# Optimal Taxation with Endogenous Wages

by

Stefanie Stantcheva

B.A., University of Cambridge (2007)
M.S., Ecole Polytechnique (2009)
M.S., Ecole d'Economie de Paris and ENSAE (2009)

Submitted to the Department of Economics
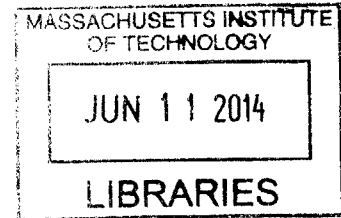in partial fulfillment of the requirements for the degrees of
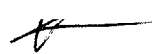
Doctor of Philosophy

and

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2014

Signature redacted

Signature of Author . . . . .
Department of Economics
15 May 2014

Signature redacted

Certified by . . . . . . . . . . . . .
James Poterba
Mitsui Professor of Economics
Thesis Supervisor

Signature redacted

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

Signature redacted

Accepted by . . . . .
Michael Greenstone
3M Professor of Environmental Economic,
Chairman, Departmental Committee on Graduate Studies

# Optimal Taxation with Endogenous Wages

by

## Stefanie Stantcheva

Submitted to the Department of Economics
on 15 May 2014, in partial fulfillment of the
requirements for the degrees of
Doctor of Philosophy
and
Master of Science

## Abstract

This thesis consists of three chapters on optimal tax theory with endogenous wages.

Chapter 1 studies optimal linear and nonlinear income taxation when firms do not know workers' abilities, and competitively screen them through nonlinear compensation contracts, unobservable to the government, in a Miyazaki-Wilson-Spence equilibrium. Adverse selection changes the optimal tax formulas because of the use of work hours as a screening tool, which for higher talent workers results in a "rat race," and for lower talent workers in informational rents and cross-subsidies. If the government has sufficiently strong redistributive goals, welfare is higher when there is adverse selection than when there is not. The model has practical implications for the interpretation, estimation, and use of taxable income elasticities, central to optimal tax design.

Chapter 2 derives optimal income tax and human capital policies in a dynamic life cycle model with risky human capital formation through monetary expenses and training time. The government faces asymmetric information regarding the stochastic ability of agents and labor supply. When the wage elasticity with respect to ability is increasing in human capital, the optimal subsidy involves less than full deductibility of human capital expenses on the tax base, and falls with age. The optimal tax treatment of training time also depends on its interactions with contemporaneous and future labor supply. Income contingent loans, and a tax scheme with deferred deductibility of human capital expenses can implement the optimum. Numerical results suggest that full dynamic risk-adjusted deductibility of expenses is close to optimal, and that simple linear age-dependent policies can achieve most of the welfare gain from the second best.

Chapter 3 considers dynamic optimal income, education, and bequest taxes in a Barro-Becker dynastic setup. Each generation is subject to idiosyncratic preference and productivity shocks. Parents can transfer resources to their children either through education investments, which improve the child's wage, or through financial bequests. I derive optimal linear tax formulas as functions of estimable sufficient statistics, robust to underlying heterogeneities in preferences. It is in general not optimal to make education expenses fully tax deductible. I also show how to derive equivalent formulas using reform-specific elasticities that can be targeted to already available estimates from existing reforms.

2

Thesis Supervisor: James Poterba
Title: Mitsui Professor of Economics

Thesis Supervisor: Iván Werning
Title: Robert M. Solow Professor of Economics

# Acknowledgments

I would like to express my enormous gratitude to my advisers, James Poterba and Iván Werning for their invaluable help, support, and guidance. Both of them, in their own special ways, always pushed me to do better, while also showing me that they believed in me. Jim has been an admirable mentor, pivotal to my professional development, and who guided me at every crossroad, big or small, during the Ph.D. He has been generous with his time and wise advice in a truly remarkable way. Iván challenged me to go deeper and further on my own and made me appreciate the unique pleasure of digging into complex problems. I received countless deep insights from him and he was always there in the important moments with lifesaving advice.

I am very grateful to Robert Townsend for our numerous interactions, our joint projects and our discussions (including during field trips in a van in Thailand), which taught me a lot about economics. I want to thank Esther Duflo who has been guiding me since my undergraduate years – when I begged her to hire me as an RA – and made me discover hands-on economic research. I owe incredibly much to Emmanuel Saez for his attention to and immense support for my academic development. Our joint projects were a joy to work on and a source of constant learning.

I benefited from many discussions with and comments from Amy Finkelstein, Nathan Hendren, Louis Kaplow, Ilyana Kuziemko, Thomas Piketty, Florian Scheuer, Matthew Weinzierl, and Victor Zhorin.

At MIT, John Arditi, Beata Shuster and Linda Woodbury repeatedly helped me with all the logistics. Special thanks to Gary King for his infinite patience and help.

My colleagues and friends Isaiah Andrews, Nina Harari, Anil Jain, Patricia Gomez, Maria Polyakova, Vincent Pons, Dan Rees, Miikka Rokkanen, Adam Sacarny, Annalisa Scognamiglio, Brad Shapiro, and Daan Struyven made the Ph.D. much more enjoyable.

Even before the Ph.D., I was lucky to encounter people who helped me, sometimes without even realizing it, and I want to thank Claude Eterstein, Michelle Baddeley, Sara Ellison, Thomas

4

On a personal note, the road that has led to the completion of this dissertation of course began long before graduate school, and I owe a lot to my parents, Elena and Pentcho Stantchev for providing me with unconditional love and support, for making many sacrifices so their children could be better off, and for giving me the (perhaps illusory but strangely helpful) sense that with hard work, I could achieve anything. My wonderful sister Zvetelina Stantcheva has been an amazing source of support and inspiration.

I am fortunate to have my dear friends whom I could always rely on: Madeleine Eickhoff, Nikola Kamburov, Juliana Kerrest, and Plamen Nenov.

Finally, I owe more than words could ever express to Ufuk Akcigit. As a fellow economist, he has been of invaluable guidance and has made endless work days downright enjoyable (dk). As my partner and best friend, he stood by me with his love and support through all these years.

# Contents

7

# List of Figures

# Introduction

This thesis contributes to the study of the optimal design of the tax system. A large theoretical literature has analyzed the tradeoff between efficiency and equity when setting taxes, while an abundant empirical literature has considered the estimation of behavioral responses to the tax system. Comparatively less attention has been given to the way taxation may affect the pretax wage distribution, or the structure of compensation. The focus has predominantly been on the response of taxable income to taxation and these reactions have typically been interpreted as labor supply adjustments or tax evasion. The typical assumption in taxation models has been that wages are exogenous because people are paid their marginal product, which is independent of taxes.

This assumption, while analytically convenient, is conceptually and empirically not very appealing. First, supposing that workers are indeed paid their marginal product, that marginal product itself need not be exogenous to taxes. Workers can invest in their own productivity throughout life through human capital acquisition. Any investment in human capital is naturally influenced by its net return, which depends on the tax system. Second, pay is not necessarily equal to marginal product. When firms are introduced as distinct agents, informational frictions between employers and employees can emerge. These frictions can lead to a divergence between pay and marginal product: firms set wages to balance the need to extract information and provide incentives to workers of different qualities. Because taxes differentially affect the returns to work of workers of different qualities, they influence the wage setting process of firms.

This thesis explores optimal taxation when wages are endogenous. I consider both of the aforementioned reasons: endogenous marginal products and a more complex endogenous wage

setting process. Taking into account the dependence of wages on the tax system leads to new and more complex implications for optimal tax design.

In Chapter 1, I introduce firms as important active agents in the optimal income tax problem. Firms do not know workers' abilities and competitively screen them through nonlinear compensation contracts unobservable to the government. Because wages and hours are used as screening tools by firms, they are interdependent across workers, and private labor contracts are endogenous to the tax system. The government must anticipate that the compensation structure itself, i.e., the pre-tax wage distribution, will emerge endogenously to taxes.

In Chapter 2, wages are endogenous because of investments in human capital. Agents can improve their productivity throughout their lives by acquiring human capital through either monetary expenses or time. Income taxation captures part of the net return to human capital investments, and hence discourages them, but also provides insurance against risky returns to human capital, which can encourage investments. Capital taxes affect the tradeoff between physical and human capital investments as alternative ways of transferring resources intertemporally.

In Chapter 3, wages are again endogenous because of human capital investments, but the intergenerational feature of human capital investments is acknowledged. It is now parents who can choose between purchasing education for their children or leaving them financial bequests. Hence, human capital investments and the wage are jointly influenced by the income tax, the bequest tax, and education subsidies.

More precisely, chapter 1 studies optimal income taxation in a setting in which firms cannot directly observe workers' talents. Unlike in the traditional Mirrlees (1971) frictionless labor market – in which firms pay workers a wage equal to their marginal product – firms must set nonlinear compensation contracts to screen high ability workers from low ability ones. The government does not observe those potentially complicated private labor market contracts, but only total income earned. When it sets income taxes, it must anticipate that the private market contracts are nested in and interacting with the government's taxation contract.

The labor market is modeled according to a Miyazaki-Wilson-Spence (hereafter, MWS) equilibrium, which is always constrained efficient, thus *a priori* minimizing the scope for government intervention. I derive new optimal linear tax formulas for a general discrete types

model and characterize the full Pareto frontiers with nonlinear taxation.

A surprising result is that, when the government has sufficiently strong redistributive goals, welfare is higher when there is adverse selection than when there is not. A sufficient condition on social preferences is that lower types are weighted cumulatively more than their cumulative proportions in the population. This result is due to the "rat race" in which high productivity workers are caught, created by firms to screen them. The use of work hours and pay as screening tools limits the scope of high types to react negatively to distortive taxation, and helps the government redistribute at lower efficiency cost. This is true for both linear and nonlinear taxation, and even when the government cannot observe the private labor contracts.

Second, the taxable income elasticity, which has been the focus of much of the empirical literature, is no longer a sufficient statistic for optimal taxes. Indeed, because of the labor distortions caused by adverse selection, there are first-order welfare effects of taxes through labor supply, which appear as two new types of terms in the optimal tax formula. The "rat race" terms measure the cost of labor supply distortions on each type's welfare, and provide a rationale for a corrective Pigouvian tax, even absent redistributive concerns. In addition, firms are already engaged in some redistribution between types through cross-subsidization between different contracts offered: this is captured in the "informational rent" terms of lower types. If the government's redistributive preferences are sufficiently strong, then at a given elasticity of taxable income and a given income distribution, as observed in the data, the optimal tax will be higher when there is adverse selection.

The two main policy implications of these findings are that, first, the interpretation, estimation, and use of taxable income elasticities are complicated by the presence of adverse selection. For instance, it is no longer straightforward to map measured elasticities into structural elasticities without knowledge of the underlying market structure. Second, a government with highly redistributive preferences might find some degree of adverse selection useful. The information structure of the economy, and by consequence, the optimal tax and welfare, can be endogenous to some widely used government policies, such as bans on discrimination, or regulations on firing and pay structures.

Chapter 2 considers optimal taxation in the presence of endogenous human capital investments, which leads to an endogenous pre-tax wage distribution. Human capital accumulation

is a key concern for many people that starts early in childhood, continues through the formal education system, and extends throughout working life in the form of job training. Investments in human capital can be shaped by tax policy. Labor income taxes discourage investment by reducing the net return to human capital, but also encourage investment by insuring against risky human capital returns. Capital taxes affect the choice between physical and human capital.

This chapter jointly determines optimal taxes and human capital policies over the life cycle. In the model, each individual's wage is a function of endogenous human capital and stochastic "ability," a comprehensive measure of the exogenous component of productivity as in the Mirrlees (1971) model. Agents' heterogeneous innate abilities are subject to persistent and privately uninsurable shocks. This leads to heterogeneous and uncertain returns to human capital. Throughout their lives, agents can invest in risky human capital by spending money (e.g., tuition or books) or time (e.g., studying or training). The government attempts to provide insurance against idiosyncratic uncertainty, as well as to redistribute across intrinsic heterogeneity. However, it faces asymmetric information about initial ability and its evolution, as well as labor effort. This necessitates the imposition of incentive compatibility constraints in the dynamic mechanism designed by the government. To describe the distortions in the resulting constrained efficient allocations, the wedges, or implicit taxes and subsidies, are analyzed.

This chapter highlights that, because of the endogenous wage, the income tax and savings tax distortions should optimally be counterbalanced by a subsidy for human capital expenses. In addition, the endogeneity of wages makes a subsidy to human capital attractive as a way to indirectly stimulate labor supply by increasing the returns to labor. Finally, the human capital subsidy can itself fulfill a redistributive role traditionally assigned to the income tax, by exploiting the differential effect of human capital on the pre-tax incomes of high and low ability people. If the wage elasticity with respect to ability is decreasing in human capital, human capital has a positive redistributive effect on after-tax income and a positive insurance value. It is then optimal to subsidize human capital expenses beyond simply insuring a neutral tax system with respect to human capital expenses, i.e., beyond making human capital expenses fully tax deductible in a dynamic, risk-adjusted fashion.

The chapter illustrates these conceptual points by calibrating a stylized model with parameters roughly corresponding to the current-day U.S. economy. The results from this modeling

13

exercise suggest that full dynamic risk-adjusted deductibility may be close to optimal and that simple linear age-dependent human capital subsidies, as well as income and savings taxes, achieve almost the entire welfare gain from the full second-best optimum for the calibrations studied.

Two ways of implementing the constrained efficient allocations are proposed: Income Contingent Loans (ICLs), the repayments of which depend on the history of earnings and human capital, and a "Deferred Deductibility" scheme in which only part of current investment in human capital can be deducted from current taxable income, while the rest is deducted from future taxable incomes.

Chapter 3 continues the investigation of endogenous human capital, but turns to an intergenerational setting. Parents are now the ones investing in the human capital of their children and can alternatively choose to transfer resources to their offspring in the form of financial bequests. Hence, children's future wages – determined by the investments in human capital – become endogenous to income and bequest taxes, as well as education subsidies. Indeed, income taxes reduced the net returns to education for children, but also redistribute resources towards low income parents who can then invest in education. Bequest taxes affect the decision to purchase education or leave bequests. In turn, bequests affect the incentives of children to work and, hence, the revenues from income taxes.

This chapter jointly determines the optimal linear income tax, bequest tax, and education subsidy in a dynamic intergenerational model. As in the standard dynastic Barro-Becker model, each generation cares about the expected discounted utility of all future generations. The future wage of each child is a function of his parents' endogenous education investments and a stochastic exogenous component.

I express the optimal formulas for education subsidies, income taxes and bequest taxes in terms of estimable statistics that are robust to heterogeneity in preferences and primitives. A crucial determinant of education subsidies and bequest taxes are their distributional incidences, i.e., how concentrated education expenses and bequests are among high marginal utility agents. I compare the results to the benchmark result by Bovenberg and Jacobs (2005) which states that income taxes and education subsidies are "Siamese Twins," i.e., they should be set equal to each other. In a static model, this corresponds to full deductibility of education expenses.

14

This no longer holds in a model with generalized preferences, heterogeneity, and uncertainty. Even more, because of the unrestricted wage and utility functions, it is not even the case that the education subsidy and the income tax should comove. I show that the optimal formulas can equivalently be derived in terms of reform-specific elasticities that capture the same trade-off between the efficiency and distributional impact of each instrument, but that can be targeted to already conducted reforms for which elasticity estimates are easier to obtain.

This thesis tries to cast light on the income tax problem with endogenous wages. The main contribution of this work is to describe how optimal income tax results need to be nuanced and modified when taxes also shape the pre-tax wage distribution. This extension to the workhorse taxation model can improve the applicability of optimal tax theory to policy.

# Chapter 1

# Optimal Income Taxation with Adverse Selection in the Labor Market

## 1.1 Introduction

For many workers, the labor market may resemble a rat race, in which they have to compete for high-paying jobs by always working harder. Indeed, if talent and ability are difficult to recognize, hard work may be the only way for employees to favorably influence the perceptions of their employers and, hence, their pay. Understanding the informational structure of the labor market, and the mechanism through which hours of work and pay are set, is crucial for many policy questions. One of them is optimal income taxation, since labor supply is a key margin on which individuals may respond to taxation. What is the optimal income tax in a setting in which firms cannot directly observe workers' talents, but instead set nonlinear compensation contracts to screen high ability from low ability ones? In this chapter, I attempt to answer this question by studying optimal linear and nonlinear income taxes with adverse selection in the labor market.

The standard income taxation model, introduced in Mirrlees' (1971) seminal paper, assumes a frictionless labor market in which firms pay workers a wage equal to their ability, i.e., their

marginal product per hour. The government, on the other hand, tries to redistribute from high to low ability workers, but does not observe abilities. It hence sets nonlinear taxes subject to incentive compatibility constraints to ensure that workers truthfully reveal their types. By contrast, in the current chapter, firms do not know workers' abilities and play an active role in determining hours of work and pay. When the government sets taxes, it must take into account the modified responses to them, due to the nonlinear, screening wage schedules facing workers. Private market contracts are nested in and interacting with the government's contract. As an added challenge, the government does not observe those potentially complicated private labor market contracts, but only total income earned. Accordingly, it must not only anticipate which contracts workers will choose out of a fixed set, but also the set of labor contracts, that is the compensation structure itself, which will emerge endogenously to taxes.[1]

To explain the functioning of the labor market, I use a Miyazaki-Wilson-Spence (hereafter, MWS) equilibrium (Spence, 1978, Wilson, 1977 and Miyazaki, 1977), which is always constrained efficient, thus *a priori* minimizing the scope for government intervention. I also discuss the Rothschild-Stiglitz (hereafter, RS) equilibrium notion (Rothschild and Stiglitz, 1976), which has its own peculiar challenges of potential non-existence and constrained inefficiency in the Appendix. I derive new optimal linear tax formulas for a general discrete types model and characterize the full Pareto frontiers with nonlinear taxation.

The most surprising result is that, when the government has sufficiently strong redistributive goals, welfare is higher when there is adverse selection than when there is not. This result is due to the "rat race" in which high productivity workers are caught, which is engineered by firms to separate them from lower productivity ones. The use of work hours and pay as screening tools limits the flexibility of high types to react adversely to distortive taxation, and helps the government redistribute.

Second, since the usual envelope conditions on labor supply no longer hold, there are first-order welfare effects from affecting it through taxes, and the optimal linear tax formula is modified to include two new types of terms. The corrective "rat race" terms capture the cost of labor supply distortions on each type's welfare, and can make the optimal tax positive

---

[1]In the standard model, the pretax *income* distribution of the economy is endogenous to taxes, but the endogeneity is driven solely by hours worked, while wages are equal to the intrinsic productivities of workers.

even absent any redistributive agenda, akin to a Pigouvian tax. In addition, firms are already performing some redistribution themselves by cross-subsidizing workers, which is captured in the "informational rent" terms of lower types. For a given elasticity of taxable income and a given income distribution, the optimal tax will be higher when there is adverse selection, provided the redistributive preferences of the government are sufficiently strong.[2]

Third, in the nonlinear tax case, I compare the Pareto frontiers under three different informational regimes: the standard Mirrlees, the Second Best with Adverse Selection – in which neither the government nor firms know workers' types, but the government observes private labor contracts – and the Adverse Selection with unobservable private contracts. The main result carries over: whenever the government wants to redistribute from high to low ability workers, the Pareto frontiers with Adverse Selection – with either observable or unobservable private contracts – are strictly above the Mirrlees frontier. A sufficient condition on social preferences is that lower types are weighted cumulatively more than their cumulative proportions in the population. When private contracts are unobservable to the government, it can still implement any Second Best allocation with observable contracts using a mix of nonlinear income taxes levied on workers, and nonlinear payroll taxes levied on firms.

I discuss the two main policy implications of these findings and draw the link to tax praxis. First, I outline how the interpretation, estimation, and use of taxable income elasticities is complicated by the presence of adverse selection – an important cautionary tale given how central the latter are in the taxation literature. In particular, it is no longer straightforward to map measured elasticities into structural elasticities without knowledge of the underlying market structure. Estimation relying on reforms as natural experiments may be affected by the interconnections of different groups through their labor contracts. Even correctly estimated, these elasticities are no longer sufficient statistics for the deadweight loss of taxation, and strict reliance on them for optimal tax design may be misleading. Secondly, the result that welfare may be higher with adverse selection suggests that a government with highly redistributive preferences might find some degree of adverse selection useful, and naturally leads to question to what extent the information structure of the economy is endogenous to government policies.

---

[2]In addition, in the RS setting, raising taxes can destroy an existing equilibrium, and hence tax policy is more constrained.

18

Some widely used labor market interventions, such as bans on discrimination, or regulations on firing and pay structures can affect the degree of adverse selection, and, by consequence, the optimal tax and welfare.

**Empirical Literature on adverse selection:** All results in this chapter are based on two empirically testable assumptions. First, there must be asymmetric information about worker quality between firms and workers, a friction that has been widely documented. Acemoglu and Pischke (1998) show that a worker's current employer has more information about his quality than other potential employers, suggesting that, at the time of hiring, quality is uncertain. Gibbons and Katz (1991) also test a model in which the incumbent employer has superior information, so that laid-off workers are perceived as lower ability.

The second assumption is that firms are screening their workers through the labor contracts offered, rather than through other direct means, such as ability tests. Although I focus on requirements on the hours of work, other productive actions which are costlier to lower ability workers, such as sophisticated training programs, or effort on specific tasks, could also serve as valid screening tools. Evidence that employers screen indirectly through training comes from Autor (2001) for Temporary Help Firms. Career concerns seem to make workers work harder in order to positively influence the perception of their employers about their talent (Holmstrom, 1978, Gibbons and Murphy, 1992, and Baker, Gibbons and Murphy, 1994). Most closely related to this chapter is the empirical study of "rat races" at large law firms by Landers, Rebitzer and Taylor (1996), who show that employees are required to work inefficiently long hours before being promoted to partners in order to distinguish those with a high propensity to work.

**Related optimal taxation literature:** This chapter contributes to the optimal taxation literature (as developed by Mirrlees, 1971, Diamond, 1998, Saez, 2001, Albanesi and Sleet, 2006, Golosov, Tsyvinski and Werning, 2006, and Weinzierl, 2011 among others), but mostly to a growing strand of it which considers the interplay of private markets and government-imposed taxation. The focus until now has generally been on private credit and insurance markets rather than on informational problems in the labor market itself. Golosov and Tsyvinski (2006) study optimal dynamic taxation when agents can secretly trade risk-free bonds, while Krueger and Perri (2010) examine the role of progressive income taxation in insuring agents when private risk sharing is imperfect. But unlike their private market equilibria, the one in this chapter

is already constrained efficient.[3] Chetty and Saez (2010) highlight that the fiscal externality generated by private sector insurance that suffers from moral hazard or adverse selection needs to be taken into account in the optimal tax formulas. Unlike them, I focus on the labor supply contract, and deal more explicitly with the private market equilibrium. Scheuer (2013a,b) considers optimal income and profit taxes with incomplete credit markets for entrepreneurs. The link to the literature on contracts, imperfect information, and hidden trades is drawn in Section 1.4.1.

The rest of the chapter is organized as follows. The next section describes the labor market, and solves for the optimal tax in the standard case with no adverse selection. Section 1.3 studies the optimal linear taxation problem with adverse selection, while Section 1.4 focuses on the optimal nonlinear tax. Section 1.5 discusses the policy implications, and Section 1.6 concludes. All proofs are in the Appendix.

## 1.2 A Model of the Labor Market with Adverse Selection

### 1.2.1 The Labor Market

Consider a perfectly competitive labor market with workers of $N$ different productivities, hired by risk-neutral competitive firms.[4] Type $i$ has productivity $\theta_i \in \Theta = \{\theta_1, ..., \theta_N\}$, with $\theta_1 < ... < \theta_i < ... < \theta_N$, and produces $f(h) = \theta_i h$ units of output for $h$ hours of work at a disutility cost of $\phi_i(h)$. The fraction of types $i$ in the population is $\lambda_i$ with $\sum_i \lambda_i = 1$. The assumptions on the cost functions required to permit screening are analogous to the ones in Spence (1978):

**Assumption 1** *i)* $\phi_i'(h) > 0$, $\phi_i''(h) > 0$ $\forall h > 0$, *and* $\phi_i(0) = \phi_i'(0) = 0$, $\forall i$

**Assumption 1** *ii)* $\phi_i(h) < \phi_{i-1}(h)$ $\forall h > 0$, $\forall i > 1$

   *iii)* $\phi_i'(h) < \phi_{i-1}'(h)$ $\forall h > 0, \forall i > 1$.

---

[3] The potentially constrained inefficient Rothschild-Stliglitz equilibrium is in the Appendix.

[4] The N-type model was introduced and solved by Spence (1978) in the context of insurance policies. I adapt it to the labor market and introduce taxes into the model.

Hence, lower productivity workers not only have a higher cost of effort, but also a higher marginal cost. The utility of a worker of type $i$ takes a simple quasilinear form:

$$U_i(c, h) = c - \phi_i(h)$$

where $c$ is net consumption, equal to total pay $y$ minus any taxes $T(y)$ paid to the government.

Firms cannot observe a worker's type, but can perfectly monitor hours of work. They hence post screening contracts specifying pairs of pay and hours $\{y_i, h_i\}_{i=1}^N$. There exist several equilibrium concepts for such hidden information settings, but no consensus about the best one. In this chapter, I focus mostly on an analytically tractable Miyazaki-Wilson-Spence foresight equilibrium (MWS), (Spence, 1977, Wilson, 1976, and Miyazaki, 1977). The Appendix contains the analysis of a Nash behavior à la Rothschild and Stiglitz (1976).

**Definition 1** *(Miyazaki-Wilson-Spence equilibrium) A set of contracts is an equilibrium if i) firms make zero profits on their overall portfolio of contracts offered, and ii) there is no other potential contract which would make positive profits, if offered, after all contracts rendered unprofitable by its introduction have been withdrawn.*

In the MWS setting, each firm is only required to break even overall on its portfolio of contracts, allowing for cross-subsidization between contracts. Firms have foresight: they anticipate that if they offer a new contract, some existing contracts might become unprofitable and be withdrawn. An equilibrium always exists and is constrained efficient (Miyazaki, 1977), thus reducing the scope for government intervention.[5]

### 1.2.2 The Optimal Linear Tax Without Adverse Selection

Here and in Section 1.3, I suppose that the only two instruments available for redistribution are a linear income tax $t$, levied on total earned income $y$, and a lump-sum transfer $T$, which

---

[5]In a two types model ($N = 2$) if $\lambda_1$, the fraction of low types, is small, then the equilibrium involves cross-subsidization from high productivity to low productivity workers. High productivity workers are paid less than their product and low types are paid more than theirs. If $\lambda_1$ is sufficiently high, the MWS and RS equilibrium allocations coincide.

ensures budget balance.[6] As a benchmark, it is useful to solve for the standard Second Best Pareto frontier, in the case without adverse selection. A weighted sum of utilities is maximized, subject to the reaction functions of the private market. For any given tax, workers of type $i$ choose a level of hours $h_i^*(t)$, referred to as the efficient level of hours for type $i$,[7] at which the marginal cost of effort just equals the net of tax return:

$$\phi_i'(h_i^*(t)) = \theta_i(1-t) \tag{1.1}$$

Earnings are $y_i(t) = \theta_i h_i^*(t)$. For a set of Pareto weights $\boldsymbol{\mu} \equiv \{\mu_i\}_{i=1}^N$, the social welfare function is:

$$SWF(\boldsymbol{\mu}) = \sum_{i=1}^N \mu_i(c_i(t) - \phi_i(h_i(t))) \tag{1.2}$$

Using that $c_i = y_i(1-t) + T$, the government's program is:

$$\left(P^{SB,N}(\boldsymbol{\mu})\right) : \max_t \left\{ \sum_{i=1}^N \mu_i(\theta_i h_i^*(t)(1-t) - \phi_i(h_i^*(t)) + T) \right\}$$

with

$$T = t \sum_{i=1}^N \lambda_i \theta_i h_i^*(t)$$

where $\{h_i^*(t)\}_i$ are the workers' reaction functions to taxes as defined in (1.1).

It is instructive to derive the optimal tax formula heuristically, using a perturbation argument as in Saez (2001). When the tax rate is raised by a marginal amount $dt$, there are three effects. The mechanical revenue effect, $dM$ – the change in tax revenue if there were no behavioral responses – is simply equal to average income, denoted $y(t) \equiv \sum_i \lambda_i y_i(t)$:

$$dM = y(t)\,dt$$

The behavioral effect, $dB$, caused by changes in agents' labor supply, is:

$$dB = t \left( \sum_i \lambda_i \theta_i \frac{dh_i^*(t)}{dt} \right) dt$$

which, after some algebraic manipulations, can be rewritten as:

$$dB = -\frac{t}{1-t} \varepsilon_y y dt$$

where $\varepsilon_y$ is the usual aggregate elasticity of taxable income to the retention rate $(1-t)$, also equal to the income-share-weighted average of individual elasticities:

$$\varepsilon_y \equiv d \log y / d \log(1-t) = \sum_i \alpha_i(t) \varepsilon_{y_i}$$

where

$$\varepsilon_{y_i} \equiv d \log(y_i) / d \log(1-t)$$

is type $i$'s taxable income elasticity, and $\alpha_i(t) \equiv \lambda_i y_i(t)/y$ is the share of total income produced by type $i$ workers.

Finally, the welfare effect $dW$ – sum of the individual welfare effects $dW_i$ – is equal to the Pareto weights weighted reduction in consumption, since the indirect effect on welfare through changes in hours of work is zero by the envelope theorem:

$$dW = \sum_i dW_i = - \left( \sum_i \mu_i y_i(t) \right) dt$$

Denote the Pareto-weights weighted income shares by $\bar{y} \equiv \sum_i \mu_i y_i(t)/y$. $\bar{y}$ measures the concentration of income relative to redistributive preferences. Whenever the social welfare function puts the same weight on each type as his proportion in the population ($\lambda_i = \mu_i, \forall i$), $\bar{y} = 1$. If Pareto weights are concentrated mostly on those with low incomes, then $\bar{y} \ll 1$. Hence,

$$dW = -(y\bar{y}) dt$$

The optimal tax is the one at which the sum of these three effects $dM + dB + dW$ is zero,

which yields the familiar (implicit) tax formula:[8]

$$\frac{t^{SB}}{1 - t^{SB}} = \frac{1 - \bar{y}}{\varepsilon_y} \tag{1.3}$$

This formula highlights the two usual forces determining the optimal tax, namely, the equity concern, proxied by the income distribution and Pareto weights in $\bar{y}$, and the efficiency concern, captured by the taxable income elasticity $\varepsilon_y$. Note that the revenue-maximizing tax rate is $t^R / \left(1 - t^R\right) = 1/\varepsilon_y$, while the Rawlsian tax rate (when $\mu_1 = 1$) is $t^{Rawls} / \left(1 - t^{Rawls}\right) = \frac{1 - \alpha_1}{\varepsilon_y}$. If $\mu_i = \lambda_i$, for all $i$, the utilitarian criterion, combined with quasilinear utility yields $t^{SB} = 0$.

## 1.3 Linear Taxes with Adverse Selection

Suppose now that firms do not know workers' types. An adverse selection problem arises if, at the first best allocation, lower type workers would like to pretend they are higher types, i.e., if and only if:

$$\theta_{i+1} h_{i+1}^* (0) - \phi_i \left(h_{i+1}^* (0)\right) > \theta_i h_i^* (0) - \phi_i \left(h_i^* (0)\right) \quad \forall i \leq N - 1 \tag{1.4}$$

where $h_i^* (0)$ is as defined in (1.1) at $t = 0$. I assume that (1.4) holds throughout this Section.

A two-stage game takes place, with the government first setting taxes $t$, and the corresponding transfer $T$, and firms then choosing what labor contracts $\{h_i (t), y_i (t)\}_{i=1}^N$ to offer (see figure $1 - 1$ below). Working backwards from the second stage, I study the reaction functions of firms to any given tax, and then solve the government's optimal tax program, taking the responses of the labor market as additional constraints. To build the intuition, I start with $N = 2$.

---

[8]With $N = 2$, this can also be rewritten as:

$$\frac{t^{SB}}{1 - t^{SB}} = \frac{(\mu - \lambda) \left(\frac{\alpha_2(t)}{1 - \lambda} - \frac{a_1(t)}{\lambda}\right)}{\varepsilon_y}$$

Note that $\alpha_2 (t) / (1 - \lambda)$ and $a_1 (t) / \lambda$ are the shares of total income per *worker* of each type respectively. The greater this difference, and the greater inequality. Whenever the low type workers are valued more at the margin than is justified by their population share ($\mu > \lambda$), the optimal tax is positive.

Figure 1-1: Timeline



| Government announces t and T. | Firms post menu of contracts $\{h_i(t), y_i(t)\}$. | Workers choose labor contract from menu. | Type $i$ works $h_i(t)$ hours, receives $y_i(t)$, pays $t y_i(t)$ and gets T. |
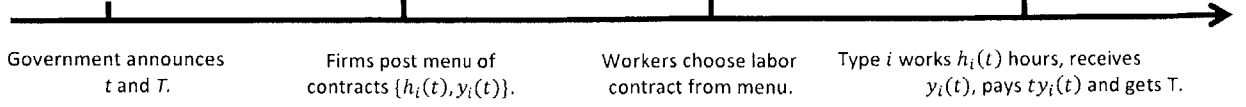
### 1.3.1 Second Stage: The Private Sector's Reaction to Taxes

**Two types model $N = 2$ :**

In a MWS setting with $N = 2$, as shown by Miyazaki (1977) for the case without taxes, firms offer a menu of contracts $\{h_i, y_i\}_{i=1}^2$ solving program $P^{MWS}(t)$, conditional on a given tax level $t$ and a transfer $T$. The transfer $T$ will not affect the firm's problem thanks to quasilinear utility and is omitted. Let $\lambda_1 = \lambda$ and $\lambda_2 = 1 - \lambda$.

$$\left(P^{MWS}(t)\right) : \max_{\{y_1, y_2, h_1, h_2\}} (1 - t) y_2 - \phi_2(h_2)$$

$$(IC_{12}) : (1 - t) y_1 - \phi_1(h_1) \geq (1 - t) y_2 - \phi_1(h_2)$$

$$(IC_{21}) : (1 - t) y_2 - \phi_2(h_2) \geq (1 - t) y_1 - \phi_2(h_1)$$

$$(profit) : \lambda y_1 + (1 - \lambda) y_2 = \lambda \theta_1 h_1 + (1 - \lambda) \theta_2 h_2$$

$$(RS_1) : (1 - t) y_1 - \phi_1(h_1) \geq (1 - t) y_1^{RS} - \phi_1\left(h_1^{RS}\right) \equiv u_1^{RS}$$

The first two constraints are the incentive compatibility constraints for the low and high type respectively, ensuring that each type self-selects into the appropriate work contract. The third one is the zero profit condition on the full portfolio of contracts. The final constraint ensures that the low productivity worker always receives at least his utility from the separating allocation,[9] defined by $h_1^{RS}(t) = h_1^*(t)$ and $y_1^{RS}(t) = \theta_1 h_1^{RS}(t)$. Note that for some $t$, $(IC_{12})$ could become slack with hours at their efficient levels, even if it would be binding at $t = 0$ (when (1.4) holds).

---

[9] This separating allocation is also known as the "Rothschild-Stiglitz" allocation. Why this constraint appears in the program is explained in detail in the original Miyazaki (1977) paper for the case without income taxes. In short, if this constraint was not satisfied, there would be a profitable deviation for some firm, consisting in offering a slightly worse contract than the fully separating one, attracting all low types, and making a positive profit.

Then we would have:

$$(1 - t) \theta_2 h_2^* (t) - \phi_1 (h_2^* (t)) < (1 - t) \theta_1 h_1^* (t) - \phi_1 (h_1^* (t)) \qquad (1.5)$$

The following proposition characterizes the private market equilibrium for any tax $t$ as a function of a threshold $\tilde{\lambda} (t)$ for the fraction of low types (defined in the Appendix).

**Proposition 1** *For a given $t$, constraint $(profit)$ is binding and $(IC_{21})$ is slack. The low type always works an efficient amount of hours $h_1^* (t)$, and there are three possible equilibrium configurations:*

*i) If $(1.5)$ holds, $h_2 (t) = h_2^* (t)$, and the allocation is equal to the Second Best one.*

*ii) If $(1.5)$ does not hold and $\lambda > \tilde{\lambda} (t)$ (called case AS1), constraint $(RS_1)$ is binding, each worker earns his marginal product, and there is full separation. $h_2 (t)$ is above the efficient level, and is the solution to $\theta_1 h_1^* (t) (1 - t) = \theta_2 h_2 (t) (1 - t) - (\phi_1 (h_2 (t)) - \phi_1 (h_1^* (t)))$.*

*iii) If $(1.5)$ does not hold and $\lambda \leq \tilde{\lambda} (t)$ (called case AS2), constraint $(RS_1)$ is not binding and there is cross-subsidization from high to low productivity workers. $h_2 (t)$ is above the efficient level, and is the solution to $\phi_2' (h_2 (t)) = (1 - t) (1 - \lambda) \theta_2 + \lambda \phi_1' (h_2 (t))$.*

*In addition,*

*iv) $dh_i (t) / dt < 0$, for $i = 1, 2$*

*v) $dh_2 (t) / d\lambda > 0$ for $\lambda \leq \tilde{\lambda} (t)$, $dh_2 (t) / d\lambda = 0$ for $\lambda > \tilde{\lambda} (t)$.*

The first case occurs if the low type no longer wants to pretend to be a high type at $t$ and is not of great interest: it is unlikely to occur with $N > 2$, and can be ruled out by assumption 2 below. More generally, low productivity workers work an efficient number of hours, but high productivity workers work excessively. There is a critical level of the fraction of low types, $\tilde{\lambda} (t)$, which determines whether firms find it profitable to cross-subsidize workers or not. The intuition is that, for a low $\lambda$, it is beneficial to reduce the distortion in the labor supply of high types in exchange for a higher cross-subsidy to low types. When the fraction of low types increases, however, this subsidy to each of them becomes too costly, and there is separation.[10]

---

[10]But each worker earns his product at the equilibrium levels of hours only. It is not the case that a worker would earn his marginal product had he chosen another level of hours, unlike in standard competitive labor markets.

Hours of work of the high type are increasing in $\lambda$ for $\lambda \leq \tilde{\lambda}(t)$ because of the standard trade-off in screening models between the distortion imposed on the high type and the informational rent forfeited to the low type (see Laffont and Martimort, 2001). The higher the fraction of bad types, the costlier it becomes for the firm to give up an informational rent to each of them. As a result, the hours of the high type must be distorted more.

An additional assumption, namely that the disutility of labor is isoelastic, simplifies the exposition but is not needed for the derivation of the optimal tax in Subsection 1.3.2.

**Assumption 2** $\phi_i(h) = a_i h^\eta$, *for $i = 1, 2$.*

To satisfy assumption 1 with this specification would require $a_1 > a_2$.

**Proposition 2** *If assumption 2 holds:*

*i) $IC_{12}$ binds at all $t$,*

*ii) $\tilde{\lambda}(t)$ is independent of $t$: $\tilde{\lambda}(t) = \tilde{\lambda}$, $\forall t$.*

Result i) states that, if at $t = 0$ there is an adverse selection problem (i.e., (1.4) holds) and assumption 2 holds, then there is an adverse selection problem at all tax levels. Then, the marginal utility of the low type from his own efficient allocation and from deviating to the high type's efficient allocation grow at the same rate, and the relative rewards from cheating versus revealing truthfully are unaffected by the tax. The second result guarantees that the type of equilibrium does not depend on the tax rate. $\tilde{\lambda}(t)$ is equal to the ratio of the marginal welfare loss of the high type $\theta_2(1 - t) - \phi_2'(h_2)$ and the marginal informational rent gain of the low type, $\theta_1(1 - t) - \phi_1'(h_1)$. As long as the cost of distortion in $h_2$ remains low relative to the informational rent ($\lambda > \tilde{\lambda}(t)$), the contract is separating. If it grows too high ($\lambda < \tilde{\lambda}(t)$), it becomes better to grant the low type a cross-subsidy rather than to keep distorting hours of work. With isoelastic disutility functions these two effects grow at the same rate with the tax, so that their ratio is independent of $t$.


**N types model, for $N \geq 2$:**

For any $t$ set by the government, define a sequence of programs $\left(P_i^{MWS}(t)\right)_{i=1}^N$ and utilities $\bar{u}_i$ such that:

$$\left(P_1^{MWS}(t)\right) : \bar{u}_1 = \max_h \theta_1 h(1 - t) - \phi_1(h)$$

For $2 \leq i \leq N$:

$$\left(P_i^{MWS}(t)\right) : \bar{u}_i = \max_{\{y_j, h_j\}_{j=1}^i} y_i(1-t) - \phi_i(h_i)$$

subject to:

$$
\begin{aligned}
y_j(1-t) - \phi_j(h_j) &\geq \bar{u}_j, \quad j < i \\
y_j(1-t) - \phi_j(h_j) &\geq y_{j+1}(1-t) - \phi_j(h_{j+1}), \quad j < i \\
\sum_{j=1}^i (\theta_j h_j - y_j)\lambda_i &= 0
\end{aligned}
$$

The equilibrium with $N$ types is the set of income and hour pairs $\{y_i, h_i\}_{i=1}^N$ which solve, for a given $t$, program $P^{MWS,N}(t) \equiv P_N^{MWS}(t)$.

**Proposition 3** *In the MWS equilibrium with $N$ types, $N \geq 2$ :*

*i) There is a number of "break agents" $k_1, k_2, ...k_n$ with $n \leq N$ such that:*

*- Firms make losses on all subsets of types of the form $\{1,..,i\}$, $\{k_1+1,...,i\}$,.., $\{k_{n-1}+1,...,i\}$ for $i \neq k_1, k_2, ...k_n$.*

*- Firms break even on the subsets of types of the form $\{1,..,k_1\}$, $\{k_1+1,...,k_2\}$,.., $\{k_{n-1}+1,...,k_n\}$, called "cross-subsidization groups."*

*- If $(IC_{j,j+1})$ is not binding for some $j$, types $j$ and $j+1$ are in two different cross-subsidization groups, called "disjoint."*

*ii) The lowest productivity agents of each disjoint cross-subsidization group (including type 1) work efficient hours.*

*iii) All other types work excessively much, i.e., $h_i(t) > h_i^*(t)$, $\forall t < 1$.*

The cross-subsidization groups are subsets of agents such that the firms breaks even on the group as a whole, but within which some types cross-subsidize others. With $N$ types, if $IC_{j,j+1}$ is not binding, then $j$ and $j+1$ are in different cross-subsidization groups (see the Appendix), and the population is split into (at least) two non-interacting sets, above and including $j+1$ and strictly below $j+1$.

## 1.3.2 First stage: The Optimal Linear Tax Problem with Two Types

In the first stage, the government chooses the optimal linear tax to maximize the weighted sum of individual utilities in (1.2), taking as given the reaction functions of the private market, $\{y_1(t), y_2(t), h_1(t), h_2(t)\}$. With $N = 2$ and $\mu_1 = \mu$, the program is:

$$\left(P^{AS}(\mu)\right) \quad : \quad \max \mu\left(y_1(t)(1-t) - \phi_1(h_1(t)) + T\right) + (1-\mu)\left(y_2(t)(1-t) - \phi_2(h_2(t)) + T\right)$$

$$s.t. \quad : \quad T = t\left(\lambda\theta_1 h_1(t) + (1-\lambda)\theta_2 h_2(t)\right)$$

The behavioral and mechanical revenue effects are still the same as in the Second Best (Subsection 1.2.2), but the welfare effects on the two types, $dW_1$ and $dW_2$, are now different. They can be decomposed into a *direct* effect from reduced consumption, and, if and only if the envelope condition does not hold, as is the case for $h_2$ here, additional *indirect* effects from changing labor supplies.

First, when taxes increase, the excessively high hours of work of the high type are reduced, which has a positive marginal effect on his own welfare, called the "rat race" effect and denoted by $\xi_2$:

$$\xi_2 \equiv (1-t)\theta_2 - \phi_2'(h_2) \leq 0$$

Second, the "informational rent" effect captures how the rent forfeited to induce the low type to reveal his true type changes with taxes, and is denoted by $\kappa_2$:

$$\kappa_2 \equiv \left[(1-t)\theta_2 - \phi_1'(h_2)\right] \leq 0$$

As is usual in screening models, there is a trade-off for the firm between reducing the informational rent of the low type and the distortion in the hours of the high type. As taxes increase, the high type is made to work less, which reduces the distortion in his labor supply, increases the rent transfer to the low type, and hence indirectly redistributes income. When firms cross-

subsidize workers, they redistribute from high to low types. Hence, the welfare effects are:

$$dW_1 \; = \; -\mu y_1 dt + \mu \left(1 - \lambda\right) I^c\left(t\right) \kappa_2 \frac{dh_2\left(t\right)}{dt} dt \tag{1.6}$$

$$dW_2 \; = \; -\left(1 - \mu\right) y_2 dt + \left(1 - \mu\right) \left\{\xi_2 - \lambda I^c\left(t\right) \kappa_2\right\} \frac{dh_2\left(t\right)}{dt} dt \tag{1.7}$$

where the indicator variable $I^c\left(t\right) = 1$ if there is cross-subsidization, and 0 otherwise. The informational rent effect only enters when there is cross-subsidization. Setting the sum $dM + dB + dW$ to zero, we obtain the optimal tax.[11]

**Proposition 4** *The optimal tax rate with adverse selection is:*

$$\frac{t^{AS}}{1 - t^{AS}} = \frac{\left(1 - \bar{y}\right)}{\varepsilon_y} + \frac{\frac{y_2/\theta_2}{y}\left(1 - \mu\right)\frac{\varepsilon_{y_2}}{1-t}}{\varepsilon_y}\left(-\xi_2\right) + \frac{\frac{y_2/\theta_2}{y} I^c\left(\mu - \lambda\right)\frac{\varepsilon_{y_2}}{1-t}}{\varepsilon_y}\left(-\kappa_2\right) \tag{1.8}$$

In general, $I^c\left(t\right)$, the elasticities, and incomes depend on $t$. The tax formula is thus as usual endogenous. Recall that assumption (2), however, makes the type of equilibrium, and hence $I^c$, independent of taxes $t$. It also guarantees that the Second Best never occurs (hence, $\xi_2 < 0$). Table 1 specializes formula (1.8) for the three possible cases.

| Case | Tax Rate Formula |
|---|---|
| Second Best (no adverse selection) | $\frac{t^{SB}}{1-t^{SB}} = \frac{1-\bar{y}}{\varepsilon_y}$ |
| Adverse Selection with full separation (case AS1) | $\frac{t^{AS1}}{1-t^{AS1}} = \frac{\left(1-\bar{y}\right)}{\varepsilon_y} - \frac{\frac{y_2/\theta_2}{y}\left(1-\mu\right)\frac{\varepsilon_{y_2}}{1-t^{AS1}}}{\varepsilon_y}\xi_2$ |
| Adverse Selection with cross-subsidization (case AS2) | $\frac{t^{AS2}}{1-t^{AS2}} = \frac{\left(1-\bar{y}\right)}{\varepsilon_y} - \frac{\frac{y_2/\theta_2}{y}\left(1-\mu\right)\frac{\varepsilon_{y_2}}{1-t^{AS2}}}{\varepsilon_y}\xi_2 + \frac{\frac{y_2/\theta_2}{y}\left(\lambda-\mu\right)\frac{\varepsilon_{y_2}}{1-t^{AS2}}}{\varepsilon_y}\kappa_2$ |

TABLE 1

---

[11]Note that $\bar{y}$, $\varepsilon_y$, and $\varepsilon_{y_2}$ are not the same *functions* of the tax as in the Second Best case.

The usual "sufficient statistics" $\bar{y}$ and $\varepsilon_y$ are no longer sufficient, as they do not capture the rat race and informational rent effects. The latter require knowledge of the underlying disutility of effort functions.

**Comparison of tax rates with and without adverse selection:** There are two complementary ways of comparing the optimal tax with and without adverse selection. The first one is to take the primitives of the model, i.e., the production and utility functions, as given; conceptually, this is akin to comparing tax rates in two economies which are exactly the same, except that one of them suffers from adverse selection while the other does not. The second possibility is to take as given the empirically measurable parameters, namely, the elasticities of taxable income and the distributional factors, and to compare the taxes which would be optimal if it was a market with adverse selection versus one without which generated them. This approach, adopted here, is more policy-relevant: it reflects the situation of a government equipped with widely available measures of elasticities and statistics about the income distribution, but unaware of the true market structure.

**Proposition 5** *At given $\varepsilon_y$ and $\bar{y}$:*

*i) In the separating equilibrium (case AS1): $t^{AS1} > t^{SB}$ and $t^{AS1} \to t^{SB}$ as $\mu \to 1$.*

*ii) With cross-subsidization (case AS2): If $\mu \geq \lambda$, $t^{AS2} > t^{SB}$ and $t^{AS2} > t^{AS1}$.*

At given $\bar{y}$ and $\varepsilon_y$, the corrective Pigouvian term in $\xi_2$ leads to a higher tax rate destined to reduce the distortion in hours of work of the high type, the more so when the government cares about the welfare loss of the high type ($\mu$ small). In addition, if there is cross-subsidization, a higher tax redistributes toward the lower type in two ways: directly through a higher transfer $T$, but also indirectly through the informational rent term $\kappa_2$. If the government puts a high weight on low type agents ($\mu \geq \lambda$), this pushes the tax up.

**Comparison of welfare with and without adverse selection:** To compare welfare, on the other hand, the individuals' utility functions are held constant, since welfare is measured relative to them.

**Proposition 6** *For the same economy: i) When the government has highly redistributive preferences ($\mu = 1$), welfare is higher when there is adverse selection in the labor market than when*

31

*firms can perfectly observe workers' types.*

*ii) When the government only cares about high type workers ($\mu = 0$), welfare is higher when there is no adverse selection in the labor market.*

*iii) The low type is always weakly better off when there is adverse selection.*

The counter-intuitive result in i) implies that the inability of firms to observe workers' productivities and their reliance on nonlinear compensation contracts for screening are not necessarily detrimental when the government wants to redistribute and can only use distortionary taxes. Like in traditional second-best theory, fixing a distortion in one place (here, adverse selection in the labor market) need not be good when there is another irremovable distortion (here, the absence of non-distortionary taxation for redistribution). If the government had lump-sum taxation available, or if it did not want to redistribute, adverse selection would only cause a deadweight loss. It is the *interaction* of the imperfect instruments available to fulfill strongly redistributive goals with adverse selection which improves welfare. This result becomes most relevant if the informational structure in this economy is endogenous, a point discussed in Section 1.5.

