})e^{2\tilde{z}}}{R(e^{\tilde{z}})}\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})}\right| \leq b(B^R)^2,\ \left|\frac{R'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})}\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})}\right| \leq bB^R,\\
\left|\frac{R'(e^{\tilde{z}})^2e^{2\tilde{z}}}{R(e^{\tilde{z}})^2}\frac{\Delta(e^{\tilde{z}})}{R(e^{\tilde{z}})}\right| \leq b(B^R)^2,\ \left|\frac{R'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})}\frac{\Delta'(e^{\tilde{z}})e^{\tilde{z}}}{R(e^{\tilde{z}})}\right| \leq b(B^R)^2
\end{gathered} \tag{2.113}$$

Combining these inequalities implies that $\widetilde{\Delta}''(\tilde{z})$ is uniformly bounded by η across all $||\Delta|| < b(\eta, R)$. To see this, note that for any $b \in (0,1)$

$$\widetilde{\Delta}''(\tilde{z}) < \frac{5b(B^R)^2 + 3bB^R}{1-b} + \frac{2bB^R}{(1-b)^2} \leq 10\frac{b}{(1-b)^2}\max(1, B^R)^2 \tag{2.114}$$

In particular, $\widetilde{\Delta}''(z) < \eta$ if $b < \min\left(\frac{1}{2}, \frac{\eta}{40\max(1,B^R)^2}\right) \equiv b(\eta, R)$.[81]

Taking stock, if $\Delta \in B_{b(\eta,R)}(\mathbf{0})$ then $v^h(\tilde{z}) + \widetilde{\Delta}(\tilde{z})$ is strictly convex on $B_{\epsilon^h}(\log z_0^h)$, where $\widetilde{\Delta}(\cdot)$ is defined as in (2.110). This implies that, for $\Delta \in B_{\min(\hat{c}(\bar{c}),b(\eta,R))}(\mathbf{0})$, h's problem at $R+\Delta$ has a unique solution. (Note that $u^h(R)$ is therefore $> -\infty$, a fact we will use throughout.)

Differentiability of labor supply function

[81] Doing out the algebra, let $\bar{b} = \min\left(\frac{1}{2}, \frac{\eta}{40\max(1,B^R)^2}\right)$. Then for any $b < \bar{b}$,

$$10\frac{b}{(1-b)^2}\max(1,B^R)^2 < 10\frac{\bar{b}}{(1-\bar{b})^2}\max(1,B^R)^2 \leq 40\bar{b}\max(1,B^R)^2 \leq \eta \tag{2.115}$$

So far, we have shown the existence of a unique labor supply function $z^h(R+\Delta) : R+B_{\min(\hat{c}(\bar{c}),b(\eta,R))}(\mathbf{0}) \to e^{B_{\epsilon^h}(\log z_0^h)}$ in a neighborhood of an initial tax schedule R. We will now show that $z^h(R+\Delta)$ is twice continuously differentiable.

To start, recall from (2.111) that

$$z^h(R+\Delta) = \underset{z \in e^{\overline{B}_{\epsilon^h}(\log z_0^h)}}{\arg\min}\; v^h(\log z) - \widetilde{\Delta}(\log z). \tag{2.116}$$

Recall we have already argued that $v^h(\cdot) - \widetilde{\Delta}(\cdot)$ is three times continuously differentiable on $B_{\epsilon^h}(\log z_0^h)$.

As we have shown that $z^h(R+\Delta)$ exists and is in $e^{B_{\epsilon^h}(\log z_0^h)}$, we have the first order condition:

$$\frac{d}{d\log z}\left(v^h - \widetilde{\Delta}\right)\left(\log z^h(R+\Delta)\right) = 0. \tag{2.117}$$

Toward an application of the implicit function theorem, define $F : (B_{\min(\hat{c}(\bar{c}),b(\eta,R))}(\mathbf{0}) \subset \mathbf{\Delta}) \times (B_{\epsilon^h}(\log z_0^h) \subset \mathbb{R}) \to \mathbb{R}$ by

$$F(\Delta, \log z) \equiv \frac{d}{d\log z}\left(v^h(\log z) - \log\left(1 + \frac{\Delta(z)}{R(z)}\right)\right) \tag{2.118}$$

First note that F is a map from a product of open subsets of Banach spaces to a Banach space. Next we claim F is twice-continuously Frechet differentiable. That F is twice-continuously differentiable in $\log z$ follows from that v^h, Δ, and R are all three-times continuously differentiable, as discussed above. It therefore suffices to show that F's derivatives in Δ (including the cross-partial with $\log z$) exist and are—at each

$\Delta, \log z$—bounded (over all directions in which the derivatives in Δ may be taken). To see this, we compute

$$
\begin{aligned}
F(\Delta, \log z) &= v^{h\prime}(\log z) - \frac{\frac{\Delta'(z)}{R'(z)} - \frac{\Delta(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \frac{d\log R(z)}{d\log z} \\
D_{\widetilde{\Delta}} F(\Delta, \log z) &= \left(-\frac{\frac{\widetilde{\Delta}'(z)}{R'(z)} - \frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} + \frac{\frac{\Delta'(z)}{R'(z)} - \frac{\Delta(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \frac{\frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \right) \frac{d\log R(z)}{d\log z} \\
D^2_{\widetilde{\Delta}\widehat{\Delta}} F(\Delta, \log z) &= \left(\frac{\frac{\widetilde{\Delta}'(z)}{R'(z)} - \frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \frac{\frac{\widehat{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} + \frac{\frac{\widetilde{\Delta}'(z)}{R'(z)} - \frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \frac{\frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} - 2\frac{\frac{\Delta'(z)}{R'(z)} - \frac{\Delta(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \frac{\frac{\widetilde{\Delta}(z)}{R(z)}\frac{\widehat{\Delta}(z)}{R(z)}}{\left(1+\frac{\Delta(z)}{R(z)}\right)^2} \right) \frac{d\log R(z)}{d\log z}
\end{aligned}
$$

$$
\begin{aligned}
D^2_{\widetilde{\Delta}\log z} F(\Delta, \log z) &= \left(-\frac{\frac{\widetilde{\Delta}'(z)}{R'(z)} - \frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} + \frac{\frac{\Delta'(z)}{R'(z)} - \frac{\Delta(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \frac{\frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \right) \frac{d^2\log R(z)}{d\log z^2} \\
&+ \left(-\frac{\frac{\widetilde{\Delta}''(z)z}{R'(z)} - \frac{\widetilde{\Delta}'(z)}{R'(z)}\frac{d\log R'(z)}{d\log z} - \frac{\widetilde{\Delta}'(z)}{R'(z)}\frac{d\log R(z)}{d\log z} + \frac{\widetilde{\Delta}(z)}{R(z)}\frac{d\log R(z)}{d\log z}}{1+\frac{\Delta(z)}{R(z)}} \right. \\
&\quad + \frac{\frac{\widetilde{\Delta}'(z)}{R'(z)} - \frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \frac{\frac{\Delta'(z)}{R'(z)}\frac{d\log R(z)}{d\log z} - \frac{\Delta(z)}{R(z)}\frac{d\log R(z)}{d\log z}}{1+\frac{\Delta(z)}{R(z)}} \\
&\quad + \frac{\frac{\Delta''(z)z}{R'(z)} - \frac{\Delta'(z)}{R'(z)}\frac{d\log R'(z)}{d\log z} - \frac{\Delta'(z)}{R'(z)}\frac{d\log R(z)}{d\log z} + \frac{\Delta(z)}{R(z)}\frac{d\log R(z)}{d\log z}}{1+\frac{\Delta(z)}{R(z)}} \frac{\frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \\
&\quad - \frac{\frac{\Delta'(z)}{R'(z)} - \frac{\Delta(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \frac{\frac{\Delta'(z)}{R'(z)}\frac{d\log R(z)}{d\log z} - \frac{\Delta(z)}{R(z)}\frac{d\log R(z)}{d\log z}}{1+\frac{\Delta(z)}{R(z)}} \frac{\frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \\
&\quad + \frac{\frac{\Delta'(z)}{R'(z)} - \frac{\Delta(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \left(\frac{\frac{\widetilde{\Delta}'(z)}{R'(z)}\frac{d\log R(z)}{d\log z} - \frac{\widetilde{\Delta}(z)}{R(z)}\frac{d\log R(z)}{d\log z}}{1+\frac{\Delta(z)}{R(z)}} \right. \\
&\qquad \left.\left. + \frac{\frac{\widetilde{\Delta}(z)}{R(z)}}{1+\frac{\Delta(z)}{R(z)}} \frac{\frac{\Delta'(z)}{R'(z)}\frac{d\log R(z)}{d\log z} - \frac{\Delta'(z)}{R'(z)}\frac{d\log R(z)}{d\log z}}{1+\frac{\Delta(z)}{R(z)}} \right) \right) \frac{d\log R(z)}{d\log z}
\end{aligned} \tag{2.119}
$$

It is immediate from Assumption 1, the definition of $\mathbf{\Delta}$, and the fact that $\Delta \in B_{\frac{1}{2}}(\mathbf{0})$ that all of the derivatives above are bounded proportionally to $||\widetilde{\Delta}||$ and $||\widehat{\Delta}||$ as appropriate. F is therefore Frechet differentiable.

Finally, recall we have shown that $D_{\log z} F(\Delta, \log z) = \frac{d^2}{d\log(z)^2}\left[v^h(\log z) - \log\left(1+\frac{\Delta(z)}{R(z)}\right)\right] > 0$. The implicit function theorem for Banach spaces[82] therefore implies that a continuously differentiable solution $\log z^h(\Delta) \in B_{\epsilon^h}(\log z_0^h)$ to the equation $F(\Delta, \log z(\Delta))$ exists on $\Delta B_{\min(\hat{c}(\bar{c}), b(\eta, R))}(\mathbf{0})$ and satisfies $D_{\widetilde{\Delta}} \log z(\Delta) = -\left(D_{\log z} F(\Delta, \log z(\Delta))\right)^{-1} D_{\widetilde{\Delta}} F(\Delta, \log z(\Delta))$. This derivative is bounded over all $\widetilde{\Delta}$ proportionally to $||\widetilde{\Delta}||$ by the arguments above, so it is a Frechet derivative. Moreover, note that—since F is twice-Frechet differentiable, $\log z^h(\Delta)$ has a second derivative given by

$$
\begin{aligned}
D^2_{\widetilde{\Delta},\widehat{\Delta}} \log z(\Delta) &= \left(D_{\log z} F(\Delta, \log z(\Delta))\right)^{-2} D_{\log z \widehat{\Delta}} F(\Delta, \log z(\Delta)) D_{\widetilde{\Delta}} F(\Delta, \log z(\Delta)) \\
&\quad - \left(D_{\log z} F(\Delta, \log z(\Delta))\right)^{-1} D_{\widetilde{\Delta}\widehat{\Delta}} F(\Delta, \log z(\Delta)).
\end{aligned} \tag{2.120}
$$

Since by the arguments above this is bounded over all $\widetilde{\Delta}, \widehat{\Delta}$ proportionally to $||\widetilde{\Delta}||||\widehat{\Delta}||$, it is a Frechet

[82] See, e.g., Martinsson (2006).

derivative as well.

Recalling from the previous step of the proof that the equation $F(\Delta, \log z) = 0$ has a unique solution on its domain, we identify $e^{\log z(\Delta)}$ with $z^h(R+\Delta)$ and conclude that $z^h(\cdot)$ is twice continuously (Frechet) differentiable in $R + B_{\min(\hat{c}(\bar{c}), b(\eta, R))}(\mathbf{0})$.

Measurability of labor supply

Fixing any $\widetilde{R} \in R + B_{\min(\hat{c}(\bar{c}), b(\gamma, R))}(\mathbf{0})$, we wish to show that $z^h(\widetilde{R})$—which recall we have just shown is well-defined—is measurable in h. This argument, written out below, is a straightforward implication of the measurable maximum theorem as stated in Aliprantis and Border (2006). The theorem states that,[83] if $\Gamma : \mathcal{H} \rightrightarrows \mathbb{R}_{>0}$ is a weakly measurable correspondence with non-empty compact values and $f : \mathbb{R}_{>0} \times \mathcal{H} \to \mathbb{R}$ is a Catheodory function (continuous in its first argument and measurable in its second), then the arg max function $\mu(h) \equiv \arg\max_{z \in \Gamma(h)} f(z, h)$ admits a measurable selector.

First define $\Gamma : \mathcal{H} \rightrightarrows \mathbb{R}_{>0}$ by $h \mapsto e^{\overline{B}_{\epsilon^h}(\log z_0^h)}$. Note this correspondence has non-empty and compact values. To see it is moreover weakly measurable, note that it may be obtained by the composition of the (by assumption and our use of the canonical product measure) measurable function $h \mapsto (\log z_0^h, \epsilon^h)$ with the correspondence $\widetilde{\Gamma} : (\tilde{z}, \epsilon) \mapsto e^{\overline{B}_\epsilon(\tilde{z})}$. It therefore suffices to show that this correspondence $\widetilde{\Gamma}$ is weakly measurable. To see this fix any (WLOG non-empty) open set $A \subset \mathbb{R}_{>0}$ and let $\widetilde{A}$ be a dense, countable subset of A.[84] Then

$$\begin{aligned} \widetilde{\Gamma}^{-1}(A) &= \left\{ (\tilde{z}, \epsilon) \in \mathbb{R} \times \mathbb{R}_0 \mid \exists a \in A,\ a \in e^{\overline{B}_\epsilon(\tilde{z})} \right\} \\ &= \left\{ (\tilde{z}, \epsilon) \in \mathbb{R} \times \mathbb{R}_0 \mid \exists \tilde{a} \in \widetilde{A},\ \tilde{a} \in e^{\overline{B}_\epsilon(\tilde{z})} \right\} \\ &= \bigcup_{\tilde{a} \in A} \Gamma^{-1}(\{\tilde{a}\}) \end{aligned} \tag{2.121}$$

where in the second line we have used A's openness and $\widetilde{A}$'s density in A. Since $\widetilde{A}$ is countable it suffices to show each $\Gamma^{-1}(\{\tilde{a}\})$ is measurable.[85] Indeed, $\Gamma^{-1}(\{\tilde{a}\}) = \emptyset$ (which is measurable) if $\tilde{a} \leq 0$ and if $\tilde{a} > 0$,

$$\begin{aligned} \widetilde{\Gamma}^{-1}(\{\tilde{a}\}) &= \{ (\tilde{z}, \epsilon) \in \mathbb{R} \times \mathbb{R}_{>0} \mid \tilde{z} - \epsilon \leq \log a \leq \tilde{z} + \epsilon \} \\ &= \bigcup_{n=0,1,2\ldots, m=1,2,3,\ldots} \left[\log \tilde{a} - n - \frac{1}{m}, \log \tilde{a} + n + \frac{1}{m} \right] \times \left[n + \frac{1}{m}, 1 + n + \frac{1}{m} \right] \\ &\qquad \cup \left[\log \tilde{a} - n - \frac{1}{m}, \log \tilde{a} + n + \frac{1}{m} \right] \times \left[-1 - n - \frac{1}{m}, -n - \frac{1}{m} \right] \end{aligned} \tag{2.122}$$

which is Lebesgue-product-measurable since it is a countable union of rectangles.

Second, define $f : \mathbb{R}_{>0} \times \mathcal{H} \to \mathbb{R}$ by $(z, h) \mapsto u^h(\widetilde{R}(z), z)$. f is well-defined by Assumption 2 since $||\widetilde{R} - R|| < 1$ and $z > 0$ implies $\widetilde{R}(z) > 0$. Moreover f is continuous since u^h and $\widetilde{R}$ are continuous, and

[83] Precisely, the Theorem also requires that $\mathbb{R}_{>0}$ is a separable metrizable space (it is), and $\mathcal{H}$ and its associated event space constitute a measurable space (they do).

[84] Any non-empty subset A of $\mathbb{R}$ admits a countable, dense subset. Here we provide a proof: Fix any $a_0 \in A$. For all $n \in \mathbb{Z}$, $m, k \in \mathbb{N}$, $k \leq m$; define $x_{m,k}^n$ as follows: If $[n + \frac{k-1}{m}, n + \frac{k}{m}]$ intersects A, let $x_{m,k}^n$ be any point in their intersection; otherwise set $x_{m,k}^n = a_0$. We claim that $\widetilde{A} = \bigcup_{n \in \mathbb{Z},\ k,m \in \mathbb{N},\ k \leq m} \{x_{m,k}^n\}$ is a countable (obviously) subset (obviously) of A which is dense in A. To see $\widetilde{A}$ is dense in A, note that for all $a \in A$, a is contained in some interval $[n, n+1]$ for $n \in \mathbb{Z}$, and moreover for all $m \in \mathbb{N}$, there exists $k \in \mathbb{N}, k \leq m$ for which $a \in [n + \frac{k-1}{m}, n + \frac{k}{m}]$. Therefore, for all $m \in \mathbb{N}$, there exists a point $x_{m,k}^n \in \widetilde{A}$ within $\frac{1}{m}$ of a. So a is a limit point of $\widetilde{A}$.

[85] This argument is similar to the one presented in Himmelberg (1975), available at `http://repository.ias.ac.in/90958/1/90958.pdf`

h-measurable since u^h is.

Theorem 18.19 of Aliprantis and Border (2006) implies that the maximization problem

$$\max_{z \in \Gamma(h)} f(z, h) \tag{2.123}$$

has an argmax correspondence that admits a measurable selector $s(h) \in \mathbb{R}_{>0}$. Since we have already established that this maximization problem has a unique solution for each h, it must be that $s(h) = z^h(\widetilde{R})$. We conclude that $z^h(\widetilde{R})$ is measurable in h.

First derivative of labor supply

We have already established each household's labor supply function $z^h(\widetilde{R})$ is well-defined and has two continuous Frechet derivatives when $\widetilde{R} \in R + B_\delta(\mathbf{0})$. While it is not strictly necessary to compute these derivatives in order to complete the proof of the Lemma, this section computes explicit expressions for them, which will be used in a later step of the proof. We therefore now fix such a $\widetilde{R}$ and a household $h \in \mathcal{H}$, and will compute these Frechet derivatives.

We compute the first derivative by totally differentiating h's MRS condition, since by continuous differentiability of preferences and taxes and since $z^h(\widetilde{R}) > 0$, this always holds at the solution to h's problem. Fixing any direction $\Delta \in \mathbf{\Delta}$ and recalling $M^h(c, z) \equiv \frac{-u_z^h(c,z)}{u_c^h(c,z)}$, we have

$$\begin{aligned}
&M^h\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right) = \widetilde{R}'(z^h(\widetilde{R})) \\
&\left.\frac{d \log M^h(c,z)}{d \log c}\right|_{\substack{c=\widetilde{R}(z^h(\widetilde{R})) \\ z=z^h(\widetilde{R})}} \left(\frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))} + \left.\frac{d \log \widetilde{R}(z)}{d \log z}\right|_{z=z^h(\widetilde{R})} D_\Delta \log z^h(\widetilde{R}) \right) \\
&+ \left.\frac{d \log M^h(c,z)}{d \log z}\right|_{\substack{c=\widetilde{R}(z^h(\widetilde{R})) \\ z=z^h(\widetilde{R})}} D_\Delta \log z^h(\widetilde{R}) \quad = \quad \frac{\Delta'(z^h(\widetilde{R}))}{\widetilde{R}'(z^h(\widetilde{R}))} + \left.\frac{d \log \widetilde{R}'(z)}{d \log z}\right|_{z=z^h(\widetilde{R})} D_\Delta \log z^h(\widetilde{R}) \\
&D_\Delta \log z^h(\widetilde{R}) = \left.\frac{-\frac{d \log M^h(c^h,z^h)}{d \log c^h} \frac{\Delta(z^h)}{\widetilde{R}(z^h)} + \frac{\Delta'(z^h)}{\widetilde{R}'(z^h)}}{\frac{d \log M^h(c^h,z^h)}{d \log c^h}\frac{d \log \widetilde{R}(z^h)}{d \log z^h} + \frac{d \log M^h(c^h,z^h)}{d \log z^h} - \frac{d \log \widetilde{R}'(z^h)}{d \log z^h}}\right|_{\substack{z^h \equiv z^h(\widetilde{R}), \\ c^h \equiv \widetilde{R}(z^h)}}
\end{aligned} \tag{2.124}$$

where we may divide through by the denominator in the last line since—were it zero—then $D_\Delta \log z^h(\widetilde{R})$ would not exist for some Δ, and we have already established that $D_\Delta \log z^h(\widetilde{R})$ exists. Note that the log of M^h, $\widetilde{R}(z^h)$, and $\widetilde{R}'(z^h)$ are well defined since $z^h(\widetilde{R}) > 0$ implies $\widetilde{R}(z^h), \widetilde{R}'(z^h) > 0$ and since by Assumption 2 this implies $u_c(c^h, z^h) > 0, u_z(c^h, z^h) < 0$.

This expression in mind, we define

$$\begin{aligned}
\eta^h(\widetilde{R}) &= \left.\frac{-\frac{d \log M^h(c^h,z^h)}{d \log c^h}}{\frac{d \log M^h(c^h,z^h)}{d \log c^h}\frac{d \log \widetilde{R}(z^h)}{d \log z^h} + \frac{d \log M^h(c^h,z^h)}{d \log z^h} - \frac{d \log \widetilde{R}'(z^h)}{d \log z^h}}\right|_{\substack{z^h \equiv z^h(\widetilde{R}), \\ c^h \equiv \widetilde{R}(z^h)}}, \\
\varepsilon^h(\widetilde{R}) &= \left.\frac{1}{\frac{d \log M^h(c^h,z^h)}{d \log c^h}\frac{d \log \widetilde{R}(z^h)}{d \log z^h} + \frac{d \log M^h(c^h,z^h)}{d \log z^h} - \frac{d \log \widetilde{R}'(z^h)}{d \log z^h}}\right|_{\substack{z^h \equiv z^h(\widetilde{R}), \\ c^h \equiv \widetilde{R}(z^h)}}.
\end{aligned} \tag{2.125}$$

Note that—by our expression for $D_\Delta \log z^h(\widetilde{R})$—this coincides with the definition in the main text, in (2.29), but formally treats the set of feasible deviations. In terms of these elasticities the first Frechet derivative of labor supply can be written as

$$D_{\Delta} \log z^h(\widetilde{R}) = \eta^h(\widetilde{R}) \frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))} + \varepsilon^h(\widetilde{R}) \frac{\Delta'(z^h(\widetilde{R}))}{\widetilde{R}'(z^h(\widetilde{R}))} \tag{2.126}$$

Aside: positivity of compensated elasticity

Since they are the second derivatives of the expenditure function, compensated elasticities are always weakly positive. For the reader who is not convinced that $\varepsilon^h(\widetilde{R})$ as defined above is indeed a compensated elasticity in the traditional sense, we suggest Scheuer and Werning (2018). However, for completeness we also provide below an explicit proof that $\widetilde{\varepsilon}^h(\widetilde{R}) \geq 0$.

Fix $h \in \mathcal{H}$. Let $z^h \equiv z^h(\widetilde{R})$. Since $z^h(\widetilde{R})$ solves h's problem at $\widetilde{R}$,

$$\begin{aligned}
0 =& \left.\frac{d}{dz}\right|_{z=z^h} u^h(\widetilde{R}(z), z) \\
=& \underbrace{u^h_c(\widetilde{R}(z^h), z^h)}_{>0} \underbrace{\left[\widetilde{R}'(z^h) - M^h(\widetilde{R}(z^h), z^h)\right]}_{=0} \\
0 \geq & \left.\frac{d^2}{dz^2}\right|_{z=z^h} u^h(\widetilde{R}(z), z) \\
=& \left.\frac{d}{dz}\right|_{z=z^h} \left(u^h_c(\widetilde{R}(z^h), z^h)\right) \underbrace{\left[\widetilde{R}'(z^h) - M^h(\widetilde{R}(z^h), z^h)\right]}_{=0} \\
&+ u^h_c(\widetilde{R}(z^h), z^h) \frac{1}{z^h} \left.\frac{d}{d\log z}\right|_{z=z^h} \left[\widetilde{R}'(z) - M^h(\widetilde{R}(z), z)\right] \\
&= \underbrace{u^h_c(\widetilde{R}(z^h), z^h)}_{>0} \underbrace{\widetilde{R}'(z^h)}_{>0} \frac{1}{z^h} \left[\frac{d\log \widetilde{R}'(z^h)}{d\log z^h} - \frac{d\log M^h(c^h, z^h)}{d\log c^h} \frac{d\log \widetilde{R}(z^h)}{d\log z^h} - \frac{d\log M^h(c^h, z^h)}{d\log z^h}\right]
\end{aligned} \tag{2.127}$$

where here we have used that $\widetilde{R}'(z^h) = M^h(\widetilde{R}(z^h), z^h) > 0$

We conclude that $\frac{d\log \widetilde{R}'(z^h)}{d\log z^h} - \frac{d\log M^h(c^h, z^h)}{d\log c^h} \frac{d\log \widetilde{R}(z^h)}{d\log z^h} - \frac{d\log M^h(c^h, z^h)}{d\log z^h} \geq 0$, implying $\varepsilon^h(\widetilde{R}) \geq 0$.

Second derivative of labor supply

To compute the second derivative of labor supply, we differentiate the first derivative along an arbitrary direction $\widetilde{\Delta} \in \mathbf{\Delta}$.

$$\begin{aligned}
D^2_{\widetilde{\Delta}\Delta} \log z^h(\widetilde{R}) &= D_{\widetilde{\Delta}} \left(\eta^h(\widetilde{R}) \frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))} + \varepsilon^h(\widetilde{R}) \frac{\Delta'(z^h(\widetilde{R}))}{\widetilde{R}'(z^h(\widetilde{R}))}\right) \\
&= D_{\widetilde{\Delta}} \left(\eta^h(\widetilde{R}^h)\right) \frac{\Delta^h}{\widetilde{R}^h} + \eta^h \left[\left(\frac{\Delta^{h\prime} z^h}{\widetilde{R}^h} - \frac{\Delta^h}{\widetilde{R}^h} \frac{d\log \widetilde{R}^h}{d\log z}\right) D_{\widetilde{\Delta}} \log z^h(\widetilde{R}) - \frac{\Delta^h}{\widetilde{R}^h} \frac{\widetilde{\Delta}^h}{\widetilde{R}^h}\right] \\
&+ D_{\widetilde{\Delta}} \left(\varepsilon^h(\widetilde{R}^h)\right) \frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}} + \varepsilon^h \left[\left(\frac{\Delta^{h\prime\prime} z^h}{\widetilde{R}^{h\prime}} - \frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}} \frac{d\log \widetilde{R}^h}{d\log z}\right) D_{\widetilde{\Delta}} \log z^h(\widetilde{R}) - \frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}} \frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}}\right]
\end{aligned} \tag{2.128}$$

where above, and in the equations below, variables are evaluated at $\widetilde{R}, z^h = z^h(\widetilde{R})$, and/or $c^h = \widetilde{R}(z^h)$ as relevant; and $\widetilde{R}^h$ is shorthand for $\widetilde{R}(z^h)$ and similar.

Next, we compute $D_{\widetilde{\Delta}}\varepsilon^h(\widetilde{R})$ and $D_{\widetilde{\Delta}}\varepsilon^h(\widetilde{R})$ in isolation by differentiating (2.125):

$$
\begin{aligned}
D_{\widetilde{\Delta}}\varepsilon^h(\widetilde{R}) = & -(\varepsilon^h)^2\left[\left(\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log c}\right)\left(\frac{\widetilde{\Delta}^h}{\widetilde{R}^h}+\frac{d\log\widetilde{R}^h}{d\log z}D_{\widetilde{\Delta}}\log z^h\right)+\frac{d}{d\log z}\left(\frac{d\log M^h}{d\log c}\right)D_{\widetilde{\Delta}}\log z^h\right)\frac{d\log\widetilde{R}^h}{d\log z}\right.\\
&+\frac{d\log M^h}{d\log c}\left(\frac{d}{d\log z}\left(\frac{d\log\widetilde{R}^h}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h+\frac{\widetilde{\Delta}^{h\prime}z^h}{\widetilde{R}^h}-\frac{d\log\widetilde{R}^h}{d\log z}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h}\right)\\
&+\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log z}\right)\left(\frac{\widetilde{\Delta}^h}{\widetilde{R}^h}+\frac{d\log\widetilde{R}^h}{d\log z}D_{\widetilde{\Delta}}\log z^h\right)+\frac{d}{d\log z}\left(\frac{d\log M^h}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h\\
&\left.-\left(\frac{d}{d\log z}\left(\frac{d\log\widetilde{R}^{h\prime}}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h+\frac{\widetilde{\Delta}^{h\prime\prime}z^h}{\widetilde{R}^{h\prime}}-\frac{d\log\widetilde{R}^{h\prime}}{d\log z}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}}\right)\right]\\
= &\quad \eta^h\varepsilon^h\left[\frac{d}{d\log z}\left(\frac{d\log\widetilde{R}^h}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h+\frac{\widetilde{\Delta}^{h\prime}z^h}{\widetilde{R}^h}-\frac{d\log\widetilde{R}^h}{d\log z}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h}\right]\\
&+(\varepsilon^h)^2\left[\frac{d}{d\log z}\left(\frac{d\log\widetilde{R}^{h\prime}}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h+\frac{\widetilde{\Delta}^{h\prime\prime}z^h}{\widetilde{R}^{h\prime}}-\frac{d\log\widetilde{R}^{h\prime}}{d\log z}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}}\right]\\
&+\varepsilon^h_{+0}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h}+\varepsilon^h_{+1}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}}\\
\text{where}\quad \varepsilon^h_{+0} = & -(\varepsilon^h)^2\left[\left(\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log c}\right)\left(1+\frac{d\log\widetilde{R}^h}{d\log z}\eta^h\right)+\frac{d}{d\log z}\left(\frac{d\log M^h}{d\log c}\right)\eta^h\right)\frac{d\log\widetilde{R}^h}{d\log z}\right.\\
&\left.+\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log z}\right)\left(1+\frac{d\log\widetilde{R}^h}{d\log z}\eta^h\right)+\frac{d}{d\log z}\left(\frac{d\log M^h}{d\log z}\right)\eta^h\right]\\
\varepsilon^h_{+1} = & -(\varepsilon^h)^3\left[\left(\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log c}\right)\frac{d\log\widetilde{R}^h}{d\log z}+\frac{d}{d\log z}\left(\frac{d\log M^h}{d\log c}\right)\right)\frac{d\log\widetilde{R}^h}{d\log z}\right.\\
&\left.+\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log z}\right)\frac{d\log\widetilde{R}^h}{d\log z}+\frac{d}{d\log z}\left(\frac{d\log M^h}{d\log z}\right)\right]
\end{aligned}
\tag{2.129}
$$

Note that the terms $\eta^h\varepsilon^h\frac{d}{d\log z}\left(\frac{d\log\widetilde{R}^h}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h$, $(\varepsilon^h)^2\frac{d}{d\log z}\left(\frac{d\log\widetilde{R}'^h}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h$, $\varepsilon^h_{+0}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h}$, and $\varepsilon^h_{+1}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}}$ correspond to changes in $\varepsilon^h(\widetilde{R})$ through changes in labor supply, whereas the complementary terms correspond to changes in $\varepsilon^h(\widetilde{R})$ directly through changes in the tax schedule. By the earlier expression for $D_{\widetilde{\Delta}}\log z^h$, we can decompose the terms operating through changes in labor supply into income and compensated effects. For later use, we define ε^{+h} to be the compensated component, i.e.

$$
\varepsilon^{+h}(\widetilde{R}) \equiv \varepsilon^h_{+1}(\widetilde{R})+\eta^h(\varepsilon^h)^2\frac{d^2\log\widetilde{R}(z)}{d\log z^2}+(\varepsilon^h)^3\frac{d^2\log\widetilde{R}'(z)}{d\log z^2} \tag{2.130}
$$

$$
\begin{aligned}
D_{\widetilde{\Delta}}\eta^h(\widetilde{R}) =& -\frac{d\log M^h}{d\log c} D_{\widetilde{\Delta}}\varepsilon^h(\widetilde{R}) - \varepsilon^h D_{\widetilde{\Delta}}\frac{d\log M^h}{d\log c} \\
=& \quad (\eta^h)^2 \left[\frac{d}{d\log z}\left(\frac{d\log \widetilde{R}^h}{d\log z}\right) D_{\widetilde{\Delta}}\log z^h + \frac{\widetilde{\Delta}' z^h}{\widetilde{R}} - \frac{d\log \widetilde{R}^h}{d\log z}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h}\right] \\
& + \eta^h\varepsilon^h \left[\frac{d}{d\log z}\left(\frac{d\log \widetilde{R}^{h\prime}}{d\log z}\right) D_{\widetilde{\Delta}}\log z^h + \frac{\widetilde{\Delta}^{h\prime\prime} z^h}{\widetilde{R}^{h\prime}} - \frac{d\log \widetilde{R}^{h\prime}}{d\log z}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}}\right] \\
& + \eta^h_{+0}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h} + \eta^h_{+1}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}} \\
\text{where}\quad \eta^h_{+0} =& -\eta^h\varepsilon^h\left[\left(\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log c}\right)\left(1+\frac{d\log \widetilde{R}^h}{d\log z}\eta^h\right) + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log c}\right)\eta^h\right)\frac{d\log \widetilde{R}^h}{d\log z}\right. \\
& \left. + \frac{d}{d\log c}\left(\frac{d\log M^h}{d\log z}\right)\left(1+\frac{d\log \widetilde{R}^h}{d\log z}\eta^h\right) + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log z}\right)\eta^h\right] \\
& - \varepsilon^h\left[\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log c}\right)\left(1+\frac{d\log \widetilde{R}^h}{d\log z}\eta^h\right) + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log c}\right)\eta^h\right] \\
=& -\left(\eta^h\frac{d\log \widetilde{R}^h}{d\log z}+1\right)\varepsilon^h\left[\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log c}\right)\left(1+\frac{d\log \widetilde{R}^h}{d\log z}\eta^h\right) + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log c}\right)\eta^h\right] \\
& - \eta^h\varepsilon^h\left[\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log z}\right)\left(1+\frac{d\log \widetilde{R}^h}{d\log z}\eta^h\right) + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log z}\right)\eta^h\right] \\
\eta^h_{+1} =& -\eta^h(\varepsilon^h)^2\left[\left(\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log c}\right)\frac{d\log \widetilde{R}^h}{d\log z} + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log c}\right)\right)\frac{d\log \widetilde{R}^h}{d\log z}\right. \\
& \left. + \frac{d}{d\log c}\left(\frac{d\log M^h}{d\log z}\right)\frac{d\log \widetilde{R}^h}{d\log z} + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log z}\right)\right] \\
& - (\varepsilon^h)^2\left[\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log c}\right)\frac{d\log \widetilde{R}^h}{d\log z} + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log c}\right)\right] \\
=& -\left(\eta^h\frac{d\log \widetilde{R}^h}{d\log z}+1\right)(\varepsilon^h)^2\left[\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log c}\right)\frac{d\log \widetilde{R}^h}{d\log z} + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log c}\right)\right] \\
& - \eta^h(\varepsilon^h)^2\left[\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log z}\right)\frac{d\log \widetilde{R}^h}{d\log z} + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log z}\right)\right] \\
=& -(\varepsilon^h)^2\left[\left(\frac{d}{d\log c}\left(\frac{d\log M^h}{d\log c}\right)\left(1+\frac{d\log \widetilde{R}^h}{d\log z}\eta^h\right) + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log c}\right)\eta^h\right)\frac{d\log \widetilde{R}^h}{d\log z}\right. \\
& \left. + \frac{d}{d\log c}\left(\frac{d\log M^h}{d\log z}\right)\left(1+\frac{d\log \widetilde{R}^h}{d\log z}\eta^h\right) + \frac{d}{d\log z}\left(\frac{d\log M^h}{d\log z}\right)\eta^h\right]
\end{aligned}
\tag{2.131}
$$

Note that $\varepsilon^h_{+0} = \eta^h_{+1}$. This reflects that they are both (essentially) second derivatives, but with opposite orders of differentiation.

Finally, we substitute these expressions back into (2.128). Simplifying, we obtain:

$$
\begin{aligned}
& D^2_{\widetilde{\Delta\Delta}}\log z^h(\widetilde{R}) \\
&= \eta^h\left[\left(\frac{\Delta^{h\prime}z^h}{\widetilde{R}^h} - \frac{\Delta^h}{\widetilde{R}^h}\frac{d\log \widetilde{R}^h}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h - \frac{\Delta^h}{\widetilde{R}^h}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h}\right] + \varepsilon^h\left[\left(\frac{\Delta^{h\prime\prime}z^h}{\widetilde{R}^{h\prime}} - \frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}}\frac{d\log \widetilde{R}^{h\prime}}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h - \frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}}\right] \\
&+ (\eta^h)^2\frac{\Delta^h}{\widetilde{R}^h}\left[\frac{d}{d\log z}\left(\frac{d\log \widetilde{R}^h}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h + \frac{\widetilde{\Delta}^{h\prime}z^h}{\widetilde{R}^h} - \frac{d\log \widetilde{R}^h}{d\log z}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h}\right] + \eta^h\varepsilon^h\frac{\Delta^h}{\widetilde{R}^h}\left[\frac{d}{d\log z}\left(\frac{d\log \widetilde{R}^{h\prime}}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h + \frac{\widetilde{\Delta}^{h\prime\prime}z^h}{\widetilde{R}^{h\prime}} - \frac{d\log \widetilde{R}^{h\prime}}{d\log z}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}}\right] \\
&+ \eta^h\varepsilon^h\frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}}\left[\frac{d}{d\log z}\left(\frac{d\log \widetilde{R}^h}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h + \frac{\widetilde{\Delta}^{h\prime}z^h}{\widetilde{R}^h} - \frac{d\log \widetilde{R}^h}{d\log z}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h}\right] + (\varepsilon^h)^2\frac{\Delta^{\prime h}}{\widetilde{R}^{\prime h}}\left[\frac{d}{d\log z}\left(\frac{d\log \widetilde{R}^{h\prime}}{d\log z}\right)D_{\widetilde{\Delta}}\log z^h + \frac{\widetilde{\Delta}^{h\prime\prime}z^h}{\widetilde{R}^{h\prime}} - \frac{d\log \widetilde{R}^{h\prime}}{d\log z}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}}\right] \\
&+ \eta^h_{+0}\frac{\Delta^h}{\widetilde{R}^h}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h} + \eta^h_{+1}\frac{\Delta^h}{\widetilde{R}^h}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}} + \varepsilon^h_{+0}\frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}}\frac{\widetilde{\Delta}^h}{\widetilde{R}^h} + \varepsilon^h_{+1}\frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}}\frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}}
\end{aligned}
\tag{2.132}
$$

$$
\begin{aligned}
&D^2_{\widetilde{\Delta}\Delta} \log z^h(\widetilde{R}) \\
&= \left[\eta^h \left(- \frac{d \log \widetilde{R}^h}{d \log z} \eta^h - 1 \right) + (\eta^h)^2 \left(\frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^h}{d \log z} \right) \eta^h - \frac{d \log \widetilde{R}^h}{d \log z} \right) + \eta^h \varepsilon^h \frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^{h\prime}}{d \log z} \right) \eta^h + \eta^h_{+0} \right] \frac{\Delta^h}{\widetilde{R}^h} \frac{\widetilde{\Delta}^h}{\widetilde{R}^h} \\
&+ \left[-\eta^h \frac{d \log \widetilde{R}^h}{d \log z} \varepsilon^h + (\eta^h)^2 \left(\frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^h}{d \log z} \right) \varepsilon^h + \frac{d \log \widetilde{R}^h}{d \log z} \right) + \eta^h \varepsilon^h \left(\frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^{h\prime}}{d \log z} \right) \varepsilon^h - \frac{d \log \widetilde{R}^{h\prime}}{d \log z} \right) + \eta^h_{+1} \right] \frac{\Delta^h}{\widetilde{R}^h} \frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}} \\
&+ \left[(\eta^h)^2 \frac{d \log \widetilde{R}^h}{d \log z} - \varepsilon^h \eta^h \frac{d \log \widetilde{R}^{h\prime}}{d \log z} + \eta^h \varepsilon^h \left(\frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^h}{d \log z} \right) \eta^h - \frac{d \log \widetilde{R}}{d \log z} \right) + (\varepsilon^h)^2 \frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^{h\prime}}{d \log z} \right) \eta^h + \varepsilon^h_{+0} \right] \frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}} \frac{\widetilde{\Delta}^h}{\widetilde{R}^h} \\
&+ \left[\eta^h \varepsilon^h \frac{d \log \widetilde{R}^h}{d \log z} - \varepsilon^h \left(\frac{d \log \widetilde{R}^{h\prime}}{d \log z} \varepsilon^h + 1 \right) + \eta^h \varepsilon^h \left(\frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^h}{d \log z} \right) \varepsilon^h + \frac{d \log \widetilde{R}^h}{d \log z} \right) + (\varepsilon^h)^2 \left(\frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^{h\prime}}{d \log z} \right) \varepsilon^h - \frac{d \log \widetilde{R}^{h\prime}}{d \log z} \right) + \varepsilon^h_{+1} \right] \frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}} \frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}} \\
&+ \quad \eta^h \varepsilon^h \frac{\Delta^{h\prime\prime} z^h}{\widetilde{R}'} \frac{\widetilde{\Delta}^h}{\widetilde{R}^h} \quad + \quad (\varepsilon^h)^2 \frac{\Delta^{h\prime\prime} z^h}{\widetilde{R}'} \frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}} \quad + \quad \eta^h \varepsilon^h \frac{\Delta^h}{\widetilde{R}^h} \frac{\widetilde{\Delta}^{h\prime\prime} z^h}{\widetilde{R}'} \quad + \quad (\varepsilon^h)^2 \frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}} \frac{\widetilde{\Delta}^{h\prime\prime} z^h}{\widetilde{R}'}
\end{aligned}
\tag{2.133}
$$

$$
\begin{aligned}
&D^2_{\widetilde{\Delta}\Delta} \log z^h(\widetilde{R}) \\
&= \left[-\eta^h - 2 \frac{d \log \widetilde{R}^h}{d \log z} (\eta^h)^2 + \frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^h}{d \log z} \right) (\eta^h)^3 + \frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^{h\prime}}{d \log z} \right) (\eta^h)^2 \varepsilon^h + \eta^h_{+0} \right] \frac{\Delta^h}{\widetilde{R}^h} \frac{\widetilde{\Delta}^h}{\widetilde{R}^h} \\
&+ \left[\frac{d \log \widetilde{R}^h}{d \log z} (\eta^h)^2 - \left(\frac{d \log \widetilde{R}^h}{d \log z} + \frac{d \log \widetilde{R}^{h\prime}}{d \log z} \right) \eta^h \varepsilon^h + \frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^h}{d \log z} \right) (\eta^h)^2 \varepsilon^h + \frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^{h\prime}}{d \log z} \right) \eta^h (\varepsilon^h)^2 + \underbrace{\eta^h_{+1}}_{=\varepsilon^h_{+0}} \right] \left(\frac{\Delta^h}{\widetilde{R}^h} \frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}} + \frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}} \frac{\widetilde{\Delta}^h}{\widetilde{R}^h} \right) \\
&+ \left[-\varepsilon^h + 2 \frac{d \log \widetilde{R}^h}{d \log z} \eta^h \varepsilon^h - 2 \frac{d \log \widetilde{R}'}{d \log z} (\varepsilon^h)^2 + \frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^h}{d \log z} \right) \eta^h (\varepsilon^h)^2 + \frac{d}{d \log z} \left(\frac{d \log \widetilde{R}^{h\prime}}{d \log z} \right) (\varepsilon^h)^3 + \varepsilon^h_{+1} \right] \frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}} \frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}} \\
&+ \quad \eta^h \varepsilon^h \left(\frac{\Delta^h}{\widetilde{R}^h} \frac{\widetilde{\Delta}^{h\prime\prime} z^h}{\widetilde{R}'} + \frac{\Delta^{h\prime\prime} z^h}{\widetilde{R}'} \frac{\widetilde{\Delta}^h}{\widetilde{R}^h} \right) \quad + \quad (\varepsilon^h)^2 \left(\frac{\Delta^{h\prime}}{\widetilde{R}^{h\prime}} \frac{\widetilde{\Delta}^{h\prime\prime} z^h}{\widetilde{R}'} + \frac{\Delta^{h\prime\prime} z^h}{\widetilde{R}'} \frac{\widetilde{\Delta}^{h\prime}}{\widetilde{R}^{h\prime}} \right)
\end{aligned}
\tag{2.134}
$$

Measurability of labor supply derivatives

Fix any $\widetilde{R} \in R + B_\delta(\mathbf{0})$ and $\Delta, \widetilde{\Delta} \in \mathbf{\Delta}$. We wish to show $D_\Delta z^h(\widetilde{R})$ and $D^2_{\widetilde{\Delta}\Delta} z^h(\widetilde{R})$ are measurable in h. Because

$$
\begin{aligned}
D_\Delta z^h(\widetilde{R}) &= z^h(\widetilde{R}) D_\Delta \log z^h(\widetilde{R}) \\
D^2_{\widetilde{\Delta}\Delta} z^h(\widetilde{R}) &= z^h(\widetilde{R}) \left[D_{\widetilde{\Delta}} \log z^h(\widetilde{R}) D_\Delta \log z^h(\widetilde{R}) + D_{\widetilde{\Delta}\Delta} \log z^h(\widetilde{R}) \right],
\end{aligned}
\tag{2.135}
$$

and we have shown $z^h(\widetilde{R})$ is measurable, it suffices to show that $D_\Delta \log z^h(\widetilde{R})$ and $D^2_{\widetilde{\Delta}\Delta} \log z^h(\widetilde{R})$ are measurable in h.

We begin with $D_\Delta \log z^h(\widetilde{R})$. Recall our expression, (2.126), for this term. Since $z^h(\widetilde{R})$ is measurable, Δ and $\widetilde{R}$ are continuous in z, and $\widetilde{R}(z^h(\widetilde{R})) > 0$ and $\widetilde{R}'(z^h(\widetilde{R})) > 0$ for all $h \in \mathcal{H}$, $\frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))}$ and $\frac{\Delta'(z^h(\widetilde{R}))}{\widetilde{R}'(z^h(\widetilde{R}))}$ are continuous in $z^h(\widetilde{R})$; so they are measurable. It therefore remains to show that $\eta^h(\widetilde{R})$ and $\varepsilon^h(\widetilde{R})$ are also measurable.