What are the sources of this welfare gain? First, the use of hours as a screening tool and the resulting rat race limit the ability of the high type to reduce his labor supply as a response to taxes. Hence, revenue is higher at any $t$, which is beneficial for the low type. Second, at a given $\mu$, with adverse selection and cross-subsidization – a form of redistribution done by firms – the optimal tax required to achieve the same level of redistribution could be lower, which is beneficial for both types.[12]

### 1.3.3   Optimal Linear Tax with N Types

With $N \geq 2$ types, the government maximizes social welfare as in (1.2), taking as given the private sector's reaction functions $\{h_i(t), y_i(t)\}_{i=1}^{N}$ derived in Proposition 3 and its proof. For

---

[12]Result i) will hold as long as $\mu \geq \tilde{\mu}$ for some threshold $\tilde{\mu}$, as explained in the general $N \geq 2$ case below.

any set of Pareto weights $\mu$, the program is:

$$\left(P^{AS,N}\left(\mu\right)\right) \quad : \quad \max_t \left\{ \sum_{i=1}^{N} \mu_i \left(y_i\left(t\right)\left(1-t\right) - \phi_i\left(h_i\left(t\right)\right) + T\right) \right\}$$

$$s.t. \quad : \quad T = t \sum_{i=1}^{N} \lambda_i y_i\left(t\right)$$

Let $I_{ij}\left(t\right)$ be the indicator function equal to 1 if $j$ is in $i$'s cross-subsidization group at tax $t$. Define $\bar{\lambda}^i$ (respectively, $\underline{\lambda}^i$) as the proportion of types strictly better (respectively, strictly worse) than $i$ in $i$'s cross-subsidization group:

$$\bar{\lambda}^i = \frac{\sum_{j>i} I_{ij}\lambda_j}{\sum_{j=1}^{N} I_{ij}\lambda_j}, \qquad \underline{\lambda}^i = \frac{\sum_{j<i} I_{ij}\lambda_j}{\sum_{j=1}^{N} I_{ij}\lambda_j}$$

Let $\bar{\mu}^i = \sum_{j>i} I_{ij}\mu_j$ (respectively, $\underline{\mu}^i = \sum_{j<i} I_{ij}\mu_j$) be the cumulative Pareto weights on types strictly better (respectively, strictly worse) than $i$ in $i$'s cross-subsidization group. As before, denote the rat race term of type $i$ by $\xi_i\left(t\right)$, and the informational rent forfeited by type $i$ by $\kappa_i\left(t\right)$:

$$\xi_i\left(t\right) \equiv \theta_i\left(1-t\right) - \phi_i'\left(h_i\right)$$

$$\kappa_i\left(t\right) \equiv \theta_i\left(1-t\right) - \phi_{i-1}'\left(h_i\right)$$

**Proposition 7** *The optimal tax for any $N \geq 2$ is:*[13]

$$\frac{t^{AS}}{1-t^{AS}} = \frac{1-\bar{y}+\Delta^{AS}}{\varepsilon_y} \tag{1.9}$$

*with*

$$\Delta^{AS} = \frac{1}{y} \sum_{j=1}^{N} \left[ \left(\bar{\mu}^j + \mu_j - \bar{\lambda}^j\right)\left(-\xi_j\right) + \left(\underline{\mu}^j - \underline{\lambda}^j\right)\left(-\kappa_j\right) \right] \varepsilon_{y_j} \frac{y_j/\theta_j}{1-t^{AS}} \tag{1.10}$$

$\varepsilon_{y_j}$, $\varepsilon_y$, $y$, and $\bar{y}$ are as defined in Subsection 1.2.2. The formula in (1.9) highlights the same basic effects that were at play in the two types case, but allows for all possible equilibria configurations that can endogenously occur. Each type $j$ (except type 1 and those at the bottom

---

[13] *To reduce notational clutter, most dependences on the tax rate $t$ are left implicit.*

33

of each disjoint cross-subsidization group) now has other types below him with binding incentive constraints, leading to an upward distortion in his hours of work, and a rat race term $\xi_j$, which tends to push the tax rate up whenever $\bar{\lambda}^j < \mu_j + \bar{\mu}^j$. Put differently, if the government cares sufficiently about the welfare of types higher than $j$, it will raise the tax to correct for their excessive work.

Similarly, each type (except $N$ and the highest types of each disjoint cross-subsidization group) now receives an informational rent, $\kappa_j$. This will tend to increase the tax if the cumulative Pareto weights exceed the fractions in the population, i.e., $\underline{\mu}^j > \underline{\lambda}^j$, which is the analog of the condition $\mu > \lambda$ for two types. The intuition lies again in the trade-off between the informational rent earned by $j$ (and, hence, all lower types) and the distortion imposed on all higher types. When the government disproportionately cares about $j$ and lower types, it wants to raise the tax, reducing the hours distortions above $j$, and simultaneously increasing the informational rent to $j$ and below.

In the limit, an elitist government with $\mu_N = 1$ mostly cares about the rat race terms of high types, while trying to minimize informational rents transferred to low types. At the other extreme, a Rawlsian government with $\mu_1 = 1$ would mostly focus on increasing the transfer and informational rents to low types. If all agents are in the same cross-subsidization group, and the population weights are equal to the Pareto weights, then all redistributive concerns drop out, and only the corrective terms for the rat race remain, yielding a positive Pigouvian tax:

$$\frac{t^{Pigou}}{1 - t^{Pigou}} = \frac{-\sum_{j=1}^{N} \lambda_j \xi_j \varepsilon_{y_j} \frac{y_j}{\theta_j y} \frac{1}{1 - t^{Pigou}}}{\varepsilon_y}$$

If some worker groups in the economy are not affected by adverse selection, they will only appear in the $1 - \bar{y}$ and $\varepsilon_y$ terms, but not in $\Delta^{AS}$. Hence, the discrepancy $\Delta^{AS}$ between the optimal Second Best and Adverse Selection taxes is directly linked to the fraction of the population affected by adverse selection.

**Proposition 8** *At given $\bar{y}$ and $\varepsilon_y$:*

*i) With fully separating contracts, $t^{AS} > t^{SB}$ and $t^{AS} \to t^{SB}$ as the Pareto weights converge to Rawlsian weights ($\mu_1 \to 1$).*

*ii) With full cross-subsidization, when all types are in the same cross-subsidization group,*

*if $\underline{\mu}^j - \underline{\lambda}^j \geq 0$ $(\forall j > 1)$, $t^{AS} > t^{SB}$.*

Condition $\underline{\mu}^j - \underline{\lambda}^j \geq 0$, $\forall j > 1$, is a generalization of the condition $\mu \geq \lambda$ in Proposition 5, with the same intuitions.[14] Proposition 9 is the direct analog of Proposition 6.

**Proposition 9** *With a Rawlsian social welfare function, in the same economy, welfare is higher when there is adverse selection than when there is not.*

This result can be extended from Rawlsian weights to weights mostly concentrated on lower productivity agents (in Section 1.4 a rigorous condition is given). The intuition is that there are two sources of welfare gain from adverse selection. First, because of the rat race, revenues raised at any tax level are higher. This effect unambiguously makes all types who are at the bottom of a disjoint cross-subsidization groups better off. All other types directly suffer from their upward distortion in work, but also indirectly benefit from the raised revenue. The net effect is ambiguous, but lower types are more likely to gain on net, especially if higher types are much more productive. Secondly, the optimal tax could be lower with adverse selection for a given set of Pareto weights, benefitting most or all agents, especially whenever the government has highly redistributive preferences (i.e., Pareto weights $\mu_i$ are concentrated on low $\theta_i$ agents), and the optimal tax in the Second Best would have been very high and costly to high types. Overall, there is a range of Pareto weights, mostly concentrated on low types, for which welfare is higher with adverse selection.

## 1.4  Nonlinear Taxation

When nonlinear income taxation is available, the goal is to compare the full Pareto Frontiers under three informational regimes, illustrated in Figure 1. The first regime is the standard Mirrlees one, in which firms pay workers their marginal products, and the government, who does not know workers' types, sets nonlinear taxation subject to truth-telling constraints. The second regime, the "Second Best with Adverse Selection," refers to a situation in which firms do not know workers' types either, but the government sees private market contracts. Alternatively,

---

[14] As when $\mu < \lambda$ in the two types case, there are intermediate cases involving different configurations of cross-subsidization groups, Pareto weights, and population weights in which $t^{AS}$ and $t^{SB}$ cannot be unambiguously ranked.

one can imagine a government-run firm which takes over private firms, and directly sets the hours and pay contracts so as to screen workers.[15] The most novel case, called Regime 3 or "Adverse Selection and unobservable private contracts," is the one in which neither firms nor the government see workers' types, but the government is in addition either unable (or unwilling) to take over private firms or to manipulate labor contracts *directly*. It only observes total realized pay, not the underlying labor contract. Unlike in the Mirrlees case, it must anticipate that workers are not free to choose their hours of work at a given wage, but face a nonlinear screening wage schedule. Unlike in the Second Best, it must ensure not only that workers self-select appropriately, but also that *firms* do not deviate by offering different types of contracts in response to taxes. The main conclusion from the linear tax case is still true with nonlinear taxation: whenever the government wants to redistribute from high to low types, adverse selection improves welfare.[16] I formulate and solve the general problem with $N \geq 2$ types, but start with the more intuitive and graphically appealing solution for $N = 2$.

Figure 1-2: Informational Regimes
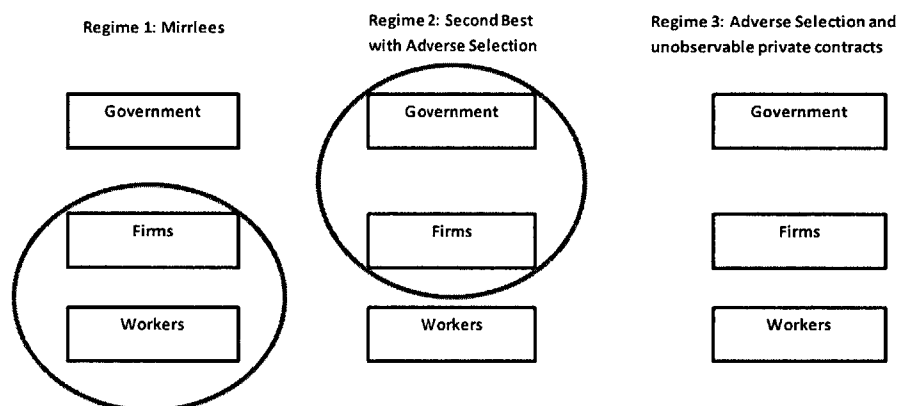


Agents within the same circle have the same information.

---

[15] This case has been studied by e.g., Prescott and Townsend (1984) for an insurance and a signaling problem, by Crocker and Snow (1985) for an insurance problem, and by Spence (1977) for a signaling problem.

[16] Because the proof of this result relies on a direct revelation mechanism, it does not make any assumptions on the tax instruments available to the government – an issue taken up again in the "Implementation" Section 1.4.3 – as long as the government cannot see abilities directly.

## 1.4.1 Characterizing the Mirrlees Frontier and the Second Best Frontier with Adverse Selection

**Mirrlees Frontier**

In the traditional Mirrlees framework (regime 1), the government sees total pay $y$ and sets a menu of contracts specifying consumption and hour pairs $(\{c_i, h_i\})_{i=1}^N$ to solve the program $\left(P^{Mirr,N}\right)$:

$$\left(P^{Mirr,N}\left(\boldsymbol{\mu}\right)\right) : \max_{\{c_i,h_i\}_{i=1}^N} \sum_{i=1}^N \mu_i \left(c_i - \phi_i\left(h_i\right)\right) \tag{1.11}$$

$$\left(IC_{i,i+1}\right) : c_i - \phi_i\left(h_i\right) \geq c_{i+1} - \phi_i\left(\frac{h_{i+1}\theta_{i+1}}{\theta_i}\right) \quad \forall i < N$$

$$\left(IC_{i+1,i}\right) : c_{i+1} - \phi_{i+1}\left(h_{i+1}\right) \geq c_i - \phi_{i+1}\left(\frac{h_i\theta_i}{\theta_{i+1}}\right) \quad \forall i < N$$

$$\left(RC\right) : \sum_{i=1}^N \lambda_i c_i \leq \sum_{i=1}^N \lambda_i \theta_i h_i$$

The constraints $\left(IC_{i,i+1}\right)$ (respectively, $\left(IC_{i+1,i}\right)$) are called "upward incentive compatibility constraints" (respectively, "downward incentive compatibility constraints"), as they ensure type $i$ does not pretend he is higher (respectively, lower) productivity. The final constraint $(RC)$ ensures aggregate resources balance. Alternatively, the government's problem can be specified as maximizing the utility of the highest type $\left(c_N - \phi_N\left(h_N\right)\right)$, subject to incentive compatibility constraints, the resource constraint $(RC)$, and minimal utility constraints on all other types. Under this formulation, for $N = 2$, the low type needs to obtain at least some threshold utility $u$, i.e., $c_1 - \phi_1\left(h_1\right) \geq u$. By varying $u$, we can trace out the whole frontier. This latter formulation will be more convenient for the graphical exploration and occasionally used.

The following Proposition characterizes the familiar Mirrlees frontier for two types.[17]

**Proposition 10** *The Mirrlees frontier can be characterized by three regions.*

*Region 1: When $\mu = \lambda$, none of the incentive constraints are binding, hours of work are efficient, and the Pareto frontier is linear in this region.*

---

[17]See also Bierbrauer and Boyer (2010). The result here is reformulated in terms of the relative proportions of types in the population and extended to N types below.

*Region 2: When $\mu > \lambda$, $(IC_{21})$ is binding, the low type works inefficiently little, the high type works an efficient number of hours, and the Pareto frontier is strictly concave.*

*Region 3: When $\mu < \lambda$, $(IC_{12})$ is binding, the low type works an efficient number of hours, the high type works inefficiently much, and the Pareto frontier is strictly concave.*

Whenever the low type is granted a disproportionate Pareto weight $(\mu > \lambda)$, the incentive constraint of the high type is binding – the most typical case in the optimal taxation literature. The threshold for $\mu$ translates into thresholds for $u$. In particular, there exist four cut-off levels $u_{\min}^{Mirr} < \underline{u} < \overline{u} < u_{\max}^{Mirr}$, defined in the Appendix, such that the regions are delimited by, respectively, $\underline{u} \leq u \leq \overline{u}$ (Region 1), $\overline{u} \leq u \leq u_{\max}^{Mirr}$ (Region 2), and $u_{\min}^{Mirr} \leq u \leq \underline{u}$ (Region 3). To interpret them, note that in Region 1, work hours are fixed at their efficient levels, and utility is transferred one-for-one (because of quasilinearity) from one type to the other, by varying only consumption. As the consumption of the low type keeps increasing, however, constraint $(IC_{21})$ will become binding. This point defines $\bar{u}$ as the utility of the low type when work hours are efficient, and constraint $(IC_{21})$ has just become binding. It is the highest utility level that can be granted to the low type without the high type wanting to mimic him, i.e., before hours $h_1$ have to be distorted. The threshold $\underline{u}$ is defined symmetrically as the utility level of type 1 when, at efficient hours, $(IC_{12})$ has just become binding.

## Second Best Frontier with Adverse Selection

In the Second Best case with Adverse Selection, neither firms nor the government know workers' types, but the government can directly set private labor contracts. The program is now:

$$\left(P^{SB,N}\left(\boldsymbol{\mu}\right)\right) : \max_{\{c_i,h_i\}_{i=1}^{N}} \sum_{i=1}^{N} \mu_i \left(c_i - \phi_i\left(h_i\right)\right) \tag{1.12}$$

$$\left(IC_{i,i+1}\right) : c_i - \phi_i\left(h_i\right) \geq c_{i+1} - \phi_i\left(h_{i+1}\right) \qquad \forall i < N$$

$$\left(IC_{i+1,i}\right) : c_{i+1} - \phi_{i+1}\left(h_{i+1}\right) \geq c_i - \phi_{i+1}\left(h_i\right) \qquad \forall i < N$$

$$\left(RC\right) : \sum_i \lambda_i c_i \leq \sum_i \lambda_i \theta_i h_i$$

The constraints look very similar to the ones in the Mirrlees model, with one crucial difference in the downward constraints $(IC_{i+1,i})$, which drives all of the subsequent results. In the Mirrlees

case, when a high productivity agent deviates to a lower level of income in response to taxes, he can take advantage of his higher productivity to generate the same level of income as the low type, but with less hours of work. In other words, he receives the same *wage per hour* for any level of hours worked. With adverse selection, this is no longer true because the wage, which serves as part of a screening mechanism, is a nonlinear function of hours worked. When a high type wants to mislead the government into thinking that he is a lower type, by producing the lower type's income level, he unavoidably also misleads the firm. The firm then pays him the lower type's wage, so that he still needs to work as many hours as the lower type to earn the same income. This makes the downward deviation less attractive. The following proposition characterizes the Second Best frontier with two types:

**Proposition 11** *The Second Best Frontier with Adverse Selection is characterized by:*

*Region 1: For $\mu = \lambda$, both incentive constraints are slack, both workers work efficient hours, and the Pareto frontier is linear.*

*Region 2: For $\mu > \lambda$, $(IC_{21})$ is binding, the high type works efficient hours, the low type works too little, and the Pareto frontier is strictly concave.*

*Region 3: For $\mu < \lambda$, $(IC_{12})$ is binding, the low type works efficient hours, the high type works too much, and the Pareto frontier is strictly concave.*

Again, there exist four thresholds for $u$, $u_{\min}^{SB} < \underline{u}' < \overline{u}' < u_{\max}^{SB}$ (defined in the Appendix) which delimit the three regions, with an interpretation analogous to the Mirrlees case.

**Comparing welfare with and without adverse selection**

Proposition (12) compares the frontiers in the Mirrlees and the Second Best with Adverse Selection cases.

**Proposition 12** *For $\mu > \lambda$, welfare is higher in the Second Best with Adverse Selection regime than in the Mirrlees regime. For $\mu < \lambda$, welfare is lower.*

Hence, whenever the government disproportionately cares about low types relative to their share in the population, welfare is higher under adverse selection. It is also instructive to rephrase this result more visually using the Pareto frontiers. There are two cases depending on whether condition $NL1$ holds or not.

39

Figure 1-3: Mirrlees and Second-Best with Adverse Selection Frontiers
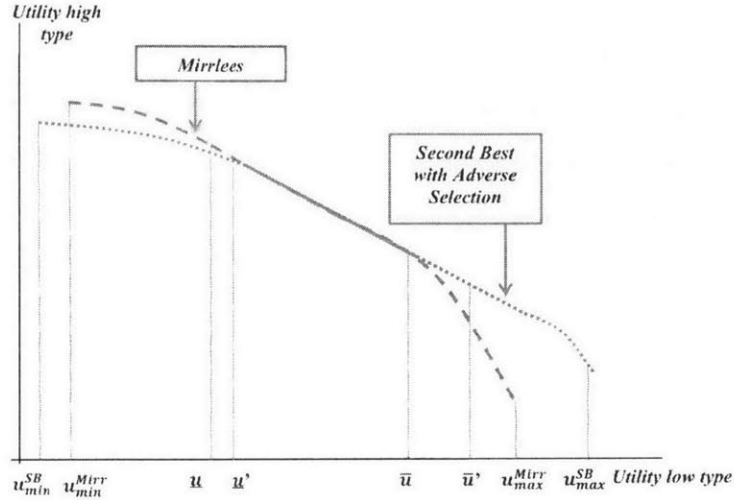


Condition **NL1**: $(\phi_1(h_1^*) - \phi_1(h_2^*)) \leq \phi_2\left(\frac{h_1^*\theta_1}{\theta_2}\right) - \phi_2(h_2^*)$ where $h_i^*$ is the first best effort level for agent of type $i$, defined by $\phi_i'(h_i^*) = \theta_i$.

If condition **NL1** holds, then i) $\underline{u} \leq \underline{u}' \leq \overline{u} \leq \overline{u}'$, ii) for $u \leq \underline{u}'$, the Mirrlees Pareto frontier is above the Adverse Selection Pareto frontier, iii) for $\underline{u}' \leq u \leq \overline{u}$, the Adverse Selection and Mirrlees frontiers coincide and are linear, and iv) for $u \geq \overline{u}$, the Adverse Selection Pareto frontier is above the Mirrlees Pareto frontier. If, condition NL1 does not hold, then i) $\underline{u} \leq \overline{u} \leq \underline{u}' \leq \overline{u}'$, ii) for $u \leq \overline{u}$, the Mirrlees Pareto frontier is above the Adverse Selection Pareto frontier, iii) for $\overline{u} \leq u \leq \underline{u}'$, it is possible to have either frontier above the other one, and iv) for $u \geq \underline{u}'$, the Adverse Selection Pareto frontier is above the Mirrlees Pareto frontier. Figure $1-3$ illustrates the relative position of the frontiers when **NL1** holds.

The welfare result can be extended to $N > 2$ types, in a sharper way than with linear taxes. To simplify the proof, we assume that at the optimum, different types are not pooled completely – in the sense of being assigned exactly the same contract.

**Proposition 13** *For $N \geq 2$, if the government cannot observe workers' types, can use nonlinear income taxation, and does not pool different types at the optimum, then:*

*i ) if $\sum_{i=1}^{j} \mu_i > \sum_{i=1}^{j} \lambda_i \; \forall j \leq (N-1)$, welfare is higher when there is adverse selection,*

*ii) if $\sum_{i=1}^{j} \mu_i < \sum_{i=1}^{j} \lambda_i \; \forall j \leq (N-1)$, welfare is lower when there is adverse selection.*

These conditions on the Pareto weights make explicit how strong or weak the government's

redistributive preferences have to be for welfare to be higher or lower with adverse selection. Highly redistributive preferences are those which place higher cumulative welfare weights up to a given type than the corresponding cumulative proportions in the population.[18] Whenever the government wants to redistribute heavily toward low types, having adverse selection in the labor market helps him do so with a lower deadweight loss. The intuition for this was already captured in the relaxed incentive compatibility constraints. It is now less attractive for any worker to try to lie to the government by pretending to have lower productivity, because, by doing so, he also misleads the firm, and is paid a lower wage per hour. The rat race reduces a worker's capacity to respond negatively to taxes.

**Link to the literature on hidden trades and screening**

Several differences with some important papers in the abundant literature on screening and hidden trades explain why adverse selection can be welfare-improving.

Prescott and Townsend (1984) essentially consider a version of regime 2 in their analysis of a Rothschild-Stiglitz insurance market. In the current chapter, there exists "double" adverse selection, namely between the government and workers, and between firms and workers, which are conflated in Prescott and Townsend. The welfare result here crucially depends on the existence of firms, with potentially more information than the government, as a middle layer between the latter and workers, something which is missing in Prescott and Townsend.[19]

There is also a literature that assesses the welfare effects of improving information when there is adverse selection and that highlights the detrimental redistributive effects and positive efficiency effects of allowing categorical discrimination in insurance markets (see Crocker and Snow, 1986, Hoy, 1984, 1989, among others). While Schmalensee (1984) cautions against the idea that more information is always welfare improving, there is general agreement that *perfect* information is better than imperfect information, unless information acquisition is costly. The

---

[18] Note that there is an intermediate range of Pareto weights such that welfare cannot be unambiguously ranked - analogous to the case $\mu = \lambda$ with $N = 2$, when the ranking of the frontiers depended on whether condition NL1 held or not.

[19] Prescott and Townsend also note that the Second Best allocation is problematic to implement competitively because of the absence of individualized prices -- but these can be imitated by nonlinear income taxes. When private firms can act in potential discordance with the government, the decentralization problem arises even with nonlinear taxation and thus also requires nonlinear prices for *firms* (through the nonlinear payroll taxes considered in the next Subsection).

big difference is that, in these papers, any information that firms discover is immediately known to the government as well. Of course, moving from a world with an uninformed social planner to one with a perfectly informed one is welfare improving (this represents a shift from the Second Best with Adverse Selection to the First Best case). But when the government wants to redistribute, while firms have other objectives, improvements in the differential information set of *firms* can be welfare-reducing and increase the efficiency cost of taxation, as represented in the move from the Second Best to the Mirrlees frontier.

Related is the wide literature on hidden trades, in which trades adjust endogenously to government policies as do the private labor contracts here. In Golosov and Tsyvinski (2007) the government tries to insure agents who can engage in hidden trades in a private insurance market. Their private market equilibrium is inefficient because of the externality imposed by a firm's contracts on other firms' contracts through the work incentives of workers. The government can correct for the externality and improve welfare using taxes and subsidies.

Within many such models, the government can create Pareto improvements relative to the competitive equilibrium using tax tools - which might sound identical to the result in this chapter. Rothschild and Stiglitz (1976) themselves showed that public subsidies for insurance contracts can be Pareto improving, by essentially replacing the cross-subsidy which guarantees efficiency in the MWS setting. Greenwald and Stiglitz (1986) show that asymmetric information generates externalities which typically cause competitive equilibria to be constrained inefficient,[20] and that linear taxation can be Pareto improving. Guesnerie (1998) and Geanakoplos and Polemarchakis (2008) focus on how differential commodity taxation can improve upon a private market equilibrium with hidden trades. But in the current chapter, the private market *is already constrained efficient*, and the government, armed with weakly less information than firms, is not generating a Pareto improvement. It merely moves the economy along the Pareto frontier. I take the redistributive preferences of the government as given and study the effects of different market structures to find that a market with adverse selection may be better for welfare when redistributive preferences are high. The distinctive result emerges by looking at adverse selection from the different angle of Mirrleesian optimal taxation.

---

[20]The MWS equilibrium satisfies all their stringent conditions for *not* being constrained inefficient.

## 1.4.2 Welfare in the Adverse Selection and Unobservable Private Contracts Regime

I now turn to regime 3, in which the government no longer sees private labor contracts or cannot directly control them. It can only see the income paid by firms to workers, but neither the underlying menus offered by firms, nor actual work hours. Hence, for each desired allocation $\{c_i, h_i\}$, it needs to set income levels and taxes $\{y_i, T_i\}$, with $T_i = y_i - c_i$, such that i) firms do not find it profitable to deviate and offer another contract outside of the menu $\{h_i, y_i\}_{i=1}^N$ ("firms' incentive compatibility constraints"), ii) workers indirectly choose the pair $\{c_i, h_i\}$ destined for them by directly choosing $\{h_i, y_i\}$, and paying income taxes $T_i = T(y_i)$ ("workers' incentive compatibility constraints"), iii) firms break even on the portfolio of contracts offered, so that

$$\sum_{i=1}^N \lambda_i y_i = \sum_{i=1}^N \lambda_i \theta_i h_i \tag{1.13}$$

and iv) the government's budget constraint holds:

$$\sum_{i=1}^N \lambda_i y_i = \sum_{i=1}^N \lambda_i c_i \tag{1.14}$$

If constraint i) could be omitted, then the constraints for problem $\left(P^{SB,N}\right)$ in (1.12) would be sufficient and setting $y_i = \theta_i h_i$ would be feasible for all $i$, so that (1.13) would be equivalent to (1.14). But firms too can deviate and offer different contracts than those the government intended. To limit such deviations, the government needs to set prohibitively high taxes (say, 100%) on incomes not in $\{y_i\}_{i=1}^N$. Even then firms can still undertake many possible deviations. A profitable deviation must i) involve only one or several of the allowed incomes in $\{y_i\}_{i=1}^N$, ii) make non-negative profits, even after other contracts rendered unprofitable by it are dropped. Formally, let $\mathcal{P}(\{1, ..., N\})$ be the power set of $\{1, ..., N\}$. Let $\bar{\theta}_{A_k}$ denote the average productivity within any subset $A_k \in \mathcal{P}(\{1, ..., N\})$. A deviation is a collection of $K$ triples $\left\{A_k, y^k, h_{k,A_k}\right\}_{k=1}^K$, specifying which groups of agents $A_k$ (potentially singletons) are targeted by a contract offering income level $y^k$ in exchange for an amount of work $h_{k,A_k}$.[21] Profits from

---

[21] To be profitable, naturally, $y^k$ must be part of the allowed income levels $\{y_i\}_{i=1}^N$.

the deviating sets of contracts must be non-negative if accepted by their targeted groups, i.e., $\sum_k h_{k,A_k} \bar{\theta}_{A_k} \geq 0$. The required work hours $h_{k,A_k}$ can be smaller than (respectively, larger than or equal to) $y^k/\bar{\theta}_{A_k}$, in which case we say that group $A_k$ is being cross-subsidized by other groups (respectively, is cross-subsidizing others or breaking even).[22] In general, all possible configurations of pooling groups and pooling income levels need to be considered. This underscores that firms still have a lot of leeway to trick the government by offering new contracts.[23] Proposition 14 shows that despite this hurdle the ranking of the frontiers is the same as when private contracts are observable (regime 2).

**Proposition 14** *For $N \geq 2$, if the government cannot observe private labor contracts, the result from Proposition 13 still holds.*

The essence of the proof is that, no matter what deviations firms consider, they can never offer workers as profitable "downward" deviation opportunities as in the Mirrlees case without making losses. Another way to gain intuition is to once more think in terms of the wage per hour. In the Mirrlees case, the wage per hour of type $i$ is equal to his marginal product $\theta_i$ for any amount of hours worked. In the Second Best with Adverse Selection, type $i$ would only be paid a wage of $\theta_j < \theta_i$ per hour were he to deviate to a lower income level $y_j < y_i$. With unobservable contracts, the situation is in between those two. Firms can potentially provide a higher wage per hour than $\theta_j$ at $y_j$ (the Second Best with Adverse Selection case), but will never be able to pay a wage of $\theta_i$ at any income level $y_j \neq y_i$ (the Mirrlees case) without violating workers' incentive constraints. The natural next question is, how closely the government can come to his desired Second Best allocation when private contracts are unobservable, or, the question of implementation.

---

[22]For example, consider the deviation which consists in pooling workers $\theta_{i+1}$ and $\theta_i$ at $y_i$. Using the newly introduced notation, $y^1 = y_i$, $A_1 = \{i, i+1\}$. The required work hours for such a deviation would have to be at least $h_{1,\{i,i+1\}} \geq y_i/\bar{\theta}_{\{i,i+1\}}$ with $\bar{\theta}_{\{i,i+1\}} = \frac{\lambda_i}{\lambda_i + \lambda_{i+1}}\theta_i + \frac{\lambda_{i+1}}{\lambda_i + \lambda_{i+1}}\theta_{i+1}$. This deviation will attract both workers if $y_i = \theta_i h_i$, because then $y_i/\bar{\theta}_{\{i,i+1\}} < h_i$, and, by the binding $(IC_{i+1,i})$, type $i+1$, who was just indifferent between $\{h_{i+1}, y_{i+1}\}$ and $\{h_i, y_i\}$, will strictly prefer the deviating contract $\{h_{1,\{i,i+1\}}, y_i\}$.

[23]However, whenever the allocation that the government desires to implement is such that the upward constraints are binding, then then there is no conflict between firms and government. See the proof of Proposition 14.

### 1.4.3 Implementation with Adverse Selection and Unobservable Private Contracts

A little thought experiment can highlight the peculiarities of this adverse selection situation, in which the government needs to ensure that firms, as well as workers, comply with its recommendations. Consider all potential choices available to the government. First, it could force firms to break even on each contract separately, and do all of the redistribution itself through taxes, so that $y_i = \theta_i h_i$ and $T_i = \theta_i h_i - c_i$ for all $i$. At the other extreme, it could let firms do all the redistribution, by assigning gross incomes to be the desired consumption levels, and setting taxes to zero, i.e., $y_i = c_i$ and $T_i = 0$. In between those two extremes, given the target allocations $\{c_i, h_i\}_{i=1}^N$, the government could set any incomes $\{y_i\}_{i=1}^N$ satisfying simultaneously (1.13) and (1.14).

The question then becomes how much of the redistribution the government can leave to firms. The choices of income and tax levels $\{y_i, T_i\}_{i=1}^N$ determine the profitable deviation opportunities available to firms. In particular, at a fully separating contract ($y_i = \theta_i h_i$), as illustrated above, firms are tempted to pool workers of type $\theta_i$ and $\theta_{i+1}$ at $y_i$. On the contrary, if income taxes are zero and firms take care of all the redistribution ($c_i = y_i, \forall\ i$), then, if agent $i$ is strongly cross-subsidizing others ($\theta_i h_i >> c_i$), the incentives for firms are toward cream-skimming worker $i$ into actuarially fairer contracts. For $N = 2$, the government can perfectly implement any Second Best allocation despite unobservable labor contracts.

**Proposition 15** *With unobservable labor contracts and $N = 2$, the government can implement any allocation from the Second Best with Adverse Selection (when labor contracts are observable) using only nonlinear income taxes.*

The second-best allocation can be implemented by assigning any income levels $(y_1, y_2)$ such that constraints (1.26), (1.27) and (1.28) in the Appendix hold. They could bear only weak relation to consumption levels – leading to a potentially unusual and non-monotone tax system – because, no matter what their assigned *income* levels are, workers only care about their final *consumption* levels. Firms on the other hand only care about the assigned income levels, and are happy to offer any pair satisfying the aforementioned constraints. The tax system is indeterminate even at equilibrium income levels, in the sense that many $(y_1, y_2)$ pairs can

sustain the second-best consumption levels. This is not true in the Mirrlees model, where, for any level of recommended hours, earned income is hours times marginal product. Here, it is always necessary for the government to let firms do some of the desired redistribution through cross-subsidization between workers (the conditions on $y_2$ imply that the high type is paid less than his product).

In general, however, it will not be possible to implement any arbitrary Second Best allocation. As the number of types increases, the requirements on each income level become more stringent, since any configuration of deviating contracts needs to be ruled out. There is no guarantee that the ranges for the income levels needed to prevent all deviations will contain non-negative values only, and yet allow firms to break even. The Appendix illustrates these difficulties for $N = 3$.

The problem can be resolved by nonlinear payroll taxes, levied on firms, and which vary with the income paid to workers. Since the Second Best allocation is resource-compatible, there always exist transfers between firms and governments which allow firms to break even. The optimal payroll taxes compensate for the net profits or losses that firms *would have made* if they offered the income levels recommended by the government. Direct profit taxation is ruled out because of the unobservable output. First, the government determines the admissible income levels, which do not allow firms any profitable deviations. It then announces a menu of payroll taxes (or transfers), $T_i^F = T^F(y_i) = \theta_i h_i - y_i$ as a function of the income paid by firms to workers, so as to either tax away a net gain or to compensate for a net loss.[24] The two-tier tax system is crucial. A pure income tax system may not allow firms to break even while satisfying their no-deviation constraints. Conversely, a pure payroll tax need not satisfy all the deviation constraints by firms.

**Proposition 16** *With $N \geq 2$ types, any allocation $\{c_i, h_i\}_{i=1}^N$ solving problem $\left(P^{SB,N}\right)$ in the Second Best with Adverse Selection can be implemented, even if private contracts are unobservable, by a sequence of incomes $\{y_i\}_{i=1}^N$, income taxes $\{T_i\}_{i=1}^N$, and payroll taxes $\left\{T_i^F\right\}_{i=1}^N$ such that:*

---

[24]Note that, since the transfers are conditional on money which actually changes hands between firms and workers (assumed to be observable, say, on the paystubs of employees), firms cannot game the system and collude with workers, by pretending to pay some income level, when they in fact pay another one, in order to get a payroll transfer that they could use to pay workers more.

$$y_i > \max_{m>i} \bar{\theta}_{\{i,..,m\}} \phi_m^{-1} (c_i - c_m + \phi_m (h_m)) \quad \forall i < N \quad \text{and } y_N < \phi_{N-1}^{-1} (c_N) \theta_N.$$

$T(y) = y - c_i$ if $y = y_i$ for some $i$, and $T(y) = 2y$ otherwise.

$T^F(y) = \theta_i h_i - y$ if $y = y_i$ for some $i$, and $T^F(y) = 2y$ otherwise.

The addition of payroll taxes to the government's toolbox can allow to implement any allocation from the Second Best with Adverse Selection, for $N \geq 2$, even if private contracts are unobservable. This setting, in which the government cannot observe private contracts or, equivalently, cannot make taxes dependent on the labor contract itself, seems more realistic. Proposition 16 highlights that policies which would incentivize firms to reveal their contracts offered would be helpful only insofar as the government did not have access to the nonlinear payroll taxes needed.

## 1.5 Empirical and Policy Implications

The findings in this chapter have two implications for policy design. The first is that estimates of labor supply and taxable income elasticities that are obtained using standard empirical methods may not capture the underlying Marshallian elasticities that are the key inputs for optimal tax calculations. The second involves the normative analysis of tax rates and their design, as well as labor market policies that may affect the degree of adverse selection.

### 1.5.1 Interpretation, Measurement, and Use of Elasticities

**Interpretation of measured elasticities:** Because the wage depends on the tax structure in equilibrium, one cannot directly map measured elasticities – the change in hours or income associated with a given change in net wages – to structural elasticities, the fundamental parameters of preferences, without a knowledge of the underlying market structure. For example, the labor supply elasticity of the high type is no longer just a function of his disutility of effort, but also of the low type's preferences, his proportion in the population, and the type of equilibrium. This is because the high type's labor supply is determined by firms in general equilibrium, subject to the low type's reaction.[25] Related, the elasticity of taxable income is not directly mapped

---

[25] This is reminiscent of other papers in which work hours are part of a job "package," and where knowledge of the market structure is required in order to map estimated elasticities to primitives (Chetty *et al.*, 2011, Altonji

into the elasticity of labor supply, since taxable income is also the result of the wage per hour, which is endogenously determined, and depends on taxes (see also Feldstein, 1999, Slemrod and Yitzhaki, 2002, and Chetty, 2009).

**Estimating taxable income elasticities:** The second lesson is that the measurement and estimation of the relevant elasticities is more difficult than typically assumed. Even policy reforms used as "natural experiments" might not be able to correctly capture the elasticities, which are determined in general equilibrium when different groups are interconnected. In the US, the largest changes in tax rates have been for the top of the income distribution, typically used as the treatment group, with lower incomes acting as control groups in a difference-in-difference analysis (for comprehensive and critical assessment of this literature see Slemrod (1998), Giertz (2004), and Saez *et al.* (2012)). Unfortunately, in the presence of adverse selection, a reform affecting high incomes (the high types of the model) will also affect the labor contracts offered to lower incomes (low types), turning the latter into an invalid control group.[26] Paradoxically, the problem is greatest when the groups are more comparable, i.e., closer in the income distribution, as they are likely to interact in the same labor market and have interdependent labor contracts.[27]

**Use of elasticities for tax design:** Taxable income elasticities may not be sufficient statistics for the welfare cost of taxation, as they do not capture the externalities arising from the distortions in labor supply (the rat race and the informational rent effects). This is similar to the limitations of the taxable income elasticity as the sole measure of the efficiency cost of taxation, when there are additional channels through which households react to taxes, such as avoidance and income shifting, which generate fiscal externalities (see Saez *et al.*, 2012, Chetty, 2009).

These three implications could also be valid for other environments in which adverse selection is thought to be a problem, such as health care insurance or markets for used durables.

---

and Paxson, 1989, Dickens and Lundberg, 1993).

[26]If there is cross-subsidization, the detrimental effect of tax hikes would be overestimated: As taxes increase and the labor supply of high incomes decreases, the pre-tax incomes of the control group increase (through their informational rents).

[27]This is reminiscent of models in which the effects of public policies act through coordinated changes in institutions (here, labor contracts), rather than only individual behaviors (see for example Lindbeck's (1995) "Social Multiplier" idea, or Alesina et al. (2005)).

### 1.5.2 Construction of Optimal Tax Schedules

This model has three important implications for the praxis of tax policy design, as well as an application to social insurance.

**Setting the optimal linear tax rate:** At any given set of measured elasticities and income distribution, adverse selection will tend to push tax rates higher, as long as the government wants to redistribute toward lower types (see Propositions 5 and 8). Knowledge of the underlying market structure is hence important for the government; a government armed with estimates of taxable income elasticities and inequality will set the tax rate too low if it wrongly assumes there is no adverse selection.

The strength of this effect depends on the proportion of markets in the economy affected by adverse selection and subject to the same income tax schedule. Markets more prone to adverse selection can be defined among others by age groups (younger workers without a track record), by type of job (more complex, multifaceted jobs), by profession (less automated jobs where worker quality matters more). Ideally, the government could set market specific taxes according to the formulas in this chapter. Age-dependent taxation could be viewed in this light: if younger people are more prone to adverse selection, their income tax schedule would optimally be shifted upwards at all income levels relative to older people.[28] But if taxes only condition on income, the optimal tax would be based on an average weighted taxable income elasticity, and the externality term $\Delta^{AS}$ would only take into account those affected by adverse selection (see Subsection 1.3.3).

**Adverse Selection is endogenous to government policy:** While imperfect information about heterogeneous workers' types might be a common feature in most markets and economies, its consequences, i.e., adverse selection *per se* and the use of screening, depend on the structure of the economy, which is endogenous to government action, mostly to regulatory policies.

*Statistical discrimination*: The government can influence firms' opportunities to engage in statistical discrimination – that is, selecting workers based on characteristics correlated with productivity - through regulations on labor contracts and anti-discrimination laws. For instance,

---

[28] Of course, this ignores considerations of age-specific labor supply elasticities or credit constraints for younger people.

through the lens of firms in my model, women with children are lower productivity workers. If direct discrimination against them is prevented – as is the case in many countries – firms will have to indirectly screen through the labor contract.[29] They might then offer a menu of contracts: a low-paying, part-time contract with shorter hours and more maternity leave, likely to be taken up by working mothers, and a high-paying, full-time contract with overtime bonuses, late-afternoon and week-end meetings, and little parental leave, likely to be taken up by workers without small children.[30]

*Firing costs*: The more difficult it is to fire a worker once his type is discovered, the costlier adverse selection will be for firms. Kugler and Saint Paul (2004) review empirical studies which find that increasing the stringency of employment protection legislation shifts the composition of employment away from young people and female workers, perceived as being of lower productivity. If however increasing firing costs are coupled with stricter regulation on ex ante statistical discrimination, screening through menus of work contracts becomes more attractive to firms.

*Pay structure*: Adverse selection cannot occur if there is perfect pay for performance, such as piece rates or purely bonus-based pay, because then the firm would directly reward the worker as a function of his output. On the contrary, contracts specifying the wage as a function of inputs (e.g., required number of hours per day, or set of obligatory tasks) are prone to adverse selection since the firm bears the full risk of having hired a low type. Most pay structures are in between these two extremes, including some pay-for-performance as well as some fixed or input-oriented components. By reducing the prevalence of pay-for-performance, the government can shift more of the risk to firms, and increase the consequences of adverse selection for them, thus augmenting their need to engage in screening through labor contracts.[31]

These policies – widely used in the real world – might not have been introduced explicitly to deal with adverse selection, but, once in place, need to be taken into account by the tax

---

[29] This assumes that the market does not unravel once direct discrimination is forbidden, i.e., firms have access to a screening tool such as hours.

[30] This result hinges on the inability of the government to leverage that information to set taxes. If the government could itself use the information extracted by firms in the tax system (for example, base taxes on gender or IQ tests performed by firms), this would pose a very different problem, in the spirit of the "tagging" literature (Akerlof, 1976).

[31] Incentive pay of course fulfills a useful role when there is moral hazard.

system. A government interested in redistribution might choose not to reduce adverse selection in the labor market, even though it had the aforementioned tools to do so.

If adverse selection were a relevant issue, higher redistribution should, all else equal, go hand in hand with more anti-discrimination policies against lower productivity groups (potentially, working mothers or inexperienced youth), more stringent employment protection, and less pay-for-performance. Indeed, it might seem, at least anecdotally, that Continental Europe, with its more rigid labor market, more generous youth and maternal employment policies, and regulated labor contracts which alleviate workers from risk can afford a higher level of redistribution at a lower efficiency cost than the US.

**The government may hamper screening:** If the government wants to redistribute, firms' abilities to indirectly screen through work hours and contracts should not be excessively restricted through, for instance, constraints on hours of work such as the 35-hour work week in France.

**Application to social insurance:** The model in this chapter can be directly applied to social insurance – such as health insurance – when it coexists with private insurance providers. If the government chooses a subsidy on health insurance expenditures to maximize a weighted sum of utilities of people with high and low health risks, redistribution toward higher health risks will be facilitated by adverse selection. Intuitively, the government and private insurers have conflicting objectives, which relieve insurees' incentive compatibility constraints. Hence, policies to reduce adverse selection through mandates or regulations may be misguided under some conditions.

## 1.6 Conclusion

Empirical evidence suggests that there is asymmetric information between firms and workers regarding the latter's ability, and that, accordingly, firms may be screening workers through nonlinear compensation contracts. Because work effort is used as a screening device for unobserved talent, labor supply decisions and responses to income taxes are different from those in the traditional optimal taxation literature. Firms have a more active role in setting hours

of work and pay than is typically assumed, while workers are more constrained in their labor supply choices.

This chapter considered the problem of optimal linear and nonlinear income taxation when there is adverse selection in the labor market because of workers' private information about their ability. Higher productivity workers are trapped in a rat race, in which they are forced to work excessively, so that firms can screen them from low productivity ones. The nonlinear wage schedule imposed by firms affects the response to taxes, with several implications for optimal tax policy. Most importantly, if the government has sufficiently strong redistributive goals, welfare is higher when there is adverse selection, both with linear and nonlinear taxes. The informational structure of the economy is potentially endogenous to government policies such as bans on discrimination or firing and pay regulations, and a government with strong redistributive goals might find some degree of adverse selection useful. Secondly, the optimal linear tax formula contains additional terms: corrective terms for the rat race distortions, as well as redistributive terms due to the informational rents, whenever firms cross-subsidize workers. At given taxable income elasticities and a given income distribution, taxes are higher with adverse selection whenever the government has highly redistributive preferences. Thirdly, the usual interpretation, estimation, and use of taxable income elasticities may be problematic when labor market contracts are interconnected, and hours of work are determined not just by workers, but also by firms.

At the most general level, the idea of this chapter is that there are endogenous private market contracts which react to, and interact with, the government's tax contract. This "contract inside a contract" setup modifies responses to taxes. In future research, it would be interesting to consider the other ways in which wages and labor supply are part of private market contracts, such as incentive or screening schemes, and their implications for optimal tax policy. The consequences of these labor market imperfections for our interpretation of the estimated taxable income elasticities would be important. It would also be useful to extend the analysis to other labor market imperfections which could affect responses to taxes, among others moral hazard or rent-seeking.

# 1.A  Appendix with Proofs

## 1.A.1  Proofs for the Linear Tax Problem

**Lemma 1** *Monotonicity: At an implementable solution, we have $h_2(t) \geq h_1(t)$:*

**Proof:** Combining the two $(IC)$ constraints yields:

$$\phi_2(h_1) - \phi_2(h_2) \geq (1 - t)(y_1 - y_2) \geq \phi_1(h_1) - \phi_1(h_2)$$

which requires: $\phi_2(h_2) - \phi_2(h_1) \leq \phi_1(h_2) - \phi_1(h_1)$, and hence: $\int_{h_1}^{h_2} \phi_2'(h)\,dh \leq \int_{h_1}^{h_2} \phi_1'(h)\,dh$. But since by the Spence-Mirrlees single-crossing condition, $\phi_1'(h) > \phi_2'(h)$ at every $h$, implementability requires that $h_2(t) \geq h_1(t)$.

**Proof of Proposition (1).**

Substituting for $y_1$ from the budget constraint, $y_1 = \frac{1}{\lambda}[(\lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2) - (1 - \lambda)y_2]$, the maximization problem of the firm is (multipliers are in brackets after the corresponding constraint):

$$\left(P^{MWS}(t)\right) : \max_{y_2, h_2, h_1} (1 - t)y_2 - \phi_2(h_2)$$

$$(IC_{12}) \quad : \quad \frac{1}{\lambda}[(\lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2) - (1 - \lambda)y_2](1 - t) - \phi_1(h_1) \geq y_2(1 - t) - \phi_1(h_2) \quad [\lambda_{12}]$$

$$(IC_{21}) \quad : \quad y_2(1 - t) - \phi_2(h_2) \geq \frac{1}{\lambda}[(\lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2) - (1 - \lambda)y_2](1 - t) - \phi_2(h_1) \quad [\lambda_{21}]$$

$$(RS_1) \quad : \quad \frac{1}{\lambda}[(\lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2) - (1 - \lambda)y_2](1 - t) - \phi_1(h_1) \geq u_1^{RS} \quad [\varphi]$$

The general FOCs are:

$$[y_2] \quad : \quad \lambda = \lambda_{12} - \lambda_{21} + (1 - \lambda)\varphi$$

$$[h_1] \quad : \quad [\theta_1(1 - t) - \phi_1'(h_1)]\lambda_{12} - \lambda_{21}((1 - t)\theta_1 - \phi_2'(h_1)) + \varphi(\theta_1(1 - t) - \phi_1'(h_1)) = 0$$

$$[h_2] \quad : \quad -\phi_2'(h_2) + \lambda_{12}\left[\frac{(1 - \lambda)}{\lambda}\theta_2(1 - t) + \phi_1'(h_2)\right] - \lambda_{21}\left[\frac{(1 - \lambda)}{\lambda}(1 - t)\theta_2 - \phi_2'(h_2)\right] + \frac{1 - \lambda}{\lambda}\theta_2(1 - t)\varphi = 0$$

Note that whenever $\varphi = 0$, we require that $\lambda_{12} > 0$, or else we would have $\lambda_{21} = -\lambda < 0$, which is not possible. Secondly, the incentive compatibility constraint of the high type $(IC_{21})$ should never be binding since the firm is trying to maximize that type's utility. Finally, it could

happen that the incentive constraint of the low type $IC_{12}$ is not binding, although this will never occur if the cost functions satisfy assumption (2). Whenever $IC_{12}$ is slack, with $\lambda_{12} = 0$, then necessarily $\varphi = \frac{\lambda + \lambda_{21}}{1-\lambda} > 0$, so that the contract is fully separating.

Hence, the three possible cases are: i) $\varphi > 0$, $\lambda_{12} = \lambda_{21} = 0$ ii) $\varphi > 0$, $\lambda_{12} > 0$, $\lambda_{21} = 0$ iii) $\varphi = 0$, $\lambda_{12} > 0$, $\lambda_{21} = 0$

**i) Case SB, Second Best:** $\varphi > 0$, $\lambda_{12} = \lambda_{21} = 0$. This is immediate, since no incentive compatibility constraint is binding, hence the allocation is as in the second best.

**ii) Case AS1, Separation:** $\varphi > 0$, $\lambda_{12} > 0$, $\lambda_{21} = 0$ : The necessary equilibrium conditions are:

$$\theta_1 h_1(t)(1-t) = \theta_2 h_2(t)(1-t) - (\phi_1(h_2(t)) - \phi_1(h_1(t))) \tag{1.15}$$

$$\phi_2'(h_2) = (\lambda - (1-\lambda)\varphi + \varphi)\frac{(1-\lambda)}{\lambda}\theta_2(1-t) + (\lambda - (1-\lambda)\varphi)\phi_1'(h_2)$$

$$y_1(t) = \theta_1 h_1(t), \; y_2(t) = \theta_2 h_2(t)$$

and the second order condition is: $\left(\lambda - \frac{\phi_2''(h_2)}{\phi_1''(h_2)}\right) < \varphi(1-\lambda)$. To rewrite the characterization as in the main text, let $\delta = \varphi + 1 > 1$, so that: $\phi_2'(h_2) = \theta_2(1-t)\delta(1-\lambda) + (1 - \delta(1-\lambda))\phi_1'(h_2)$. Note that there could be several solutions to equation (1.15). If this were the case, the one we pick is the one which yields the highest utility to the type 2. For this case, we need $\varphi > 0$. Since the high type is supplying more labor than $h_2^*(t)$, we need:

$$\lambda > \widetilde{\lambda}(t) = \frac{(1-t)\theta_2 - \phi_2'(h_2)}{(1-t)\theta_2 - \phi_1'(h_2)} \tag{1.16}$$

Note that this is indeed a well-defined threshold since $h_2$ (and hence the right-hand side) does not depend on $\lambda$ in case AS1. To check that $(IC_{21})$ is indeed slack, note that from the binding $(IC_{12}) : \theta_2 h_2(t)(1-t) - \theta_1 h_1(t)(1-t) = (\phi_1(h_2(t)) - \phi_1(h_1(t)))$. Combined with the Spence Mirrlees single-crossing condition $\phi_1(h_2(t)) - \phi_1(h_1(t)) \geq \phi_2(h_2) - \phi_2(h_1)$, this guarantees that $(IC_{21})$ is slack.

**iii) Case AS2, Cross-subsidization:** $\varphi = 0$, $\lambda_{12} > 0$, $\lambda_{21} = 0$. The FOCs become:

$$\lambda = \lambda_{12}, \quad \theta_1(1-t) = \phi_1'(h_1), \quad \phi_2'(h_2) = (1-\lambda)\theta_2(1-t) + \lambda\phi_1'(h_2) \tag{1.17}$$

54

and the income levels are determined by the binding $(IC_{12})$ :

$$y_1(t) = \lambda\theta_1 h_1(t) + (1-\lambda)\theta_2 h_2(t) - \frac{1-\lambda}{1-t}(\phi_1(h_2(t)) - \phi_1(h_1(t)))$$

$$y_2(t) = \lambda\theta_1 h_1(t) + (1-\lambda)\theta_2 h_2(t) + \frac{\lambda}{(1-t)}(\phi_1(h_2(t)) - \phi_1(h_1(t)))$$

The Second order condition is: $\lambda\phi_1''(h_2) < \phi_2''(h_2)$. This case can apply only if the last equation in (1.17) has a solution at the given $t$, which requires that $\lambda \leq \tilde{\lambda}(t)$. To check that $(IC_{21})$ is indeed slack, note that from the expression for $y_2$ above: $y_2(1-t) - \phi_2(h_2) = (1-t)\lambda\theta_1 h_1(t) + (1-t)(1-\lambda)\theta_2 h_2(t) + \lambda(\phi_1(h_2(t)) - \phi_1(h_1(t))) - \phi_2(h_2)$ which is this greater than:

$$y_1(1-t) - \phi_1(h_1) = (1-t)\lambda\theta_1 h_1(t) + (1-t)(1-\lambda)\theta_2 h_2(t) - (1-\lambda)(\phi_1(h_2(t)) - \phi_1(h_1(t))) - \phi_2(h_1)$$

because, by the Spence-Mirrlees single crossing condition, and by monotonicity in the hours of work $h_2(t) \geq h_1(t)$, $(\phi_1(h_2(t)) - \phi_1(h_1(t))) \geq \phi_2(h_2) - \phi_2(h_1)$. At *any* solution to (1.17) the high type is supplying too much labor relative to the second best case, since: $(1-t)(1-\lambda)\theta_2 = \phi_2'(h_2) - \lambda\phi_1'(h_2) \leq (1-\lambda)\phi_2'(h_2)$, so that $(1-t)\theta_2 \leq \phi_2'(h_2)$, and since the cost function $\phi_2()$ is convex, this implies that $h_2(t) \geq h_2^*(t)$.

**iv)** The result is straightforward for the low type. For the high type, consider the two cases separately. In case $AS2$, $(\lambda \leq \tilde{\lambda}(t))$, $\frac{dh_2}{dt} = -\frac{(1-\lambda)\theta_2}{\phi_2''(h_2) - \lambda\phi_1''(h_2)}$, which is negative by the SOC. In case $AS1$, $\left(\lambda \geq \tilde{\lambda}(t)\right)$,

$$\frac{dh_2}{dt} = \frac{\theta_2[\phi_2'(h_2) - \phi_1'(h_2)]}{[(1-t)\theta_2 - \phi_2'(h_2)]\phi_1''(h_2) - [(1-t)\theta_2 - \phi_1'(h_2)]\phi_2''(h_2)}$$

which is again negative by the SOC and the Spence-Mirrlees condition.

**v)** In case $AS1$, $\left(\lambda \geq \tilde{\lambda}(t)\right)$, $h_2$ is obtained directly from the binding $(IC_{12})$, and, hence, does not depend on $\lambda$. In case $AS2$, $(\lambda \leq \tilde{\lambda}(t))$, $dh_2/d\lambda = (\phi_1'(h_2) - \theta_2(1-t))/\phi_2''(h_2) - \lambda\phi_1''(h_2)$. This is positive because the high type's excessive labor supply and the Spence-Mirrlees condition together imply that $\phi_1'(h_2) > \phi_2'(h_2) > \theta_2(1-t)$. The denominator is positive because of the SOC.

**Proof of Proposition (2):**

i) Let us show that $(IC_{12})$ is binding, if condition (4) and assumption (2) hold. Suppose by contradiction that at some level of taxes $t$, $(IC_{12})$ was slack. In that case, necessarily $\varphi > 0$, as explained above. With both incentive constraints slack, we know that both $h_1$ and $h_2$ would be at their efficient levels so that: $h_i^*(t) = (\theta_i(1-t)/(a_i\eta))^{\frac{1}{\eta-1}}$ and a slack $(IC_{12})$ would imply, whenever $t < 1$:

$$(1-t)\,\theta_1 \left(\frac{\theta_1(1-t)}{a_1\eta}\right)^{\frac{1}{\eta-1}} - a_1 \left(\frac{\theta_1(1-t)}{a_1\eta}\right)^{\frac{\eta}{\eta-1}} > (1-t)\,\theta_2 \left(\frac{\theta_2(1-t)}{a_2\eta}\right)^{\frac{1}{\eta-1}} - a_1 \left(\frac{\theta_2(1-t)}{a_2\eta}\right)^{\frac{\eta}{\eta-1}}$$

$$\Leftrightarrow \quad \theta_1 \left(\frac{\theta_1}{a_1\eta}\right)^{\frac{1}{\eta-1}} - a_1 \left(\frac{\theta_1}{a_1\eta}\right)^{\frac{\eta}{\eta-1}} > \theta_2 \left(\frac{\theta_2}{a_2\eta}\right)^{\frac{1}{\eta-1}} - a_1 \left(\frac{\theta_2}{a_2\eta}\right)^{\frac{\eta}{\eta-1}}$$

which exactly violates condition (4).

ii) With isoelastic utility of the form $\phi_i(h) = a_i h^\eta$, the $(IC_{12})$ constraint implies that: $\theta_2(1-t)h_2 - a_1 h_2^\eta = K(t)$ where $K(t) = (1-t)\theta_1 h_1^* - a_1 h_1^\eta$. But from the FOC of the low type, $\eta a_1 h_1^*(t)^{\eta-1} = (1-t)\theta_1$ hence $h_1^*(t) = [(1-t)\theta_1/a_1\eta]^{\frac{1}{\eta-1}}$ so that $K(t) = (\eta-1)a_1\left((1-t)\theta_1/a_1\eta\right)^{\frac{\eta}{\eta-1}}$. Hence the $(IC_{12})$ constraint implies that: $\theta_2(1-t)h_2 - a_1 h_2^\eta = (\eta-1)a_1 \left(\frac{\theta_1}{a_1\eta}\right)^{\frac{\eta}{\eta-1}}(1-t)^{\frac{1}{\eta-1}}$. For this to hold at any value of $t$, we require that: $h_2 = M(1-t)^{\frac{1}{\eta-1}}$, for a constant $M$, independent of $t$. In that case, $\tilde{\lambda}$ becomes independent of $t$ :

$$\tilde{\lambda}(t) = \frac{(1-t)\theta_2 - a_2 M^{\eta-1}(1-t)}{(1-t)\theta_2 - a_1 M^{\eta-1}(1-t)} = \frac{\theta_2 - a_2 M^{\eta-1}}{\theta_2 - a_1 M^{\eta-1}}$$

**Proof of Proposition (3) :**

In problem $P^{MWS,N}(t)$, let $\varphi_j$ be the multiplier on the constraint guaranteeing utility $\bar{u}_j$ for type $j$, $\beta_{j,j+1}$ the multiplier on the incentive constraint ensuring that $j$ does not pretend to be type $j+1$, and $\delta$ the multiplier on the resource constraint. The FOCs are:

$$[h_i] \quad : \quad -\phi_i'(h_i)\varphi_i - \beta_{i,i+1}\phi_i'(h_i) + \beta_{i-1,i}\phi_{i-1}'(h_i) + \lambda_i\theta_i\delta = 0$$

$$[y_i] \quad : \quad \varphi_i + \beta_{i,i+1} - \beta_{i-1,i} - \frac{1}{1-t}\delta\lambda_i = 0$$

By convention, normalize $\varphi_N = 1$, $\beta_{0,1} = 0$, and $\beta_{N,N+1} = 0$. Define the modified multipliers

56

$\beta_{i,i+1}/\delta \equiv \beta_i$ and $\eta_i \equiv \left(\varphi_i + \beta_{i,i+1}\right)/\delta$, so that the FOCs become:

$$\eta_i = \frac{1}{1-t}\lambda_i + \beta_{i-1}, \quad \frac{\eta_i}{\beta_{i-1}} = \frac{\left[\phi'_{i-1}\left(h_i\right) - (1-t)\,\theta_i\right]}{\left[\phi'_i\left(h_i\right) - (1-t)\,\theta_i\right]}$$

i) It is immediately clear that for the lowest type $h_1\left(t\right) = h_1^*\left(t\right)$, $\forall t$, since $\beta_{0,1} = 0$. For all other groups, hours of work are inefficiently high since $\left[\phi'_i\left(h_i\right) - (1-t)\,\theta_i\right] > 0$, unless $\varphi_i = \beta_i = 0$, in which case the market splits into two non-interacting groups strictly below and weakly above agent of type $i+1$.

ii) Whenever $\varphi_j = 0$, firms lose money on all groups $1, ...., j$. On the other hand, whenever $\varphi_j > 0$, firms break even on agents $1, ..., j$ as a group. To see why, note that firms cannot make strictly positive profits on any subset of agents. Else, it would be possible for some firm to enter, offer slightly lower hours of work at the same pay, and still make a positive profit. Whenever $\varphi_j > 0$, we have $y_j\left(1-t\right) - \phi_j\left(h_j\right) = \bar{u}_j$. By definition of $\bar{u}_j$, firms can then break even on agents $1, ..., j$, and they will not provide those agents with additional utility (all surplus resources could instead be used to increase type $N$'s utility). Whenever $\varphi_j = 0$, $y_j\left(1-t\right) - \phi_j\left(h_j\right) > \bar{u}_j$, and by the definition of $\bar{u}_j$, this means that the firm is losing money on the subset $1, ..., j$. Hence, the cross-subsidization groups referred to in the main text are defined by the break points $k_1, ..., k_n$ at which $\varphi_k > 0$.

**Generic solution for the equilibrium income levels of each type (no assumptions on which constraints are binding):**

Note that whenever $\beta_i = 0$, given that $\eta_i > 0$, we must have $\varphi_i > 0$ (if $IC_{i,i+1}$ is not binding, agent $i$ must be a break agent and is not part of agent $i+1$'s cross-subsidization group). Thus, within a cross-subsidization group, all ICs bind. Let $N_i$ be the highest index of the types who are together with $i$ in a cross-subsidization group (if $m$ and $k$ are in the same cross-subsidization group, then $N_m = N_k$). Symmetrically, let $n_i$ be the smallest index in the cross-subsidization group. Let $I_{ij}$ be the indicator function equal to 1 if $j$ is in $i's$ cross-subsidization group. $\bar{\theta h}_i$ is the average production in $i$'s cross-subsidization group, i.e., $\bar{\theta h}_i = \sum_j \lambda_j I_{ij} \theta_j h_j / \sum_j \lambda_j I_{ij}$. Let $\bar{\lambda}_j^i = \left(\lambda_{j+1} + ... + \lambda_{N_i}\right) / \sum_j I_{ij}\lambda_j$ (respectively, $\underline{\lambda}_j^i = \left(\lambda_{j-1} + ... + \lambda_{n_j}\right) / \sum_j I_{ij}\lambda_j$) denote the population weights on those strictly above $j$ (respectively, strictly below $j$) in $i$'s cross subsidization group. Using the binding ICs and setting the weighted profit in each cross-

subsidization group to 0 allows to write each type's income as:

$$y_i = \bar\theta \bar h_i + \sum_{j=i+1}^{N_i} \frac{\bar\lambda_{j-1}^i}{(1-t)} \left[\phi_{j-1}(h_{j-1}) - \phi_{j-1}(h_j)\right] - \sum_{m=n_i}^{i-1} \frac{\lambda_{m+1}^i}{(1-t)} \left[\phi_m(h_m) - \phi_m(h_{m+1})\right] \quad (1.18)$$

**Proof of Proposition (4) and (7):**

The welfare effect on $i$ of changing taxes is:

$$dW_i = -\mu_i y_i dt + \mu_i dt \left\{ \sum_j \frac{dy_i}{dh_j}\frac{dh_j}{dt}(1-t) - \phi_i'(h_i)\frac{dh_i}{dt} \right\}$$

$$\frac{dy_i}{dh_j} = \frac{\lambda_j}{\sum_m I_{im}\lambda_m}\theta_j + \frac{\bar\lambda_j^i}{1-t}\phi_j'(h_j) - \frac{\bar\lambda_{j-1}^i}{1-t}\phi_{j-1}'(h_j) \quad \text{for } j > i,\ I_{ij} = 1$$

$$\frac{dy_i}{dh_j} = \frac{\lambda_j}{\sum_m I_{im}\lambda_m}\theta_j + \frac{\lambda_j^i}{1-t}\phi_{j-1}'(h_j) - \frac{\lambda_{j+1}^i}{1-t}\phi_j'(h_j) \quad \text{for } j < i,\ I_{ij} = 1$$

$$\frac{dy_i}{dh_j} = \frac{\lambda_j}{\sum_m I_{im}\lambda_m}\theta_j + \frac{\bar\lambda_i^i}{(1-t)}\phi_i'(h_i) + \frac{\lambda_i^i}{(1-t)}\phi_{i-1}'(h_i) \quad \text{for } j = i,\ I_{ij} = 1$$

Hence:

$$dW_i = -\mu_i y_i dt + \mu_i dt \left\{ \begin{array}{l} (1-t)\sum_{j=1} I_{ij}\frac{\lambda_j}{\sum_k I_{ik}\lambda_k}\theta_j\frac{dh_j}{dt} + \sum_{j>i} I_{ij}\left(\bar\lambda_j^i\phi_j'(h_j) - \bar\lambda_{j-1}^i\phi_{j-1}'(h_j)\right)\frac{dh_j}{dt} \\ + \sum_{j<i} I_{ij}\left(\underline\lambda_j^i\phi_{j-1}'(h_j) - \underline\lambda_{j+1}^i\phi_j'(h_j)\right)\frac{dh_j}{dt} + \left(\bar\lambda_i^i\phi_i'(h_i) - \underline\lambda_i^i\phi_{i-1}'(h_i)\right)\frac{dh_i}{dt} - \phi_i'(h_i)\frac{dh_i}{dt} \end{array} \right\}$$

Using the definitions for $\xi_i$ and $\kappa_i$ from the main text, some cumbersome algebra yields:

$$dW_i = -\mu_i y_i dt - \mu_i dt \left\{ \sum_{j>i}\left(\bar\lambda_j^i\xi_j - \bar\lambda_{j-1}^i\kappa_j\right)\frac{dh_j}{dt} - \sum_{j\le i}\left(\underline\lambda_{j+1}^i\xi_j - \underline\lambda_j^i\kappa_j\right)\varepsilon_{y_j}\frac{y_j}{1-t}\frac{1}{\theta_j} \right\}$$

(with $\varepsilon_{y_j} \equiv d\log y_j / d\log(1-t)$). Define $\bar\mu^i \equiv \sum_{j>i} I_{ij}\mu_j$ (respectively, $\underline\mu^i \equiv \sum_{j<i} I_{ij}\mu_j$). Then:

$$\frac{dW}{dt} = \sum_i \frac{dW_i}{dt} = -\sum_j \left[\left(\bar\lambda_j^j - \bar\mu_j^j - \mu_j\right)\xi_j + \left(\underline\lambda_j^j - \underline\mu_j^j\right)\kappa_j\right]\varepsilon_{y_j}\frac{y_j}{1-t}\frac{1}{\theta_j}$$

With an abuse in notation, let $\bar\lambda^j \equiv \bar\lambda_j^j$, $\underline\lambda^j \equiv \underline\lambda_j^j$, $\bar\mu^j \equiv \bar\mu_j^j$, and $\underline\mu^j \equiv \underline\mu_j^j$. The behavioral and mechanical revenue effects are as in Subsection 1.2.2, $dB = -y\frac{t}{1-t}\varepsilon_y dt$ and $dM = ydt$. Setting

$dW + dB + dM = 0$ yields the formula in the proposition.

**Proof of Proposition (5) :**

Both results follow from the fact that $\xi_2 < 0$, $\kappa_2 < 0$, and the terms $\bar{y}$ and $\varepsilon_y$ are held constant in the comparison.

**Proof of Proposition (6) :**

i) When $\mu = 1$ : In case $AS1$ $\left(\lambda \geq \widetilde{\lambda}(t)\right)$, welfare is

$(1 - t)\theta_1 h_1(t) - \phi_1(h_1(t)) + t(\lambda\theta_1 h_1(t) + (1 - \lambda)\theta_2 h_2(t))$, which is higher at any tax level since the $h_1(t)$ function is the same while the $h_2(t)$ function is higher at any tax level. In case $AS2$, $\left(\lambda \geq \widetilde{\lambda}(t)\right)$, the difference in welfare with the second best is:

$$W^{SB} - W^{AS2} = t(1 - \lambda)\theta_2\left(h_2^*(t) - h_2(t)\right) + (1 - \lambda)\left[(1 - t)\theta_1 h_1^*(t) - \phi_1(h_1^*(t)) - ((1 - t)\theta_2 h_2(t) - \phi_1(h_2(t)))\right]$$

$$(1.19)$$

But by cross-subsidization and the binding $(IC_{12})$, we have that:

$$[(1 - t)\theta_1 h_1^*(t) - \phi_1(h_1^*(t)) - ((1 - t)\theta_2 h_2(t) - \phi_1(h_2(t)))]$$

$$\leq \quad [(1 - t)y_1(t) - \phi_1(h_1^*(t)) - ((1 - t)\theta_2 h_2(t) - \phi_1(h_2(t)))]$$

$$\leq \quad [(1 - t)y_1(t) - \phi_1(h_1^*(t)) - ((1 - t)y_2(t) - \phi_1(h_2(t)))] = 0$$

so that the last term in (1.19) is negative. The first term is negative, since the high type is working more under adverse selection than in the second best. Hence $W^{SB} \leq W^{AS2}$.

ii) If $\mu = 0$: In both the second best case and the adverse selection cases, the government maximizes the utility of the high type exclusively. Hence, even with adverse selection, it acts as a single agent with the firms. The addition of the incentive compatibility constraint makes the best achievable allocation with adverse selection for the high type worse than in the second best because: 1) hours are distorted relative to the second best level $h_2^*(t)$ (the level that maximizes the high type's utility), and 2) because pay is weakly lower than the true product, for any level of hours, i.e., $y_2(t) \leq \theta_2 h_2(t)$. The second best allocation is no longer feasible for the high type with the added incentive compatibility constraint.

iii) When there is adverse selection, the low type is always working the same amount, yet consuming weakly more due to the higher transfer $T$ and the cross-subsidization transfer.

Hence, he must be better off.

**Proof of Proposition (8) :**

At fixed $\varepsilon_y$ and $\bar{y}$, $t^{SB}$ and $t^{AS}$ differ only by the term $\Delta^{AS}$, so the result will follow if we show $\Delta^{AS} > 0$.

i) With fully separating contracts, $\Delta^{AS}$ becomes simply $-\frac{1}{y}\sum_{j=1}^{N}\mu_j\xi_j\varepsilon_{y_j}\frac{y_j/\theta_j}{1-t} > 0$.

ii) With a single cross-subsidization group, $\sum_m I_{jm}\omega_m = \sum_m I_{jm}\lambda_m = 1$. $\Delta^{AS}$ can be rewritten as: $\Delta^{AS} = \frac{1}{y}\sum_j\left[\left(\underline{\mu}^j - \underline{\lambda}^j\right)\left(\xi_j - \kappa_j\right) - \lambda_j\xi_j\right]\varepsilon_{y_j}\frac{y_j/\theta_j}{1-t}$. In this case $\forall j$, $\xi_j < 0$, $\kappa_j < 0$, $\left(\xi_j - \kappa_j\right) = \phi'_{j-1}\left(h_j\right) - \phi'_j\left(h_j\right) > 0$ (by assumption 1), and hence $\Delta^{AS} > 0$ follows from the condition in the Proposition.

**Proof of Proposition (9):**

Identical to the proof of Proposition (6), since the lowest type works the same hours, but benefits from more revenues from the increased work of all other types.

## 1.A.2 Proofs for the Nonlinear Tax Problem

**Proof of Proposition (10):**

This proof is similar to Bierbrauer and Boyer (2010), adapted to the case at hand. I first reformulate the problem as maximizing type 2's utility subject to the low type's utility constraint. Multipliers are in brackets on the line of the constraint they apply to:

$$\left(P^{Mirr}\left(u\right)\right) : \max_{\{c_1,c_2,h_1,h_2\}} c_2 - \phi_2\left(h_2\right)$$

$$(IC_{12}) : c_1 - \phi_1\left(h_1\right) \geq c_2 - \phi_1\left(\frac{h_2\theta_2}{\theta_1}\right) \quad [\beta_1]$$

$$(IC_{21}) : c_2 - \phi_2\left(h_2\right) \geq c_1 - \phi_2\left(\frac{h_1\theta_1}{\theta_2}\right) \quad [\beta_2]$$

$$(RC) : \lambda c_1 + (1 - \lambda)c_2 \leq \lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2 \quad [\delta]$$

$$c_1 - \phi_1\left(h_1\right) \geq u \quad [\gamma]$$

The Pareto weights from the main text can be mapped into the multipliers of the utility constraints using $\gamma = \mu_1/\mu_2 = \mu/\left(1 - \mu\right)$. Note that if the Pareto frontier is linear in some regions, then the same set of Pareto weights could correspond to several different levels of $u$. I

60

simultaneously solve for the thresholds for $\mu$ in the Proposition and the thresholds for $u$ from the text.

Denote the set of admissible allocations, i.e., the allocations $\{c_1, c_2, h_1, h_2\}$ which satisfy both incentive compatibility and the resource constraint, by $B^{Mirr}(u)$ and let $V_2^{Mirr}(u) = \max_{c_1,c_2,h_1,h_2 \in B^{Mirr}(u)} (c_2 - \phi_2(h_2))$. When the value of $u$ varies, the function $V_2^{Mirr}(u)$ traces out all possible values for the utility of the high type. The Pareto frontier consists of all pairs $\left(u, V_2^{Mirr}(u)\right)$ such that $\frac{\partial}{\partial u} V_2^{Mirr}(u) < 0$.

The general solution to this problem is then characterized by the following necessary conditions:

$$[c_2] : 1 + \beta_2 - \beta_1 - (1 - \lambda)\delta = 0$$

$$[c_1] : \beta_1 - \beta_2 + \delta(1 - \lambda) - 1 = 0$$

$$\gamma = \delta - 1$$

$$[h_1] : -(\beta_1 + (\delta - 1))\phi_1'(h_1) + \beta_2 \frac{\theta_1}{\theta_2}\phi_2'\left(\frac{h_1\theta_1}{\theta_2}\right) + \delta\theta_1\lambda = 0$$

$$[h_2] : -\phi_2'(h_2) + \beta_1 \frac{\theta_2}{\theta_1}\phi_1'\left(\frac{h_2\theta_2}{\theta_1}\right) + \delta\theta_2(1 - \lambda) = 0$$

$$\lambda c_1 + (1 - \lambda)c_2 = \lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2, \; c_1 - \phi_1(h_1) = u$$

There are three possible cases to consider: either one of $(IC_{12})$ or $(IC_{21})$ is binding, or none is. It is never optimal to have both binding, as has been shown several times in the literature (e.g., Bierbrauer and Boyer, 2010).

**Region 1.** Suppose that both constraints are slack. This case occurs when $\beta_1 = 0$, $\beta_2 = 0$, and hence $\gamma = \overline{\gamma} \equiv \frac{\lambda}{(1-\lambda)}$. In this region, hours of work are at their efficient levels, defined by $\phi_i'(h_i) = \theta_i$. The interval of levels of utility $u$ for which this can occur, denoted by $[\underline{u}, \overline{u}]$, is derived as follows. Suppose that $u$ increases. At some level, it will become attractive for the high type to pretend to be a low type, and the $(IC_{21})$ will just start binding. One can easily check that this will occur exactly at level $\overline{u}$, defined by: $\overline{u} \equiv \overline{c_2} - \phi_2(h_2^*) + \phi_2\left(\frac{h_1^*\theta_1}{\theta_2}\right) - \phi_1(h_1^*)$, where $\overline{c_2} \equiv \lambda\theta_1 h_1^* + (1 - \lambda)\theta_2 h_2^* + \lambda\left(\phi_2(h_2^*) - \phi_2\left(\frac{h_1^*\theta_1}{\theta_2}\right)\right)$. Similarly, if $u$ decreases too much, the low type will start wanting to pretend to be a high type and $(IC_{12})$ will just become binding. This occurs exactly at utility level $\underline{u}$ defined by $\underline{u} \equiv \underline{c_2} - \phi_1\left(\frac{h_2^*\theta_2}{\theta_1}\right)$, where $\underline{c_2} \equiv$

$\lambda \theta_1 h_1^* + (1-\lambda)\theta_2 h_2^* + \lambda\left(\phi_1\left(\frac{\theta_2 h_2^*}{\theta_1}\right) - \phi_1\left(h_1^*\right)\right)$. Note that in this region, the Pareto frontier is linear, since $V_2^{Mirr}(u) = c_2 - \phi_2\left(h_2^*\right) = \frac{\lambda}{(1-\lambda)}\theta_1 h_1 + \theta_2 h_2 - \frac{\lambda}{(1-\lambda)}u - \frac{\lambda}{(1-\lambda)}\phi_1\left(h_1\right) - \phi_2\left(h_2^*\right)$, so that $\partial V_2^{Mirr}(u)/\partial u = -\lambda/(1-\lambda)$.

**Region 2.** Suppose that $(IC_{21})$ is binding, then $(IC_{12})$ must be slack, so that $\beta_1 = 0$. The solution is then characterized by:

$$\gamma \phi_1'(h_1) = \gamma\theta_1 + [\gamma - \gamma\lambda - \lambda]\left[\frac{\theta_1}{\theta_2}\phi_2'\left(\frac{h_1\theta_1}{\theta_2}\right) - \theta_1\right] \tag{1.20}$$

$$c_1 = \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2 - (1-\lambda)\left(\phi_2(h_2) - \phi_2\left(\frac{h_1\theta_1}{\theta_2}\right)\right)$$

$$c_2 = \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2 + \lambda\left(\phi_2(h_2) - \phi_2\left(\frac{h_1\theta_1}{\theta_2}\right)\right)$$

$$h_2 = h_2^* \equiv \left(\phi_2'\right)^{-1}(\theta_2)$$

$$u + \phi_1(h_1) = \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2^* - (1-\lambda)\left(\phi_2(h_2^*) - \phi_2\left(\frac{h_1\theta_1}{\theta_2}\right)\right) \tag{1.21}$$

Note that this case is consistent (i.e., $\beta_2 \geq 0$) if and only if $\gamma \geq \bar{\gamma}$, or equivalently, if and only if $u \geq \bar{u}$. There is a downward distortion in the labor supply of the low type, as can be seen from the FOCs for $h_1$.

The maximal value for $u$ in this region, denoted by $u_{\max}^{Mirr}$, is achieved when $h_1$ is at the level which maximizes $c_1 - \phi_1(h_1)$ subject to $(IC_{21})$, or equivalently, at the level which obtains when $\gamma \to \infty$ and $\phi_1'(h_1) = \lambda\theta_1 + (1-\lambda)\frac{\theta_1}{\theta_2}\phi_2'\left(\frac{h_1\theta_1}{\theta_2}\right)$.

In this region, $\frac{\partial}{\partial u}V_2^{Mirr}(u) = -\gamma < 0$. The Lagrange multiplier $\gamma$ is given by:

$$\gamma = \lambda\frac{\left[\theta_1 - \frac{\theta_1}{\theta_2}\phi_2'\left(\frac{h_1\theta_1}{\theta_2}\right)\right]}{\left(\phi_1'(h_1) - \lambda\theta_1 - [1-\lambda]\frac{\theta_1}{\theta_2}\phi_2'\left(\frac{h_1\theta_1}{\theta_2}\right)\right)}$$

Note that since $\theta_2 = \phi_2'(h_2)$, we have: $\phi_2'\left(\frac{h_1\theta_1}{\theta_2}\right) \leq \phi_2'(h_1) \leq \phi_2'(h_2) = \theta_2$ so that the numerator of $\gamma$ is positive and hence the denominator must also be positive.

To show that the Pareto frontier is concave in this region, note that:

$$\frac{\partial^2}{\partial u^2}V_2^{Mirr}(u) = \lambda\frac{\partial h_1}{\partial u}\left(\frac{\theta_1}{\theta_2}\right)\frac{\left(\frac{\theta_1}{\theta_2}\right)\phi_2''\left(\frac{h_1\theta_1}{\theta_2}\right)(\phi_1'(h_1) - \theta_1) + \left[\theta_2 - \phi_2'\left(\frac{h_1\theta_1}{\theta_2}\right)\right]\phi_1''(h_1)}{\left(\phi_1'(h_1) + \lambda\theta_1 - [1-\lambda]\frac{\theta_1}{\theta_2}\phi_2'\left(\frac{h_1\theta_1}{\theta_2}\right)\right)^2} \tag{1.22}$$

Using expression in (1.21), the fact that $dh_2/du = 0$ in this region, and that the denominator of $\gamma$ is positive, we can see that $dh_1/du < 0$. In addition, the numerator in (1.22) is positive from the SOC with respect to $h_1$. Hence, $\partial^2 V_2^{Mirr}(u)/\partial u^2 < 0$.

**Region 3.** Suppose that $(IC_{12})$ is binding and $(IC_{21})$ is slack, with $\beta_2 = 0$. The solution is characterized by:

$$h_1 = h_1^* \equiv \left(\phi_1'\right)^{-1}(\theta_1), \quad \phi_2'(h_2) = \theta_2 + (-\gamma + \lambda\gamma + \lambda)\left[\frac{\theta_2}{\theta_1}\phi_1'\left(\frac{\theta_2 h_2}{\theta_1}\right) - \theta_2\right]$$

$$c_1 = \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2 - (1-\lambda)\left(\phi_1\left(\frac{\theta_2 h_2}{\theta_1}\right) - \phi_1(h_1)\right)$$

$$c_2 = \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2 + \lambda\left(\phi_1\left(\frac{\theta_2 h_2}{\theta_1}\right) - \phi_1(h_1)\right)$$

$$c_1 - \phi_1(h_1) = u$$

This case is consistent (i.e., $\beta_1 \geq 0$) if and only if $\gamma \leq \bar{\gamma}$ which is equivalent to $u \leq \underline{u}$. The minimal level of utility $u$ in this regime, denoted by $u_{min}^{Mirr}$, is achieved when $h_2$ is solution to $\max\left(c_2 - \phi_2(h_2)\right)$ subject to $(IC_{12})$, or equivalently, when $\gamma \to 0$ and $\phi_2'(h_2) = (1-\lambda)\theta_2 + \lambda\left[\frac{\theta_2}{\theta_1}\phi_1'\left(\frac{\theta_2 h_2}{\theta_1}\right)\right]$. The argument to show concavity is as for Region 2, except that the SOC for $h_2$ is used.

The proof is complete by recalling that $\gamma = \mu/(1-\mu)$ so that $\gamma > \bar{\gamma} = \lambda/(1-\lambda) \Leftrightarrow \mu > \lambda$.

**Proof of Proposition (11) :**

The problem, indexed by the utility level of the low type and specialized to $N = 2$ is reformulated here, with multipliers in brackets after each constraint:

$$\left(P^{SB}(u)\right) : \max c_2 - \phi_2(h_2)$$

$$(IC_{12}) : c_1 - \phi_1(h_1) \geq c_2 - \phi_1(h_2) \quad [\beta_1]$$

$$(IC_{21}) : c_2 - \phi_2(h_2) \geq c_1 - \phi_2(h_1) \quad [\beta_2]$$

$$(RC) : \lambda c_1 + (1-\lambda)c_2 \leq \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2 \quad [\delta]$$

$$c_1 - \phi_1(h_1) \geq u \quad (\gamma) \quad [\gamma]$$

Again, we can map problem $P^{SB}(\mu)$ to $P^{SB}(u)$ using $\gamma = \mu/(1-\mu)$, taking into account that in linear regions of the Pareto frontier, several values of $u$ correspond to the same value of $\mu$.

63

Denote the set of allocations satisfying both incentive compatibility and the resource constraint by $B^{SB}(u)$. Let $V_2^{SB}(u) = \max_{c_1,c_2,h_1,h_2 \in B^{SB}(u)}(c_2 - \phi_2(h_2))$. When the value of $u$ varies, the function $V_2^{SB}(u)$ traces out all possible values for the utility of the high type. The Pareto frontier consists of all pairs $(u, V_2^{SB}(u))$ such that $\frac{\partial}{\partial u}V_2^{SB}(u) < 0$. In the general case, the FOCs are:

$$[c_2] \quad : \quad 1 + \beta_2 - \beta_1 - \delta(1 - \lambda) = 0$$

$$[c_1] \quad : \quad \beta_1 + \gamma - \beta_2 - \delta\lambda = 0$$

$$[h_1] \quad : \quad -\beta_1\phi_1'(h_1) + \beta_2\phi_2'(h_1) + \delta\lambda\theta_1 - \gamma\phi_1'(h_1) = 0$$

$$[h_2] \quad : \quad -(1 + \beta_2)\phi_2'(h_2) + \beta_1\phi_1'(h_2) + \delta(1 - \lambda)\theta_2 = 0$$

$$\lambda c_1 + (1 - \lambda)c_2 = \lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2, \quad c_1 - \phi_1(h_1) = u$$

**Region 1:** Suppose that the utility level $u$ is in $[\underline{u}', \overline{u}']$ where $\underline{u}' \equiv \underline{c_2} - \phi_1(h_2^*)$, $\underline{c_2} \equiv \lambda\theta_1 h_1^* + (1 - \lambda)\theta_2 h_2^* + \lambda(\phi_1(h_2^*) - \phi_1(h_1^*))$, $\overline{u}' \equiv \overline{c_2} - \phi_2(h_2^*) + \phi_2(h_1^*) - \phi_1(h_1^*)$, and $\overline{c_2} \equiv c_2 = \lambda\theta_1 h_1^* + (1 - \lambda)\theta_2 h_2^* + \lambda(\phi_2(h_2^*) - \phi_2(h_1^*))$. One can check that it is possible to set hours at their efficient levels, and $c_1$ and $c_2$ such that:

$$\lambda c_1 + (1 - \lambda)c_2 = \lambda\theta_1 h_1^* + (1 - \lambda)\theta_2 h_2^*, \quad c_1 = u + \phi_1(h_1)$$

and to have both incentive constraints slack. Any distortion in the hours of work would imply a reduced welfare. Hence, both $\beta_1 = \beta_2 = 0$, and $\gamma = \overline{\gamma} = \lambda/(1 - \lambda)$. The Pareto frontier is linear and decreasing in $u$, since $\frac{\partial}{\partial u}V_2^{SB}(u) = \frac{-\lambda}{1-\lambda}$.

**Region 2:** Suppose that $u \geq \overline{u}'$. Suppose by contradiction that $(IC_{21})$ is slack, so that $h_1 = h_1^*$. Since $u_1 = c_1 - \phi_1(h_1^*)$, the inequality on $u$ implies that: $c_1 \geq \lambda\theta_1 h_1^* + (1 - \lambda)\theta_2 h_2^* + \lambda(\phi_2(h_2^*) - \phi_2(h_1^*)) - \phi_2(h_2^*) + \phi_2(h_1^*)$. The utility of the high type from pretending to be low type would be:

$$c_1 - \phi_2(h_1^*) \geq \lambda\theta_1 h_1^* + (1 - \lambda)\theta_2 h_2^* + \lambda(\phi_2(h_2^*) - \phi_2(h_1^*)) - \phi_2(h_2^*) = \overline{c_2} - \phi_2(h_1^*) \geq c_2 - \phi_2(h_2^*)$$

where the last inequality follows from the inequality on $c_1$, which implies that $c_2 \leq \overline{c_2}$ by the resource constraint. Hence, the high type agent has an incentive to deviate if $h_1 = h_1^*$, so that in fact, we need to have $h_1 < h_1^*$. But then, $(IC_{21})$ needs to be binding, or we could increase $h_1$ by some small $dh_1$ without violating $(IC_{21})$ and generate more output than would be necessary to compensate the low type for the increased effort (that is, $\theta_1 dh_1 > \phi_1'(h_1) dh_1$). In addition, the constraint of the low type, $(IC_{12})$, is slack since:

$$
\begin{aligned}
c_2 - \phi_1(h_2^*) &\leq \lambda \theta_1 h_1^* + (1 - \lambda) \theta_2 h_2^* + \lambda (\phi_2(h_2^*) - \phi_2(h_1^*)) - \phi_1(h_2^*) \\
&\leq c_1 + \phi_2(h_2^*) - \phi_2(h_1^*) - \phi_1(h_2^*) < c_1 - \phi_1(h_1^*)
\end{aligned}
$$

where we used that $\phi_2(h_2^*) - \phi_2(h_1^*) < \phi_1(h_2^*) - \phi_1(h_1^*)$, implied by the Spence-Mirrlees condition. Hence $\beta_1 = 0$ and the solution in Region 2 is characterized by:

$$
\begin{aligned}
&\beta_2 = \gamma - \lambda\gamma - \lambda, \ \ \gamma \geq \frac{\lambda}{(1 - \lambda)}, \ \delta = \gamma + 1, \ \ u = c_1 - \phi_1(h_1) \\
&\gamma \phi_1'(h_1) = (\gamma - \lambda\gamma - \lambda)(\phi_2'(h_1) - \theta_1) + \gamma\theta_1 \\
&\theta_2 = \phi_2'(h_2), \\
&c_2 = \lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2 + \lambda(\phi_2(h_2) - \phi_2(h_1)) \\
&c_1 = \lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2 - (1 - \lambda)(\phi_2(h_2) - \phi_2(h_1))
\end{aligned}
$$

The maximal level of utility $u$ in this region, denoted by $u_{\max}^{SB}$, is the level achieved when $h_1$ is set to maximize $c_1 - \phi_1(h_1)$ subject to $(IC_{21})$, or equivalently, the level of $u$ when $\gamma \to \infty$ and $\phi_1'(h_1) = (1 - \lambda)\phi_2'(h_1) + \lambda\theta_1$. The Pareto frontier is decreasing since $\partial V_2^{SB}(u)/\partial u = -\gamma < 0$. In addition, it is concave. Indeed,

$$
\frac{\partial^2 V_2^{SB}(u)}{\partial u^2} = -\frac{\partial\gamma}{\partial u} = -\lambda\frac{\partial h_1}{\partial u}\frac{\phi_2''(h_1)(\theta_1 - \phi_1'(h_1)) - (\theta_1 - \phi_2'(h_1))\phi_1''(h_1)}{[\phi_1'(h_1) - (1 - \lambda)\phi_2'(h_1) - \lambda\theta_1]^2}
$$

First note that $\gamma = \lambda\frac{\theta_1 - \phi_2'(h_1)}{[\phi_1'(h_1) - (1-\lambda)\phi_2'(h_1) - \lambda\theta_1]}$. Since the numerator is positive, the denominator must be positive too. From the constraints: $u = c_1 - \phi_1(h_1) = \lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2^* - (1 - \lambda)(\phi_2(h_2^*) - \phi_2(h_1)) - \phi_1(h_1)$. Given that $\frac{\partial h_2}{\partial u} = 0$ in this region and that the denominator of $\gamma$ is positive, the previous expression shows that $\frac{\partial h_1}{\partial u} < 0$. From the SOC in $h_1$ and

the positive denominator of $\gamma$, we have $\partial^2 V_2^{SB}(u)/\partial u^2 < 0$.

**Region 3:** Suppose that $u$ is such that $u \leq \underline{u}'$, where $\underline{u}'$ is defined as $\underline{u}' \equiv \underline{c_2} - \phi_1(h_2^*)$ and $\underline{c_2} \equiv \lambda\theta_1 h_1^* + (1-\lambda)\theta_2 h_2^* + \lambda(\phi_1(h_2^*) - \phi_1(h_1^*))$. Define $\underline{c_1} \equiv \lambda\theta_1 h_1^* + (1-\lambda)\theta_2 h_2^* - (1-\lambda)(\phi_1(h_2^*) - \phi_1(h_1^*))$ from the budget constraint. In this case, it must be that $(IC_{12})$ is violated at the first best hour levels, since:

$$
\begin{aligned}
c_1 - \phi_1(h_1^*) &\leq \underline{c_1} - \phi_1(h_1^*) = \lambda\theta_1 h_1^* + (1-\lambda)\theta_2 h_2^* - (1-\lambda)(\phi_1(h_2^*) - \phi_1(h_1^*)) - \phi_1(h_1^*) \\
&= \underline{c_2} - \lambda(\phi_1(h_2^*) - \phi_1(h_1^*)) - (1-\lambda)(\phi_1(h_2^*) - \phi_1(h_1^*)) - \phi_1(h_1^*) = \underline{c_2} - \phi_1(h_2^*)
\end{aligned}
$$

Hence, $h_2^*$ is distorted upwards, and $(IC_{12})$ must be binding. If it were not, we could decrease $h_2^*$ and $c_2$ simultaneously so as not to change the utility of the high type and still create a surplus to be given to the low type. Thus, $\beta_1 \geq 0$. In addition, $(IC_{21})$ is slack. From the $(IC_{12})$, replace $c_2$ as a function of $c_1$ and note that $c_2 - \phi_2(h_2) = c_1 + \phi_1(h_2) - \phi_1(h_1^*) - \phi_2(h_2) \geq c_1 - \phi_2(h_1^*)$, where the last inequality follows from the Spence-Mirrlees single crossing condition. Hence $\beta_2 = 0$. The solution in this case is fully characterized by:

$$
\begin{aligned}
&\beta_1 = (\gamma+1)\lambda - \gamma, \quad \gamma < \frac{\lambda}{(1-\lambda)}, \quad \phi_1'(h_1) = \theta_1, \; u = c_1 - \phi_1(h_1) \\
&\phi_2'(h_2) = (\gamma\lambda + \lambda - \gamma)(\phi_1'(h_2) - \theta_2) + \theta_2 \\
&c_2 = \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2 + \lambda(\phi_1(h_2) - \phi_1(h_1)) \\
&c_1 = \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2 - (1-\lambda)(\phi_1(h_2) - \phi_1(h_1))
\end{aligned}
$$

The minimal level of utility $u$ in this region, denoted by $u_{\min}^{SB}$, is the utility level achieved when $h_2$ is set to maximize $c_2 - \phi_2(h_2)$ subject to $(IC_{12})$, or equivalently, the level of $u$ when $\gamma \to 0$ and $\phi_2'(h_2) = \lambda\phi_1'(h_2) + (1-\lambda)\theta_2$. The proof of concavity of the frontier is exactly as for Region 2.

Finally, the proof is complete by recalling that $\gamma = \mu/(1-\mu)$ so that $\gamma > \bar\gamma = \lambda/(1-\lambda) \Leftrightarrow \mu > \lambda$.

**Proof of Proposition (12):**

The proofs of Propositions (10) and (11) showed that whenever $\mu > \lambda$, $IC_{21}$ is binding both in the Mirrlees and SB with Adverse Selection case. But $IC_{21}$ in the Mirrlees case is more

stringent than in the SB with Adverse Selection: namely, for each $\{c_i, h_i\}_{i=1}^2$, $c_1 - \phi_2\left(\frac{h_1\theta_1}{\theta_2}\right) >$ $c_1 - \phi_2(h_1)$. The set of incentive compatible allocations is hence smaller and welfare is lower. The exact opposite applies when $\mu < \lambda$, as then $IC_{12}$ – which is more stringent in the SB with Adverse Selection – is binding.

**Proof of Proposition (13):**

The problem is reformulated conditional on the set of utilities to be provided to types lower than $N$, $\mathbf{u} = \{u_i\}_{i=1}^{N-1}$. Multipliers are in brackets.

$$\left(P^{SB,N}\left(\mathbf{u}\right)\right) : \max_{\{c_i, h_i\}} c_N - \phi_N\left(h_N\right)$$

$$(IC_{i,i+1}) : c_i - \phi_i\left(h_i\right) \geq c_{i+1} - \phi_i\left(h_{i+1}\right) \quad \left[\beta_{i,i+1}\right] \quad i = 1, .., N-1$$

$$(IC_{i+1,i}) : c_{i+1} - \phi_{i+1}\left(h_{i+1}\right) \geq c_i - \phi_{i+1}\left(h_i\right) \quad \left[\beta_{i+1,i}\right] \quad i = 1, .., N-1$$

$$(RC) : \sum_i \lambda_i c_i \leq \sum_i \lambda_i \theta_i h_i \quad [\delta]$$

$$c_i - \phi_i\left(h_i\right) \geq u_i \quad [\gamma_i] \quad i = 1, .., N-1$$

The weights from the main text can be mapped into the multipliers of the utility constraints using: $\mu_i = \gamma_i / \sum_{j=1}^N \gamma_j$, and the normalization $\gamma_N = 1$. Note that if the Pareto frontier is a linear hyperplane along some dimensions in some regions, then the same set of Pareto weights could correspond to several different threshold utilities $\mathbf{u}$. The FOCs are:

$$[c_i] \quad : \quad \beta_{i,i+1} + \beta_{i,i-1} - \beta_{i-1,i} - \beta_{i+1,i} - \lambda_i\delta + \gamma_i = 0$$

$$[h_i] \quad : \quad -\phi_i'\left(h_i\right)\beta_{i,i+1} - \phi_i'\left(h_i\right)\beta_{i,i-1} + \beta_{i-1,i}\phi_{i-1}'\left(h_i\right) + \phi_{i+1}'\left(h_i\right)\beta_{i+1,i} + \lambda_i\theta_i\delta - \phi_i'\left(h_i\right)\gamma_i = 0$$

$$[c_N] \quad : \quad 1 + \beta_{N,N-1} - \beta_{N-1,N} - \lambda_N\delta = 0$$

$$[h_N] \quad : \quad -\phi_N'\left(h_N\right) - \beta_{N,N-1}\phi_N'\left(h_N\right) + \phi_{N-1}'\left(h_N\right)\beta_{N-1,N} + \lambda_N\theta_N\delta = 0$$

$$[c_1] \quad : \quad \beta_{1,2} - \beta_{2,1} - \delta\lambda_1 + \gamma_1 = 0 \qquad [h_1] : -\phi_1'\left(h_1\right)\beta_{1,2} + \phi_2'\left(h_1\right)\beta_{2,1} + \delta\lambda_1\theta_1 - \phi_1'\left(h_1\right)\gamma_1 = 0$$

**Lemma 2** *In the Second Best with Adverse Selection, if $\sum_{i=1}^j \mu_i > \sum_{i=1}^j \lambda_i$, $\forall j \leq (N-1)$, all downward incentive compatibility constraints $IC_{i+1,i}$ are binding, and all upward incentive compatibility constraints $IC_{i,i+1}$ are slack for all $i \leq N-1$.*

**Proof.** Given the mapping from Pareto weights to multipliers, the condition $\sum_{i=1}^j \mu_i >$

$\sum_{i=1}^{j} \lambda_i$, $\forall i \leq N - 1$ corresponds to:

$$\sum_{k=i}^{N} \lambda_k > \sum_{k=i}^{N} \frac{\gamma_k}{\sum_{j=1}^{N} \gamma_j}, \quad \forall i \geq 2 \tag{1.23}$$

Suppose that, for all $i \geq 2$, condition (1.23) holds. First, let us show that there cannot be any upward binding constraint. Start from $i = N - 1$, and suppose by contradiction that constraint $IC_{N-1,N}$ is binding, so that $\beta_{N-1,N} > 0$, and $\beta_{N,N-1} = 0$ (since we assumed that pooling is not optimal). The FOC for $c_{N-1}$ would imply that $\beta_{N-1,N} - \beta_{N-2,N-1} = \lambda_{N-1}\delta - \gamma_{N-1}$ (with $\beta_{N-2,N-1}$ either strictly positive or zero), while the FOC for $c_N$ implies that: $1 - \beta_{N-1,N} = \lambda_N \delta$. Adding these two expressions yields:

$$-\beta_{N-2,N-1} = \lambda_N \delta + \lambda_{N-1}\delta - \gamma_{N-1} - 1 \tag{1.24}$$

But by the assumption on the parameters in (1.23), $\lambda_N\delta + \lambda_{N-1}\delta - \gamma_{N-1} - 1 > 0$, which implies $\beta_{N-2,N-1} < 0$, a contradiction. Hence, $\beta_{N-1,N} = 0$.