To do so, recall the definitions of $\eta^h(\widetilde{R})$ and $\varepsilon^h(\widetilde{R})$ in (2.125). Their tax curvature terms, $\frac{d \log \widetilde{R}(z^h(\widetilde{R}))}{d \log z^h}$ and $\frac{d \log \widetilde{R}'(z^h(\widetilde{R}))}{d \log z^h}$, are measurable because the tax schedule is three-times continuously differentiable by Assumption 1, $\widetilde{R}(z^h(\widetilde{R}))$ and $\widetilde{R}'(z^h(\widetilde{R})) > 0$ for all $h \in \mathcal{H}$, and $z^h(\widetilde{R})$ is measurable. Their MRS curvature terms are measurable because they are continuous functions of the first and second derivatives of preferences, both of which are continuous and measurable in (h, c, z)—and because for any (c, z)-continuous and (h, c, z)-

measurable function $f(h,c,z)$, $f(h,\widetilde{R}(z^h(\widetilde{R})),z^h(\widetilde{R}))$ is measurable in h.[86] Since—as we have also already argued—the denominators of both expressions in (2.125) are non-zero for all $h\in\mathcal{H}$, both elasticities are continuous functions of measurable functions and so are measurable.

We now argue that $D^2_{\Delta\widetilde{\Delta}}\log z^h(\widetilde{R})$ is measurable in h. Recall our expression, (2.134), for this term. On top of the terms we have already considered in showing the first derivative is measurable, we now must also show that $\frac{\Delta''(z^h(\widetilde{R}))z^h(\widetilde{R})}{\widetilde{R}'(z^h(\widetilde{R}))}$, $\frac{\widetilde{\Delta}''(z^h(\widetilde{R}))z^h(\widetilde{R})}{\widetilde{R}'(z^h(\widetilde{R}))}$, and each of the super-elasticities are measurable. $\frac{\Delta''(z^h(\widetilde{R}))z^h(\widetilde{R})}{\widetilde{R}'(z^h(\widetilde{R}))}$ (and similarly for $\widetilde{\Delta}$) is measurable because Δ'' and $\widetilde{R}'$ are continuous and $z^h(\widetilde{R})$ are measurable. To show the super-elasticities are measurable, we refer to their definitions in (2.129) and (2.131). All of the elasticities are continuous functions of functions which we have already shown to be measurable, except for $\frac{d}{d\log c}\left(\frac{d\log M^h(\widetilde{R}(z^h(\widetilde{R})),z^h(\widetilde{R}))}{d\log z}\right)$ and $\frac{d}{d\log c}\left(\frac{d\log M^h(\widetilde{R}(z^h(\widetilde{R})),z^h(\widetilde{R}))}{d\log z}\right)$. Note that these are both continuous functions of the first, second, and third derivatives of preferences, all of which are continuous and measurable in (h,c,z). By our earlier observations, this implies these terms are measurable as well.

Properties of post-tax income

It remains to show that, for $h\in\mathcal{H},\widetilde{R}\in R+B_\delta(\mathbf{0})$, post-tax income $\widetilde{R}(z^h(\widetilde{R}))$ is well-defined, measurable, and strictly positive, and that it has two continuous and $\mathcal{H}$-measurable Frechet derivatives in $\widetilde{R}$.

These properties are all inherited from those of $z^h(\widetilde{R})$. Well-definedness is immediate. Measurability follows from that $z^h(\widetilde{R})$ is measurable and $\widetilde{R}$ is continuous. Positivity follows from that $z^h(\widetilde{R})>0$, our earlier observation that $R(z)>0$ for all $z>0$, and that $\delta<\frac{1}{2}$. Finally, the existence, continuity, and measurability of Frechet derivatives follows from that $\widetilde{R}$ is twice-continuously differentiable and that $z^h(\widetilde{R})$ has two continuous and $\mathcal{H}$-measurable Frechet derivatives.

2.9.6.3 Proof of Lemma 3

We complete the proof in several steps: First we provide bounds on the first and second derivatives of log labor supply, second we show that labor supply and its derivatives are integrable, and third we show that post-tax income and its derivatives are bounded and integrable. Throughout we fix $\delta>0$ to be smaller than the values referred to in both Assumption 4 and Lemma 2.

First and second derivatives of labor supply: bounding

We now show that there exist uniform upper bounds on the Frechet derivatives $D\log z^h(\widetilde{R})$ and $D^2\log z^h(\widetilde{R})$ (as linear maps) across all $\widetilde{R}\in R+B_\delta(\mathbf{0})$, $h\in\mathcal{H}$. WLOG we will fix $\Delta,\widetilde{\Delta}\in\mathbf{\Delta}$ with $||\Delta||=||\widetilde{\Delta}||=1$ and the show existence of uniform bounds on $D_\Delta\log z^h(\widetilde{R})$ and $D^2_{\Delta\widetilde{\Delta}}\log z^h(\widetilde{R})$.

By the expressions (2.126) and (2.134) for the first and second derivatives of labor supply, it suffices to provide uniform bounds on

- all elasticities and super-elasticities;
- across all $z>0$, the log-tax-change terms $\frac{\Delta(z)}{\widetilde{R}(z)}$, $\frac{\Delta'(z)}{\widetilde{R}'(z)}$, and $\frac{\Delta''(z)z}{\widetilde{R}'(z)}$ (those for $\widetilde{\Delta}$ are analogous); and

[86] For any measurable set A in the image of f, $f^{-1}(A)$ is a measurable set consisting of (h,c,z) triples. Since $\widetilde{R}(z^h(\widetilde{R}))$ and $z^h\widetilde{R}$ are measurable, and since component-wise measurable functions are measurable with respect to the product measure, the set of triples (h_1,h_2,h_3) such that $\left(h_1,\widetilde{R}(z^{h_2}(\widetilde{R})),z^{h_3}(\widetilde{R})\right)\in f^{-1}(A)$ is also measurable. It remains to show that the diagonal of the set of these triples (h_1,h_2,h_3) is a measurable set in $\mathcal{H}$. For this it suffices to show that the diagonal function $h\mapsto(h,h,h)$ is measurable. To see this note that it suffices to check on a generic generating element of the product measure, e.g. $H\times H'\times H''$, whose inverse image is simply $H\cap H\cap H'\cap H''$; this is measurable if H, H', and H'' each are.

- across all $z > 0$, the tax curvature terms $\frac{d\log\widetilde{R}(z)}{d\log z}$, $\frac{d\log\widetilde{R}'(z)}{d\log z}$, $\frac{d}{d\log z}\left(\frac{d\log\widetilde{R}(z)}{d\log z}\right)$, and $\frac{d}{d\log z}\left(\frac{d\log\widetilde{R}'(z)}{d\log z}\right)$.

The first bullet is ensured by Assumption 4.

Next, consider the second bullet. Since (from the Lemma we are trying to prove) we need only show this is true for sufficiently small δ, we may WLOG assume $\delta < \frac{1}{2}$. In this case, $\widetilde{R} \in B_\delta(\mathbf{0})$ and $||\Delta||= 1$ imply

$$\left|\frac{\Delta(z)}{\widetilde{R}(z)}\right| \leq \left|\frac{\Delta(z)}{\frac{1}{2}R(z)}\right| \leq 2, \qquad \left|\frac{\Delta'(z)}{\widetilde{R}'(z)}\right| \leq \left|\frac{\Delta'(z)}{\frac{1}{2}R'(z)}\right| \leq 2, \qquad \left|\frac{\Delta''(z)z}{\widetilde{R}'(z)}\right| \leq B^R\left|\frac{R'(z)}{\widetilde{R}'(z)}\right| \leq B^R\left|\frac{R'(z)}{\frac{1}{2}R'(z)}\right| = 2B^R. \tag{2.136}$$

Finally, consider the third bullet. Again, using that $\delta < \frac{1}{2}$, we have

$$\begin{aligned}
\left|\frac{d\log\widetilde{R}(z)}{d\log z}\right| &= \left|\frac{\widetilde{R}'(z)z}{\widetilde{R}(z)}\right| \leq \left|\frac{\frac{3}{2}R'(z)z}{\frac{1}{2}R(z)}\right| \leq 3B^R,\\
\left|\frac{d\log\widetilde{R}'(z)}{d\log z}\right| &= \left|\frac{\widetilde{R}''(z)z}{\widetilde{R}'(z)}\right| \leq \left|\frac{\frac{3}{2}R''(z)z}{\frac{1}{2}R'(z)}\right| \leq 3B^R\\
\left|\frac{d}{d\log z}\left(\frac{d\log\widetilde{R}(z)}{d\log z}\right)\right| &= \left|\frac{\widetilde{R}''(z)z^2}{\widetilde{R}(z)} + \frac{\widetilde{R}'(z)z}{\widetilde{R}(z)} - \left(\frac{\widetilde{R}'(z)z}{\widetilde{R}(z)}\right)^2\right| \leq \left|\frac{\widetilde{R}''(z)z^2}{\widetilde{R}(z)}\right| + \left|\frac{\widetilde{R}'(z)z}{\widetilde{R}(z)}\right| + \left|\frac{\widetilde{R}'(z)z}{\widetilde{R}(z)}\right|^2\\
&\leq \left|\frac{(B^R+\frac{1}{2})R'(z)z}{\frac{1}{2}R(z)}\right| + 3B^R + 9(B^R)^2 \leq (2B^R+1)B^R 3B^R + 9(B^R)^2\\
\left|\frac{d}{d\log z}\left(\frac{d\log\widetilde{R}'(z)}{d\log z}\right)\right| &= \left|\frac{\widetilde{R}'''(z)z^2}{\widetilde{R}'(z)} + \frac{\widetilde{R}''(z)z}{\widetilde{R}'(z)} - \left(\frac{\widetilde{R}''(z)z}{\widetilde{R}'(z)}\right)^2\right| \leq \left|\frac{\frac{d^2\widetilde{R}'(z)}{d\log z^2}}{\widetilde{R}'(z)}\right| + 2\left|\frac{\widetilde{R}''(z)z}{\widetilde{R}'(z)}\right| + \left|\frac{\widetilde{R}''(z)z}{\widetilde{R}'(z)}\right|^2\\
&\leq \left|\frac{(B^R+\frac{1}{2})R'(z)}{\frac{1}{2}R'(z)}\right| + 6B^R + 9(B^R)^2 \leq (2B^R+1) + 6B^R + 9(B^R)^2
\end{aligned} \tag{2.137}$$

where on the penultimate line, we have used that $\frac{d^2\widetilde{R}'(z)}{d\log z^2} = z\widetilde{R}''(z) + z^2\widetilde{R}'''(z)$.

We conclude that the level and first two derivatives of log labor supply (appropriately normalized) are uniformly bounded across $h \in \mathcal{H}$ and $\widetilde{R} \in R + B_\delta(\mathbf{0})$.

Integrability of labor supply and its derivatives

In this section we argue that—for any $\widetilde{R} \in R + B_\delta(\mathbf{0})$, $\Delta, \widetilde{\Delta} \in \boldsymbol{\Delta}$—$z^h(\widetilde{R})$, $D_\Delta z^h(\widetilde{R})$, and $D^2_{\widetilde{\Delta}\Delta} z^h(\widetilde{R})$ are bounded across all $h \in \mathcal{H}$ by constants times z_0^h and are integrable. Since by Lemma 2, $z^h(\widetilde{R})$, $D_\Delta z^h(\widetilde{R})$, and $D^2_{\widetilde{\Delta}\Delta} z^h(\widetilde{R})$ are measurable, and by Assumption 4, z_0^h is integrable, the bounding is sufficient for integrability. To this end, a useful observation is that:

$$\begin{aligned}
D_\Delta z^h(\widetilde{R}) &= z^h(\widetilde{R}) D_\Delta \log z^h(\widetilde{R})\\
D^2_{\widetilde{\Delta}\Delta} z^h(\widetilde{R}) &= z^h(\widetilde{R})\left[D_{\widetilde{\Delta}} \log z^h(\widetilde{R}) D_\Delta \log z^h(\widetilde{R}) + D_{\widetilde{\Delta}\Delta} \log z^h(\widetilde{R})\right]
\end{aligned} \tag{2.138}$$

Since we have shown that $D_{\widetilde{\Delta}} \log z^h(\widetilde{R})$, $D_\Delta \log z^h(\widetilde{R})$, and $D_{\widetilde{\Delta}\Delta} \log z^h(\widetilde{R})$ are uniformly bounded, it remains to show $z^h(\widetilde{R})$ is bounded in absolute value by a constant times $z_0^h = z^h(R)$.

Indeed, for all $\widetilde{R} \in R + B_\delta(\mathbf{0})$, since $z^h(\widetilde{R}) > 0$ and $z^h(\widetilde{R})$ is continuously differentiable, we may apply the fundamental theorem of calculus along a path between R and $\widetilde{R}$, giving us

$$\begin{aligned}
\log z^h(\widetilde{R}) &= \log z^h(R) + \int_0^{||\widetilde{R}-R||} D_{\widetilde{R}-R} \log z^h\left(R + \alpha\frac{\widetilde{R}-R}{||\widetilde{R}-R||}\right) d\alpha\\
\left|\log z^h(\widetilde{R}) - \log z^h(R)\right| &= ||\widetilde{R}-R|| 2M \leq 2M\delta
\end{aligned} \tag{2.139}$$

where $D_{\widetilde{R}-R}$ is the Frechet along the path between R and $\widetilde{R}$,[87] and where we have used the bounds on $\log z^h(\widehat{R})$'s first derivative derived in the previous step. We conclude that $|z^h(\widetilde{R})| \leq e^{2\delta M}|z^h(R)| = e^{2\delta M} z^h(R)$, i.e. $z^h(R)$ bounds $z^h(\widetilde{R})$ as desired.

Integrability of retained income and its derivatives

In this section we argue that—for any $\widetilde{R} \in R + B_\delta(\mathbf{0})$, $\Delta, \widetilde{\Delta} \in \mathbf{\Delta}$—$c^h(\widetilde{R}) = \widetilde{R}(z^h(\widetilde{R}))$, $D_\Delta \widetilde{R}(z^h(\widetilde{R}))$, and $D^2_{\widetilde{\Delta}\Delta}\widetilde{R}(z^h(\widetilde{R}))$ are bounded by c^h_0 and are integrable. WLOG, we show this in the case where $||\Delta|| = ||\widetilde{\Delta}|| = 1$. Since by Lemma 2 and the continuity of R and $\widetilde{R}$, we know that $c^h_0 = R(z^h(R))$, $c^h(\widetilde{R}) = \widetilde{R}(z^h(\widetilde{R}))$, $D_\Delta \widetilde{R}(z^h(\widetilde{R}))$, and $D^2_{\widetilde{\Delta}\Delta}\widetilde{R}(z^h(\widetilde{R}))$ are measurable, and since by Assumption 4 c^h_0 is integrable, the bounding is sufficient for integrability. To this end, a useful observation is that:

$$
\begin{aligned}
D_\Delta \log \widetilde{R}(z^h(\widetilde{R})) &= \frac{d \log \widetilde{R}(z^h(\widetilde{R}))}{d \log z^h} D_\Delta \log z^h(\widetilde{R}) + \frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))} \\
D^2_{\widetilde{\Delta}\Delta} \log \widetilde{R}(z^h(\widetilde{R})) &= \frac{d \log \widetilde{R}(z^h(\widetilde{R}))}{d \log z^h} \left(\left(\frac{d \log \widetilde{R}'(z^h(\widetilde{R}))}{d \log z^h} + 1 - \frac{d \log \widetilde{R}(z^h(\widetilde{R}))}{d \log z^h} \right) D_{\widetilde{\Delta}} \log z^h(\widetilde{R}) D_\Delta \log z^h(\widetilde{R}) + D_{\widetilde{\Delta}\Delta} \log z^h(\widetilde{R}) \right) \\
&\quad + \frac{d \log \widetilde{R}(z^h(\widetilde{R}))}{d \log z^h} \left(\left(\frac{\widetilde{\Delta}'(z^h(\widetilde{R}))}{\widetilde{R}'(z^h(\widetilde{R}))} - \frac{\widetilde{\Delta}(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))} \right) D_\Delta \log z^h(\widetilde{R}) + \left(\frac{\Delta'(z^h(\widetilde{R}))}{\widetilde{R}'(z^h(\widetilde{R}))} - \frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))} \right) D_{\widetilde{\Delta}} \log z^h(\widetilde{R}) \right) \\
&\quad - \frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))} \frac{\widetilde{\Delta}(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))}
\end{aligned}
\tag{2.140}
$$

We have already provided—in earlier steps of the proof—uniform bounds (across $h \in \mathcal{H}$ and $\widetilde{R} \in R + B_\delta(\mathbf{0})$) on $\frac{d \log \widetilde{R}(z^h(\widetilde{R}))}{d \log z^h}$, $\frac{d \log \widetilde{R}'(z^h(\widetilde{R}))}{d \log z^h}$, $\frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))}$, $\frac{\Delta'(z^h(\widetilde{R}))}{\widetilde{R}'(z^h(\widetilde{R}))}$, $\frac{\widetilde{\Delta}(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))}$, $\frac{\widetilde{\Delta}'(z^h(\widetilde{R}))}{\widetilde{R}'(z^h(\widetilde{R}))}$, $D_{\widetilde{\Delta}} \log z^h(\widetilde{R})$, $D_\Delta \log z^h(\widetilde{R})$, and $D_{\widetilde{\Delta}\Delta} \log z^h(\widetilde{R})$, i.e. all of the terms above. The same argument as in the previous step of this proof (connecting the bounds on $\log z^h(\widetilde{R})$'s derivatives to those on $z^h(\widetilde{R})$'s derivatives) then implies that the first two Frechet derivatives of $c^h(\widetilde{R}) = \widetilde{R}(z^h(\widetilde{R}))$ are bounded across $\widetilde{R} \in R + B_\delta(\mathbf{0})$ by c^h_0 times a constant which is uniform across $h \in \mathcal{H}$.

2.9.6.4 Proof of Lemma 4

Here we show that within some neighborhood around R, $w^h \circ V^h(\widetilde{R})$ is finite, $\mathcal{H}$-measurable, and has two continuous and $\mathcal{H}$-measurable Frechet derivative in $\mathcal{R}$.

To begin, fix a standard social objective $((w^h)_{h \in \mathcal{H}}, G)$ and take $\hat{\delta} > 0$ small enough that Lemma 2 applies (so the definition of a standard social objective is well-defined), and so that conditions in the definition of a standard social objective hold.

Fixing any $h \in \mathcal{H}$, we now proceed to verify the conditions in the Lemma statement. First, finiteness is immediate from the definition of a standard social objective. Next, existence and continuity of derivatives: By Assumption 2 and the positivity of consumption and labor supply (Lemma 2), u^h's first two derivatives in (c, z) are continuous at all $c^h(\widetilde{R}), z^h(\widetilde{R})$ with $\widetilde{R} = R + B_\delta(\mathbf{0})$. Since by Lemma 2, $z^h(\widetilde{R})$ and $c^h(\widetilde{R})$ and their first two derivatives are continuous in $\widetilde{R} \in R + B_\delta(\mathbf{0})$, this is therefore also true of $V^h(\widetilde{R}) = u^h(c^h(\widetilde{R}), z^h(\widetilde{R}))$. Since by the argument above $V^h(\widetilde{R})$ is finite on this domain, the definition of a standard social objective implies that $w^h \circ V^h(\widetilde{R})$ is twice-continuously differentiable in $\widetilde{R} \in R + B_\delta(\mathbf{0})$. Finally, measurability of levels and derivatives: This is stated directly in the definition of a standard social objective.

[87] Note this is defined since $\widetilde{R} - R \in \Delta$

2.9.6.5 Proof of Lemma 5

We complete the proof in several steps. We begin by deriving explicit formulas for the first two Frechet derivatives of $w^h \circ V^h(\widetilde{R})$, and then argue that each is integrable.

First derivative of welfare

Throughout the proof we fix δ small enough that Lemmas 2 and 4 and the conditions in Definitions 1 and 3 of standard and regular social objectives apply.

By Lemma 4, we have that for all $h \in \mathcal{H}$, $\widetilde{R} \in R + B_\delta(\mathbf{0})$ $w^h \circ V^h(\widetilde{R})$ is twice-continuously Frechet differentiable. Since $w^h \circ V^h(\widetilde{R}) = w^h \circ u^h(\widetilde{c}^h(\widetilde{R}), z^h(\widetilde{R}))$, and by Lemma 2 $c^h(\widetilde{R})$ and $z^h(\widetilde{R})$ are twice-continuously Frechet differentiable, it is straightforward to compute the derivatives of $w^h \circ V^h$ (which exist by Assumption 2 and the definition of a standard social objective.)

For the first derivative, we have, for any $\Delta \in \mathbf{\Delta}$,

$$\begin{aligned}
D_\Delta w^h \circ V^h(\widetilde{R}) &= (w^h \circ u^h)_c\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right)\widetilde{R}(z^h(\widetilde{R}))D_\Delta \log \widetilde{R}(z^h(\widetilde{R}))\\
&\quad + (w^h \circ u^h)_z\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right) z^h(\widetilde{R})D_\Delta \log z^h(\widetilde{R})\\
&= (w^h \circ u^h)_c\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right)\widetilde{R}(z^h(\widetilde{R}))\left[\cancel{\frac{d\log \widetilde{R}(z^h(\widetilde{R}))}{d\log z^h}D_\Delta \log z^h(\widetilde{R})} + \frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))}\right]\\
&\quad \cancel{+ (w^h \circ u^h)_z\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right) z^h(\widetilde{R})D_\Delta \log z^h(\widetilde{R})}\\
&= (w^h \circ u^h)_c\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right)\widetilde{R}(z^h(\widetilde{R}))\frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))},
\end{aligned} \tag{2.141}$$

where the cancellation is by the household's first-order condition for labor supply (see (2.106) in the proof of Lemma 2). This generates an intuitive envelope expression.

Second derivative of welfare

To compute the second derivative of each household h's contribution to welfare, we differentiate (2.141). We have, for any $\Delta, \widetilde{\Delta} \in \mathbf{\Delta}$,

$$\begin{aligned}
D^2_{\widetilde{\Delta}\Delta} w^h \circ V^h(\widetilde{R}) &= \Bigg[(w^h \circ u^h)_{cc}\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right)\widetilde{R}(z^h(\widetilde{R}))\left(\frac{d\log \widetilde{R}(z^h(\widetilde{R}))}{d\log z}D_{\widetilde{\Delta}}\log z^h(\widetilde{R}) + \frac{\widetilde{\Delta}(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))}\right)\\
&\qquad + (w^h \circ u^h)_{zc}\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right) z^h(\widetilde{R})D_{\widetilde{\Delta}}\log z^h(\widetilde{R})\Bigg]\Delta(z^h(\widetilde{R}))\\
&\quad + (w^h \circ u^h)_c\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right)\Delta'(z^h(\widetilde{R}))z^h(\widetilde{R})D_{\widetilde{\Delta}}\log z^h(\widetilde{R})\\
&= (w^h \circ u^h)_{cc}\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right)\widetilde{\Delta}(z^h(\widetilde{R}))\Delta(z^h(\widetilde{R}))\\
&\quad + (w^h \circ u^h)_c\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right)\frac{1}{\varepsilon^h(\widetilde{R})}\widetilde{R}'(z^h(\widetilde{R}))z^h(\widetilde{R})\\
&\qquad \left[\varepsilon^h(\widetilde{R})\frac{\Delta'(z^h(\widetilde{R}))}{\widetilde{R}'(z^h(\widetilde{R}))} + \eta^h(\widetilde{R})\frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))}\right]D_{\widetilde{\Delta}}\log z^h(\widetilde{R})\\
&= (w^h \circ u^h)_{cc}\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right)\widetilde{R}(z^h(\widetilde{R}))^2\frac{\widetilde{\Delta}(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))}\frac{\Delta(z^h(\widetilde{R}))}{\widetilde{R}(z^h(\widetilde{R}))}\\
&\quad + (w^h \circ u^h)_c\left(\widetilde{R}(z^h(\widetilde{R})), z^h(\widetilde{R})\right)\widetilde{R}(z^h(\widetilde{R}))\frac{d\log \widetilde{R}(z^h(\widetilde{R}))}{d\log z}\frac{1}{\varepsilon^h(\widetilde{R})}D_{\widetilde{\Delta}}\log z^h(\widetilde{R})D_\Delta \log z^h(\widetilde{R}).
\end{aligned} \tag{2.142}$$

The second equality follows from that

$$
\begin{aligned}
M^h(c^h, z^h) &= \frac{-u_z^h(c^h, z^h)}{u_c^h(c^h, z^h)} \\
\implies M^h(c^h, z^h)\frac{d}{dc^h}(w^h \circ u^h)(c^h, z^h) &= -\frac{d}{dz^h}(w^h \circ u^h)(c^h, z^h) \\
\implies \underbrace{M_c^h}_{=-\eta^h/\varepsilon^h \cdot M^h/c^h} \cdot (w^h \circ u^h)_c + M^h \cdot (w^h \circ u^h)_{cc} &= -(w^h \circ u^h)_{zc} \\
-\frac{\widetilde{R}' z^h}{\widetilde{R}}\frac{\eta^h}{\varepsilon^h}(w^h \circ u^h)_c + \frac{\widetilde{R}' z^h}{\widetilde{R}}(w^h \circ u^h)_{cc} c^h &= -(w^h \circ u^h)_{zc} z^h,
\end{aligned}
\tag{2.143}
$$

where here $z^h = z^h(\widetilde{R})$, $c^h = \widetilde{R}(z^h)$, elasticities and utility are are evaluated at $\widetilde{R}$, and we haveusing the definitions of ε^h and η^h and the fact that $\widetilde{R}'(z^h) = M^h(c^h, z^h)$ at the initial equilibrium.

Boundedness and integrability of welfare and its derivatives

Finally, we will argue that each of welfare $w^h \circ V^h(\widetilde{R})$ and its first two Frechet derivatives are, as linear maps, bounded across all $\widetilde{R} \in R + B_\delta(\mathbf{0})$ by some linear combination of the functions $b_n(h)$ from Definition 3 (of local regularity).[88] As we have already argued that each is measurable, this also implies by dominated convergence that each is integrable.

The case of welfare itself (not its derivatives) is immediate from Definition 3. Next, consider the first derivative of welfare. By our expression (2.141) for the first derivative, the definition of $||\cdot||$, and Definition 3, we have that $\left|D_\Delta w^h \circ V^h(\widetilde{R})\right| \leq ||\Delta|| b_1(h)$ for all $h \in \mathcal{H}, \widetilde{R} \in B_\delta(\mathbf{0})$, as desired. Finally, consider the second derivative of welfare. By our expression (2.142) for the second derivative, the definition of $||\cdot||$, Definition 3, Assumptions 1 and 4, and the fact that any $|D_\Delta \log z^h(\widetilde{R})|$ is bounded by a constant times $||\Delta||$ (by Lemma 3; see Footnote 56), we have that $\left|D^2_{\Delta\widetilde{\Delta}} w^h \circ V^h(\widetilde{R})\right| \leq ||\Delta|| ||\widetilde{\Delta}|| (b_2(h) + const \cdot b_1(h))$ for all $h \in \mathcal{H}, \widetilde{R} \in B_\delta(\mathbf{0})$, as desired.

2.9.6.5.1 Supporting details for Theorem 1

In this section we cover various details omitted from the proof of Theorem 1 in order to focus on the main points. Each supporting detail is referenced in the main proof.

2.9.6.5.2 Properties of aggregate revenue and welfare

By Lemmas 3 and 5, there exists $\delta > 0$ such that on $R + B_\delta(\mathbf{0}) \ni \mathcal{R}$,

- $z^h(\widetilde{R})$, $c^h(\widetilde{R})$ are well-defined and have two Frechet derivatives in $\widetilde{R}$ that are continuous, $\mathcal{H}$-integrable, and bounded by linear combinations of z_0^h and c_0^h (which are integrable by Assumption 4). By the linearity of differentiation, the same is true for tax revenue $z^h(\widetilde{R}) - c^h(\widetilde{R})$.
- $w^h \circ V^h(\widetilde{R})$ is well-defined and has two Frechet derivatives in $\widetilde{R}$ that are continuous, $\mathcal{H}$-integrable, and bounded by linear combinations the functions $b_n(h)$ in Definition 3 (which are integrable by assumption).

[88] That is, for any $\Delta, \widetilde{\Delta} \in \mathbf{\Delta}$, $w^h \circ V^h(\widetilde{R}) \leq (a_0^0 b_0(h) + a_1^0 b_1(h) + a_2^0 b_2(h))$, $D_\Delta w^h \circ V^h(\widetilde{R}) \leq ||\Delta||(a_0^1 b_0(h) + a_1^1 b_1(h) + a_2^1 b_2(h))$, and $D^2_{\Delta\widetilde{\Delta}} w^h \circ V^h(\widetilde{R}) \leq ||\Delta|| ||\widetilde{\Delta}|| (a_0^2 b_0(h) + a_1^2 b_1(h) + a_2^2 b_2(h))$ for some constants $a_j^k \in \mathbb{R}$.

Theorem 1 of Kammar (2016) therefore implies that—for $f^h(\widetilde{R}) = z^h(\widetilde{R}) - c^h(\widetilde{R})$ and for $f^h(\widetilde{R}) = w^h \circ V^h(\widetilde{R})$—$\int f^h(\widetilde{R})d\mu$ is two-times Frechet differentiable on $R + B_\delta(\mathbf{0})$ and that

$$D \int f^h(\widetilde{R})d\mu = \int Df^h(\widetilde{R})d\mu \qquad \text{and} \qquad D^2 \int f^h(\widetilde{R})d\mu = \int D^2 f^h(\widetilde{R})d\mu. \tag{2.144}$$

Moreover—by dominated convergence—$\int f^h(\widetilde{R})d\mu$, $D\int f^h(\widetilde{R})d\mu$, and $D^2\int f^h(\widetilde{R})d\mu$ are continuous in $\widetilde{R}$ because $f^h(\widetilde{R}), Df^h(\widetilde{R})$, and $D^2 f^h(\widetilde{R})$ are, by Lemma 2, and because each is dominated by an integrable function (a linear combination of z_0^h, c_0^h, and the functions $b_n(h)$).[89]

2.9.6.5.3 Measure-theoretic steps used to change variables

The following steps are implicit in the change of variables carried out in (2.54):

$$\begin{aligned}
&\int \left[(1 - R'(z_0^h))z_0^h \left(\eta^h(R)\frac{\Delta(z_0^h)}{R(z_0^h)} + \varepsilon^h(R)\frac{\Delta'(z_0^h)}{R'(z_0^h)}\right) - \Delta(z_0^h)\right] d\mu \\
= &\int_{\{z_0^h \in \operatorname{supp} g\}} \left[(1 - R'(z_0^h))z_0^h \left(\eta^h(R)\frac{\Delta(z_0^h)}{R(z_0^h)} + \varepsilon^h(R)\frac{\Delta'(z_0^h)}{R'(z_0^h)}\right) - \Delta(z_0^h)\right] d\mu \\
= &\int_{\{z_0^h \in \operatorname{supp} g\}} \left[(1 - R'(z_0^h))z_0^h \left(\eta(z_0^h)\frac{\Delta(z_0^h)}{R(z_0^h)} + \varepsilon(z_0^h)\frac{\Delta'(z_0^h)}{R'(z_0^h)}\right) - \Delta(z_0^h)\right] d\mu \\
= &\int_{\operatorname{supp} g} \left[(1 - R'(z))z \left(\eta(z)\frac{\Delta(z)}{R(z)} + \varepsilon(z)\frac{\Delta'(z)}{R'(z)}\right) - \Delta(z)\right] d\mu_z \\
= &\int_{\operatorname{supp} g} \left[(1 - R'(z))z \left(\eta(z)\frac{\Delta(z)}{R(z)} + \varepsilon(z)\frac{\Delta'(z)}{R'(z)}\right) - \Delta(z)\right] g(z)dz
\end{aligned} \tag{2.146}$$

Above, the second line is by the absolute continuity of the integral and the fact that $\mu(\{z_0^h \in \operatorname{supp} g\}) = 1$.[90] The third line follows from breaking the integral into each of its additively-separable components—each of which are integrable by bounding arguments in the first part of the proof of Lemma 3—then applying the definition of conditional expectation—and then adding these integrals back together them (by linearity of Lebesgue integration). The fourth line follows from changing variables and letting μ_z denote the measure on $\mathbb{R}_{\geq 0}$ induced by $z_0^{(\mathcal{H})}$.[91] The fifth line is by the Radon-Nikodyn / change of measure theorem and the definition of the density as a Radon-Nikodyn derivative; the density exists by Assumption 5.[92]

[89] To see why this is sufficient, consider any continuous-in-$\widetilde{R}$ and integrable-over-h function $\rho(\widetilde{R}, h)$ which is bounded across all $\widetilde{R}$ by some integrable $\overline{\rho}(h)$. For any sequence $\widetilde{R}_n \to \widetilde{R}$, we therefore have—by continuity, the integrable bound, and dominated convergence—that

$$\rho(\widetilde{R}_n, h) \to \rho(\widetilde{R}, h) \qquad \Longrightarrow \qquad \int \rho(\widetilde{R}_n, h)d\mu \to \int \rho(\widetilde{R}, h)d\mu. \tag{2.145}$$

[90] See Chapter 3, Theorem 2.5, Shorack (2000).
[91] See Chapter 3, Theorem 2.6, Shorack (2000).
[92] See Chapter 4, Theorem 2.2, Shorack (2000).

2.9.6.5.4 Application of optimization-theoretic results

Recall the optimization problem (2.69):

$$\begin{aligned}
&\mathbf{0} \in \underset{\Delta \in \mathbf{\Delta}}{\arg\max}\ \ F(\Delta) \quad \text{s.t.} \quad H(\Delta) \in \mathbb{R}_{\geq 0} \\
&\text{where } F(\Delta) \equiv \begin{cases} \int w^h \circ u^h \left((R+\Delta)(z^h(R+\Delta)), z^h(R+\Delta)\right) d\mu, & \text{if } \Delta \in B_\delta(\mathbf{0}) \\ F(\mathbf{0}) - 1 & \text{if } \Delta \notin B_\delta(\mathbf{0}) \end{cases} \\
&H(\Delta) \equiv \begin{cases} \int \left[z^h(R+\Delta) - (R+\Delta)(z^h(R+\Delta))\right] d\mu - G, & \text{if } \Delta \in B_\delta(\mathbf{0}) \\ 0, & \text{if } \Delta \notin B_\delta(\mathbf{0}) \end{cases}
\end{aligned} \tag{2.147}$$

where $\delta > 0$ is small enough that F and H are well-defined and within $B_\delta(\mathbf{0})$ have well-defined and continuous first and second Frechet derivatives (see Appendix 2.9.6.5.2).

We wish to apply results from optimization theory on Banach spaces results to the problem above. In particular, we leverage Theorems 3.2 and 3.3 (part 2) of Maurer and Zowe (1979) in the special case of a one-dimensional constraint, restated below:

Appendix Theorem 1. *Let X be a real Banach space, $\bar{x}$ a point in X, and $F : X \to \mathbb{R}$ and $H : X \to \mathbb{R}$ functions whose first and second Frechet derivatives exist at $\bar{x}$. Suppose that*

$$\bar{x} \in \underset{H(x) \geq 0}{\arg\max}\, F(x), \tag{2.148}$$

$H(\bar{x}) = 0$, and for some $h \in X$, $D_h H(\bar{x}) \neq 0$ ($\bar{x}$ is optimal and full-rank / regular, and H binds.).[93]

Then there exists $\kappa \in \mathbb{R}_{\geq 0}$ such that $DF(\bar{x}) + \kappa DH(\bar{x}) = 0$ and, for all non-zero $d \in X$ satisfying $D_d H(\bar{x}) = 0$, $D^2_{d,d} F(\bar{x}) + \kappa D^2_{d,d} H(\bar{x}) \leq 0$.

In order to apply Appendix Theorem 1, we must verify that the optimization problem (2.69) satisfies several conditions. It is immediate from the setup above that:

- F is a functional defined on a real Banach space, $(\mathbf{\Delta}, ||\cdot||)$.[94]
- G is a map from $\mathbf{\Delta}$ into $\mathbb{R}$, a real Banach space (w.r.t. the standard norm).
- The feasible range for G is $\mathbb{R}_{\geq 0}$.
- F restricted to $G^{-1}(\mathbb{R}_{\geq 0})$ achieves a local (indeed, global) maximum at $\mathbf{0}$.
- The first and second Frechet derivatives of F and G exist at $\mathbf{0}$.

In order to reach the result stated in the main proof of Theorem 1, it remains to verify that $H(\mathbf{0}) = 0$, i.e. the revenue constraint binds. To see this suppose not. Since H is continuous local to $\mathbf{0}$, and F is differentiable local to $\mathbf{0}$, it suffices to show that for any $\epsilon > 0$, there exists $\Delta \in \mathbf{\Delta}$ with $F(\epsilon\Delta) > F(\mathbf{0})$. To see this, define $\Delta = R$ and note that $R(z_0^h) > 0$ for all $h \in \mathcal{H}$ (this follows from the fact shown in the proof of Lemma 2 that $R(z) > 0$ for all $z > 0$ and from Assumption 3).[95] By Assumption 2,

[93] These facts are sufficient to guarantee the strong version of Equation (2.3) of Maurer and Zowe (1979) used in their Theorem 3.3. Specifically, they imply that for any set $K' \subset R$ containing 0, $0 \in \text{int}(H(\bar{x}) + DH(\bar{x})X - K')$, where int is the topological interior, "$+$" and "$-$" denote addition and subtraction of set elements, and $DH(\bar{x})X$, is the image of $DH(\bar{x})$ (recall the Frechet derivative $DH(\bar{x})$ is a linear map $X \to \mathbb{R}$).

[94] Recall we have already shown this is a real Banach space; see Lemma 1.

[95] It is immediate from the defintion of $\mathbf{\Delta}$ that $R \in \mathbf{\Delta}$.

$V^h(R+\epsilon\Delta) > u^h\left(c_0^h + \epsilon R(z_0^h), z_0^h\right) > V^h(R)$. So by Definition 1, $w^h \cdot V^h(R+\epsilon\Delta) > w^h \circ V^h(R)$ for a positive measure of households h; so $F(\epsilon\Delta) > F(\mathbf{0})$.

2.9.6.5.5 Zero-rank case of (2.69)

Consider the optimization problem (2.69) and suppose that $DH(\mathbf{0}) = 0$, i.e. $D_\Delta(\mathbf{0}) = 0$ for all $\Delta \in \boldsymbol{\Delta}$. We claim in this case (ABC) holds with equality for all $z \in \operatorname{supp} g$, so the theorem holds.

The argument is identical to that of the section "First-order condition" in the main proof of Theorem 1, except that (a) one may start at (2.70), which now holds with equality, and (b) the argument in the next paragraph implies $\psi(z) = 0$, because of (a).

2.9.6.5.6 Example of first-order tax deviation

The main proof of Theorem 1 uses the existence of a weakly positive function $\Delta(z)$ that is strictly positive on a non-zero-measure sub-interval of $[\underline{z}, \overline{z}]$, zero outside of $[\underline{z}, \overline{z}]$, and is contained in $\boldsymbol{\Delta}$. We now give an example of such a function.

We begin by defining the "deviation function", for any $\tilde{z} > 0, \tilde{\delta} \in (0, \tilde{z})$, by

$$\widetilde{\Delta}(z; \tilde{z}, \tilde{\delta}) \equiv \begin{cases} -\left(\frac{z-(\tilde{z}-\tilde{\delta})}{\tilde{\delta}}\right)^5 \left(\frac{z-(\tilde{z}+\tilde{\delta})}{\tilde{\delta}}\right)^5 & \text{if } z \in B_{\tilde{\delta}}(\tilde{z}) \\ 0 & \text{otherwise.} \end{cases} \tag{2.149}$$

It is easy to verify that $\widetilde{\Delta}(\cdot; \tilde{z}, \tilde{\delta})$ has four continuous and bounded derivatives, is strictly positive in $B_{\tilde{\delta}}(\tilde{z})$ and zero elsewhere, and has $\widetilde{\Delta}(z; \tilde{z}, \tilde{\delta}) \leq \widetilde{\Delta}(\tilde{z}; \tilde{z}, \tilde{\delta}) = 1$. Moreover note that $R(z)$ and $R'(z)$ are bound strictly above zero in $\overline{B}_{\tilde{\delta}}(\tilde{z})$ because $R(z), R'(z) > 0$ at all $z > 0$—as argued in the proof of Lemma 2—and R and R' are continuous by Assumption 1. Together, these observations imply $\widetilde{\Delta}(\cdot; \tilde{z}, \tilde{\delta}) \in \boldsymbol{\Delta}$.

To obtain the desired deviation Δ, consider $\Delta \equiv \widetilde{\Delta}\left(\cdot; \frac{\overline{z}+\underline{z}}{2}, \frac{\overline{z}-\underline{z}}{2}\right)$.

2.9.6.5.7 Example of second-order tax deviation

The main proof of Theorem 1 uses the existence of, for any $k > 0$, $r > 0$, $\hat{z} \in \operatorname{supp} g$, a tax change Δ that is in $\boldsymbol{\Delta}$, is zero outside of the interval $B_r(\hat{z})$, and satisfies $\int_{\hat{z}-r}^{\hat{z}+r} \Delta'(\hat{z})^2 dz > k \int_{\hat{z}-r}^{\hat{z}+r} \Delta(z)^2 dz$. Here the idea is to take $\Delta(\cdot; \hat{z}, r, k)$ to be a sufficiently narrow "bump" centered at z. We now provide an example of some such function.

To construct the example, we use the "deviation function" defined in Appendix 2.9.6.5.6. Specifically, we take $\Delta_N \equiv \widetilde{\Delta}(\cdot; \frac{\overline{z}+\underline{z}}{2}, r/N)$ for some $N \geq 1$. We have already established in Appendix 2.9.6.5.6 that $\Delta_N \in \boldsymbol{\Delta}$ and Δ_N is zero outside of the interval $B_r(\hat{z})$. To see that there exists N for which we obtain

$$\int_{\hat{z}-r}^{\hat{z}+r} \Delta_N'(\hat{z})^2 dz > k \int_{\hat{z}-r}^{\hat{z}+r} \Delta_N(z)^2 dz, \tag{2.150}$$

note that as $N \to \infty$, the RHS converges to 0, whereas the LHS diverges to ∞.

2.9.6.6 Supporting details for Theorem 2

In this section we cover various details omitted from the proof of Theorem 2 in order to focus on the main points. Each supporting detail is referenced in the main proof.

2.9.6.6.1 Properties of $\widehat{\lambda}(z)$ and $\widehat{\lambda}^h$

We make several observations about $\widehat{\lambda}(z)$ and $\widehat{\lambda}^h$.

First, $\widehat{\lambda}(z)$ is continuously differentiable on $\operatorname{supp} g$. This is immediate from the definition of $\Pi_{ABC}(z)$ and Assumptions 2, 5, and 6.

Second, there exists $M > 0$ such that for all $z \in \operatorname{supp} g$, $z|\widehat{\lambda}'(z)| \leq M|\lambda(z)|$. This is immediate from the conditions on $\Pi_{ABC}(z)$ in the statement of the theorem.

Third, there exist $\overline{\lambda}_c, \overline{\lambda}_z \in \mathbb{R}_{>0}$ such that for all $z \in \operatorname{supp} g$, $R(z)\widehat{\lambda}(z) \leq \overline{\lambda}_c R(z) + \overline{\lambda}_z z$. This can be seen by considering the definition of $\Pi_{ABC}(z)$ and making the following observations:

- $R'(z)z$ is bounded across all $z \in \mathbb{R}_{\geq 0}$ by $B^R R(z)$, by Assumption 2.
- $\eta(z), \varepsilon(z)$ are bounded across $\operatorname{supp} g$ by a constant by Assumption 4.
- $\alpha(z)$, $\frac{d \log \varepsilon(z)}{d \log z}\varepsilon(z) = z\varepsilon'(z)$, and $\frac{R(z)}{R'(z)z}$ are bounded across $\operatorname{supp} g$ by constants, by Assumption 6.
- For all $z \in \operatorname{supp} g$, $R(z)\left|\frac{d}{d\log z}\left(\frac{1-R'(z)}{R'(z)}\right)\right| = R(z)\left(\left|\frac{R''(z)z}{R'(z)}\right| + \left|\frac{1-R'(z)}{R'(z)}\right|\left|\frac{R''(z)z}{R'(z)}\right|\right) \leq B^R(1+1)R(z) + B^R kz$ for some constant k by Assumption 1 and the boundedness of $\frac{R(z)}{R'(z)z}$.

Fourth, $\widehat{\lambda}^h$ is measurable in $\mathcal{H}$. This follows from that (a) $\varepsilon^h(R)$ is measurable (see the second-to-last section of the proof of Lemma 2), (b) within $\operatorname{supp} g$, $p_{\leq}(z;\epsilon)$ is continuous in z and strictly positive by assumption, (c) z_0^h is measurable by Assumption 3, (d) $\widehat{\lambda}(z)$ is piece-wise continuous, and (e) $\operatorname{supp} g$ is measurable by Lemma 9 in Appendix 2.9.6.7.