Proceeding recursively, consider agent $N - 2$ and suppose that constraint $IC_{N-2,N-1}$ binds, so that $\beta_{N-2,N-1} > 0$ and $\beta_{N-1,N-2} = 0$. The FOC for $c_{N-2}$ then implies that $\beta_{N-2,N-1} - \beta_{N-3,N-2} - \lambda_{N-2}\delta + \gamma_{N-2} = 0$ with $\beta_{N-3,N-2}$ either strictly positive or zero. The FOC for $c_{N-1}$ implies $-\beta_{N-2,N-1} - \beta_{N,N-1} - \lambda_{N-1}\delta + \gamma_{N-1} = 0$. The FOC for $c_N$ implies: $\beta_{N,N-1} - \lambda_N\delta + \gamma_N = 0$. Adding these three expressions, we get: $-\beta_{N-3,N-2} = \left(\lambda_N\delta + \lambda_{N-2}\delta + \lambda_{N-1}\delta - \gamma_N - \gamma_{N-1} - \gamma_{N-2}\right) > 0$ (by condition (1.23)). Hence, $\beta_{N-3,N-2} < 0$, a contradiction. We can continue in this fashion up to type 1 to show that no constraint of the form $IC_{i,i+1}$ binds, hence $\beta_{i,i+1} = 0$.

To show that the downward constraints are not slack but binding, let us now show that it is not optimal to have both $IC_{i,i+1}$ and $IC_{i+1,i}$ slack for some $i$. Start from agent $i = N - 1$ and suppose that $IC_{N-1,N}$ and $IC_{N,N-1}$ are both slack, so that $\beta_{N-1,N} = \beta_{N,N-1} = 0$. Then, the FOC for $c_N$ implies that $1 = \lambda_N\delta$, which violates the strict inequality in (1.23). Continuing recursively, suppose that $IC_{N-2,N-1}$ and $IC_{N-1,N-2}$ are both slack. Then, we can decrease $c_N$ and $c_{N-1}$ by the same small amount $dc > 0$ (leaving constraint $IC_{N,N-1}$ unaffected) and increase all $c_i$ for $i \leq N - 2$ by the same amount $dc'$ such that the resource constraint is

unaffected (this leaves all incentive constraints for types below $N - 2$ unaffected as well):

$$dc' \left( \lambda_1 + .... + \lambda_{N-2} \right) = dc \left( \lambda_{N-1} + \lambda_N \right)$$

The change in welfare from this resource neutral transfer is:

$$dc \left( - \left( \gamma_N + \gamma_{N-1} \right) + \frac{\left( \lambda_{N-1} + \lambda_N \right)}{\left( \lambda_1 + .... + \lambda_{N-2} \right)} \left( \gamma_1 + ... + \gamma_{N-2} \right) \right)$$

Which is positive, from the assumption on parameters in (1.23). ∎

**Lemma 3** *In the Second Best with Adverse Selection, if $\sum_{i=1}^{j} \mu_i < \sum_{i=1}^{j} \lambda_i \; \forall j \leq (N - 1)$, all upward incentive compatibility constraints $IC_{i,i+1}$ are binding, and all downward incentive compatibility constraints $IC_{i+1,i}$ are slack for all $i \leq N - 1$.*

**Proof.** The proof is symmetric to the one above, starting from the opposite strict inequality than in (1.23) and proceeding recursively from type $i = 1$, using the condition on multipliers:

$$\sum_{k=i}^{N} \lambda_k < \sum_{k=i}^{N} \frac{\gamma_k}{\sum_{j=1}^{N} \gamma_j}, \; i \geq 2 \tag{1.25}$$

∎

**Lemma 4** *In the Mirrlees regime, if $\sum_{i=1}^{j} \mu_i > \sum_{i=1}^{j} \lambda_i \; \forall j \leq (N - 1)$, all constraints $IC_{i+1,i}$ are binding and all constraints $IC_{i,i+1}$ are slack for all $i \leq N - 1$.*

*If $\sum_{i=1}^{j} \mu_i < \sum_{i=1}^{j} \lambda_i \; \forall j \leq (N - 1)$, all constraints $IC_{i,i+1}$ are binding and all constraints $IC_{i+1,i}$ are slack for all $i \leq N - 1$.*

**Proof.** The program of the Planner with N types can also be reformulated as maximizing the utility of type $N$, conditional on the utilities of other types being above some thresholds, and subject to the same $(IC_{i,i+1})$, $(IC_{i+1,i})$, and $(RC)$ as in $\left( P^{Mirr,N} \left( \mu \right) \right)$ in the text. The proof is then exactly as for Lemmas 2 and 3, since the only thing that differs between the Mirrlees and the Second Best with Adverse Selection cases is how hours of work enter the incentive compatibility constraints, but the aforementioned proofs only used the FOCs with respect to consumption levels $\{c_i\}_{i=1}^{N}$. ∎

Thus, when condition (1.23) holds, the downward incentive compatibility constraints ($IC_{i+1,i}$ $\forall i \leq N - 1$) are binding both in the Mirrlees and Adverse Selection case. But these constraints are more stringent in $\left(P^{Mirr,N}\left(\boldsymbol{\mu}\right)\right)$ than in $\left(P^{SB,N}\left(\boldsymbol{\mu}\right)\right)$. Namely, for each $\{c_i, h_i\}$, $\forall i$, $\phi_i\left(\frac{h_{i-1}\theta_{i-1}}{\theta_i}\right) < \phi_i\left(h_{i-1}\right)$. The incentive compatible set of allocations is hence smaller and welfare is lower. Inversely, when condition (1.25) holds, the upward constraints ($IC_{i,i+1}, \forall i \leq N - 1$) are binding, and are more stringent in $\left(P^{SB,N}\left(\boldsymbol{\mu}\right)\right)$ than in $\left(P^{Mirr,N}\left(\boldsymbol{\mu}\right)\right)$.

**Proof of Proposition (14) :**

i) Any second best allocation $\{h_i, c_i\}_{i=1}^N$ for which the upward incentive compatibility constraints ($IC_{i,i+1}, \forall i \leq N - 1$) are binding can be implemented by assigning $y_i = \theta_i h_i$ and $T_i = c_i - y_i$ (and prohibitively high tax levels on all other incomes $y \notin \{y_i\}_{i=1}^N$). The maximization program of firms then becomes the same as the government's and, since the Second Best allocation was optimal, there is no possible deviation which could make some type better off without violating the ICs. Hence, welfare in this region is equal to welfare in the Second Best case, and we showed in Proposition 13 that in this region, the Second Best frontier is below the Mirrlees frontier.

ii) The proof proceeds by finding a lower bound for the Pareto frontier when $\sum_{j=1}^i \mu_j > \sum_{j=1}^i \lambda_j \ \forall i \leq (N - 1)$. In this case, Proposition 13 showed that in the Second Best all downward incentive constraints are binding. We already know that, if the only ICs are those from $\left(P^{SB,N}\right)$, the Pareto frontier is above the Mirrlees one in that region. The incentive compatibility constraints in $\left(P^{SB,N}\right)$ are still necessary with unobserved contracts. Are the additional constraints needed to prevent firms from deviating, if any, weaker than those in $\left(P^{Mirr,N}\right)$?

Suppose that the government artificially strengthens constraint (1.13) to $y_i = \theta_i h_i \ \forall i$, limiting its choice variables to only $h_i$ and $T_i$. Starting from $i = N$, we will now rule out all deviations which involve type $i$ being offered a new contract (together with a pool of types), in which he is either cross-subsidizing other deviating agents or earning exactly his product. By doing this for all $i$, no $i$ can be attracted to a deviating contract in which he is cross-subsidized, since the types made to cross-subsidize him will never join any such deviation.

Start with agent $N$ and suppose firms try to attract him to a pool with some subset $A_k$ of workers at income level $y^k = y_j = \theta_j h_j (< y_N)$ for some of the available income levels $y_j$. This requires hours of work of at least $h_{k,A_k} \geq y_j/\bar{\theta}_{A_k} = \left(\theta_j/\bar{\theta}_{A_k}\right) h_j$. In the Mirrlees case, on the

other hand, $N$ would have had to work only $y_j/\theta_N = (\theta_j/\theta_N)\, h_j < y_j/\bar{\theta}_{A_k} \leq h_{k,A_k} \; \forall A_k$, for the same pay. Thus, ruling out even the most attractive (non-loss making) pool $\{A_k, y_j, h_{k,A_k}\}$ for type $N$ is strictly easier than to rule out his most attractive deviation in the Mirrlees case: the incentive compatibility constraint for type $N$ has to be strengthened relative to $IC_{N,N-1}$ in program $P^{SB,N}$, but it will never have to be strengthened as much as to become stricter than $IC_{N,N-1}$ in $P^{Mirr,N}$.

Continue with agent $N-1$. Given that we have ruled out even the most attractive deviation for agent $N$, any deviation offered to agent $N-1$ must have him as the highest type in any pool he is part of. Again, no matter at which income level $y_m$ the pool occurs, a deviation which is not cross-subsidized by another contract (which we are ruling out for each type) cannot be more profitable than in the Mirrlees case, since agent $N-1$ will necessarily be pooled with lower types and his pay per hour diluted to some $\bar{\theta}_{A_k}$ for some $A_k$. Continuing recursively this way, we see that for every desired allocation $\{c_i, h_i\}_i$, the downward binding IC for each $i$ will be easier to satisfy than in the Mirrlees case. Removing the artificially imposed constraint $y_i = \theta_i h_i$ will then allow the government to reach even higher social welfare. Hence, *a fortiori*, welfare will be higher than in the Mirrlees case.

**Proof of Proposition (15):**

Suppose the government wants to implement the second best consumption and hour levels, $(\{h_i, c_i\})_{i=1}^{2}$, which are characterized by (using the same notation as in the previous section):

$$c_2 = \lambda\theta_1 h_1 + (1-\lambda)\,\theta_2 h_2 + \lambda\left(\phi_2\left(h_2\right) - \phi_2\left(h_1\right)\right)$$

$$c_1 = \lambda\theta_1 h_1 + (1-\lambda)\,\theta_2 h_2 - (1-\lambda)\left(\phi_2\left(h_2\right) - \phi_2\left(h_1\right)\right)$$

$$h_2 = h_2^*, \quad \left(\gamma - \lambda\gamma - \lambda\right)\left(\phi_2'\left(h_1\right) - \theta_1\right) + \gamma\theta_1 = \gamma\phi_1'\left(h_1\right)$$

Take any arbitrarily assigned income levels $(y_1, y_2)$, such that firms' break-even constraints holds. There are no profitable deviations attracting only the low type, who is already weakly subsidized in the second-best allocation; any such contract making him better off would make a loss. It is also not possible to attract both types to a pooling contract at income level $y_2$. If it were, then this contract would have made both types better off, be budget feasible for the government, and, hence, would violate the Pareto optimality of the second best allocation. The

only deviations that a firm could make are hence:

1) Offer to pay $y_2$ for hours of work $h_2 = y_2/\theta_2$. This will be accepted by the high type if $y_2$, was originally not actuarially fair, i.e., $y_2 < \theta_2 h_2$. Under the MWS assumption, other firms will then drop the loss-making cross-subsidization contract. If the low type joins this new contract, it becomes unprofitable. Hence an equilibrium requires that $c_2 - \phi_1\left(\frac{y_2}{\theta_2}\right) \geq 0$ so the low type prefers joining the deviating contract rather than staying out of the market (at utility 0).

2) Pool both types at $y_1$ with hours $h'_1$ such that: $\bar{\theta}_{\{1,2\}}h'_1 = y_1$. The low type prefers this if $h'_1 \leq h_1$ (since he would be working less for the same pay). The constraint needed to render this deviation unprofitable is hence: $h_1 \leq \frac{y_1}{\bar{\theta}_{\{1,2\}}}$. In order to implement the second-best allocation, we thus only need to find two assigned income levels $y_1$ and $y_2$ such that (using that from the break even requirement, $y_1 = \frac{c}{\lambda} - \frac{(1-\lambda)}{\lambda}y_2$):

$$\lambda y_1 + (1 - \lambda)\, y_2 = \lambda\theta_1 h_1 + (1 - \lambda)\,\theta_2 h_2 = \lambda c_1 + (1 - \lambda)\, c_2 := c \qquad (1.26)$$

$$0 \leq c_2 - \phi_1\left(\frac{y_2}{\theta_2}\right) \qquad (1.27)$$

$$h_1 \leq \frac{\frac{c}{\lambda} - \frac{(1-\lambda)}{\lambda}y_2}{\bar{\theta}_{\{1,2\}}} \qquad (1.28)$$

where $c := \lambda c_1 + (1 - \lambda)\, c_2$.[32] Thus, it is sufficient to find a $y_2$ such that:

$0 \leq y_2 \leq \min\left\{\phi_1^{-1}(c_2)\,\theta_2, \frac{c}{(1-\lambda)} - \frac{h_1\bar{\theta}_{\{1,2\}}}{(1-\lambda)}\lambda\right\}$. Such a level will exist if and only if $\frac{c}{(1-\lambda)} - \frac{h_1\bar{\theta}_{\{1,2\}}}{(1-\lambda)}\lambda \geq 0$, or alternatively, if $c \geq h_1\bar{\theta}_{\{1,2\}}\lambda$. Using the resource constraint, this requires $\lambda\theta_1 h_1 + (1 - \lambda)\,\theta_2 h_2 = \lambda c_1 + (1 - \lambda)\, c_2 \geq h_1\lambda\,(\lambda\theta_1 + (1 - \lambda)\,\theta_2) \Leftrightarrow h_1\theta_1\lambda + \theta_2\,(h_2 - h_1\lambda) \geq 0$, which is always true since $h_2 \geq h_1$ (and hence $h_2 \geq \lambda h_1$) in the Second Best.

**Proof of Proposition (16):**

*Illustration with $N = 3$*

The income level for type 3, $y_3$, must be such that firms are not tempted to cream-skim type 3. This requires that, if it occurs, and the unprofitable contracts of type 1 and 2 are dropped in response, type 2 (at least) joins the new contract, i.e., $\theta_3\phi_2^{-1}(c_3) > y_3$. Income level $y_2$ must be such that no pooling of 2 and 3 can occur, either because 1 would then

---

[32]To ensure that $y_1$ is non-negative, we need $y_2 \leq \frac{c}{1-\lambda}$ but this is guaranteed by the third constraint, $y_2 \leq \frac{c}{(1-\lambda)} - \frac{h_1\bar{\theta}_{\{1,2\}}}{(1-\lambda)}\lambda \leq \frac{c}{(1-\lambda)}$.

join the pool and a pool with 1 would not be profitable for 3 to join (i.e., $\bar{\theta}_{\{2,3\}} \phi_1^{-1}(c_2) > y_2 > \bar{\theta}_{\{1,2,3\}} h_2$), or because the pooling would not attract type 3 in the first place, i.e., $y_2 > h_2 \bar{\theta}_{\{2,3\}}$. Income level $y_1$ must be such that pooling 1 and 2 at 1 is not profitable, and pooling all 3 is impossible because 3 would not join such a pool. Hence, we need $y_1 > \max\left\{h_1 \bar{\theta}_{\{1,2\}}, \bar{\theta}_{\{1,2,3\}} \phi_3^{-1}(\phi_3(h_2) - (c_2 - c_1))\right\}$. Finally, all income levels must be non-negative, and firms must break even on average: $\sum_{i=1}^3 \lambda_i y_i = \sum_{i=1}^3 \lambda_i \theta_i h_i$. In general, we cannot ensure that there are non-negative income levels $\{y_i\}_{i=1}^N$ which will satisfy all these constraints.

*General implementation with $N \geq 2$ :*

We now focus on the implementation of Second Best allocations $\{c_i, h_i\}_{i=1}^N$, in which the downward ICs are binding. First, suppose that type $m > i$ is attracted by a deviating contract, paying $y_i$ for $y_i / \bar{\theta}_{\{i,m\}}$ hours of work. Using the binding $(IC_{m,m-1})$, we have $c_i - \phi_m\left(y_i / \bar{\theta}_{\{i,m\}}\right) > c_m - \phi_m(h_m) = c_{m-1} - \phi_m(h_{m-1})$. But, by the Spence-Mirrlees single crossing condition in 1, if type $\theta_m$ prefers the allocation with less work and less consumption $\left\{c_i, y_i / \bar{\theta}_{\{i,m\}}\right\}$ to the one with more work and consumption $\{c_{m-1}, h_{m-1}\}$, then so must type $m - 1$. Hence, if $m$ is attracted by this deviation, so is $m - 1$. Repeating this argument iteratively, all types $i, .., m - 1$ will be attracted if $m$ is. Thus, we only need to consider connected intervals from $i$ to $m$, for some $m > i$. For each $y_i$, pick the type for whom a deviation to $y_i$ would be most attractive, and set $y_i$ such that the deviation is not preferred to the type's own allocation:

$$y_i > \max_{m \geq i} \bar{\theta}_{\{i,..,m\}} \phi_m^{-1}(c_i - c_m + \phi_m(h_m)) \quad \forall i < N$$

The constraint on $y_i$ implies that there is no profitable pool that could attract all workers of types $i$ through $m$ to income level $y_i$, for any $m$, even if they were just made to work sufficient hours for the firm to break even. By extension, this also implies that no contract could be offered which allowed to cross-subsidize other contracts. For type $N$, $y_N$ must be set sufficiently low, so that it would be attractive for type $N - 1$ to join if type $N$ was offered an actuarially fair contract with $y_N / \theta_N$ hours of work for a pay $y_N$, i.e., we need $y_N < \phi_{N-1}^{-1}(c_N) \theta_N$ (this assumes that $N$ is not alone in a cross-subsidization group. If $N$ is disjoint from other types, there is no profitable deviation for firms to start with). Given these income levels, income taxes are set according to $c_i = y_i - T_i > 0$, $\forall i$, and to 100% for income levels not in the recommended set.

The income levels thus specified are potentially very large, and would cause losses for the firms overall. The government can rebate the losses or tax away the profits from each individual contract, by setting a payroll tax schedule $\left\{T_i^F\right\}_{i=1}^N$ such that $T_i^F = T^F(y_i) = \theta_i h_i - y_i$, and $T^F(y) = 2y$ for $y \notin \{y_i\}_{i=1}^N$.

## 1.B Rothschild and Stiglitz Equilibrium Type

I now consider the alternative equilibrium definition of Rothschild and Stiglitz (RS), but limit myself to the $N = 2$ model, because of the existence problems with $N > 2$. In RS, firms offer only a single labor contract and have to break even.[33]

**Definition 2** *(Rothschild-Stiglitz equilibrium) A set of contracts offered is a Rothschild-Stiglitz equilibrium if i) firms make zero profit on each contract and ii) there is no other potential contract which would make positive profits, if offered.*

The Rothschild and Stiglitz (1976) equilibrium notion has been used almost exclusively for two type models $(N = 2)$. The authors show that no pooling equilibrium can exist, and that, for a sufficiently low fraction of low productivity workers, $\lambda_1$, no equilibrium can exist at all. Whenever a separating equilibrium exists, low types work an efficient number of hours, but high types work excessively, so that firms can separate them from low types. Hence, high ability workers are caught in a "rat race" (Akerlof, 1976).

### 1.B.1 Linear Taxation

In a RS-type equilibrium, firms are constrained to break even on each contract offered, so that, at any potential equilibrium allocation, pay is equal to a worker's total product: $y_i = (1 - t_i)\theta_i h_i$.[34] Whenever a separating equilibrium exists, the tax formula will be the same as in case $AS1$. However, an equilibrium may not exist, because it might be possible for some firm to offer an alternative pooling contract which would attract all workers. The existence of an equilibrium depends on taxes. Rather than a smooth response of the private market, there may

---

[33]Since firms need to break even on each contract, they might as well be offering a pair of contracts. But it is without loss of generality to assume that those contracts are offered by two different firms.

[34]Recall that a pooling equilibrium cannot exist in RS, even in the presence of linear taxes.

be an abrupt shift, at some tax level, from a separating equilibrium to non-existence, which makes the optimal tax problem more complicated.

One can check that a separating equilibrium $\left(h_2^{RS}, y_2^{RS}\right)$ will be stable,[35] if the fraction of low types is larger than $\lambda^{RS}(t)$, defined as the threshold value of $\lambda$ for which the indifference curve of the high type worker going through the equilibrium allocation is just tangent to the pooling line with slope $\theta^{RS}(t) = \left(\lambda^{RS}(t)\theta_1 + \left(1 - \lambda^{RS}(t)\right)\theta_2\right)$. Denote this tangency point by $h(t)$. In this case, there is no possible pooling region, that is, no allocations which are pooling, make non-negative profits, and make both workers better off than the separating equilibrium allocation. If, at a given tax $t$, $\lambda$ is already close to the threshold $\lambda^{RS}(t)$ and the government increases the tax, then this *could* open up a possible pooling region and destroy the equilibrium altogether. This phenomenon is illustrated in Figure $1 - 4$ where, starting from a situation with low taxes (solid indifference curves), taxes are increased (dotted indifference curves) and a pooling region is created. The following proposition describes when this can occur.

**Proposition 17** *With a general utility function, if the utility of the high type at the candidate separating equilibrium, denoted $u_2^{RS}$, is sufficiently strongly decreasing in taxes, that is if*

$$du_2^{RS}/dt \leq -\theta^{RS}(t)h(t)$$

*then raising taxes could destroy an existing separating equilibrium (in the sense of pushing up the critical threshold $\lambda^{RS}(t)$).*

*With an isoelastic utility function, the critical threshold $\lambda^{RS}(t)$ is always increasing in $t$ and higher taxes make the existence of a separating equilibrium less likely.*

If nonexistence of an equilibrium is an undesirable state, then, the optimal tax rate must be set subject to the additional constraint $\lambda^{RS}(t) \leq \lambda$.[36]

---

[35] Stability means that the separating equilibrium cannot be broken by a pooling equilibrium, which is the definition used in the original Rothschild and Stiglitz (1976) paper.

[36] The detailed analysis of this problem is not particularly enlightening, given already performed for the MWS case.

Figure 1-4: Non Existence of the RS equilibrium
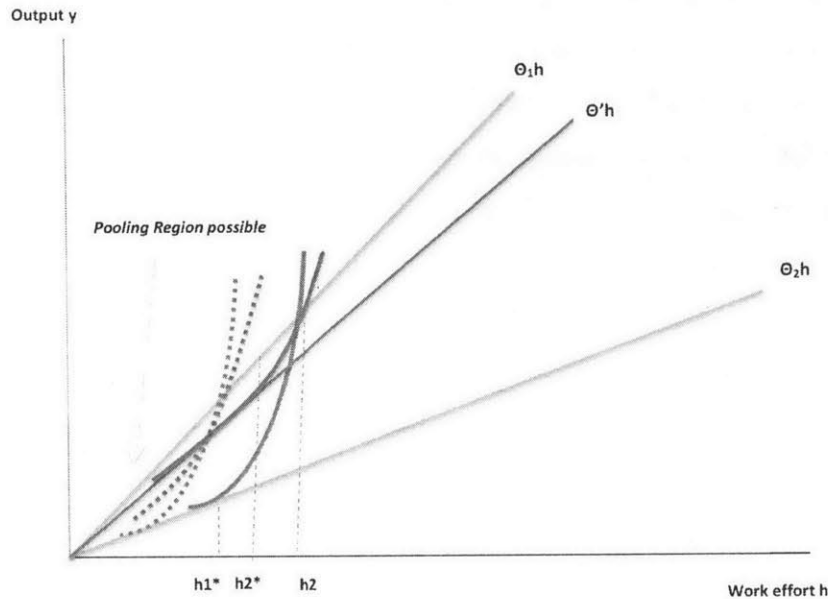


## 1.B.2 Nonlinear Taxation

Is it still possible to implement all Second Best allocations in an RS setting like it was for the MWS setting with $N = 2$ (i,e, regime 3 in the main text, called "Adverse Selection with unobservable private contracts"), despite the government being unable to see private labor contracts?[37] We focus on the more interesting case in which the high productivity worker's incentive constraint would be binding in the Second Best with Adverse Selection and show that this allocation can no longer be an equilibrium with unobservable private contracts.

Suppose again, that the government imposes confiscatory taxes on all income levels other than the recommended ones, $y_1$ and $y_2$. This reduces to only three the potential deviations that firms can make. First, they could try to pool both workers at $y_2$. However, no such contract would attract high types, since they would have to work more for the same pay than under the original contract, to compensate for the low type's poor productivity. Secondly, firm could not possibly "invert" the separating equilibrium by offering contracts $\{y_1, y_1/\theta_2\}$ and $\{y_2, y_2/\theta_1\}$, because that would violate the monotonicity condition on hours. There is however one other

---

[37] Again, any Second Best allocation $\{c_i, h_i\}_{i=1}^{2}$ at which the incentive compatibility constraint of the low type is binding can still be implemented. It is sufficient to set taxes equal to $T_i = y_i - c_i$ at income levels $y_i = \theta_i h_i$, $i = 1, 2$, and $T = y$ for all other income levels.

profitable deviation, namely to offer a pooling contract at $y_1$. To see this, start from the candidate allocation at which the incentive compatibility constraint of the high type is binding:

$$y_2 - T_2 - \phi_2(h_2) = y_1 - T_1 - \phi_2(h_1) \tag{1.29}$$

Consider a pooling contract paying $y_1$ in exchange for $h_1'$ hours of work, where $h_1'$ is determined by the zero profit condition:

$$\lambda \theta_1 h_1' + (1 - \lambda) \theta_2 h_1' = y_1 \tag{1.30}$$

There then exists a slightly higher level of hours, $h^* = h_1' + \varepsilon$ (for some very small $\varepsilon > 0$), such that $\lambda \theta_1 h^* + (1 - \lambda) \theta_2 h^* > y_1$, and a new contract $(y_1, h^*)$ which would yield strictly positive profits if both types accepted it. And indeed, both types will accept it. By (1.29), the high type was just indifferent between his allocation and the original allocation of the low type, $(y_1, h_1)$. Furthermore, by (1.30), it is clear that $h_1' < h_1$, since $\theta_1 h_1 = y_1$ so that $\theta_2 h_1 > y_1$. If $\varepsilon$ is small enough, we also have $h^* < h_1$. Thus, the high type now strictly prefers the allocation $(y_1, h^*)$ to $(y_2, h_2)$. The low type also prefers this allocation, since he earns the same total pay but works less. The original allocation can thus not have been an equilibrium.[38] Intuitively, the government is trying to force the private market to do the opposite of what it would normally do, namely to reduce the welfare of the high types for the benefit of the low types. But competition among firms makes them exploit the loophole, created by the inability of the government to see hours worked, to try to make the high type as well off as possible. Therefore, we need to add a new incentive constraint for *firms* to not be able to deviate to that pooling contract, which is more stringent than the standard incentive constraint for the high type:

$$y_2 - T_2 - \phi_2\left(\frac{y_2}{\theta_2}\right) \geq y_1 - T_1 - \phi_2(h_1') \tag{1.31}$$

The Pareto frontier for the RS case is thus obtained by solving program $P^{RS}(\mu)$ and is char-

---

[38] Note that this pooling allocation to which firms are tempted to deviate might not be an equilibrium either, because of the familiar cream-skimming argument which also precludes the existence of a pooling equilibrium in the original Rothschild and Stiglitz paper. But it still is a profitable deviation. Indeed, if it was an equilibrium and it was making both types better off, it would have been on the Pareto frontier.

acterized in the next proposition.

$$\left(P^{RS}\left(\mu\right)\right) \quad : \quad \max_{c_1,c_2,h_1,h_2} \mu\left(c_1 - \phi_1\left(h_1\right)\right) + \left(1 - \mu\right)\left(c_2 - \phi_2\left(h_2\right)\right)$$

$$\left(IC_{12}\right) : c_1 - \phi_1\left(h_1\right) \geq c_2 - \phi_1\left(h_2\right)$$

$$\left(IC_{21}\right) : c_2 - \phi_2\left(h_2\right) \geq c_1 - \phi_2\left(\frac{\theta_1 h_1}{\left(1 - \lambda\right)\theta_2 + \lambda\theta_1}\right)$$

$$\left(RC\right) : \lambda c_1 + \left(1 - \lambda\right)c_2 \leq \lambda\theta_1 h_1 + \left(1 - \lambda\right)\theta_2 h_2$$

**Proposition 18** *The Pareto frontier in the RS setting is characterized by:*

*Region 1: When $\mu = \lambda$, both $(IC_{12})$ and $(IC_{21})$ are slack and both workers work efficient hours. The Pareto frontier is linear.*

*Region 2: When $\mu > \lambda$, $(IC_{21})$ is binding, the high type works efficient hours, the low type works too little, and the Pareto frontier is strictly concave. The Rothschild-Stiglitz frontier is above the Mirrlees frontier, but below the Second Best with Adverse Selection frontier.*

*Region 3: When $\mu < \lambda$, $(IC_{12})$ is binding, the low type works efficient hours, the high type works too much, and the Pareto frontier is strictly concave. The Rothschild-Stiglitz frontier is below the Mirrlees frontier and coincides with the Second Best with Adverse Selection frontier.*

Again, the three Regions can be mapped into three regions for the utility of the low type, $u$, with the four thresholds defined in the Appendix, $u_{\min}^{RS} < \underline{u}'' < \overline{u}'' < u_{\max}^{RS}$, and whose interpretation is exactly as in the main text. Region 1 corresponds to $\underline{u}'' \leq u \leq \overline{u}''$, Region 2 to $\overline{u}'' \leq u \leq u_{\max}^{RS}$, and Region 3 to $u_{\min}^{RS} \leq u \leq \underline{u}''$. In the proofs, I provide a ranking of all four frontiers with two types (the RS setting, the Mirrlees case, the Second Best with Adverse Selection case, and the Adverse Selection with unobservable private contracts case). Figure $1 - 5$ illustrates this relation for when condition **NL1** holds.

It is clear that, while adverse selection still helps the government redistribute, the Rothschild-Stiglitz case lies between the Second Best with Adverse Selection and the Mirrlees case. Firms try to attract the high type worker to a pooling allocation, by offering him an hourly wage equal to the average productivity $\lambda\theta_1 + \left(1 - \lambda\right)\theta_2$. This makes a deviation more attractive than in the Second Best, but still less attractive than in the Mirrlees case. Note the difference to regime 3 with $N = 2$: it is the fact that firms are constrained to break even on each contract

78

Figure 1-5: Rothschild-Stiglitz, Mirrlees, and Second-Best with Adverse Selection Frontiers
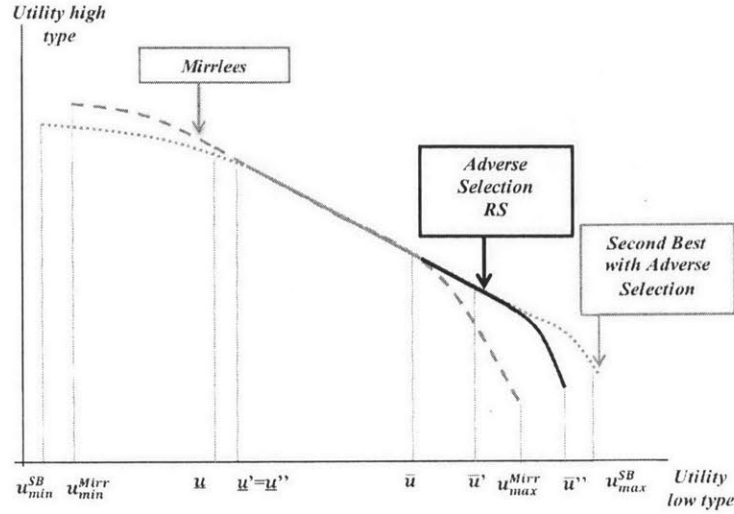


offered which prevents the government from implementing the Second Best allocation.[39]

## 1.B.3 Proofs

**Proof of Proposition (17):**

At any intersection of the indifference curve with the pooling line $y = \theta h$ we should have $\left(u_2^{RS} + \phi_2\left(h\right)\right) / \left(1 - t\right) = \theta h$. In addition to obtain an exact tangency, the slopes must be equal so that: $\phi_2'\left(h_2\right) / \left(1 - t\right) = \theta$.

*Special case: power disutility function* (under assumption 2) $\phi_i\left(h_i\right) = a_i h^\eta$.

Then: $u_1^{RS} = \left[\theta_1\left(1 - t\right) / a_1\eta\right]^{\frac{\eta}{\eta - 1}}\left[a_1\eta - 1\right]$ is the utility of the low type at the candidate $RS$ allocation, (assume that $a_1\eta \geq 1$). Then by the $(IC_{12})$ being binding:

$$
\begin{aligned}
u_2^{RS} &= u_1^{RS} + \left(h_2^{RS}\right)^\eta\left(a_1 - a_2\right) \\
u_2^{RS} &= \left[\frac{\theta_1\left(1 - t\right)}{a_1\eta}\right]^{\frac{\eta}{\eta - 1}}\left[a_1\eta - 1\right] + \left(h_2^{RS}\right)^\eta\left(a_1 - a_2\right)
\end{aligned}
$$

---

[39] As a side remark, it is interesting to consider what happens if instead of having the Rothschild-Stiglitz Nash Equilibrium behavior, firms exhibit a behavior characterized by Wilson's (1976) foresight assumption. Under this assumption, firms will consider a deviation to another contract, if and only if that deviation still remains profitable after all contracts which have been rendered unprofitable by it have been dropped. The Wilson notion can be used to justify why a pooling equilibrium may persist in the market. However, it is straightforward to check that it does not alter the Pareto frontier characterization. Indeed, the deviation to pooling at income level $y_1$ is still profitable, unless the appropriate incentive constraints from the RS setting hold.

The intersection and tangency conditions become: $\eta a_2 h^{\eta-1} = (1-t)\theta$ and $u_2^{RS} + a_2 h^{\eta} = (1-t)\theta h$. We can rewrite: $h = \left(u_2^{RS}/(\eta-1)a_2\right)^{\frac{1}{\eta}}$ and solve for $h$:

$$h = \left(\frac{\left[\frac{\theta_1(1-t)}{a_1\eta}\right]^{\frac{\eta}{\eta-1}}[a_1\eta - 1] + \left(h_2^{RS}\right)^{\eta}(a_1 - a_2)}{(\eta-1)a_2}\right)^{\frac{1}{\eta}}$$

To obtain the threshold $\lambda^{RS}(t)$, use that: $\theta = \eta a_2 h^{\eta-1}/(1-t)$, plug in the value for $h$ which leads to, after some algebra:

$$\lambda^{RS}(t) = -\frac{\theta_2}{(\theta_1 - \theta_2)} + \frac{1}{(\theta_1 - \theta_2)}\frac{\eta a_2}{(1-t)}\left(\frac{\left[\frac{\theta_1(1-t)}{a_1\eta}\right]^{\frac{\eta}{\eta-1}}[a_1\eta - 1] + \left(h_2^{RS}\right)^{\eta}(a_1 - a_2)}{(\eta-1)a_2}\right)^{\frac{\eta-1}{\eta}}$$

Taking the derivative of $\lambda^{RS}(t)$ with respect to $t$:

$$
\begin{aligned}
\frac{d\lambda}{dt} &= \frac{-1}{(\theta_1 - \theta_2)}\frac{\eta a_2}{(1-t)^2}\left(\frac{\left[\frac{\theta_1(1-t)}{a_1\eta}\right]^{\frac{\eta}{\eta-1}}[a_1\eta - 1] + \left(h_2^{RS}\right)^{\eta}(a_1 - a_2)}{(\eta-1)a_2}\right)^{\frac{\eta-1}{\eta}} \\
&+ \frac{1}{(\theta_1 - \theta_2)}\frac{\eta a_2}{(1-t)}\frac{\eta-1}{\eta} \\
&\times \left(\frac{1}{(\eta-1)a_2}\left(-\frac{\eta}{\eta-1}[a_1\eta - 1]\frac{\theta_1}{a_1\eta}\left[\frac{\theta_1(1-t)}{a_1\eta}\right]^{\frac{1}{\eta-1}} + (a_1 - a_2)\eta\frac{dh_2^{RS}}{dt}\left(h_2^{RS}\right)^{\eta-1}\right)\right) \\
&\times \left(\frac{\left[\frac{\theta_1(1-t)}{a_1\eta}\right]^{\frac{\eta}{\eta-1}}[a_1\eta - 1] + \left(h_2^{RS}\right)^{\eta}(a_1 - a_2)}{(\eta-1)a_2}\right)^{\frac{-1}{\eta}}
\end{aligned}
$$

all terms of which are positive. Hence, with isoelastic utility, the critical threshold for existence is increasing in taxes, which means that imposing taxes makes it harder for an equilibrium to exist and the government runs the risk of destroying the equilibrium.

*General case (no isoelastic utility):*

The hours of type 1 are efficient: $h_1^* = \left(\phi_1'\right)^{-1}((1-t)\theta_1)$. From $(IC_{12})$ binding, we get $u_2^{RS} = \theta_2(1-t)h_2 - \phi_2(h_2) = u_1^{RS} + \phi_1(h_2) - \phi_2(h_2)$. The condition for intersection of the indifference curve of the high type at the candidate separating allocation with the pooling line $y = \theta h$ is $u_2^{RS} + \phi_2(h) = (1-t)\theta h$, while the condition for tangency is $\phi_2'(h) = \theta(1-t)$.

Hence, combining these two gives a (implicit) solution for $h$:

$$u_2^{RS} + \phi_2(h) = \phi_2'(h)h \qquad (1.32)$$

$\lambda^{RS}(t)$ is the solution to: $\lambda^{RS} = \phi_2'(h) / (1-t)(\theta_1 - \theta_2) - \theta_2 / (\theta_1 - \theta_2)$ and taking the derivative: $d\lambda/dt = -\frac{dh}{dt}\phi_2''(h)(1-t) - \phi_2'(h) / (1-t)^2 (\theta_2 - \theta_1)$. Differentiate totally equation (1.32), to obtain $dh/dt = \left(du_2^{RS}/dt\right)\left(1/\left(\phi_2''(h)h\right)\right)$, so that $d\lambda/dt = \left(-\frac{du_2^{RS}}{dt}\frac{1}{h} - \theta\right) / \left[(1-t)(\theta_2 - \theta_1)\right]$. Hence, for $d\lambda/dt \geq 0$, we need: $\frac{du_2^{RS}}{dt} \leq -\theta h$, which is the condition in the main text.

**Proof of Proposition (18):**

The problem, indexed by the utility level of the low type is restated here (multipliers for each constraint are in brackets next to it):

$$\left(P^{RS}(u)\right) : \max_{c_1,c_2,h_1,h_2} c_2 - \phi_2(h_2)$$

$$(IC_{12}) : c_1 - \phi_1(h_1) \geq c_2 - \phi_1(h_2) \quad [\beta_1]$$

$$(IC_{21}) : c_2 - \phi_2(h_2) \geq c_1 - \phi_2\left(\frac{\theta_1 h_1}{(1-\lambda)\theta_2 + \lambda\theta_1}\right) \quad [\beta_2]$$

$$(RC) : \lambda c_1 + (1-\lambda)c_2 \leq \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2 \quad [\delta]$$

$$u \leq c_1 - \phi_1(h_1) \quad [\gamma]$$

Denote the set of admissible allocations, that is, those allocations which satisfy both incentive compatibility constraints and the resource constraint by $B^{RS}(u)$ and denote by $V_2^{RS}(u) = \max_{c_1,c_2,h_1,h_2 \in B^{RS}(u)}(c_2 - \phi_2(h_2))$. When the value of $u$ varies, the function $V_2^{RS}(u)$ traces out all possible values for the utility of the high type. The Pareto frontier is made of all pairs $\left(u, V_2^{RS}(u)\right)$ such that $\frac{\partial}{\partial u}V_2^{Mirr}(u) < 0$. The FOCs are:

$$[c_1] : 1 + \beta_2 - \beta_1 - \delta(1-\lambda) = 0$$

$$[c_2] : \beta_1 + \gamma - \beta_2 - \delta\lambda = 0$$

$$[h_1] : -\beta_1\phi_1'(h_1) + \frac{\theta_1}{(1-\lambda)\theta_2 + \lambda\theta_1}\beta_2\phi_2'\left(\frac{\theta_1}{(1-\lambda)\theta_2 + \lambda\theta_1}h_1\right) + \delta\lambda\theta_1 - \gamma\phi_1'(h_1) = 0$$

$$[h_2] : -(1+\beta_2)\phi_2'(h_2) + \beta_1\phi_1'(h_2) + \delta(1-\lambda)\theta_2 = 0$$

$$\lambda c_1 + (1-\lambda)c_2 = \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2, \quad c_1 - \phi_1(h_1) = u$$

**Region 1:** Suppose that $u \in [\underline{u}'', \overline{u}'']$ where the thresholds are defined by $\underline{u}'' := \underline{c_2} - \phi_1(h_2^*)$, where $\underline{c_2} := \lambda\theta_1 h_1^* + (1-\lambda)\theta_2 h_2^* + \lambda(\phi_1(h_2^*) - \phi_1(h_1^*))$ and $\overline{u}'' := \overline{c_2} - \phi_2(h_2^*) + \phi_2\left(\frac{h_1^*\theta_1}{\lambda\theta_1 + (1-\lambda)\theta_2}\right) - \phi_1(h_1^*)$ where $\overline{c_2} := \lambda\theta_1 h_1^* + (1-\lambda)\theta_2 h_2^* + \lambda\left(\phi_2(h_2^*) - \phi_2\left(\frac{\theta_1 h_1^*}{(1-\lambda)\theta_2 + \lambda\theta_1}\right)\right)$. Then it is possible to set the hours at their efficient levels, $h_1^*$ and $h_2^*$ and $c_1$ and $c_2$ such that:

$$\lambda c_1 + (1-\lambda)c_2 = \lambda\theta_1 h_1^* + (1-\lambda)\theta_2 h_2^*$$

$$c_1 - \phi_1(h_1^*) = u$$

and have both incentive constraints slack. Then, $\gamma = \frac{\lambda}{1-\lambda}$ and $\beta_1 = \beta_2 = 0$. In this region, the Pareto frontier is linear.

**Region 2:** Suppose $u \geq \overline{u}''$. A parallel reasoning to the one in the proof of Proposition (11) shows that in this region, $(IC_{21})$ is binding and $(IC_{12})$ is slack so that $\beta_1 = 0$. Then, the solution is characterized by:

$$\beta_2 = \gamma - \lambda\gamma - \lambda, \quad \gamma \geq \frac{\lambda}{(1-\lambda)}, \quad \gamma + 1 = \delta$$

$$\gamma\phi_1'(h_1) = (\gamma - \lambda\gamma - \lambda)\left(\frac{\theta_1}{(1-\lambda)\theta_2 + \lambda\theta_1}\phi_2'\left(\frac{\theta_1}{(1-\lambda)\theta_2 + \lambda\theta_1}h_1\right) - \theta_1\right) + \gamma\theta_1$$

$$\theta_2 = \phi_2'(h_2), \quad u = c_1 - \phi_1(h_1)$$

$$c_1 = \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2 - (1-\lambda)\left(\phi_2(h_2) - \phi_2\left(\frac{\theta_1 h_1}{(1-\lambda)\theta_2 + \lambda\theta_1}\right)\right)$$

$$c_2 = \lambda\theta_1 h_1 + (1-\lambda)\theta_2 h_2 + \lambda\left(\phi_2(h_2) - \phi_2\left(\frac{\theta_1 h_1}{(1-\lambda)\theta_2 + \lambda\theta_1}\right)\right)$$

There is a maximal level of utility for the low type achievable here, denoted $u_{\max}^{RS}$ which is the utility level attained when $h_1$ is the solution to $\max c_1 - \phi_1(h_1)$ subject to $(IC_{21})$, or equivalently, when $\gamma \to \infty$ in the above FOCs, that is when

$$\phi_1'(h_1) = (1-\lambda)\frac{\theta_1}{(1-\lambda)\theta_2 + \lambda\theta_1}\phi_2'\left(\frac{\theta_1}{(1-\lambda)\theta_2 + \lambda\theta_1}h_1\right) + \lambda\theta_1$$

A completely symmetric proof to Proposition (11) shows that the Pareto frontier is concave in this region.

**Region 3:** Suppose that $u \leq \underline{u}''$. Applying the argument from the proof of Proposition

(11), we can show that $(IC_{12})$ is binding and $(IC_{21})$ is slack, so that $\beta_2 = 0$. Then the solution is characterized by:

$$\beta_1 = (\gamma + 1)\lambda - \gamma, \quad \gamma \le \frac{\lambda}{(1 - \gamma)}, \quad \phi_1'(h_1) = \theta_1, \quad u = c_1 - \phi_1(h_1)$$

$$\phi_2'(h_2) = (\gamma\lambda + \lambda - \gamma)\left(\phi_1'(h_2) - \theta_2\right) + \theta_2$$

$$c_2 = \lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2 + \lambda\left(\phi_1(h_2) - \phi_1(h_1)\right)$$

$$c_1 = \lambda\theta_1 h_1 + (1 - \lambda)\theta_2 h_2 - (1 - \lambda)\left(\phi_1(h_2) - \phi_1(h_1)\right)$$

There is a minimum level of utility for the low type achievable here, denoted $u_{\min}^{RS}$ which is the utility level attained when $h_2$ is the solution to $\max_h c_2 - \phi_2(h)$ subject to $(IC_{12})$, or equivalently, when $\gamma \to 0$ in the above FOCs, that is when $\phi_2'(h_2) = \lambda\phi_1'(h_2) + (1 - \lambda)\theta_2$. A completely symmetric proof to before shows that the Pareto frontier is concave in this region.

The logic for this result is as follows: For $u$ small enough (below $\underline{u}''$) and hence $\gamma$ small enough (below $\lambda/(1 - \lambda)$), $(IC_{12})$ is binding and the high type is distorted upwards. As $u$ grows, $h_2$ is allowed to fall, until the utility $u$ reaches exactly the threshold $\underline{u}''$ (and correspondingly, $\gamma$ reaches the threshold $\frac{\lambda}{1-\lambda}$) and $h_2$ reaches its first best level. After that, as $u$ grows, through the transfer of consumption from agent 2 to agent 1, the $(IC_{12})$ becomes more and more slack while the $(IC_{21})$ eventually becomes binding, which occurs exactly when $u$ reaches the upper threshold $\overline{u}''$ and $\gamma$ becomes larger than $\lambda/(1 - \lambda)$.

**Proposition 19** *Let the thresholds $\overline{u}', \overline{u}, \overline{u}'', \underline{u}', \underline{u}''$ and $\underline{u}$ be defined as in the proofs.*

*For $u \le \overline{u}''$, the Second Best (and hence the Adverse Selection with unobservable private contracts frontier) and Rothschild-Stiglitz frontiers coincide. For $u \ge \overline{u}''$, the Rothschild-Stiglitz frontier is strictly below the Second Best frontier.*

*Furthermore, there are three regions:*

**Case 1:** *If condition $(NL1)$ holds, the thresholds for utilities are ranked as*

$$\underline{u} \le \underline{u}' = \underline{u}'' \le \overline{u} \le \overline{u}'' \le \overline{u}'$$

*Region 1: For $u \leq \underline{u}'$, the Mirrlees frontier (M) is above the Second Best (SB)*[40] *and Rothschild-Stiglitz (AS-RS) frontiers*

*Region 2: For $\underline{u}' \leq u \leq \overline{u}$, all frontiers coincide and are linear.*

*Region 3: For $u \geq \overline{u}$, the Second Best and Rothschild Stiglitz frontiers are both above the Mirrlees frontier.*

**Case 2:** *If condition $(NL1)$ does not hold, then:*

$$\underline{u} \leq \overline{u} \leq \underline{u}' = \underline{u}'' \leq \overline{u}'' \leq \overline{u}'$$

*Region 1: For $u \leq \overline{u}$, the Mirrlees frontier is above the Second Best and Rothschild-Stiglitz frontiers.*

*Region 2: For $\overline{u} \leq u \leq \underline{u}'$, either the Second Best and Rothschild-Stiglitz or the Mirrlees frontier could be higher.*

*Region 3: For $u \geq \underline{u}'$, both the Second Best and the Rothschild-Stiglitz frontiers are above the Mirrlees frontier.*

**Proof of Proposition (19):**

This proof involves only cumbersome algebra on the thresholds for utilities from the various regimes. Then, for a given ranking of those thresholds, the argument for which frontier is above the others is based solely upon considering the binding constraints and under which regime the constrained set is larger. To rank the thresholds, $\underline{u}, \overline{u}$ (for the Mirrlees frontier), $\underline{u}', \overline{u}'$ (for the Second Best), and $\underline{u}'', \overline{u}''$ (for the Adverse Selection case) consider the following calculations: $\underline{u}' - \underline{u} = \lambda \left( \phi_1 \left( h_2^* \right) - \phi_1 \left( h_1^* \right) \right) - \phi_1 \left( h_2^* \right) - \lambda \left( \phi_1 \left( \frac{\theta_2 h_2^*}{\theta_1} \right) - \phi_1 \left( h_1^* \right) \right) + \phi_1 \left( \frac{h_2^* \theta_2}{\theta_1} \right) = -(1 - \lambda) \left( \phi_1 \left( h_2^* \right) - \phi_1 \left( \frac{\theta_2 h_2^*}{\theta_1} \right) \right) \geq 0$

$\overline{u} - \overline{u}' = \left( \phi_2 \left( h_1^* \right) - \phi_2 \left( \frac{h_1^* \theta_1}{\theta_2} \right) \right) \lambda + \left( \phi_2 \left( \frac{h_1^* \theta_1}{\theta_2} \right) - \phi_2 \left( h_1^* \right) \right) = -(1 - \lambda) \left( \phi_2 \left( h_1^* \right) - \phi_2 \left( \frac{h_1^* \theta_1}{\theta_2} \right) \right) \leq 0$. Hence, it is always the case that: $\overline{u} \leq \overline{u}'$ and $\underline{u}' \geq \underline{u}$. In addition, we need to compare $\underline{u}'$ to

---

[40] *Recall that the Second Best Frontier coincides with the Adverse Selection with unobservable private contracts frontier.*

84

$\overline{u}$:

$$
\begin{aligned}
\underline{u}' \quad \leq \quad & \overline{u} \Leftrightarrow \lambda\theta_1 h_1^* + (1-\lambda)\,\theta_2 h_2^* + \lambda\left(\phi_1\left(h_2^*\right) - \phi_1\left(h_1^*\right)\right) - \phi_1\left(h_2^*\right) \leq \lambda\theta_1 h_1^* + (1-\lambda)\,\theta_2 h_2^* \\
& +\lambda\left(\phi_2\left(h_2^*\right) - \phi_2\left(\frac{h_1^*\theta_1}{\theta_2}\right)\right) - \phi_2\left(h_2^*\right) + \phi_2\left(\frac{h_1^*\theta_1}{\theta_2}\right) - \phi_1\left(h_1^*\right) \\
\Leftrightarrow \quad & (1-\lambda)\left(\phi_1\left(h_1^*\right) - \phi_1\left(h_2^*\right)\right) \leq -\,(1-\lambda)\left(\phi_2\left(h_2^*\right) - \phi_2\left(\frac{h_1^*\theta_1}{\theta_2}\right)\right) \\
\Leftrightarrow \quad & \left(\phi_1\left(h_1^*\right) - \phi_1\left(h_2^*\right)\right) \leq \phi_2\left(\frac{h_1^*\theta_1}{\theta_2}\right) - \phi_2\left(h_2^*\right)
\end{aligned}
$$

If this condition holds, (condition **NL1** in the main text) then $\underline{u}' \leq \overline{u}$. Comparing $\overline{u}''$ and $\overline{u}'$, as well as $\overline{u}''$ and $\overline{u}$ yields:

$$
\overline{u}'' \leq \overline{u}' \Leftrightarrow \phi_2\left(\frac{h_1^*\theta_1}{\lambda\theta_1 + (1-\lambda)\,\theta_2}\right) \leq \phi_2\left(h_1^*\right)
$$

$$
\overline{u}'' \geq \overline{u} \Leftrightarrow \phi_2\left(\frac{\theta_1 h_1^*}{(1-\lambda)\,\theta_2 + \lambda\theta_1}\right) \geq \phi_2\left(\frac{h_1^*\theta_1}{\theta_2}\right)
$$

which are both always true. In addition, note that $\underline{u}' = \underline{u}''$. Hence, there are two possible cases. If $(NL1)$ holds, then $\underline{u} \leq \underline{u}' = \underline{u}'' < \overline{u} \leq \overline{u}'' \leq \overline{u}'$. Else, $\underline{u} \leq \overline{u} \leq \underline{u}' = \underline{u}'' \leq \overline{u}'' \leq \overline{u}'$.

# Chapter 2

# Optimal Taxation and Human Capital Policies over the Life Cycle

## 2.1 Introduction

Investments in human capital, in the form of both time and money, play a key role in most people's lives. Children and young adults acquire education, and human capital accumulation continues throughout life through job training. There is a two-way interaction between human capital and the tax system. On the one hand, investments in human capital are influenced by tax policy – a point recognized early on by Schultz (1961).[1] Taxes on labor income discourage investment by capturing part of the return to human capital, yet also help insure against earnings risk, thereby encouraging investment in risky human capital. Capital taxes affect the choice between physical and human capital. On the other hand, investments in human capital directly impact the available tax base and are a major determinant of the pre-tax income distribution. Policies to stimulate human capital acquisition, which vary greatly across countries, shape the skill distribution of workers –a crucial input into optimal income taxation models.

---

[1] "Our tax laws everywhere discriminate against human capital. Although the stock of such capital has become large and even though it is obvious that human capital, like other forms of reproducible capital, depreciates, becomes obsolete and entails maintenance, our tax laws are all but blind on these matters." (Schultz, 1961, pg. 17).

This two-way feedback calls for a joint analysis of optimal income taxes and optimal human capital policies over the life cycle, which is the goal of this chapter. The vast majority of optimal tax research assumes that productivity is exogenously determined, instead of being the product of investment decisions made throughout life.[2] Therefore, this chapter addresses the following questions. First, how, if at all, should the tax and social insurance system take into account human capital acquisition? Should human capital expenses be tax deductible, and how should the opportunity cost of time be treated? Second, how does the optimal tax system change when human capital acquisition is unobservable to the government? Third, what parameters are important for setting optimal human capital policies, such as subsidies, and how do optimal policies evolve over time? Finally, what combination of policy instruments implements the optimum? Can simple policies yield a level of welfare close to that achieved with complex systems?

Specifically, this chapter jointly determines optimal tax and human capital policies over the life cycle, and incorporates essential characteristics of the human capital acquisition process. First, human capital pays off over long periods of time and thus returns are inherently uncertain: skills can be rendered obsolete by unpredictable changes in technology, industry shocks, or macroeconomic contractions.[3] Yet, private markets for insurance against personal productivity shocks are limited. Second, the investment has two types of costs, namely financial costs (e.g., tuition or books) and time costs (e.g., opportunity cost of foregone wages and loss of leisure while studying or training),[4] the relative importance of which can change over the life cycle.[5] Finally, individuals have heterogeneous intrinsic abilities, which may differentially affect their returns to human capital investment.[6]

Accordingly, in the model, each individual's wage is a function of endogenous human capital and stochastic "ability." Ability, as in the standard Mirrlees (1971) income taxation model,

[2] See Mirrlees (1971), Saez (2001), Golosov et al. (2006), Albanesi and Sleet (2006), Farhi and Werning (2013). Recent papers have relaxed the assumption of exogenous ability and are reviewed below.

[3] See Goldin and Katz (2008) and the literature reviewed later in the introduction.

[4] OECD countries spend on average 6.3% of GDP on education; the US spends 7.3% (OECD, 2013). Heckman (1976a) indirectly infers that 30% of time in early working years is spent training. 40% of adults participate in formal and/or non-formal education in a given year in OECD countries, and receive on average 988 hours of instruction in non-formal education during life, out of which 715 are job-related (OECD, 2013).

[5] See Dynarski (2003) and Heckman et al. (2005) on, respectively, the effects of tuition and of disutility from education on school enrollment decisions.

[6] The empirical evidence on this issue is reviewed in section 2.5.1.

is a comprehensive measure of the exogenous component of productivity. Agents have heterogeneous innate abilities, which are subject to persistent and privately uninsurable shocks. Throughout their lives, they can invest in human capital with risky returns, through either monetary expenses or time spent training at a disutility cost. The government maximizes a standard social welfare function under asymmetric information about any agent's ability – both its initial level and its evolution over life – as well as labor effort. This requires the imposition of incentive compatibility constraints in the dynamic mechanism designed by the government. To describe the distortions in the resulting constrained efficient allocations, the wedges, or implicit taxes and subsidies, are analyzed.

The lifecycle model in this chapter highlights the different forces at play. The implicit subsidy for human capital expenses is determined by three goals. The first is to counterbalance the distortions to human capital from implicit income and savings taxes. The second is to indirectly stimulate labor supply by increasing the wage, i.e., the returns to labor. The third is to redistribute, taking into account the differential effect of human capital on the pre-tax income of high and low ability people. When the wage elasticity with respect to ability is decreasing in human capital, human capital has a positive redistributive effect on after-tax income and a positive insurance value. It is then optimal to subsidize human capital expenses beyond simply insuring a neutral tax system with respect to human capital expenses, i.e., beyond making human capital expenses fully tax deductible in a dynamic, risk-adjusted fashion. In this case, the human capital wedge drifts up with age, and the drift is amplified by the magnitude of the wage elasticity with respect to ability. The persistence of ability shocks directly translates into a persistence of the optimal human capital wedge over time.

When human capital can be acquired through training time at a disutility cost, as well as by spending money on tuition or other inputs, there is an additional, direct interaction of training with both contemporaneous and future labor supply, which affects optimal policies. Training may increase the marginal disutility cost from working, the case labeled "Learning-or-Doing," or decrease it, labeled "Learning-And-Doing."

If, instead, human capital is unobservable to the government, the optimal labor and savings wedges need to indirectly provide the right incentives for human capital acquisition. The labor wedge may increase or decrease relative to the observable human capital case, depending, again,

on whether human capital has a positive redistributive or insurance effect.

This chapter considers two ways of implementing the constrained efficient allocations: Income Contingent Loans (ICLs), and a "Deferred Deductibility" scheme. For ICLs, the loan repayment schedules are contingent on the past history of earnings and human capital investments. In the Deferred Deductibility scheme, only part of current investment in human capital can be deducted from current taxable income. The remainder is deducted from future taxable incomes, to account for the risk and the nonlinearity of the tax schedule.

I calibrate the model based on US data to illustrate the optimal policies under different assumptions regarding the complementarity between human capital and ability. When human capital has a positive redistributive or insurance value, the net stimulus to human capital is small and positive, and grows with age. It is not optimal to deviate much from a neutral system with respect to human capital, a type of "production efficiency" result and, hence, full dynamic risk-adjusted deductibility is close to optimal. The optimal labor wedge rises with age, but the age trend is less steep than in standard models with no human capital. Labor taxes are progressive in the short run when human capital has a positive insurance value, but regressive otherwise. Simple linear age-dependent human capital subsidies, as well as income and savings taxes, achieve almost the entire welfare gain from the full second-best optimum for the calibrations studied.

### 2.1.1 Related Literature

The complex process of human capital acquisition has been studied in a long-standing literature, starting with Becker (1964), Ben-Porath (1967), and Heckman (1976a,b). The model in this chapter tries to adopt, in a stylized way, some of this literature's main findings. The structural branch of the literature (Cunha et al. 2006, Cunha and Heckman, 2007, 2008) emphasizes that human capital acquisition occurs throughout life, underscoring the need for a life cycle model. Both ex ante heterogeneity in the returns to human capital and uncertainty matter. Disutility or psychic costs, above and beyond pure opportunity costs, are important in explaining human capital investment decisions (Heckman, Todd and Lochner, 2005). A large body of empirical work documents the importance of human capital as a determinant of earnings (Card, 1995a, Goldin and Katz, 2008, Chetty et al., 2011, Huggett and Kaplan, 2011, Acemoglu and Autor,

2011), and the financial and other factors shaping individuals' decisions to acquire human capital (Lochner and Monge-Naranjo, 2011, Altonji, Blom and Meghir, 2012, Avery *et al.*, 2013). The subset of this literature which studies the interaction between ability and schooling for earnings – a crucial consideration for optimal policies in this chapter – is reviewed in detail in section 2.5.1.

On the other hand, the optimal taxation literature, dating back to Mirrlees (1971), and developed more recently by Saez (2001), Kocherlakota (2005), Albanesi and Sleet (2006), Golosov *et al.* (2006), Werning (2007a,b,c), Battaglini and Coate (2008), Scheuer (2013a), Golosov, Tsyvinski and Troshkin (2013), and Farhi and Werning (2013) typically assumes exogenous ability, thus abstracting from endogenous human capital investments. Therefore, this chapter builds on the lifecycle framework in Farhi and Werning (2013), and introduces endogenous stochastic productivity as the result of human capital acquisition by agents.

A series of papers, evolving from static to dynamic, have considered optimal taxation jointly with education policies. Bovenberg and Jacobs (2005), using a static taxation model, find that education subsidies and income taxes are "Siamese Twins" and should always be set equal to each other, which is equivalent to making human capital expenses fully tax deductible.[7] A few subsequent static papers emphasize the importance of the complementarity between intrinsic ability and human capital (Maldonado, 2007, with two types, Bovenberg and Jacobs, 2011 with a continuum of types), or between risk and human capital (DaCosta and Maestri, 2007).

Several recent dynamic optimal tax papers examine the impact of taxation on human capital, with important differences to the current chapter. Previous dynamic models allowed for heterogeneity across agents, but not uncertainty (Bohacek and Kapicka, 2008, Kapicka, 2013a), or uncertainty, but not heterogeneity (Anderberg, 2009, Grochulski and Piskorski, 2010), which precludes a discussion of redistributive policies. Findeisen and Sachs (2013) include both heterogeneity and uncertainty, but focus on a one-shot investment during "college," before the work life of the agent starts, with a one-time realization of uncertainty. By contrast, this chapter features life cycle investment in human capital, through both expenses and time, and a progressive realization of uncertainty throughout life. A complementary analysis is Kapicka and Neira (2013), who posit a different human capital accumulation process with time invest-

---

[7]Kaplow (1994) also considers the tax treatment of human capital.

ments and a fixed ability, and consider the case in which effort spent to acquire human capital is unobservable. Also complementary is the work by Krueger and Ludwig (2013), who adopt a Ramsey approach by specifying ex ante the instruments available to the government, in contrast to the Mirrlees approach adopted here, which considers an unrestricted direct revelation mechanism. In their overlapping generations general equilibrium model, "education" is a binary decision that occurs exclusively before entry into the labor market.[8] The lifecycle analysis also addresses the issue of age-dependent taxation, as explored in Kremer (2002), Weinzierl (2011), and Mirrlees at al. (2011).[9]

The rest of the chapter is organized as follows. Section 2.2 presents the dynamic lifecycle model and the full information benchmark. Section 2.3 sets up a recursive mechanism design program using the first-order approach. Section 2.4 solves for the optimal policies when human capital is observable and describes the forces shaping those policies, as well as their evolution over life. Section 2.5 contains the numerical analysis, which brings out the main mechanics behind the optimal policies, and illustrates their evolution over life. Section 2.6 discusses the implementation of the optimal policies using Income Contingent Loans (ICLs) and a Deferred Deductibility scheme. Section 2.7 considers the case with unobservable human capital. Section 2.8 concludes and discusses three alternative applications of the model: to intergenerational transfers and bequest taxation, to entrepreneurial taxation, and to health investments.

## 2.2 Lifecycle Model of Human Capital Acquisition and Labor Supply

This section starts with the setup of a lifecycle model with risky human capital.

The economy consists of agents who live for $T$ years, during which they work and acquire human capital. Agents who work $l_t \geq 0$ hours in period $t$ at a wage rate $w_t$ earn a gross income

---

[8] Benabou (2002) jointly analyses taxes and education in a dynastic Ramsey model.

[9] This paper is more generally related to the dynamic mechanism design literature, as developed by, among many others, Baron and Besanko (1984), Spear and Srivastava (1987), Fernandes and Phelan (2000), and Doepke and Townsend (2006), which also underscores the challenges posed by general forms of persistence of shocks. The closest methodological link is to Farhi and Werning's (2013) first-order approach to such problems (see also Pavan, Segal and Toikka, 2013). The case with unobservable human capital builds on models with hidden savings (Cole and Kocherlakota, 2001, Werning, 2002, Kocherlakota, 2004, Abraham and Pavoni, 2008) or hidden trades (Golosov and Tsyvinski, 2007).

$y_t = w_t l_t$. Each period, agents can build their stock of human capital in two ways. The first one is by spending money. A monetary investment of amount $M_t(e_t)$ generates an increase in human capital $e_t \geq 0$. The cost function satisfies: $M'_t(e) > 0$, $\forall e_t > 0$; $M'_t(0) = 0$; $M''_t(e_t) \geq 0$, $\forall e_t \geq 0$. These monetary investments add to a stock of human capital acquired by expenses ("expenses" for short), $s_t$, which evolves according to $s_t = s_{t-1} + e_t$.[10] Expenses can be thought of as the necessary material inputs into the production of human capital, such as books, tuition fees, or living and board costs while at college, net of the cost of living elsewhere. Second, agents can spend time training. This may represent the time spent in formal classes or training courses, time spent reading, or on-the-job training. The disutility cost to an agent who provides $l_t$ units of work and spends $i_t$ units in training is $\phi_t(l_t, i_t)$, with $i_t \geq 0$. This second type of investment leads to a stock of training time $z_t$, which evolves according to $z_t = z_{t-1} + i_t$. $\phi_t$ is strictly increasing and convex in each of its arguments ($\frac{\partial \phi(l,i)}{\partial i}, \frac{\partial \phi(l,i)}{\partial l} > 0$, $\forall l, i > 0$; $\frac{\partial^2 \phi(l,i)}{\partial l^2}, \frac{\partial^2 \phi(l,i)}{\partial i^2} \geq 0$, $\forall l, i$); however, no assumption is yet made on the cross partial $\frac{\partial^2 \phi(l,i)}{\partial l \partial i}$.[11]

These two forms of human capital acquisition might vary in importance over the life cycle. For example, young people may face only low-wage opportunities, hence a low opportunity cost of time, but high tuition fees and financial costs. On the other hand, for working adults with college education and high opportunity cost of time, the costs of human capital investments in the form of on-the-job training might mostly be time and disutility costs.

The wage rate $w_t$ is determined by the stocks of human capital built until time $t$ and stochastic ability $\theta_t$:

$$w_t = w_t(\theta_t, s_t, z_t)$$

$w_t$ is strictly increasing and concave in each of its arguments ($\frac{\partial w}{\partial m} > 0$, $\frac{\partial^2 w}{\partial m^2} \leq 0$ for $m = \theta, s, z$).[12] Importantly though, no restrictions are placed on the cross-partials. This formulation allows for different types of human capital to affect the wage differently over the lifecycle.

Agents are born at time $t = 1$ with a heterogeneous earning ability $\theta_1$ with distribution $f^1(\theta_1)$. Earning ability in each period is private information, and evolves according to a Markov

---

[10]The agent cannot wilfully destroy human capital, hence $e_t \geq 0$. There is no explicit depreciation of human capital, but the ability shock $\theta$ described right below can partially account for stochastic depreciation. Deterministic depreciation each period is easy to add and dropped for notational convenience.

[11]Appendix A.1. links this model to the standard Ben-Porath model.

[12]Equivalently, we could define human capital as a composite function $h_t$ of all expenses made and time spent training $h_t = h_t(s_t, z_t)$, with wage $\tilde{w}_t = w_t(\theta_t, h_t)$.

process with a time-varying transition function $f^t(\theta_t|\theta_{t-1})$, over a fixed support $\Theta \equiv [\underline{\theta}, \bar{\theta}]$. There are several possible interpretations for $\theta_t$, such as stochastic productivity or stochastic returns to human capital. For example, with a separable wage form $w_t = \theta_t + h_t(s_t, z_t)$, for some function $h_t$, $\theta_t$ resembles a stochastic version of productivity from the static Mirrlees (1971) model. With a wage such as $w_t = \theta_t h_t(s_t, z_t)$, $\theta_t$ is perhaps more naturally interpreted as the stochastic return to human capital. To keep with the tradition in the literature, $\theta_t$ will be called "ability" throughout. Ability to earn income can be stochastic among others because of health shocks, individual labor market idiosyncrasies or luck.

The agent's per period utility is separable in consumption and effort (both labor and training):

$$\tilde{u}_t(c_t, y_t, s_t, z_t; \theta_t, z_{t-1}) = u_t(c_t) - \phi_t\left(\frac{y_t}{w_t(\theta_t, s_t, z_t)}, z_t - z_{t-1}\right)$$

$u_t$ is increasing, twice continuously differentiable, and concave.

Denote by $\theta^t$ the history of ability shocks up to period $t$, by $\Theta^t$ the set of possible histories at $t$, and by $P(\theta^t)$ the probability of a history $\theta^t$, $P(\theta^t) \equiv f^t(\theta_t|\theta_{t-1})...f^2(\theta_2|\theta_1)f^1(\theta_1)$. An allocation $\{x_t\}_t$ specifies consumption, output, and expenses and training stocks for each period $t$, conditional on the history $\theta^t$, i.e., $x_t = \{x(\theta^t)\}_{\Theta^t} = \{c(\theta^t), y(\theta^t), s(\theta^t), z(\theta^t)\}_{\Theta^t}$. The expected lifetime utility from an allocation, discounted by a factor $\beta$, is given by:

$$U\left(\{c(\theta^t), y(\theta^t), s(\theta^t), z(\theta^t)\}\right)$$
$$= \sum_{t=1}^{T}\int \beta^{t-1}\left[u_t\left(c(\theta^t)\right) - \phi_t\left(\frac{y(\theta^t)}{w_t(\theta_t, s(\theta^t), z(\theta^t))}, z(\theta^t) - z(\theta^{t-1})\right)\right]P(\theta^t)\,d\theta^t \quad (2.1)$$

where, with some abuse of notation, $d\theta^t \equiv d\theta_t...d\theta_1$.

Let $w_{m,t}$ denote the partial of the wage function with respect to argument $m$ ($m \in \{\theta, s, z\}$), and $w_{mn,t}$ the second order partial with respect to arguments $m, n \in \{\theta, s, z\} \times \{\theta, s, z\}$. Two important parameters are the Hicksian coefficients of complementarity between ability and

93

human capital (of type $s$ or $z$) in the wage function at time $t$, denoted by $\rho_{\theta s,t}$ and $\rho_{\theta z,t}$:[13]

$$\rho_{\theta s} \equiv \frac{w_{\theta s}w}{w_s w_\theta}, \quad \rho_{\theta z} \equiv \frac{w_{\theta z}w}{w_z w_\theta} \qquad (2.2)$$

A positive Hicksian complementarity between human capital $s$ and ability $\theta$ means that higher ability agents have a higher marginal benefit from human capital ($w_{\theta s} \geq 0$).[14] Put differently, human capital compounds the exposure of the agent to stochastic ability and to risk. A Hicksian complementarity greater than 1 means that higher ability agents have a higher proportional benefit from human capital, i.e., the wage elasticity with respect to ability is increasing in human capital, i.e., $\frac{\partial}{\partial s}\left(\frac{\partial w}{\partial \theta}\frac{\theta}{w}\right) \geq 0$.[15] The same applies to training time $z$.

A separable wage function of the form $w_t = \theta_t + h_t\left(s_t, z_t\right)$ for some function $h_t$ implies that $\rho_{\theta s,t} = \rho_{\theta z,t} = 0$. A multiplicative form $w_t = \theta_t h_t\left(s,z\right)$, the one typically used in the taxation literature, implies that $\rho_{\theta s,t} = \rho_{\theta z,t} = 1$. Finally, with a CES wage function, of the form

$$w_t = \left[\alpha_{1t}\theta^{1-\rho_t} + \alpha_{2t}s_t^{1-\rho_t} + \alpha_{3t}z^{1-\rho_t}\right]^{\frac{1}{1-\rho_t}} \qquad (2.3)$$

ability and human capital can be substituted one for the other at a fixed, but potentially time-varying rate: $\rho_{\theta s,t} = \rho_{\theta z,t} = \rho_t$.

For training time, another relevant parameter is $\rho_{lz}^\phi$, the Hicksian complementarity coefficient between labor and training in the disutility function $\phi_t\left(l_t, z_t - z_{t-1}\right)$. Let $\phi_{z,t} = \frac{\partial \phi_t}{\partial z_t}$, $\phi_{l,t} = \frac{\partial \phi_t}{\partial l_t}$, $\phi_{lz,t} = \frac{\partial^2 \phi}{\partial l \partial z}$. Then:

$$\rho_{lz}^\phi \equiv \frac{\phi_{lz}\phi}{\phi_l \phi_z} \qquad (2.4)$$

## 2.3 The Planning Problem

The informational problem that the planner faces is that he cannot see ability $\theta_t$ in any period. Hence, he also does not know an agent's wage $w_t\left(\theta_t, s_t, z_t\right)$, or labor supply $l_t = y_t/w_t$. Output $y_t$ and consumption $c_t$ on the other hand are observable. Human capital investments are

---

[13] On studies of the Hicksian complementarity as see Hicks (1970), Samuelson (1973), Bovenberg and Jacobs (2011).

[14] $\rho_{\theta s}$ is also the Hicksian complementarity coefficient between education and ability in earnings $y$.

[15] Equivalently, the wage elasticity with respect to human capital is increasing in ability.

also observable to the government but the case with unobservable human capital expenses is discussed in Section 2.7.

This technical section sets up the planning problem, starting from the sequential problem, and defining incentive compatibility. It then goes through two steps to make this problem tractable, following the recent procedure proposed for dynamic Mirrlees models by Farhi and Werning (2013), augmented here with human capital. First, a relaxed problem based on the first-order approach is written out, which replaces the full set of incentive compatibility constraints by the agent's envelope condition. This relaxed program is then turned into a recursive dynamic programing problem through a suitable definition of state variables. The results on optimal policies are in the subsequent section.

### 2.3.1 Incentive Compatibility

To solve for the constrained efficient allocations, suppose that the planner designs a direct revelation mechanism, in which, each period, agents have to report their ability $\theta_t$. Denote a reporting strategy, specifying a reported type $r_t$ after each history by $r = \left\{ r_t \left( \theta^t \right) \right\}_{t=1}^{T}$. Let $\mathcal{R}$ be the set of all possible reporting strategies and $r^t = \left\{ r_1 \left( \theta_1 \right), ..., r_t \left( \theta^t \right) \right\}$ be the history of reports generated by reporting strategy $r$. Because output, savings, and human capital are observable, the planner can directly specify allocations as functions of the history of reports, according to some allocation rules $c \left( r^t \right)$, $y \left( r^t \right)$, $s \left( r^t \right)$, $z \left( r^t \right)$.[16] Let the continuation value after history $\theta^t$ under a reporting strategy $r$, denoted by $\omega^r \left( \theta^t \right)$, be the solution to:

$$\omega^r \left( \theta^t \right) = u_t \left( c \left( r^t \left( \theta^t \right) \right) \right) - \phi_t \left( \frac{y \left( r^t \left( \theta^t \right) \right)}{w_t \left( \theta_t, s \left( r^t \left( \theta^t \right) \right), z \left( r^t \left( \theta^t \right) \right) \right)}, z \left( r^t \left( \theta^t \right) \right) - z \left( r^{t-1} \left( \theta^{t-1} \right) \right) \right)$$
$$+ \beta \int \omega^r \left( \theta^{t+1} \right) f^{t+1} \left( \theta_{t+1} | \theta_t \right) d\theta_{t+1}$$

The continuation value under truthful revelation, $\omega \left( \theta^t \right)$, is the unique solution to:

---

[16]Hours of work are determined residually by $l \left( r^t \right) = y \left( r^t \right) / w \left( \theta_t, s \left( r^t \right), z \left( r^t \right) \right)$ and hours of training by $i \left( r^t \right) = z \left( r^t \right) - z \left( r^{t-1} \right)$.

$$\omega\left(\theta^{t}\right) = u_{t}\left(c\left(\theta^{t}\right)\right) - \phi_{t}\left(\frac{y\left(\theta^{t}\right)}{w_{t}\left(\theta_{t}, s\left(\theta^{t}\right), z\left(\theta^{t}\right)\right)}, z\left(\theta^{t}\right) - z\left(\theta^{t-1}\right)\right) + \beta \int \omega\left(\theta^{t+1}\right) f^{t+1}\left(\theta_{t+1}|\theta_{t}\right) d\theta_{t+1}$$

Incentive compatibility requires that truth-telling yields a weakly higher continuation utility than any reporting strategy $r$:

$$(IC): \omega\left(\theta_{1}\right) \geq \omega^{r}\left(\theta_{1}\right) \quad \forall\theta_{1}, \forall r \tag{2.5}$$

Denote by $X^{IC}$ the set of allocations which satisfy incentive compatibility condition (2.5).

## 2.3.2 Planning Problem

The analysis is in partial equilibrium, or assuming a small open economy with fixed gross interest rate $R$. The planner's objective is to minimize the expected discounted cost of providing an allocation, subject to incentive compatibility as defined in (2.5), and to expected lifetime utility being above a threshold $\underline{U}$. The planning problem in sequential form is:

$$\min_{\{c,y,s,z\}} \Pi\left(\{c,y,s,z\}\right) = \left[\sum_{t=1}^{T}\left(\frac{1}{R}\right)^{t-1}\int_{\Theta^{t}}\left(c\left(\theta^{t}\right) - y\left(\theta^{t}\right) + M_{t}\left(s\left(\theta^{t}\right) - s\left(\theta^{t-1}\right)\right)\right) P\left(\theta^{t}\right) d\theta^{t}\right] \tag{2.6}$$

$$\text{s.t.}: U\left(\{c,y,s,z\}\right) \geq \underline{U}$$

$$y\left(\theta^{t}\right) \geq 0, s\left(\theta^{t}\right) \geq s\left(\theta^{t-1}\right), z\left(\theta^{t}\right) \geq z\left(\theta^{t-1}\right), c\left(\theta^{t}\right) \geq 0$$

$$\{c,y,s,z\} \text{ is incentive compatible}$$

**First order approach:** To solve this dynamic problem, a version of the first order approach is used. The following assumptions are needed and maintained throughout the chapter.

**Assumption 3** *i)* $\tilde{u}_{t}\left(c,y,s,z;\theta,z_{-}\right)$ *and* $\frac{\partial\phi(l,i)}{\partial l}\frac{\partial w(\theta,s,z)}{\partial\theta}\frac{l}{w}$ *are bounded.* *ii)* $\frac{\partial f^{t}(\theta_{t}|\theta_{t-1})}{\partial\theta_{t-1}}$ *exists and is bounded.*[17] *iii)* $f^{t}\left(\theta_{t}|\theta_{t-1}\right)$ *has full support on* $\Theta$.

---

[17] *For some distributions, this derivative is not bounded and assumption 3 in Kapicka (2013) could be used instead, namely that for* $F^{t}\left(\theta_{t}|\theta_{t-1}\right) \equiv \int_{\underline{\theta}}^{\theta} f^{t}\left(\theta_{s}|\theta_{t-1}\right) d\theta_{s}$, *we have* $\frac{\partial}{\partial\theta_{t-1}}F^{t}\left(\theta_{t}|\theta_{t-1}\right) \leq 0$ *and* $F^{t}\left(\theta_{t}|\theta_{t-1}\right)$ *either concave or convex.*

Suppose the agent has witnessed a history of shocks $\theta^t$. Consider one particular deviation strategy $\tilde{r}_t$, under which he reports truthfully until period $t$ ($\tilde{r}_s\left(\theta^s\right) = \theta_s \ \forall s \leq t-1$), and lies in period $t$ by reporting $\tilde{r}_t\left(\theta^t\right) = \theta' \neq \theta_t$. The continuation utility under this strategy is the solution to:

$$\omega^{\tilde{r}}\left(\theta^t\right) = u_t\left(c\left(\theta^{t-1},\theta'\right)\right) - \phi_t\left(\frac{y\left(\theta^{t-1},\theta'\right)}{w_t\left(\theta_t, s\left(\theta^{t-1},\theta'\right), z\left(\theta^{t-1},\theta'\right)\right)}, z\left(\theta^{t-1},\theta'\right) - z\left(\theta^{t-1}\right)\right)$$
$$+ \beta \int \omega^{\tilde{r}}\left(\theta^{t-1},\theta',\theta_{t+1}\right) f^t\left(\theta_{t+1}|\theta_t\right) d\theta_{t+1}$$

Incentive compatibility in (2.5) implies that, after almost all $\theta^t$, the temporal incentive constraint holds:

$$\omega\left(\theta^t\right) = \max_{\theta'} \omega^{\tilde{r}}\left(\theta^t\right) \tag{2.7}$$

Inversely, if (2.7) holds after all $\theta^{t-1}$ and for almost al $\theta_t$, then (2.5) also holds (see Kapicka (2013b), Lemma 1). The first order approach consists in replacing (2.7) by the envelope condition of the agent:

$$\dot{\omega}\left(\theta^t\right) := \frac{\partial \omega\left(\theta^t\right)}{\partial \theta_t} = \frac{w_{\theta,t}}{w_t} l\left(\theta^t\right) \phi_{l,t}\left(l\left(\theta^t\right), z_t\left(\theta^t\right) - z_{t-1}\left(\theta^{t-1}\right)\right) + \beta \int \omega\left(\theta^{t+1}\right) \frac{\partial f^{t+1}\left(\theta_{t+1}|\theta_t\right)}{\partial \theta_t} d\theta_{t+1}$$
$$\tag{2.8}$$

or its integral version:

$$\omega\left(\theta^t\right) = \omega\left(\theta^{t-1},\underline{\theta}\right) + \int_{\underline{\theta}}^{\theta_t} \left(\frac{w_{\theta,t}}{w_t} l\left(\theta^{t-1},\theta_s\right) \phi_{l,t}\left(l\left(\theta^{t-1},\theta_s\right), z_t\left(\theta^{t-1},\theta_s\right) - z_{t-1}\left(\theta^{t-1}\right)\right)\right.$$
$$\left. + \beta \int \omega\left(\theta^{t-1},\theta_s,\theta_{t+1}\right) \frac{\partial f^{t+1}\left(\theta_{t+1}|\theta_s\right)}{\partial \theta_s} d\theta_{t+1}\right) d\theta_s \quad (2.9)$$

Let $X^{FOA}$ denote the set of allocations which satisfy the envelope condition (2.9). It can be shown that this is a necessary condition for incentive compatibility, i.e., $X^{IC} \subseteq X^{FOA}$.[18] The

---

[18]An application of Theorem 2 in Milgrom and Segal (2002), under assumption 3.

relaxed planning problem, denoted by $P^{FOA}$ is:

$$\min_{\{c,y,s,z\}} \Pi\left(\{c,y,s,z\}\right)$$

$$\text{s.t: } U\left(\{c,y,s,z\}\right) \geq \underline{U}$$

$$\{c,y,s,z\} \in X^{FOA}$$

### 2.3.3   Recursive Formulation of the Relaxed Program

To write the problem recursively, let the second term in the envelope condition in differential form – which governs the future incentives of the agent – be denoted by:

$$\Delta\left(\theta^t\right) \equiv \int \omega\left(\theta^{t+1}\right) \frac{\partial f^{t+1}\left(\theta_{t+1}|\theta_t\right)}{\partial \theta_t} d\theta_{t+1} \tag{2.10}$$

The envelope condition can then be rewritten as:

$$\dot{\omega}\left(\theta^t\right) = \frac{w_{\theta,t}}{w_t} l\left(\theta^t\right) \phi_{l,t}\left(l\left(\theta^t\right), z_t\left(\theta^t\right) - z_{t-1}\left(\theta^{t-1}\right)\right) + \beta\Delta\left(\theta^t\right) \tag{2.11}$$

Let $v\left(\theta^t\right)$ be the expected future continuation utility:

$$v\left(\theta^t\right) \equiv \int \omega\left(\theta^{t+1}\right) f^{t+1}\left(\theta_{t+1}|\theta_t\right) d\theta_{t+1} \tag{2.12}$$

Continuation utility $\omega\left(\theta^t\right)$ can hence be rewritten as:

$$\omega\left(\theta^t\right) = u_t\left(c\left(\theta^t\right)\right) - \phi_t\left(\frac{y\left(\theta^t\right)}{w_t\left(\theta_t, s\left(\theta^t\right), z\left(\theta^t\right)\right)}, z\left(\theta^t\right) - z\left(\theta^{t-1}\right)\right) + \beta v\left(\theta^t\right) \tag{2.13}$$

Define the expected continuation cost of the planner at time $t$, given $v_{t-1}, \Delta_{t-1}, \theta_{t-1}, s_{t-1}$, and $z_{t-1}$:

$$K\left(v, \Delta, \theta_{t-1}, s_{t-1}, z_{t-1}, t\right) = \min\left[\sum_{\tau=t}^{T}\left(\frac{1}{R}\right)^{\tau-t} \int \left(c_\tau\left(\theta^\tau\right) - y_\tau\left(\theta^\tau\right) + M_\tau\left(s_\tau\left(\theta^\tau\right) - s_\tau\left(\theta^{\tau-1}\right)\right)\right) P\left(\theta^{\tau-t}\right) d\theta^{\tau-t}\right]$$

where, with some abuse of notation, $d\theta^{\tau-t} = d\theta_\tau d\theta_{\tau-1}...d\theta_t$, and $P\left(\theta^{\tau-t}\right) = f^\tau\left(\theta_\tau|\theta_{\tau-1}\right)...f^t\left(\theta_t|\theta_{t-1}\right)$. The minimization is over continuation allocations after period $t$, $\{c, y, s, z, \omega, v, \Delta\}_{\tau \geq t}$, and subject to the constraints $(2.10)$, $(2.11)$, $(2.12)$, and $(2.13)$, evaluated at $t-1$.

A recursive formulation of the relaxed program is then for $t \geq 2$:

$$
K\left(v, \Delta, \theta_-, s_-, z_-, t\right) = \min \int (c\left(\theta\right) + M_t\left(s\left(\theta\right) - s_-\right) - w_t\left(\theta, s\left(\theta\right), z\left(\theta\right)\right) l\left(\theta\right)
$$
$$
+ \frac{1}{R} K\left(v\left(\theta\right), \Delta\left(\theta\right), \theta, s\left(\theta\right), z\left(\theta\right), t+1\right)) f^t\left(\theta|\theta_-\right) d\theta \quad (2.14)
$$

subject to:

$$
\omega\left(\theta\right) = u_t\left(c\left(\theta\right)\right) - \phi_t\left(l\left(\theta\right), z\left(\theta\right) - z_-\right) + \beta v\left(\theta\right)
$$
$$
\dot{\omega}\left(\theta\right) = \frac{w_{\theta,t}}{w_t} l\left(\theta\right) \phi_{l,t}\left(l\left(\theta\right), z\left(\theta\right) - z_-\right) + \beta \Delta\left(\theta\right)
$$
$$
v = \int \omega\left(\theta\right) f^t\left(\theta|\theta_-\right) d\theta
$$
$$
\Delta = \int \omega\left(\theta\right) \frac{\partial f^t\left(\theta|\theta_-\right)}{\partial \theta_-} d\theta
$$

where the maximization is over $\left(c\left(\theta\right), l\left(\theta\right), s\left(\theta\right), z\left(\theta\right), \omega\left(\theta\right), v\left(\theta\right), \Delta\left(\theta\right)\right)$.

For period $t = 1$, the problem needs to be reformulated only if the initial shock $\theta_1$ is interpreted as heterogeneity, rather than uncertainty, and the objective is not utilitarian, so as to derive the full Pareto frontier. Suppose all agents have identical initial human capital levels $s_0$ and $z_0$. The problem for $t = 1$ is then indexed by $(\underline{U}\left(\theta\right))_\Theta$, the set of target lifetime utilities $\underline{U}\left(\theta\right)$ for each type $\theta$:

$$
K\left((\underline{U}\left(\theta\right))_\Theta, 1\right) = \min \int (c\left(\theta\right) + M_1\left(s\left(\theta\right) - s_0\right) - w_1\left(\theta, s\left(\theta\right), z\left(\theta\right)\right) l\left(\theta\right)
$$
$$
+ \frac{1}{R} K\left(v\left(\theta\right), \Delta\left(\theta\right), \theta, s\left(\theta\right), z\left(\theta\right), 2\right)) f^1\left(\theta\right) d\theta
$$

99

$$\text{s.t.} \quad : \quad \omega(\theta) = u_1(c(\theta)) - \phi_1(l(\theta), z(\theta) - z_0) + \beta v(\theta)$$

$$\dot{\omega}(\theta) = \frac{w_{\theta,1}}{w_1} l(\theta) \phi_{l,1}(l(\theta), z(\theta)) + \beta \Delta(\theta)$$

$$\omega(\theta) \geq \underline{U}(\theta)$$

where the maximization is again over $(c(\theta), l(\theta), \omega(\theta), s(\theta), z(\theta), v(\theta), \Delta(\theta))$. The set of constrained efficient allocations that solve the Planning problem is indexed by the set of utilities $(\underline{U}(\theta))_\Theta$ and denoted by $X^{*,FOA}((\underline{U}(\theta))_\Theta)$.

### 2.3.4 Validity of the First-order Approach and Assumptions

The solution to the relaxed program might not be a solution to the full program, because the envelope condition is only a necessary condition. In the static taxation model (Mirrlees, 1971), the validity of the first-order approach is guaranteed if the utility function satisfies the standard Spence-Mirrlees single-crossing property and a simple monotonicity condition on the allocation. However, in the dynamic case, the conditions imposed on the allocations are more involved (see Battaglini and Lamba, 2013, Golosov, Troshkin and Tsyvinski, 2013 or Pavan, Segal and Toikka, 2013), and do not always provide much simplification. Hence, for the proposed calibrations in section 2.5, incentive compatibility of the candidate allocation, as well as any omitted non-negativity constraints, are checked numerically, using a procedure in the spirit of Werning (2007) and Farhi and Werning (2013).[19]

**Technical assumptions:** The following assumptions are used as sufficient conditions only to determine the sign of the optimal wedges. All formulas derived still apply without them.[20]

**Assumption 4** *i)* $\frac{\partial v(\theta)}{\partial \theta} > 0$ *for all* $\theta$, *ii)* $\int_{\underline{\theta}}^{\theta'} f^t(\theta|\theta_b) d\theta \leq \int_{\underline{\theta}}^{\theta'} f^t(\theta|\theta_s) d\theta$, $\forall t, \theta'$, *and* $\theta_b > \theta_s$; *iii)* $\frac{\partial}{\partial \theta_t} \left( \frac{\partial f^t(\theta_t|\theta_{t-1})}{\partial \theta_{t-1}} \frac{1}{f^t(\theta_t|\theta_{t-1})} \right) \geq 0$, $\forall t, \forall \theta_{t-1}$; *iv)* $\frac{\partial}{\partial v} K \geq 0$ *and* $\frac{\partial^2}{\partial v^2} K \geq 0$.

---

[19] An alternative could be the random contracts or "lotteries" approach (e.g.: Karaivanov and Townsend, 2013), which circumvents the sufficiency problem, and has explored increasingly sophisticated methods to increase computational efficiency, and counter the "curse of dimensionality," which arises from the exponential growth of incentive constraints when adding hidden states or actions. For optimal tax analysis, however, it is appealing to have analytical formulas, which the first order approach, when valid, can deliver, and which build the intuition for the solution.

[20] In addition, all theoretical results on the signs of the wedges are indeed satisfied in the simulations (section 2.5).

Assumption i) guarantees that the expected continuation utility is increasing in the type. The first-order stochastic dominance assumption in ii) ensures that, for any given future payoff function increasing in $\theta$, higher types today have a higher expected payoff. Assumption iii) introduces a form of monotone likelihood ratio property. Assumption iv) states that the resource cost is increasing and convex in promised utility.

To reduce notational clutter throughout the chapter, the dependence on the full history is often left implicit, e.g.: $c_t = c\left(\theta^t\right)$ and $\tau_{Lt} = \tau_{Lt}\left(\theta^t\right)$. Similarly, function arguments are sometimes left out, e.g.: $w_{z,t} = \frac{\partial}{\partial z}w_t\left(\theta_t, s\left(\theta^t\right), z\left(\theta^t\right)\right)$. $E_t$ denotes the expectation at time $t$, conditional on $\theta_t$.

## 2.4 Optimal Human Capital Policies

This section characterizes the allocations, obtained as solutions to the relaxed program $P^{FOA}$ above, using wedges, or implicit taxes and subsidies. Unless otherwise stated, "optimal" refers to the optimum of program $P^{FOA}$.

### 2.4.1 The Wedges or Implicit Taxes and Subsidies

In the second best, marginal distortions in agents' choices can be described using wedges which represent the implicit local marginal tax and subsidy rates. At a given allocation and history $\theta^t$, define the intratemporal wedge on labor $\tau_L\left(\theta^t\right)$, the intertemporal wedge on savings or capital $\tau_K\left(\theta^t\right)$, and the human capital wedges $\tau_S\left(\theta^t\right)$ and $\tau_Z\left(\theta^t\right)$ as follows:

$$\tau_L\left(\theta^t\right) \equiv 1 - \frac{\phi_{l,t}\left(l_t, z_t - z_{t-1}\right)}{w_t\left(\theta_t, s_t, z_t\right)u_t'\left(c_t\right)} \tag{2.15}$$

$$\tau_K\left(\theta^t\right) \equiv 1 - \frac{1}{R\beta}\frac{u_t'\left(c_t\right)}{E_t\left(u_t'\left(c_{t+1}\right)\right)} \tag{2.16}$$

$$\tau_S\left(\theta^t\right) \equiv -\left(1 - \tau_L\left(\theta^t\right)\right)w_{s,t}l_t + \left[M_t'\left(s_t - s_{t-1}\right) - \beta E_t\left(\frac{u_{t+1}'\left(c_{t+1}\right)}{u_t'\left(c_t\right)}M_{t+1}'\left(s_{t+1} - s_t\right)\right)\right] \tag{2.17}$$

$$\tau_Z\left(\theta^t\right) \equiv -\left(1 - \tau_L\left(\theta^t\right)\right)w_{z,t}l_t + \left[\frac{\phi_{z,t}}{u_t'\left(c_t\right)} - \beta E_t\left(\frac{u_{t+1}'\left(c_{t+1}\right)}{u_t'\left(c_t\right)}\frac{\phi_{z,t+1}}{u_{t+1}'\left(c_{t+1}\right)}\right)\right] \tag{2.18}$$

The labor wedge $\tau_L$ is defined as the gap between the marginal rate of substitution and the marginal rate of transformation between consumption and labor. The savings or capital wedge

$\tau_K$ is defined as the difference between the expected marginal rate of intertemporal substitution and the return on savings. The implicit subsidy on human capital expenses $\tau_S$ can be thought of as the reduction in marginal cost for an additional dollar of spending on human capital: the agent's net marginal cost is $M'_t(e_t) - \tau_{St}$.[21] Similarly, the "bonus" for training $\tau_Z$ can be viewed as the incremental pay received by an agent for one more unit of training time.

Intuitively, conditional on a history, wedges are akin to locally linear subsidies and taxes, and would be zero absent government intervention. Indeed, the chapter will sometimes loosely refer to the wedges as taxes and subsidies, and often appeal to intuitions related to a standard tax system. The relation between wedges and explicit taxes is studied in the section on implementation (section 2.6).

The following definitions will be needed for the formulas below. For any variable $x$, define the "insurance factor" of $x$, $\xi_{x,t+1}$ :

$$\xi_{x,t+1} \equiv Cov\left(-\beta\frac{u'_{t+1}}{u'_t}, x_{t+1}\right) \Big/ \left(E_t\left(\beta\frac{u'_{t+1}}{u'_t}\right)E_t(x_{t+1})\right)$$

with $\xi_{x,t+1} \in [-1,1]$. If $x$ is a flow to the agent, it is a good hedge if $\xi < 0$, and a bad hedge otherwise. With some abuse of notation, define also:

$$\xi'_{x,t+1} \equiv -Cov\left(\frac{\beta u'_{t+1}}{u'_t} - 1, x_{t+1}\right) \Big/ \left(E_t\left(\frac{\beta u'_{t+1}}{u'_t} - 1\right)E_t(x_{t+1})\right)$$

which, up to an additive constant, captures the same risk properties as $\xi_{x,t+1}$.

Denote by $\varepsilon_{xy,t}$ the elasticity of $x_t$ to $y_t$, $\varepsilon_{xyt} \equiv d\log(x_t)/d\log(y_t)$. Let $\varepsilon^u_t$ and $\varepsilon^c_t$ be the uncompensated and compensated labor supply elasticities to the net wage, holding both savings and training fixed.[22]

_____

[21] The wedges are here written with a recursive shorthand notation, replacing the future stream of marginal benefits by the future marginal cost, which holds at interior solutions.

[22] I.e., $\varepsilon^c$ and $\varepsilon^u$ are defined as in the static framework (e.g., Saez, 2001), at constant savings and training time:

$$\varepsilon^u = \frac{\phi_l(l,i)/l + \frac{\phi_l(l,i)^2}{u'(c)^2}u''(c)}{\phi_{ll}(l,i) - \frac{\phi_l(l)^2}{u'(c)^2}u''(c)} \qquad \varepsilon^c = \frac{\phi_l(l,i)/l}{\phi_{ll}(l,i) - \frac{\phi_l(l)^2}{u'(c)^2}u''(c)}$$

With per-period utility separable in consumption and labor, $\frac{\varepsilon^c_t}{1+\varepsilon^u_t-\varepsilon^c_t}$ is the Frisch elasticity of labor.

## 2.4.2 Optimal Subsidy for Human Capital Expenses

**The Net Subsidy**

To achieve insurance and redistribution in the presence of incentive compatibility constraints, the optimal allocation features a labor wedge and an intertemporal wedge, both of which affect human capital investments. Intuitively, an increase in the implicit labor tax captures part of the return to the agent's investment, while the implicit savings tax encourages substitution from physical capital to human capital as an indirect way to transfer resources intertemporally. Hence, part of the implicit subsidy on human capital expenses merely acts to counterbalance the distortions in the human capital choice caused by other wedges. A useful benchmark is the subsidy which just ensures that the tax system is neutral with respect to human capital. Hence, I define a "net subsidy," as the gross subsidy from which we filter out all the parts that only go toward compensating for the other distortions.

**Definition 3** *Define the net wedge on human capital expenses, $t_{st}$, as:*

$$t_{st} \equiv \frac{\tau_{St} - \tau_{Lt} M_t^{\prime d} + P_t}{\left(M_t^{\prime d} - \tau_{St}\right)\left(1 - \tau_{Lt}\right)} \tag{2.19}$$

$M_t^{\prime d} \equiv M_t^{\prime} - \frac{(1-\xi_{M^{\prime}})}{R(1-\tau_K)} E_t\left(M_{t+1}^{\prime}\right)$ *denotes the dynamic, risk-adjusted cost.*

$P_t \equiv \frac{\tau_K}{R(1-\tau_K)}\left(1 - \tau_{Lt}\right)\left(1 - \xi_{M^{\prime}}^{\prime}\right) E_t\left(M_{t+1}^{\prime}\right)$ *captures the risk-adjusted savings distortion.*

A zero net wedge $t_{st}$ means that the tax system is neutral with respect to human capital. It is equivalent to full dynamic, risk-adjusted deductibility.[23] For instance, in the last period $T$, or in a static problem, $t_{st} = 0$ means that $\tau_{ST} = \tau_{LT} M_T^{\prime}$, i.e., full local contemporaneous deductibility of expenses – akin to the type of deductions we are familiar with (see Bovenberg and Jacobs, 2005). In this dynamic setting, however, three additional effects need to be filtered out. First, the effective marginal cost of investing is dynamic $(M_t^{\prime d})$, since a higher investment at $t$ means a lower cost of reaching any given human capital stock at $t + 1$. Second, the intertemporal resource transfer from human capital investment is also taken into account in $P_t$, which would be zero with no savings distortion or with linear utility. Finally, uncertainty

---

[23]It is not strictly speaking a standard tax deductibility, because wedges are not equivalent to explicit taxes, but this intuition is helpful to grasp the underlying logic, and will often be referred to.

requires a risk-adjustment, embodied in the insurance factor $\xi_{M'}$, since a given deduction is not worth the same in different states given that marginal utility is not constant.