Fifth, $R(z_0^h)\widehat{\lambda}^h$ is integrable. We will show this by an application of Fatou's Lemma:[96] Define, for $n \in \mathbb{N}$, $\widehat{\lambda}_n^h \equiv \widehat{\lambda}(z_0^h)\mathbb{1}_{\varepsilon^h(R)\leq\epsilon} \min\left[n, \frac{1}{p_{\leq}(z_0^h;\epsilon)}\right]$, and note that $\widehat{\lambda}_n^h \geq 0$ and that $\widehat{\lambda}_n^h \to \widehat{\lambda}^h$ pointwise. Moreover, $R(z_0^h)\widehat{\lambda}_n^h$ is integrable since is measurable (by a similar argument to that used for $R(z_0^h)\widehat{\lambda}^h$) and bounded by $nR(z_0^h)\widehat{\lambda}(z_0^h)$, which recall is itself bounded by $n\overline{\lambda}_c R(z_0^h) + n\overline{\lambda}_z z_0^h$ (which is integrable by Assumption 4). Further, note that

$$\begin{aligned}
\mathbb{E}[R(z_0^h)\widehat{\lambda}_n^h] = \mathbb{E}\left[\mathbb{E}[R(z_0^h)\widehat{\lambda}_n^h \mid z_0^h]\right] &= \mathbb{E}\left[\mathbb{E}\left[R(z_0^h)\widehat{\lambda}(z_0^h)\mathbb{1}_{\varepsilon^h(R)\leq\epsilon}\min\left[n, \frac{1}{p_{\leq}(z_0^h;\epsilon)}\right] \;\middle|\; z_0^h\right]\right] \\
&= \mathbb{E}\left[R(z_0^h)\widehat{\lambda}(z_0^h)\min\left[n, \frac{1}{p_{\leq}(z_0^h;\epsilon)}\right]\mathbb{E}\left[\mathbb{1}_{\varepsilon^h(R)\leq\epsilon} \;\middle|\; z_0^h\right]\right] \\
&= \mathbb{E}\left[R(z_0^h)\widehat{\lambda}(z_0^h)\min\left[n, \frac{1}{p_{\leq}(z_0^h;\epsilon)}\right]p_{\leq}(z_0^h;\epsilon)\right] \\
&\leq \mathbb{E}\left[R(z_0^h)\widehat{\lambda}(z_0^h)\right]
\end{aligned} \tag{2.151}$$

where here we have used z_0^h's measurability to take the conditional expectation, then used the tower property of conditional expectations, and then taken advantage of $\mathbb{1}_{\varepsilon^h(R)\leq\epsilon}$'s integrability to pull out z_0^h-measurable terms from the conditional expectation.[97] Since $R(z_0^h)\widehat{\lambda}^h$, and $R(z_0^h)\widehat{\lambda}_n^h$ are all non-negative, we have by Fatou's Lemma (and then the bounds above) that

$$\int R(z_0^h)\widehat{\lambda}^h d\mu = \int \liminf_{n\to\infty} R(z_0^h)\widehat{\lambda}_n^h d\mu \leq \liminf_{n\to\infty}\int R(z_0^h)\widehat{\lambda}_n^h d\mu \leq \mathbb{E}\left[\mathbb{1}_{z_0^h\in\operatorname{supp} g}R(z_0^h)\widehat{\lambda}(z_0^h)\right] < \infty \tag{2.152}$$

In particular, $R(z_0^h)\widehat{\lambda}^h$ is integrable.

[96] Fatou's Lemma is a standard result in measure theory. In words, it says that, for non-negative real-valued random variables, the lim-inf of an expectation is less than the expectation of a lim-inf.

[97] See, e.g., Theorems 34.3 and 34.4 of Billingsley (2008).

2.9.6.6.2 Conditional expectations related to $\widehat{\lambda}^h$

We first argue that several conditional expectations of interest—namely those of $R(z_0^h)\widehat{\lambda}^h$, $R(z_0^h)\widehat{\lambda}^h\eta^h(R)$, $R(z_0^h)\widehat{\lambda}^h\varepsilon^h(R)$, and $R(z_0^h)\widehat{\lambda}^h\frac{\eta^h(R)^2}{\varepsilon^h(R)}$ conditional on income—exist. Second, we show how these moments can be related to conditional expectations of $\widehat{\lambda}^h$ $\eta^h(R), \varepsilon^h(R)$ and $\frac{\eta^h(R)^2}{\varepsilon^h(R)}$. Finally, we argue that each conditional expectation is continuous in income and that the conditional expectation of $R(z_0^h)\widehat{\lambda}^h\eta^h(R)$ is continuously differentiable in income. These arguments rely on facts shown in Appendix 2.9.6.6.1.

To begin, we argue that the following conditional expectation functions exist:

$$\mathbb{E}\left[R(z_0^h)\widehat{\lambda}^h \mid z_0^h\right], \quad \mathbb{E}\left[R(z_0^h)\widehat{\lambda}^h\eta^h(R) \mid z_0^h\right], \quad \mathbb{E}\left[R(z_0^h)\widehat{\lambda}^h\varepsilon^h(R) \mid z_0^h\right], \quad \mathbb{E}\left[R(z_0^h)\widehat{\lambda}^h\frac{\eta^h(R)^2}{\varepsilon^h(R)} \mid z_0^h\right]. \tag{2.153}$$

Since by Assumption 3, z_0^h is measurable, this follows so long as $R(z_0^h)\widehat{\lambda}^h$, $R(z_0^h)\widehat{\lambda}^h\eta^h(R)$, $R(z_0^h)\widehat{\lambda}^h\varepsilon^h(R)$, and $R(z_0^h)\widehat{\lambda}^h\frac{\eta^h(R)^2}{\varepsilon^h(R)}$ are each integrable. Indeed, each is measurable by the observations (above) that $\widehat{\lambda}^h$ is measurable and (see the second-to-last step of the proof of Lemma 2) that elasticities are measurable; and each is bounded by an integrable function, by Assumption 4 and our observation above that $R(z_0^h)\widehat{\lambda}^h$ is integrable.

Next we observe that, for $x^h = 1, \eta^h(R), \varepsilon^h(R), \frac{\eta^h(R)^2}{\varepsilon^h(R)}$; for any conditional expectations $\mathbb{E}\left[R(z_0^h)\widehat{\lambda}^h x^h \mid z_0^h\right]$ and $\mathbb{E}\left[x^h \mid z_0^h, \mathbb{1}_{\varepsilon^h(R)\le\epsilon}\right]$;[98] and with probability one,

$$\begin{aligned}
\mathbb{E}\left[R(z_0^h)\widehat{\lambda}^h x^h \mid z_0^h\right] &= \mathbb{E}\left[R(z_0^h)\widehat{\lambda}(z_0^h)\frac{\mathbb{1}_{\varepsilon^h(R)\le\epsilon}}{p_{\le}(z_0^h;\epsilon)}x^h \,\middle|\, z_0^h\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[R(z_0^h)\widehat{\lambda}(z_0^h)\frac{\mathbb{1}_{\varepsilon^h(R)\le\epsilon}}{p_{\le}(z_0^h;\epsilon)}x^h \,\middle|\, z_0^h, \mathbb{1}_{\varepsilon^h(R)\le\epsilon}\right] \,\middle|\, z_0^h\right] \\
&= \mathbb{E}\left[R(z_0^h)\widehat{\lambda}(z_0^h)\frac{\mathbb{1}_{\varepsilon^h(R)\le\epsilon}}{p_{\le}(z_0^h;\epsilon)}\mathbb{E}[x^h|\, z_0^h, \mathbb{1}_{\varepsilon^h(R)\le\epsilon}] \,\middle|\, z_0^h\right] \\
&= \mathbb{E}\Bigg[R(z_0^h)\widehat{\lambda}(z_0^h)\frac{\mathbb{1}_{\varepsilon^h(R)\le\epsilon}}{p_{\le}(z_0^h;\epsilon)} \\
&\qquad \cdot \left(\mathbb{E}[x^h|\, z_0^h, \mathbb{1}_{\varepsilon^h(R)\le\epsilon}=1]\mathbb{1}_{\varepsilon^h(R)\le\epsilon} + \mathbb{E}[x^h|\, z_0^h, \mathbb{1}_{\varepsilon^h(R)\le\epsilon}=0]\mathbb{1}_{\varepsilon^h(R)>\epsilon}\right) \,\Bigg|\, z_0^h\Bigg] \\
&= \mathbb{E}\left[R(z_0^h)\widehat{\lambda}(z_0^h)\frac{\mathbb{1}_{\varepsilon^h(R)\le\epsilon}}{p_{\le}(z_0^h;\epsilon)}\mathbb{E}[x^h|\, z_0^h, \mathbb{1}_{\varepsilon^h(R)\le\epsilon}=1] \,\middle|\, z_0^h\right] \\
&= R(z_0^h)\widehat{\lambda}(z_0^h)\frac{1}{\cancel{p_{\le}(z_0^h;\epsilon)}}\mathbb{E}[x^h|\, z_0^h, \mathbb{1}_{\varepsilon^h(R)\le\epsilon}=1]\cdot\cancel{\mathbb{E}[\mathbb{1}_{\varepsilon^h(R)\le\epsilon} \mid z_0^h]} \\
&= R(z_0^h)\widehat{\lambda}(z_0^h)x_{\le}(z_0^h;\epsilon)
\end{aligned} \tag{2.154}$$

where recall $x_{\le}(z_0^h;\epsilon) = \mathbb{E}[x^h|z_0^h = z, \varepsilon^h(R) \le \epsilon]$ is as defined in Assumption 6 and where for $x^h = 1$, $x_{\le}(z_0^h;\epsilon)$ simply denotes 1. Above, the second equality holds (with probability one) by the tower property,[99] for any inner conditional expectation.[100] The third and sixth equalities hold since the pulled-out terms are measurable with respect to $(z_0^h, \mathbb{1}_{\varepsilon^h\le\epsilon})$ and z_0^h in the respective cases.[101] The third line also uses that $\mathbb{E}[x^h|\, z_0^h, \mathbb{1}_{\varepsilon^h(R)\le\epsilon}]$ exists, since x^h is integrable and z_0^h and $1_{\varepsilon^h(R)\le\epsilon}$ are measurable; the sixth uses that

[98] The latter conditional expectation exists since each x^h is integrable (by observations above and Assumption 4) and since $\mathbb{1}_{\varepsilon^h(R)\le\epsilon}$ is measurable since $\varepsilon^h(R)$ is.

[99] See, e.g., Theorem 34.4 of Billingsley (2008)

[100] The inner conditional expectation exists by the bounding arguments above and since $\mathbb{1}_{\varepsilon^h(R)<\epsilon}$ is measurable.

[101] See, e.g. Theorem 34.3 of Billingsley (2008).

$\mathbb{E}[\mathbb{1}_{\varepsilon^h(R)\leq\epsilon} \mid z_0^h]$ exists, since $\varepsilon^h(R)$ and z_0^h are measurable and any indicator is bounded. The fourth equality is definitional and the fifth is immediate. The cancellations on the second-to-last line are with probability one, since $p_{\leq}(z_0^h;\epsilon)$ is a conditional expectation of $\mathbb{1}_{\varepsilon^h(R)\leq\epsilon}$ on z_0^h. The seventh equality is because all conditional expectations of the same variables are equal with probability one.

Since—for any choice of conditional expectation (conditional expectations are only unique up to differences on measure zero sets)—$\mathbb{E}[R(z_0^h)\widehat{\lambda}^h x^h \mid z_0^h]$ and $R(z_0^h)\widehat{\lambda}(z_0^h)x_{\leq}(z_0^h;\epsilon)$ coincide with probability one, and since $R(z_0^h)\widehat{\lambda}(z_0^h)x_{\leq}(z_0^h;\epsilon)$ is z_0^h-measurable, it follows from the definition of conditional expectation that $R(z_0^h)\widehat{\lambda}(z_0^h)x_{\leq}(z_0^h;\epsilon)$ is a conditional expectation for $R(z_0^h)\widehat{\lambda}^h x^h$ conditional on z_0^h. For the remainder of the proof, we will work with this particular choice of conditional expectation.

Another consequence of our observations is that, for $x^h = \varepsilon^h(R), \eta^h(R), \frac{\eta^h(R)^2}{\varepsilon^h(R)}$, there exist functions $(\widehat{\lambda}x) : \operatorname{supp} g \to \mathbb{R}$—namely $(\widehat{\lambda}x)(z) \equiv \widehat{\lambda}(z)x_{\leq}(z,\epsilon)$ —that are equal to $\frac{\mathbb{E}[R(z_0^h)\widehat{\lambda}^h x^h \mid z_0^h = z]}{R(z)}$ for all $z \in \operatorname{supp} g$. Moreover, the continuous differentiability of $\widehat{\lambda}(z)$ (shown above) and Assumption 6 imply that $(\widehat{\lambda}\varepsilon)(z)$ and $\left(\widehat{\lambda}\frac{\eta^2}{\varepsilon}\right)(z)$—as well as, by Assumption 1, $R(z)(\widehat{\lambda}\varepsilon)(z)$ and $R(z)\left(\widehat{\lambda}\frac{\eta^2}{\varepsilon}\right)(z)$—are continuous in $z \in \operatorname{supp} g$ and $(\widehat{\lambda}\eta)(z)$—as well as $R(z)(\widehat{\lambda}\eta)(z)$—is continuously differentiable in z.

2.9.6.6.3 Properties of $(\widehat{\lambda\gamma})(z)$ and $\widehat{\gamma}^h$

We make several observations about $(\widehat{\lambda\gamma})(z)$ and $\widehat{\gamma}^h$.

First, $(\widehat{\lambda\gamma})(z)$ is continuous on $\operatorname{supp} g$. This follows from $\widehat{\lambda}(z)$'s continuity, Lemma 6, and Assumptions 2, 4, and 6.

Second, there exist $\overline{\lambda\gamma}_c, \overline{\lambda\gamma}_z > 0$ such that for all $z \in \mathbb{R}_{\geq 0}$, $|R(z)(\widehat{\lambda\gamma})(z)| \leq \overline{\lambda\gamma}_c R(z) + \overline{\lambda\gamma}_z z$. This follows from the following observations:

- $(\widehat{\lambda\gamma})(z) = 0$ for $z \notin \operatorname{supp} g$.
- By Lemma 6, $\Psi_0(z)$ is bounded on $\operatorname{supp} g$ by a linear combination of z and $R(z)$.
- $\frac{d\log R(z)}{d\log z} \leq B^R$ by Assumption 1.
- By our observations in Appendix 2.9.6.6.2, $R(z)(\widehat{\lambda}x)(z) = R(z)\widehat{\lambda}(z)x_{\leq}(z;\epsilon)$ is—on $\operatorname{supp} g$—bounded by a linear combination of z and $R(z)$ since—for $x = \eta, \varepsilon, \frac{\eta^2}{\varepsilon}$—$x_{\leq}(z;\epsilon)$ is bounded (by Assumption 4) and $R(z)\widehat{\lambda}(z) \leq \overline{\lambda}_c R(z) + \overline{\lambda}_z z$ (see Appendix 2.9.6.6.1).
- By Assumption 6, $\alpha(z)$ is bounded in $\operatorname{supp} g$.
- By Assumptions 4 and 6 and by our observations in Appendices 2.9.6.6.1 and 2.9.6.6.2, there exist constants M, M', M'' such that for all $z \in \operatorname{supp} g$

$$\begin{aligned} R(z)z(\widehat{\lambda}\eta)'(z) &= R(z)\frac{d}{dz}\left(\widehat{\lambda}(z)\eta_{\leq}(z;\epsilon)\right) = R(z)\left(z\widehat{\lambda}'(z)\eta_{\leq}(z;\epsilon) + \widehat{\lambda}(z)z\eta'_{\leq}(z;\epsilon)\right) \\ &\leq R(z)\left(M\widehat{\lambda}(z)\eta_{\leq}(z;\epsilon) + M'\widehat{\lambda}(z)\right) \leq M''R(z)\widehat{\lambda}(z) \end{aligned} \tag{2.155}$$

  which is integrable.

Third, $\widehat{\gamma}^h$ is measurable in h. This follows from that (a) as shown above and noted / proved in Appendix 2.9.6.6.1, $(\widehat{\lambda\gamma})(z_0^h)$, $\widehat{\lambda}^h$, $\varepsilon^h(R)$, and $p(z_0^h;\epsilon)$ are all measurable functions of h; (b) by Lemma 9 in Appendix 2.9.6.7, $\operatorname{supp} g$ is a measurable set; and (c) for all h with $z_0^h \in \operatorname{supp} g$ and $\varepsilon^h(R) \leq \epsilon$, $\widehat{\lambda}^h > 0$ and $p(z_0^h;\epsilon) > 0$ (by Assumption 6).

Fourth $R(z_0^h)\widehat{\lambda}^h\widehat{\gamma}^h$ is integrable. To see this first note that $R(z_0^h)(\widehat{\lambda\gamma})(z_0^h)$ is integrable, which follows from (a) the continuity of R (Assumption 1); (b) the measurability of z_0^h, $\widehat{\lambda}^h$, and $\widehat{\gamma}^h$ (Assumption 3),

Appendix 2.9.6.6.1, above); and (c) the fact that $|R(z_0^h)(\widehat{\lambda\gamma})(z_0^h)| \leq \overline{\lambda\gamma}_c R(z_0^h) + \overline{\lambda\gamma}_z z_0^h$ for all $h \in \mathcal{H}$, where note the RHS is integrable by Assumption 4. One may then apply the same Fatou's Lemma argument used in Appendix 2.9.6.6.2.

Fifth, we claim $R(z_0^h)(\widehat{\lambda\gamma})(z_0^h)$ is a conditional expectation for $R(z_0^h)\widehat{\lambda}^h\widehat{\gamma}^h$ given z_0^h. To see this, note that since z_0^h is measurable, we may consider the conditional expectation

$$\begin{aligned}\mathbb{E}\left[R(z_0^h)\widehat{\lambda}^h\widehat{\gamma}^h \middle| z_0^h\right] &= \mathbb{E}\left[R(z_0^h)(\widehat{\lambda\gamma})(z_0^h)\frac{\mathbb{1}_{\varepsilon^h(R)\leq\epsilon}}{p(z_0^h;\epsilon)}\middle| z_0^h\right] \\ &= R(z_0^h)(\widehat{\lambda\gamma})(z_0^h)\,\frac{1}{p(z_0^h;\epsilon)}\,\mathbb{E}\left[\mathbb{1}_{\varepsilon^h(R)\leq\epsilon}\middle| z_0^h\right]\end{aligned} \tag{2.156}$$

where the second equality holds with probability one (same logic as in (2.154)). By the same logic as in Appendix 2.9.6.6.2, $R(z_0^h)(\widehat{\lambda\gamma})(z_0^h)$ is a conditional expectation for $R(z_0^h)\widehat{\lambda}^h\widehat{\gamma}^h$ given z_0^h. For the remainder of the proof, we work with this particular choice of conditional expectation.

2.9.6.6.4 Characterization of $\hat{c}^h(u)$

The main proof of Theorem 2 relies on a characterization of the compensating-consumption function $\hat{c}^h(u) \equiv u^h(\cdot, z_0^h)^{-1}(u)$. While useful, this characterization is very tedious, and so we (further) relegate its proof to Appendix 2.9.6.8. Here, we simply state the result, i.e. Lemma 11:

There exists $\delta > 0$ small enough that the function

$$\hat{c}^h(u) \equiv u^h(\cdot, z_0^h)^{-1}(u) \tag{2.157}$$

is, for all $h \in \mathcal{H}$, well-defined and strictly positive when $u = V^h(\widetilde{R})$ for some $\widetilde{R} \in R + B_\delta(\mathbf{0})$; moreover, $\hat{c}^h(V^h(\widetilde{R}))$ is $\mathcal{H}$-measurable. Further, there exists $\bar{m} > 0$ such that for all $h \in \mathcal{H}, \widetilde{R} \in R + B_\delta(\mathbf{0})$ and—for all real-valued functions ϕ^h that are defined and twice differentiable in a neighborhood around $V^h(\widetilde{R})$ and satisfy $\phi^{h\prime}(V^h(\widetilde{R})) > 0$—we have[102]

$$\begin{aligned}&\left|\log \hat{c}^h\left(V^h(\widetilde{R})\right) - \log c^h\left(\widetilde{R}\right)\right| \leq \bar{m} \\ &\left|\log\left[(\phi^h \circ u^h)_c\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)\right] - \log\left[(\phi^h \circ u^h)_c\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)\right]\right| \leq \bar{m} \\ \text{and}\quad &\left|\frac{d\log}{d\log c}(\phi^h \circ u^h)_c\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right) - \frac{d\log}{d\log c}(\phi^h \circ u^h)_c\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)\right| \leq \bar{m}.\end{aligned} \tag{2.158}$$

2.9.6.6.5 Properties of $w^h(\cdot)$

In this section, we establish that w^h is well-defined and that $((w^h)_{h\in\mathcal{H}}, G)$ is a standard, regular social objective.

Proof w^h is well-defined

First, we argue w^h is well-defined. First, all terms in the integrand within the definition of w^h are

[102] The fact that $\hat{c}^h(V^h(\widetilde{R})) > 0$ implies that u^h twice differentiable and has strictly positive first consumption derivative at all inputs where evaluated above, by Assumption 2. ϕ^h is twice differentiable by assumption. Finally, since $\phi^{h\prime}(V^h(\widetilde{R})) > 0$ by assumption. Together, these observations imply all derivatives and logs used in the Lemma statement are well-defined.

defined for $[V_0^h, u]$ because—as $V^h(R + B_\delta(\mathbf{0}))$ is convex[103]—any $\tilde{u} \in [V_0^h, u]$ is equal to $V^h(\widetilde{R})$ for some $\widetilde{R} \in V^h(R + B_\delta(\mathbf{0}))$, which by Lemma 11 implies $\hat{c}^h(\tilde{u})$ is defined and strictly positive; $u_c^h(\hat{c}^h(\tilde{u}), z_0^h)$ is therefore defined and strictly positive by Assumption 2. Next, note that this integrand is continuous in $\tilde{u}$, by Assumption 2 and the fact that $\hat{c}^h(\tilde{u})$ is continuous (which follows from that $\hat{c}^h(u) = u^h(\cdot, z_0^h)^{-1}(u)$ and—since $\hat{c}^h(u) > 0$—u^h is is locally differentiable with $u_c^h > 0$, by Assumption 2.) Since the integrand is continuous on a closed, bounded domain, the integral in the definition of w^h exists (and is finite).

Proof of standard-ness

We now argue that $((w^h)_{h\in\mathcal{H}}, G)$ is a *standard social objective* in the sense of Definition 1.

First, for all $h \in \mathcal{H}$, w^h is a function from $\mathrm{Im}(u^h)$ to $\mathbb{R} \cup \{-\infty\}$. Moreover, the definition of w^h and the argument for its well-definition in the previous step imply that w^h is always finite-valued.

Second, we claim that for all $h \in \mathcal{H}$ and $\widetilde{R} \in R + B_\delta(\mathbf{0})$, $w^h(u)$ is twice-continuously differentiable on the domain $u \in V^h(R + B_\delta(\mathbf{0}))$. To see that $w^h(u)$ is once-continuously differentiable on this domain follows from the fundamental theorem of calculus and the observation—established in the previous step of the proof—that the integrand in the definition of w^h is continuous in $\tilde{u}$; indeed,

$$w^{h\prime}(u) = \widehat{\lambda}^h \frac{e^{\Phi\left(\widehat{\gamma}^h\left(\log \hat{c}^h(u) - \log c_0^h\right)\right)}}{u_c^h(\hat{c}^h(u), z_0^h)}. \tag{2.159}$$

To establish that $w^{h\prime}(u)$ is itself continuously differentiable, it suffices—given Assumption 2 and the observation in the proof of w^h's well-defined-ness $\hat{c}^h(u) > 0$ for $u \in V^h(R + B_\delta(\mathbf{0}))$—to show that $\hat{c}^h(u)$ is continuously differentiable in u. This in turn follows from the implicit function theorem, $u^h(\hat{c}^h(u), z_0^h)$'s continuous differentiability (from Assumption 2 and $\hat{c}^h(u) > 0$), and that $u_c^h(\hat{c}^h(u), z_0^h) > 0$.

Third, we claim that for all $\widetilde{R} \in R + B_\delta(\mathbf{0})$, $w^h(u)$ is weakly increasing in u for all $h \in \mathcal{H}$ and strictly increasing for a finite measure of $h \in \mathcal{H}$. From (2.159) it is clear that $w^h(u)$ is weakly and moreover strictly increasing for all $h \in \mathcal{H}$ with $\widehat{\lambda}^h \geq 0$ and moreover $\widehat{\lambda}^h > 0$, respectively. It therefore suffices to show $\widehat{\lambda}^h \geq 0$ for all $h \in \mathcal{H}$ and $\widehat{\lambda}^h > 0$ for a finite measure of h. This follows from the definition of $\widehat{\lambda}^h$ and the observation in Appendix 2.9.6.6.1 that $\widehat{\lambda}(z) > 0$ for all $z \in \operatorname{supp} g$.

In order to show that $((w^h)_{h\in\mathcal{H}}, G)$ is a standard social objective, it remains to show that for all $\widetilde{R} \in R + B_\delta(\mathbf{0})$ and all $\Delta, \widetilde{\Delta} \in \boldsymbol{\Delta}$, $w^h(V^h(\widetilde{R}))$, $D_\Delta w^h(V^h(\widetilde{R}))$, and $D^2_{\Delta\widetilde{\Delta}} w^h(V^h(\widetilde{R}))$ are measurable in h. We begin by computing the three terms explicitly, using w^h's twice-continuous differentiability and using the fact that—since $c^h(\widetilde{R})$ and $z^h(\widetilde{R})$ are twice-continuous differentiable by Lemma 2 and strictly positive by integrating the bounds on their derivatives in Lemma 3, and since u^h is twice-continuously differentiable on

[103] This follows from that $R + B_\delta(\mathbf{0})$ is convex and—by Assumption 2 and Lemma 2, plus the fact that within $R + B_\delta(\mathbf{0})$, $c^h(\widetilde{R}), z^h(\widetilde{R})$ are strictly positive (this follows from integrating the bounds on their logs from Lemma 3)—$V^h(\widetilde{R}) = u^h(c^h(\widetilde{R}), z^h(R))$ is continuous in $\widetilde{R}$.

$\mathbb{R}_{>0} \times \mathbb{R}_{>0}$ by 2—$V^h(\widetilde{R}) = u^h(c^h(\widetilde{R}), z^h(\widetilde{R}))$ is twice-continuously Frechet differentiable.

$$\begin{aligned}
w^h(V^h(\widetilde{R})) &= \widehat{\lambda}^h \int\limits_{[V_0^h, V^h(\widetilde{R})]} \frac{e^{\Phi\left(\widehat{\gamma}^h\left(\log \hat{c}^h(\tilde{u}) - \log c_0^h\right)\right)}}{u_c^h(\hat{c}^h(\tilde{u}), z_0^h)} d\tilde{u} \\
D_\Delta w^h(V^h(\widetilde{R})) &= \widehat{\lambda}^h \frac{e^{\Phi\left(\widehat{\gamma}^h\left(\log \hat{c}^h(V^h(\widetilde{R})) - \log c_0^h\right)\right)}}{u_c^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h)} D_\Delta V^h(\widetilde{R}) \\
D^2_{\Delta\widetilde{\Delta}} w^h(V^h(\widetilde{R})) &= \widehat{\lambda}^h \frac{e^{\Phi\left(\widehat{\gamma}^h\left(\log \hat{c}^h(V^h(\widetilde{R})) - \log c_0^h\right)\right)}}{u_c^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h)} \\
&\quad \cdot \Bigg[\widehat{\gamma}^h \Phi'\left(\widehat{\gamma}^h\left(\log \hat{c}^h(V^h(\widetilde{R})) - \log c_0^h\right)\right) \frac{D_{\widetilde{\Delta}} V^h(\widetilde{R}) D_\Delta V^h(\widetilde{R})}{u_c^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h) \hat{c}^h(V^h(\widetilde{R}))} \\
&\qquad - \frac{u_{cc}^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h)}{u_c^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h)} \frac{D_{\widetilde{\Delta}} V^h(\widetilde{R}) D_\Delta V^h(\widetilde{R})}{u_c^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h)} + D^2_{\Delta\widetilde{\Delta}} V^h(\widetilde{R}) \Bigg]
\end{aligned} \tag{2.160}$$

where in the last two lines we have used that—by the implicit function theorem—$\hat{c}^{h\prime}(V^h(\widetilde{R})) = u_c^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h)^{-1}$.

Using the expressions above, the facts that $D_\Delta w^h(V^h(\widetilde{R}))$ and $D^2_{\Delta\widetilde{\Delta}} w^h(V^h(\widetilde{R}))$ are measurable follow from that

- $\widehat{\lambda}^h$ and $\widehat{\gamma}^h$ are measurable (see Appendices 2.9.6.6.1 and 2.9.6.6.3),
- $c^h(V^h(\widetilde{R}))$ is measurable (see Appendix 2.9.6.6.4),
- by Assumption 2 the first two derivatives of u^h in c and/or z are measurable,
- the first two Frechet derivatives of $V^h(\widetilde{R})$ are products of the first two derivatives of $u^h(c^h(\widetilde{R}), z^h(\widetilde{R}))$ (which are measurable by Assumption 2) and the first two Frechet derivatives of $c^h(\widetilde{R})$ and $z^h(\widetilde{R})$ (which are measurable by Lemma 2), and
- products, sums, and (with non-zero denominators) quotient of measurable functions measurable, as are compositions of measurable functions with continuous functions and/or with measurable functions.

It remains to show that $w^h(V^h(\widetilde{R}))$ is measurable. To see this, note that by substituting $\tilde{u} = V_0^h + \alpha(V^h(\widetilde{R}) - V_0^h)$, it may be rewritten as

$$w^h(V^h(\widetilde{R})) = \widehat{\lambda}^h \left(V^h(\widetilde{R}) - V_0^h\right) \int\limits_{[0,1]} \frac{e^{\Phi\left(\widehat{\gamma}^h\left(\log \hat{c}^h(V_0^h + \alpha(V^h(\widetilde{R}) - V_0^h)) - \log c_0^h\right)\right)}}{u_c^h(\hat{c}^h(V_0^h + \alpha(V^h(\widetilde{R}) - V_0^h)), z_0^h)} d\alpha \tag{2.161}$$

Note that $V^h(R)$ and V_0^h are measurable as they are the composition of $u^h(c, z)$ (which is jointly measurable in (c, z, h)) with $c^h(\widetilde{R})$ and $z^h(\widetilde{R})$ and $c^h(R)$ and $z^h(R)$, respectively, all of which are measurable by Lemma 2. This fact, the observation—made in the proof that w^h is well-defined—that $V^h(R + B_\delta(\mathbf{0}))$ is convex, and the arguments used above to establish the measurability of $w^h(V^h(\widetilde{R}))$'s derivatives imply that the *integrand* in (2.161) is measurable in h. It remains to argue that the *integral* in (2.161) is measurable. To see this note that the integrand is continuous in α; this follows from same the argument used (in the proof that w^h is well-defined) to argue that the integrand in w^h's definition is continuous. Lemma 4.51 of Aliprantis and Border (2006) then implies the inner integrand is jointly measurable in (α, h). Finally, the Fubini-Tonelli theorem (as stated in 2.37 of Folland (1999)) then implies that the integral of (2.161) is measurable in h, as desired.

Proof of regularity, part 1

We next establish that $((w^h)_{h\in\mathcal{H}}, G)$ is *regular* in the sense of Definition 3. Specifically, we establish in this step the existence of integrable functions $b_0, b_1, b_2 : \mathcal{H} \to \mathbb{R}$ such that for all $\widetilde{R} \in R + B_\delta(\mathbf{0})$,

$$\begin{aligned} \left|(w^h \circ u^h)(c^h(\widetilde{R}), z^h(\widetilde{R}))\right| &\leq b_0(h), \\ \left|c^h(\widetilde{R})(w^h \circ u^h)_c(c^h(\widetilde{R}), z^h(\widetilde{R}))\right| &\leq b_1(h), \\ \text{and} \qquad \left|c^h(\widetilde{R})^2(w^h \circ u^h)_{cc}(c^h(\widetilde{R}), z^h(\widetilde{R}))\right| &\leq b_2(h); \end{aligned} \tag{2.162}$$

We establish the other part of the definition of regularity in the next step.

We begin by proving the bound on the level of welfare *using the bounds on the first two derivatives*; we then *independently* establish the bounds on the derivatives. To start, note that by the definition of w^h, $(w^h \circ u^h)(c_0^h, z_0^h) = 0$ for all $h \in \mathcal{H}$. So applying Taylor's theorem to the path between R and any $\widetilde{R} \in R + B_\delta(\mathbf{0})$—which we may do since $w^h(V^h(\widetilde{R}))$ is twice-continuously Frechet differentiable for $\widetilde{R} \neq R \in R + B_\delta(\mathbf{0})$—gives us

$$w^h(V^h(\widetilde{R})) \;=\; D_{\widetilde{R}-R} w^h(V^h(R)) \;+\; \frac{1}{2} D^2_{\widetilde{R}-R,\widetilde{R}-R} w^h(V^h(R + \alpha(\widetilde{R} - R))) \tag{2.163}$$

for some $\alpha \in [0, 1]$, where here we have used that $\widetilde{R} - R \in \mathbf{\Delta}$. By the expressions (2.141) and (2.142) for the derivatives of $w^h(V^h(R))$ in the proof of Lemma 5, by Assumptions 1 and 4, by the existence of a uniform bound on $D \log z^h(\widetilde{R})$ (from Lemma 3; see Footnote 56), by the definition of $||\cdot||$, and finally by the existence of the desired bounds on the derivatives of $(w^h \circ c^h)$, there exist constants k_0^{11}, k_0^{21}, and k_0^{22} such that, for all $h \in \mathcal{H}$, $\widetilde{R} \in R + B_\delta(\mathbf{0})$, non-zero $\Delta \in \mathbf{\Delta}$

$$\begin{aligned} |D_{\widetilde{R}-R} w^h(V^h(R))| &\leq k_0^{11} ||\widetilde{R} - R|| b_1(h) \\ |D^2_{\widetilde{R}-R,\widetilde{R}-R} w^h(V^h(R))| &\leq ||\widetilde{R} - R||^2 \big(k_0^{21} b_1(h) + k_0^{22} b_2(h)\big) \end{aligned} \tag{2.164}$$

Combining these bounds with (2.163), we have

$$|(w^h \circ u^h)(c^h(\widetilde{R}), z^h(\widetilde{R}))| \leq \; \delta k_0^{11} b_1(h) \;+\; \big(k_0^{21} b_1(h) + k_0^{22} b_2(h)\big) \frac{\delta^2}{2} \;\equiv\; b_0(h) \tag{2.165}$$

We now proceed to the first derivative. From (2.159), we have

$$\begin{aligned} \left|c^h(\widetilde{R})\; (w^h \circ u^h)_c(c^h(\widetilde{R}), z^h(\widetilde{R}))\right| &= \left|c^h(\widetilde{R})\; w^{h\prime}(V^h(\widetilde{R}))\; u_c^h(c^h(\widetilde{R}), z^h(\widetilde{R}))\right| \\ &= c_0^h \widehat{\lambda}^h \left|\frac{c^h(\widetilde{R})}{c_0^h}\right| \left|\frac{u_c^h(c^h(\widetilde{R}), z^h(\widetilde{R}))}{u_c^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h)}\right| \underbrace{e^{\Phi\left(\widehat{\gamma}^h\left(\log \hat{c}^h(V^h(\widetilde{R})) - \log c_0^h\right)\right)}}_{\leq e^{\overline{\Phi}}} \end{aligned} \tag{2.166}$$

Now note that $\left|\frac{c^h(\widetilde{R})}{c_0^h}\right|$ is uniformly bounded across $h \in \mathcal{H}$ and $\widetilde{R} \in R + B_\delta(\mathbf{0})$; this follows from integrating the fact that, by Lemma 3 (see Footnote 56), $|D \log c^h(\widetilde{R})|$ is uniformly bounded. Moreover, $\left|\frac{u_c^h(c^h(\widetilde{R}), z^h(\widetilde{R}))}{u_c^h(\hat{c}^h(u), z_0^h)}\right|$

is uniformly bounded by Lemma 11 (see Appendix 2.9.6.6.4). So there exists some uniform constant k_1 such that for all $h \in \mathcal{H}$, $\widetilde{R} \in R + B_\delta(\mathbf{0})$,

$$\left| c^h(\widetilde{R})(w^h \circ u^h)_c(c^h(\widetilde{R}), z^h(\widetilde{R})) \right| \leq k_1 c_0^h \widehat{\lambda}^h \equiv b_1(h). \tag{2.167}$$

As we have established in the first step of this proof that $c_0^h \widehat{\lambda}^h \equiv b_1(h)$ is integrable, this establishes the desired bound for the first derivative.

Finally, consider the second derivative. We start with the observation that, by (2.159),

$$(w^h \circ u^h)_c(c^h(\widetilde{R}), z^h(\widetilde{R})) = \widehat{\lambda}^h \frac{u_c^h\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)}{u_c^h\left(\hat{c}^h\left(u^h(c^h(\widetilde{R}), z^h(\widetilde{R}))\right), z_0^h\right)} e^{\Phi\left(\widehat{\gamma}^h\left(\log \hat{c}^h(u^h(c^h(\widetilde{R}), z^h(\widetilde{R}))) - \log c_0^h\right)\right)} \tag{2.168}$$

An important trick for the remainder of this step is to note that for any differentiable function $\phi^h : V^h(R + B_\delta(\mathbf{0})) \to \mathbb{R}$ with $\phi^{h\prime}\left(u^h(c^h(\widetilde{R}), z^h(\widetilde{R}))\right) > 0$ we have

$$\frac{u_c^h\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)}{u_c^h\left(\hat{c}^h\left(u^h(c^h(\widetilde{R}), z^h(\widetilde{R}))\right), z_0^h\right)} = \frac{(\phi^h \circ u^h)_c\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)}{(\phi^h \circ u^h)_c\left(\hat{c}^h\left(u^h(c^h(\widetilde{R}), z^h(\widetilde{R}))\right), z_0^h\right)} \tag{2.169}$$

We will work with the following function ϕ^h, which (as we will show) has several useful properties:

$$\phi^h(u) = \int\limits_{[u_0^h, u]} \exp\left[\int\limits_{[u_0^h, u']} \frac{1 - \frac{u_{cc}^h(\hat{c}^h(\tilde{u}), z_0^h)\hat{c}^h(\tilde{u})}{u_c^h(\hat{c}^h(\tilde{u}), z_0^h)}}{u_c^h(\hat{c}^h(\tilde{u}), z_0^h)\ \hat{c}^h(\tilde{u})} d\tilde{u}\right] du' \tag{2.170}$$

To see that $\phi^h(u)$ is well-defined, we first note that (by the same argument used for the well-definedness of w^h) $\hat{c}^h(\cdot)$ is well-defined and strictly positive where evaluated in the definition. That the inner integral used in the definition exists follows from that the integrand is defined—since $\hat{c}^h(\tilde{u}) > 0$ and therefore $u_c^h(\hat{c}^h(\tilde{u}), z_0^h) > 0$ (by Assumptions 2 and 3)—and that the integrand is continuous in $\tilde{u}$—since $\hat{c}^h(\tilde{u})$ is continuous in $\tilde{u}$ (because $u_c^h > 0$ by Assumption 2) and by Assumption 2 u^h is twice-continuously differentiable. That the outer integral used in the definition exists follows from that the outer integral exists (since the inner integral is defined) and is continuous in u' (since the inner integrand's continuity allows us to apply the fundamental theorem of calculus).

The same fundamental-theorem-of-calculus arguments also imply that $\phi^h(u)$ is continuously differentiable with

$$\phi^{h\prime}(u) = \exp\left[\int\limits_{[u_0^h, u]} \frac{1 - \frac{u_{cc}^h(\hat{c}^h(\tilde{u}), z_0^h)\hat{c}^h(\tilde{u})}{u_c^h(\hat{c}^h(\tilde{u}), z_0^h)}}{u_c^h(\hat{c}^h(\tilde{u}), z_0^h)\ \hat{c}^h(\tilde{u})} d\tilde{u}\right] \tag{2.171}$$

and that $\log \phi^{h\prime}(u)$ is continuously differentiable with

$$\frac{d}{du} \log \phi^{h\prime}(u) = \frac{1 - \frac{u_{cc}^h(\hat{c}^h(u), z_0^h)\hat{c}^h(u)}{u_c^h(\hat{c}^h(u), z_0^h)}}{u_c^h(\hat{c}^h(u), z_0^h)\ \hat{c}^h(u)}. \tag{2.172}$$

Note that $\log \phi^{h\prime}(u)$ and $\exp[\cdot]$ are both continuously differentiable, so is $\exp\left[\log \phi^{h\prime}(u)\right] = \phi^{h\prime}(u)$, i.e. $\phi^{h\prime}(u)$ is twice-continuously differentiable. Also, by (2.171), $\phi^{h\prime}(u) > 0$ for all u in ϕ^h's domain. Finally, note that for all $h \in \mathcal{H}$, $\widetilde{R} \in R + B_\delta(\mathbf{0})$,

$$\begin{aligned}
&\frac{(\phi^h \circ u^h)_{cc}\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)\ \hat{c}^h(V^h(\widetilde{R}))}{(\phi^h \circ u^h)_c\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)} \\
&= \frac{\phi^{h\prime\prime}(V^h(\widetilde{R}))}{\phi^{h\prime}(V^h(\widetilde{R}))} u_c^h\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)\hat{c}^h(V^h(\widetilde{R})) + \frac{u_{cc}^h\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)\hat{c}^h(V^h(\widetilde{R}))}{u_c^h\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)} \qquad (2.173)\\
&= 1
\end{aligned}$$

by (2.172).

Returning to (2.168), we now substitute using (2.169) and differentiate in order to compute $c^h(\widetilde{R})^2\ (w^h \circ u^h)_{cc}(c^h(\widetilde{R}), z^h(\widetilde{R}))$:

$$\begin{aligned}
&c^h(\widetilde{R})^2\ (w^h \circ u^h)_{cc}(c^h(\widetilde{R}), z^h(\widetilde{R})) \\
&\quad = c_0^h \widehat{\lambda}^h \frac{c^h(\widetilde{R})}{c_0^h} \frac{(\phi^h \circ u^h)_c\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)}{(\phi^h \circ u^h)_c\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)} e^{\Phi\left(\widehat{\gamma}^h\left(\log \hat{c}^h(V^h(\widetilde{R})) - \log c_0^h\right)\right)} \\
&\qquad \cdot \left[\frac{(\phi^h \circ u^h)_{cc}\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)\ c^h(\widetilde{R})}{(\phi^h \circ u^h)_c\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)}\right. \qquad (2.174)\\
&\qquad\quad - \frac{(\phi^h \circ u^h)_{cc}\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)\ \hat{c}^h(V^h(\widetilde{R}))}{(\phi^h \circ u^h)_c\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)} \frac{c^h(\widetilde{R})}{\hat{c}^h(V^h(\widetilde{R}))} \cancel{\hat{c}^{h\prime}(V^h(\widetilde{R}))}\ \cancel{u_c^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h)} \\
&\qquad\quad \left. + \widehat{\gamma}^h \Phi'\left(\widehat{\gamma}^h\left(\log \hat{c}^h(V^h(\widetilde{R})) - \log c_0^h\right)\right) \frac{c^h(\widetilde{R})}{\hat{c}^h(V^h(\widetilde{R}))} \cancel{\hat{c}^{h\prime}(V^h(\widetilde{R}))}\ \cancel{u_c^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h)}\right]
\end{aligned}$$

where the cancellations are by our earlier observation that $\hat{c}^{h\prime}(V^h(\widetilde{R})) = u_c^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h)^{-1}$. Next, recall that by Lemma 11 (see 2.9.6.6.4), there exists $\bar{m}$ such that for all $h \in \mathcal{H}, R \in R + B_\delta(\mathbf{0})$,

$$\begin{aligned}
&\left|\log \hat{c}^h\left(V^h(\widetilde{R})\right) - \log c^h\left(\widetilde{R}\right)\right| \le \bar{m} \\
&\left|\log\left[(\phi^h \circ u^h)_c\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)\right] - \log\left[(\phi^h \circ u^h)_c\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)\right]\right| \le \bar{m} \qquad (2.175)\\
\text{and} \quad &\left|\frac{d\log}{d\log c}(\phi^h \circ u^h)_c\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right) - \frac{d\log}{d\log c}(\phi^h \circ u^h)_c\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)\right| \le \bar{m}.
\end{aligned}$$

Also recalling (as argued in the bound of the first derivative), there exists some uniform bound $\hat{m}$ on $\left|\frac{c^h(\widetilde{R})}{c_0^h}\right|$, we may use (2.174) to bound

$$
\begin{aligned}
\Big| c^h(\widetilde{R})^2 \quad & (w^h\circ u^h)_{cc}(c^h(\widetilde{R}), z^h(\widetilde{R}))\Big| \\
&\leq c_0^h \widehat{\lambda}^h \cdot \hat{m} e^{\bar{m}} e^{\overline{\Phi}} \left[\bar{m} + \frac{(\phi^h\circ u^h)_{cc}\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)\ \hat{c}^h(V^h(\widetilde{R}))}{(\phi^h\circ u^h)_{c}\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)} \right. \\
&\qquad \left. + \frac{(\phi^h\circ u^h)_{cc}\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)\ \hat{c}^h(V^h(\widetilde{R}))}{(\phi^h\circ u^h)_{c}\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)} e^{\bar{m}} \quad + \quad \left|\widehat{\gamma}^h\right| \overline{\Phi} e^{\bar{m}} \right] \\
&= c_0^h \widehat{\lambda}^h \cdot \hat{m} e^{\bar{m}} e^{\overline{\Phi}} \left(\bar{m} + 1 + e^{\bar{m}}\right) \quad + \quad \left| c_0^h \widehat{\lambda}^h \widehat{\gamma}^h \right| \cdot \hat{m} e^{2\bar{m}} e^{\overline{\Phi}} \overline{\Phi} \quad \equiv \quad b_2(h)
\end{aligned}
\tag{2.176}
$$

which is integrable since we have—in earlier steps of this proof—shown that $c_0^h\widehat{\lambda}^h$ and $c_0^h\widehat{\lambda}^h\widehat{\gamma}^h$ are integrable. To reach the last line we have used that $c_0^h, \widehat{\lambda}^h \geq 0$ and we have used (2.173).

Proof of regularity, part 2

We now establish that $((w^h)_{h\in\mathcal{H}}, G)$ satisfies the second component of Definition 3. To do so, we must show that several income-conditional expectations are continuous or continuously differentiable in income.