The net wedge is positive if the government wants to encourage human capital acquisition on net, i.e., above and beyond simply ensuring a neutral tax system with respect to human capital. The next proposition shows how the net wedge is set in a recursive fashion at the optimum.

**Proposition 20** *i) At the optimum, the net wedge is given by:*

$$t_{st}^* \left( \theta^t \right) = \frac{\mu \left( \theta^t \right) u_t' \left( c \left( \theta^t \right) \right)}{f^t \left( \theta_t | \theta_{t-1} \right)} \left( 1 - \rho_{\theta s, t} \right) \frac{\varepsilon_{w\theta, t}}{\theta_t} \tag{2.20}$$

*where the multiplier $\mu \left( \theta_t \right)$ can be written recursively as:*

$$\mu \left( \theta^t \right) = \kappa \left( \theta^t \right) + \eta \left( \theta^t \right) \tag{2.21}$$

$$\kappa \left( \theta^t \right) = \int_{\theta_t}^{\bar{\theta}} \left( 1 - g_s \right) \frac{1}{u_t' \left( c \left( \theta^{t-1}, \theta_s \right) \right)} f \left( \theta_s | \theta_{t-1} \right) d\theta_s \tag{2.22}$$

$$\text{with } g_s = u_t' \left( c \left( \theta^{t-1}, \theta_s \right) \right) \lambda_{t-1} \text{ and } \lambda_{t-1} = \int_{\underline{\theta}}^{\bar{\theta}} \frac{1}{u_t' \left( c \left( \theta^{t-1}, \theta_m \right) \right)} f \left( \theta_m | \theta_{t-1} \right) d\theta_m$$

$$\eta \left( \theta^t \right) = t_{st-1}^* \left( \theta^{t-1} \right) \left[ \frac{R\beta}{u_{t-1}' \left( c \left( \theta^{t-1} \right) \right)} \frac{1}{\left( 1 - \rho_{\theta s, t-1} \right)} \frac{\theta_{t-1}}{\varepsilon_{w\theta, t-1}} \left( \int_{\theta_t}^{\bar{\theta}} \frac{\partial f \left( \theta_s | \theta_{t-1} \right)}{\partial \theta_{t-1}} d\theta_s \right) \right]$$

*ii) If assumption (4) holds, then at each history $\theta^t$.[24]*

$$t_{st}^* \left( \theta^t \right) \geq 0 \Leftrightarrow \rho_{\theta s, t} \leq 1$$

*iii) $t_{st}^* \left( \theta^t \right) = 0$ if $u_t' \left( c_t \right) = 1$ and $\theta_t$ is iid.*

*iv) $t_{st}^* \left( \theta^{t-1}, \bar{\theta} \right) = t_{st}^* \left( \theta^{t-1}, \underline{\theta} \right) = 0$, $\forall t$.*

The multiplier $\mu \left( \theta^t \right)$ on the envelope condition is split into two parts. The insurance motive is captured in $\kappa \left( \theta^t \right)$, familiar from the static taxation literature. It would be zero with linear

---

[24] *Assumption 4 can be replaced by the condition $\tau_{Lt} \left( \theta^t \right) > 0$. That $t_{st} = 0$ when $\rho_{\theta s} = 1$ and corollary 1 apply even when the first-order approach is not valid (see the Appendix).*

utility. $g_s$ is the marginal social welfare weight on an agent of type $\theta_s$, measuring the social value of one more dollar transferred to that individual, and $1/\lambda_{t-1}$ is the social cost of public funds.

The novel term $\eta\left(\theta^t\right)$ captures the previous period's net wedge, hence indirectly the previous period's insurance motive, weighted by a measure of ability persistence. Recall that there can be a redistributive motive in the first period if there is initial heterogeneity.[25] This motive remains effective through $\eta\left(\theta^t\right)$, the more so if types are more persistent, but vanishes as skills become less persistent. In the limit, if $\theta_t$ is identically and independently distributed (iid), only the contemporaneous insurance motive $\kappa\left(\theta_t\right)$ matters. If, in addition to iid shocks, utility is linear in consumption, the optimal net subsidy is 0.[26]

The zero distortion at the bottom and top result, familiar for the labor wedge from the static literature, holds here for the net human capital wedge. It does not hold for the gross wedge $\tau_{St}$ (see below), underscoring again that the true incentive effects are captured by $t_{st}$, not $\tau_{St}$.

Importantly, the proposition underlines that the net wedge on human capital is positive if and only if $\rho_{\theta s,t} \le 1$, a point discussed next.

**The Redistributive and Insurance Values of Human Capital**

Together with the aforementioned insurance and persistence channels, the Hicksian coefficient of complementarity $\rho_{\theta s}$ plays a crucial role. The optimal net wedge results from the balance of two effects that human capital has on social welfare. First, because it increases returns to work, it makes leisure less attractive and encourages labor supply, the "Labor Supply Effect." This is a positive effect in the presence of a distortion in the labor decision.[27] At the same time, if $\rho_{\theta s} > 0$, which is equivalent to ability being complementary to human capital in the wage

---

[25]With a non-utilitarian objective, if $\theta_1$ is interpreted as heterogeneity (see section 2.3.3):

$$\mu\left(\theta_1\right) = \int_{\theta_1}^{\bar{\theta}} \frac{1}{u_1'\left(c_1\left(\theta_s\right)\right)} \left(1 - \lambda_0\left(\theta_s\right) u_1'\left(c_1\left(\theta_s\right)\right)\right) f\left(\theta_s\right)$$

where $\lambda_0\left(\theta_s\right)$ is the multiplier (scaled by $f\left(\theta_s\right)$) on type $\theta_s$ target utility. With linear utility: $1 = \int_{\underline{\theta}}^{\theta} \lambda_0\left(\theta_s\right) f\left(\theta_s\right)$.

[26]Except at $t = 1$ with a non-utilitarian objective, see the previous footnote.

[27]This effect would disappear if there were no distortion on labor, ie., if $\mu\left(\theta^t\right) = 0$. This is different from the direct effect of human capital on earnings, through the increase in wage, which would exist even with no distortion on labor, or with fixed labor supply, and which is filtered out from the net subsidy.

$\left( \frac{\partial^2 w}{\partial \theta \partial s} > 0 \right)$, human capital mostly benefits already able agents, and hence compounds existing inequality due to intrinsic differences in $\theta_t$. The opposite occurs if $\rho_{\theta s} < 0$, in which case human capital reduces inequality. This effect will be labeled the "Inequality effect." Because ability is stochastic, it is equivalent to say that if $\rho_{\theta s} > 0$, human capital increases exposure to risk, because it mostly benefits agents when they receive high productivity shocks.

Technically, the inequality effect is the result of a rent transfer, which arises from the need to satisfy agents' incentive compatibility constraints. If high productivity agents benefit more from a marginal increase in human capital than lower productivity agents ($\rho_{\theta s} > 0$), an increase in their human capital tightens their incentive constraints. What matters for social welfare is whether, when encouraging human capital, the increase in resources from more labor is completely absorbed by the compensation forfeited to high productivity agents to satisfy their incentive compatibility constraints, or whether there are resources left over. The answer is that, when $\rho_{\theta s} < 1$, i.e., when high ability agents do not disproportionately benefit from human capital, the positive labor supply effect dominates any potential inequality effect, and generates net resources to be used for redistribution and insurance of all agents. Hence, it will be beneficial to encourage human capital on net. In this case, human capital is said to have a positive insurance effect, or a positive redistributive effect on after-tax income inequality.[28] Put differently, if $\rho_{\theta s} > 0$, a human capital subsidy increases pre-tax income inequality, but as long as $\rho_{\theta s} \leq 1$, after-tax income inequality will be reduced thanks to the additional resources generated by human capital investment.

Although the human capital and labor literatures have long studied the interaction between human capital and ability (see the summary in section 2.5.1), the optimal taxation literature has mostly adopted the more restrictive wage form $w = \theta s$.[29] This special case entails $\rho_{\theta s} = 1$ and,

---

[28] Importantly, the social objective assigns non-negative weights to all agents, and hence all consumption gains arising from higher resources are positively weighted, even if they increase inequality. Any Pareto improving rise in human capital would be stimulated. But the subsidy does not encourage rises in human capital, which benefit some agents at the expense of having to draw resources from other agents, with a resulting negative change in total social welfare.

[29] With the exception of the references in the introduction (Bovenberg and Jacobs, 2011, Maldonado, 2007, DaCosta and Maestri, 2007). This generalized dynamic model nests their findings. In particular, it emphasizes the role of the distribution of productivity and its persistence over time, which can modulate, amplify, or fully dampen the effects of the complementarity between human capital and ability. In addition, I show below how the *evolution* of the net wedge is affected by the coefficient of complementarity, something which could not be seen in a static model.

hence, a null net wage at the optimum. This is reminiscent of the Atkinson and Stiglitz (1976) result on the non-optimality of differential commodity taxation if preferences satisfy a form of separability between goods and labor. Note that the zero net subsidy result for $\rho_{\theta s} = 1$ does not depend on the optimality of the labor or intertemporal wedges.[30] Indeed, when $\rho_{\theta s} = 1$, the choice of education does not reveal any additional information on ability, as all types benefit equally from it in proportional terms. This benchmark case is discussed next.

**Full Risk-adjusted Dynamic Deductibility**

With any multiplicatively separable wage of the form $w(\theta, s, z) = w^1(\theta) w^2(s, z)$, where $w^1$ and $w^2$ are increasing functions, the gross wedge $\tau_{St}$ only compensates for the distortion caused by income and savings taxes. The following corollary shows the implicit dynamic deductibility relation that the gross wedge, the labor tax, and the capital tax need to satisfy at the optimum.[31]

**Corollary 1** *If $\rho_{\theta s} = 1$, the constrained efficient human capital wedge $\tau_{St}^*$, labor wedge $\tau_{Lt}^*$, and capital wedge $\tau_{Kt}^*$ must satisfy:*

$$\tau_{St}^* = \left( M_t' - \frac{1}{R} E_t\left(M_{t+1}'\right) \right) \tau_{Lt}^* - \frac{\tau_{Kt}^*}{R\left(1 - \tau_{Kt}^*\right)} \left(1 - \xi_{M',t+1}'\right) E_t\left(M_{t+1}'\right) \tag{2.23}$$

Applying this corollary to a static model or to time $t = T$ shows why Bovenberg and Jacobs (2005) find that education subsidies and income taxes are "Siamese Twins," and that the optimal (linear) subsidy on education is equal to the optimal (linear) tax rate. In a lifecycle setting with uncertainty, additional considerations, namely the dynamic nature of the investment cost, the intertemporal wedge, and the insurance factor, complicate this simple contemporaneous deductibility. Section 2.6 maps this result into a deferred deductibility scheme with explicit tax instruments.

**The Capital Wedge with Observable Human Capital**

The presence of observable human capital does not change a standard result in dynamic moral hazard models with observable savings and separable utility, namely the Inverse Euler Equation.

---

[30] As in the direct proofs of the Atkinson-Stiglitz result by Kaplow (2006) and Laroque (2005).

[31] See the general formula for $\rho_{\theta s} \neq 1$, as well as an alternative formulation in the Appendix.

**Proposition 21** *At the optimum, the inverse Euler Equation holds:*

$$\frac{R\beta}{u_t'\left(c\left(\theta^t\right)\right)} = \int_{\underline{\theta}}^{\bar{\theta}} \frac{1}{u_{t+1}'\left(c\left(\theta^{t+1}\right)\right)} f^{t+1}\left(\theta_{t+1}|\theta_t\right) d\theta_{t+1} \tag{2.24}$$

Equation (2.24), combined with Jensen's inequality, implies that the savings wedge in (2.16) is positive, i.e., savings are discouraged. Indeed, saving and shirking in the future are complements. Human capital is an alternative way to transfer resources to the future, and a substitute to savings through physical capital (Heckman, 1976). Hence, in the absence of additional redistributive or insurance effects from human capital ($\rho_{\theta s} = 1$), formula (2.23) shows that the human capital subsidy and the savings wedge co-move inversely.

### 2.4.3   The Optimal Labor Wedge with Observable Human Capital

The optimal labor wedge is very similar to the one in dynamic taxation models without human capital, except for the novel wage function.

**Proposition 22** *i) At the optimum, the labor wedge is equal to:*

$$\frac{\tau_{Lt}^*\left(\theta^t\right)}{1 - \tau_L^*\left(\theta^t\right)} = \frac{\mu\left(\theta^t\right) u_t'\left(c\left(\theta^t\right)\right)}{f^t\left(\theta_t|\theta_{t-1}\right)} \frac{\varepsilon_{w\theta,t}}{\theta_t} \frac{1 + \varepsilon_t^u}{\varepsilon_t^c} \tag{2.25}$$

*with* $\mu\left(\theta^t\right) = \eta\left(\theta^t\right) + \kappa\left(\theta^t\right)$ *as in (2.21), where* $\eta\left(\theta^t\right)$ *can be rewritten recursively as a function of the past labor wedge,* $\tau_{L,t-1}$:

$$\eta\left(\theta^t\right) = \frac{\tau_{Lt-1}^*\left(\theta^{t-1}\right)}{1 - \tau_{L,t-1}^*\left(\theta^{t-1}\right)} \left[\frac{R\beta}{u_{t-1}'\left(c\left(\theta^{t-1}\right)\right)} \frac{\varepsilon_{t-1}^c}{1 + \varepsilon_{t-1}^u} \frac{\theta_{t-1}}{\varepsilon_{w\theta,t-1}} \int_{\theta_t}^{\bar{\theta}} \frac{\partial f\left(\theta_s|\theta_{t-1}\right)}{\partial \theta_{t-1}} d\theta_s\right]$$

*ii)* $\tau_{Lt}^*\left(\theta^{t-1}, \bar{\theta}\right) = \tau_{Lt}^*\left(\theta^{t-1}, \underline{\theta}\right) = 0, \forall t.$

Unlike in the standard taxation model, the wage elasticity with respect to ability $\varepsilon_{w\theta,t}$ is not constant at 1; a higher elasticity amplifies the labor wedge as it increases the value of insurance and redistribution.[32] However, the standard zero distortion at the bottom and the top results

---

[32]E.g.: with a CES wage as in (2.3), $\varepsilon_{w\theta} = \left(\frac{w_t(\theta_t, s_t, z_t)}{\theta_t}\right)^{\rho-1}$; human capital indirectly enters the optimal tax formula.

from the static Mirrlees model continue to apply in the presence of observable human capital. The labor wedge at any age is inversely related to the elasticity of labor supply that prevails at that time.

In this setting, both the labor wedge and the net human capital wedge are tools for redistribution and insurance. The following relation shows that they are set according to a type of inverse elasticity rule, i.e., inversely proportional to their efficiency costs:

**Corollary 2** *At the optimum, the labor wedge and human capital wedge need to satisfy the following relation after each history:*

$$t_{st}^* = \left( \frac{\tau_{Lt}^*}{1 - \tau_{Lt}^*} \right) \frac{\varepsilon_t^c}{1 + \varepsilon_t^u} \left( 1 - \rho_{\theta s, t} \right) \tag{2.26}$$

While the labor wedge always has a positive redistributive or insurance value,[33] the net human capital wedge only has a positive redistributive value if and only if $(1 - \rho_{\theta s}) > 0$. The optimal policies must be consistent with each other: if the labor wedge is higher so as to provide more insurance, the net human capital wedge must also be higher if and only if $(1 - \rho_{\theta s}) > 0$.

In the special case of a CES wage function as in (2.3) and a separable isoelastic disutility:

$$\phi(l, i) = \frac{1}{\gamma} l^\gamma + \frac{1}{\eta} i^\eta \quad (\gamma > 1, \eta > 1) \tag{2.27}$$

the ratio of the net human capital wedge and labor wedge is constant over time and across all agents, and equal to:

$$t_{st}^* / \left( \frac{\tau_{Lt}^*}{1 - \tau_{Lt}^*} \right) = \frac{(1 - \rho)}{\gamma}$$

### 2.4.4 Age-dependency

This section focuses on the lifetime evolution of the optimal wedges. In formulas (2.20) and (2.25), the optimal wedges after each history were expressed recursively as a function of the previous period's wedges. Instead of these point-wise expressions, one can also rewrite the formulas in terms of a weighted expectation across types at time $t$, using some weighting function $\pi(\theta)$. Different weighting functions $\pi(\theta)$ lead to different recursive relations, which

---

[33] As long as $\tau_{Lt} \geq 0$, i.e., $\mu(0^t) \geq 0$, which is true under assumption (4).

must hold at the optimum, and some weighting functions draw out particularly enlightening effects.[34]

For the exposition only, ability is assumed to follow a log autoregressive process:

$$\log(\theta_t) = p\log(\theta_{t-1}) + \psi_t \tag{2.28}$$

where $\psi_t$ has density $f^\psi(\psi|\theta_{t-1})$, with $E(\psi|\theta_{t-1}) = 0$. The general formula without this assumption is in the Appendix (formula (2.56)):

**Corollary 3** *The labor wedge evolves over time according to:*

$$E_{t-1}\left(\frac{\tau_{Lt}}{(1-\tau_{Lt})}\frac{\varepsilon_{w\theta,t-1}}{\varepsilon_{w\theta,t}}\frac{\varepsilon_t^c}{1+\varepsilon_t^u}\frac{1+\varepsilon_{t-1}^u}{\varepsilon_{t-1}^c}\left(\frac{1}{R\beta}\frac{u'_{t-1}}{u'_t}\right)\right)$$
$$= \varepsilon_{w\theta,t-1}\frac{1+\varepsilon_{t-1}^u}{\varepsilon_{t-1}^c}Cov\left(\frac{1}{R\beta}\frac{u'_{t-1}}{u'_t},\log(\theta_t)\right) + p\frac{\tau_{Lt-1}}{(1-\tau_{Lt-1})} \tag{2.29}$$

Dynamic incentive compatibility constraints cause a positive covariance between consumption growth and productivity: by promising them higher consumption growth, the government induces higher ability agents to truthfully reveal their types. This however makes insurance valuable and is captured by the drift term. The insurance motive is scaled down by the efficiency cost of the labor wedge, $\varepsilon_t^c/(1+\varepsilon_t^u)$, which makes provision of insurance more expensive, and magnified by the sensitivity of the wage to stochastic ability $\varepsilon_{w\theta,t}$, which increases the value of insurance. The persistence of the shock $p$ translates into a persistence for the labor wedge. But the relation is not one-for-one: the autocorrelation of the labor wedge is modulated by the growth in efficiency costs and changes in the sensitivity of the wage to ability over time. Age-varying labor supply elasticities are a commonly stated argument in favor of age-dependent taxation. A time-varying elasticity of the wage to ability can provide a further rationale for age-contingent taxes.

A similar recursive formulation can be derived for the lifetime evolution of the net human capital wedge. Because of its tight link with the labor wedge (as explained in formula (2.26)), many of the same forces driving the evolution of the labor wedge are also important for the net

---

[34]Such a reformulation for the optimal labor wedge formula was proposed by Farhi and Werning (2013) for a convenient weighting function $\pi(\theta) = 1$.

subsidy.

**Corollary 4** *The optimal net subsidy evolves over time according to:*

$$E_{t-1}\left(t_{st}\frac{\varepsilon_{w\theta,t-1}}{\varepsilon_{w\theta,t}}\frac{\left(1-\rho_{\theta s,t-1}\right)}{\left(1-\rho_{\theta s,t}\right)}\left(\frac{1}{R\beta}\frac{u'_{t-1}}{u'_{t}}\right)\right)$$

$$=\varepsilon_{w\theta,t-1}\left(1-\rho_{\theta s,t-1}\right)Cov\left(\frac{1}{R\beta}\frac{u'_{t-1}}{u'_{t}},\log\left(\theta_{t}\right)\right)+pt_{st-1}\quad(2.30)$$

Formula (2.30) again exhibits a drift term capturing insurance concerns, magnified by the sensitivity of the wage to ability $\varepsilon_{w\theta,t}$, and the redistributive or insurance factor of human capital $(1-\rho_{\theta s})$. The drift term inherits the sign of the latter; when $\rho_{\theta s}\leq1$, human capital has a positive insurance effect, which caters well to the rising need for insurance. The simulations in section 2.5 show that, indeed, the net subsidy rises with age when human capital has a positive insurance value ($\rho_{\theta s}\leq1$), and falls when it has a negative insurance value ($\rho_{\theta s}\geq1$). They also highlight a "subsidy smoothing" result: the net subsidy becomes more strongly correlated over time as age increases, because the variance of consumption growth falls to zero, which makes the drift term in the formula above vanish.

The Hicksian complementarity $\rho_{\theta s}$ might vary over life, as suggested by the empirical literature in section 2.5.1. If it is decreasing faster, the net subsidy will rise faster or fall slower over the lifecycle. The sensitivity of the wage amplifies the effect of the Hicksian complementarity coefficient in either direction.

Finally, note that all results from Propositions 20, 21, and 22, and Corollaries 1, 2, 3, and 4 still apply with a moving support $\left[\underline{\theta}_{t}\left(\theta_{t-1}\right),\bar{\theta}_{t}\left(\theta_{t-1}\right)\right]\subseteq\Theta$, except the zero distortions at the top and bottom results for $\tau_{L}$ and $t_{s}$.[35]

### 2.4.5 Optimal Bonus for Training Time

The foregoing results described the optimal subsidy for human capital expenses. Should training time be treated differently than expenses? This section analyzes the optimal bonus for training time, and focuses on the peculiarities of training relative to expenses, namely its direct

---

[35]See Farhi and Werning (2013) for the analysis of the optimal tax with moving support, which can easily be extended here to the model with human capital.

interaction with labor supply.

As in subsection 2.4.2, a net wedge, the "net bonus," is defined. It measures the deviation away from a tax system that is neutral with respect to training time. However, the cost being deducted is agents' disutility cost, converted into a monetary cost $(\phi_{z,t}/u')$, and again accounting for the dynamic nature of investment.

**Definition 4** *The net wedge or net bonus on training time $t_{zt}$ is defined as:*

$$t_{zt} \equiv \frac{\tau_{Zt} - \tau_{Lt}\left(\frac{\phi_{z,t}}{u'_t(c_t)}\right)^d + P_{Zt}}{\left(\left(\frac{\phi_{z,t}}{u'_t(c_t)}\right)^d - \tau_{Zt}\right)(1 - \tau_{Lt})} \tag{2.31}$$

$\left(\frac{\phi_{z,t}}{u'_t(c_t)}\right)^d \equiv \frac{\phi_{z,t}}{u'_t(c_t)} - \frac{1}{R(1-\tau_K)}\left(1 - \xi_{\phi'/u',t+1}\right)E_t\left(\frac{\phi_{z,t+1}}{u'_{t+1}(c_{t+1})}\right)$ *is the dynamic risk-adjusted disutility cost, converted into monetary units.*

$P_{Zt} \equiv \frac{\tau_K}{R(1-\tau_K)}\left(1 - \xi'_{\phi'/u',t+1}\right)(1 - \tau_{Lt})E_t\left(\frac{\phi'_{z,t+1}}{u'_{t+1}(c_{t+1})}\right)$ *is the risk-adjusted savings distortion.*

The following proposition characterizes the net bonus at the optimum.

**Proposition 23** *At the optimum, the net bonus is given by:*

$$t^*_{zt}\left(\theta^t\right) = \frac{\tau^*_{Lt}\left(\theta^t\right)}{1 - \tau^*_{Lt}\left(\theta^t\right)}\frac{\varepsilon^c_t}{1 + \varepsilon^u_t}\left(1 - \rho_{\theta z,t} - \frac{\varepsilon_{\phi z,t}}{\varepsilon_{wz,t}}\rho^\phi_{lz,t}\right) + \frac{1}{R}E_t\left(\frac{\tau^*_{Lt+1}\left(\theta^{t+1}\right)}{1 - \tau^*_{Lt}\left(\theta^t\right)}\frac{\varepsilon^c_{t+1}}{1 + \varepsilon^u_{t+1}}\frac{w_{z,t+1}}{w_{z,t}}\frac{l_{t+1}}{l_t}\frac{\varepsilon_{\phi z,t+1}}{\varepsilon_{wz,t+1}}\rho^\phi_{lz,t+1}\right) \tag{2.32}$$

*with $\varepsilon_{\phi z,t} = d\log\left(\phi_t\right)/d\log\left(z_t\right)$ the elasticity of disutility with respect to training time. $\rho^\phi_{lz,t}$ is the Hicksian coefficient of complementarity between $l$ and $z$ in the disutility $\phi_t$. $\tau^*_{Lt}\left(\theta^t\right)$ and $\tau^*_{Lt+1}\left(\theta^{t+1}\right)$ are at their optimal levels from (2.25).*

First, the Labor Supply Effect and the Inequality Effect are again present, with exactly the same interpretation as in subsection 2.4.2. But the bonus on training time has an additional direct interaction effect with current and future labor supply through the disutility function.

A first natural conjecture is that contemporaneous training and labor supply are substitutes: time spent acquiring human capital cannot be spent working. In addition, "fatigue" effects can set in. In this case, the Hicksian complementarity between training and labor in the disutility

112

function is positive $(\rho^{\phi}_{lz,t} > 0)$, a case labeled "Learning-or-Doing." However, because $\phi(l,i)$ is a disutility cost – not purely an opportunity cost of time – it might also be that labor supply and training are complements at least over some range $(\rho^{\phi}_{lz,t} < 0)$. For example, the guarantee of regular training, which gives workers a prospective for progress, might boost motivation and make labor effort seem less painful, a situation called "Learning-and-Doing."

In the formula, $\left(1 - \frac{\varepsilon_{\phi z}}{\varepsilon_{wz}}\rho^{\phi}_{lz,t}\right)$ is the total effect of the bonus on labor, i.e., the sum of the Labor Supply Effect, which simulates labor because of the higher wage, and either one of the Learning-and-Doing or Learning-or-Doing effects, which can increase or decrease work. The total effect of a training subsidy on contemporaneous labor is positive if and only if

$$\varepsilon_{wz} > \varepsilon_{\phi_l z} \quad \text{with } \varepsilon_{\phi_l z} \equiv \partial \log(\phi_l) / \partial \log(z)$$

i.e., if and only if the wage is more sensitive to training than the marginal disutility of work is.[36]

Even if training diverts time away from contemporaneous labor supply, the effects on future labor supply can motivate a positive net subsidy. Whatever the pattern of complementarity or substitutability between contemporaneous labor supply and training, the relation between current training and future labor supply is its mirror image. If the former are complements, the latter are substitutes, and vice versa, because investing in human capital today means having to invest less tomorrow to reach any given level.[37]

If there is Learning-or-Doing, the net bonus co-moves positively with the future income tax rate $\tau_{Lt+1}$, but negatively with the current tax rate $\tau_{Lt}$. Intuitively, if there is a higher contemporaneous wedge on labor, there already is an indirect stimulus to training because training is a substitute for labor; hence the need for an additional stimulus through a net bonus is reduced. However, a higher future labor wedge is a stimulus for training in the future, which is a substitute for training today. Hence, to stimulate training today, there is a need for a higher net bonus.

Corollary 5 illustrates two cases, among other possible ones.

---

[36] Note that: $(\varepsilon_{\phi z,t}/\varepsilon_{wz,t})\,\rho^{\phi}_{lz,t} = \varepsilon_{\phi_l z,t}/\varepsilon_{wz,t}$.

[37] This is because only the flow $i_t = z_t - z_{t-1}$, and not the stock $z_{t-1}$ enters the disutility. This reasoning is also only valid at interior solutions for training.

**Corollary 5** *Under assumption 4:*

*i) The net bonus is positive if* $\left(1 - \frac{\varepsilon_{\phi z}}{\varepsilon_{wz}} \rho_{lz,t}^{\phi}\right) \geq \rho_{\theta z,t}$ *and* $\rho_{lz,t}^{\phi} \geq 0.$

*ii) The net bonus is negative if* $\left(1 - \frac{\varepsilon_{\phi z}}{\varepsilon_{wz}} \rho_{lz,t}^{\phi}\right) \leq \rho_{\theta z,t}$ *and* $\rho_{lz,t}^{\phi} \leq 0.$

The sufficient conditions in i) guarantee that, although training diverts time away from contemporaneous labor effort, it stimulates labor supply in the future and has a sufficiently large positive redistributive or insurance effect $(1 - \rho_{\theta z})$. On the other hand, the sufficient conditions in ii) imply that despite stimulating contemporaneous labor supply through Learning-and-Doing, training disproportionately benefits high productivity workers, and has a negative redistributive or insurance effect.

What this analysis also highlights is that time and money costs are not equivalent in the presence of asymmetric information and incentive problems on unobservable labor. Unlike in the full information case, disutility costs cannot simply be converted into monetary units and treated as equivalent to expenses. An exception case is when the disutility function is separable in labor and human capital $\left(\frac{\partial^2 \phi}{\partial z \partial l} = 0\right)$. Then, as corollary (6) shows, the forces governing the net bonus on training are the same as those driving the net subsidy on human capital expenses.

**Corollary 6** *i) If* $\frac{\partial^2 \phi_t}{\partial z \partial l} = 0$ *and assumption 4 holds:*

$$t_{zt}^* \left(\theta^t\right) \geq 0 \Leftrightarrow \rho_{\theta z,t} \leq 1$$

*ii) The following relations between the wedges hold at the optimum at every history* $\theta^t$:

$$t_{zt}^* = \frac{\tau_{Lt}^*}{1 - \tau_{Lt}^*} \frac{\varepsilon_t^c}{1 + \varepsilon_t^u} \left(1 - \rho_{\theta z,t}\right) \tag{2.33}$$

$$\frac{t_{zt}^*}{t_{st}^*} = \frac{\left(1 - \rho_{\theta z,t}\right)}{\left(1 - \rho_{\theta s,t}\right)} \tag{2.34}$$

*ii) If in addition the wage function is CES with parameter* $\rho$ *as in* (2.3):

$$t_{st}^* = t_{zt}^* \tag{2.35}$$

First, when there no longer is a direct interaction with labor supply, the net stimulus to

114

training, i.e., the net bonus, is positive if and only if training has a positive insurance effect $(\rho_{\theta z} \leq 1)$. Formula (2.33) shows that in this case, the training wedge and the labor tax evolve simply according to a type of inverse elasticity rule. Formula (2.34) highlights that both types of human capital need to be encouraged proportionally to their redistributive and insurance effects $(1 - \rho_{\theta z})$ and $(1 - \rho_{\theta s})$. With a CES wage function, these redistributive effects are the same: there is then no reason to stimulate one type of human capital more than the other. At the optimum, the net wedges in (2.35) fully equate the incentives to invest in both types of human capital.

## 2.5 Numerical Analysis

In the numerical analysis in this Section, I take a middle stand between a simple illustration of the qualitative features of the optimal mechanism and a more careful calibration with quantitative implications for the optimal wedges and their lifecycle patterns. The focus is on human capital expenses. Additional details of the computational procedure and the calibration, as well as results for alternative calibrations, are in the online Computational Appendix. Before presenting simulation results, the empirical evidence on the crucial parameter of the model, namely the complementarity between human capital and ability in the wage is discussed.

### 2.5.1 Empirical Evidence on the Complementarity between Human Capital and Ability

The formulas for the optimal net wedge and its evolution (see (2.19) and (2.30)) highlighted the importance of the complementarity between ability and human capital. The benchmark Mincer model of returns to schooling yields a log-linear relation between the wage and schooling, with homogeneous returns across ability levels. The Becker (1964) model, reformulated and developed by Card (1995a), allows for returns to education to vary with ability. When put to work empirically though, the goal is frequently to circumvent the problem of unobserved ability. Instrumental variables related to institutional changes for schooling (Angrist and Krueger, 1991, Card, 1995b), or matching between siblings or twins (see Ashenfelter and Krueger, 1994, Ashenfelter and Rouse, 1998) have been exploited. Ashenfelter and Rouse (1998) suggest that

lower ability children benefit more from schooling, a finding also in line with Cunha *et al.* (2006) for early childhood interventions. This is consistent with $\rho_{\theta s} \leq 1$ for childhood investments.[38]

Importantly, however, the Hicksian complementarity can change over life, as suggested by the structural literature on human capital formation (see among others Cunha *et al.*, 2005, Cunha and Heckman, 2007, 2008), so that estimates for primary and secondary schooling could be of limited use for the analysis of higher education or job training. Several studies show that college education might mostly benefit already able students, implying that $\rho_{\theta s} > 0$, and that $\rho_{\theta s} > 1$ is possible. Cunha *et al.* (2006) (hereafter, CHLM, 2006) estimate that the return to one year of college is around 16% at the $5^{th}$ percentile of the math test scores distribution, as opposed to 26% at the $95^{th}$ percentile. There is only scarce evidence on the complementarity between on-the-job training and ability, but the same authors show that on-the-job training is mostly taken up by those with higher AFQT scores, which might, all else equal, signal that they have a higher marginal return from it.[39] The OECD (2004) reports that training mostly benefits skilled workers in terms of higher wages, but benefits low-skilled workers in terms of job security. Huggett *et al.* (2011) use a multiplicatively separable functional form for the wage in their structural model of time investments in human capital (implying $\rho_{\theta z} = 1$), which generates a lifecycle path of earnings that matches the data well.

### 2.5.2 Calibration

**Calibration to US data:** To calibrate the model, I construct a "baseline economy." The baseline economy has the same primitives as in Section 2.2, but no social planner and, hence, no optimal tax system. Instead, the linear labor taxes, capital taxes, and human capital subsidies are set to their current averages in the US. Public and private subsidies in the US cover around 50% of total resource costs of formal higher education. However, in the model, human capital expenses are a comprehensive measure, including all types of formal and informal investments, and some mostly unsubsidized expenses (e.g.: textbooks, computers). In addition, there are practically no subsidies beyond the initial 4 years of higher education. Hence, the linear subsidy

---

[38] However, since schooling - as well as college and job training – builds the two stocks of human capital $z$ and $s$, this finding could be explained by several configurations of the parameters $\rho_{\theta s}$ and $\rho_{\theta z}$.

[39] It is also not evident that the test scores used as measures of "ability" are themselves exogenous, especially at later ages.

in the baseline model, applicable to all expenses, is reduced to 35% for the first 2 periods and to 0 thereafter.[40] The linear labor tax rate is set to 13%, and the savings tax to 25%.[41] Agents can borrow and save at a constant interest rate, subject to a debt limit. In this baseline economy, some parameters are set exogenously based on the existing literature, while others are set to endogenously match two key moments from the data, namely, a wage premium and a ratio of human capital expenses to lifetime income, as explained in more detail next.

**Functional Forms:** Agents live for $T = 30$ periods, each representing roughly 2 years of life. They work for 20 periods and spend 10 years in retirement. Preferences during working life are given by:

$$\tilde{u}\left(c_t, y_t, s_t, \theta_t\right) = \log\left(c_t\right) - \frac{\kappa}{\gamma}\left(\frac{y_t}{w_t\left(\theta, s\right)}\right)^{\gamma}, \quad \kappa > 0, \gamma > 1$$

During retirement, utility is simply $\tilde{u}^R\left(c_t\right) = \log\left(c_t\right)$. The aforementioned literature seemed to indicate that $\rho_{\theta s} > 0$. Given the uncertainty about the magnitude of this parameter though, one of the goals in the simulations will be to explore different values, and, in particular, how the results depend on whether it is above or below its critical threshold of 1. To this end, the wage is assumed to be CES

$$w_t\left(\theta, s\right) = \left(\theta^{1-\rho} + c_s s^{1-\rho}\right)^{\frac{1}{1-\rho}} \tag{2.36}$$

with a constant parameter $\rho_{\theta s} = \rho$, where $c_s$ is a scaling factor. Two values of $\rho$ are studied in the main text, namely, $\rho = 0.2$ and $\rho = 1.2$, with additional values considered in the Computational Appendix.[42]

The cost function of human capital contains an adjustment term and takes the form:

$$M_t\left(s_{t-1}, e_t\right) = c_l e_t + c_a \left(\frac{e_t}{s_{t-1}}\right)^2$$

with $c_l$ and $c_a$ the linear and the adjustment components of cost. Consistent with the high

---

[40] See indicators B, in OECD "Education At A Glance, 2013."

[41] Only interest income and short term capital gains are taxed at ordinary income rates. Taxes on a lot of dividends and long term capital gains stop at 20% (plus possibly 2.9% for the new Medicare tax above $250K of total income, and various state taxes).

[42] Parameter $\rho$ affects the scale of the wage. To make the two values of $\rho$ most comparable, for $\rho > 1$, the CES form used is instead: $w_t\left(\theta, s\right) = \left(\theta^{\rho-1} + c_s s^{\rho-1}\right)^{\frac{1}{\rho-1}}$ with $\rho = 1.45$, which implies $\rho_{\theta s} = \left(\frac{2-\rho}{\rho-1}\right) = 1.2$.

persistence in earnings documented in Storesletten *et al.* (2004), ability is assumed to follow a geometric random walk:

$$\log \theta_t = \log \theta_{t-1} + \psi_t$$

with $\psi_t \sim N\left(-\frac{1}{2}\sigma_\psi^2, \sigma_\psi^2\right)$.

*Exogenously calibrated parameters:* The baseline model has $\gamma = 3$ and $\kappa = 1$, which implies a Frisch elasticity of 0.5 (Chetty, 2012). The discount factor is set to $\beta = 0.95$, and the net interest rate to 5%. The adjustment cost is normalized to $c_a = 2$. There is large variation across empirical estimates of the variance of productivity $\sigma_\psi^2$. A medium-range value of 0.0095 (Heathcote *et al.*, 2005, hereafter HSV 2005) is adopted. Although the qualitative features of the solution are unaffected for a wide range of parameters, the quantitative results are. Therefore, several alternative calibrations are explored in the online Computational Appendix.

*Endogenously matched parameters:* The scaling factor for human capital $c_s$ and the linear cost parameter $c_l$ are set to match two statistics in the data: a wage premium and a ratio of human capital expenses to income. One complication which arises is that the model does not *a priori* restrict investments to occur only during the traditional "college" years. Indeed, one of the motivations for this study is the lifelong nature of human capital investments. However, most available estimates for wage premiums in the literature are for college education, with scarce evidence on job training. This difficulty can be overcome by redefining "college" appropriately for the model. Following Autor *et al.* (hereafter AKK, 1998) who find 42.7% full-time college equivalents, the top 42.7% in the population of the baseline economy, ranked by educational expenses, are assumed to represent the real-life college-goers. Their average wage relative to the bottom 42.7% is set to match the wage premium for college estimated in the literature.[43] These estimates range from 1.58 in Murphy and Welch (1992), to between 1.66 and 1.73 for AKK (1998), and above 1.80 in Heathcote *et al.* (2010). The calibration targets a mid-range value of 1.7.

Turning to the second target, the net present value of higher education expenses over the net present value of lifetime income in the data is 13%.[44] However, since agents invest beyond

---

[43]The middle 14.6% are omitted to clearly delineate between college-goers and others. Indeed, because of the continuous investments in the model, there is no sharp distinction between "college" and "no-college," and in particular, no notion of "a degree."

[44]Computed using data on attendance at different types of colleges and costs (Chang Wei, 2010 US Department

118

traditional college years in the model, there needs to be an allowance for later-in-life investments. It is assumed that college costs represent 2/3 of all lifetime investments in human capital, so that the target ratio of the net present value of lifetime human capital expenses and the net present value of lifetime income is 19%. Table 1 summarize the resulting parameters.

Using the computed policy functions, a Monte Carlo simulation with 100,000 draws is performed for each value of $\rho_{\theta s}$. The initial states are set to yield a zero present value resource cost for the allocation. This ensures comparability across simulations, and gives a sense of outcomes achievable without outside government revenue.

|  | Definition | Sim 1 | Sim 2 | Source/Target |
|---|---|---|---|---|
| *Exogenously calibrated or normalized* | | | | |
| $\rho$ | Hicksian complementarity | 0.2 | 1.2 | CHLM (2006) |
| $\kappa$ | Disutility of work scale | 1 | 1 | |
| $\gamma$ | Disutility elasticity | 3 | 3 | Chetty (2012) |
| $\sigma_\psi^2$ | Variance of productivity | 0.0095 | 0.0095 | HSV (2005) |
| $T$ | Working periods | 20 | 20 | |
| $T_r$ | Retirement periods | 10 | 10 | |
| $\beta$ | Discount factor | 0.95 | 0.95 | |
| $R$ | Gross interest rate | 1.053 | 1.053 | |
| *Endogenously calibrated in baseline economy* | | | | |
| $c_s$ | Scale of $HC$ in wage | 0.09 | 0.1 | Wage premium |
|  |  |  |  | (AKK, 1998) |
| $c_l$ | Linear cost | 0.5 | 0.5 | (OECD, 2013, |
|  |  |  |  | US Dept. Educ, 2010) |

TABLE 1

## 2.5.3 Results

A brief note on the presentation of the wedges is required. When agents are at the corner solution of no investment, which occurs in the baseline calibration for most simulated paths

---

of Education, OECD, 2013). See the detailed calculations in the online Computational Appendix. For an earlier period, Caucutt and Kumar (2003) find a slightly lower number of 12%.

119

approximately after period 13 (after 26 years), the subsidy is indeterminate, as long as it remains below an upper bound that does not induce agents to invest. Hence, the policy function for the gross wedge is set to zero for agents after they stop investing.[45] To make it most comparable to an explicit subsidy, it is presented as a fraction of the marginal cost, i.e., $\tau_{St}/C'_t(e_t)$.

**The optimal gross and net wedges as functions of** $\rho_{\theta s}$ : Figure 2-1 presents the human capital wedges. For this figure only, the focus should be on periods $t \leq 13$, as many agents no longer invest in human capital later in life, and wedges are normalized thereafter, as just explained. The theoretical results from Section 2.4 are illustrated here: First, panel (a) shows that the optimal wedge on human capital is higher and grows faster when human capital has a positive insurance or redistributive effect ($\rho_{\theta s} \leq 1$). When $\rho_{\theta s} = 0.2$, the wedge starts from 1% and grows to 19%; for $\rho_{\theta s} = 1.2$, it instead starts at -2% and grows to only 13%. Panel (b) illustrates that with $\rho_{\theta s} = 1.2$, the net subsidy is negative, so that human capital expenses are made less than fully deductible. Conversely, when $\rho_{\theta s} = 0.2$, human capital expenses are subsidized on net beyond pure deductibility. The comparison between panels (a) and (b) once more highlights that the true incentive effect is different from the gross wedge when there are several wedges present. Finally, the net subsidy is growing when $\rho_{\theta s} \leq 1$ and declining otherwise, as seen in the drift term of formula (2.30).

However, the net wedges are very small. Hence the overall system remains very close to neutrality with respect to human capital expenses. Put differently, full dynamic risk-adjusted deductibility is very close to optimal.[46] This is akin to a "production efficiency" result: human capital is an intertemporal decision, with persistent effects, and distorting it for redistributive or insurance reasons is relatively costly (see also Diamond and Mirrlees, 1971, Chamley, 1986, and Judd, 1985) unless the redistributive or insurance effects are very strong.

Because of this, the values chosen for the complementarity between human capital and ability and for the volatility of ability clearly matter. Moving further away from a multiplicatively separable wage, i.e., increasing the complementarity coefficient further away from 1 in either

---

[45]The net wedge cannot be innocuously normalized to zero; hence, when agents no longer invest, it is set to the right hand side in (2.19), a level that will not artificially induce agents to invest.

[46]The gross wedge is correspondingly smaller than 50% real world subsidy. Bear in mind, however, that the gross wedge here is more comprehensive, as it covers all human capital expenses, even those unsubsidized in the real world, and is available throughout life, not exclusively for college years.

direction, leads to larger net wedges in absolute value. Similarly, a higher volatility increases the value of insurance and yields a higher optimal net wedge if human capital has a positive insurance value ($\rho_{\theta s} < 1$) and a lower optimal net wedge if not. Unfortunately, the estimates of these parameters from the literature are surrounded by large uncertainty. Therefore, the online Computational Appendix provides alternative calibrations.

The production efficiency result also stems from the fact that the government jointly optimizes the full system, and insurance can also be achieved through the labor wedge. In particular, the planner is able through transfers to endogenously relieve "credit constraints," which might be motivating high subsidies in the real world, without having to distort human capital acquisition *at the margin*.[47]

In addition, absent from this model are positive spillovers from human capital on growth, political stability, or social cohesion, often studied in the literature, and pointed to by policy makers as reasons to subsidize education. Finally, the belief about $\rho_{\theta s}$ held by policy makers, and ultimately, society, could be very different from the current estimates of $\rho_{\theta s}$ in the literature. In particular, one way to rationalize the higher subsidies in the real world is that human capital is believed to have very strong redistributive and insurance values, and strongly benefit lower ability people. If the model is taken literally, a very negative $\rho_{\theta s}$ would be needed. Related to this, education is often perceived as a basic human right.

**Optimal labor wedges:** Figure $2 - 2$ panel (a) explains why the gross and net human capital wedges differ so starkly: the optimal labor wedge rises over time to provide insurance against widening income dispersion, and a large part of the growing human capital wedge merely goes towards compensating for this growing disincentive to accumulate human capital. The labor wedge is compared to the one in a standard dynamic taxation model without human capital, in which the wage is equal to exogenous ability, $w_t = \theta_t$. As is intuitive, the wedge grows slower in the presence of human capital, particularly when $\rho_{\theta s} \leq 1$. The labor wedge has a disincentive effect on human capital acquisition, which is undesirable, and the more so when human capital itself has positive insurance and redistributive effects ($\rho_{\theta s} \leq 1$).

---

[47] A related reason is that, although it is optimal to invest more in human capital early in life in the model, there is no sharp exogenous constraint which forces investments to occur exclusively during college years. In practice, it seems that the signaling value of a college degree makes it critical to acquire human capital in a very concentrated manner at the beginning of working life.

## Figure 2-1: Average human capital wedges over time

### (a) Gross Wedge $\tau_{St}$



### (b) Net Wedge $t_{st}$
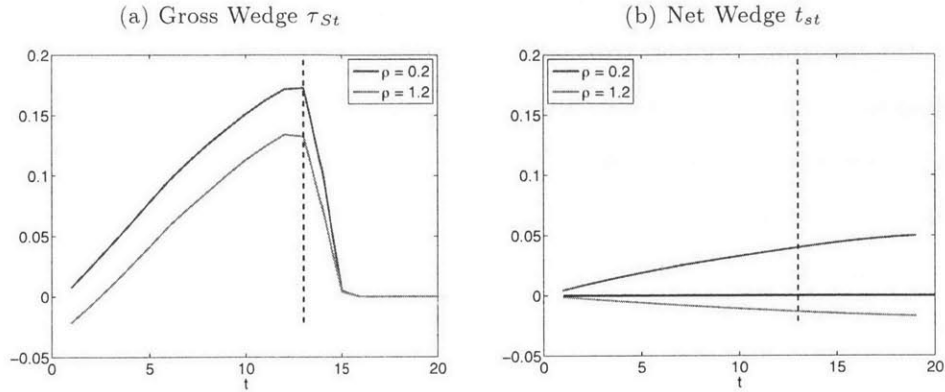


Dashed lines mark the time after which most agents no longer invest in human capital. (a) The gross wedge is higher and grows faster when human capital has positive redistributive and insurance effects ($\rho_{\theta s} < 1$). The wedge is normalized to zero for zero investment (a corner solution). (b) The provision of dynamic incentives also creates a value for insurance. If $\rho_{\theta s} < 1$, human capital has positive redistributive and insurance values, and expenses are subsidized on net at a rising rate. Conversely, if $\rho_{\theta s} > 1$, expenses are only partially deductible, and deductibility decreases over time.

## Figure 2-2: Average labor and capital wedges over time

### (a) Labor Wedge
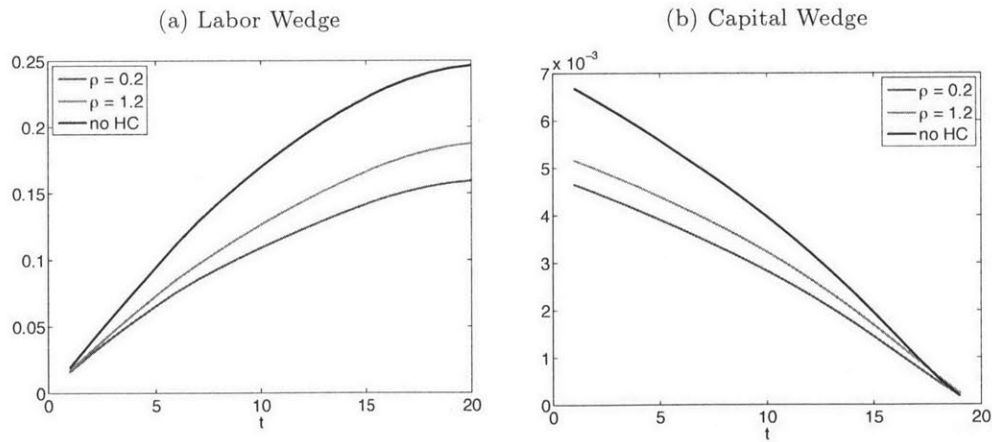


### (b) Capital Wedge



"No HC" denotes the case without human capital. (a) The labor wedge – which represents a disincentive for human capital investments but also insures agents against risky human capital returns – grows slower over time in the presence of human capital, the more so if human capital has positive redistributive and insurance effects ($\rho_{\theta s} < 1$). (b) The capital wedge is still positive, but lower in the presence of human capital.

122

**Optimal capital wedges:** Figure 2−2 panel (b) plots the capital wedge over time. It starts at 0.5% of the gross interest on savings, which corresponds to a 10% tax on net interest, and declines to zero.[48] The capital wedge arises at the optimum from the inverse Euler equation in (2.24), and is a standard feature in dynamic moral hazard models, in which savings are complementary to future shirking. However, savings are less distorted in the presence of human capital. One possible intuition for this is that human capital investments improve productivity in the future, which to some extent counters incentives to shirk. The capital wedge without human capital is 0.7% of gross interest, equivalent to 14% of net interest.

**Subsidy Smoothing:** Figure 2−3 plots the net subsidy in period $t$ against the net subsidy at $t−1$, for young adults ($t = 5$ in panel (a)), and for middle-aged workers ($t = 13$ in panel (b)). Earlier in life, the net wedge is more volatile from one period to the next, but becomes more deterministic over time, leading to a "subsidy smoothing" result. The dynamic taxation literature has highlighted a similar "tax smoothing" result for the labor wedge, which also applies in the presence of human capital (see the online Computational Appendix). The intuitions for these results are the same. A persistent productivity shock early in life has repercussions over many periods, leading to a larger present value change in the income flow than a later shock. Consumption in early years will react strongly to unexpected changes in ability, as the agent attempts to smooth out the shock. Accordingly, the variance of consumption growth is initially large, but decreases to zero over time. The drift term in the net subsidy formula (see (2.30)), which is proportional to the covariance between ability and consumption growth, tends to zero towards retirement. Then, only the autoregressive term of the random walk remains.

**Allocations and Insurance:** Figure 2-6 plots the average allocations over time. Average human capital investments are almost flat and highest early in life, before declining with age. Mean consumption is constant, a result due to the Inverse Euler equation in (2.24), and log utility with $\beta = \frac{1}{R}$, which imply that consumption is a martingale: $E_{t-1}(c_t) = c_{t-1}$. Mean output is increasing, despite the rising labor wedge, because of the growing productivity of agents driven by their endogenous human capital investments.

Figure 2-4 panel (a) shows that lifetime spending on human capital is more tightly linked

---

[48]The equivalent tax on net interest, $\tilde{\tau}_{Kt}$ solves $(1 + (R - 1)(1 - \tilde{\tau}_{Kt})) = R(1 - \tau_{Kt})$.

## Figure 2-3: Subsidy smoothing over life

(a) $t_{st}$ against $t_{st-1}$ at $t = 5$

(b) $t_{st}$ against $t_{st-1}$ at $t = 13$
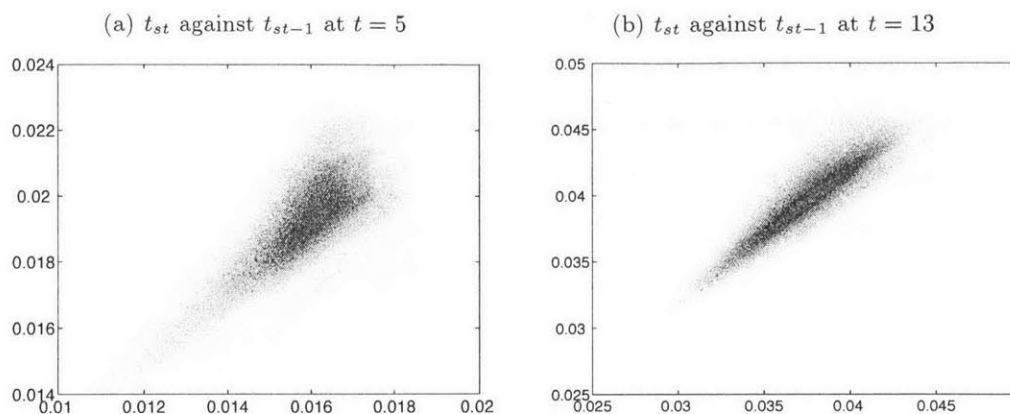


The net human capital wedge becomes more correlated from one period to the next as age increases, because the variance of consumption growth, which drives changes in the subsidy over time, vanishes. Figures are for

$$\rho_{\theta s} = 0.2.$$

## Figure 2-4: Human Capital and insurance over the life cycle

(a) Human capital expenses and lifetime income

(b) Consumption and education expenses
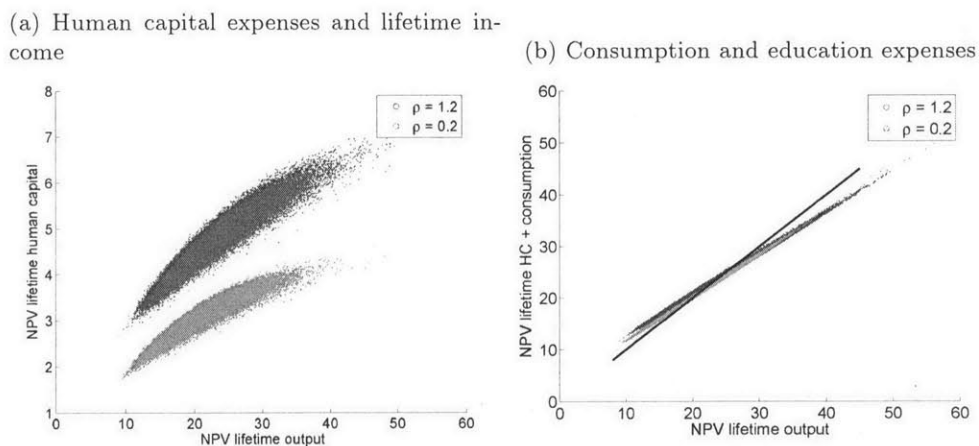


(a) Lifetime income is positively correlated with lifetime human capital expenses, the more so when $\rho_{\theta s} > 1$. There is a two-way causality: Higher ability people both acquire more human capital and have higher earnings potential. At the same time, human capital increases earnings.
(b) The figure shows the present value of consumption and human capital expenses against the present value of lifetime income. The laissez-faire outcome is represented by the 45 degree line. Clockwise pivots of the line represent more insurance.

124

## Figure 2-5: Variance and Risk

(a) $\rho_{\theta s} = 0.2$          (b) $\rho_{\theta s} = 1.2$
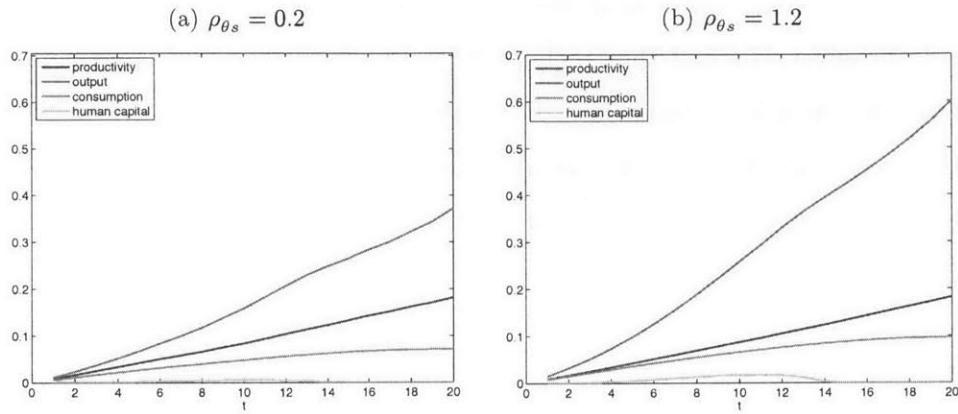


The figure shows cross-sectional variances over time. Output is more volatile than consumption. Its volatility grows at an increasing rate, driven both by ability shocks and differential investments in human capital. But pre-tax income inequality does not fully translate into consumption inequality. All outcomes are more volatile when human capital has a negative insurance value ($\rho_{\theta s} > 1$).

## Figure 2-6: Allocations

(a) $\rho_{\theta s} = 0.2$          (b) $\rho_{\theta s} = 1.2$
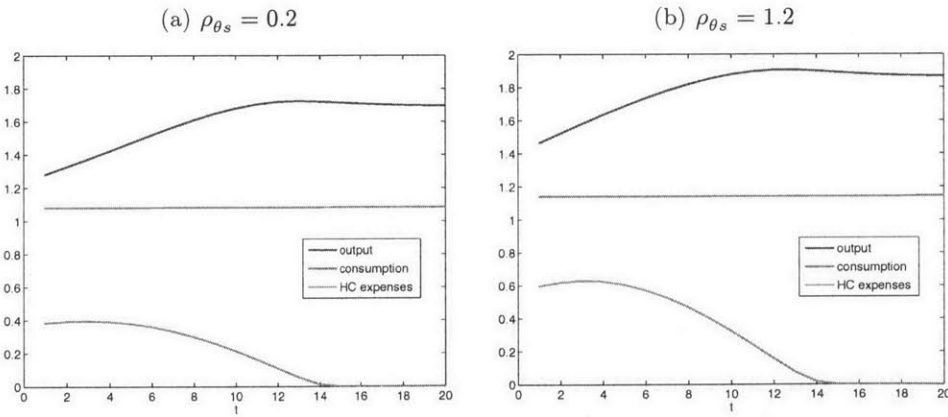


Agents optimally invest in human capital early in life. Consumption is a martingale; hence, average consumption is perfectly flat. Output and consumption are higher when human capital disproportionately benefits higher ability agents.

125

to lifetime income when human capital disproportionately benefits high ability agents. The causality goes both ways: higher ability agents both acquire more human capital and earn more. In turn, human capital increases earnings even further. When $\rho_{\theta s} > 1$, both effects are amplified. However, the fact that higher ability people also acquire more human capital does not mean that there is no insurance. Panel (b) highlights the extent of lifecycle insurance at the optimum by plotting the net present value of lifetime spending (consumption plus human capital expenses) against the net present value of lifetime income. Clockwise pivots of the line represent higher insurance.

Finally, figure 2-5 describes the cross-sectional variances of output, human capital, consumption, and ability over time. The variance of output is driven not only by stochastic ability, but also by differential investments in human capital at different ability levels. Output is much more volatile than consumption. Hence, pre-tax income inequality grows at an increasing rate, but the provision of insurance prevents this from translating fully into consumption inequality. In addition, while consumption variance grows, it does so at a decreasing rate, echoing the tax and subsidy smoothing results described above.

**Progressivity:** Figure 2-7 show the implicit progressivity of the labor wedge, by plotting $\tau_{Lt}$ against the contemporaneous productivity shock, $\theta_t$, at $t = 19$. When $\rho_{\theta s} > 1$, the labor wedge is regressive in the short run, which is true for a similar parameterization of the problem without human capital. On the other hand, when $\rho_{\theta s} < 1$, the labor wedge exhibits a short-run progressivity.

The reason for this reverse pattern is that both the labor wedge and the net subsidy are tools to insure against earnings risk. Along the optimal path, they need to evolve consistently, according to the "modified inverse elasticity rule" in (2.26). The labor wedge always has positive insurance and redistributive effects. The same is true for the net subsidy only if $\rho_{\theta s} < 1$. Accordingly, the two instruments co-move positively when $\rho_{\theta s} < 1$ and negatively when $\rho_{\theta s} \geq 1$. The net subsidy is always regressive when $\rho_{\theta s} > 0$ because higher ability people benefit more from human capital (see figure 2-8 for $t = 13$). The labor wedge will hence exhibit the inverse relation when $\rho_{\theta s} \leq 1$ (which is equivalent to a progressive labor wedge). Despite any short-run regressivity, the system is progressive overall and does provide insurance, as was shown in figure 2-4.

Figure 2-7: Progressivity and regressivity of the labor wedge

(a) $\tau_{Lt}$ against $\theta_t$ for $\rho_{\theta s} = 0.2$

(b) $\tau_{Lt}$ against $\theta_t$ for $\rho_{\theta s} = 1.2$
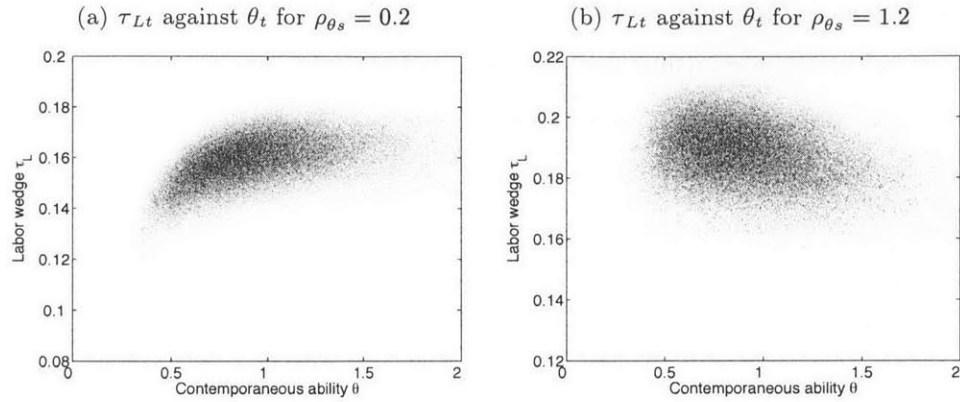


The labor wedge exhibits short-run progressivity when $\rho_{\theta s} < 1$, but short-run regressivity when $\rho_{\theta s} > 1$.

Figure 2-8: Regressivity of the net human capital wedge

(a) $t_{st}$ against $\theta_t$ for $\rho_{\theta s} = 0.2$

(b) $t_{st}$ against $\theta_t$ for $\rho_{\theta s} = 1.2$
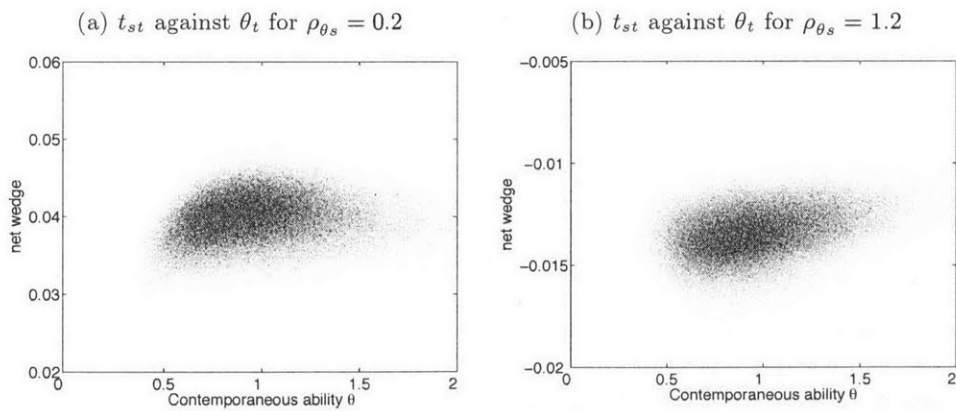


The net wedge is always regressive in the short run, but more so when $\rho_{\theta s} > 1$.

## 2.6 Implementation

The wedges, as functions of histories of types, do not immediately translate into explicit taxes and subsidies on income, human capital or savings. This section considers the implementation of the optimal allocations using decentralized instruments. For clarity, the focus is on educational expenses only, but observable training time can be added as an additional conditioning variable. The first implementation is through income contingent loans (hereafter, ICLs), which are compared and contrasted to certain types of ICLs currently used in several countries. The second is through a Deferred Deductibility scheme, which can be applied in the special case when shocks are independently and identically distributed.

### 2.6.1 Income Contingent Loans

Before presenting the ICLs, the decentralized economy is described and some notation introduced. In the decentralized economy, agents choose their human capital expenses $e_t$, income $y_t$, and savings $b_t$ in a risk-free account at a gross rate $R$. Initial wealth is zero and initial human capital is $s_0$.[49] The government can observe and keep record of the histories of consumption, output, human capital, and wealth.

Denote by $m_t^*\left(\theta^t\right)$ the optimal allocation of the social planner's problem after history $\theta^t$ for any choice variable $m \in \{c, y, b, e\}$ (the Appendix shows how to construct them from the recursive allocations derived above). For any history $\theta^t$ and subset of variables $\mathbf{m} \subset \{c, y, b, e\}$, let $Q_{\mathbf{m}}^t(\theta^{t-1})$ be the set of values for these variables at time $t$, which could arise in the planner's problem after history $\theta^{t-1}$, i.e., such that for some $\theta \in \Theta$, $\mathbf{m}_t = \mathbf{m}_t^*\left(\theta^{t-1}, \theta\right)$. For a history of observed choices $\mathbf{m}^t$, denote by $\Theta^t\left(\mathbf{m}^t\right)$ the set of all histories $\theta^t$ consistent with these choices, i.e., all $\theta^t$ such that $\mathbf{m}_s = \mathbf{m}_s^*\left(\theta^s\right)$ for all $s \leq t$. Assumption 5 guarantees that in the planner's problem, the histories $\left(y^t, e^t\right)$ can be uniquely inverted to identify the history of abilities, $\theta^t$.

**Assumption 5** $\Theta^t\left(y^t, e^t\right)$ *is either the empty set or a singleton for all histories* $\left(y^t, e^t\right)$.

In the proposed ICL scheme, loans are combined with a standard income tax based on contemporaneous income $T_Y\left(y_t\right)$, and a history-independent savings tax $T_K\left(b_t\right)$. In each period,

---

[49] Initial wealth and human capital can be heterogeneous as long as they are observable, and will enter the proposed repayment schedule as additional conditioning variables.

the agent is offered a government loan $L_t(e_t)$ as a function of his human capital expenses, and is required to make a history contingent repayment $D_t\left(L^{t-1}, y^{t-1}, e_t, y_t\right)$, as a function of the full history of past loans and earnings, as well as current income, and human capital expenses. The agent's problem is then to select the supremum over $\left\{c_t\left(\theta^t\right), y_t\left(\theta^t\right), b_t\left(\theta^t\right), e_t\left(\theta^t\right)\right\}_{\theta^t}$ in:

$$V_1(b_0, \theta_0) = \sup \sum_{t=1}^{T} \int \left[u_t\left(c_t\left(\theta^t\right)\right) - \phi_t\left(\frac{y_t\left(\theta^t\right)}{w_t\left(\theta_t, s_{t-1}\left(\theta^{t-1}\right) + e_t\left(\theta^t\right)\right)}\right)\right] P\left(\theta^t\right) d\theta^t \tag{2.37}$$

$$\text{s.t:} \quad c_t\left(\theta^t\right) + \frac{1}{R} b_t\left(\theta^t\right) + M_t\left(e_t\left(\theta^t\right)\right) - b_{t-1}\left(\theta^{t-1}\right) - L_t\left(e_t\left(\theta^t\right)\right)$$

$$\leq y_t\left(\theta^t\right) - D_t\left(L^{t-1}\left(\theta^{t-1}\right), y^{t-1}\left(\theta^{t-1}\right), e_t\left(\theta^t\right), y_t\left(\theta^t\right)\right) - T_Y\left(y_t\left(\theta^t\right)\right) - T_K\left(b_t\left(\theta^t\right)\right)$$

$$s_t\left(\theta^t\right) = s_{t-1}\left(\theta^{t-1}\right) + e_t\left(\theta^t\right), \quad s_0 \text{ given}, e_t\left(\theta^t\right) \geq 0, b_0 = 0, b_T \geq 0$$

The construction of the ICL schedule, explained formally in the Appendix, is intuitively as follows. First, the loan is set to exactly cover the cost of human capital:

$$L_t(e_t) = M_t(e_t) \quad \forall t, \forall e_t \tag{2.38}$$

The savings tax $T_K(b_t)$ is constructed to guarantee zero private wealth holdings.[50] The repayment schedule $D$ and income tax $T_Y$ are such that, along the equilibrium path, the optimal allocations from the social planner's problem are affordable for each agent after all histories, given zero asset holdings:

$$D_t\left(L^{t-1}, y^{t-1}, e_t^*\left(\theta^{t-1}, \theta\right), y_t^*\left(\theta^{t-1}, \theta\right)\right) + T_Y\left(y_t^*\left(\theta^{t-1}, \theta\right)\right) = y_t^*\left(\theta^{t-1}, \theta\right) - c_t^*\left(\theta^{t-1}, \theta\right)$$

for all $\left(L^{t-1}, y^{t-1}\right)$ such that $\theta^{t-1} \in \Theta^{t-1}\left(\left\{M_1^{-1}(L_1), ..., M_{t-1}^{-1}(L_{t-1})\right\}, y^{t-1}\right) \neq \emptyset$, and all $\theta \in \Theta$, where the history of education $e^{t-1}$ is inverted from $L^{t-1}$ using (2.38). The repayment schedule on off-equilibrium allocations – those allocations which are not optimally assigned to any type in the social planner's program – is set to be sufficiently unattractive, to ensure that agents do not select them. Intuitively then, conditional on entering a period with no savings,

---

[50]The construction in the Appendix builds on Werning (2007), who shows also that the savings tax can be redefined to implement non zero savings, at the expense of modifying the repayment schedule. The repayment scheme could also allow for private savings, and directly condition on their history $(b^{t-1})$. See the next implementation proposed, with non-zero private wealth holdings.

and with a given history of loans and output, agents only face the choice of allocations available in the planner's problem after ability histories which, up to this period, are consistent with the observed choices. By the temporal incentive compatibility of the constrained efficient allocation, they will choose the allocation designed for them.

This set of instruments defines a decentralized allocation rule, which, to an agent with past history $\left(L^{t-1}, y^{t-1}\right)$, assigns an allocation:

$$\left\{ \hat{c}_t\left(L^{t-1}, y^{t-1}, \theta_t\right), \hat{y}_t\left(L^{t-1}, y^{t-1}, \theta_t\right), \hat{b}_t\left(L^{t-1}, y^{t-1}, \theta_t\right), \hat{e}_t\left(L^{t-1}, y^{t-1}, \theta_t\right) \right\}_{t, \theta_t}$$

The equilibrium allocation as a function of ability histories $\left\{ \hat{c}\left(\theta^t\right), \hat{y}_t\left(\theta^t\right), \hat{b}\left(\theta^t\right), \hat{e}\left(\theta^t\right) \right\}$ can be deduced from the decentralization rule using the recursive relation:

$$\hat{m}_t\left(\theta^t\right) = \hat{m}_t\left(L^{t-1}\left(\theta^{t-1}\right), y^{t-1}\left(\theta^{t-1}\right), \theta_t\right) \text{ for } m \in \{c, y, b, e\}$$

where $\theta^{t-1} \in \Theta^{t-1}\left(\left\{ M_1^{-1}\left(L_1\right), ..., M_{t-1}^{-1}\left(L_{t-1}\right)\right\}, y^{t-1}\right)$ is unique by assumption 5. The decentralization rule is said to implement the optimum from the planner's problem for a given set of promised utilities $(\underline{U}\left(\theta\right))_\Theta$ if, for all $t$ and $\theta^t$, the decentralized allocations under this rule coincide with the social planner's optimal allocations, i.e., $\hat{m}_t\left(\theta^t\right) = m_t^*\left(\theta^t\right)$ for $m \in \{c, y, b, e\}$.[51]

**Proposition 24** *The optimum can be implemented through human capital loans $L_t\left(e_t\right)$, with repayments $D_t\left(L^{t-1}, y^{t-1}, e_t, y_t\right)$, contingent on the history of loans and earnings, current income, and human capital expenses, together with a history-independent savings tax $T_K\left(b_t\right)$, and an income tax on contemporaneous income $T_Y\left(y_t\right)$.*

**Income Contingent Loans:** Figure 2-9 illustrates the implementation through ICLs, by plotting the average loan received and average consolidated payment made as a fraction of contemporaneous income. The loan received naturally declines over life, as less human capital investments are needed. The repayment rises as a fraction of income until late in life, illustrating the insurance provided by the contingent repayment schedule and the increasing ability to pay over life, driven by human capital investments.

---

[51] In the planner's problem, savings are indeterminate when consumption is controlled, and without loss of generality, agents could be saving zero.

Figure 2-9: Income history contingent loans
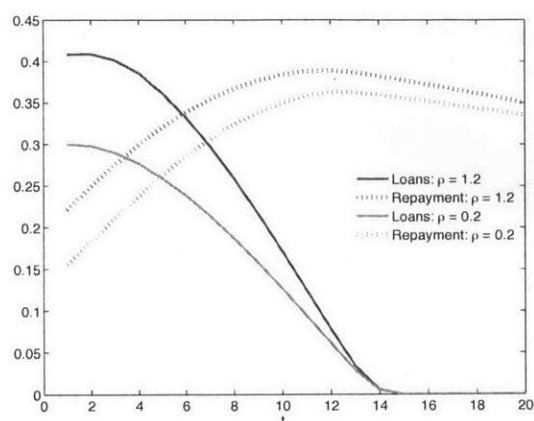


Loans and Repayments as a fraction of contemporaneous income.
Loans are high early in life, while repayments increase to provide insurance.

Figure 2-10 further highlights the insurance role of the repayment schedule, by showing a snapshot of repayment, as a fraction of income, against the contemporaneous realization of the productivity shock, at an arbitrarily chosen time ($t = 15$). Repayments increase on average in higher productivity states and decrease in lower productivity ones. The income-history contingent nature of repayments is clearly seen in their large dispersion at a given $\theta_t$: repayments depend on the full past, not only on current productivity. Indeed, a positive productivity shock is correlated with higher repayments over several future periods.

## Comparison of the Proposed Implementation to Existing ICLs

Certain types of ICLs for college education are used in several countries, including the US, New Zealand, Australia, the UK, Chile, South Africa, Sweden, and Thailand, and have been growing in popularity as a tool to reduce public spending on education, while guaranteeing equality of access, and providing partial insurance in economic hardship (see Chapman, 2005, for an empirical overview).[52] The loans sometimes depend on the level of education acquired,

---

[52] In the US, an important rationale seems to have been the fear that fixed repayment loans would discourage students from careers in the lower-paying public sector (Brody, 1994). The first schemes introduced in 1994 were Income Contingent Repayments (ICRs) for public sector jobs. In 1997, the College Cost Reduction and Access Act (CCRAA), introduced Income Based Repayment (IBR) beyond public sector jobs. Australia is one of the success stories since 1989 with its nationwide scheme ("Higher Education Contribution Scheme") that

Figure 2-10: Insurance through contingent repayments

(a) repayment against $\theta_t$ for $\rho_{\theta s} = 0.2$        (b) repayment against $\theta_t$ for $\rho_{\theta s} = 1.2$
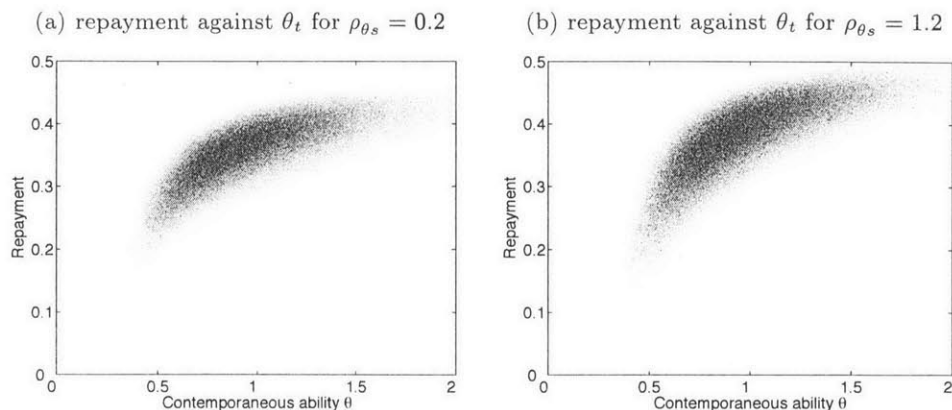


The repayment schedule provides insurance: repayments are higher when productivity realizations are higher. The history-contingent nature is seen in the dispersion of the repayment at a given contemporaneous ability realization.

the type of degree or field, and are indexed to the costs of education, which mirrors the proposed loan above. In several countries, such as Australia or New Zealand, repayments are directly collected through the tax system, as in the optimal integrated system suggested. The coercive tax power of the government is required, together with full-commitment to announced policies, to prevent agents from dropping out ex post after the realization of their incomes. This is made clear by the prominent failure of the so-called "Yale Plan," an attempt at risk-pooling within cohorts of students by Yale University in the 1970s. Because more successful earners would ex post face higher repayments, and hence cross-subsidize less successful ones, the plan suffered from a typical adverse selection problem: students with the best earnings prospects did not join or dropped out (Palacios, 2004).

There are four main differences between the ICL proposed here and existing ones. First, in the real world, the total present value of repayments is closely linked to the amount borrowed, except for subsidized interest rates or exceptional loan forgiveness. In the model, repayments are consolidated repayments for all past loans and need not, in any way, be equal to the total loan amount for a given agent; there is an implicit subsidy or tax fully integrated into the optimal system. Second, in ICL schemes observed in practice, the focus is almost exclusively

---

automatically enrolls students in an ICL, with repayments collected directly through the tax system.

on the downside, so that repayments can be deferred or forgiven in times of economic hardship. The optimal scheme is also focused on the upside, with repayments potentially increasing after a good history of earnings. Third, loans are optimally made available throughout life – not only for young adults in University – for instance, for expenses related to job training or continuing education. Fourth, the optimal repayment schedules depend not only on current income and the outstanding loan balance, but rather on the full history of earnings and past loans.[53] There is very little history-dependence in real-world ICLs, possibly with the exception of Sweden, which uses a two-year averaging of earnings to determine repayments. Interestingly, there are other real-world policies which exhibit exactly the kind of history dependence which is required, for instance, social security and some types of tax-free savings accounts. However, the numerical analysis below shows that the gain from history-dependent policies, relative to simpler history-independent (but age-dependent) policies is not very large for the calibration chosen, implying that history-independent ICLs might be close to optimal.

### 2.6.2 Implementation with iid Shocks and Wealth Dependence

A natural question is when the history dependence of the optimal policies proposed above can be reduced. In the special case of independently and identically distributed (iid) shocks, wealth and the starting stock of human capital each period can serve as sufficient statistics for the full past. This yields history-independent policies, explored in this subsection. A similar implementation for physical capital, in the absence of human capital, is studied in Albanesi and Sleet (2006).

The recursive problem with iid shocks is nested in the formulation in section 2.2, if the states $\theta_{t-1}$ and $\Delta_{t-1}$, which account for persistence, are omitted and the distribution of shocks is $f(\theta)$ each period. Allocations can be expressed as functions of the reduced state space $(v_{t-1}, s_{t-1})$ for each $\theta_t$, and the government's continuation cost is $K(v_{t-1}, s_{t-1}, t)$.

**The decentralization rule with iid shocks and taxes:** For this implementation, interpret the initial ability $\theta_1$ as uncertainty, like all other shocks $\theta_t$, rather than intrinsic heterogeneity.[54] The government selects an initial promised utility $U_1$. All agents again start with

---

[53] In particular, the sum of past loans is not a sufficient statistic for the full sequence of loans $L^{t-1}$.

[54] This incidentally makes the model directly comparable to Albanesi and Sleet (2006). Unlike Atkeson and

the same human capital $s_0$, and receive an initial wealth level $b_0$ assigned by the government.[55] The government also sets borrowing limits $\underline{b}_t$, and wealth is constrained by $b_t \in B_t \equiv [\underline{b}_t, \infty)$. The proposed decentralization rule allocates $\left(\hat{c}_t, \hat{y}_t, \hat{b}_t, \hat{e}_t\right)$ to each agent type, following some mappings from observed initial wealth $b_{t-1}$ and human capital $s_{t-1}$:

$$\hat{c}_t, \hat{y}_t, \hat{e}_t : B_{t-1} \times \mathbb{R}_+ \times \Theta \to \mathbb{R}_+ \quad \text{and} \quad \hat{b}_t : B_{t-1} \times \mathbb{R}_+ \times \Theta \to B_t$$

Starting from an initial wealth level $b_0$, the recursive decentralization rule can be mapped into a sequential allocation for all $\theta^t$.

Let $V_t(b, s)$ denote the value of an agent with beginning-of-period wealth $b$ and human capital $s$. A decentralization rule $\left(\hat{c}_t, \hat{y}_t, \hat{b}_t, \hat{e}_t\right)$, an initial assignment of wealth $b_0$, a sequence of borrowing limits $\{\underline{b}_t\}_{t=1}^T$, and initial human capital level $s_0$ form a decentralized equilibrium if, in all periods, $\left(\hat{c}_t, \hat{y}_t, \hat{b}_t, \hat{e}_t\right)$ attains the supremum in the agent's problem in (2.39) :

$$V_t(b, s) = \sup_{c, y, b', e} \int \left( u_t(c(\theta)) - \phi_t \left( \frac{y(\theta)}{w(\theta, s + e(\theta))} \right) + \beta V_{t+1}(b'(\theta), s + e(\theta)) \right) f(\theta) \, d\theta \quad (2.39)$$

$$\text{s.t:} \quad c(\theta) + M_t(e(\theta)) + \frac{1}{R} b'(\theta) = y(\theta) - T_t(b, s, y(\theta), e(\theta)) + b \quad \forall \theta$$

$$c, y, e \quad : \quad \Theta \to \mathbb{R}_+ \text{ and } b' : \Theta \to B_t = [\underline{b}_t, \infty) \text{ with } V_{T+1} \equiv 0, \ b_T \equiv 0.$$

A constrained efficient allocation from the planner's problem is implemented as a decentralized equilibrium if it arises as an equilibrium choice of agents in the above problem, and delivers expected lifetime utility $V(b_0, s_0) = U_1$.

**Proposition 25** *If $\theta$ is iid, the optimum can be implemented in an economy with borrowing constraints, an initial assignment of wealth, and an income tax schedule $T_t(b_{t-1}, s_{t-1}, y_t, e_t)$ that depends on the beginning-of-period wealth and human capital stocks, as well as on contemporaneous income and human capital investment.*

The intuition for this result is that, conditional on human capital $s_{t-1}$, there is a di-

---

Lucas (1995) and Albanesi and Sleet (2006), the analysis is in partial equilibrium analysis, with agents, as well as the government facing a constant gross rate $R$.

[55]This distribution of initial wealth need not be degenerate if there is a non-degenerate distribution of initial utility promises, from which we abstract here.

rect mapping between the social planner's cost of providing the optimal allocation to an agent with promised utility $v_{t-1}$ and the beginning-of-period wealth $b_{t-1}$ of the agent in the decentralized equilibrium. That is, for all $(v_{t-1}, s_{t-1})$, there is an associated wealth level $b_{t-1} = K(v_{t-1}, s_{t-1}, t)$. Taxes are designed such that the problem of an agent starting with wealth level $b_{t-1} = K(v_{t-1}, s_{t-1}, t)$ and human capital $s_{t-1}$ is the dual of the planner's problem with promised utility $v_{t-1}$, facing an agent with human capital $s_{t-1}$.

This recursive implementation allows a relatively simple map between the derivatives of the proposed tax function $T_t$ and the optimal wedges. Let $\tau_{Lt}(v_{t-1}, s_{t-1}, \theta_t)$ and $\tau_{St}(v_{t-1}, s_{t-1}, \theta_t)$ stand for the labor wedge and the gross human capital wedge (as defined in (2.15) and (2.17), respectively) evaluated at the optimal allocation for $(v_{t-1}, s_{t-1}, \theta_t)$. The labor wedge is immediately linked to the marginal income tax through:

$$\tau_{Lt}(v_{t-1}, s_{t-1}, \theta_t) = \frac{\partial T_t(K(v_{t-1}, s_{t-1}, t), s_{t-1}, y_t^*, e_t^*)}{\partial y_t}$$

with $y_t^* = y_t^*(v_{t-1}, s_{t-1}, \theta_t)$, and $e_t^* = e_t^*(v_{t-1}, s_{t-1}, \theta_t)$ the optimal allocations from the planner's problem. The human capital wedge is linked to the marginal subsidy more indirectly:

$$\tau_{St}(v_{t-1}, s_{t-1}, \theta_t) = \frac{-\partial T_t(K(v_{t-1}, s_{t-1}, t), s_{t-1}, y_t^*, e_t^*)}{\partial e_t} + \beta E_t \left( \frac{u_{t+1}'(c_{t+1}^*)}{u_t'(c_t^*)} \left( \frac{\partial T_{t+1}}{\partial e_{t+1}} - \frac{\partial T_{t+1}}{\partial s_t} \right) \right)$$

where $c_t^* = c_t^*(v_{t-1}, s_{t-1}, \theta_t)$, $c_{t+1}^* = c_{t+1}^*(v_t^*, s_{t-1} + e_t^*, \theta_{t+1})$, and $T_{t+1} = T_{t+1}(b_t, s_t, y_{t+1}, e_{t+1})$ is evaluated at:

$$v_t^* = v_t^*(v_{t-1}, s_{t-1}, \theta_t), \qquad b_t = K(v_t^*, s_{t-1} + e_t^*, t+1) \qquad s_t = s_{t-1} + e_t^*$$

$$y_{t+1} = y_{t+1}^*(v_t^*, s_{t-1} + e_t^*, \theta_{t+1}), \qquad e_{t+1} = e_{t+1}^*(v_t^*, s_{t-1} + e_t^*, \theta_{t+1})$$

The wedge is not in general equal to the marginal subsidy, that is, the expected reduction in tax from an incremental investment in human capital. A positive wedge does also not necessarily imply a positive marginal subsidy. This can be made clear by rewriting the wedge as:

$$\tau_{St} = -\left( \frac{\partial T_t}{\partial e_t} \right) + E_t \left( \beta \frac{u_{t+1}'}{u_t'} \right) E_t \left( \frac{\partial T_{t+1}}{\partial e_{t+1}} - \frac{\partial T_{t+1}}{\partial s_t} \right) + Cov \left( \beta \frac{u_{t+1}'}{u_t'}, \frac{\partial T_{t+1}}{\partial e_{t+1}} - \frac{\partial T_{t+1}}{\partial s_t} \right)$$

A positive wedge on human capital can be engineered either directly through expected positive marginal subsidies, or, instead, more indirectly through the risk properties of the optimal tax schedule. If the marginal tax reduction from human capital $\left( \frac{\partial T_{t+1}}{\partial e_{t+1}} - \frac{\partial T_{t+1}}{\partial s_t} \right)$ is high when marginal utility of consumption is high, human capital is a good hedge, and the covariance term is positive. It is then possible in theory that the overall wedge is positive even if expected marginal subsidies are zero.[56]

**Means-tested Human Capital Grants with iid shocks:** As an immediate corollary of Proposition 25, the tax system can instead be reformulated as a means-tested grant. In period $t$, the agent's assets and stock of human capital are verified, and he receives a grant such that:

$$G_t \left( y_t, e_t | b_{t-1}, s_{t-1} \right) = -T_t \left( b_{t-1}, s_{t-1}, y_t, e_t \right)$$

Means-testing based only on contemporaneous assets and income is hence optimal if shocks are iid. It is interesting that, although formally just a reformulation of the tax from Proposition 25, means-tested grants for higher education are very common in many countries, while wealth contingent income taxes are not. In the US for instance, Pell grants take assets as well as contemporaneous income into account.[57]

**ICLs with iid shocks:** The implementation through ICLs from subsection 2.6.1 can be modified to use wealth instead of the full history of earnings and loans to determine the repayments (hence abandoning the savings tax proposed above). In particular, the loans are again equal to the cost of human capital acquisition in each period, $L_t \left( e_t \right) = M_t \left( e_t \right)$, and the wealth and income contingent repayment schedules $D_t \left( b_{t-1}, s_{t-1}, y_t, e_t \right)$, are such that:

$$T_Y \left( y_t \right) - L_t \left( e_t \right) + D_t \left( b_{t-1}, s_{t-1}, y_t, e_t \right) = T_t \left( b_{t-1}, s_{t-1}, y_t, e_t \right)$$

---

[56] See Kocherlakota (2005) for the case of savings. In Kocherlakota (2005), the income tax depends on the full history of incomes, and wealth carries no additional information about the past. However, in general, expected marginal subsidies are not zero. Human capital, like wealth in Albanesi and Sleet (2006), carries information value about the past, which restricts the marginal subsidies at the optimum.

[57] Again, these grants are typically limited to tertiary education only. Grants for job training exist for the unemployed (hence, somewhat means-tested), for youth at risk ("YouthBuild" in the US), or for difficult to employ seniors (the "Senior Aide Program"). They are most often in the form of a direct provision of training. Some programs do provide funds for training based on need, such as the "Adult and Dislocated Worker Program" or the "Trade Adjustment Assistance."

This implementation is closer in spirit to the means-tested college loans, where the amount of funds provided is conditional on the contemporaneous resources of the child or her parents, and potentially on the highest education already attained, rather than on the history of earnings.

This subsection highlights that, while the optimal allocation derived in the direct revelation mechanism in section 2.4 is generally unique, there are many possible combinations of subsidies, loans, grants, and income taxes that can implement it.[58] If grants, repayments and subsidies were set at the appropriate levels, the difference between a grant-intensive and heavily subsidized education system such as in Continental Europe, and a loan-intensive, high-tuition system as in the US might be more apparent than real.[59] The real focus should be on the parameters entering the optimal formulas, such as the Hicksian complementarity $\rho_{\theta s}$, which determines to what extent human capital should be subsidized on net. The exact mix of instruments used is more a matter of administrative capabilities and infrastructure, which might very well be country-specific.

### 2.6.3 A Deferred, Risk-adjusted Human Capital Expense Deductibility Scheme

This implementation directly addresses the debate about whether education expenses should be tax deductible (Boskin, 1977, Blomquist, 1982, Bovenberg and Jacobs, 2005). It has been argued that a true economic depreciation of educational expenses, for which the net present value of the deduction is equal to the expense, would recover neutrality of the tax system with respect to human capital. Even starker is the argument by Bovenberg and Jacobs (2005) that a purely contemporaneous deduction of education expenses from taxable income is sufficient.

Assume first that $\rho_{\theta s} = 1$ so that the optimal net wedge is zero, and the focus is on the human capital subsidy that aims to neutralize the distortionary effects of income and savings wedges, which is the case most analogous to full deductibility. In this case, equation (2.23) must hold at the optimum. Accordingly, if the proposed tax system $T_t (b_{t-1}, s_{t-1}, y_t, e_t)$ is differentiable in all its arguments (at least over the range of equilibrium path values), then it

---

[58] On the tax incentives for higher education in the US, see for instance Hoxby (1998).

[59] Given the discrepancies in the net cost of education in Continental Europe and the US, it is clear that the financing systems are far from equivalent at their current levels. The argument proposed here is that they *could* be made equivalent while still preserving their general structure (loans-based vs. subsidy and grants-based).

must satisfy:[60]

$$-\frac{\partial T_t}{\partial e_t} + \beta E_t \left( \frac{u'_{t+1}}{u'_t} \left( \frac{\partial T_{t+1}}{\partial e_{t+1}} - \frac{\partial T_{t+1}}{\partial s_t} \right) \right)$$

$$= \frac{\partial T_t}{\partial y_t} \left( M'_t - \frac{1}{R} E_t \left( M'_{t+1} \right) \right) - \frac{1}{R} \left( 1 - \xi'_{M',t+1} \right) E_t \left( \beta R \frac{u'_{t+1}}{u'_t} \frac{\partial T_{t+1}}{\partial b_t} \right) E_t \left( M'_{t+1} \right) \quad (2.40)$$

On the other hand, the pure contemporaneous deductibility scheme proposed by Bovenberg and Jacobs (2005) would imply that at all dates $\frac{\partial T_t}{\partial y_t} = M'_t(e_t) \frac{\partial T_t}{\partial e_t}$, $\forall t$.[61] In this dynamic model, this is only true for the last period of investment, $T$, in which agents face a static problem. In all earlier periods, first, the true cost is not simply the static cost $M'_t$, but instead the dynamic expected cost $M'_t - \frac{1}{R} E_t \left( M'_{t+1} \right)$. Second, there is uncertainty and risk aversion; as already mentioned above, a positive wedge can arise if the tax burden on human capital is lower in states with high marginal utility of consumption. Third, there is also a (stochastic) distortion to physical capital accumulation, which needs to be taken into account. The (risk-adjusted) human capital subsidy and the (risk-adjusted) intertemporal wedge need to co-move inversely. This is because, if the tax on physical capital increases, there is a substitution towards human capital, and a lower subsidy is needed to stimulate it. Finally, adding back $\rho_{\theta s} \neq 1$ would push the optimum even further away from pure contemporaneous deductibility, as an additional net encouragement or discouragement of human capital would be desirable.

The right way to implement full risk-adjusted dynamic deductibility, which is the optimal policy when $\rho_{\theta s} = 1$, is a risk-adjusted deferred deductibility scheme. For the sake of the exposition, assume that $\beta = \frac{1}{R}$ and $M'_t(e_t) = 1$. Start from period $T$, in which a simple deductibility of expenses is sufficient with $-\frac{\partial T_T}{\partial e_T} = \frac{\partial T_T}{\partial y_T}$, and work backwards in order to rewrite the total change in the tax burden from an incremental investment at $t$ as:

---

[60]Obtained by applying formula (2.23) to this tax system. See Appendix formula (2.58) for a rewriting.

[61]The discrepancy in this recommendation to the one in Bovenberg and Jacobs (2005) does not come from the restrictive wage function assumed there, because the argument made in this subsection is for the case in which $\rho_{\theta s} = 1$. It arises instead from the dynamics and risk.

$$-\frac{\partial T_t}{\partial e_t} = (1 - \beta) \sum_{j=1}^{T-t} \beta^{j-1} E_t \left( \frac{u'_{t+j-1}}{u'_t} \frac{\partial T_{t+j-1}}{\partial y_{t+j-1}} \right)$$

$$+ \beta^{T-t} E_t \left( \frac{u'_T}{u'_t} \left( \frac{\partial T_T}{\partial y_T} \right) \right) - \sum_{j=1}^{T-t} \beta^j E_t \left( \frac{u'_{t+j}}{u'_t} \left( \frac{\partial T_{t+j}}{\partial b_{t+j-1}} - \frac{\partial T_{t+j}}{\partial s_{t+j-1}} \right) \right) \quad (2.41)$$

The optimal subsidy is hence equivalent to a deferred deductibility scheme, in which a fraction $(1 - \beta)$ of the human capital expense of time $t$ are deducted from taxable income in each subsequent period. Intuitively, with changing income tax rates $\frac{\partial T_t}{\partial y_t}$, a non-dynamic deductibility scheme would mean that the expense of time $t$ would be deducted at time $t'$s marginal tax rate, but the returns to this investment would accrue in the future when the agent faces potentially different marginal tax rates. If income is growing over time and marginal tax rates are increasing, as in a progressive tax system, there would be insufficient incentives to invest in human capital. A poor student would see little benefit from deducting his tuition fees from his low income, only to pay high marginal tax rates in the future. In addition, there is a "no arbitrage" term, (the last term in (2.41)), which takes into account the relative shift in the future tax schedule from more physical capital stock versus more human capital stock. Since physical and human capital are two ways to transfer resources intertemporally, there should at the optimum be no incentive to substitute from one to the other because of a tax advantage. If full deductibility is not the target, (i.e., $\rho_{\theta s} \neq 1$), implementing the optimal allocation requires adding back the optimal net subsidy each period, on top of this scheme.[62]

Tax incentives in the form of deduction schemes for higher education expenses are common, but are usually contemporaneous to the expense. In the US, the American Opportunity Credit and the Lifetime Learning Credit allow families to claim a deduction up to a certain level per student per year for college, as well as for books, supplies, and required equipment.

The deferred deductibility scheme sets the right incentives in expected, discounted utility terms. Figure 2-11 illustrates that the scheme is naturally progressive and provides insurance, by plotting the fraction of the net present value of human capital expenses that the agent cannot deduct against the net present value of lifetime income. Lower income agents hence end up

---

[62]The linear cost is for exposition only, since we want interior solutions in general. See Appendix formula (2.57) for the general case.

Figure 2-11: Progressivity of the deferred deductibility scheme
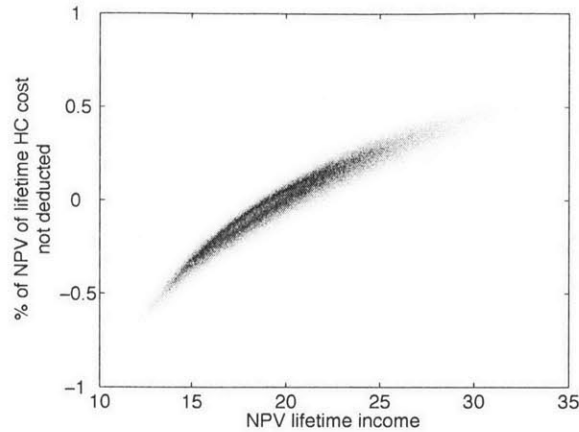


Lower income people can deduct more than they spend on human capital, while higher income people deduct less.

deducting more than they actually spent on human capital, while higher income agents end up deducting less and hence implicitly cross-subsidizing lower income agents.

## 2.6.4 Welfare Gains and Simple Age-dependent Policies

What are the welfare gains from the optimal mechanism, and how do they compare to the welfare gains from simpler, linear, but age-dependent policies? The first line of table 2 shows the welfare gains from the second best relative to the laissez-faire economy, with no taxes or subsidies, in which agents are unconstrained to borrow and save at the gross interest rate $R$. Four cases are distinguished, according to the value of the complementarity coefficient between human capital and ability $\rho_{\theta s}$ and the volatility of the productivity shock $\sigma_\psi^2$.[63] Welfare gains are expressed as the percentage increase in consumption which, if received every year after all histories, would yield the same gain in lifetime utility.

Although comparable, welfare gains are higher when human capital has negative insurance and redistributive effects ($\rho_{\theta s} > 1$). In this case, the laissez-faire economy features both more cross-sectional consumption inequality and higher consumption volatility over time, which makes the insurance from the constrained efficient mechanism more valuable. Naturally, welfare

---

[63]The high volatility (0.0161) is from Storesletten *et al.* (2004).

gains are higher when productivity is more volatile.

TABLE 2: WELFARE GAINS

| | $\rho_{\theta s} = 0.2$ | | $\rho_{\theta s} = 1.2$ | |
|---|---|---|---|---|
| Volatility | Medium | High | Medium | High |
| Welfare gain from second best | 0.85% | 1.60% | 0.98% | 1.76% |
| Welfare gain from linear age-dependent policies | 0.79% | 1.53% | 0.94% | 1.74% |
| as % of second best | 93% | 95.6% | 95.5% | 98.5% |

Medium volatility is 0.0095, high volatility is 0.0161. Line 1 expresses the gain from the second best, relative to the laissez-faire economy, in terms of the equivalent increase in consumption after all histories. Welfare gains are higher when human capital has negative redistributive and insurance values ($\rho_{\theta s} > 1$). Line 2 shows the gain from linear age-dependent policies relative to the laissez-faire, while line 3 expresses this gain as a fraction of the gain from the second best.

Age-dependent linear policies achieve a very large fraction of the welfare gain from the second best.

Given the clear age trends in the above figures and in the optimal formulas, it is natural to compare the full optimum to simple age-dependent policies. The policy under consideration sets the linear human capital subsidy, the linear income tax rate, and the linear capital tax rate at each age equal to their cross-sectional averages at that age. It is numerically challenging to precisely optimize over age-dependent tax rates, given the number of periods and the presence of three instruments; hence, this procedure delivers a lower bound for the welfare gains. It turns out, however, that even this lower bound is very tight. Indeed, the third line in table 2 shows that welfare gains as a fraction of the second best gains range from 93% for a low-volatility and low $\rho_{\theta s}$ case to a surprising 98.5% for a high volatility and high $\rho_{\theta s}$ scenario. This suggests that – for these particular calibrations – the history-dependent policies can be informative about simpler, history-independent policies, and that the bulk of the gain comes from the age-trend of optimal policies.[64]

---

[64] A word of caution is needed. Given the order of magnitude of 10e-5, it is actually very challenging to estimate these welfare comparisons between the second best and age-dependent policies precisely, especially over longer horizons. The numbers should only be taken as evidence for small welfare gains, not precise welfare calculations. The author is currently studying which other factors are most important for the welfare gains. For instance, stepping away from log utility and from a log-normal process for $\theta$ can increase welfare gains.