Since these expectations include products with $\lambda^h(R) \equiv (w^h\circ u^h)_c(c_0^h, z_0^h)$ and $(\lambda\gamma)^h(R) \equiv c_0^h (w^h \circ u^h)_{cc}(c_0^h, z_0^h)$, we first compute expressions of these variables. From (2.168) and (2.174), and because $\hat{c}^h(V_0^h) = c_0^h$, we have

$$
\lambda^h(R) = (w^h\circ u^h)_c(c_0^h, z_0^h) = \widehat{\lambda}^h \qquad \text{and} \qquad (\lambda\gamma)^h(R) = c_0^h\ (w^h\circ u^h)_{cc}(c_0^h, z_0^h) = \widehat{\lambda}^h\widehat{\gamma}^h \tag{2.177}
$$

Now, recall from Appendices 2.9.6.6.2 and 2.9.6.6.3 the expectations of $R(z_0^h)\widehat{\lambda}^h$, $R(z_0^h)\widehat{\lambda}^h$, $R(z_0^h)\widehat{\lambda}^h\eta^h(R)$, $R(z_0^h)\widehat{\lambda}^h\varepsilon^h(R)$, $R(z_0^h)\widehat{\lambda}^h\frac{\eta^h(R)^2}{\varepsilon^h(R)}$, $R(z_0^h)\widehat{\lambda}^h\widehat{\gamma}^h$ conditional on income z_0^h exist and are equal to $R(z_0^h)\widehat{\lambda}(z_0^h)$, $R(z_0^h)\widehat{\lambda}(z_0^h)\eta_{\leq}(z_0^h;\epsilon)$, $R(z_0^h)\widehat{\lambda}(z_0^h)\varepsilon_{\leq}(z_0^h;\epsilon)$ and $R(z_0^h)\widehat{\lambda}(z_0^h)\left(\frac{\eta^2}{\varepsilon}\right)_{\leq}(z_0^h;\epsilon)$, and $R(z_0^h)(\widehat{\lambda\gamma})(z_0^h)$ respectively (where the moments $x_{\leq}(z;\epsilon)$ are as in Assumption 6). From our observations in those sections as well as Assumption 1, it moreover follows that $R(z)\widehat{\lambda}(z)$, $R(z)\widehat{\lambda}(z)\varepsilon_{\leq}(z;\epsilon)$ and $R(z)\widehat{\lambda}(z)\left(\frac{\eta^2}{\varepsilon}\right)_{\leq}(z;\epsilon)$, and $R(z)(\widehat{\lambda\gamma})(z)$ are continuous in $z\in\operatorname{supp} g$; and that $R(z)\widehat{\lambda}(z)\eta_{\leq}(z;\epsilon)$ are continuously differentiable in $z\in\operatorname{supp} g$.

Combining these observations with (2.177) implies that expectations of $R(z_0^h)\lambda^h(R)$, $R(z_0^h)\lambda^h(R)\eta^h(R)$, $R(z_0^h)\lambda^h(R)\varepsilon^h(R)$, $R(z_0^h)\lambda^h(R)\frac{\eta^h(R)^2}{\varepsilon^h(R)}$, $R(z_0^h)(\lambda\gamma)^h(R)$ conditional on income z_0^h exist and are equal to $R(z_0^h)\widehat{\lambda}(z_0^h)$, $R(z_0^h)\widehat{\lambda}(z_0^h)\eta_{\leq}(z_0^h;\epsilon)$, $R(z_0^h)\widehat{\lambda}(z_0^h)\varepsilon_{\leq}(z_0^h;\epsilon)$ and $R(z_0^h)\widehat{\lambda}(z_0^h)\left(\frac{\eta^2}{\varepsilon}\right)_{\leq}(z_0^h;\epsilon)$, and $R(z_0^h)(\widehat{\lambda\gamma})(z_0^h)$ respectively, and so have the continuity and differentiability properties described above on $\operatorname{supp} g$.

2.9.6.6.6 Strengthening Lemmas 6 and 7

We claim that—under Assumption 6 and for the welfare function defined in the main proof of Theorem 2—Lemmas 6 and 7 hold for all $\Delta \in \boldsymbol{\Delta}$. From the lemmas' proofs it is clear that we need only show that (i) for all $\Delta\in\boldsymbol{\Delta}$, each additive term of (2.54), (2.57), and (2.65) is integrable over $\operatorname{supp} g$ in isolation, and (ii) the *integration by parts* steps are valid for all $\Delta\in\boldsymbol{\Delta}$ (rather than just the specific Δs described in the lemmas). To do so, first we deal with (i) by providing appropriate bounds on each term and second deal with (ii) by (a) providing a general result about integration by parts on $\operatorname{supp} g$, (b) applying it to Lemma 6, and (c) applying it to Lemma 7.

Bounds on revenue and welfare derivative terms

We claim that each additive term of the integral expressions in (2.54), (2.57), and (2.65) is integrable in isolation.

By dominated convergence—and since by the definition of $\boldsymbol{\Delta}$, Assumption (1), and the arguments in the proofs of Lemmas 6 and 7 each term is continuously differentiable and therefore measurable on $\operatorname{supp} g$—it suffices to show that each is bounded by an integrable function. Because (a) by Assumption 4, $zg(z)$ and $R(z)g(z)$ are integrable on $\operatorname{supp} g$, and (b) by the definition of $\boldsymbol{\Delta}$, $\frac{\Delta(z)}{R(z)}$, $\frac{\Delta'(z)}{R'(z)}$ and $\frac{\Delta''(z)z}{R'(z)}$ are bounded by $||\Delta||$ on $\operatorname{supp} g$, it suffices to show that (i) the terms $(1-R'(z))z\eta(z)$, $(1-R'(z))z\varepsilon(z)$, and $\Delta(z)$ from (2.54), (ii) the terms $A(z), B(z), C(z), D(z), E(z)$ in (2.57), and (iii) the terms $A(z), B(z), C(z)$ in (2.65) are each bounded by linear combinations of z and $R(z)$.

In the case of (2.54), this is immediate from Assumptions 1, 4, and the definition of $\boldsymbol{\Delta}$.

In the case of (2.57), this is immediate from Assumptions 1 and 4, and the definition of the conditional expectations used in the definitions of $A(z), B(z), C(z), D(z), E(z)$.

For the case of (2.65), we use the observations proven in Appendices 2.9.6.6.1, 2.9.6.6.2, and 2.9.6.6.3 that, over all $z \in \operatorname{supp} g$, $(\lambda\gamma)(z) = (\widehat{\lambda}\gamma)(z)$ is bounded by a linear combination of z and $R(z)$; $(\lambda\varepsilon)(z) = (\widehat{\lambda}\varepsilon)(z), (\lambda\eta)(z) = (\widehat{\lambda}\eta)(z), (\lambda\frac{\eta^2}{\varepsilon})(z) = (\widehat{\lambda}\frac{\eta^2}{\varepsilon})(z) \leq \epsilon\widehat{\lambda}(z) = \epsilon\lambda(z)$ for ϵ the constant defined in the main proof of Theorem 2; and $\lambda(z) = \widehat{\lambda}(z)$ is bounded by a linear combination of z and $R(z)$.[104] Combining these observations with Assumption 1 gives the desired bounds.

General integration by parts on $\operatorname{supp} g$

Lemma 8. *Suppose that $F(z), G(z) : \operatorname{supp} g \to \mathbb{R}$ are continuously differentiable. Moreover suppose $F(z)$ and $G(z)$ are bounded except possibly in limits as $z \to 0$ and/or $z \to \infty$,[105] and that for all sequences $(z_n) \subset \operatorname{supp} g$ that converge either to 0 or to ∞, $\lim_{n\to\infty} F(z_n)G(z_n)g(z_n) = 0$. Then if $F'(z)G(z)g(z)$ and $F(z)\frac{d(G(z)g(z))}{dz}$ are bounded on $\operatorname{supp} g$ by integrable functions,[106] we have*

$$\int\limits_{\operatorname{supp} g} F'(z)G(z)g(z)dz = -\int\limits_{\operatorname{supp} g} F(z)\frac{d\left(G(z)g(z)\right)}{dz}dz. \tag{2.178}$$

Proof. We will use throughout that, by Assumption 5, g is continuously differentiable.

To start, note that because—by the continuous differentiability of $F(z)$, $G(z)$, and $g(z)$ on $\operatorname{supp} g$—$F'(z)G(z)g(z)$ and $F(z)\frac{d(G(z)g(z))}{dz}$ are measurable on $\operatorname{supp} g$, and because they are by assumption bounded by integrable functions, they are integrable on $\operatorname{supp} g$ by dominated convergence.

Now, recalling Lemma 9, let $\mathcal{B}$ be a countable set of disjoint, open, positive intervals, so that $\operatorname{supp} g = \cup_{I\in\mathcal{B}} I$. By the countable additivity of Lebesgue integration,

$$\begin{aligned}\int\limits_{\operatorname{supp} g} F'(z)G(z)g(z)dz &= \sum_{I\in\mathcal{B}}\int\limits_{I} F'(z)G(z)g(z)dz \\ \text{and} \quad \int\limits_{\operatorname{supp} g} F(z)\frac{d\left(G(z)g(z)\right)}{dz}dz &= \sum_{I\in\mathcal{B}}\int\limits_{I} F(z)\frac{d\left(G(z)g(z)\right)}{dz}dz\end{aligned} \tag{2.179}$$

[104] Our ability to interchange λ and $\widehat{\lambda}$ is established in Appendix 2.9.6.6.5.

[105] More formally, for all $a, b \in \mathbb{R}_{>0}$, $F(z)$ and $G(z)$ are bounded within $\operatorname{supp} g \cap [a, b]$.

[106] I.e. functions that integrable as random variables on the measure space defined by restricting the standard measure space on $\mathbb{R}$ to $\operatorname{supp} g$ (which note by Lemma 9 is measurable).

To complete the proof it therefore suffices to show that for each $I \in \mathcal{I}$,

$$\int\limits_I F'(z)G(z)g(z)dz = -\int\limits_I F(z)\frac{d\left(G(z)g(z)\right)}{dz}dz. \tag{2.180}$$

To this end, fix any $I = (a, b) \in \mathcal{B}$. For any decreasing sequence $a_n \to a$ and increasing sequence $b_n \to b$ so that $a < a_n < b_n < b$, the fact that $F(z)$ and $G(z)g(z)$ are defined and continuously differentiable on $[a_n, b_n]$ implies we may integrate by parts:

$$\int\limits_{[a_n,b_n]} F'(z)G(z)g(z)dz = F(z)G(z)g(z)\Big|_{a_n}^{b_n} - \int\limits_{[a_n,b_n]} F(z)\frac{d\left(G(z)g(z)\right)}{dz}dz \tag{2.181}$$

where note the integral on the RHS exists because the integrand is continuous on the closed interval over which it is integrated. Taking the limit as $a_n \to a$ and $b_n \to b$, we have by the continuity of Lebesgue integration (and that both integrands are integrable) that

$$\int\limits_{(a,b)} F'(z)G(z)g(z)dz = \lim_{z \xrightarrow{-} b} F(z)G(z)g(z) - \lim_{z \xrightarrow{+} a} F(z)G(z)g(z) - \int\limits_{(a,b)} F(z)\frac{d\left(G(z)g(z)\right)}{dz}dz \tag{2.182}$$

so long as the first two limits on the RHS exist. We begin with the limit from below to b. If $b < \infty$, then since $b > 0$ by construction, since $h(b) = 0$ and h is continuous, and since by assumption $F(z)G(z)$ is bounded in the vicinity of b, the limit is zero. Alternatively, if $b = \infty$, then the limit is zero by assumption. Next, consider the limit from above to a. If $a > 0$, then since $a < \infty$ by construction, the same argument used in the $b < \infty$ case implies the limit is zero. If instead $a = 0$, then the limit is zero by assumption. So both limits exist and equal zero, completing the proof. □

Application to proof of Lemma 6

It suffices to apply Lemma 8 to each instance of integration by parts in the proof.

We begin with (2.55). We take $F(z) = \Delta(z)$ and $G(z) = g(z)\frac{1-R'(z)}{R'(z)}\varepsilon(z)z$. The proof of Lemma 8 argues that F and G are continuously differentiable. That they are bounded except in limits to 0 or ∞ follows from (a) $\varepsilon(z)$'s boundedness (Assumption 4), (b) $g(z)$'s continuity (Assumption 5), (c) $R'(z)$'s continuity (Assumption 1) and (d) Δ's continuity. Now, consider the limit condition: To see it, note that because $|\Delta(z)| \leq ||\Delta|||R(z)|$, $\varepsilon(z)$ is bounded, and by Assumption 6 $\frac{R(z)}{R'(z)z}$ is bounded on $\operatorname{supp} g$, we have that, for all $z \in \operatorname{supp} g$ $|F(z)G(z)g(z)| \leq k_1 R(z)zg(z) + k_2 z^2 g(z)$ for some $k_1, k_2 \in \mathbb{R}_{\geq 0}$. By Assumption 6, these bounds go to 0 in limits as $z \to 0$ or $z \to \infty$. It remains to show $F'(z)G(z)g(z)$ and $F(z)\frac{d(G(z)g(z))}{dz}$ are bounded on $\operatorname{supp} g$ by integrable functions. To see this, note that by the observations above and Assumptions 1 and 6, there exist constants $k_n \in \mathbb{R}_{\geq 0}$ such that for all $z \in \operatorname{supp} g$,[107]

$$\begin{aligned}
&|F'(z)G(z)g(z)| \leq k_3 zg(z) + k_4 R(z)g(z) \\
&F(z)\frac{d\left(G(z)g(z)\right)}{dz} = \varepsilon(z)\Bigg[-\frac{1-R'(z)}{R'(z)}\alpha(z) + \underbrace{\frac{d}{d\log z}\left(\frac{1-R'(z)}{R'(z)}\right)}_{=-\frac{1}{R'(z)}\frac{d\log R'(z)}{d\log z}} + \frac{1-R'(z)}{R'(z)}\frac{d\log\varepsilon(z)}{d\log z}\Bigg]\Delta(z)g(z) \\
&\left|F(z)\frac{d\left(G(z)g(z)\right)}{dz}\right| \leq \left(k_5 + k_6\frac{1}{R'(z)}\right)R(z)g(z) \leq k_5 R(z)g(z) + k_7 zg(z)
\end{aligned} \tag{2.183}$$

[107]Here, we use that $\alpha(z) \equiv -\frac{d\log zg(z)}{d\log z}$ is well-defined since h is differentiable by Assumption 5 and $g(z) > 0$ on $\operatorname{supp} g$; $\frac{d\log\varepsilon(z)}{d\log z}$ is well-defined on $\operatorname{supp} g$ since $\varepsilon(z)$ is continuously differentiable by Assumption 5 and strictly positive by Assumption 4 and the fact that compensated elasticities are always positive (see the proof of Lemma 2).

Note that $zg(z)$ and $R(z)g(z)$ are integrable functions on $\operatorname{supp} g$, since their integrals over $\operatorname{supp} g$ correspond to those of z_0^h and c_0^h over $h \in \mathcal{H}$ and the latter are integrable by Assumption 4.

Second, we consider the $B(z)$ term in (2.58). We take $F(z) = \frac{\Delta(z)^2}{2}$ and $G(z) = \frac{B(z)}{R(z)R'(z)}$. Note that $F(z)$ and $G(z)$ are continuously differentiable on $\operatorname{supp} z$ by Assumption 1, $B(z)$'s continuous differentiability (see the proof of Lemma 6), the definition of $\mathbf{\Delta}$, and the fact that $R(z)$ and $R'(z)$ are strictly positive on $\mathbb{R}_{>0}$ (see the proof of Lemma 2). Next, note that since $R(z)$ and $R'(z)$ are also continuous in z (Assumption 1), $R(z), R'(z) > 0$ on $\mathbb{R}_{>0}$, $|\Delta(z)| \leq ||\Delta|||R(z)|$, and $B(z)$ is bounded by a linear combination of z and $R(z)$ (shown above), $F(z)$ and $G(z)$ are both bounded on any domain $\operatorname{supp} g \cap [a, b]$, $a, b \in \mathbb{R}_{>0}$. These observations, combined with the fact that $\frac{R(z)}{R'(z)z}$ is bounded on $\operatorname{supp} g$ (Assumption 6), imply $|F(z)G(z)g(z)|$ is bounded on $\operatorname{supp} g$ by a linear combination of $z^2 g(z)$ and $zR(z)g(z)$; so Assumption 6 ensures that for any sequence $(z_n) \subset \operatorname{supp} g$ such that $z_n \to 0$ or ∞, we have $\lim_{n\to\infty} F(z_n)G(z_n)g(z_n) = 0$. Differentiating the expression (2.57) for $B(z)$ reveals that $zB'(z)$ is bounded on $\operatorname{supp} g$ by a linear combination of z and $R(z)$.[108] This fact, the bound on $B(z)$, the fact that $|\Delta(z)| \leq ||\Delta|||R(z)|$ and $|\Delta'(z)| \leq ||\Delta|||R'(z)|$, and Assumptions 1 and 6, give us that

$$
\begin{aligned}
|F'(z)G(z)g(z)| &= \left|\Delta(z)\Delta'(z)\frac{B(z)}{R(z)R'(z)}g(z)\right| \leq k_1 zg(z) + k_2 R(z)g(z) \\
F(z)\frac{d\left(G(z)g(z)\right)}{dz} &= \left(\frac{B'(z)z}{R(z)R'(z)z} + \frac{B(z)}{R(z)R'(z)z}\frac{d\log}{d\log z}\left(\frac{zg(z)}{R(z)R'(z)z}\right)\right)\frac{\Delta(z)^2}{2}g(z) \\
&= \frac{R(z)}{R'(z)z}\left(\frac{B'(z)z}{R(z)^2} - \frac{B(z)}{R(z)^2}\left(\alpha(z) + 1 + \frac{d\log R(z)}{d\log z} + \frac{d\log R'(z)}{d\log z}\right)\right)\frac{\Delta(z)^2}{2}g(z) \\
\left|F(z)\frac{d\left(G(z)g(z)\right)}{dz}\right| &\leq k_3 zg(z) + k_4 R(z)g(z)
\end{aligned}
\tag{2.185}
$$

for all $z \in \operatorname{supp} g$, for various constants $k_n \in \mathbb{R}_{\geq 0}$. By Assumption 4, these bounds are integrable, as desired.

Third, we consider the $E(z)$ term in (2.58). We take $F(z) = \frac{\Delta'(z)^2}{2}$ and $G(z) = \frac{zE(z)}{R'(z)^2}$. Note that $F(z)$ and $G(z)$ are continuously differentiable on $\operatorname{supp} z$ by Assumptions 1, $E(z)$'s continuous differentiability (see the proof of Lemma 6), the definition of $\mathbf{\Delta}$, and the fact that $R'(z)$ is strictly positive on $\mathbb{R}_{>0}$ (see the proof of Lemma 2). Since $R(z)$ and $R'(z)$ are also continuous in z (Assumption 1), $|\Delta'(z)| \leq ||\Delta|||R'(z)|$, and $E(z)$ is bounded by a linear combination of z and $R(z)$ (shown above), $F(z)$ and $G(z)$ are both bounded on any domain $\operatorname{supp} g \cap [a, b]$, $a, b \in \mathbb{R}_{>0}$. These observations imply $|F(z)G(z)g(z)|$ is bounded on $\operatorname{supp} g$ by a linear combination of $z^2 g(z)$ and $zR(z)g(z)$; so Assumption 6 ensures that for any sequence $(z_n) \subset \operatorname{supp} g$ such that $z_n \to 0$ or ∞, we have $\lim_{n\to\infty} F(z_n)G(z_n)g(z_n) = 0$. Differentiating the expression (2.57) for $E(z)$ reveals that $zE'(z)$ is bounded on $\operatorname{supp} g$ by a linear combination of z and $R(z)$.[109] This fact, the bound on

[108] More explicitly, we have

$$
\begin{aligned}
zB'(z) = {} & \left[2\left(z - zR'(z)\right) - 2z^2R''(z)\right]\left[\frac{d\log R(z)}{d\log z}\eta^2(z) + \left(1 - \frac{d\log R(z)}{d\log z} - \frac{d\log R'(z)}{d\log z}\right)(\eta\varepsilon)(z)\right. \\
& \left. + \frac{d}{d\log z}\left(\frac{d\log R(z)}{d\log z}\right)(\eta^2\varepsilon)(z) + \frac{d}{d\log z}\left(\frac{d\log R'(z)}{d\log z}\right)(\eta\varepsilon^2)(z) + \eta_{+1}(z)\right] \\
& + 2\left(z - zR'(z)\right)\left[\frac{d}{d\log z}\frac{d\log R(z)}{d\log z}\eta^2(z) + \frac{d\log R(z)}{d\log z}z(\eta^2)'(z)\right. \\
& + \left(-\frac{d}{d\log z}\frac{d\log R(z)}{d\log z} - \frac{d}{d\log z}\frac{d\log R'(z)}{d\log z}\right)(\eta\varepsilon)(z) + \left(1 - \frac{d\log R(z)}{d\log z} - \frac{d\log R'(z)}{d\log z}\right)z(\eta\varepsilon)'(z) \\
& + \frac{d^2}{d\log z^2}\left(\frac{d\log R(z)}{d\log z}\right)(\eta^2\varepsilon)(z) + \frac{d}{d\log z}\left(\frac{d\log R(z)}{d\log z}\right)z(\eta^2\varepsilon)'(z) \\
& \left. + \frac{d^2}{d\log z^2}\left(\frac{d\log R'(z)}{d\log z}\right)(\eta\varepsilon^2)(z) + \frac{d}{d\log z}\left(\frac{d\log R'(z)}{d\log z}\right)z(\eta\varepsilon^2)'(z) + z\eta'_{+1}(z)\right] \\
& - 2\left(zR'(z) + z^2R''(z)\right)\left[\eta(z) + \frac{d\log R'(z)}{d\log z}(\eta\varepsilon)(z)\right] - 2zR'(z)\left[z\eta'(z) + \frac{d}{d\log z}\frac{d\log R'(z)}{d\log z}(\eta\varepsilon)(z) + \frac{d\log R'(z)}{d\log z}z(\eta\varepsilon)'(z)\right]
\end{aligned}
\tag{2.184}
$$

By Assumptions 1, 4, and 6, this implies $zB'(z)$ is bounded by a linear combination of z and $R(z)$.

[109] More explicitly, we have

$$
E'(z) = 2(1 - R'(z))(\varepsilon^2)(z) - 2zR''(z)(\varepsilon^2)(z) + 2z(1 - R'(z))(\varepsilon^2)'(z). \tag{2.186}
$$

By Assumptions 1, 4, and 6, this implies $zE'(z)$ is bounded by a linear combination of z and $R(z)$.

$E(z)$, the fact that $|\Delta'(z)| \leq |R'(z)|$ and $|\Delta''(z)z| \leq |R'(z)|$, and Assumptions 1 and 6 imply

$$\begin{aligned}
|F'(z)G(z)g(z)| &= \left|\Delta'(z)\Delta''(z)\frac{zE(z)}{R'(z)^2}g(z)\right| \leq k_1 z g(z) + k_2 R(z) g(z) \\
F(z)\frac{d\,(G(z)g(z))}{dz} &= \left(\frac{zE'(z)}{R'(z)^2} + \frac{E(z)}{R'(z)^2}\frac{d\log}{d\log z}\left(\frac{zg(z)}{R'(z)^2}\right)\right)\frac{\Delta'(z)^2}{2}g(z) \\
&= \left(\frac{zE'(z)}{R'(z)^2} - \frac{E(z)}{R'(z)^2}\left(\alpha(z) + 2\frac{d\log R'(z)}{d\log z}\right)\right)\frac{\Delta'(z)^2}{2}g(z) \\
\left|F(z)\frac{d\,(G(z)g(z))}{dz}\right| &\leq k_3 z g(z) + k_4 R(z) g(z)
\end{aligned} \tag{2.187}$$

for all $z \in \operatorname{supp} g$, for various constants $k_n \in \mathbb{R}_{\geq 0}$. By Assumption 4, these bounds are integrable, as desired.

Fourth, we consider the first term on the RHS of the third line of (2.58) (the first line proportional to $D(z)$). We begin by noting that

$$\int_{\operatorname{supp} g} g(z)D(z)\frac{\Delta(z)}{R(z)}\frac{\Delta''(z)z}{R'(z)}dz = \int_{\operatorname{supp} g} g(z)\frac{D(z)z}{R(z)R'(z)}\underbrace{\left[\Delta(z)\Delta''(z) + \Delta'(z)^2\right]}_{\frac{d}{dz}[\Delta(z)\Delta'(z)]}dz - \int_{\operatorname{supp} g} g(z)D(z)\frac{R'(z)z}{R(z)}\left(\frac{\Delta'(z)}{R'(z)}\right)^2 dz \tag{2.188}$$

where the integrals on the RHS by Assumption 1, the definition of $\boldsymbol{\Delta}$, and the fact shown above that $D(z)$ is bounded on $\operatorname{supp} g$ by a linear combination of z and $R(z)$. We set aside the second term and integrate the first by parts, setting $F(z) = \Delta(z)\Delta'(z)$ and $G(z) = \frac{zD(z)}{R(z)R'(z)}$. Note that $F(z)$ and $G(z)$ are continuously differentiable on $\operatorname{supp} z$ by Assumption 1, $D(z)$'s continuous differentiability (see the proof of Lemma 6), the definition of $\boldsymbol{\Delta}$, and the fact that $R(z)$ and $R'(z)$ are strictly positive on $\mathbb{R}_{>0}$ (see the proof of Lemma 2). Since $R(z)$ and $R'(z)$ are continuous in z (Assumption 1), $R(z), R'(z) > 0$ on $\mathbb{R}_{>0}$ (proof of Lemma 2), $|\Delta(z)| \leq ||\Delta|||R(z)|$ and $|\Delta'(z)| \leq ||\Delta|||R'(z)|$, and $D(z)$ is bounded by a linear combination of z and $R(z)$, $F(z)$ and $G(z)$ are both bounded on any domain $\operatorname{supp} g \cap [a,b]$, $a, b \in \mathbb{R}_{>0}$. These bounds also ensure $|F(z)G(z)g(z)|$ is bounded on $\operatorname{supp} g$ by a linear combination of $z^2 g(z)$ and $zR(z)g(z)$; so by Assumption 6, $\lim_{n\to\infty} F(z_n)G(z_n)g(z_n) = 0$ for any sequence $(z_n) \subset \operatorname{supp} g$ such that $z_n \to 0$ or ∞. Differentiating the expression (2.57) for $D(z)$ reveals that $zD'(z)$ and $z^2D''(z)$ are bounded on $\operatorname{supp} g$ by a linear combination of z and $R(z)$.[110] This fact; the fact that $|\Delta(z)| \leq |R(z)|$, $|\Delta'(z)| \leq |R'(z)|$, and $|\Delta''(z)z| \leq |R'(z)|$; and Assumptions 1 and 6 imply

$$\begin{aligned}
|F'(z)G(z)g(z)| &= \left|\left(\Delta(z)\Delta''(z) + \Delta'(z)^2\right)z\frac{D(z)}{R(z)R'(z)}g(z)\right| \\
&\leq \left|\left(\frac{\Delta''(z)z}{R'(z)} + \frac{\Delta'(z)}{R'(z)}\frac{R'(z)z}{R(z)}\right)D(z)g(z)\right| \leq k_1 z g(z) + k_2 R(z) g(z) \\
F(z)\frac{d\,(G(z)g(z))}{dz} &= \left(\frac{zD'(z)}{R(z)R'(z)} + \frac{D(z)}{R(z)R'(z)}\frac{d\log}{d\log z}\left(\frac{zg(z)}{R(z)R'(z)}\right)\right)\Delta(z)\Delta'(z)g(z) \\
&= \left(\frac{zD'(z)}{R(z)R'(z)} - \frac{D(z)}{R(z)R'(z)}\left(\alpha(z) + \frac{d\log R(z)}{d\log z} + \frac{d\log R'(z)}{d\log z}\right)\right)\Delta(z)\Delta'(z)g(z) \\
\left|F(z)\frac{d\,(G(z)g(z))}{dz}\right| &\leq k_8 z g(z) + k_9 R(z) g(z)
\end{aligned} \tag{2.190}$$

for all $z \in \operatorname{supp} g$, for various constants $k_n \in \mathbb{R}_{\geq 0}$. By Assumption 4, these bounds are integrable, as desired.

Fifth, we consider the first term on the RHS of the last line of (2.58). Take $F(z) = \frac{\Delta(z)^2}{2}$ and $G(z) =$

[110] More explicitly, we have

$$\begin{aligned}
D'(z) &= 2(1 - R'(z))(\eta\varepsilon)(z) - 2zR''(z)(\eta\varepsilon)(z) + 2z(1 - R'(z))(\eta\varepsilon)'(z) \\
D''(z) &= -2R''(z)(\eta\varepsilon)(z) + 2(1 - R'(z))(\eta\varepsilon)'(z) \\
&\quad - 2R''(z)(\eta\varepsilon)(z) - 2zR'''(z)(\eta\varepsilon)(z) - 2zR''(z)(\eta\varepsilon)'(z) \\
&\quad + 2(1 - R'(z))(\eta\varepsilon)'(z) - 2zR''(z)(\eta\varepsilon)'(z) + 2z(1 - R'(z))(\eta\varepsilon)''(z)
\end{aligned} \tag{2.189}$$

By Assumption 1 (and the observation that $z^2R'''(z) = \frac{d^2R'(z)}{d\log z^2} - zR''(z)$), Assumption 4, and Assumption 6, this implies $zD'(z)$ and $z^2D''(z)$ are bounded by a linear combinations of z and $R(z)$.

$\frac{1}{g(z)}\frac{d}{dz}\left[\frac{g(z)zD(z)}{R(z)R'(z)}\right]$. Before beginning to verify the conditions of Lemma 8, it is helpful to note that

$$G(z) = \frac{1}{R(z)R'(z)}\left[zD'(z) - D(z)\left(\alpha(z) + \frac{d\log R(z)}{d\log z} + \frac{d\log R'(z)}{d\log z}\right)\right]. \tag{2.191}$$

Now, note that $F(z)$ and $G(z)$ are continuously differentiable on $\operatorname{supp} z$ by Assumptions 1 and 5 (for $\alpha(z)$), $D(z)$'s twice-continuous differentiability (see the proof of Lemma 6), the definition of $\boldsymbol{\Delta}$, and the fact that $R(z)$ and $R'(z)$ are strictly positive on $\mathbb{R}_{>0}$ (see the proof of Lemma 2). Since $R(z)$ is twice-continuously differentiable (Assumption 1), $R(z), R'(z) > 0$ on $\mathbb{R}_{>0}$ (proof of Lemma 2), $|\Delta(z)| \leq ||\Delta|||R(z)|$ (definition of $||\cdot||$), and $D(z)$ and $zD'(z)$ are bounded by linear combinations of z and $R(z)$ (shown above), $F(z)$ and $G(z)$ are both bounded on any domain $\operatorname{supp} g \cap [a,b]$, $a, b \in \mathbb{R}_{>0}$. Moreover, because $|\Delta(z)| \leq ||\Delta|||R(z)|$, because $\frac{d\log R(z)}{d\log z}$ and $\frac{d\log R'(z)}{d\log z}$ are bounded by Assumption 1, because $\alpha(z)$ and $\frac{R(z)}{R'(z)z}$ are bounded on $\operatorname{supp} g$ by Assumption 6, and because $D(z)$ and $zD'(z)$ are bounded by linear combinations of z and $R(z)$, $|F(z)G(z)g(z)|$ is bounded by a linear combination of $z^2 g(z)$ and $zR(z)g(z)$ over all $z \in \operatorname{supp} g$. So by Assumption 6, $\lim_{n\to\infty} F(z_n)G(z_n)g(z_n) = 0$ for any sequence $(z_n) \subset \operatorname{supp} g$ such that $z_n \to 0$ or ∞. Finally, by our bounding observations on $D(z)$, $zD'(z)$, and $z^2D''(z)$ (see above); the fact that $|\Delta(z)| \leq ||\Delta|||R(z)|$; and Assumptions 1 and 6, we have

$$\begin{aligned}
|F'(z)G(z)g(z)| &\leq k_1 zg(z) + k_2 R(z)g(z) \qquad \text{(done in previous integration-by-parts argument)} \\
F(z)\frac{d\left(G(z)g(z)\right)}{dz} &= \frac{\Delta(z)^2}{2}\frac{d}{dz}\left[\frac{g(z)}{R(z)R'(z)}\left(zD'(z) - \left(\alpha(z) + \frac{d\log R(z)}{d\log z} + \frac{d\log R'(z)}{d\log z}\right)D(z)\right)\right] \\
&= \frac{\Delta(z)^2}{2z}\frac{d}{d\log z}\left[\frac{g(z)}{R(z)R'(z)}\right]\left(zD'(z) - \left(\alpha(z) + \frac{d\log R(z)}{d\log z} + \frac{d\log R'(z)}{d\log z}\right)D(z)\right) \\
&+ \frac{\Delta(z)^2}{2z}\frac{g(z)}{R(z)R'(z)}\frac{d}{d\log z}\left[zD'(z) - \left(\alpha(z) + \frac{d\log R(z)}{d\log z} + \frac{d\log R'(z)}{d\log z}\right)D(z)\right] \\
&= -\frac{1}{2}\left(\frac{\Delta(z)}{R(z)}\right)^2\frac{R(z)}{R'(z)z}\left(1 + \alpha(z) + \frac{d\log R(z)}{d\log z} + \frac{d\log R'(z)}{d\log z}\right)\left(zD'(z) - \left(\alpha(z) + \frac{d\log R(z)}{d\log z} + \frac{d\log R'(z)}{d\log z}\right)D(z)\right) \\
&\quad + \frac{1}{2}\left(\frac{\Delta(z)}{R(z)}\right)^2\frac{R(z)}{R'(z)z}\left[zD'(z) + z^2D''(z) - \left(\alpha(z) + \frac{d\log R(z)}{d\log z} + \frac{d\log R'(z)}{d\log z}\right)zD'(z)\right. \\
&\qquad\qquad \left. - \left(z\alpha'(z) + \frac{d}{d\log z}\frac{d\log R(z)}{d\log z} + \frac{d}{d\log z}\frac{d\log R'(z)}{d\log z}\right)D(z)\right] \\
\left|F(z)\frac{d\left(G(z)g(z)\right)}{dz}\right| &\leq k_3 zg(z) + k_4 R(z)g(z)
\end{aligned} \tag{2.192}$$

for all $z \in \operatorname{supp} g$, for various constants $k_n \in \mathbb{R}_{\geq 0}$. By Assumption 4, these bounds are integrable, as desired.

Application to proof of Lemma 7

Consider the term proportional to $(\lambda\eta)(z)$ in (2.65). Take $F(z) = \frac{\Delta(z)^2}{2}$ and $G(z) = \frac{z(\lambda\eta)(z)}{R(z)} = \frac{zR(z)(\lambda\eta)(z)}{R(z)^2}$. Note that $F(z)$ and $G(z)$ are continuously differentiable on $\operatorname{supp} z$ by Assumption 1, the fact that $((w^h)_{h\in\mathcal{H}}, G)$ is a regular social objective (see Appendix 2.9.6.6.5), the definition of $\boldsymbol{\Delta}$, and the fact that $R(z)$ is strictly positive on $\mathbb{R}_{>0}$ (see the proof of Lemma 2). Next, note that since $R(z)$ is continuous in z (Assumption 1), $R(z) > 0$ on $\mathbb{R}_{>0}$, $|\Delta(z)| \leq ||\Delta|||R(z)|$ (definition of $||\cdot||$), $R(z)|(\lambda\eta)(z)| \leq MR(z)\lambda(z)$ for some constant M (by Assumption 4), and for some $b_c, b_z > 0$, $R(z)\lambda(z) \leq b_c R(z) + b_z z$ (recalling from Appendix 2.9.6.6.5 that $\lambda(z) = \widehat{\lambda}(z)$, see Appendix 2.9.6.6.1), we have that $F(z)$ and $G(z)$ are both bounded on any domain $\operatorname{supp} g \cap [a,b]$, $a, b \in \mathbb{R}_{>0}$. Next, note that because $|\Delta(z)| \leq |R(z)|$ and $R(z)|(\lambda\eta)(z)| \leq MR(z)\lambda(z)$ for some constant M, $|F(z)G(z)g(z)|$ is bounded over all $z \in \operatorname{supp} g$ by a constant times $zR(z)\lambda(z)g(z)$; so Assumption 6 ensures that $\lim_{n\to\infty} F(z_n)G(z_n)g(z_n) = 0$ for any sequence $(z_n) \subset \operatorname{supp} g$ such that $z_n \to 0$ or ∞. Finally, by Assumptions 1, 4, and 6, and the fact that $|\Delta(z)| \leq ||\Delta|||R(z)|$ and $|\Delta'(z)| \leq ||\Delta|||R'(z)|$, we have

$$\begin{aligned}
|F'(z)G(z)g(z)| &= \left|\Delta(z)\Delta'(z)\frac{R(z)}{R(z)}\frac{R'(z)}{R'(z)}\frac{z(\lambda\eta)(z)}{R(z)}g(z)\right| \leq k_1 R(z)\lambda(z)g(z) \\
F(z)\frac{d\left(G(z)g(z)\right)}{dz} &= \left(\frac{z(\lambda\eta)'(z)}{R(z)} + \frac{(\lambda\eta)(z)}{R(z)}\frac{d\log}{d\log z}\left(\frac{zg(z)}{R(z)}\right)\right)\frac{\Delta(z)^2}{2}g(z) \\
\left|F(z)\frac{d\left(G(z)g(z)\right)}{dz}\right| &\leq k_2 R(z)\lambda(z)g(z)
\end{aligned} \tag{2.193}$$

for all $z \in \operatorname{supp} g$, for various constants $k_n \in \mathbb{R}_{\geq 0}$. Since $R(z)\lambda(z) = R(z)\widehat{\lambda}(z)$ is bounded across $\operatorname{supp} g$ by a linear combination of z and $R(z)$ (recalling from Appendix 2.9.6.6.5 that $\lambda(z) = \widehat{\lambda}(z)$, see Appendix 2.9.6.6.1), these bounds are integrable, as desired.

2.9.6.7 Characterization of $\operatorname{supp} g$

Lemma 9. $\operatorname{supp} g$ *is a countable union of disjoint, open, positive intervals.*

Proof. First, note that for all $z \in \operatorname{supp} g$, there exists by h's continuity (Assumption 5) some $a_z, b_z \in \mathbb{Q}$ such that $z \in (a_z, b_z) \subset \operatorname{supp} g$. Since $h(0) = 0$ by Assumption 3, we may take $a_z, b_z \geq 0$. Since $(a_z, b_z) \in \mathbb{Q}^2$, which is countable, we conclude that

$$\operatorname{supp} g = \bigcup_{n\in\mathcal{B}} I_n, \tag{2.194}$$

for I_n positive, open intervals and $\mathcal{B}$ countable.

Next, define an equivalence relation on $\mathcal{B}$ by

$$n \sim m \quad \Longleftrightarrow \quad \exists i_1, ..., i_k \quad \text{s.t.} \quad \forall j = 1, ..., k-1, \quad I_{i_j} \cap I_{i_{j+1}} \neq \emptyset \tag{2.195}$$

Letting $\mathcal{E}$ be the (countable) set of equivalence classes E of $\mathcal{B}$ under $\sim$, we now claim that each union $S_E \equiv \cup_{n\in E} I_n$ is an open interval. S_E is open because it is the union of open sets. To see that S_E is an interval, it suffices to show it is connected. To see this in turn, suppose not, i.e. $S_E = A \cup B$ with $A, B \neq \emptyset$ and $\overline{A} \cap B = A \cap \overline{B} = \emptyset$. This implies that for each $n \in E$, I_n is contained in either A or B, since it must be contained in $A \cup B$ and if both $A \cap I_n$ and $B \cap I_n$ are non-empty, then I_n is not connected, contradicting that it is an interval. As a consequence, each $I_{n\in E} \subset A$ can only be in the same equivalence class as other $I_{m\in E} \subset A$ (and similarly for B); otherwise, some interval $I_{j\in E}$ on the path between them must contain points in both A and B, violating that $\overline{A} \cap B = A \cap \overline{B} = \emptyset$. Since A and B are both non-empty and $S_E = A \cup B$, we may therefore take $n, m \in E$ with $I_n \subset A$ and $I_m \subset B$, implying $n \not\sim m$; this violates $n, m \in E$, a contradiction.

Finally, note that for any distinct $E, E' \in \mathcal{E}$, S_E and $S_{E'}$ are disjoint, since if they intersect then they contain intervals I_n in the same equivalence class. We conclude that

$$\operatorname{supp} g = \bigcup_{n\in\mathcal{B}} I_n = \bigcup_{E\in\mathcal{E}} \bigcup_{n\in E} I_n = \bigcup_{E\in\mathcal{E}} S_E, \tag{2.196}$$

where we have shown S_E are disjoint, open, positive intervals and $\mathcal{E}$ is countable. ☐

2.9.6.8 Indifference curve lemmas

Below, we prove two useful but tedious lemmas that establish uniform bounds on various ratios of derivatives of household utilities. In particular, these lemmas allow us to bound such ratios along households' indifference curves locally to their consumption-labor profile at the initial equilibrium. These bounds are essential to establishing regularity conditions on the social objective constructed in the proof of sufficiency.

The first lemma below shows that—roughly—given any initial, local tax schedule, there exist other local tax schedules that cause a household to experience any local consumption-labor pair on the same indifference curve as the initial consumption-labor pair.

The second lemma leverages the first to bound ratios of utility derivatives at certain points along indifference curves local to the initial tax schedule.

Lemma 10. *There exists $\delta, \tilde{\delta} > 0$ with $\delta < \tilde{\delta} < \frac{1}{2}$ both small enough that*

- *Assumption 4 and Lemmas 2 and 3 apply at $2\tilde{\delta}$,*[111] *and*
- *for all $h \in \mathcal{H}, \widetilde{R} \in R + B_\delta(\mathbf{0}), z \in \left[\min[z_0^h, z^h(\widetilde{R})], \max[z_0^h, z^h(\widetilde{R})]\right]$, there exists a tax schedule $\widehat{R}^h(\cdot; z, \widetilde{R})$ such that*
  - $\widehat{R}^h(\cdot; z, \widetilde{R}) \in \overline{B}_{\tilde{\delta}}(\mathbf{0})$,
  - $z^h\left(\widehat{R}^h(\cdot; z, \widetilde{R})\right) = z$, *and*
  - $V^h\left(\widehat{R}^h(\cdot; z, \widetilde{R})\right) = V^h\left(\widetilde{R}\right)$.

Proof. We complete the proof in several steps. As a preliminary step, we introduce a convenient tax deviation that will be used throughout. Then we show how to select δ and $\tilde{\delta}$. We proceed to fix any $h \in \mathcal{H}, \widetilde{R} \in R + B_\delta(\mathbf{0})$, and $z \in [\min[z_0^h, z^h(\widetilde{R})], \max[z_0^h, z^h(\widetilde{R})]]$ and then show the existence of an desired tax schedule $\widehat{R}(\cdot; z, \widetilde{R})$ by constructing a sequence of tax changes that converge to it.

Preliminary step: a useful deviation

We begin by defining and studying the properties of a useful tax deviation that we will use later on.

For any $\hat{z} \in \mathbb{R}_{>0}$, define $\widehat{\Delta}(\cdot; \hat{z}) : \mathbb{R}_{\geq 0} \to \mathbb{R}$ by

$$\begin{aligned} &\widehat{\Delta}(z; \hat{z}) \equiv \hat{z} R'(\hat{z}) p(z; \hat{z}) \\ \text{where} \quad &p(z; \hat{z}) \equiv \begin{cases} (\log z - (\log \hat{z} - 1))^6 (\log z - \log \hat{z})(\log z - (\log \hat{z} + 1))^6 & \text{if } z \in \left[e^{\log z - 1}, e^{\log z + 1}\right] \\ 0 & \text{else} \end{cases} \end{aligned} \tag{2.197}$$

One may easily verify the following properties of $p(\cdot; \hat{z}) : \mathbb{R}_{\geq 0} \to \mathbb{R}$:

- $p(\cdot; \hat{z})$ is three-times continuously differentiable.
- For $n = 0, 1, 2, 3$, the n^{th} log-derivative of $p(\cdot; \hat{z})$—i.e. $\frac{d^n}{d \log z^n} p(z; \hat{z})$—achieves a maximum absolute value that is independent of $\hat{z}$. Note this implies the existence of $\hat{z}$-independent upper bounds $\overline{p}^1$, $\overline{p}^2$, and $\overline{p}^3$ on the absolute value of $p'(z; \hat{z}) z$ $\frac{dp'(z; \hat{z})}{d \log z} z$ and $\frac{d^2 p'(z; \hat{z})}{d \log z^2} z$, respectively.
- $p(\hat{z}; \hat{z}) = 0$ and $p'(\hat{z}; \hat{z}) = \frac{1}{\hat{z}}$

We now consider the implications of these facts for $\widehat{\Delta}(\cdot; \hat{z})$. In particular, we claim there exists $\widehat{B} > 0$ such that for all $\hat{z} \in \mathbb{R}_{>0}$,

- $\widehat{\Delta}(\hat{z}; \hat{z}) = 0$
- $\widehat{\Delta}'(\hat{z}; \hat{z}) = R'(\hat{z})$

[111] In particular, this guarantees that labor supply and household utility are well-defined at all tax schedules $\widetilde{R} \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$, which facilitates the rest of the Lemma statement.

- $\widehat{\Delta}(\cdot;\hat{z}) \in \mathbf{\Delta}$ and $||\widehat{\Delta}(\cdot;\hat{z})|| \leq \widehat{B}$

The first two bullets are immediate. As our observations on $p(\cdot;\hat{z})$ imply $\widehat{\Delta}(\cdot;\hat{z})$ is three-times continuously differentiable, it remains—in order to prove that $\widehat{\Delta}(\cdot;\hat{z}) \in \mathbf{\Delta}$—to show that $\exists B \in \mathbb{R}$ s.t. $\forall z \in \mathbb{R}_{\geq 0}$

$$|\widehat{\Delta}(z;\hat{z})| \leq B|R(z)|, \quad |\widehat{\Delta}'(z;\hat{z})| \leq B|R'(z)|, \quad \left|\frac{d\widehat{\Delta}'(z;\hat{z})}{d\log z}\right| \leq B|R'(z)|, \quad \text{and} \quad \left|\frac{d^2\widehat{\Delta}'(z;\hat{z})}{d\log z^2}\right| \leq B|R'(z)| \tag{2.198}$$

Our observations about the derivatives of $p(z;\hat{z})$—along with the fact that $p(z;\hat{z})$=0 outside of $[e^{\log\hat{z}-1}, e^{\log\hat{z}+1}]$—imply that it suffices to show there exists $B \in \mathbb{R}$ such that (a) in the levels case we have for all $\log z \in B_1(\log\hat{z})$, that $\bar{p}^0 \underbrace{\frac{R'(\hat{z})\hat{z}}{R(\hat{z})}}_{\leq B^R} R(\hat{z}) \leq BR(z)$ and (b) for $n = 1,2,3$, we have for all $\log z \in B_1(\log\hat{z})$ that $\bar{p}^n R'(\hat{z})\frac{\hat{z}}{z} \leq BR'(z)$. To see that such a B exists note that by Assumption 1 $\frac{R(\hat{z})}{R(z)}, \frac{R'(\hat{z})}{R'(z)} \in [e^{-B^R}, e^{B^R}]$ for all $\log z \in B_1(\log\hat{z})$; similarly, $|\frac{\hat{z}}{z}| \in e^1$. This implies that we have (2.198) for $B = \widehat{B} \equiv \max[\bar{p}^0 B^R, \bar{p}^1, \bar{p}^2, \bar{p}^3]e^{B^R+1} > 0$. We conclude that for any $\hat{z} \in \mathbb{R}_{>0}$, $\widehat{\Delta}(\cdot;\hat{z}) \in \mathbf{\Delta}$, and $||\widehat{\Delta}(\cdot;\hat{z})|| \leq \widehat{B}$.

Choosing appropriate δ, $\tilde{\delta}$

We set $\tilde{\delta} \in (0, \frac{1}{2})$ small enough that Assumption 4 and Lemmas 2 and 3 apply at $2\tilde{\delta}$. Note this guarantees that labor supply $z^h(\widetilde{R})$ and household utility $V^h(\widetilde{R})$ are well-defined and twice-continuously differentiable at all tax schedules $\widetilde{R} \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$, which we will use throughout.[112]

Next, note that by Lemma 3 (in particular see Footnote 56) there exist $d_z, d_c > 0$ such that for all $h \in \mathcal{H}, \widetilde{R} \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$, and non-zero $\Delta \in \mathbf{\Delta}$,

$$|D_\Delta \log z^h(\widetilde{R})| \leq d_z||\Delta|| \qquad \text{and} \qquad |D_\Delta \log c^h(\widetilde{R})| \leq d_c||\Delta|| \tag{2.199}$$

Note that integrating these bounds imply $\log z^h(\widetilde{R}) \in \overline{B}_{d_z\tilde{\delta}}(\log z_0^h)$ and $\log c^h(\widetilde{R}) \in \overline{B}_{d_c\tilde{\delta}}(\log c_0^h)$; we use these facts later on.

Finally, we take $\delta > 0$ small enough that $\delta + \frac{d_z}{\underline{\varepsilon}/2}\widehat{B}\delta < \tilde{\delta}$, where for all $\underline{\varepsilon} > 0$ is a lower bound on compensated elasticities across all $h \in \mathcal{H}$, $\widetilde{R} \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$; this exists by Assumption 4.

Setup for main claim

Now fix any $h \in \mathcal{H}, \widetilde{R} \in R + B_\delta(\mathbf{0})$, and $z \in \left[\min[z_0^h, z^h(\widetilde{R})], \min[z_0^h, z^h(\widetilde{R})]\right]$. Note that if $z = z^h(\widetilde{R})$, we may set $\widehat{R}^h(\cdot; z, \widetilde{R}) = \widetilde{R}$ and we are done. We will therefore prove the claim in the statement of the Lemma assuming $z^h(\widetilde{R}) < z \leq z_0^h$; the complementary case where $z_0^h \leq z < z^h(\widetilde{R})$ is analogous.

Before proceeding, recall our goal: Having fixed $h, \widetilde{R}$, and $z \in (z^h(\widetilde{R}), z_0^h)$, we wish to show the existence of a tax schedule $\widehat{R}$ such that

$$\widehat{R} \in \overline{B}_{\tilde{\delta}}(\mathbf{0}), \qquad z^h(\widehat{R}) = z, \qquad \text{and} \qquad V^h(\widehat{R}) = V^h(\widetilde{R}). \tag{2.200}$$

The remainder of the proof shows the existence of $\widehat{R}$ in two basic steps. First, we construct a sequence of tax schedules $\left(\widehat{R}_N\right)_{N\in\mathbb{N}}$ meant to satisfy these conditions in the limit as $N \to \mathbb{N}$. Second, we argue the limit actually exists and does in fact satisfy the desired conditions.

Sequence of tax schedules

[112]For labor supply, this is from Lemma 2; for indirect utility, it is from Assumptions 1 and 2, and the twice-continuous differentiability of labor supply.

As a first step toward constructing the desired sequence of tax schedules $(\widehat{R}^N)_{N\in\mathbb{N}}$, define—for every $N \in \mathbb{N}$—$\epsilon_N \equiv \frac{\log z_0^h - \log z^h(\widetilde{R})}{N\frac{\varepsilon/2}{\widehat{B}}}$ and for all $n = 1, ..., N$ and $\epsilon \in [0, \epsilon_N]$, define

$$\widehat{R}_n^N(\cdot;\epsilon) : \mathbb{R}_{\geq 0} \to \mathbb{R}, \qquad \tilde{z} \mapsto \widehat{R}_{n-1}^N(\tilde{z};\epsilon_N) + \frac{\widehat{\Delta}\left(\tilde{z}; z^h(\widehat{R}_{n-1}^N(\cdot;\epsilon_N))\right)}{\left|\left|\widehat{\Delta}\left(\cdot; z^h(\widehat{R}_{n-1}^N(\cdot;\epsilon_N))\right)\right|\right|}\epsilon, \tag{2.201}$$

where $\widehat{R}_0^N(\cdot;\epsilon_N) \equiv \widetilde{R}$. Note that—as can be seen by iterating the definition above and applying the triangle inequality—we have that for all $N \in \mathbb{N}, n = 1, ..., N, \epsilon \in [0, \epsilon_N]$, $\widehat{R}_n^N(\cdot;\epsilon) \in \widetilde{R} + \overline{B}_{n\epsilon_N}(\mathbf{0})$. More strongly, we in fact have, $\widehat{R}_n^N(\cdot;\epsilon) \in \widetilde{R} + \overline{B}_{n\epsilon_N}(\mathbf{0}) \subset R + B_{\tilde{\delta}}(\mathbf{0})$[113] since $\widetilde{R} \in B_\delta(\mathbf{0})$; since by the definition of ϵ_N, $n\epsilon_N \leq N\epsilon_N \leq \frac{d_z\delta}{(\varepsilon/2)/\widehat{B}}$; and since by the definition of δ, $\delta + \frac{d_z}{\varepsilon/2}\widehat{B}\delta < \tilde{\delta}$. Finally, note that by the definition of $\widehat{\Delta}(\cdot;\hat{z})$ —and our observation in the second step of this proof that for all $\check{R} \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$, $\log z^h(\check{R}) \in B_{\tilde{\delta}d_z}(\mathbf{0})$—we always have $\widehat{R}_n^N(\tilde{z};\epsilon) = \widetilde{R}(\tilde{z})$ whenever $\log\tilde{z} \notin [\log z_0^h - d_z\tilde{\delta} - 1, \log z_0^h + d_z\tilde{\delta} + 1]$.

We next establish two properties of the tax schedules $\widehat{R}_n^N(\cdot;\epsilon)$ that hold for all sufficiently large N. Specifically, for large enough N, for all $n = 1, ..., N$, $\epsilon \in [0, \epsilon_N]$,

- $\log z^h(\widehat{R}_n^N(\cdot;\epsilon)) \geq \log z^h(\widehat{R}_{n-1}^N(\cdot;\epsilon_N)) + \frac{\varepsilon/2}{\widehat{B}}\epsilon$
- $\left|V^h(\widehat{R}_n^N(\cdot;\epsilon)) - V^h(\widehat{R}_{n-1}^N(\cdot;\epsilon_N))\right| \leq \widetilde{B}\frac{\epsilon^2}{2}$

for some constant $\widetilde{B}$ independent of n, N, ϵ. Both of these facts are consequences of Taylor's theorem. More concretely, by the two-times continuous differentiability of $z^h(\cdot)$ and $V^h(\cdot)$ within $R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$, we have that—for some $\tilde{\epsilon}, \hat{\epsilon} \in [0, \epsilon_N]$—

$$\begin{aligned}
\log z^h(\widehat{R}_n^N(\cdot;\epsilon)) &= \log z^h(\widehat{R}_n^N(\cdot;0)) + D_{\overline{\Delta}_n}\log z^h(\widehat{R}_n^N(\cdot;0))\epsilon + D_{\overline{\Delta}_n}^2\log z^h(\widehat{R}_n^N(\cdot;\tilde{\epsilon}))\frac{\epsilon^2}{2} \\
&\leq \log z^h(\widehat{R}_{n-1}^N(\cdot;\epsilon_N)) + \frac{\varepsilon}{\widehat{B}}\epsilon - \overline{d}_2\frac{\epsilon^2}{2} \\
V^h(\widehat{R}_n^N(\cdot;\epsilon)) - V^h(\widehat{R}_n^N(\cdot;0)) &= D_{\overline{\Delta}_n}V^h(\widehat{R}_n^N(\cdot;0))\epsilon + D_{\overline{\Delta}_n}^2V^h(\widehat{R}_n^N(\cdot;\hat{\epsilon}))\frac{\epsilon^2}{2} \\
\left|V^h(\widehat{R}_n^N(\cdot;\epsilon)) - V^h(\widehat{R}_{n-1}^N(\cdot;\epsilon_N))\right| &\leq 0 + \left(\overline{u}_{cc}^h\left(c_0^he^{\frac{1}{2}}\right)^2 + \overline{u}_c^h\ c_0^h\ e^{\frac{1}{2}}\ B^R\ \frac{1}{\underline{\varepsilon}}(\overline{d}_1)^2\right)\frac{\epsilon^2}{2}
\end{aligned} \tag{2.202}$$

where $\overline{\Delta}_n = \frac{\widehat{\Delta}\left(\cdot;z^h(\widetilde{R}_{(n)}(\cdot;0))\right)}{\left|\left|\widehat{\Delta}\left(\cdot;z^h(\widetilde{R}_{(n)}(\cdot;0))\right)\right|\right|}$, where $\overline{d}_1$ and $\overline{d}_2$ are upper bounds on the first and second derivatives of log labor supply across all $\widetilde{R} \in B_{2\tilde{\delta}}(\mathbf{0})$ (by Lemma 3; see Footnote 56), and where[114]

$$\begin{aligned}
\overline{u}_c^h &= \max_{c,z} u_c^h(c,z) \quad \text{s.t.} \quad \log c \in \overline{B}_{d_c\tilde{\delta}}(\log c_0^h), \log z \in \overline{B}_{d_z\tilde{\delta}}(\log z_0^h) \\
\overline{u}_{cc}^h &= \max_{c,z} u_{cc}^h(c,z) \quad \text{s.t.} \quad \log c \in \overline{B}_{d_c\tilde{\delta}}(\log c_0^h), \log z \in \overline{B}_{d_z\tilde{\delta}}(\log z_0^h)
\end{aligned} \tag{2.203}$$

The first inequality follows from Assumption 4 and the definitions of $\widehat{\Delta}(\cdot;\tilde{z},\widetilde{R})$ and $\widehat{B}$ (from the first section of this proof). The second inequality follows from the definition of $\widehat{\Delta}(\cdot;\tilde{z},\widetilde{R})$, the expressions (2.141) and (2.142) for the first and second derivatives of welfare in the proof of Lemma 5, Assumptions 1 and 4, the fact that $\tilde{\delta} < \frac{1}{2}$, and the observation (from the second section of this proof) that for all $\check{R} \in \overline{B}_{\tilde{\delta}}(\mathbf{0})$, $\log z^h(\check{R}) \in \overline{B}_{d_z\tilde{\delta}}(\log z_0^h)$ and $\log c^h(\check{R}) \in \overline{B}_{d_c\tilde{\delta}}(\log c_0^h)$. (2.202) implies that claims above (those in bullets)

[113]Note that this justifies our usage of $z^h(\widehat{R}_{n-1}^N(\cdot;\epsilon_N))$ in the definition of $\widehat{R}_n^N(z;\epsilon)$.

[114]The maxima defined below exist by Assumption 2.

hold so long as ϵ is sufficiently small, which—since $\epsilon \in [0, \epsilon_N]$ and $\epsilon_N \to 0$ as $N \to \infty$—holds for all N larger than some $\widehat{N}$.

We are now almost ready to define $\widehat{R}^N$, the N^{th} term of our tax sequence of interest. To do this, note that iterating the first bullet established above implies that, for large enough N, $\log z^h(\widehat{R}_N^N(\cdot;\epsilon_N)) \geq \log z^h(\widetilde{R}) + \frac{\varepsilon/2}{\widehat{B}}\varepsilon_N N = \log z^h(\widetilde{R}) \left(\log z_0^h - \log z^h(\widetilde{R})\right) = \log z_0^h$ by the definition of ϵ_N. Since, for this fixed N, stringing the series of tax schedules $R_n^N(\cdot;\epsilon)$ generates a continuous path of tax schedules in $R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$, since $z^h(\cdot)$ is continuous on this domain, and since $z \in (z^h(\widetilde{R}), z_0^h]$, the intermediate value theorem implies there exists $n^*(N) \leq N$ and $\epsilon^*(N) \in [0, \epsilon_N]$ such that $z^h(\widehat{R}^N_{n^*(N)}(\cdot;\epsilon^*(N))) = z$. We therefore define

$$\widehat{R}^N = \begin{cases} \widehat{R}^N_{n^*(N)}(\cdot;\epsilon^*(N)) & \text{if } N > \widehat{N} \\ \widetilde{R} & \text{else.} \end{cases} \tag{2.204}$$

By construction, $\widehat{R}^N \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$. Moreover note that iterating the second bullet proved above implies $\left|V^h(\widehat{R}^N) - V^h(\widetilde{R})\right| \leq \widetilde{B}\frac{\epsilon_N N}{2}\epsilon_N \to 0$ as $N \to \infty$.

Taking stock

Let us take stock. So far, we have

- defined some $\delta, \tilde{\delta} > 0$,
- fixed arbitrary $h \in \mathcal{H}, \widetilde{R} \in R + B_\delta(\mathbf{0})$,
- assumed WLOG that $z^h(\widetilde{R}) < z_0^h$ and fixed $z \in \left(z^h(\widetilde{R}), z_0^h\right]$, and
- shown the existence of a sequence of tax schedules $\left(\widehat{R}^N\right)_{N\in\mathbb{N}}$ such that
  - $\widehat{R}^N \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$,
  - $z^h(\widehat{R}^N) = z$,
  - $V^h(\widehat{R}^N) \to V^h(\widetilde{R})$ as $N \to \infty$, and
  - $\widehat{R}^N(\tilde{z}) = \widetilde{R}(\tilde{z})$ whenever $\log \tilde{z} \notin B_{\tilde{\delta}+1}(\log z_0^h)$.

Recall our goal is to show the existence of a tax schedule $\widehat{R}$ such that

$$\widehat{R} \in \overline{B}_{\tilde{\delta}}(\mathbf{0}), \qquad z^h(\widehat{R}) = z, \qquad \text{and} \qquad V^h(\widehat{R}) = V^h(\widetilde{R}). \tag{2.205}$$

In order to do this, the final step of our proof will argue that the sequence $\left(\widehat{R}^N\right)_{N\in\mathbb{N}}$ has a subsequence $\left(\widehat{R}^{N_k}\right)_{k\in\mathbb{N}}$ that is Cauchy with respect to the metric $||\cdot||$. Because $\mathbf{\Delta}$ is complete (see Lemma 1) R^{N_k} converges, and by our earlier observations, converges to some $\widehat{R} \in \overline{B}_{\tilde{\delta}}(\mathbf{0})$. Finally, the continuity of $z^h(\cdot)$ implies that $z^h(\widehat{R}) = z$ and the continuity of $V^h(\cdot)$ implies $V^h(\widehat{R}) = V^h(\widetilde{R})$, completing the proof.

Existence of a Cauchy subsequence

Finally, we argue that $\left(\widehat{R}^N\right)_{N\in\mathbb{N}}$ has a subsequence which is Cauchy (in the metric $||\cdot||$). We will show this by first arguing a subsequence has uniformly convergent third derivatives, and then argue this implies the sequence is Cauchy.

To the point on third derivatives, recall that for all $N \in \mathbb{N}$, $\widehat{R}^N(\tilde{z}) = \widetilde{R}(\tilde{z})$—and in particular $\widehat{R}^{N\prime\prime\prime}(\tilde{z}) = \widetilde{R}'''(\tilde{z})$—for all $\tilde{z} \notin \left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]$. So to show that some subsequence has uniformly convergent third derivatives it suffices to do so only at $\tilde{z} \in \left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]$. On this domain (a finite interval), this is implied by the Arzelà-Ascoli theorem[115] applied to the sequence $\left(\widehat{R}^{N\prime\prime\prime} - \widetilde{R}'''\right)_{N\in\mathbb{N}}$ because

- $\left|\widehat{R}^{N\prime\prime\prime}(\tilde{z}) - \widetilde{R}'''(\tilde{z})\right|$ is **uniformly bounded** across $N \in \mathbb{N}$, $\tilde{z} \in \left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]$. To see this, note that since $\widetilde{R}, \widehat{R}^N \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$, the definition of $||\cdot||$ implies $\frac{(\widehat{R}^{N\prime\prime\prime}(\tilde{z}) - \widetilde{R}'''(\tilde{z}))\tilde{z}^2}{R'(\tilde{z})} \in \overline{B}_{4\tilde{\delta}}(\mathbf{0})$.[116] Because (see the proof of Lemma 2) $R'(\tilde{z}) > 0$ for all $\tilde{z} \in \mathbb{R}_{>0}$ and since $R'(\tilde{z})$ is continuous by Assumption 2, it and $\tilde{z}$ achieve strictly positive upper and lower bounds $\overline{R}'$ and $\underline{z}$ on $\left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]$. We conclude that $\left|\widehat{R}^{N\prime\prime\prime}(\tilde{z}) - \widetilde{R}'''(\tilde{z})\right| \leq 4\tilde{\delta}\overline{R}'/\underline{z}$.

- The sequence $\left(\widehat{R}^{N\prime\prime\prime}(\cdot) - \widetilde{R}'''(\cdot)\right)_{N\in\mathbb{N}}$ is **equicontinuous** on $\left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]$. To see this, first note that since the function $\widehat{\Delta}(\cdot;\hat{z})$ (see the first step of the proof) is four-times continuously differentiable, the construction of $\widehat{R}^{N\prime\prime\prime}(\cdot)$ implies that $\widehat{R}^{N\prime\prime\prime}(\cdot) - \widetilde{R}'''(\cdot)$ is continuously differentiable. To show equicontinuity, it suffices to show this derivative is uniformly bounded across $N \in \mathbb{N}$ and $\tilde{z} \in \left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]$. To see this, note in turn that—again, by the construction of $\widehat{R}^N$—

$$\frac{d}{d\tilde{z}}\left|\widehat{R}^{N\prime\prime\prime}(\tilde{z}) - \widetilde{R}'''(\tilde{z})\right| \leq N\epsilon_N \max_{\substack{\hat{z} \in \left[e^{\log z_0^h - \tilde{\delta}d_z}, e^{\log z_0^h + \tilde{\delta}d_z}\right] \\ \check{z} \in \left[e^{\log z_0^h - \tilde{\delta}d_z - 1}, e^{\log z_0^h + \tilde{\delta}d_z + 1}\right]}} \frac{d^4}{d\check{z}^4}\widehat{\Delta}(\check{z};\hat{z}) \tag{2.206}$$

  That the supremum above exists is evident from the definition of $\widehat{\Delta}(\check{z};\hat{z})$ in (2.197). Finally, note that the RHS of the equation above is bounded since $N\epsilon_N = \frac{\log z_0^h - \log z^h(\widetilde{R})}{\frac{\varepsilon/2}{\widehat{B}}}$ which is independent of N and z.

So far, we have shown the existence of a subsequence N_k such that $\widehat{R}^{N_k\prime\prime\prime}$ converges to some function—call it $\check{R}_3 : \mathbb{R}_{\geq 0} \to \mathbb{R}$—uniformly in z. Note that (as each $\widehat{R}^{N_k}$ is three-times continuously differentiable) this implies $\check{R}_3$ is continuous and that $\check{R}^3(\tilde{z}) = \widetilde{R}'''(\tilde{z})$ for $\tilde{z} \notin \left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]$. Now, define $\check{R}^2 : \mathbb{R}_{\geq 0} \to \mathbb{R}$ by

$$\check{R}^2(\tilde{z}) = \widetilde{R}''(0) + \int_0^{\tilde{z}} \check{R}^3(\hat{z})\hat{z}, \tag{2.207}$$

which note is well-defined since $R^3(\hat{z})$ is continuous. Note that $R^3(\tilde{z}) = R^{2\prime}(\tilde{z})$. Noting that for each N_k,

$$\widehat{R}^{N_k\prime\prime}(\tilde{z}) = \widetilde{R}''(0) + \int_0^{\tilde{z}} \widehat{R}^{N_k\prime\prime\prime}(\hat{z})\hat{z}, \tag{2.208}$$

we claim that as $k \to \infty$, $R^{N_k\prime\prime} \to \check{R}^2$ uniformly in $\tilde{z}$. To see this fix any $\epsilon > 0$. Since for sufficiently high k, $|\widehat{R}^{N_k\prime\prime\prime}(\tilde{z}) - \check{R}^3(\tilde{z})| \leq \frac{\epsilon}{e^{\log z_0^h + d_z\tilde{\delta}+1} - e^{\log z_0^h - d_z\tilde{\delta}-1}}$ for all $\tilde{z} \notin \left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]$ and $= 0$ otherwise, we

[115] The Arzelà-Ascoli theorem is a standard result in functional analysis. It provides conditions under which a sequence of functions has a uniformly convergent subsequence.

[116] The definition of $||\cdot||$ gives us $\frac{d(\widehat{R}^N - R)'}{d\log z} = z(\widehat{R}^N - R)''(z) \leq \overline{\delta}R'(z)$ and $\frac{d^2(\widehat{R}^N - R)'}{d\log z^2} = z(\widehat{R}^N - R)''(z) + z^2(\widehat{R}^N - R)'''(z) \leq \overline{\delta}R'(z)$, which imply $\frac{z^2(\widehat{R}^N - R)'''(z)}{R'(z)} \leq 2\overline{\delta}$ (here we have used that $R'(z) > 0$, see the proof of Lemma 2). Combining this with the same observation for $\widetilde{R}$ gives us the desired conclusion.

have

$$\left|\check{R}^2(\tilde{z}) - \widehat{R}^{N_k}{}''(\tilde{z})\right| \leq \int_0^{\tilde{z}} \left|\check{R}^3(\hat{z}) - \widehat{R}^{N_k}{}'''(\hat{z})\right| \hat{z} \leq \frac{e^{\log z_0^h + d_z\tilde{\delta}+1} - e^{\log z_0^h - d_z\tilde{\delta}-1}}{e^{\log z_0^h + d_z\tilde{\delta}+1} - e^{\log z_0^h - d_z\tilde{\delta}-1}}\epsilon = \epsilon. \tag{2.209}$$

Repeating this argument again for the first and zeroth derivatives of the subsequence $\widehat{R}^{N_k}$, we have shown the existence of a function $\check{R}^0 : \mathbb{R}_{\geq 0} \to \mathbb{R}$ such that for $n = 0, 1, 2, 3$, $\frac{d^n}{dz^n}\widehat{R}^{N_k} \to \frac{d^n}{dz^n}\check{R}^0$ uniformly as $k \to \infty$.

Finally, we wish to conclude that $\widehat{R}^{N_k}$ is Cauchy with respect to the metric $||\cdot||$. To this end, fix $\epsilon > 0$. Take k large enough that for all $k' > k$, $|\widehat{R}^{N_{k'}} - \check{R}^0|_\infty \leq \overline{R}\frac{\epsilon}{2}$, $|(\widehat{R}^{N_{k'}})' - (\check{R}^0)'|_\infty \leq \overline{R}'\frac{\epsilon}{2}$, $|(\widehat{R}^{N_{k'}})'' - (\check{R}^0)''|_\infty \leq \frac{\overline{R}'}{e^{\log z_0^h + d_z\tilde{\delta}+1}}\frac{\epsilon}{4}$, and $|(\widehat{R}^{N_{k'}})''' - (\check{R}^0)'''|_\infty \leq \frac{\overline{R}'}{\left(e^{\log z_0^h + d_z\tilde{\delta}+1}\right)^2}\frac{\epsilon}{4}$,[117] where here we have defined $\overline{R} = \underset{\tilde{z}\in\left[e^{log\ z_0^h - d_z\tilde{\delta}-1}, e^{log\ z_0^h + d_z\tilde{\delta}+1}\right]}{max} R(\tilde{z}) > 0$ and $\overline{R}' = \underset{\tilde{z}\in\left[e^{log\ z_0^h - d_z\tilde{\delta}-1}, e^{log\ z_0^h + d_z\tilde{\delta}+1}\right]}{max} R'(\tilde{z}) > 0$; both exist and are strictly positive since $R(\tilde{z}), R'(\tilde{z})$ are continuous and strictly positive for $\tilde{z} \in \mathbb{R}_{>0}$ by Assumption 2 and the proof of Lemma 2. To show that for all $k', k'' > k$, $||\widehat{R}^{N_{k'}} - R^{N_{k''}}|| < \epsilon$, we will (by the triangle inequality) show $||\widehat{R}^{N_{k'}} - \check{R}^0|| < \epsilon/2$.[118] Indeed, since $\widehat{R}^{(N_{k'})}(\tilde{z}) = \check{R}^0(\tilde{z}) = \widetilde{R}(\tilde{z})$ for all $z \notin \left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]$

$$\begin{aligned}
\left|\left|\widehat{R}^{N_{k'}} - \check{R}^0\right|\right| \leq & \sup_{\tilde{z}\in\left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]} \max\Bigg(\frac{|\widehat{R}^{N_{k'}}(\tilde{z}) - \check{R}^0(\tilde{z})|}{|R(\tilde{z})|}, \frac{|(\widehat{R}^{N_{k'}})'(\tilde{z}) - (\check{R}^0)'(\tilde{z})|}{|R'(\tilde{z})|}, \\
& \frac{|(\widehat{R}^{N_{k'}})''(\tilde{z}) - (\check{R}^0)''(\tilde{z})|\tilde{z}}{|R'(\tilde{z})|}, \\
& \frac{|(\widehat{R}^{N_{k'}})'''(\tilde{z}) - (\check{R}^0)'''(\tilde{z})|\tilde{z}^2}{|R'(\tilde{z})|} + \frac{|(\widehat{R}^{N_{k'}})''(\tilde{z}) - (\check{R}^0)''(\tilde{z})|\tilde{z}}{|R'(\tilde{z})|}\Bigg) \\
= & \sup_{\tilde{z}\in\left[e^{\log z_0^h - d_z\tilde{\delta}-1}, e^{\log z_0^h + d_z\tilde{\delta}+1}\right]} \max\Bigg(\frac{|\widehat{R}^{N_{k'}}(\tilde{z}) - \check{R}^0(\tilde{z})|}{\overline{R}}, \frac{|(\widehat{R}^{N_{k'}})'(\tilde{z}) - (\check{R}^0)'(\tilde{z})|}{\overline{R}'}, \\
& \frac{|(\widehat{R}^{N_{k'}})''(\tilde{z}) - (\check{R}^0)''(\tilde{z})|e^{\log z_0^h + d_z\tilde{\delta}+1}}{\overline{R}'}, \\
& \frac{|(\widehat{R}^{N_{k'}})'''(\tilde{z}) - (\check{R}^0)'''(\tilde{z})|\left(e^{\log z_0^h + d_z\tilde{\delta}+1}\right)^2}{\overline{R}'} + \frac{|(\widehat{R}^{N_{k'}})''(\tilde{z}) - (\check{R}^0)''(\tilde{z})|e^{\log z_0^h + d_z\tilde{\delta}+1}}{\overline{R}'},\Bigg) \\
\leq & \epsilon/2.
\end{aligned} \tag{2.210}$$

□

Lemma 11. *There exists $\delta > 0$ small enough that the function*

$$\hat{c}^h(u) \equiv u^h(\cdot, z_0^h)^{-1}(u) \tag{2.211}$$

is, for all $h \in \mathcal{H}$, well-defined and strictly positive when $u = V^h(\widetilde{R})$ for some $\widetilde{R} \in R + B_\delta(\mathbf{0})$; moreover, $\hat{c}^h(V^h(\widetilde{R}))$ is $\mathcal{H}$-measurable. Further, there exists $\bar{m} > 0$ such that for all $h \in \mathcal{H}, \widetilde{R} \in R + B_\delta(\mathbf{0})$ and—for all real-valued functions ϕ^h that are defined and twice differentiable in a neighborhood around $V^h(\widetilde{R})$ and

[117] $|\cdot|_\infty$ denotes the sup-norm.

[118] Although it is a slight abuse of notation to apply $||\cdot||$ to $\widehat{R}^{N_{k'}} - \check{R}^0$ without having shown that $\check{R}^0 \in R + \Delta$, it is easy to see that the argument does not depend on this.

satisfy $\phi^{h\prime}(V^h(\widetilde{R})) > 0$*—we have*[119]

$$\left|\log \hat{c}^h\left(V^h(\widetilde{R})\right) - \log c^h\left(\widetilde{R}\right)\right| \leq \bar{m}$$

$$\left|\log\left[(\phi^h \circ u^h)_c\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right)\right] - \log\left[(\phi^h \circ u^h)_c\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)\right]\right| \leq \bar{m} \tag{2.212}$$

$$\text{and} \qquad \left|\frac{d\log}{d\log c}(\phi^h \circ u^h)_c\left(\hat{c}^h(V^h(\widetilde{R})), z_0^h\right) - \frac{d\log}{d\log c}(\phi^h \circ u^h)_c\left(c^h(\widetilde{R}), z^h(\widetilde{R})\right)\right| \leq \bar{m}.$$

Proof. We complete the proof in three steps. First, we situate the claim in the statement of the Lemma in the context of Lemma 10 and establish the existence of the function $\hat{c}^h(u)$ and the measurability claim; along the way we establish the bound on the levels of log consumption. Second, we prove the bound concerning the first derivatives of $\phi^h \circ u^h$, and third, we prove the bound concerning the second derivatives of $\phi^h \circ u^h$.

Note that it suffices to establish the three bounds in (2.212) for three distinct bounds $\bar{m}_1$, $\bar{m}_2$, and $\bar{m}_3$, as we may subsequently take their minimum; we may therefore prove each bound in isolation.

Indifference curve path

To begin, take δ and $\tilde{\delta} > \delta$ as in Lemma 10; recall that Assumption 4 and Lemmas 2 and 3 apply $2\tilde{\delta}$, so that household labor supply, consumption, and indirect utility are defined at all tax schedules $\widetilde{R} \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$. Recall moreover that (from the Lemma) for all $\widetilde{R} \in R + B_\delta(\mathbf{0}), h \in \mathcal{H}, z \in [\min[z_0^h, z^h(\widetilde{R})], \max[z_0^h, z^h(\widetilde{R})]]$, there exists $\widehat{R}^h(\cdot; z, \widetilde{R}) \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$ such that $z^h(\widehat{R}^h(\cdot; z, \widetilde{R})) = z$ and

$$V^h(\widehat{R}^h(\cdot; z, \widetilde{R})) = u^h\Big(\underbrace{c^h(\widehat{R}^h(\cdot; z, \widetilde{R})))}_{=\widehat{R}^h(z; z, \widetilde{R}))}, z\Big) = V^h(\widetilde{R}). \tag{2.213}$$

We define $\check{c}^h(z, \widetilde{R}) \equiv c^h(R^h(\cdot; z, \widetilde{R})))$. In particular, note that Note that $\check{c}^h(z, \widetilde{R}) = u^h(\cdot, z)^{-1}(V^h(\widetilde{R}))$ by (2.213)—implying that $u^h(\cdot, z)^{-1}(V^h(\widetilde{R}))$ exists—and that $\check{c}^h(z, \widetilde{R}) > 0$ since it is contained in $\left[c_0^h e^{-d_c\tilde{\delta}}, c_0^h e^{d_c\tilde{\delta}}\right] > 0$, where d_c is an upper bound on the first derivative of log labor supply across households and local tax schedules (by Lemma 3; see Footnote 56). Lastly, note that—by the implicit function theorem and since by Assumption 2, $u_c^h(\check{c}^h(z, \widetilde{R}), z) > 0$ and u^h is continuously differentiable—$\check{c}^h(z; \widetilde{R})$ is continuously differentiable in $z \in [\min[z_0^h, z^h(\widetilde{R})], \max[z_0^h, z^h(\widetilde{R})]]$. Totally differentiating $u^h(\check{c}^h(z, \widetilde{R}), z) = V^h(\widetilde{R})$ implies $c_z^h(z, \widetilde{R}) = -\frac{u_z^h(\check{c}^h(z, \widetilde{R}), z)}{u_c^h(\check{c}^h(z, \widetilde{R}), z)}$.

One specific implication of these observations is that the function $\hat{c}^h(u)$ referred to in (2.211) of the Lemma statement exists (take $z = z_0^h$), satisfies

$$\log \hat{c}^h(V^h(\widetilde{R})) \in B_{d_c\tilde{\delta}}(\log c_0^h) \tag{2.214}$$

(which note establishes the desired bound in the Lemma statement) and so is strictly positive, and whenever $u = V^h(\widetilde{R} \in R + B_\delta(\mathbf{0}))$, is equal to $\check{c}^h(z_0^h, \widetilde{R})$. Note also that (by Assumption 2), $c^h(z^h(\widetilde{R}), \widetilde{R}) = c^h(\widetilde{R})$.

To see that—for any $\widetilde{R} \in R + B_\delta(\mathbf{0})$—$\hat{c}^h(\widetilde{R})$ is measurable in $h \in \mathcal{H}$, we apply the measurable maximum theorem as stated in Aliprantis and Border (2006).[120] Specifically, define $\Gamma : \mathcal{H} \rightrightarrows \mathbb{R}_{>0}$ by $\Gamma(h) =$

[119]The fact that $\hat{c}^h(V^h(\widetilde{R})) > 0$ implies that u^h twice differentiable and has strictly positive first consumption derivative at all inputs where evaluated above, by Assumption 2. ϕ^h is twice differentiable by assumption. Finally, since $\phi^{h\prime}(V^h(\widetilde{R})) > 0$ by assumption. Together, these observations imply all derivatives and logs used in the Lemma statement are well-defined.

[120]We specialize the theorem to our setting as in the "Measurability of labor supply" step of the proof of Lemma 2. Specifically, we use the following result: If $\Gamma : \mathcal{H} \rightrightarrows \mathbb{R}_{>0}$ is a weakly measurable correspondence with non-empty compact values and

$\left[c_0^h e^{-d_c\tilde{\delta}}, c_0^h e^{d_c\tilde{\delta}}\right]$; this is a non-empty- and compact-valued correspondence by construction. The fact that Γ is weakly measurable, follows as a special case of the argument made for the correspondence used in the 'Measurability of labor supply" step of the proof of Lemma 2. Next define $f : \mathbb{R}_{>0} \times \mathcal{H}$ by $(c,h) \mapsto -\left(u^h(c, z_0^h) - V^h(\widetilde{R})\right)^2$. f is continuous in c by Assumption 2 and measurable in h because z_0^h is by Assumption 3, because $u^h(\cdot, z_0^h)$ therefore is by Assumption 2 and the composition of measurable functions, and because $V^h(\widetilde{R}) = u^h(c^h(\widetilde{R}), z^h(\widetilde{R}))$ is by Assumptions 2 and Lemma 2 and the composition of measurable functions. By the measurable maximum theorem, the argmax $\arg\max_{c\in\Gamma(h)} f(c,h)$ has a measurable selector. However note that $f(c,h)$ is uniquely maximized by $\hat{c}^h(V^h(\widetilde{R}))$, since $u^h(\hat{c}^h(V^h(\widetilde{R})), z_0^h) = V^h(\widetilde{R})$ and u^h is strictly increasing in consumption. So $\hat{c}^h(\widetilde{V}^h(\widetilde{R}))$ is measurable in h.

First derivative bounds

We now consider the first bound in (2.212). To start, note that for any $h \in \mathcal{H}, \widetilde{R} \in R + B_\delta(\mathbf{0}), z \in [\min[z_0^h, z^h(\widetilde{R})], \max[z_0^h, z^h(\widetilde{R})]]$, we have

$$\begin{aligned}
\frac{d}{dz}\log\left(\phi^h \circ u^h\right)_c\left(\check{c}^h(z,\widetilde{R}),z\right) &= \frac{(\phi^h \circ u^h)_{cc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h \circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}\frac{\check{c}^{h\prime}(z;\widetilde{R})}{\check{c}^h(z,\widetilde{R})} + \frac{(\phi^h \circ u^h)_{cz}\left(\check{c}^h(z,\widetilde{R}),z\right)}{(\phi^h \circ u^h)\left(\check{c}^h(z,\widetilde{R}),z\right)} \\
&= \left(-\frac{d\log M^h(\check{c}^h(z,\widetilde{R}),z)}{d\log c} + \frac{(\phi^h \circ u^h)_{zc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h \circ u^h)_z\left(\check{c}^h(z,\widetilde{R}),z\right)}\right)\frac{\check{c}^{h\prime}(z;\widetilde{R})}{\check{c}^h(z,\widetilde{R})} + \frac{(\phi^h \circ u^h)_{cz}\left(\check{c}^h(z,\widetilde{R}),z\right)}{(\phi^h \circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)} \\
&= \frac{\eta^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}\frac{\check{c}^{h\prime}(z;\widetilde{R})}{\check{c}^h(z,\widetilde{R})} + (\phi^h \circ u^h)_{zc}\left(\check{c}^h(z,\widetilde{R}),z\right)\left(\frac{\check{c}^{h\prime}(z;\widetilde{R})}{(\phi^h \circ u^h)_z\left(\check{c}^h(z,\widetilde{R}),z\right)} + \frac{1}{(\phi^h \circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}\right) \\
\left|\frac{d}{d\log z}\log\left(\phi^h \circ u^h\right)_c\left(\check{c}^h(z,\widetilde{R}),z\right)\right| &\leq \frac{\overline{\eta}}{\underline{\varepsilon}}\left|\frac{d\log\check{c}^h(z,\widetilde{R})}{d\log z}\right|
\end{aligned} \tag{2.215}$$

where we may take logs since $\phi^{h\prime} > 0$ by the statement of the Lemma and $u_c^h > 0$ by Assumption 2, where we have used that $M^h(c,z) = -\frac{u_z^h(c,z)}{u_c^h(c,z)} = -\frac{(\phi^h \circ u^h)_z(c,z)}{(\phi^h \circ u^h)_c(c,z)}$, where we have used the definition of η^h and ε^h in (2.125), where $\overline{\eta}$ and $\underline{\varepsilon} > 0$ are upper and lower bounds on the magnitude of elasticities—per Assumption 4 and the fact that all $\widehat{R}(\cdot;z,\widetilde{R}) \in R + \overline{B}_{\tilde{\delta}}(\mathbf{0})$—, and where the cancellation is since, by the design of the path,

$$\frac{d}{dz}(\phi^h \circ u^h)\left(\check{c}^h(z,\widetilde{R}),z\right) = (\phi^h \circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^{h\prime}(z;\widetilde{R}) + (\phi^h \circ u^h)_z\left(\check{c}^h(z,\widetilde{R}),z\right) = 0. \tag{2.216}$$

Next—for any $h \in \mathcal{H}, \widetilde{R} \in R + B_\delta(\mathbf{0})$—the continuous differentiability of $c^h(\cdot, \widetilde{R})$ and $\phi^h \circ u^h$ (by the

$f : \mathbb{R}_{>0} \times \mathcal{H} \to \mathbb{R}$ is a Catheodory function (continuous in its first argument and measurable in its second), then the arg max function $\mu(h) \equiv \arg\max_{z\in\Gamma(h)} f(z,h)$ admits a measurable selector.

conditions of the Lemma and Assumption 2) allow us to apply the fundamental theorem of calculus:

$$\begin{aligned}
&\left|\log\left[(\phi^h\circ u^h)_c\left(\hat{c}^h(V^h(\widetilde{R})),z_0^h\right)\right]-\log\left[(\phi^h\circ u^h)_c\left(c^h(\widetilde{R}),z^h(\widetilde{R})\right)\right]\right|\\
&=\left|\log\left[(\phi^h\circ u^h)_c\left(\check{c}^h(z_0^h,\widetilde{R}),z_0^h\right)\right]-\log\left[(\phi^h\circ u^h)_c\left(\check{c}^h(z^h(\widetilde{R}),\widetilde{R}),z^h(\widetilde{R})\right)\right]\right|\\
&=\left|\int\limits_{\min\left[\log z_0^h,\log z^h(\widetilde{R})\right]}^{\max\left[\log z_0^h,\log z^h(\widetilde{R})\right]}\frac{d}{d\log z}\log\left[(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)\right]d\log z\right|\\
&=\int\limits_{\min\left[\log z_0^h,\log z^h(\widetilde{R})\right]}^{\max\left[\log z_0^h,\log z^h(\widetilde{R})\right]}\left|\frac{d}{d\log z}\log\left[(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)\right]\right|d\log z \qquad (2.217)\\
&\leq\frac{\overline{\eta}}{\underline{\varepsilon}}\int\limits_{\min\left[\log z_0^h,\log z^h(\widetilde{R})\right]}^{\max\left[\log z_0^h,\log z^h(\widetilde{R})\right]}\left|\frac{d\log\check{c}^h(z,\widetilde{R})}{d\log z}\right|d\log z=\frac{\overline{\eta}}{\underline{\varepsilon}}\int\limits_{\min\left[\log z_0^h,\log z^h(\widetilde{R})\right]}^{\max\left[\log z_0^h,\log z^h(\widetilde{R})\right]}\frac{d\log\check{c}^h(z,\widetilde{R})}{d\log z}d\log z\\
&=\frac{\overline{\eta}}{\underline{\varepsilon}}\left|\log\underbrace{\check{c}^h(z_0^h,\widetilde{R})}_{=c^h(\widehat{R}^h(\cdot;z_0^h,\widetilde{R}))}-\log\underbrace{\check{c}^h(z^h(\widetilde{R}),\widetilde{R})}_{=c^h(\widetilde{R})}\right|\leq\frac{\overline{\eta}}{\underline{\varepsilon}}d_c\tilde{\delta}.
\end{aligned}$$

In the second-to-last step, we have used that $\check{c}^h(z,\widetilde{R})$ is increasing in z, by Assumption 2. In the final step, d_c—and so the entire bound—is constant across all $h\in\mathcal{H}$ and $\widetilde{R}\in R+B_\delta(\mathbf{0})$ (by Lemma 3; see Footnote 56), and we have used that $\widehat{R}^h(\cdot;z_0^h,\widetilde{R})),\widetilde{R}\in R+B_{\tilde{\delta}}(\mathbf{0})$.

Second derivative bounds

Finally, we consider the second bound in (2.212). To start, note that for any $h\in\mathcal{H},\widetilde{R}\in R+B_\delta(\mathbf{0}),z\in$

$[\min[z_0^h, z^h(\widetilde{R})], \max[z_0^h, z^h(\widetilde{R})]]$, we have

$$
\begin{aligned}
&\frac{d}{dz}\frac{(\phi^h\circ u^h)_{cc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}\\
=&\left(\frac{(\phi^h\circ u^h)_{ccc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})^2}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}+\frac{(\phi^h\circ u^h)_{cc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}-\left(\frac{(\phi^h\circ u^h)_{cc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}\right)^2\right)\frac{\check{c}^{h\prime}(z;\widetilde{R})}{\check{c}^h(z,\widetilde{R})}\\
&+\left(\frac{(\phi^h\circ u^h)_{ccz}\left(\check{c}^h(z,\widetilde{R}),z\right)}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}-\frac{(\phi^h\circ u^h)_{cc}\left(\check{c}^h(z,\widetilde{R}),z\right)(\phi^h\circ u^h)_{cz}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)^2}\right)\\
=&\left(\frac{(\phi^h\circ u^h)_{zcc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})^2}{(\phi^h\circ u^h)_z\left(\check{c}^h(z,\widetilde{R}),z\right)}+\frac{(\phi^h\circ u^h)_{cz}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_z\left(\check{c}^h(z,\widetilde{R}),z\right)}\right.\\
&\qquad\left.-\left(\frac{(\phi^h\circ u^h)_{cz}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_z\left(\check{c}^h(z,\widetilde{R}),z\right)}\right)^2-\frac{d}{d\log c}\frac{d\log M^h(\check{c}^h(z,\widetilde{R}),z)}{d\log c}\right)\frac{\check{c}^{h\prime}(z;\widetilde{R})}{\check{c}^h(z,\widetilde{R})}\\
&+\left(\frac{(\phi^h\circ u^h)_{ccz}\left(\check{c}^h(z,\widetilde{R}),z\right)}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}-\frac{(\phi^h\circ u^h)_{cc}\left(\check{c}^h(z,\widetilde{R}),z\right)(\phi^h\circ u^h)_{cz}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)^2}\right)\\
=&-\frac{d}{d\log c}\left(\frac{d\log M^h(\check{c}^h(z,\widetilde{R}),z)}{d\log c}\right)\frac{\check{c}^{h\prime}(z;\widetilde{R})}{\check{c}^h(z,\widetilde{R})}+(\phi^h\circ u^h)_{zcc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})\left(\cancel{\frac{\check{c}^{h\prime}(z;\widetilde{R})}{(\phi^h\circ u^h)_z\left(\check{c}^h(z,\widetilde{R}),z\right)}+\frac{1}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}}\right)\\
&+\left(\frac{(\phi^h\circ u^h)_{cz}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_z\left(\check{c}^h(z,\widetilde{R}),z\right)}-\left(\frac{(\phi^h\circ u^h)_{cz}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_z\left(\check{c}^h(z,\widetilde{R}),z\right)}\right)^2\right)\underbrace{\frac{\check{c}^{h\prime}(z;\widetilde{R})}{\check{c}^h(z,\widetilde{R})}}_{=-\frac{(\phi^h\circ u^h)_z\left(\check{c}^h(z,\widetilde{R}),z\right)}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}}\\
&-\frac{(\phi^h\circ u^h)_{cc}\left(\check{c}^h(z,\widetilde{R}),z\right)(\phi^h\circ u^h)_{cz}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)^2}\\
=&-\frac{d}{d\log c}\left(\frac{d\log M^h(\check{c}^h(z,\widetilde{R}),z)}{d\log c}\right)\frac{\check{c}^{h\prime}(z;\widetilde{R})}{\check{c}^h(z,\widetilde{R})}\\
&+\frac{(\phi^h\circ u^h)_{cz}\left(\check{c}^h(z,\widetilde{R}),z\right)}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}\left(-1+\underbrace{\frac{(\phi^h\circ u^h)_{zc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{u_z\left(\check{c}^h(z,\widetilde{R}),z\right)}-\frac{(\phi^h\circ u^h)_{cc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}}_{\frac{d\log M^h\left(\check{c}^h(z,\widetilde{R}),z\right)}{d\log c}}\right)
\end{aligned}
\tag{2.218}
$$

where above we have again made use of the fact that $M^h(c,z)=-\frac{u_z^h(c,z)}{u_c^h(c,z)}=-\frac{(\phi^h\circ u^h)_z(c,z)}{(\phi^h\circ u^h)_c(c,z)}$.

Next, we use the facts that

$$
\begin{aligned}
&\frac{d\log M^h\left(\check{c}^h(z,\widetilde{R}),z\right)}{d\log c}=-\frac{\eta^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}\\
\text{and}\qquad&\frac{d}{d\log c}\frac{d\log M^h\left(\check{c}^h(z,\widetilde{R}),z\right)}{d\log c}=\frac{-\eta_{+0}^h(\widehat{R}(\cdot;z,\widetilde{R}))+2\frac{\eta^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}\varepsilon_{+0}^h(\widehat{R}(\cdot;z,\widetilde{R}))-\left(\frac{\eta^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}\right)^2\varepsilon_{+1}^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}
\end{aligned}
\tag{2.219}
$$

which are both easily verified from the formulae (2.125), (2.129), and (2.131) for elasticities and super-elasticities in the proof of Lemma 2. Note that by $\check{c}^h(z,\widetilde{R})$'s continuity in z and Assumption 2, both terms are continuous in z.

Finally—since by $\check{c}^h(z,\widetilde{R})$'s continuity in z, ϕ^h's twice-continuous differentiability, and Assumption 2,

$\frac{d}{d\log z}\frac{(\phi^h\circ u^h)_{cc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}$ is continuous in z—we may integrate:

$$
\begin{aligned}
&\left|\frac{(\phi^h\circ u^h)_{cc}\left(\hat{c}^h(V^h(\widetilde{R})),z_0^h\right)\hat{c}^h(V^h(\widetilde{R}))}{(\phi^h\circ u^h)_c\left(c^h(V^h(\widetilde{R})),z_0^h\right)}-\frac{(\phi^h\circ u^h)_{cc}\left(c^h(\widetilde{R}),z^h(\widetilde{R})\right)\hat{c}^h(\widetilde{R})}{(\phi^h\circ u^h)_c\left(c^h(\widetilde{R}),z^h(\widetilde{R})\right)}\right| \\
&=\left|\int_{\log z_0^h}^{\log z^h(\widetilde{R})}\frac{d}{d\log z}\frac{(\phi^h\circ u^h)_{cc}\left(\check{c}^h(z,\widetilde{R}),z\right)\check{c}^h(z,\widetilde{R})}{(\phi^h\circ u^h)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}d\log z\right| \\
&\le\left|\int_{\log z_0^h}^{\log z^h(\widetilde{R})}\frac{d}{d\log c}\left(\frac{d\log M^h(\check{c}^h(z,\widetilde{R}),z)}{d\log c}\right)\frac{d\log\check{c}^h(z,\widetilde{R})}{d\log z}+\left(\underbrace{\frac{d\log M^h\left(\check{c}^h(z,\widetilde{R}),z\right)}{d\log c}}_{\frac{\eta^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}}-1\right)\underbrace{\frac{d\log\left(\phi^h\circ u^h\right)_c\left(\check{c}^h(z,\widetilde{R}),z\right)}{d\log z}}_{=\frac{\eta^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}\frac{d\log\check{c}^h(z,\widetilde{R})}{d\log z}}d\log z\right| \\
&\le\left|\int_{\log z_0^h}^{\log z^h(\widetilde{R})}\left(\frac{d}{d\log c}\left(\frac{d\log M^h(\check{c}^h(z,\widetilde{R}),z)}{d\log c}\right)+\left(\frac{\eta^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}-1\right)\frac{\eta^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}\right)\frac{d\log\check{c}^h(z,\widetilde{R})}{d\log z}d\log z\right| \\
&\le\int_{min[\log z_0^h,\log z^h(\widetilde{R})]}^{\max[\log z_0^h,\log z^h(\widetilde{R})]}\left|\left(\frac{d}{d\log c}\left(\frac{d\log M^h(\check{c}^h(z,\widetilde{R}),z)}{d\log c}\right)+\left(\frac{\eta^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}-1\right)\frac{\eta^h(\widehat{R}(\cdot;z,\widetilde{R}))}{\varepsilon^h(\widehat{R}(\cdot;z,\widetilde{R}))}\right)\right|\left|\frac{d\log\check{c}^h(z,\widetilde{R})}{d\log z}\right|d\log z \\
&\le\int_{\min[\log z_0^h,\log z^h(\widetilde{R})]}^{\max[\log z_0^h,\log z^h(\widetilde{R})]}\left|\frac{\overline{\eta}_{+0}+2\frac{\overline{\eta}}{\underline{\varepsilon}}\overline{\varepsilon}_{+0}+\left(\frac{\overline{\eta}}{\underline{\varepsilon}}\right)^2\overline{\varepsilon}_{+1}}{\underline{\varepsilon}}+\left(\frac{\overline{\eta}}{\underline{\varepsilon}}\right)^2+\frac{\overline{\eta}}{\underline{\varepsilon}}\right|\frac{d\log\check{c}^h(z,\widetilde{R})}{d\log z}d\log z \\
&=\left|\frac{\overline{\eta}_{+0}+2\frac{\overline{\eta}}{\underline{\varepsilon}}\overline{\varepsilon}_{+0}+\left(\frac{\overline{\eta}}{\underline{\varepsilon}}\right)^2\overline{\varepsilon}_{+1}}{\underline{\varepsilon}}+\left(\frac{\overline{\eta}}{\underline{\varepsilon}}\right)^2+\frac{\overline{\eta}}{\underline{\varepsilon}}\right|\left|\log\check{c}^h(z_0^h,\widetilde{R})-\log\check{c}^h(z^h(\widetilde{R}),\widetilde{R})\right|
\end{aligned}
\tag{2.220}
$$

where $\overline{\eta}_{+0}$, $\overline{\varepsilon}_{+0}$, $\overline{\varepsilon}_{+1}$, $\overline{\eta}$, and $\underline{\varepsilon}$ are upper and lower bounds (as indicated by the notation) of the corresponding super-elasticities and elasticities, per Assumption 4 and the fact that all $\widehat{R}(\cdot;z,\widetilde{R})\in R+\overline{B}_{\tilde{\delta}}(\mathbf{0})$; and where above we have used the fact that $\check{c}^h(z,\widetilde{R})$ is increasing in z, by its definition and Assumption 2.

Since we have already shown the last term is uniformly bounded across $h\in\mathcal{H}$ and $\widetilde{R}\in R+B_\delta(\mathbf{0})$, we have the desired conclusion. □

2.10 Additional Tables and Figures

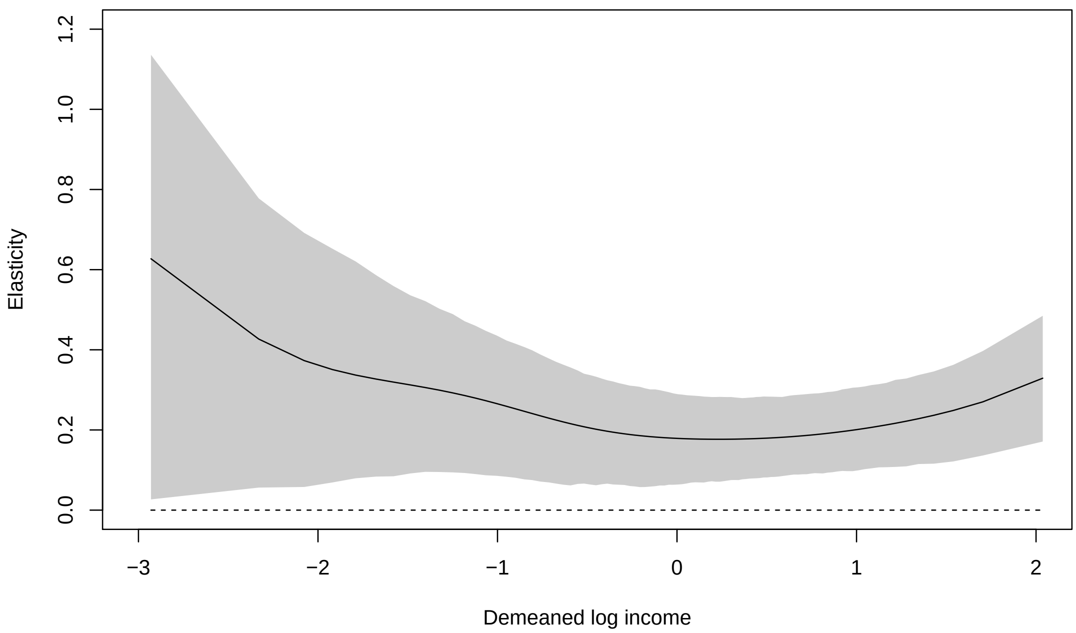


Figure 2.12: Estimates of mean elasticity by year-demeaned log income, 95% confidence bands

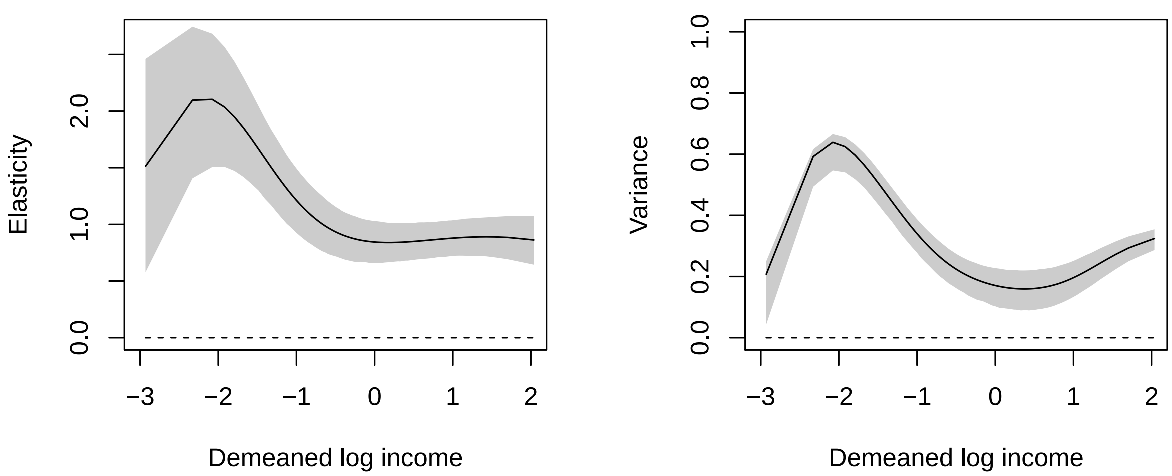


Figure 2.13: Left panel: Estimates of the difference in mean elasticities between itemization status, by year-demeaned log income Right panel: Itemization-implied lower bound on variance, by year-demeaned log income. 95% confidence bands.

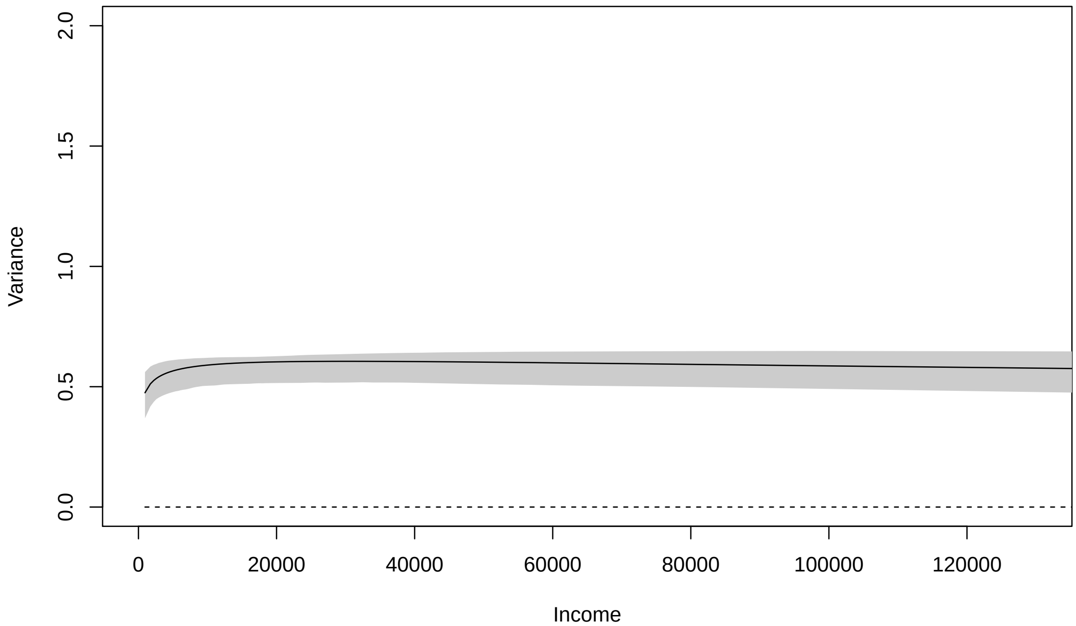


Figure 2.14: Structural estimates of variance in elasticities by income level, when maximum elasticity is capped at 5, 1990, 95% confidence bands

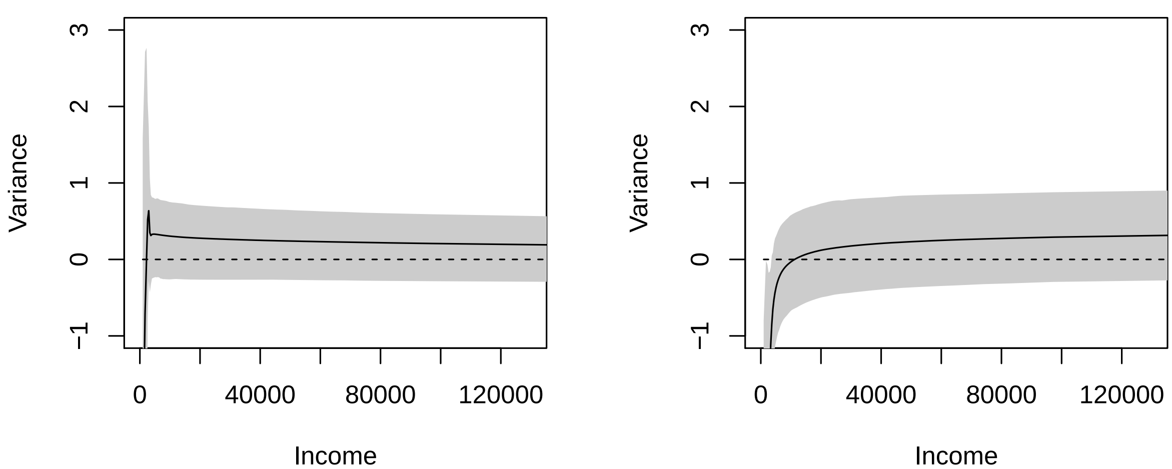


Figure 2.15: Left panel: Income-conditional variance implied by elasticity differences across low- and high-tax states. Right Panel:Income-conditional variance implied by elasticity differences across low- and high-tax state-year pairs. 1990, 95% confidence bands

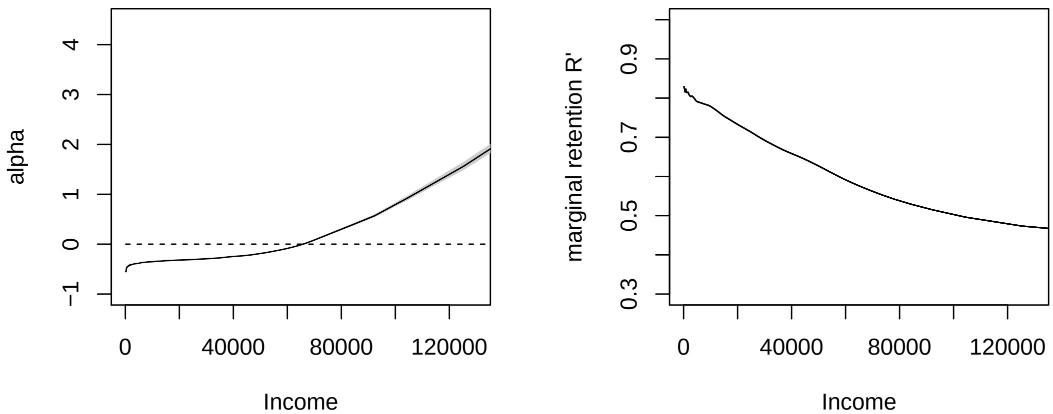


Figure 2.16: 1979. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

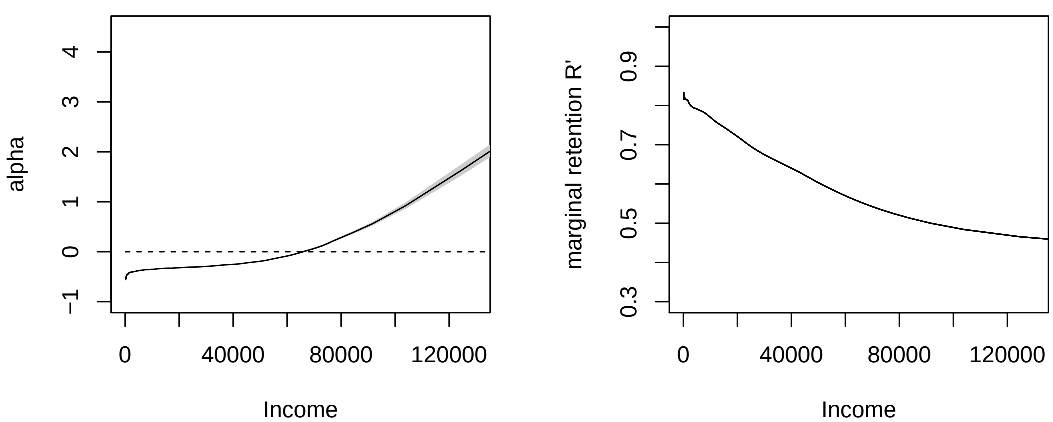


Figure 2.17: 1980. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

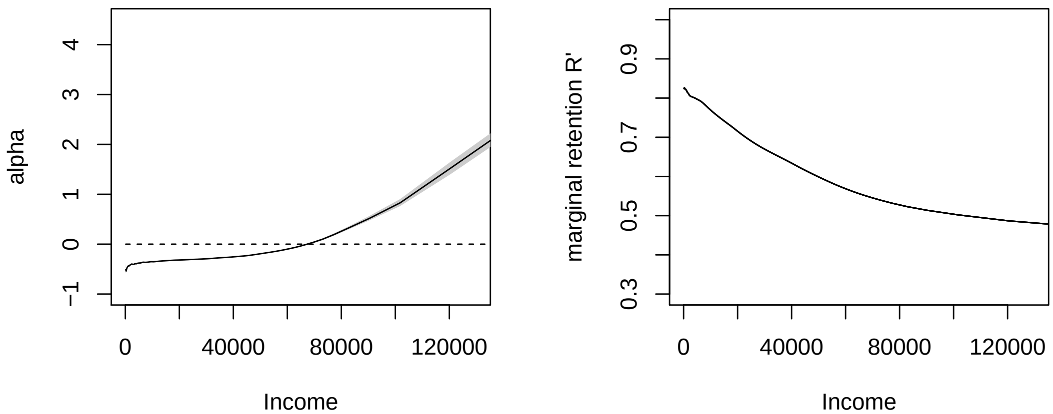


Figure 2.18: 1981. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

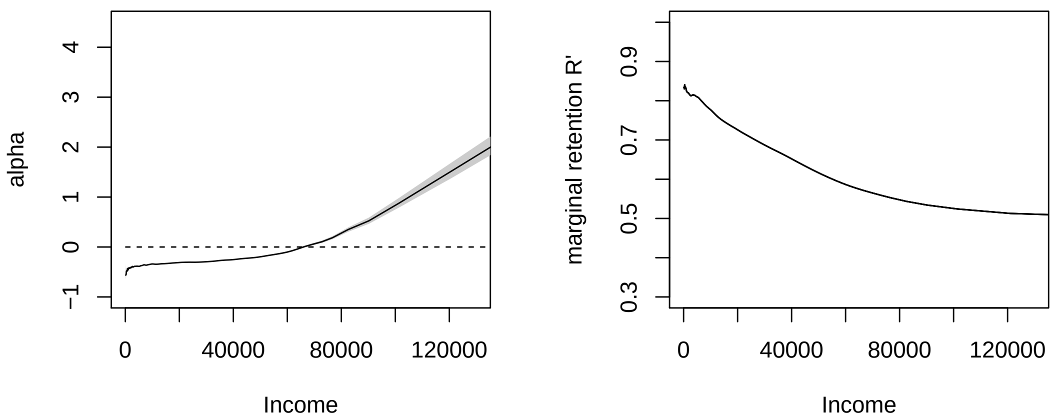


Figure 2.19: 1982. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

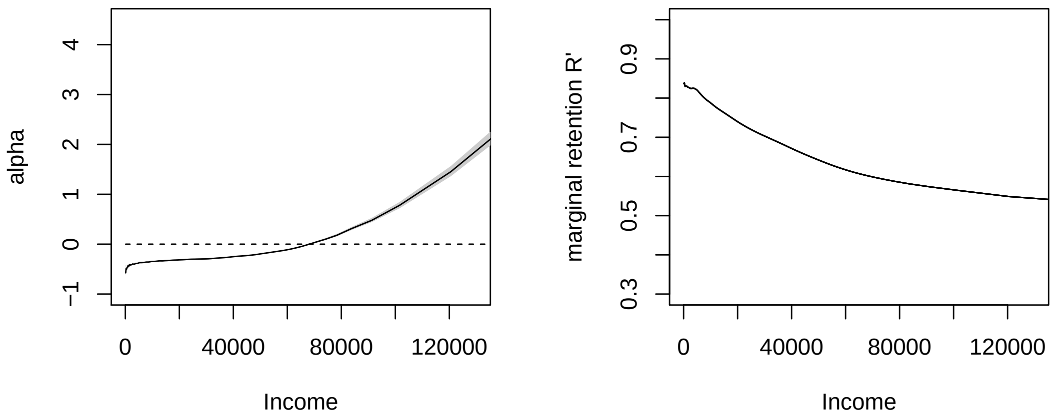


Figure 2.20: 1983. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

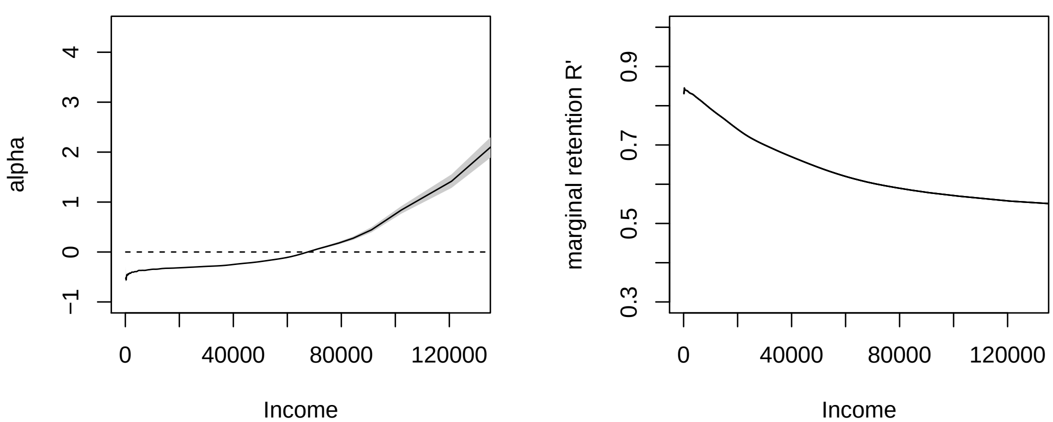


Figure 2.21: 1984. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

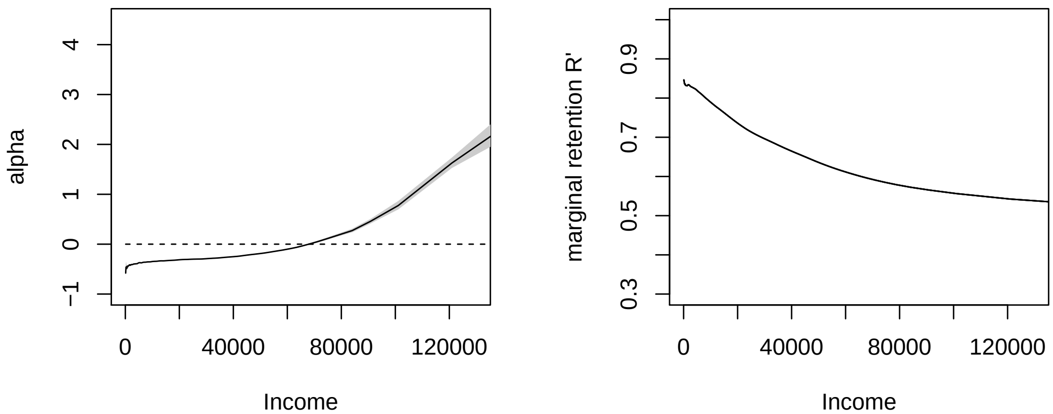


Figure 2.22: 1985. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

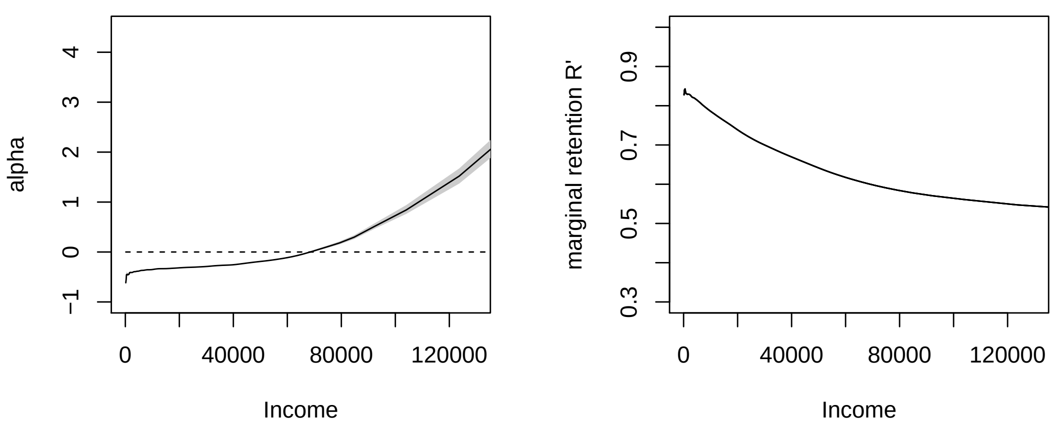


Figure 2.23: 1986. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

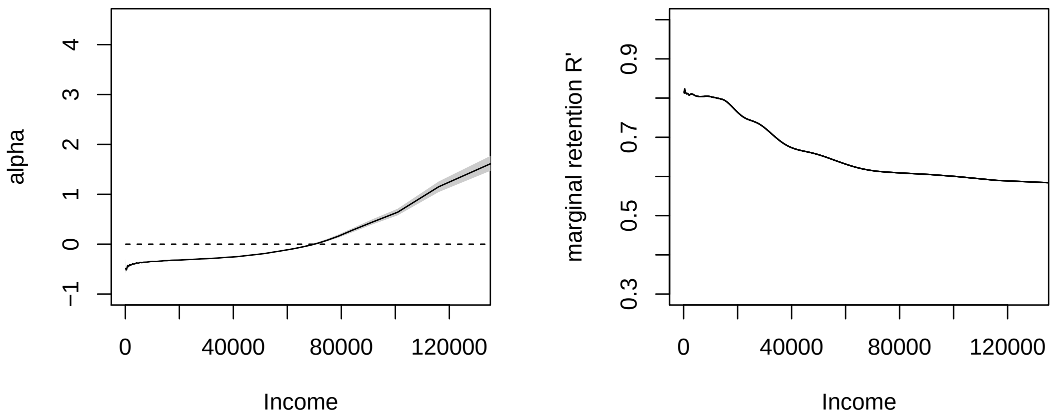


Figure 2.24: 1987. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

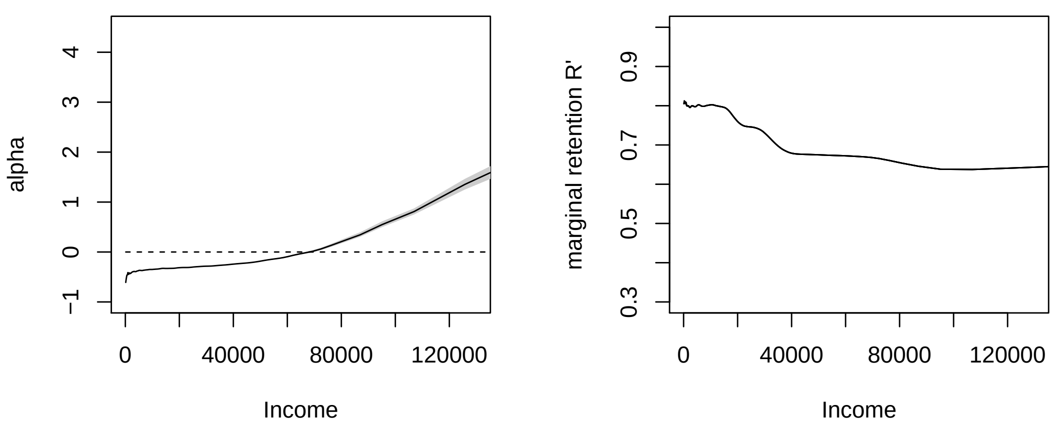


Figure 2.25: 1988. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

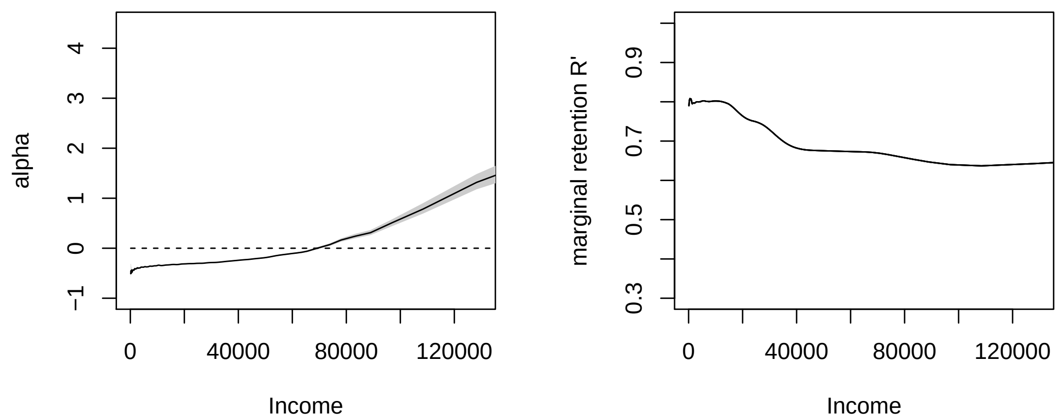


Figure 2.26: 1989. Left: Local shape of income distribution. Right: Marginal retention. 95% confidence bands

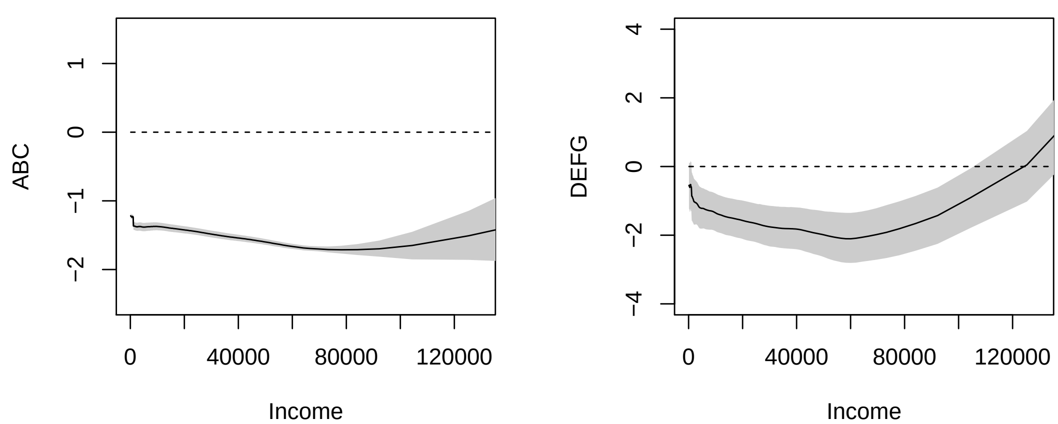


Figure 2.27: 1979. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

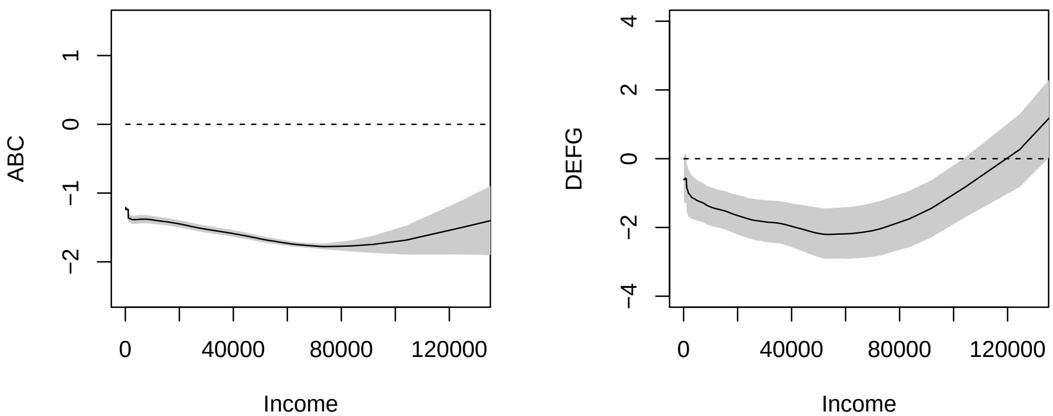


Figure 2.28: 1980. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

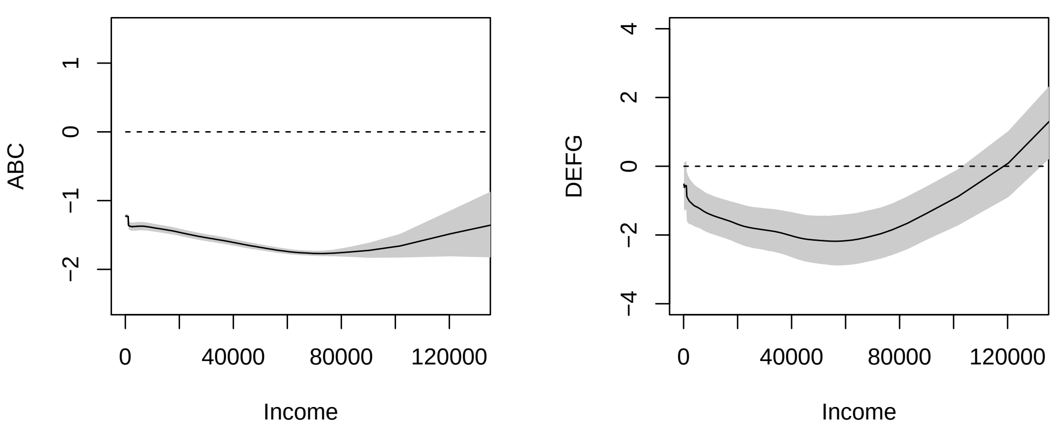


Figure 2.29: 1981. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

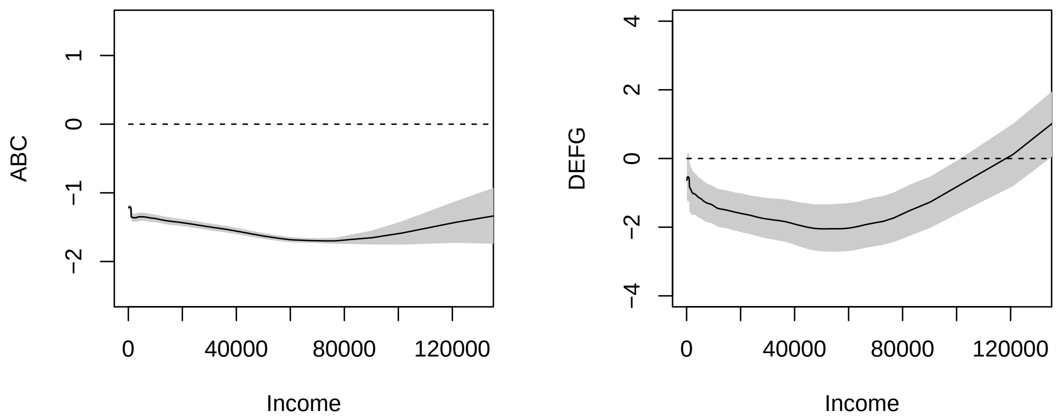


Figure 2.30: 1982. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

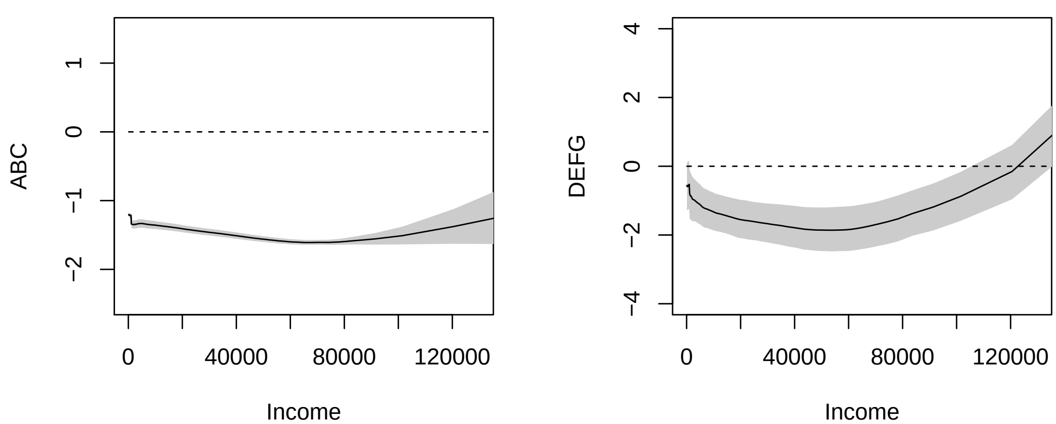


Figure 2.31: 1983. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

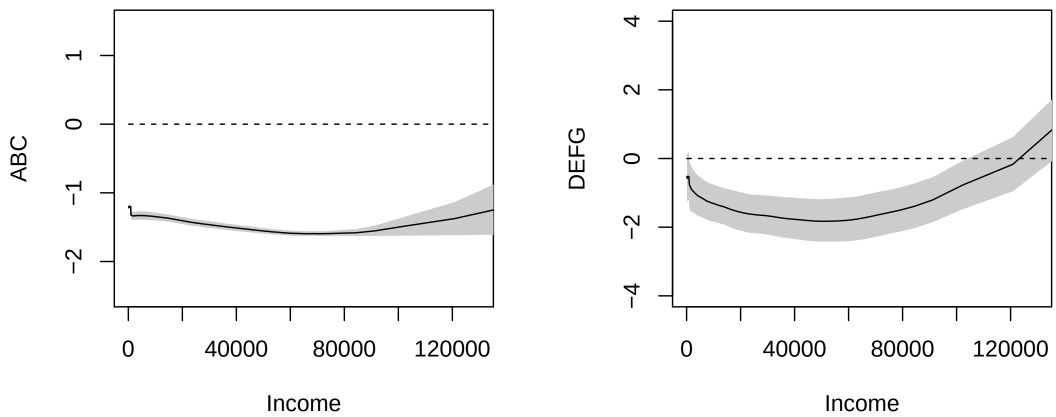


Figure 2.32: 1984. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

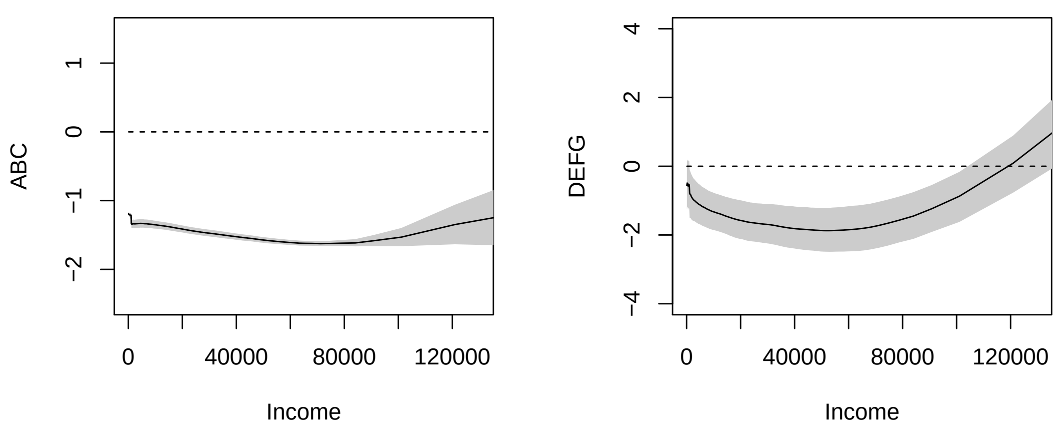


Figure 2.33: 1985. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

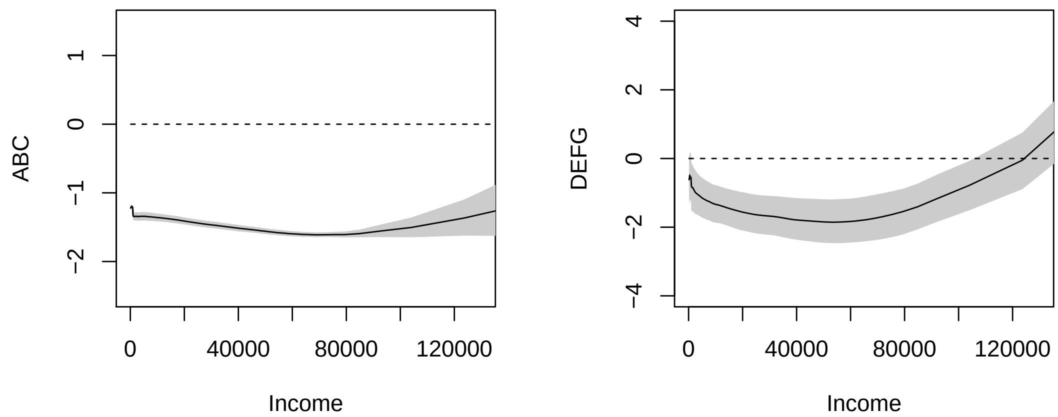


Figure 2.34: 1986. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

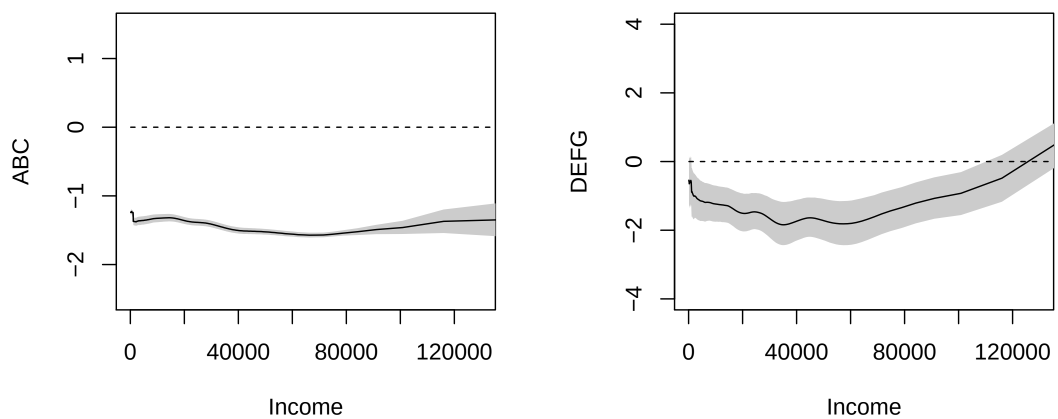


Figure 2.35: 1987. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

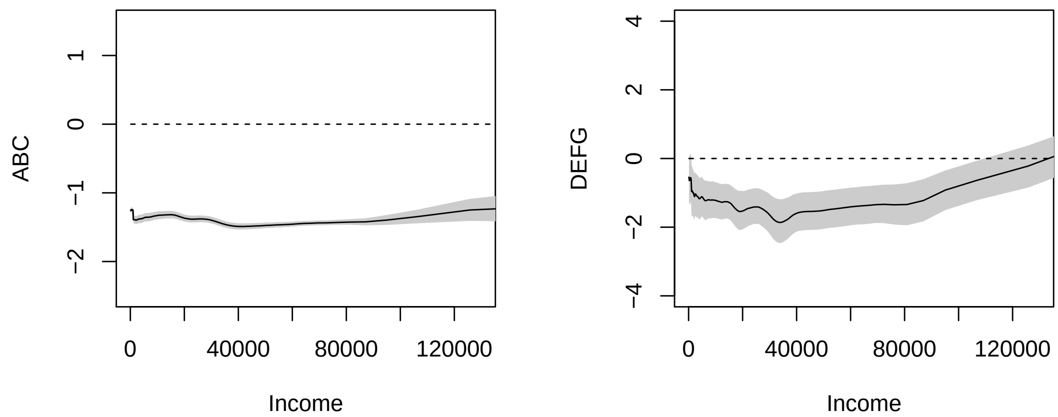


Figure 2.36: 1988. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

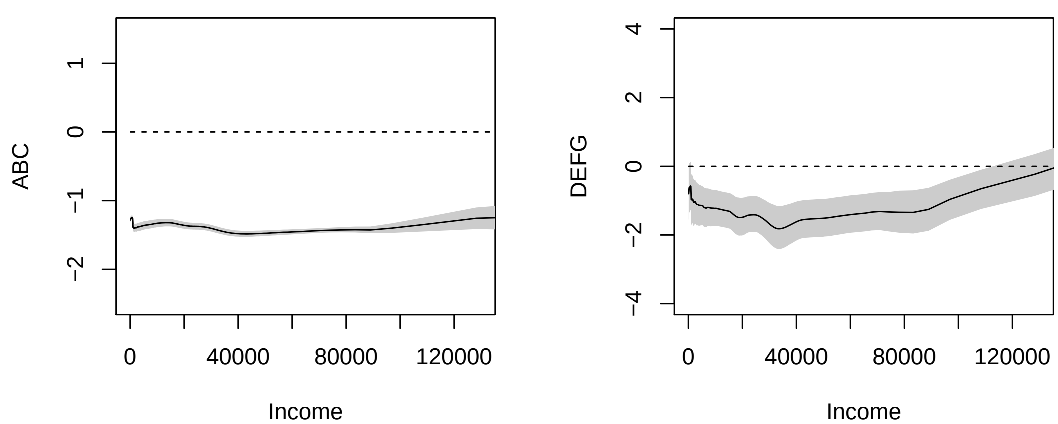


Figure 2.37: 1989. Left Panel: ABC test. Right panel: DEFG test. 90% confidence bands

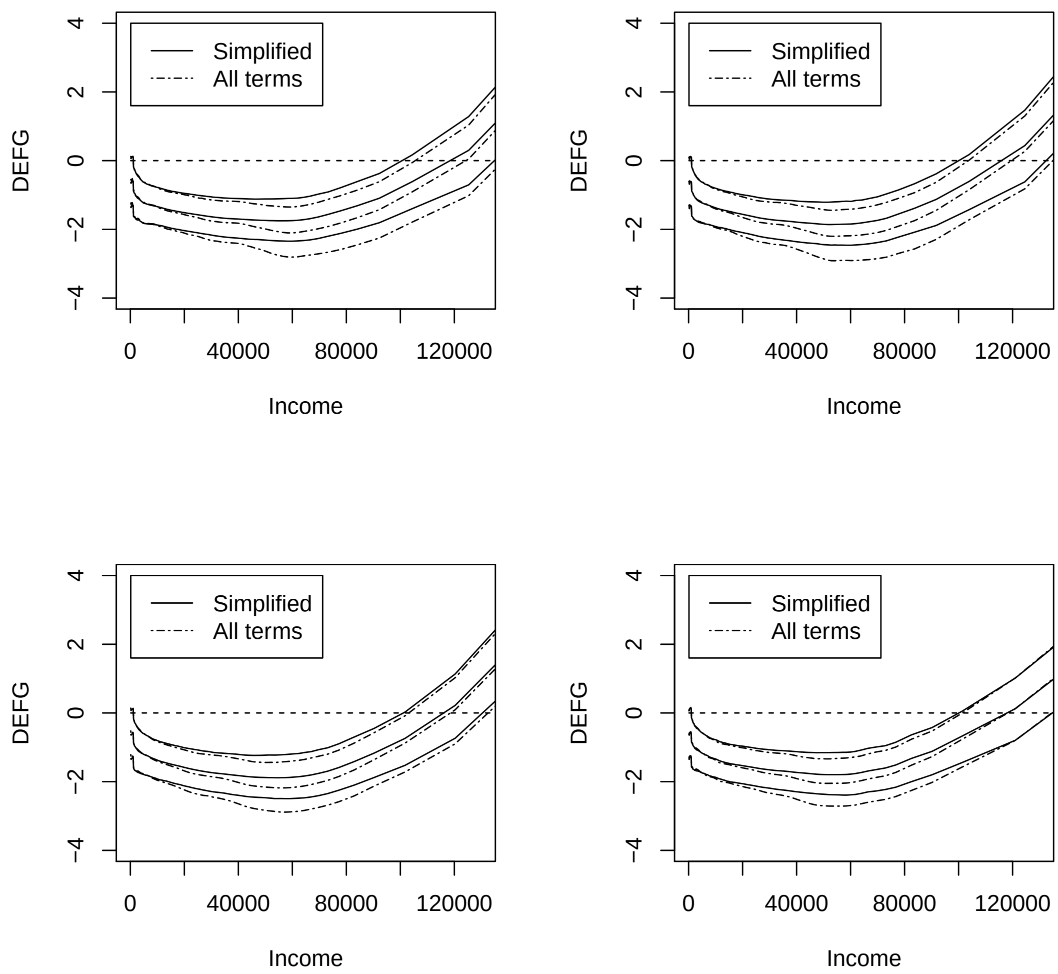


Figure 2.38: DEFG test evaluated from 1979 (top left), to 1982 (bottom right), with and without the two final terms of (2.20). 90% confidence bands

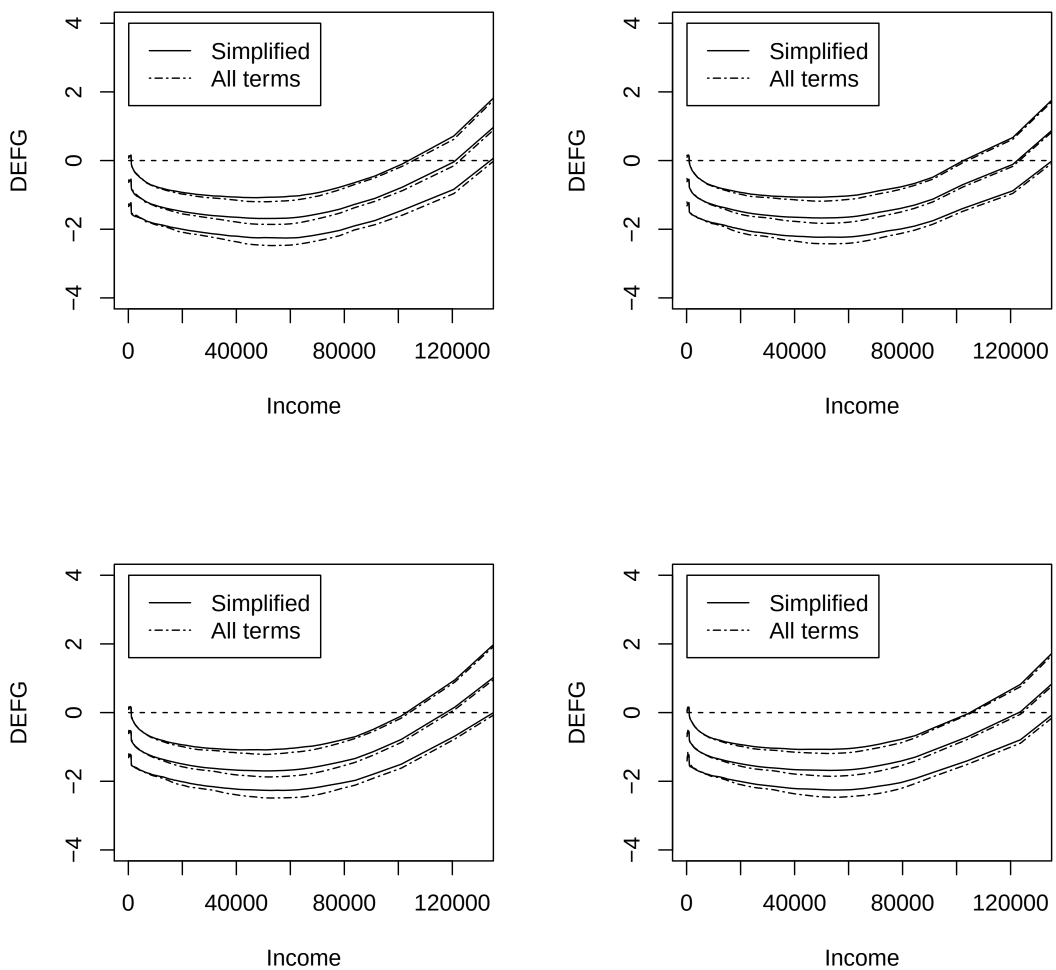


Figure 2.39: DEFG test evaluated from 1983 (top left), to 1986 (bottom right), with and without the two final terms of (2.20). 90% confidence bands

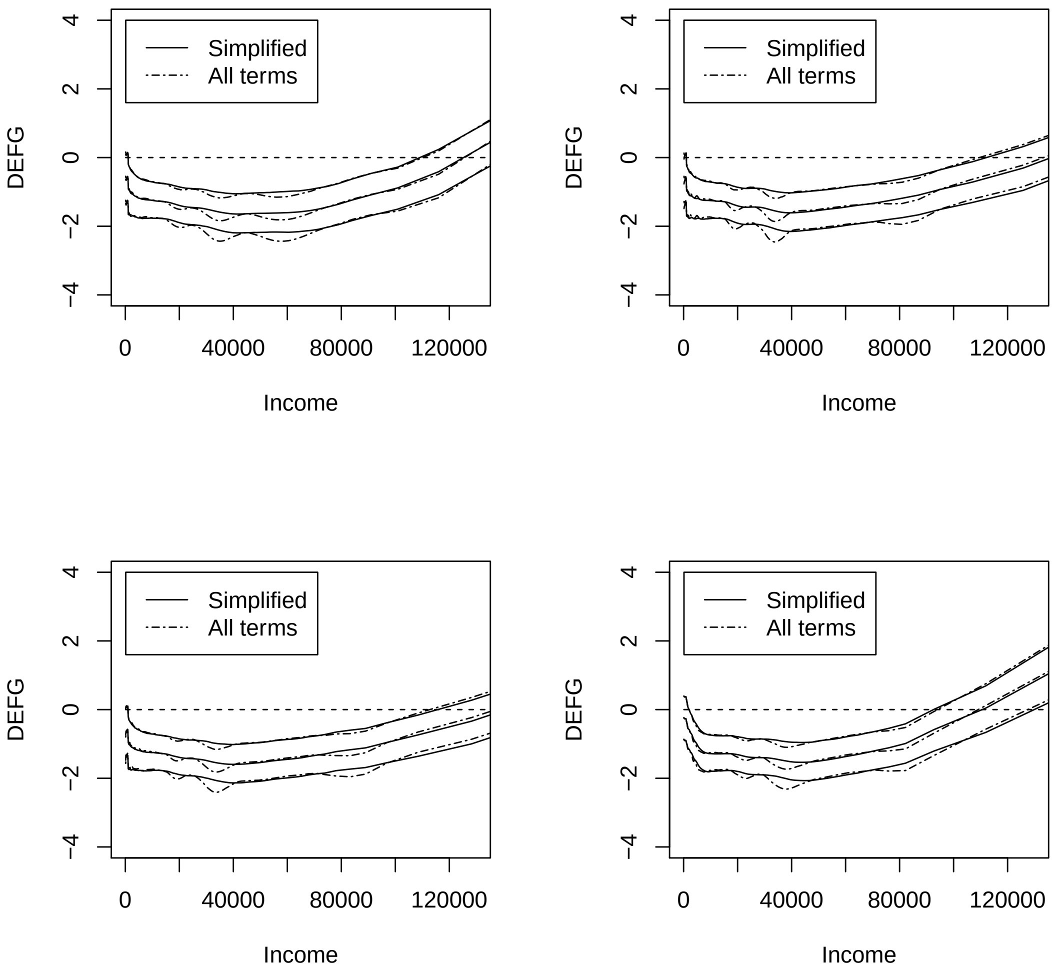


Figure 2.40: DEFG test evaluated from 1987 (top left), to 1990 (bottom right), with and without the two final terms of (2.20). 90% confidence bands

Chapter 3

Optimal Credit Scores Under Adverse Selection

Joint with Nicole Immorlica and Robert Townsend

3.1 Introduction

Data is becoming increasingly available and more easily processed. New data and methods are useful across many economic sectors and applications, including credit markets. There is a large population of potential borrowers who have short credit histories and thus are unable to receive credit (Bricker et al., 2017). The new methods would allow banks to identify the creditworthy among these potential borrowers, giving credit to those who perhaps need it the most (Jagtiani and Lemieux, 2019).

Because of this increased capacity to identify creditworthy individuals, one may hope that the inefficiencies coming from information asymmetries would progressively disappear. However, a key reason that makes information asymmetries generate inefficiencies in these thin credit market segments is adverse selection: as the price of the loans decreases (or interest rates increase), the pool of borrowers can get progressively worse. Those who would be more likely to repay are only willing to borrow at higher prices. The credit market unravels, resulting in too few or no transactions happening. As long as there is some information asymmetry, some heterogeneity in expected repayment rates that lenders cannot observe, there can still be adverse selection problems.

Data owners, such as data-intensive firms and platforms, may hope that by making their data available to financial providers they will improve credit access. However, this hope lacks a theoretical justification. The inefficiencies arising from adverse selection do not necessarily get better with more information, and indeed, may as well get worse. As shown in Levin (2001), more information does not necessarily increase the number of transactions and the realized gains from trade. More information can prevent implicit cross-subsidization between different types, making the previously subsidized types leave the market. Hence, as more data and improved technologies for processing data arrive, there remain key issues concerning how much data to share.

To answer the question of how much data to share, we build on the literature on information design and formulate the optimal disclosure problem of a partially informed intermediary with commitment, maximizing the probability of successful transactions weighted by the size of gains from trade. This formulation allows us to answer the question of which variables in a dataset should be shared with financial providers – for example, whether geographic information should be shared and at which level of granularity, or whether only an index that combines different pieces of information should be shared. We construct an optimal disclosure system and derive new conditions for the optimality of a disclosure system in terms of local sufficient statistics. The optimal policy should satisfy three simple properties: i) generically messages should combine at most two signals; ii) there should be an increasing relationship between the price elasticities of the value of the loans to investors and the prices of these loans; and iii) when different signals are combined into a single message, there should be a decreasing relationship between these elasticities and the prices these loans would have if the signals were unbundled.

We apply our results to the rural credit markets in Thailand. This is a particularly fitting setup for at least four different reasons. First, these credit markets are thin and there is not much risk sharing across villages, so the potential welfare gains are large. Second, there is evidence of intensive risk sharing within villages, which makes us think that they, through a platform acting on their behalf, are able to organize and commit to an optimal disclosure policy. Third, a unique feature of this setup benefits us from an identification perspective. There is a main lender, the Bank of Agriculture and Agricultural Cooperatives, a government-owned bank, holding a significant fraction of the market for agricultural loans. This bank uses a rigid set of rules to set interest rates. We explore variation in these rules as a source of identification for slopes of supply and average value curves. These slopes are key ingredients in the computation of the optimal credit scores and appear as sufficient statistics in the necessary conditions we derive for the optimality of disclosure systems.[1] Fourth, we benefit from rich data from Townsend Thai Project, including detailed information on consumption, income and its different sources, crops, livestock, loans, and interest rates. This allows the construction of detailed balance sheets, income and cash flow statements for each household, as well as their credit histories. Assuming that the platform has access to the detailed information in this dataset, while investors do not, we show what pieces of information should be made available to investors and how, effectively constructing "optimal credit scores."

We find that the optimal disclosure policy substantially improves the gains from trade relative to a simple full disclosure policy, with the size of gains being of the order of 0.45% the size of a typical loan per household per month. Moreover, we find that the optimal policy puts higher weight than full disclosure credit scores on variables seemingly related to the solvency of farmers relative to variables that are informative about their current liquidity. Our findings can be instrumental in improving credit access in places where it is most needed by making better use of data.

Outline of the paper: The remainder of the paper is structured as follows. Section 3.2 discusses the related literature, and Section 3.3 presents the model and a simple motivating example, Section 3.4 presents our theoretical results. Section 3.5 presents the data, followed by the discussion of the empirical strategy and empirical results in Sections 3.6 and 3.7. Section 3.8 concludes.

3.2 Related literature

This paper is related to at least three different strands of the literature. There is a long line of research that has studied adversely selected markets, starting with Akerlof (1970), including Glaeser and Kallal (1997) and Attar et al. (2011). More recently, a large literature has appeared on empirical tests for adverse selection and estimation of supply and demand in adversely selected markets (Finkelstein and Poterba, 2004; Cohen and

[1] In Appendix 3.9.3 we complement this approach with an alternative identification strategy that leverages the evidence that there is intensive risk sharing within villages in Thailand.

Einav, 2007; Einav et al., 2010; Hendren, 2013, 2017; Handel and Kolstad, 2015; Finkelstein and Notowidigdo, 2019; Cabral et al., 2019). Relative to the empirical literature on adverse selection, our contribution is to combine machine learning methods to estimate slopes of supply and demand for many different markets simultaneously. To that goal, we borrow methods from the literature on the estimation of heterogeneous elasticities (Athey and Wager, 2019; Athey et al., 2019; Davis and Heller, 2017; Wager and Athey, 2018).

Another strand in the literature has studied information design problems (Lerner and Tirole, 2006; Ostrovsky and Schwarz, 2010; Rayo and Segal, 2010; Bergemann and Morris, 2013; Bergemann et al., 2015; Kamenica and Gentzkow, 2011). Our main contributions to this literature include relating the optimal policy to sufficient statistics that can be estimated, clarifying the economic mechanisms behind simple disclosure policies, and allowing for a fairly general multidimensional distribution of types.

Within this literature, we share the linear programming approach from Kolotilin (2018), and extend and derive new results that are closest to their conditions and to the pairwise signals condition found in Kolotilin and Wolitzky (2020)[2]. In the context of adversely selected markets as in Akerlof (1970), Levin (2001) has analyzed conditions under which more information increases trade volume, besides providing examples where welfare is not monotonic in the amount of information. In the context of insurance markets, an optimal disclosure algorithm assuming a fully informed intermediary, has been developed by Garcia et al. (2018). We in contrast do not suppose the intermediary is fully informed, but rather sees a signal and then decides on an information disclosure policy. Kartik and Zhong (2019) have characterized the set of feasible payoff vectors for a buyer and a seller across all possible information structures when the seller posts the price. Besides featuring a different market structure – perfect competition on the buyers (investors) side – we focus on a different question, which is what is the best information structure an intermediary can design, when it is constrained to a limited information set.

A recent literature on information design, including some of the articles cited above, has studied in detail what is called "linear" persuasion models, where the payoffs of senders (in our case, the platform) and receivers (investors) are linear in a single dimensional state variable (Dworczak and Martini, 2019; Arieli et al., 2020; Kolotilin et al., 2017; Dizdar and Kováč, 2020). The problem we analyze does not fit these assumptions, as we allow for more general payoff structures and multidimensional types.

Finally, this paper is related to the literature that has linked consumption patterns to risk sharing and insurance, including Gruber (1997); Ahlin and Townsend (2007); Giné (2011), and Townsend (1994, 1995). In the context of Thai villages, Chiappori et al. (2014), in particular, have documented intensive consumption risk sharing at the village level.

3.3 Model

The model features three types of agents: there are potential borrowers, investors, and a platform. Borrowers sell up to a one dollar claim to investors in the present, with a promise of repaying it in the future. The platform mediates these sales and specifies the information available to the other agents. If a borrower sells a claim for a price of $x \leq 1$, then the borrower receives a loan of x dollars today and promises to repay \$1 in the future. They can sell fractional shares of this claim to multiple investors, and investors can buy claims from multiple borrowers.

Borrowers have private types $\omega \in \Omega$. The types could be, for example, the borrower's default probability. The borrower's type ω determines the borrower's opportunity cost of selling a claim $b(\omega)$. More concretely,

[2]The relationship between their results and ours is explained in more details in Section 3.4

in the Appendix 3.9.3 we present a setup where $b(\omega)$ reflects the implicit option value that selling the claim offers to the borrower, coming from the possibility that the borrower would not repay the loan in the states of the world where the marginal value of consumption is high.

Investors are homogenous, each having value $a(\omega)$ for a unit claim from borrowers of type ω. Investors share a common prior over borrower types but do not observe the type of any given borrower. Assuming investors are risk neutral, this value is simply the discounted expected repayment conditional on the type ω of the borrower.

The platform has access to signals $z(\omega)$ that are partially informative about the borrowers' types, and commits to an information disclosure policy $m(\cdot)$ which is a randomized mapping from signal realizations $z(\omega)$ to an arbitrary message. For succinctness, we will write $m(\omega) \equiv m(z(\omega))$ and note that it must be measureable with respect to the signal realizations $z(\omega)$.

After observing the signal realization, the platform sends a message $m(\omega)$ to the investors, and the investors respond by offering contracts to borrowers. As shown in Appendix 3.9.1, under a few technical assumptions, it is without loss to assume contracts are specified by a price $x(m(\omega))$ offered by investors to borrowers conditional on receiving message m. We focus on a competitive equilibrium in which investors make zero profits. If there are multiple prices that would guarantee investors break even, we take the highest price, which is the best price for the borrowers. More formally, we adopt the following definition.

Definition 1. *A competitive equilibrium is a set of prices $x^*(m(\omega))$ and allocations such that:*

- *investors break even and prices are borrower-optimal:*

$$x^*(m(\omega)) = sup(x|x = E[a(\omega)|x \geq b(\omega), m(\omega)]),$$

- *borrower types for which $b(\omega) \leq x^*(m(\omega))$ borrow up to the borrowing limit of* \$1*,*
- *and borrower types for which $b(\omega) > x^*(m(\omega))$ do not borrow.*

In general, there can be multiple solutions to the fixed-point equation $x = E[a(\omega)|x \geq b(\omega), m(\omega)])$. Regions where the expectation is decreasing as price decreases correspond to adverse selection: the borrowers willing to sell claims at this lower price are worse borrowers (and there are fewer of them), so the value of the investor goes down and the market may unravel resulting in no trade. Similarly, regions where the expectation is increasing as the price decreases correspond to advantageous selection and can result in unraveling in which all trades, even inefficient ones, happen.

The key idea behind Definition 1 is that, i) investors need to take into account that the price affects the pool of the types that are willing to borrow; and, ii) if there are multiple prices that make the investors break even, and if the equilibrium were not the highest among them, some investor could offer a price in a neighborhood of the highest, attracting the borrowers and generating positive profits.

We illustrate our model and implications for optimal disclosure in the following example adapted from Levin (2001). In particular, this example demonstrates that information can decrease welfare.

Example 1. *There are three types of borrowers $\omega \in \Omega$, associated with their repayment probabilities: low (L), medium (M) or high (H). The three types occur with equal probability $\rho(\omega) = 1/3$. The platform has access to binary signal realizations. If $\omega = H$, the signal $z(\omega)$ is $\{H\}$. Else, if $\omega = L$ or $\omega = M$, the signal $z(\omega)$ is $\{M, L\}$. In other words, the platform can differentiate H borrowers from M and L borrowers, but can not differentiate M and L borrowers from each other. The borrowers' and investor values $b(\omega)$ and $a(\omega)$ are as in Table 3.1.*

Table 3.1: Joint distribution of values in the example

	L	M	H
value for investors (as buyers) $a(\omega)$	0.12	0.3	0.84
value for borrower (as sellers) $b(\omega)$	0.06	0.24	0.36
probability $\rho(\omega)$	1/3	1/3	1/3

Consider the following disclosure policies.

- *In the full disclosure policy, the platform sends message* $m(\omega) = z(\omega)$*. In this case, in the* H *market there would be no information asymmetry. Competition between the investors would drive the price up to their willingness to pay,* 0.84*, and these gains from trade would be realized by the borrowers. In the* L *or* M *market there is asymmetric information. There cannot be a pooling equilibrium in this market, where a bundle of the two types would be sold together, because investors' willingness to pay is less than the reservation value of a borrower of type* M*, that is,* $E[a(\omega)|\omega \in \{M,L\}] = 0.21 < 0.24$*. Thus,* M*'s would not be willing to sell, and the bundle unravels to* L*. So there will be a separating equilibrium, and only the* L *will types trade, at a price of* 0.12*, which we can call the full disclosure price of the bundle* $\{M,L\}$*. The total surplus generated by the sales of* L *and* H *borrower claims would be* $(0.06+0.48)/3 = 0.18$*, and not all gains from trade would be realized, as the* M *types are left out of the market.*

- *In the no-disclosure policy, the platform sends a null message. Then there is a pooling equilibrium because the investor willingness to pay for the whole bundle is higher than the highest reservation value for the borrowers —* $E[a(\omega)] = 0.42 > 0.36$*, and we can call analogously* 0.42 *the no-disclosure price of the market. The welfare now increases to* $(0.06+0.06+0.48)/3 = 0.2$*, and all the gains from trade are realized.*

Thus, in this example welfare decreases when the platform reveals more information, which is illustrative of why more information is not always better.[3]

3.4 Optimal Disclosure Policies

In this section, we formulate the optimal disclosure problem of the intermediary as a linear programming problem. Then we derive necessary conditions for the optimality of garbling and separating signals. These conditions give economic content to what optimal disclosure policies do and tell us simple properties that they need to satisfy. These properties are summarized in three rules-of-thumb, relating prices and elasticities of the value for investors: i) generically messages should combine at most two signals; ii) there should be an increasing relationship between the price elasticities of the value of the loans to investors and the prices of these loans; and iii) when different signals are combined into a single message, there should be a decreasing relationship between these elasticities and the prices these loans would have if the signals were unbundled. We illustrate these properties with an example at the end of this section.

[3]Moreover, there is a non-monotonic relationship between information and welfare: suppose the platform has full information, i.e., $z(\omega) = \omega$. Then welfare would increase again with disclosure policy $m(\omega) = z(\omega) = \omega$. Hence full disclosure would be as good as no disclosure which, in turn, is better than partial disclosure of the form described above.

3.4.1 Characterization

As demonstrated by Example 1, the choice of disclosure policy impacts welfare. The optimal disclosure policy can be described by a linear program whose constraints characterize the competitive equilibria of Definition 1 and whose objective maximizes welfare.

This linear program can be simplified by noticing two facts: First, messages can be identified by their equilibrium prices. This is because, if investors break even at the same price for two messages, they still break even if these two messages are combined into a single message. Further, no higher price generates positive profits. Therefore we can label each message by its equilibrium price. Second, because welfare is increasing in prices, the supremum from Definition 1 can be dropped. That is, we can simply confine $x = E[a(\omega)|x \geq b(\omega), m(\omega)]$ and the objective will guarantee we select the highest such x.

Thus, for an arbitrary distribution of signals $G(z)$[4], we can formulate the optimal informational disclosure problem as follows. Note the choice variables are the conditional distribution of messages (equivalently, equilibrium prices) given signals. By assigning different probabilities of the signals to different messages, the platform affects the objective function because for each of these messages there will be a different equilibrium price and a different pool of borrowers signing a loan contract.

Proposition 1. *The optimal informational disclosure problem can be formulated as:*

$$\begin{aligned}
&\max_{\phi(x|z)} \int_{X \times Z} \sigma(x|z) d\phi(x|z) dG(z) \\
&\text{st.} \int_{\tilde{X} \times Z} \pi(x|z) d\phi(x|z) dG(z) = 0 \qquad \textit{for any measurable set } \tilde{X} \subset X \\
&\int_X d\phi(x|z) = 1 \qquad \textit{almost surely } \forall z \in Z
\end{aligned}$$

$$\begin{aligned}
\textit{where } \sigma(x|z) &\equiv \mathbb{E}[\mathbb{1}(x \geq b)(x-b)|z] \quad \textit{Borrowers' surplus / total surplus} \\
\textit{and } \pi(x|z) &\equiv \mathbb{E}[\mathbb{1}(x \geq b)(a-x)|z] \quad \textit{Investors' surplus}
\end{aligned}$$

Where we have denoted the expected borrowers' surplus for a prevailing price x and conditional on the signal z by $\sigma(x|z)$. Similarly, we denote the expected investors profits for a prevailing price x and conditional on signal z by $\pi(x|z)$.

3.4.2 Conditions for optimality of separating signals

By starting from a given information policy, and considering revealing less information, we can arrive at simple conditions, presented in Proposition 2[5] and Proposition 3.

To state these conditions, it will be convenient to discretize the model. Let m index a price, and j and k index signals. Consider a discrete approximation of the infinite linear programming problem, with a finite number of signals and prices, and such that for each signal z_j, its full disclosure price $x^*(z_j)$ is included in the discretization of the prices. Denote its index by $x(j)$. Let $\sigma_{mj} = \mathbb{E}[(x_m - b)\mathbb{1}(x_m \geq b)|z_j]\hat{g}(z_j)$ and

[4] To make the notation cleaner, in what follows, we drop the explicit dependence of z, a, b on ω.

[5] The conditions in Propostion 2 are analogous to the full disclosure conditions derived in Kolotilin (2018) under a similar but different set of assumptions. In particular, we do not assume a single-crossing condition on types.

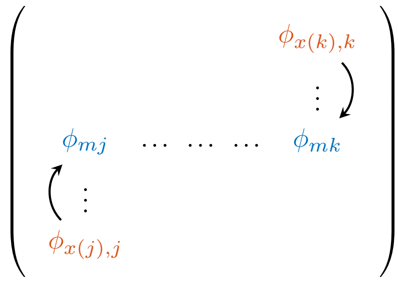


Figure 3.1: Deviations from full disclosure

$\pi_{mj} = \mathbb{E}[(a - x_m)\mathbb{1}(x_m \geq b)|z_j]\hat{g}(z_j)$, where $\hat{g}(z_j)$ is the probability of signal z_j. Let the decision variables $\phi_{m,j}$ denote the conditional probability of price m given signal j.

The first proposition says that full disclosure, which fully separates all signals, is optimal if and only if there is no benefit to pooling pairs of signals. We state this proposition as a characterization for the optimality of full-disclosure, it should be noted that it applies more generally. For any policy to be optimal, it must be the case that garbling two messages when feasible does not increase welfare. The sufficiency of this statement only holds for full disclosure because for full disclosure garbling is the only feasible perturbation.

Proposition 2. *Full disclosure is optimal if and only if*

$$\left(\sigma_{mj} - \sigma_{x(j),j} - \frac{\pi_{mj}}{\pi_{mk}}(\sigma_{mk} - \sigma_{x(k),k})\right) \leq 0 \quad \textit{for all } m,j,k \textit{ such that } \pi_{mj} > 0 \textit{ and } \pi_{mk} < 0.$$

The necessity part of the proof is based on the following argument. The expression above is the change in welfare from the following deviation: move some probability of signal z_j from being assigned to its full disclosure price to another price x_m, and make sure investors break even by moving some probability of another signal z_k to the same price x_m in the right proportion. This deviation is illustrated in matrix form in Figure 3.1. If the policy is optimal, then this change in welfare must be negative. To show that this condition is also sufficient we notice that any feasible direction can be written as conical combinations of feasible directions of the form above. The formal proof appears in Appendix 3.9.2.

To gain further economic insight into the content of this proposition and the properties of optimal disclosure policies, we further specialize to deviations of full disclosure where signals that have nearby full disclosure prices are combined.[6] This results in a simple monotonicity condition, relating prices and the price elasticity of the average value to investors. Recall that the price elasticity of value is defined as the percent change in value in response to a percent change in price, i.e.:

$$\epsilon_{V,x}(x^*(z),z) \equiv \frac{\partial E[a|b \leq x,z]}{\partial x} \frac{x}{E[a|b \leq x,z]}\bigg|_{x=x^*(z)}.$$

Proposition 3. *Suppose there is an interval $[\underline{x},\overline{x}]$ such that the full disclosure price $x^*(z)$ is dense over it. If full disclosure is optimal over this interval, then for z and z' such that $x(z) \in [\underline{x},\overline{x}]$ and $x(z') \in [\underline{x},\overline{x}]$, the signal with the higher price must have higher elasticity:*

$$x^*(z) > x^*(z') \Rightarrow \epsilon_{V,x}(x^*(z),z) \geq \epsilon_{V,x}(x^*(z'),z')$$

[6] Again, this result applies more generally than only full disclosure policies, by instead of combining signals the platform considers combining messages with nearby prices.

where $\epsilon_{V,x}(x^*(z), z)$ *is the price elasticity of the value for investors.*

This proposition has a simple graphical interpretation. Presented in Figure 3.2 is a case where the state necessary condition fails. That is, $x^*(z) > x^*(z')$ but $\epsilon_{V,x}(x^*(z), z) < \epsilon_{V,x}(x^*(z'), z')$. We will show that we can garble these signals an increase the planners' objective.

First note our equilibrium condition requires that investors break even, i.e., $x = E[a|b \leq x, z]$, and so the price elasticity of investor value reduces to:

$$\epsilon_{V,x}(x^*(z), z) = \left.\frac{\partial E[a|b \leq x, z]}{\partial x}\right|_{x=x^*(z)}$$

for equilibrium prices $x^*(z)$. That is, the elasticity at an equilibrium price is simply the slope of the average value curve at that point. In the figure, we denote this average value curve by $AVG_V(x, z) = E[a|b \leq x, z]$. In general, this can be an arbitrary function; in the figure we draw it linearly as we imagine the full disclosure prices of the signals are close and hence a linear curve is a good approximation. Points in the forty-five degree line (dashed) correspond to prices where investors break-even, that is where they pay exactly what that claim is worth for them. In the figure there are two signals, z and z'. The solid lines correspond to the average value for investors as a function of prices given those corresponding signals. The average value curve for z crosses the forty-five degree line at a higher price than the average value curve for z', that is, the full disclosure price of z is higher than z'.

Now, starting from the full disclosure configuration, let's consider the consequences of creating a message that partially garbles the two signals. In particular, let's combine units of these two signals into a new message, such that the resulting price (x_m) is the midpoint of the two full disclosure prices.

In order to do that, we need to combine them at the right proportions to make sure the investors still break even. When we lower the price of a unit of z, investors are going to face a profit of the size of the blue bar, which is the difference between the average value and the price. The average value changes as the composition of borrowers that take up the loan changes. When we do the same for a unit of z', raising their prices, investor are going to face a loss of the size of the red segment. For investors to break even, we take a number of units of z and z' that is inversely proportional to the size of these segments [7]. Because the condition fails in the proposition fails (the signal with the higher price has the flatter curve), the blue segment is larger than the red segment. Therefore, we will be lowering the price of fewer units of z than we are raising the price of units of z'. This means that on average we have increased prices. Because to a first order, the change in welfare is the change in prices times the number of units for which we are changing prices, this garbling procedure has increased welfare.

In contraposition, if the condition in Proposition 3 holds (with strict inequality), it cannot be optimal to pool two signals z and z' into a price $\bar{x}$ where $x^*(z) > \bar{x} > x^*(z')$, as depicted in Figure 3.3.

3.4.3 Conditions for optimality of pooling signals

Instead of starting from a full disclosure configuration and analyzing the consequences of garbling signals, Proposition 4 below starts from a no disclosure configuration, i.e., one in which all signals are pooled, and checks the consequences of revealing more information.

Let x_0 be the no disclosure price, and 0 its index. Let m, m' index messages, and i, j, i', j' index signals.

[7] Because we are near the region where the average value of investors is equal to prices, we can ignore the change in profits that would come from the change in quantities, because they are to a first order zero.

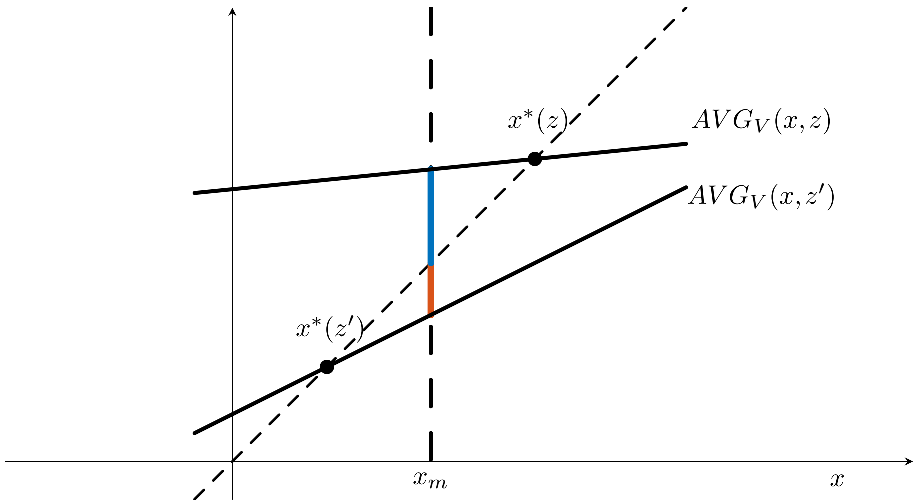


Figure 3.2: Example where condition for full disclosure to be optimal fails

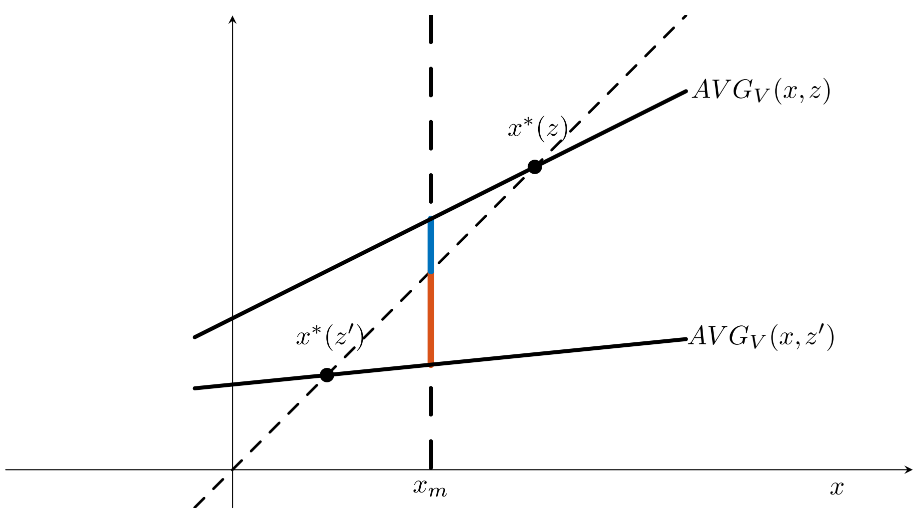


Figure 3.3: Example where condition for full disclosure to be optimal holds

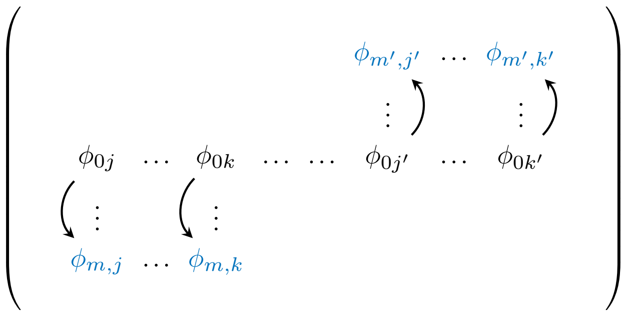


Figure 3.4: Deviations from no disclosure

For notational convenience, let

$$\Delta_\pi(m,j,k) \equiv \frac{\pi_{mj}}{\pi_{mk}}\pi_{0k} - \pi_{0j}.$$

Proposition 4. *No disclosure is optimal if and only if:*

$$\left(\sigma_{mj} - \sigma_{0j} - \frac{\pi_{mj}}{\pi_{mk}}(\sigma_{mk} - \sigma_{0k})\right) - \frac{\Delta_\pi(m,j,k)}{\Delta_\pi(m',j',k')}\left(\sigma_{m'j'} - \sigma_{0j'} - \frac{\pi_{m'j'}}{\pi_{m'k'}}(\sigma_{m'k'} - \sigma_{0k'})\right) \leq 0$$

for all m,j,k,m',j',k' such that $\pi_{mj},\pi_{m'j'},\Delta_\pi(m,j,k) \geq 0$, and $\pi_{mk},\pi_{m'k'},\Delta_\pi(m',j',k') < 0$. If there are no m,j,k,m',j',k' such that $\pi_{mk},\pi_{m'k'},\Delta_\pi(m',j',k') < 0$, then no disclosure is optimal if and only if:

$$\left(\sigma_{mj} - \sigma_{0j} - \frac{\pi_{mj}}{\pi_{mk}}(\sigma_{mk} - \sigma_{0k})\right) \leq 0$$

for all m,j,k,m',j',k', such that $\pi_{mj},\pi_{m'j'},\Delta_\pi(m,j,k) = 0$.

This is the change in welfare from the following deviation: move some probability of signal z_j away from the no disclosure price x_0 and to some other price x_m, make sure investors break even by moving some probability of another signal z_k to the same price x_m, in the right proportion. But now, investors do not break even at the price x_0; to fix this, repeat the procedure above for a price x'_m, and signals z'_j,z'_k and combine the two procedures in the right proportion. This deviation in illustrated in Figure 3.4. The quantity $\Delta_\pi(m,j,k)$ plays an analogous role to π_{mj}. It says when two signals j and k are assigned to price m in a proportion that makes investors break even at m, what is the size of the profit that investors will be making at the original price x_0.

To gain further economic insight into under what conditions revealing more increases the total value of transactions, we can specialize Proposition 4 above to the cases where disclosing information results in infinitesimally small changes in prices. Define $\pi(x,z) = \mathbb{E}[(a-x)\mathbb{1}(x \geq b)|z]g(z)$, where $g(z)$ is the density of z or the probability mass of z.

Proposition 5. *All signals z that are pooled together in x_0 must satisfy:*

$$\frac{x_0}{E[a|b \leq x_0, z] - x_0} = \kappa(x_0)\epsilon_{P,x}(x_0,z) + \gamma(x_0)$$

for some constants $\kappa(x_0)$ and $\gamma(x_0)$, and where $\epsilon_{P,x}(x_0,z) = \frac{\partial \pi(x_0,z)}{\partial x}\frac{x_0}{\pi(x_0,z)}$

The proposition implies that at any x_0, few signals should be pooled together. Most of the time, complete no disclosure is not optimal. The result arises from the following observation: in general, if there are three

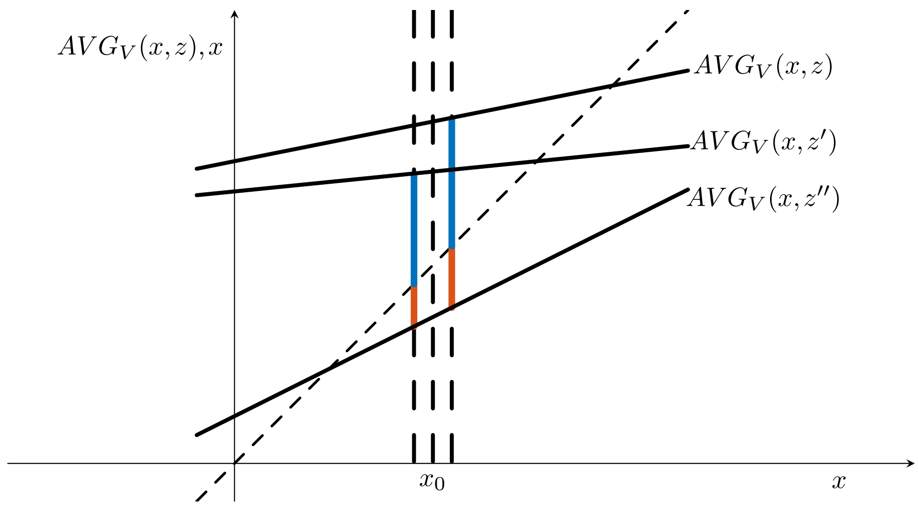


Figure 3.5: Example where the condition for no disclosure to be optimal fails

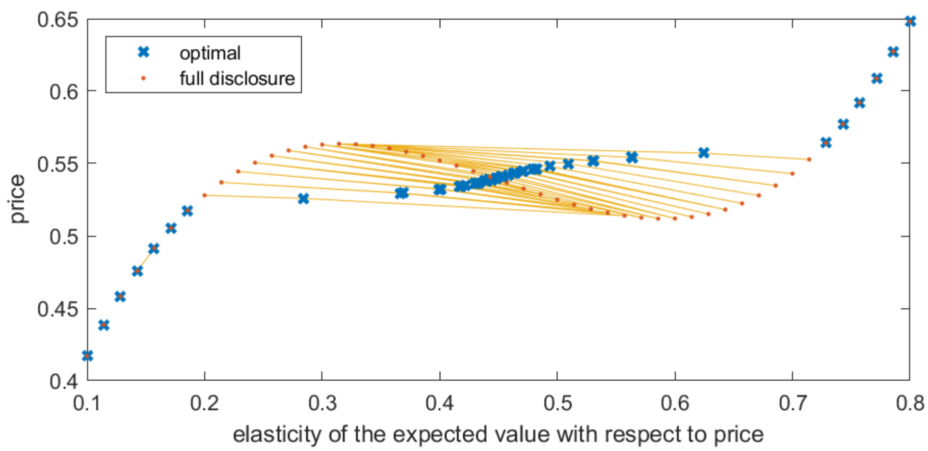


Figure 3.6: Example - Optimal information disclosure policy

or more signals combined into the same price, picking two signals and increasing their prices, and picking another pair and decreasing their prices (making sure the investors break even) either decreases or increases welfare. These possibilities are illustrated in Figure 3.5, where three signals are assigned to the same price x_0, and one tries deviate from the original configuration by raising the prices of a pair and lowering the prices of another [8]. Flipping the pairs, the planner can in general increase welfare, unless it happens that for any pair of signals the change in welfare from feasible price increases is exactly the same. This indifference can only be met if the affine relationship is satisfied. For single-dimensional or discretely distributed types, given an arbitrary x_0, this is generically satisfied only for pairs of types. For multidimensional types in $\mathbb{R}^d$ this defines a subspace of dimension $d-1$. [9]

[8] Although we plot the average values, the change in profits in this case also depends on the change in quantities, which are not plotted.

[9] A similar result appears in Kolotilin and Wolitzky (2020) stating that is generically optimal to use pairwise signal structures. In their case, payoffs are more general, but there is a single dimension of heterogeneity. Our result additionally states that signals should satisfy this affine relationship, which for multidimensional types in $\mathbb{R}^d$ defines a subspace of dimension $d-1$.

3.4.4 Economic implications

The conclusions from Propositions 3 and 5 can be summarized in three simple rules of thumb. To arrive at those, we also notice that Proposition 3 applies more generally, not only for signals that the platform sees, but also to messages that combines multiple signals. Thus, these rules of thumb state that:

1. Markets with slightly higher prices should have higher price elasticities of the expected value for investors;

2. If two signals with nearby full disclosure prices are combined into a single message, the signal with a higher full disclosure price should feature a lower elasticity. In the example, whenever signals are combined, the signal with the high full disclosure price has a low elasticity;

3. Generically, each message should combine few signals. In the discrete case, one or two signals.

Figure 3.6 illustrates these rules in a hypothetical example. In this example, we postulate that conditional on each signal, supply and demand are such that: i) conditional on each signal, elasticities are constant as a function of prices; ii) elasticities of supply are equal to one for every signal; and iii) the signals are uniformly distributed, that is $\rho(z) = 1/\#Z$.

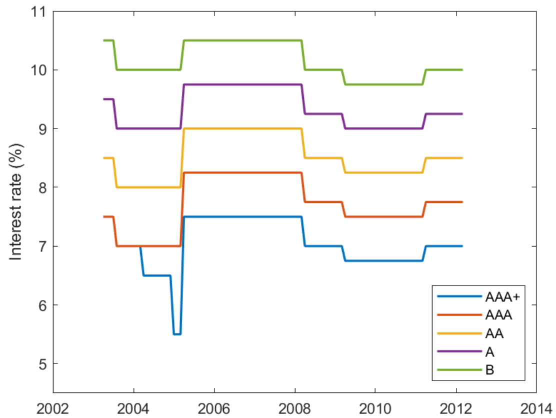
In Figure 3.6, each red dot is a different z, with their full disclosure price displayed on the vertical axis, and the elasticity of the expected value for the investors on the horizontal axis. Blue crosses represent the optimal policy, each cross is a different message x, and on the horizontal axis is their average elasticity.

In line with the rules of thumb, one can see that the arrangement of the blue crosses are such that there is an increasing relationship between prices and elasticities of the value for the investors, as implied by rule (1); the yellow lines connecting the messages to the signals that are assigned to them with positive probability are downward sloping, as implied by rule (2); and each blue cross is connected to at most two red dots, as implied by rule (3).

3.5 Data

The Townsend Thai monthly survey is an intensive monthly survey initiated in 1998. The analysis presented in this paper is based on 156 months from January 1999 (month 5) to December 2011 (month 160), which coincides with 13 calendar years. The four provinces of Thailand from which the sample is drawn are Chachoengsao and Lopburi in a more developed central region and Buriram and Srisaket located in the less developed northeastern region. The sampled townships (counties) of these provinces were part of an initial larger baseline initiative in 1997. The data utilized here are the continuously sampled households, those present in the survey throughout the 156 months. As we are concerned in this paper with rural credit markets, we include only households that generated income from farm and nonfarm business activities and drop the households whose income was almost exclusively from wage earnings. In the end, there are 541 households in the sample: 129 from Chachoengsao, 140 from Lopburi, 131 from Buriram, and 141 from Srisaket.

Notably, the monthly data have been used to create for each of these business households complete consistent financial accounts: Income Statement, including revenue, expenses and disposition of income (e.g., consumption as if dividends from profits and internal investment); Balance Sheet, including assets and liabilities, both real and financial, with net worth as the residual; and Statement of Cash Flow, with flows



Symbol	Record of loan repayment
AAA+	full and in time for the past 4 years
AAA	full and in time for the past 3 years
AA	full and in time for the past 2 years
A	full and in time in the previous fiscal year
B	overdue in the previous year or a new client

Figure 3.7: BAAC Interest rate structure for farmers

for production, consumption, investment, and financing [10]. The data set also has a loan form each loan. When the loan is initiated it specifies the lender and term, then the loan is placed on a roster and tracked each month until it is fully repaid, if ever.

A key assumption of this paper is that the platform has access to all the variables of the Townsend Thai monthly survey including the financial accounts, including disaggregated to the sectoral categories: fish/shrimp, farm, business, and livestock, as well as calendar time and geographic data.

3.6 Empirical Strategy

In this section, we present our main empirical strategy using the Thai Data, which uses exogenous variation in interest rates to identify slopes of supply and demand in these adversely selected markets. In Appendix 3.9.3 we present an alternative identification strategy, leveraging the documented fact that there is intensive within-village risk sharing, to estimate a structural model where consumption risk is shared at the village level while default decisions are idiosyncratic.

The Bank of Agriculture and Agricultural Cooperatives (BAAC) uses rigid rules to set interest rates. We can use changes in the rules as a source of exogenous variation in interest rates, to estimate the slopes of supply and average value curves as in Einav et al. (2010), Cabral et al. (2019), among others. The key idea is that the variation in interest, rates together with data on take-up, allows us to infer the supply of bonds from the borrowers. Similarly, using data on repayment rates, and assuming that investors are risk neutral, as interest rates change we can trace how repayment rates and thus the average value of the new loans for investors change.

Because we are not only interested in the average slopes across the different households, but also in the observable heterogeneity in these slopes, we use causal forests (Wager and Athey, 2018; Athey et al., 2019) to estimate the levels of take-up and expected returns and, importantly, the heterogenous elasticities of take-up and of the average value for investors.

With these estimates we can, first, test whether full disclosure is optimal (or if no disclosure or the current scoring system would be optimal under a competitive equilibrium) and, second, compute the optimal policy.

[10] See Samphantharak and Townsend (2010) for details

The econometric model we will estimate can be described by the following set of equations:

$$Q_i = X_i \cdot \gamma_i + \xi_i \qquad R_i = X_i \cdot \beta_i + \zeta_i$$

where Q_i denotes whether the household i takes up a loan at a particular month, and R_i is the repayment rate conditional on taking the loan. We allow for different individuals to respond differently to changes in prices X_i. We are interested in $\gamma(z) = \mathbb{E}[\gamma_i|Z_i = z]$ and $\beta(z) = \mathbb{E}[\beta_i|Z_i = z]$, average slopes of take-up and repayment rates with respect to price x, conditional on observable characteristics Z (which correspond to what we refer to as signals in the model), and the levels $\mathbb{E}[Q_i|Z_i = z, X_i = x]$, and $\mathbb{E}[R_i|Z_i = z, X_i = x]$. These are the average slopes (with respect to price) of take-up and repayment rates, respectively, over all households in the sample who generate signal z. The interpretation of these average slopes more concretely is in how repayment rates, in percentage points, change when there is a \$1 increase in the price of a \$100 bond ($\beta(z)$); and how take-up rates change, in percentage points, when there is a \$1 increase in the price of a \$100 bond ($\gamma(z)$).

Towards estimating these slopes, we want to make the following comparison: For households with observable characteristics z, how are repayment and take-up rates different when price is x versus when the price is x'? No two households have exactly the same observable characteristics, so we need to find a way to define who is similar to whom. Fortunately, causal forests (Athey and Wager, 2019; Wager and Athey, 2018) provide us with a solution to this question. The method uses forest-based weights to group together as a function of the observables households with similar slopes, separating households with distinctly different slopes. The approach solves a 'curse of dimensionality' problem, efficiently grouping households to estimate conditional elasticities. In doing that the algorithm effectively redefines the signals z in an economically meaningfully way.

The key identification assumption that will allows to use the method is that conditional on observables, X_i is exogenous, $\{\beta_i, \zeta_i, \gamma_i, \xi_i\} \perp\!\!\!\perp X_i|Z_i$. Figure 3.7 provides justification for this assumption. The BAAC uses very coarse rules for setting interest rates, which are seldom updated. Roughly, the BAAC considers the cost of funds in the Bangcok money market but, because it receives government subsidies, it can wait wait before making changes. It is also reluctant for public relations reasons as chartered development bank to increase interest rates. Thus, the updates in those coarse rules are unlikely to track changes in the pool of borrowers in these particular sample of villages. Therefore, we use the time variation in interest rates to estimate how take-up and repayment rates change in response to interest rate changes[11].

After estimating how repayment rates change with prices, we use additional assumptions to map those into reservation values for the investor. Assuming investors are risk neutral and given a discount factor, the average value for the investors of a unit bond they can buy is given by the simple relationship $AVG_V(x, z) = \mathbb{E}[a|b \leq x, z] = \text{repayment rate} \times \text{discount rate}$. Furthermore, because take-up and repayment rates above one hundred percent or below zero would be nonsensical, we postulate that take-up and value curves are piece-wise linear, bounded by zero and one, implying that the reservation values for investors are bounded by zero and the discount rate δ, as depicted in Figure 3.8.

The estimated slopes are presented in Figure 3.9. On the left side of the figure is a histogram of the estimates of slopes of take-up $\mathbb{E}[\gamma|z]$, evaluated at each z_i in the dataset. The right hand side figure is the histogram for the estimates of slopes of repayment rates $\mathbb{E}[\beta|z]$. We can see that estimated take-up

[11] Because we can construct from the monthly surveys the individual credit histories, we can create a proxy for the BAAC rating, and conditioning on this rating, use only the time variation in interest rates to estimate the slopes of supply and repayment rates.

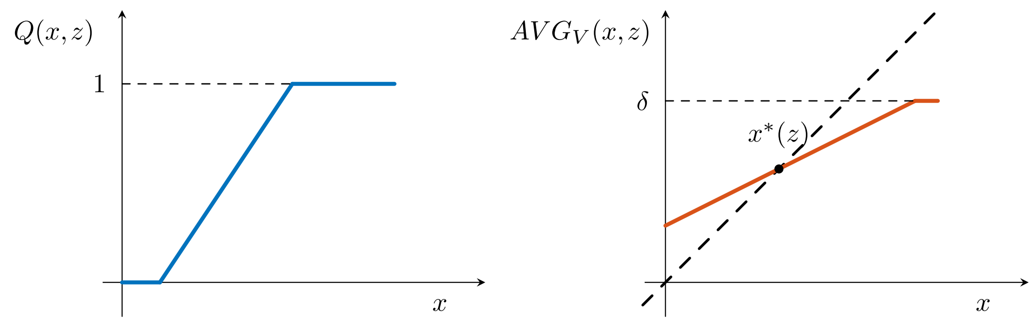


Figure 3.8: Supply and average value curves

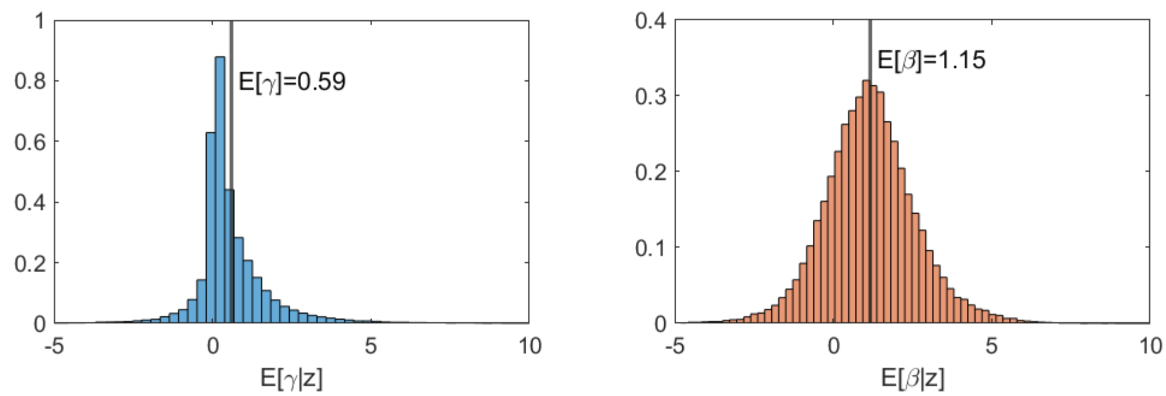


Figure 3.9: Histograms of slope estimates for take-up and repayment curves

elasticities are mostly above zero, with a mean of 0.59, and most of the estimates lying between zero and five. The slopes of repayment rates are much more spread out, showing evidence that in these rural credit markets, if all information were used, there would be both adverse selection ($\beta(z) > 0$), with higher prices leading to higher repayment rates, and advantageous selection ($\beta(z) < 0$), with higher prices leading to lower repayment rates.

To be able to compute the optimal credit scores, all we need are estimates of the joint distribution of values for investors and borrowers. Even though the estimated slopes and levels can be readily translated into this joint distribution, we take an additional step that is meant to make the problem computationally easier and interpretable. Given estimates of slopes and intercepts for supply and average value curves, we cluster them into k different groups, using a k-means algorithm[12]. The results of this estimation procedure will be used in the next section to give empirical substance to the optimal credit scores that were theoretically discussed above.

3.7 Empirical results

In this section, we combine the estimates for the joint distribution of values we presented in the previous section to arrive at the theoretically optimal credit scores. As shown in Section 3.3, it is without loss to identify messages with prices. We interchangeably refer to these messages or prices as credit scores.

[12]Alternatively, we could for example have used the empirical distribution of z for the joint distribution. However, this would require later solving a very large linear programming problem.

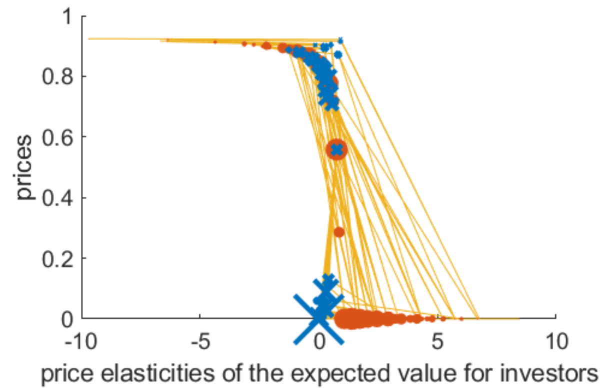


Figure 3.10: Optimal and full disclosure credit scores

Figure 3.10 illustrates the three rules of thumb we established for the optimal disclosure policy. We have shown that: markets with higher prices should have higher elasticities of the expected value for investors; in general, at most two signals should be combined in a single message; and when these two signals are combined into the same message, the one that has the higher full disclosure price should feature the lowest price elasticity of the value for investors.

These three rules of thumb imply that, in Figure 3.10, i) downward sloping yellow lines should connect two red dots to a blue cross, reflecting the second and third rules-of-thumb; ii) blue crosses should be upward sloping, reflecting the first rule-of-thumb. These rules-of-thumb, however, rely on local comparisons between adjacent prices and signals that do not need to hold when dots and crosses are further apart, and that can get blurred by the nature of the discrete approximation and the mechanics of linear programming algorithms [13]. In spite of it, we can see that most of yellow lines are downward sloping, they connect two red dots to a blue cross, and blue crosses are (mostly) upward sloping.

To compute welfare gains from the optimal policy relative to full disclosure, while avoiding counting the gains from overfitting, we use a simple form of sample splitting. We split the sample in two halves, repeating the procedure twenty times. At each time, in one half of the sample we compute the credit scoring rule, while in the second half, we apply this scoring rule and evaluate the welfare gains. This procedure shows that there would be monthly welfare gains from moving from full disclosure to the optimal policy of 0.45 % the size of a typical loan per household. This is approximately $3 per household per month.

To get a better sense of how optimal credit scores differ in practice from standard scores that would be exclusively targeted at predicting repayment probabilities, in Figure 3.11, we rank the variables in the dataset by their relevance in predicting repayment probabilities (red, on the left) and optimal credit scores (blue, on the right).

The relevance measure is a weighted sum of how many times each variable was used to create a split on the trees of the random forest, with weights proportional to the depth where the split was created. Interestingly, only two variables appear in the top ten most important variables at both of these two rankings: "total net income" and "costs: other". The most important variables for predicting repayment probabilities include for how long the individual has been a client from the BAAC, and how long since the last repayment was overdue – which are the two variables that the BAAC uses to create its own rating for the farmers. These two variables do not appear on the top ten variables of the optimal credit score. Instead, variables from the

[13] Moreover, the assumption that average value curves are piecewise linear implies that elasticities can change abruptly as prices change.

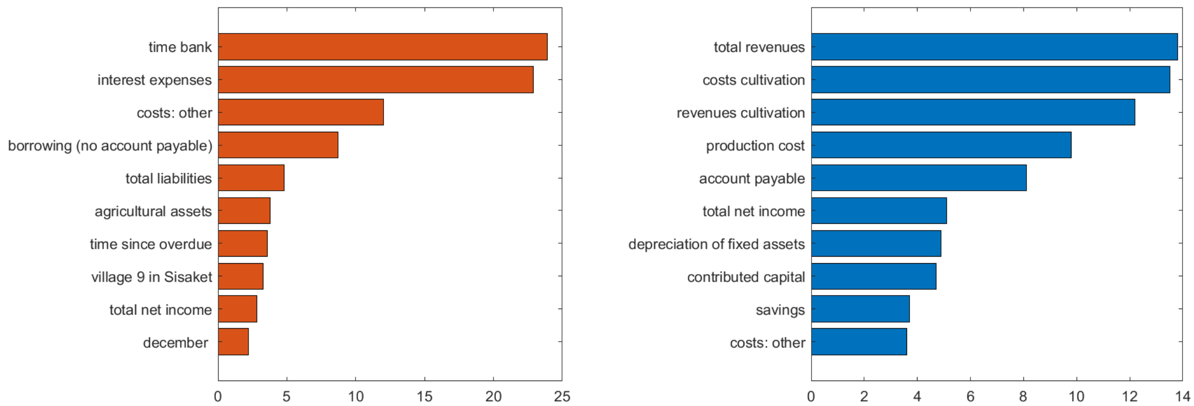


Figure 3.11: Variable importance comparison: Full disclosure vs optimal scores

balance sheet and cash flow statements of the farmers appear more prominently, with total revenues, and cultivation costs and revenues being the three most important. Surprisingly, the comparison of these two lists of variables seems to suggest that the optimal credit score would give higher weight to the solvency of farmers, instead of their short term liquidity conditions.

3.8 Conclusion

In this paper, we presented results on how to compute optimal information disclosure policies and, more generally, what they look like, focusing in on credit markets with asymmetric information.

We presented simple rules of thumb that describe the solution of the optimal disclosure problem in terms of local sufficient statistics that can be estimated. Additionally, we presented a portable empirical strategy to estimate these sufficient statistics relying on exogenous variation in interest rates[14]. We found that there are economically meaningful welfare gains from pursuing optimal disclosure policies in rural credit markets. Our estimates indicate that the monthly gains from moving from full disclosure to the optimal policy are of the order of 0.45% the size of a typical loan, or approximately $3 per household per month.

Our framework can be applied to other setups, where in principle one can think that more information could ameliorate adverse selection. Besides credit scores, the method can be applied to markets ranging from health insurance to unemployment, disability insurance and workers compensation.

[14] A second empirical strategy, presented in Appendix 3.9.3 leverages the evidence of intensive risk sharing within villages in rural Thailand.

Bibliography

Ahlin, C. and Townsend, R. M. (2007). Using repayment data to test across models of joint liability lending. *The Economic Journal*, 117(517):F11–F51.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.

Arieli, I., Babichenko, Y., Smorodinsky, R., and Yamashita, T. (2020). Optimal persuasion via bi-pooling. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 641–641.

Athey, S., Tibshirani, J., and Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178.

Athey, S. and Wager, S. (2019). Estimating treatment effects with causal forests: An application. *Observational Studies*, 5(2):37–51.

Attar, A., Mariotti, T., and Salanié, F. (2011). Nonexclusive competition in the market for lemons. *Econometrica*, 79(6):1869–1918.

Bergemann, D., Brooks, B., and Morris, S. (2015). The limits of price discrimination. *American Economic Review*, 105(3):921–57.

Bergemann, D. and Morris, S. (2013). Robust predictions in games with incomplete information. *Econometrica*, 81(4):1251–1308.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Bricker, J., Dettling, L. J., Henriques, A., Hsu, J. W., Jacobs, L., Moore, K. B., Pack, S., Sabelhaus, J., Thompson, J., and Windle, R. A. (2017). Changes in us family finances from 2013 to 2016: Evidence from the survey of consumer finances. *Fed. Res. Bull.*, 103:1.

Cabral, M., Cui, C., and Dworsky, M. (2019). What is the rationale for an insurance coverage mandate? evidence from workers' compensation insurance.

Chiappori, P.-A., Samphantharak, K., Schulhofer-Wohl, S., and Townsend, R. M. (2014). Heterogeneity and risk sharing in village economies. *Quantitative economics*, 5(1):1–27.

Cohen, A. and Einav, L. (2007). Estimating risk preferences from deductible choice. *American economic review*, 97(3):745–788.

Davis, J. and Heller, S. B. (2017). Using causal forests to predict treatment heterogeneity: An application to summer jobs. *American Economic Review*, 107(5):546–50.

Dizdar, D. and Kováč, E. (2020). A simple proof of strong duality in the linear persuasion problem. *Games and Economic Behavior*, 122:407–412.

Dworczak, P. and Martini, G. (2019). The simple economics of optimal persuasion. *Journal of Political Economy*, 127(5):1993–2048.

Einav, L., Finkelstein, A., and Cullen, M. R. (2010). Estimating welfare in insurance markets using variation in prices. *The quarterly journal of economics*, 125(3):877–921.

Finkelstein, A. and Notowidigdo, M. J. (2019). Take-up and targeting: Experimental evidence from SNAP. *The Quarterly Journal of Economics*, 134(3):1505–1556.

Finkelstein, A. and Poterba, J. (2004). Adverse selection in insurance markets: Policyholder evidence from the uk annuity market. *Journal of Political Economy*, 112(1):183–208.

Garcia, D., Teper, R., and Tsur, M. (2018). Information design in insurance markets: Selling peaches in a market for lemons. *Available at SSRN 3024500*.

Giné, X. (2011). Access to capital in rural thailand: An estimated model of formal vs. informal credit. *Journal of Development Economics*, 96(1):16–29.

Glaeser, E. L. and Kallal, H. D. (1997). Thin markets, asymmetric information, and mortgage-backed securities. *Journal of Financial Intermediation*, 6(1):64–86.

Gruber, J. (1997). The consumption smoothing benefits of unemployment insurance. *The American Economic Review*, 87(1):192–205.

Handel, B. R. and Kolstad, J. T. (2015). Health insurance for" humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8):2449–2500.

Hendren, N. (2013). Private information and insurance rejections. *Econometrica*, 81(5):1713–1762.

Hendren, N. (2017). Knowledge of future job loss and implications for unemployment insurance. *American Economic Review*, 107(7):1778–1823.

Jagtiani, J. and Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending: evidence from the lendingclub consumer platform. *Financial Management*, 48(4):1009–1029.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Kartik, N. and Zhong, W. (2019). Lemonade from lemons: Information design and adverse selection.

Kolotilin, A. (2018). Optimal information disclosure: A linear programming approach. *Theoretical Economics*, 13(2):607–635.

Kolotilin, A., Mylovanov, T., Zapechelnyuk, A., and Li, M. (2017). Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964.

Kolotilin, A. and Wolitzky, A. (2020). Assortative information disclosure.

Lerner, J. and Tirole, J. (2006). A model of forum shopping. *American Economic Review*, 96(4):1091–1113.

Levin, J. (2001). Information and the market for lemons. *RAND Journal of Economics*, pages 657–666.

Ostrovsky, M. and Schwarz, M. (2010). Information disclosure and unraveling in matching markets. *American Economic Journal: Microeconomics*, 2(2):34–63.

Rayo, L. and Segal, I. (2010). Optimal information disclosure. *Journal of political Economy*, 118(5):949–987.

Samphantharak, K. and Townsend, R. M. (2010). *Households as corporate firms: an analysis of household finance using integrated household surveys and corporate financial accounting*. Number 46. Cambridge University Press.

Townsend, R. M. (1994). Risk and insurance in village India. *Econometrica*, 62(3):539–591.

Townsend, R. M. (1995). Consumption insurance: An evaluation of risk-bearing systems in low-income economies. *Journal of Economic perspectives*, 9(3):83–102.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

3.9 Appendix

3.9.1 Omitted definitions and propositions

As discussed in the main text, definition 1 corresponds to the outcome of a game where the investors offer arbitrary menus of contracts to the potential borrowers without exclusivity, under the assumption 1 below (Attar et al., 2011).

Assumption 1.

- *The distribution P_m of b, conditional on any signal m, has bounded support.*
- *$E[a|b,m]$ exists for every b and m and it is finite.*
- *If b is an atom of P_m, then $E[a|b,m] \geq b$*
- *If $x > x^*(m)$ then $\pi(x;m) < 0$, where $\pi(x;m) = \int_{b \leq x}(a-x)\ d\pi(a,b|m)$. In other words, at any higher price than the equilibrium price for the message m, investors would make a loss.*

Denote a contract investor i offers by a vector (t_i, q_i), where t_i is how much the borrower would receive in the first period, and q_i is how much the borrower would pay the investor back in the second period.

Proposition 6. *Under Assumption 1, the non-exclusive competition between the investors result in all contracts that are traded being of the form $t_i = x^*(m)\ q_i$, where $x^*(m) = sup(x|x = E[a|x \leq b,m])$ and investors exactly break even.*

Proof. Assumption 1 is exactly parallel to the assumptions in Attar et al. (2011) and the result follows from their Propositions 1 and 2. □

3.9.2 Omitted proofs

Proposition 2. *Full disclosure is optimal if and only if*

$$\left(\sigma_{mj} - \sigma_{x(j),j} - \frac{\pi_{mj}}{\pi_{mk}}(\sigma_{mk} - \sigma_{x(k),k})\right) \leq 0 \quad \textit{for all } m,j,k \textit{ such that } \pi_{mj} > 0 \textit{ and } \pi_{mk} < 0.$$

Proof. A full disclosure configuration p is optimal if and only if for any direction $\Delta\overrightarrow{p} \in C$, where

$$C = \{\Delta\overrightarrow{p} \in \mathbb{R}^{n\times m} | \Delta p_{x(j),j} \leq 0,\ \Delta p_{m\neq x(j),j} \geq 0,\ \Delta p_{x(j),j} = -\sum_{m\neq x(j)} \Delta p_{mj} \text{ and } 0 = \sum_j \pi_{mj}\Delta p_{mj}\}$$

, the resulting change in the objective function should be non positive, that is, $\Delta W = d \cdot \Delta\overrightarrow{p} \leq 0$.

For all pair of indexes except those of the form $x(m), m$, divide the indexes in two groups, one such that $\pi_{mj} > 0$. Then construct vectors v_{mjk} such that for each index in the first group each vector v_{mjk} has a one entry at the index mj, a $-\frac{\pi_{mj}}{\pi_{mk}}$ entry for an index mk in the second group and an index -1 in the index $x(j), j$ and $\frac{\pi_{mj}}{\pi_{mk}}$ entry in the diagonal index $x(k), k$.

Notice that by construction $V \subset C$, that is, deviations in any direction of these vectors $v \in V$ are feasible. Thus if p is optimal, then any of these deviations should generate a non positive change in welfare, which is exactly the condition in the claim. This proves that it is necessary.

To prove sufficiency, first define S as the set of vectors v such that for each rs with $\pi_{rs} = 0$, there is vector $v_{rs} \in S$ with a plus one entry at rs and a minus one entry at $x(s), s$. Notice that the definition of competitive price implies that for the vectors in S, the resulting change in welfare is non positive. We are going to show that conical combinations of the vectors in $V \cup S$ generate all the feasible directions $d \in C$, starting from the full disclosure benchmark.

Notice that any direction $d \in C$, with a positive index mj with $\pi_{mj} > 0$ must have another entry mk with $\pi_{mk} < 0$, otherwise it is not a feasible direction.

The proof now proceeds by induction. Given a vector d with N positive entries r, s such that $\pi_{rs} \neq 0$, take a pair of entries mj and mk with $\pi_{mj} > 0$ and $\pi_{mk} < 0$ if there are any, and the corresponding vector $v_{mjk} \in V$ which has these positive entries. Define $\alpha_{mjk} = min(\sigma_{mj}, -\frac{\sigma_{mj}}{\pi_{mk}}\sigma_{mk})$, and build a new vector $\tilde{d} = d - \alpha_{mjk} \cdot v_{mjk}$. This new vector now has at least one zero entry less and it is still lies in C. Inductively, one can repeat the procedure until all entries rs with $\pi_{rs} \neq 0$ are zero. For the remaining entries with $\pi_{rs} = 0$ we can use the simpler vectors in S which have a plus one entry at rs and a minus one entry at $x(s), s$. We concluded implies that we can write d as a conical combination of vectors in $V \cup S$. Thus we conclude that the stated condition is sufficient.

□

Proposition 3. *Suppose there is an interval $[\underline{x}, \overline{x}]$ such that the full disclosure price $x^*(z)$ is dense over it. If full disclosure is optimal over this interval, then for z and z' such that $x(z) \in [\underline{x}, \overline{x}]$ and $x(z') \in [\underline{x}, \overline{x}]$, the signal with the higher price must have higher elasticity:*

$$x^*(z) > x^*(z') \Rightarrow \epsilon_{V,x}(x^*(z), z) \geq \epsilon_{V,x}(x^*(z'), z')$$

where $\epsilon_{V,x}(x^*(z), z)$ is the price elasticity of the value for investors.

Proof. For two signals z and z' with $x(z)$ close to $x(z')$ and $x(z) > x(z')$, the expression in proposition 1 can be rearranged as:

$$\frac{\frac{\partial \pi(x(z),z)}{dx}}{\frac{\partial \sigma(x(z),z)}{\partial x}} \geq \frac{\frac{\partial \pi(x(z'),z')}{dx}}{\frac{\partial \sigma(x(z'),z')}{\partial x}}$$

$$\text{Notice that: } \frac{\partial \sigma(x(z),z)}{\partial x} = P(z, b \leq x(z))$$

$$\text{and } \frac{\partial \pi(x(z),z)}{dx} = P(z, b \leq x(z)) \left[\frac{\partial E[a|b \leq x, z]}{\partial x} - 1 \right] + \frac{\partial P(z, b \leq x(z))}{\partial x} \left(E[a|b \leq x, z] - x \right)$$

$$= P(z, b \leq x(z)) \left[\frac{\partial E[a|b \leq x, z]}{\partial x} - 1 \right]$$

$$\Rightarrow \epsilon_V^z(x(z)) \geq \epsilon_V^{z'}(x(z'))$$

□

Proposition 4. *No disclosure is optimal if and only if:*

$$\left(\sigma_{mj} - \sigma_{0j} - \frac{\pi_{mj}}{\pi_{mk}}(\sigma_{mk} - \sigma_{0k}) \right) - \frac{\Delta_\pi(m,j,k)}{\Delta_\pi(m',j',k')} \left(\sigma_{m'j'} - \sigma_{0j'} - \frac{\pi_{m'j'}}{\pi_{m'k'}}(\sigma_{m'k'} - \sigma_{0k'}) \right) \leq 0$$

for all m, j, k, m', j', k' such that $\pi_{mj}, \pi_{m'j'}, \Delta_\pi(m,j,k) \geq 0$, and $\pi_{mk}, \pi_{m'k'}, \Delta_\pi(m',j',k') < 0$.

If there are no m, j, k, m', j', k' such that $\pi_{mk}, \pi_{m'k'}, \Delta_\pi(m',j',k') < 0$, then no disclosure is optimal if *and only if:*

$$\left(\sigma_{mj} - \sigma_{0j} - \frac{\pi_{mj}}{\pi_{mk}}(\sigma_{mk} - \sigma_{0k})\right) \leq 0$$

for all m, j, k, m', j', k'*, such that* $\pi_{mj}, \pi_{m'j'}, \Delta_\pi(m,j,k) = 0$.

Proof. The proof follows the same logic of Proposition 3. The expression is the change in welfare that results from deviating from no disclosure to a feasible direction. This vector combines two other generally infeasible vectors of the same form of the previous proposition. The first has has an one entry in the mj position, a minus one entry in the $0j$ position, a $-\frac{\pi_{mj}}{\pi_{mk}}$ entry in the mk position and $\frac{\pi_{mj}}{\pi_{mk}}$ entry in the $0k$ position. The second has analogous entries in the $m'j'$, $0j'$, $m'k'$ and $0k'$ positions. They are generally infeasible because π_{0s} is generally different than zero, while in the full disclosure case $\pi_{x(s),s}$ is zero by definition. By definition this infeasible vector should increase the expected value for investors conditional on the no disclosure price, while the second infeasible vector should decrease it. Then those are combined in the right proportion so that the zero profit constraint of the investors holds with equality, that is, at the ratio $-\frac{\frac{\pi_{mj}}{\pi_{mk}}\pi_{0k} - \pi_{0j}}{\frac{\pi_{m'j'}}{\pi_{m'k'}}\pi_{0k'} - \pi_{0j'}}$. This shows that this direction is feasible, and therefore the condition is necessary. The qualification says that if any of these ratios turn to be zero, then it is not necessary to find another entry or vector that would compensate for the violation in the zero profit condition. Observe that if on the other hand there is an index with a strictly positive π_{mj} but for this price there is no signal such that $\pi_{mk} < 0$, then it is not feasible to increase mj.

Now, to prove sufficiency, notice that any feasible direction can be decomposed as conical combinations of these directions, using the same argument in the proof of Proposition 2.

□

Proposition 5. *All signals* z *that are pooled together in* x_0 *must satisfy:*

$$\frac{x_0}{E[a|b \leq x_0, z] - x_0} = \kappa(x_0)\epsilon_{P,x}(x_0, z) + \gamma(x_0)$$

for some constants $\kappa(x_0)$ *and* $\gamma(x_0)$*, and where* $\epsilon_{P,x}(x_0,z) = \frac{\partial \pi(x_0,z)}{\partial x}\frac{x_0}{\pi(x_0,z)}$

Proof. The condition on proposition 4 can be rewritten as:

$$\frac{\left(\frac{\sigma_{mj}-\sigma_{0j}}{\pi_{mj}} - \frac{\sigma_{mk}-\sigma_{0k}}{\pi_{mk}}\right)}{\frac{\pi_{0k}}{\pi_{mk}} - \frac{\pi_{0j}}{\pi_{mj}}} \leq \frac{\left(\frac{\sigma_{m'j'}-\sigma_{0j'}}{\pi_{m'j'}} - \frac{\sigma_{m'k'}-\sigma_{0k'}}{\pi_{m'k'}}\right)}{\frac{\pi_{0k'}}{\pi_{m'k'}} - \frac{\pi_{0j'}}{\pi_{m'j'}}}$$

By assumption, $\pi_{mj} > 0$, and $\pi_{mk} < 0$. If we take x_m to be close to x_0, then $\pi_{0j} > 0$ and $\pi_{0k} < 0$. At the same time, taking this limit with $x_m > x_0$, implies that the denominator on the left side of the inequality is proportional to $\epsilon_{P,x}(x_0,z) - \epsilon_{P,x}(x_0,z')$, which by assumption is positive. Notice, however, that taking the under a sequence where $x_m < x_0$ would flip the sign of the denominator, which implies that the same pair of signals could play the role of j' and k' as long as x'_m approaches x_0 from below.

With that in mind, and taking the limit in both sides of the inequality we conclude that the condition can be written as:

$$\frac{\frac{x_0}{E[a|b\leq x_0,z_j]-x_0} - \frac{x_0}{E[a|b\leq x_0,z_k]-x_0}}{\epsilon_{P,x}(x_0,z_j) - \epsilon_{P,x}(x_0,z_k)} = \frac{\frac{x_0}{E[a|b\leq x_0,z'_j]-x_0} - \frac{x_0}{E[a|b\leq x_0,z'_k]-x_0}}{\epsilon_{P,x}(x_0,z'_j) - \epsilon_{P,x}(x_0,z'_k)}$$

Which implies that

$$\frac{\frac{x_0}{E[a|b\leq x_0,z]-x_0}-\frac{1}{E[a|b\leq x_0,z']-x_0}}{\epsilon_{P,x}(x_0,z)-\epsilon_{P,x}(x_0,z')}=k_{x_0}$$

for some constant k_{x_0}. Further, we can rewrite this expression as:

$$\frac{x_0}{E[a|b\leq x_0,z]-x_0}=\kappa(x_0)\epsilon_{P,x}(x_0,z)+\gamma(x_0)$$

□

3.9.3 Alternative identification strategy: Within village risk-sharing

In this section, we present our alternative empirical strategy to estimate the joint distribution of values for borrowers and investors. The empirical strategy leverages the evidence that there is intensive consumption risk sharing at the village level (Chiappori et al., 2014). We explore this idea, with a simple, more structural, model: individuals share consumption risk, but each has its own cut-off for which if the marginal utility of consumption at the second period is above it, she defaults.

Concretely, we use data on the time series of consumption to estimate the distribution of marginal utilities, [15]. We will assume that loans are infinitesimal, so that marginal utilities of consumption do not change with borrowing decisions, both at the individual level and at the village level. Then we turn to data on default and take-up to to estimate distribution of cut-offs.

With some convenient parametric assumptions, the model is just identified without exogenous variation in interest rates[16]. As such, to the extent that we may worry about the changes in interest rates by the BAAC not being completely exogenous to the demand and creditworthiness of the borrowers, this approach provides an alternative source of identification.

The model will generate two key equations defining the reservation values for borrowers and for investors in terms of types (that is the idiosyncratic cutoffs that trigger default) and the empirical processes for marginal utilities of consumption.

At the village-level, a household type ω is associated with their their idiossyncratic cutoff in terms of marginal utilities of consumption that would trigger default $l(\omega)$. Writing this cutoff in units of current period marginal utilities of consumption we can express the reservation value of a potential borrower as:

$$b(\omega)=\delta E_y\left[min\left\{y,\tilde{l}(\omega)\right\}\right]$$

where $\tilde{l}(\omega)$ is the idiosyncratic penalty and the cutoff that triggers default, y is common at the village-level ratio of marginal utilities of consumption between period 1 and 2, and δ is the discount rate. Intuitively, because the household does not repay the loan when the marginal utility is of consumption is higher, preferring instead to incur the non-pecuniary penalty, the loan works as an insurance device. It insures them against the states of the world where the village-level income is low, and marginal utility of consumption are high.

Assuming the investors are risk neutral, their reservation value, conditional on ω can be written as:

$$a(\omega)=\delta P_y\left(y<\tilde{l}(\omega)\right)$$

[15] We will assume, for simplicity, a common power utility, with risk aversion coefficient of three

[16] With exogenous variation in interest rates, as in the previous section, the model is non-parametrically identified

That is, for investors, the value of a unit claim to be paid in the future is the discount factor, times the probability of repayment. The probability of repayment is just the probability that the marginal utility of consumption is lower than the idiosyncratic cutoff that triggers default.

More formally, these expressions are derived from the following model: There are two periods. In the first period, each household decides whether to borrow from investors. After these decisions, the resources are pooled and a centralized decision is made with respect to consumption. In the second period, uncertainty in wealth and preferences are realized, and households decide whether to default or not, and if they default, they suffer a non-pecuniary idiosyncratic penalty. Then, within village risk-sharing takes place, that is, a centralized planner makes the second round of consumption decisions subject to a feasibility constraint.

That is, in each period $t = 1,2$, the planner solves

$$\max_{c_t^h} \sum \lambda^h u^h(c_t^h) \quad \text{s.t. } \sum c_t^h = W_t$$

The solution of this problem implies that $\lambda^h u^{h'}(c_t^h) = \mu_t$ and therefore $c_t^h = u^{h'^{-1}}\left(\frac{\mu_t}{\lambda^h}\right) = f^h(W_t)$, where μ_t is the Lagrange multiplier on the village resource constraint, and we can define $U^h(W_t) = u^h(f^h(W_t))$. For all households and every state of nature the rate of marginal utilities at time t and t' is the same, that is, there is full consumption risk sharing.

Households individually decide whether to default or not, facing a household specific cost k^h per unit of loan. Default decisions happen before consumption is realized, and the resources the household did or did not pay are added to the village level wealth W_t. Therefore the indebted household with debt b solves at the second period:

$$\max_{repay,\ default} \{U^h(W_2 - q), U^h(W_2) - l^h q\}$$

Where W_2 incorporates the equilibrium responses for the other households. This formulation implies that the equilibrium strategy is such that each household has a particular cutoff for village wealth below which she defaults. In the limit of a small loan q, the household will default if $U'^h(W_t) > l^h$. Moreover, if the total volume of loans at the village level also goes to zero, then the process for marginal utilities of consumption do not depend on the disclosure policy. If we know the distribution of penalties l^h, and the distribution of marginal utilities we can compute the distribution of default probabilities and reservation values for investors. Assuming they are risk neutral and that they have a common discount factor δ, the reservation value for investors, conditional on the household type ω, is the probability of repayment times the discount factor, that is $a(\omega) = \delta P(U'^{h,\omega}(W_t) \leq l^{h,\omega})$.

Likewise, we can find the reservation value for the potential borrowers. A borrower selling claims q and receiving t has an utility function:

$$V^h(b,t;\omega) = U_1^{h,\omega}(W_1 + t) + \delta E\left[max\{U^{h,\omega}(W_2 - q), U^{h,\omega}(W_2) - l^{h,\omega} q\}\right]$$

Which in the limit of a small loan can be written as:

$$\tilde{V}^h(b,t;\omega) = t - \delta E\left[min\left\{\frac{U'^{h,\omega}(W_2)}{U'^{h,\omega}(W_1)}, \tilde{l}^{h,\omega}\right\}\right] q \quad \text{where } \tilde{l}^{h,\omega} = \frac{l^{h,\omega}}{U'^{h,\omega}(W_1)}$$

Thus, the reservation value for a borrower of type ω is $b(\omega) = \delta E\left[min\left\{\frac{U'^{h,\omega}(W_2)}{U'^{h,\omega}(W_1)}, \tilde{l}^{h,\omega}\right\}\right]$

Given this formulation we can relate moments in the data to the moments in the model by the following set of equations.

First, the observed price x can be related to type that is just indifferent $\bar{l}$ between borrowing or not by the equation:

$$x(z) = \delta E_y \left[min \left\{ y, \bar{l} \right\} | z \right]$$

Second, all types below this cutoff $\bar{l}$ have a lower reservation value, and therefore will sell the claim. Thus the observed takeup rates ($takeup(z)$), are related to the model moments by the following equation:

$$takeup(z) = P_l(l \leq \bar{l} | z)$$

Third, the average repayment rate is given by the average probability that marginal utility of consumption is below the idiosyncratic cutoff of those who decided to borrow:

$$repayment(z) = P_{l,y}(y \leq l | l \leq \bar{l}, z)$$

With additional parametric assumptions, these three equations are going to be sufficient to identify the joint distribution of values. [17] We will assume y and l are each log normally distributed. Thus, conditional on the distribution of marginal utilities of consumption, for each z and x, we will have three equations and three unknowns ($\bar{l}(x|z)$, $\mu_{l|z}$, $\sigma_{l|z}$).

The whole model has five moments and five parameters for each z. Two moments and parameters are associated with the data on consumption – the mean and variance of marginal utility of consumption. Then there are the mean and variance of cutoffs which are associated with take-up and repayment rates, and an incidental parameter which is the type who is just indifferent $\bar{l}(x|z)$ at a price x. These last three are the moments we are referencing here.

Concretely, the first equation above, that defines the type who is just indifferent $\bar{l}(x|z)$ at a price x, becomes:

$$x(z) = \delta \int_{-\infty}^{\bar{l}(x|z)} exp(w) \phi \left(\frac{w - \mu_{u|z}}{\sigma_{u|z}} \right) dw + \delta exp(\bar{l}(x|z)) \cdot \left(1 - \Phi \left(\frac{\bar{l}(x|z) - \mu_{u|z}}{\sigma_{u|z}} \right) \right)$$

The second equation, describing the take-up rate, becomes:

$$takeup(z) = \Phi \left(\frac{\bar{l}(x|z) - \mu_{l|z}}{\sigma_{l|z}} \right)$$

And the third equation, describing the average repayment rate, becomes:

$$repayment(z) = \int_{-\infty}^{\bar{l}(x|z)} \Phi \left(\frac{w - \mu_{u|z}}{\sigma_{u|z}} \right) \frac{\phi \left(\frac{w - \mu_{l|z}}{\sigma_{l|z}} \right)}{\Phi \left(\frac{\bar{l}(x|z) - \mu_{l|z}}{\sigma_{l|z}} \right)} dw$$

Having related the theoretical and empirical moments, it remains to estimated the latter. That is, we need to estimate take-up, and repayment rates as functions of the observable characteristics z. To accomplish

[17] If we jointly consider the model equations and assume there is exogenous variation in prices, than we would be overidentified. Indeed, without the structural assumptions, we are already non-parametrically identified.

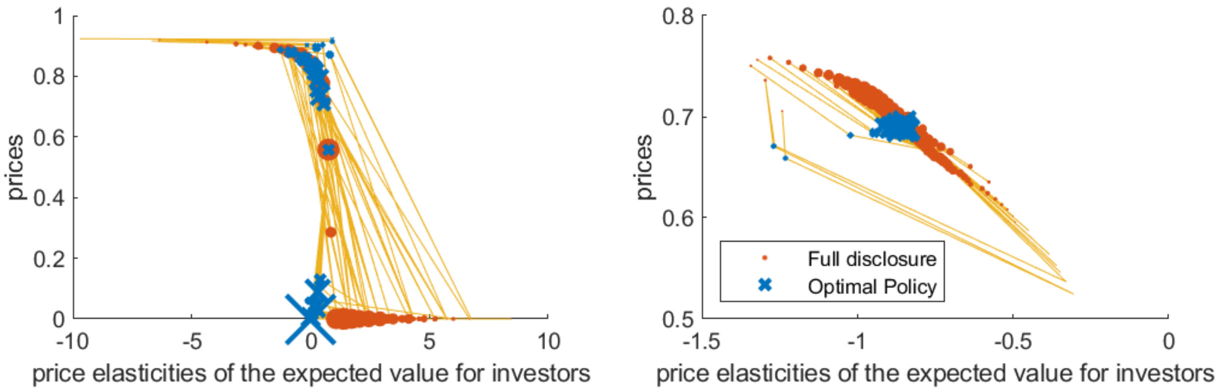


Figure 3.12: Reduced form and structural model

this task, we use random forests (Breiman, 2001; Athey and Wager, 2019). [18].

These moments are translated into estimated functions $\mu_l(z)$, $\sigma_l(z)$, $\mu_u(z)$, $\sigma_u(z)$, using the equations above. As in the previous estimation strategy, even though we could from these estimators infer the joint distributions of values and observable characteristics,we take an additional step that is meant to make the problem computationally easier and interpretable. To that goal, we use a k-means algorithm to classify households into different clusters, according to their estimated $\mu_l, \sigma_l, \mu_u, \sigma_u$.

Given the estimates for the joint distribution of values, we can compute optimal credit scores as in the section 3.7. Figure 3.12 compares these scores to the ones we derived using exogenous variation in interest rates. We can see that credit scores and elasticities of the value for investors are roughly centered around the means, but much less spread out.

[18]Random forests had the best out-of-sample performance among a variety of other machine learning methods, such as boosted trees, lasso, and linear random forests.