These findings are reminiscent of Mirrlees' (1971) conclusion that static optimal income tax schedules appear close to linear. They also echo closely recent findings from the dynamic taxation literature (Farhi and Werning, 2013), suggesting that the addition of human capital per se does not change this result.

## 2.7 Extension: Unobservable Human Capital

Until now, human capital was observable to the government, which might be a strong assumption under some circumstances. This section considers unobservable human capital investments, and focuses on how optimal policies are adjusted relative to the observable case. It starts with a simplified version of the model without training, but with unobservable human capital expenses. An augmented program is set up, in which the agent's human capital choice needs to be incentive compatible, in addition to his labor effort and type revelation. This combines dynamic moral hazard models with hidden savings and models with hidden persistent types,[65] both of which unobservably modify the agent's future response to incentives. Then, the complete model is analyzed, assuming that training time is not observable, but human capital expenses are. In this case, consumption is again controlled by the planner, and hidden training is akin to an unobservable effort with persistent effects over time.

### 2.7.1 Unobservable Monetary Investments in Human Capital

**Planning Problem**

When monetary investments in human capital are unobservable, the planner can no longer directly control consumption $c_t$. However, savings or physical capital investments remain observable, and hence the planner controls total financial resources transferred to an agent each period. The agent can both misreport his type and spend a different amount on human capital than the planner would choose. As in subsection 2.3.1, a reporting strategy $r = \{r_t(\theta^t)\}$ specifies a report after each history. In addition, a human capital strategy $\sigma = \{s_t(\theta^t)\}$

---

[65] On hidden savings, see for instance Cole and Kocherlakota (2001), Werning (2002), Kocherlakota (2004), Abraham and Pavoni (2008). On hidden persistent types, see Farhi and Werning (2013), Kapicka (2013) among others.

specifies human capital choices. The planner allocates output $y\left(r^t\right)$ and resources, denoted by $c^a\left(r^t\right)$, as functions of the history of reports by the agent. The agent then unobservably chooses his human capital investment $s\left(\theta^t\right)$. Hours of work are determined residually by $l\left(r^t\right) = y\left(r^t\right)/w\left(\theta_t, s\left(\theta^t\right)\right)$. The continuation value after history $\theta^t$ from reporting and human capital strategies $r$ and $\sigma$ can be written as:

$$\omega^{r,\sigma}\left(\theta^t\right) = u_t\left(c\left(r^t\left(\theta^t\right), s_t\left(\theta^t\right), s_{t-1}\left(\theta^{t-1}\right)\right)\right) - \phi_t\left(\frac{y\left(r^t\left(\theta^t\right)\right)}{w_t\left(\theta_t, s_t\left(\theta^t\right)\right)}\right) + \beta \int \omega^{r,\sigma}\left(\theta^{t+1}\right) f^{t+1}\left(\theta_{t+1}|\theta_t\right) d\theta_{t+1}$$

where actual consumption is equal to

$$c\left(r^t\left(\theta^t\right), s_t\left(\theta^t\right), s_{t-1}\left(\theta^{t-1}\right)\right) = c^a\left(r^t\left(\theta^t\right)\right) - M_t\left(s_t\left(\theta^t\right) - s_{t-1}\left(\theta^{t-1}\right)\right)$$

Similarly, let $\omega\left(\theta^t\right)$ be the continuation value after history $\theta^t$ from reporting truthfully and following the planner's recommended human capital strategy. Incentive compatibility requires that, after any history $\theta^t$, for all reporting and human capital strategies $r$ and $\sigma$ :

$$(IC): \omega\left(\theta^t\right) \geq \omega^{r,\sigma}\left(\theta^t\right) \quad \forall r, \sigma, \theta^t \tag{2.42}$$

The first-order approach replaces incentive constraint (2.42) by two local necessary conditions along the equilibrium path. Again, the agent's envelope condition must hold:

$$\frac{\partial \omega\left(\theta^t\right)}{\partial \theta_t} = \frac{w_{\theta,t}\left(\theta_t, s_t\left(\theta^t\right)\right)}{w_t\left(\theta_t, s_t\left(\theta^t\right)\right)^2} y\left(\theta^t\right) \phi_t'\left(\frac{y\left(\theta^t\right)}{w_t\left(\theta_t, s_t\left(\theta^t\right)\right)}\right) + \beta \int \omega\left(\theta^{t+1}\right) \frac{\partial f^{t+1}\left(\theta_{t+1}|\theta_t\right)}{\partial \theta_t} d\theta_{t+1} \tag{2.43}$$

(since there is no training in this subsection, $\phi_t'\left(l_t\right)$ denotes the derivative with respect to labor). Second, since human capital is now unobservable, the agent's first order condition with respect to $s_t$, i.e., the "Euler Equation for human capital," must also hold, which is equivalent to a zero gross wedge $\tau_{St}$ in every period:

$$M_t' \left( s_t \left( \theta^t \right) - s_{t-1} \left( \theta^{t-1} \right) \right) u_t' \left( c \left( \theta^t \right) \right) \;\; = \;\; \frac{w_{s,t} \left( \theta_t, s_t \left( \theta^t \right) \right)}{w_t \left( \theta_t, s_t \left( \theta^t \right) \right)^2} y \left( \theta^t \right) \phi_t' \left( \frac{y \left( \theta^t \right)}{w_t \left( \theta_t, s_t \left( \theta^t \right) \right)} \right) \qquad (2.44)$$

$$+ \beta \int M_{t+1}' \left( s_{t+1} \left( \theta^{t+1} \right) - s_t \left( \theta^t \right) \right) u_{t+1}' \left( c \left( \theta^{t+1} \right) \right) f^{t+1} \left( \theta_{t+1} | \theta_t \right)$$

The focus is again on interior solutions (see the discussion on this in subsection 2.3.2).

Unobservable human capital raises the challenge of guaranteeing that the first-order approach is valid. In standard hidden savings problems, a sufficiently strong complementarity between shirking and savings can invalidate it (Kocherlakota, 2004). But there are still many cases in which it remains valid (see Abraham and Pavoni, 2008), and the gain in tractability from it is even larger relative to direct approaches, as the number of potential deviations grows. The goal of this section is to highlight analytically how optimal policies need to be adjusted relative to the observable human capital case, in circumstances under which the first-order approach can be applied. Studying general conditions for the validity of this approach with hidden human capital is left for future work.

The notation is as in subsection 2.3.3. In addition, a new endogenous state variable is introduced. Let $\Delta^s \left( \theta^t \right)$ be the negative of the expected marginal cost of education in utility units:

$$\Delta^s \left( \theta^t \right) = - \int M_{t+1}' \left( s_{t+1} \left( \theta^{t+1} \right) - s_t \left( \theta^t \right) \right) u_{t+1}' \left( c \left( \theta^{t+1} \right) \right) f^{t+1} \left( \theta_{t+1} | \theta_t \right) \qquad (2.45)$$

$\Delta^s \left( \theta^t \right)$ captures the expected future cost reduction from an investment at time $t$, or, equivalently, at interior solutions, the future stream of benefits from investment. The Euler Equation from (2.44) can be rewritten as:

$$M_t' \left( s_t \left( \theta^t \right) - s_{t-1} \left( \theta^{t-1} \right) \right) u_t' \left( c \left( \theta^t \right) \right) = \frac{w_{s,t}}{w_t} l \left( \theta^t \right) \phi_t' \left( l \left( \theta^t \right) \right) - \beta \Delta^s \left( \theta^t \right) \qquad (2.46)$$

An additional "promise-keeping" constraint with respect to the future benefit of human capital needs to be added to the program

$$\Delta^s \left( \theta^{t-1} \right) = - \int M_t' \left( s_t \left( \theta^t \right) - s_{t-1} \left( \theta^{t-1} \right) \right) u_t' \left( c \left( \theta^t \right) \right) f^t \left( \theta_t | \theta_{t-1} \right) d\theta_t$$

The full recursive formulation of the program is in the Appendix.

## The Optimal Labor Wedge with Unobservable Human Capital

With unobservable human capital, both the income tax and the savings tax will be adjusted to indirectly provide incentives for HC accumulation. A lower labor tax and a higher capital tax can stimulate human capital investments and mimic a positive net subsidy $t_s$, which is the optimal policy when $\rho_{\theta s} < 1$.

The optimal formula for the labor wedge is very complicated, because distorting labor downwards has several indirect feedback effects on unobservable human capital. First, with lower hours of work, there is less incentive for the agent to acquire (hidden) human capital. The wage is lower, which tends to reduce labor further. While this effect is also present with observable human capital, the planner can no longer directly counter it by controlling human capital incentives: the realized net incentive $t_{st}$ will be different than the optimal $t_{st}^*$ with observable human capital in (2.20). The effective labor distortion is hence larger than $\tau_{Lt}$ as defined in (2.15) if $t_{st} < 0$. Second, the income effect from a change in the labor distortion modifies the agent's incentive to invest through his Euler equation for human capital (in (2.46)). Finally, changing labor in one period has repercussions on future human capital choices, through the change in the contemporaneous human capital stock, because of the nonlinear cost of human capital $M_t$. All these additional feedback effects of the labor distortion are summed up for notational purposes in an adjusted labor wedge, $\tilde{\tau}_{Lt}$:[66]

**Definition 5** *Define the adjusted labor wedge $\tilde{\tau}_{Lt}$:*

$$\frac{\tilde{\tau}_{Lt}}{1 - \tilde{\tau}_{Lt}} \equiv \frac{\tau_{Lt}}{1 - \tau_{Lt}} - \frac{1}{(1 - \rho_{ss,t})} \frac{1 + \varepsilon_t^u}{\varepsilon_t^c} t_{st} + \left( \tilde{\gamma}_t^E \left( \theta^t \right) M_t' u_t'' - \frac{\left( \frac{1}{R} E_t \left( \tilde{\gamma}_{t+1}^E M_{t+1}'' u_{t+1}' \right) - \tilde{\gamma}_t^E M_t'' u_t' \right)}{w_{s,t} l_t \left( \theta^t \right) \left( 1 - \tau_{Lt} \right) \left( 1 - \rho_{ss,t} \right)} \frac{1 + \varepsilon_t^u}{\varepsilon_t^c} \right)$$
$$(2.47)$$

*where $t_{st}$ is as defined in (2.19) evaluated at $\tau_{St} \equiv 0$; $\rho_{ss,t} \equiv \left( \frac{\partial^2 w_t}{\partial s_t} w_t \right) / \left( \frac{\partial w_t}{\partial s_t} \right)^2 < 0$; $\tilde{\gamma}_t^E \left( \theta^t \right) \equiv \tilde{\gamma}_t^E \left( \theta^t \right) / f^t \left( \theta_t | \theta_{t-1} \right)$ is the normalized multiplier on the Euler for human capital (2.46).*

---

[66] The term in brackets in the definition in (2.47) measures how the cost side of the Euler for human capital is relaxed or tightened. If the planner wants to stimulate human capital beyond what agents would themselves chose, then it can be shown that $\tilde{\gamma}_t^E \left( \theta \right) < 0$. The adjusted labor wedge is just a re-definition of a distortion that conveniently captures several terms. But the formulas can of course be re-stated in terms of the $\tau_{Lt}$ wedge, with the additional terms from $\tilde{\tau}_{Lt}$ appearing explicitly on the right hand side.

The following proposition shows how the adjusted labor wedge is set at the optimum.

**Proposition 26** *i) The optimal adjusted labor wedge* $\tilde{\tau}_{Lt}$ *satisfies:*

$$\frac{\tilde{\tau}_{Lt}\left(\theta^t\right)}{1-\tilde{\tau}_{Lt}\left(\theta^t\right)} = \frac{\mu\left(\theta^t\right)}{f^t\left(\theta_t|\theta_{t-1}\right)}\frac{\varepsilon_{w\theta}}{\theta_t}u_t'\left(c\left(\theta^t\right)\right)\frac{1+\varepsilon_t^u}{\varepsilon_t^c}\left(1-\frac{\left(1-\rho_{\theta s,t}\right)}{\left(1-\rho_{ss,t}\right)}\right) \quad (2.48)$$

*where the multiplier* $\mu\left(\theta_t\right)$ *can be written recursively as:*

$$\mu\left(\theta^t\right) = \kappa^u\left(\theta^t\right) + \eta^u\left(\theta^t\right)$$

$$\kappa^u\left(\theta^t\right) = \int_{\theta_t}^{\bar{\theta}}\left(1-g_s\right)\frac{\left(1-\tilde{\gamma}^E\left(\theta_s\right)M_t'\left(e\left(\theta^{t-1},\theta_s\right)\right)u_t''\left(c\left(\theta^{t-1},\theta_s\right)\right)\right)}{u_t'\left(c\left(\theta^{t-1},\theta_s\right)\right)}f\left(\theta_s|\theta_{t-1}\right)d\theta$$

$$\text{with } g_s = \frac{u_t'\left(c\left(\theta^{t-1},\theta_s\right)\right)}{\left(1-\tilde{\gamma}^E\left(\theta_s\right)M_t'\left(e\left(\theta^{t-1},\theta_s\right)\right)u_t''\left(c\left(\theta^{t-1},\theta_s\right)\right)\right)}\lambda_- \ ,$$

$$\lambda_- = \int_{\underline{\theta}}^{\bar{\theta}}\frac{\left(1-\tilde{\gamma}^E\left(\theta_m\right)M_t'\left(e\left(\theta^{t-1},\theta_s\right)\right)u_t''\left(c\left(\theta^{t-1},\theta_m\right)\right)\right)}{u_t'\left(c\left(\theta^{t-1},\theta_m\right)\right)}f\left(\theta_m|\theta_{t-1}\right)d\theta_m$$

$$\eta^u\left(\theta^t\right) = \frac{\mu\left(\theta^{t-1}\right)}{f\left(\theta_{t-1}|\theta_{t-2}\right)}\left[R\beta\left(\int_{\theta_t}^{\bar{\theta}}\frac{\partial f\left(\theta_s|\theta_{t-1}\right)}{\partial\theta_{t-1}}d\theta_s\right)\right]$$

*ii) Distortions at the bottom and the top are not zero:* $\tau_{Lt}\left(\theta^{t-1},\bar{\theta}\right)\neq 0$, $\tau_{Lt}\left(\theta^{t-1},\underline{\theta}\right)\neq 0$.

Formula (2.48) highlights the two roles that the income tax needs to fulfill with unobservable human capital, namely to insure and redistribute income through the standard channel, but also to partially substitute for the missing human capital subsidy. Because of the latter role, the factors which were previously entering the optimal net human capital wedge now affect the labor wedge formula. Among others, a new scaling factor appears, $\left(1-\frac{\left(1-\rho_{\theta s,t}\right)}{\left(1-\rho_{ss,t}\right)}\right)$, which can be greater than 1 when human capital has negative redistributive and insurance effects $\left(\rho_{\theta s}\geq 1\right)$.[67] The scaling factor is dampened by the labor elasticity, since by manipulating the labor wedge to indirectly provide human capital incentives, the labor choice is also distorted, which is costlier if labor is more elastic. It is not always the case that the optimal labor wedge

---

[67] Because all terms in this formula are endogenous, it is not evident whether, after any given history, $\tau_{Lt}$ is higher or lower than in an equivalent world with observable human capital (i.e., a world with the same primitives, but not the same allocations).

should be lower when human capital is unobservable. Intuitively, this is more likely to occur when $\rho_{\theta s}$ is small, in which case the planner would favor a larger net subsidy on human capital if it were observable, but agents do not factor in the full effect of their investment on social welfare.

The (unadjusted) labor wedge $\tau_{Lt}$ will no longer be zero at the bottom and at the top with unobservable human capital. Instead, it is the *adjusted* wedge $\tilde{\tau}_{Lt}$ which is zero.

The adjusted wedge can be rewritten recursively to study its life cycle evolution, again under the assumption that $\log(\theta_t)$ follows the auto-regressive process given by (2.28), as in subsection 2.4.4.[68]

$$E_{t-1}\left[\frac{\tilde{\tau}_L}{(1-\tilde{\tau}_{Lt})}\frac{1}{R\beta}\frac{\varepsilon_{w\theta,t-1}}{\varepsilon_{w\theta,t}}\left(\frac{\varepsilon_t^c}{1+\varepsilon_t^u}\right)\left(\frac{1+\varepsilon_{t-1}^u}{\varepsilon_{t-1}^c}\right)\frac{(\rho_{\theta s,t-1}-\rho_{ss,t-1})}{(\rho_{\theta s,t}-\rho_{ss,t})}\frac{(1-\rho_{ss,t})}{(1-\rho_{ss,t-1})}\frac{u'_{t-1}(c_{t-1})}{u'_t(c_t)}\right]$$

$$=\frac{1}{R\beta}\varepsilon_{w\theta,t-1}\frac{1+\varepsilon_{t-1}^u}{\varepsilon_{t-1}^c}\frac{(\rho_{\theta s,t-1}-\rho_{ss,t-1})}{(1-\rho_{ss,t-1})}Cov\left(\frac{u'_{t-1}(c_{t-1})}{u'_t(c_t)}\left(1-\tilde{\gamma}_t^E M'_t u''_t\right),\log(\theta_t)\right)+p\left(\frac{\tilde{\tau}_{Lt-1}}{1-\tilde{\tau}_{Lt-1}}\right)$$

In addition to the familiar factors, the labor wedge evolution is now also loaded with the factors which would determine the evolution of the net human capital wedge $t_{st}$, were human capital observable (formula (2.30)). Notably, if the insurance and redistributive effects of hidden human capital grow over time ($\rho_{\theta s,t}$ decreases), the income tax will rise more slowly (or fall more quickly) to provide the needed stimulus for human capital. Unobservable human capital, with potentially varying redistributive implications throughout life, is another element to take into account when thinking about age-dependent taxation.

### The Optimal Capital Wedge with Unobservable Human Capital Expenses

Unobservable human capital expenses are a hidden channel of intertemporal resource transfer, invalidating the standard Inverse Euler equation (in (2.24)). This is consistent with previous work with hidden consumption (e.g., Townsend, 1982, Cole and Kocherlakota, 2001), or hidden preferences (Atkeson and Lucas, 1992). The setup here is a hybrid of the observable and unobservable consumption cases, because the planner does control total financial resources

---

[68]The proof is the same as for Corollary 3, except that $\tau_{Lt}$ is replaced by $\tilde{\tau}_{Lt}$, and the point-wise formula is (2.48).

allocated per period, but does not control how these resources get split by the agent between consumption and human capital expenses. Therefore, the agent's standard Euler for savings or physical capital does not need to hold, but the agent's Euler equation for human capital (in (2.44)) does. This imposes a restriction on the marginal utilities across periods. Attempting to manipulate assigned consumption in different periods carries an additional cost for satisfying the agent's Euler equation for human capital, because of an income effect. In this case, a modified Inverse Euler equation applies at the optimum:[69]

$$\frac{\beta R(1 - \tilde{\gamma}_t^E \left(\theta^t\right) M_t' \left(e_t\right) u_t'' \left(c_t\right))}{u_t' \left(c_t\right)} = \int_{\underline{\theta}}^{\bar{\theta}} \frac{\left(1 - \tilde{\gamma}_{t+1}^E \left(\theta^{t+1}\right) M_{t+1}' \left(e_{t+1}\right) u_{t+1}'' \left(c_{t+1}\right)\right)}{u_{t+1}' \left(c_{t+1}\right)} f^{t+1} \left(\theta_{t+1} | \theta_t\right)$$

(2.49)

### 2.7.2 Unobservable Training Time

This subsection reintroduces training time. Suppose that training time is unobservable to the government, but human capital expenses are observable. This is most consistent with the standard taxation framework, as training is another type of effort, such as labor, potentially difficult to keep track of. Expenses on tuition fees or formal education costs on the other hand are tangible, recordable transactions – similar to income. Under this scenario, the net subsidy on human capital expenses is another tool available to the government, in addition to the labor wedge and the savings wedge, to indirectly encourage or discourage training. The following assumption simplifies the analysis in this subsection.[70]

**Assumption 6** $\phi_{lz,t} = 0 \ \forall t$.

Despite this assumption, training still affects labor supply through the wage. Recall that $\tau_{Lt}^*$ and $t_{st}^*$ are the optimal labor and human capital wedges when training is observable, as given in (2.25) and (2.20) respectively.[71] The relation between the optimal labor tax and the

---

[69]The capital wedge must satisfy a "no-arbitrage" condition:

$$\left(1 - \tau_{Kt}\right) \beta R E_t \left(u_{t+1}'\right) = \frac{w_{s,t}}{w_t} l_t \frac{\phi' \left(l_t\right)}{M_t'} + \beta E_t \left(u_{t+1}' \frac{M_{t+1}'}{M_t'}\right)$$

[70]The setup of the program and the general derivations are in the Appendix.

[71]These are endogenous functions of the allocations, hence their *values* here might not coincide with the optimal ones from the previous section.

net wedge on observable expenses is given by the following proposition.

**Proposition 27** *At the optimum, the deviations of the labor wedge and net human capital wedge from their optimal levels when training is observable must satisfy the following relation:*

$$\left(\frac{\tau_{Lt}}{1-\tau_{Lt}} - \frac{\tau_{Lt}^*}{1-\tau_{Lt}^*}\right) = \frac{1}{\left(1-\rho_{zs,t}\right)}\frac{1+\varepsilon_t^u}{\varepsilon_t^c}\left(t_{st} - t_{st}^*\right) \tag{2.50}$$

*where $\rho_{zs} \equiv \frac{w_{sz}w}{w_s w_z}$ is the Hicksian coefficient of complementarity between human capital and training in the wage function.*

Hence, the labor wedge and the net subsidy will both be adjusted away from their optimal levels which would apply if training were observable. They will be modified in proportion to their effectiveness in affecting training time, that is, respectively, $1/\frac{\varepsilon_t^c}{1+\varepsilon_t^u}$ and $\left(1-\rho_{zs,t}\right)$. This leads to a variation on the inverse elasticity rule, in which the sharpest instrument needs to be used to indirectly provide incentives for an unobservable choice.

Should human capital expenses be distorted additionally to take into account hidden training? According to the above relation, in general, they should be, unless the elasticity of the wage to training does not depend on expenses $\left(\rho_{zs,t} = 1\right)$. Note once more that this is different from training and expenses being separable in the wage $(\rho_{zs,t} = 0)$. The labor and human capital wedges will co-move if $\rho_{zs,t} \leq 1$, and move inversely otherwise. With a CES wage as in (2.3) with $\rho_{zs} = \rho_{\theta s} = \rho$ and isoelastic disutility as in (2.27):

$$\left(\frac{\tau_{Lt}}{1-\tau_{Lt}} - \frac{\tau_{Lt}^*}{1-\tau_{Lt}^*}\right) = \frac{\gamma}{(1-\rho)}\left(t_{st} - t_{st}^*\right)$$

With unobservable training, the standard Inverse Euler equation from (2.24) holds again, since consumption is controlled by the planner.

## 2.8 Conclusion

This chapter studies optimal dynamic taxation and human capital policies over the life cycle in a dynamic Mirrlees model with heterogeneous, stochastic, and persistent ability. Agents invest in human capital throughout life, by either spending money or time. The government aims

149

to provide redistribution and insurance against adverse draws from the ability distribution. However, the government faces asymmetric information about agents' ability – both its initial level and its stochastic evolution over life – and about labor supply. The constrained efficient allocations were obtained using a dynamic first order approach, and are characterized by wedges or implicit taxes and subsidies. Formulas for the optimal labor and human capital wedges, as well as for their evolution over time are derived.

A crucial consideration for the design of optimal policies is whether human capital has overall positive redistributive and insurance effects. If human capital subsidies stimulate labor supply, and hence generate additional resources more than they amplify existing pre-tax inequality, they reduce after-tax income inequality on balance. This occurs when the elasticity of the wage with respect to ability is decreasing in human capital. In this case, the optimal net subsidy on human capital expenses is positive and increasing over time. When considering the optimal subsidies on training time, the additional interactions of training with both contemporaneous and future labor supply need to be taken into account. The optimal allocations can be implemented with income contingent loans, the repayment schedules of which depend on the full history of human capital investments and earnings. If shocks to ability are independently and identically distributed, a Deferred Deductibility scheme, in which part of current human capital expenses can be deducted from future years' incomes, can also implement the optimum.

The simulations reveal that the optimal net human capital wedges are small, which implies that neutrality of the tax system relative to human capital expenses is close to optimal for the proposed calibrations. In addition, simple age-dependent linear taxes and subsidies can achieve almost the entire welfare gain from the full second-best relative to the laissez-faire outcome. Further numerical work could shed light on whether this result remains true with different preferences, in particular with higher risk aversion.

There are three alternative questions for which this analysis can provide some answers. First, should the tax system preserve neutrality with respect to the choice between bequests and human capital spending, two important ways in which parents can transfer resources to their children? The life cycle can be reinterpreted as a dynastic household, in which parents finance their children's human capital, with persistence in stochastic ability, and partial or full depreciation of human capital across generations. A reinterpretation of the optimal formulas

150

derived here shows that the optimal subsidy on parents' investments in children's human capital increases with their children's expected labor taxes, and decreases with the bequest tax. If the elasticity of the children's wage with respect to ability is decreasing in human capital, the positive redistributive and insurance effects of human capital transfers on the next generation push the optimal subsidy higher.[72] Second, should productivity-enhancing investments by entrepreneurs or the self-employed be made tax-deductible? Workers in the model can instead be viewed as entrepreneurs or self-employed, who can invest in their businesses' productivity through expenses for research, knowledge acquisition, or training of the workforce, generating risky and persistent profits.[73] If innovation and productivity expenses disproportionately increase risk, they should be made less than fully deductible.[74] Finally, the analysis could inform the study of optimal policies towards people's investments in health – another type of human capital – which also involves both time and monetary costs, as well as heterogeneity and uncertainty over life.

This theoretical research points to two important empirical explorations that could shed light on the mechanisms behind, and the magnitudes of, the optimal policies. First, how does the complementarity between ability and human capital change over life? In contrast to schooling or higher education, there is little evidence on this for human capital investments later in life, such as job training. This creates a fruitful link between optimal taxation and the long-standing empirical labor literature on this issue. Secondly, it has not been documented entirely yet how strongly people react to current and future expected taxes when making their human capital investment decisions. While challenging, estimating the long-term effects of taxation on human capital accumulation appears very important.

---

[72] Higher than the subsidy needed to simply guarantee tax neutrality with respect to the choice between bequests and human capital transfers.

[73] These productivity investments, embodied in people, are distinct from investment in physical machines or financial assets.

[74] In the sense that the elasticity of business earnings to risk is increasing in innovative activity. In practice, many expenses for the self-employed can be deducted, but there is no special category for innovation or productivity-enhancing expenses.

# 2.A Appendix with Proofs

## 2.A.1 Observable Human Capital

**Notation:**

$w_{\theta,t}, w_{s,t}, w_{z,t}$ are the partials of the wage with respect to ability, human capital $s$ and $z$ respectively. $w_{\theta s,t}, w_{\theta z,t}, w_{ss,t}, w_{zs,t}, w_{zz,t}$ are the second order derivatives. $\phi_{l,t}$ and $\phi_{z,t}$ are the partials of the disutility with respect to labor and training respectively, and $\phi_{zz,t}, \phi_{lz,t}, \phi_{ll,t}$ the corresponding second order derivatives. $\varepsilon_{xyt}$ is the elasticity of variable $x$ with respect to variable $y$, $\varepsilon_{xyt} = d\log(x_t)/d\log(y_t)$, hence for instance, $\varepsilon_{w\theta,t}$ is the elasticity of the wage with respect to ability $\theta$. When clear, the history dependence of allocations is omitted, e.g., $c_t$ denotes $c\left(\theta^t\right)$.

**Link to the Ben-Porath model in discrete time.**

In the Ben-Porath model, the general accumulation process for human capital is $z_{t+1} = H(z_t, \alpha_t) + (1-\delta)z_t$ with $\alpha_t$ the fraction of human capital reinvested into human capital acquisition and $\delta$ the depreciation rate of human capital. Labor supply is equal to $l_t = 1 - \alpha_t$, the residual time from human capital acquisition. The rental rate of human capital $w_t$ is constant and earnings are $y_t = w_t z_t l_t$. The most general functional form used is $z_{t+1} = \beta_0 \alpha_t^{\beta_1} (\alpha_t z_t)^{\beta_2} + (1-\delta)z_t$, so that $w_t \alpha_t z_t$ is the opportunity cost of human capital acquisition.

In this chapter, a variation of this model is used, which is better adapted to a Mirrleesian analysis. The accumulation process is simplified along some dimensions and made more complex along others. First, instead of $\alpha_t$, the input into human capital is effort or time $i_t$. $i_t$ plays the same role of $\alpha_t$ but has a disutility cost, potentially nonseparable from the disutility cost of labor $\phi_t(l, i)$, and not just an opportunity cost of time. This is because, first, disutility costs from learning are empirically relevant (see Heckman et al. 2005) and second, with observable human capital, labor would also be observable if we only had: $l_t = 1 - i_t$. Second, the parameters are fixed so that only time, but not the previous stock of human capital matters for the accumulation, and there is no depreciation: $\beta_1 = 1 = \beta_0$ and $\beta_2 = \delta = 0$. This leads to a linear accumulation process $z_{t+1} = z_t + i_t$. The diminishing returns to human capital are instead captured in the more complex rental rate for human capital, i.e., the wage, which is

heterogeneous, nonlinear, and stochastic $w_t = w_t(\theta_t, z_t, s_t)$ (and also depends on human capital expenses).

**Generalized model a la Ben-Porath:** It is possible to generalize the model to feature an accumulation process close to the one in Ben-Porath. The key change is to make the accumulation process depend on the human capital stock. Hence, let $z_t = Z_t(z_{t-1}, i_t) = z_{t-1} + H_t(z_{t-1}, i_t)$ be a more general production function without depreciation (which is without loss of generality). Define the inverse function with respect to $i$, $I_t(z_t, z_{t-1}) \equiv Z_t^{-1,i}(z_{t-1}, z_t)$. The accumulation process can be time-dependent. Note now that with $\frac{\partial \phi_t}{\partial z_t} = \frac{\partial \phi_t}{\partial i_t} \frac{\partial i_t}{\partial z_t} = \frac{\partial \phi_t}{\partial i_t} \frac{\partial I_t}{\partial z_t}$, the definition for the net bonus $t_{zt}$ is unchanged. The same formula for the optimal bonus on training time as in (2.32) applies.

**Derivations and proofs for Propositions** (20) and (22) :

The expenditure function: $\tilde{c}(l, \omega - \beta v, i, \theta)$ defines consumption indirectly as a function of labor $l$, current period utility ($\tilde{u} = \omega - \beta v$), training, and the current realization of the type (note that conditional on these variables, consumption does not depend on human capital $s$). Then, $\omega(\theta) = u_t(c(\theta)) - \phi_t(l(\theta), z(\theta) - z_-) + \beta v(\theta)$ becomes redundant as a constraint, and the choice variables are $(l(\theta), s(\theta), z(\theta), \omega(\theta), v(\theta), \Delta(\theta))$. Let the multipliers in program (2.14) be (in the order of the constraints there) $\mu(\theta)$, $\lambda_-$, and $\gamma_-$. The problem is solved using the optimal control approach where the "types" play the role of the running variable, $\omega(\theta)$ is the state (and $\dot{\omega}(\theta)$ its law of motion), and the controls are $l(\theta), v(\theta), s(\theta)$ and $\Delta(\theta)$. The Hamiltonian is:

$$(\tilde{c}(l(\theta), \omega(\theta) - \beta v(\theta), z(\theta) - z_-, \theta) + M_t(s(\theta) - s_-) - w_t(\theta, s(\theta), z(\theta)) l(\theta)) f^t(\theta|\theta_-)$$

$$+ \frac{1}{R} K(v(\theta), \Delta(\theta), \theta, s(\theta), z(\theta), t+1) f^t(\theta|\theta_-)$$

$$+ \lambda_- \left[ v - \omega(\theta) f^t(\theta|\theta_-) \right] + \gamma_- \left[ \Delta - \omega(\theta) \frac{\partial f^t(\theta|\theta_-)}{\partial \theta_-} \right] + \mu(\theta) \left[ \frac{w_{\theta,t}}{w_t} l(\theta) \phi_{l,t}(l(\theta), z(\theta) - z_-) + \beta \Delta(\theta) \right]$$

with boundary conditions:

$$\lim_{\theta \to \bar{\theta}} \mu(\theta) = \lim_{\theta \to \underline{\theta}} \mu(\theta) = 0$$

Part i) Taking the first order conditions (hereafter, FOC) of the recursive planning problem yields (the variable with respect to which the FOC is taken appears in brackets):

$$[l(\theta)] : \frac{\tau_L^*(\theta)}{1 - \tau_L^*(\theta)} = \frac{\mu(\theta)}{f^t(\theta|\theta_-)} \frac{w_{\theta,t}}{w_t} u_t'(c(\theta)) \left( 1 + \frac{l(\theta) \phi_{ll,t}(l(\theta), i(\theta))}{\phi_{l,t}(l(\theta), i(\theta))} \right)$$

153

using the definitions of $\varepsilon^c$, $\varepsilon^u$ and $\varepsilon$ in the text:

$$\frac{\tau_L^*(\theta)}{1 - \tau_L^*(\theta)} = \frac{\mu(\theta)}{f^t(\theta|\theta_-)} \frac{\varepsilon_{w\theta}}{\theta} u_t'(c(\theta)) \frac{1 + \varepsilon^u}{\varepsilon^c}$$

$$[s(\theta)] \quad : \quad -M_t'(s(\theta) - s_-) + l(\theta) w_{s,t} + \frac{1}{R} \frac{\partial K(v(\theta), \Delta(\theta), \theta, s(\theta), z(\theta), t+1)}{\partial s(\theta)}$$

$$= \frac{\mu(\theta)}{f^t(\theta|\theta_-)} l(\theta) \phi_{l,t}(l(\theta), i(\theta)) \frac{1}{w_t^2} w_{\theta,t} w_{s,t} (\rho_{\theta s,t} - 1)$$

where: (letting $\theta'$ and $s'$ be the next period's type and human capital respectively):

$$\frac{\partial K}{\partial s(\theta)} = \int M_{t+1}'(s'(\theta') - s(\theta)) f^{t+1}(\theta'|\theta) d\theta'$$

so that:

$$-M_t'(s(\theta) - s_-) + l(\theta) w_{s,t} + \frac{1}{R} \int M_{t+1}'(s'(\theta') - s(\theta)) f^{t+1}(\theta'|\theta) d\theta'$$

$$= \frac{\mu(\theta)}{f^t(\theta|\theta_-)} l(\theta) \phi_{l,t}(l(\theta), i(\theta)) \frac{1}{w_t^2} w_{\theta,t} w_{s,t} (\rho_{\theta s,t} - 1)$$

Use the expression for $w_{s,t} l_t$ from the definition of the human capital wedge $\tau_{St}$ in (2.17) to write the first order condition as as a function of the modified wedge $t_{st}$:

$$t_{st}^*(\theta^t) = \frac{\mu(\theta^t)}{f^t(\theta_t|\theta_{t-1})} u'(c_t(\theta^t)) \frac{\varepsilon_{w\theta,t}}{\theta_t} (1 - \rho_{\theta s,t})$$

From this, we can immediately deduce the relation between the modified wedge and the tax rate in the text:

$$t_{st} = \frac{\tau_{Lt}}{(1 - \tau_{Lt})} (1 - \rho_{\theta s,t}) \frac{\varepsilon_t^c}{1 + \varepsilon_t^u}$$

The law of motion for the co-state $\mu(\theta)$ comes from the first-order condition with respect to the state variable $\omega(\theta)$:

$$[\omega(\theta)] : \left( -\frac{1}{u_t'(c(\theta))} + (\lambda_-) + (\gamma_-) \frac{\partial f^t(\theta|\theta_-)}{\partial \theta_-} \frac{1}{f^t(\theta|\theta_-)} \right) f^t(\theta|\theta_-) = \dot{\mu}(\theta) \qquad (2.51)$$

Integrating this and using the boundary condition $\mu\left(\bar{\theta}\right) = 0$, yields:

$$\mu\left(\theta\right) = \int_{\theta}^{\bar{\theta}}\left(\frac{1}{u_t'\left(c\left(\theta\right)\right)} - \left(\lambda_-\right) - \left(\gamma_-\right)\frac{\partial f^t\left(\theta|\theta_-\right)}{\partial\theta_-}\frac{1}{f^t\left(\theta|\theta_-\right)}\right)f^t\left(\theta|\theta_-\right) \tag{2.52}$$

Integrating and using both boundary conditions yields:

$$\lambda_- = \int_{\underline{\theta}}^{\bar{\theta}}\frac{1}{u'\left(c\left(\theta\right)\right)}f^t\left(\theta|\theta_-\right)d\theta \tag{2.53}$$

Using the envelope conditions $\frac{\partial K(v(\theta),\Delta(\theta),\theta,s(\theta),z(\theta),t+1)}{\partial v(\theta)} = \lambda\left(\theta\right)$ and $\frac{\partial K(v(\theta),\Delta(\theta),\theta,s(\theta),z(\theta),t+1)}{\partial\Delta(\theta)} = -\gamma\left(\theta\right)$, the first-order conditions with respect to $v\left(\theta\right)$ and $\Delta\left(\theta\right)$ respectively lead to:

$$\left[v\left(\theta\right)\right]: \frac{1}{u'\left(c\right)} = \frac{\lambda\left(\theta\right)}{R\beta} \tag{2.54}$$

and

$$\left[\Delta\left(\theta\right)\right]: -\frac{\gamma\left(\theta\right)}{R\beta} = \frac{\mu\left(\theta\right)}{f^t\left(\theta|\theta_-\right)} \tag{2.55}$$

Using (2.53) and (2.55) in the expression for $\mu\left(\theta\right)$ from (2.52) yields: $\mu\left(\theta^t\right) = \kappa\left(\theta^t\right) + \eta\left(\theta^t\right)$ where

$$\kappa\left(\theta^t\right) = \int_{\theta_t}^{\bar{\theta}}\frac{1}{u'\left(c\left(\theta\right)\right)}\left(1 - u'\left(c\left(\theta\right)\right)\int_{\underline{\theta}}^{\bar{\theta}}\frac{1}{u'\left(c\left(m\right)\right)}f\left(m|\theta_-\right)dm\right)f^t\left(\theta|\theta_-\right)$$

$$\eta\left(\theta^t\right) = -\left(\gamma_-\right)\int_{\theta_t}^{\bar{\theta}}\frac{\partial f^t\left(\theta|\theta_-\right)}{\partial\theta_-}d\theta = \frac{\mu\left(\theta^{t-1}\right)}{f\left(\theta_{t-1}|\theta_{t-2}\right)}R\beta\int_{\theta_t}^{\bar{\theta}}\frac{\partial f^t\left(\theta|\theta_-\right)}{\partial\theta_-}d\theta$$

where the last equality uses the lag of (2.55). The multiplier is replaced by the last period's $t_{st-1}^*$ (respectively, $\tau_{Lt-1}^*$) using the optimal formulas to obtain the expressions in proposition (20) (respectively, (22)).

Part ii) The proof is immediate by inspection as long as $\mu\left(\theta^t\right) \geq 0$ $\forall t$, $\forall\theta^t$, which is now proved.

**Lemma 5** *Under assumption* (4), $\mu\left(\theta^t\right) \geq 0$ $\forall t$, $\forall\theta^t$.

**Proof of Lemma 5:**

The proof is close to the one in Golosov, Tsyvinski and Troshkin (2011), for a separable utility function and with human capital. From the envelope condition and the FOC for $v\left(\theta\right)$ in

(2.54):

$$\frac{\partial K}{\partial v} = \lambda(\theta) = \frac{R\beta}{u_t'(c(\theta))}$$

Since by assumption $v(\theta)$ is increasing in $\theta$ and $K()$ is increasing and convex in $v$, it must be that $\frac{\partial K}{\partial v}$ is increasing in $\theta$, so that $\frac{1}{u_t'(c(\theta))}$ as well is increasing in $\theta$.

Start in period $t = 1$. In this case, since $\theta_0$ has a degenerate distribution, $\frac{\partial f^1}{\partial \theta_0}(\theta_1|\theta_0) = 0$ and

$$\mu(\theta_1) = \int_{\theta_1}^{\bar{\theta}} \left( \frac{1}{u_1'\left(c\left(\tilde{\theta}_1\right)\right)} - \lambda_- \right) f^1\left(\tilde{\theta}_1\right) d\tilde{\theta}_1$$

Choose the $\theta'$ such that $\frac{1}{u'(c(\theta'))} = \lambda_-$ . Since $\frac{1}{u'(c(\theta))}$ is increasing in $\theta$, for $\theta \geq \theta'$, $\mu(\theta) \geq 0$ (integrating over non-negative numbers only). Using the boundary condition $\mu(\underline{\theta}) = 0$, $\mu(\theta_1)$ can also be rewritten as:

$$\mu(\theta_1) = \int_{\underline{\theta}}^{\theta_1} \left( -\frac{1}{u_1'\left(c\left(\tilde{\theta}_1\right)\right)} + \lambda_- \right) f^1\left(\tilde{\theta}_1\right) d\tilde{\theta}_1$$

Since for $\theta_1 \leq \theta_1'$, $\frac{1}{u'(c(\theta_1))} \leq \lambda_-$, we again have $\mu(\theta_1) \geq 0$. Thus, for all $\theta_1$, $\mu(\theta_1) \geq 0$.

By the first-order condition for $\Delta$ in (2.55):

$$-\frac{\gamma(\theta_1)}{R\beta} = \frac{\mu(\theta_1)}{f(\theta_1)}$$

so that $\gamma(\theta_1) \leq 0$, for all $\theta_1$. Note that $\mu(\theta_2)$ is equal to:

$$\mu(\theta_2) = \int_{\theta_2}^{\bar{\theta}} \left( \frac{1}{u_2'\left(c\left(\tilde{\theta}_2\right)\right)} - (\lambda_-) - \frac{\partial f\left(\tilde{\theta}_2|\theta_1\right)}{\partial \theta_1} \frac{(\gamma_-)}{f\left(\tilde{\theta}_2|\theta_1\right)} \right) f\left(\tilde{\theta}_2|\theta_1\right)$$

Since by assumption (4) iii), $\frac{g(\tilde{\theta}_2|\theta_1)}{f(\tilde{\theta}_2|\theta_1)}$ is increasing in $\tilde{\theta}_2$, and we already showed that $\frac{1}{u'(c(\tilde{\theta}_2))}$ is increasing in $\tilde{\theta}_2$, there is a $\theta_2'$ such that

$$\frac{1}{u_2'\left(c\left(\theta_2'\right)\right)} - \frac{\partial f\left(\theta_2'|\theta_1\right)}{\partial \theta_1} \frac{(\gamma_-)}{f\left(\theta_2'|\theta_1\right)} = \lambda_-$$

and such that for $\theta_2 \geq \theta_2'$, $\mu(\theta^2) \geq 0$ (since integrating over non-negative numbers only). Rewriting $\mu(\theta^2)$ as an integral from $\underline{\theta}$ to $\theta_2$ and using the boundary condition $\underline{\theta} = 0$, we can

again show that $\mu(\theta_2) \geq 0$ also for $\theta_2 \leq \theta'_2$. Proceeding in the same way for all periods up to $T$ shows the result.

**Proof that $t_{st} = 0$ when $\rho_{\theta_s} = 1$ without using the first-order approach:**

Consider a separable wage $w = \theta s$, a history $\theta^t$ and a perturbation of the allocation for all $\tilde{\theta}^t$ in a neighborhood of $\theta^t$, $|\theta^t - \tilde{\theta}^t| \leq \eta$ such that $s^\delta\left(\tilde{\theta}^t\right) = s\left(\tilde{\theta}^t\right) + \delta$ and $y^\delta\left(\tilde{\theta}^t\right) = y\left(\tilde{\theta}^t\right) + dy\left(\tilde{\theta}^t\right)$ such that $\frac{s^\delta\left(\tilde{\theta}^t\right)}{y^\delta\left(\tilde{\theta}^t\right)} = \frac{s\left(\tilde{\theta}^t\right)}{y\left(\tilde{\theta}^t\right)}$, i.e., $dy\left(\tilde{\theta}^t\right) = \delta\frac{y\left(\tilde{\theta}^t\right)}{s\left(\tilde{\theta}^t\right)}$. This perturbation leaves utilities and incentive compatibility constraints unaffected. The change in the resource cost must hence be zero in this neighborhood, and letting $\eta \to 0$, we obtain: $-\frac{y(\theta)}{s(\theta)} + \left(M'_t\left(e\left(\theta\right)\right) - \frac{1}{R}E_t\left(M'_{t+1}\left(e\left(\theta'\right)\right)\right)\right) = 0$, which is equivalent to $t_{st} = 0$ for the multiplicative wage.

Part iii) If $\theta$ is iid, $\gamma_- = 0$ and $\eta\left(\theta^t\right) = 0$ for all $t$. In addition, if $u'_t\left(c_t\right) = 1$ $\forall t$, then $\kappa\left(\theta^t\right) = 0$ as well.

Part iv) is immediate from the boundary conditions $\mu\left(\underline{\theta}\right) = \mu\left(\bar{\theta}\right) = 0$.

**Proof of Corollary (1):**

The general formula for $\rho_{\theta_s} \neq 1$ is obtained from the FOC for $s_t$, together with the definition of the gross wedge in (2.17), and replacing the multiplier $\mu\left(\theta^t\right)$ by the expression from the optimal tax formula:

$$\tau_{St}\left(\theta^t\right) = \left(1 + \left(1 - \rho_{\theta s, t}\right)\frac{\varepsilon^c_t}{1 + \varepsilon^u_t}\right)\left(M'^d_t - \tau_{St}\left(\theta^t\right)\right)\left[\frac{\tau_{Lt}\left(\theta^t\right)}{1 - \tau_{Lt}\left(\theta^t\right)}\right]$$
$$- \frac{1}{R}\left(1 - \xi'_{M'_t}\right)E_t\left(M'_{t+1}\left(s_{t+1}\left(\theta^t, \theta_{t+1}\right) - s_t\left(\theta^t\right)\right)\right)\left[\frac{\tau_{Kt}\left(\theta^t\right)}{1 - \tau_{Kt}\left(\theta^t\right)}\right]$$

When $\rho_{\theta s, t} = 1$ $\forall t$, the expression becomes $\tau^*_{St} = \tau^*_{Lt}M'^d_t - \frac{\tau^*_{Kt}}{R\left(1 - \tau^*_{Kt}\right)}\left(1 - \tau^*_{Lt}\right)\left(1 - \xi'_{M', t+1}\right)E_t\left(M'_{t+1}\right)$, equivalent to the one in the corollary.

**Proof of Proposition (21):**

Taking integral of $\mu\left(\theta\right)$ in equation (2.51) between the two boundaries, $\bar{\theta}$ and $\underline{\theta}$, and using the boundary conditions $\mu\left(\bar{\theta}\right) = \mu\left(\underline{\theta}\right) = 0$, as well as the expression for $\lambda_-$ from (2.54), lagged by one period, yields the inverse Euler equation in (2.24).

**Proof of Proposition (23) :**

Taking the first order condition of the Hamiltonian in the proof of proposition (20) with respect to $z_t$ yields:

$$[z_t] : \left( -\frac{\phi_{z,t}}{u'(c_t)} + w_{z,t} l_t + \int \frac{1}{R} \frac{\phi_{z,t+1}}{u'(c_{t+1})} f^{t+1}\left(\theta_{t+1}|\theta_t\right) + \frac{1}{R} E_t \left( \frac{\mu\left(\theta_{t+1}\right)}{f^{t+1}\left(\theta_{t+1}|\theta_t\right)} \frac{w_{\theta,t+1}}{w_{t+1}} l_{t+1} \phi_{lz,t+1} \right) \right)$$

$$- \frac{\mu\left(\theta_t\right)}{f^t\left(\theta_t|\theta_{t-1}\right)} \left( l_t \phi_{l,t}\left(l_t, z_t - z_{t-1}\right) \frac{1}{w_t} w_{\theta z,t} - l_t \phi_{l,t} w_{\theta,t} w_{z,t} \frac{1}{w_t^2} + \frac{w_{\theta,t}}{w_t} l_t \phi_{lz,t} \right) = 0$$

Using the definition of the training time bonus, $\tau_{Zt}$ to replace for $w_{z,t} l_t$ yields:

$$-\frac{\phi_{z,t}}{u'(c_t)} + w_{z,t} l_t + \int \frac{1}{R} \frac{\phi_{z,t+1}}{u'(c_{t+1})} f^{t+1}\left(\theta_{t+1}|\theta_t\right) = -\frac{1}{\left(1 - \tau_{Lt}\right)} \left( \tau_{Zt} - \tau_{Lt} \left(\frac{\phi_z}{u'(c_t)}\right)^d + P_{Zt} \right)$$

with $P_{Zt}$ and $\left(\frac{\phi_z}{u'(c_t)}\right)^d$ as defined in the text. The FOC then becomes:

$$[z_t] \quad : \quad t_{zt} = \frac{\left( \tau_{zt} - \tau_{Lt} \left(\frac{\phi_{z,t}}{u'(c_t)}\right)^d + P_{Zt} \right)}{\left(1 - \tau_{Lt}\right) \left[ \left(\frac{\phi_{z,t}}{u'(c_t)}\right)^d - \tau_{zt} \right]} = \frac{\mu\left(\theta_t\right)}{f^t\left(\theta_t|\theta_{t-1}\right)} u'(c_t) \frac{\varepsilon_{w\theta,t}}{\theta_t} \left(1 - \rho_{\theta z,t}\right)$$

$$- \frac{1}{\left[ \left(\frac{\phi_{z,t}}{u'(c_t)}\right)^d - \tau_{zt} \right]} \left( \frac{\mu\left(\theta_t\right)}{f^t\left(\theta_t|\theta_{t-1}\right)} \frac{\varepsilon_{w\theta,t}}{\theta_t} l_t \phi_{lz,t} - \frac{1}{R} E_t \left( \frac{\mu\left(\theta_{t+1}\right)}{f^{t+1}\left(\theta_{t+1}|\theta_t\right)} \frac{\varepsilon_{w\theta,t+1}}{w_{t+1}} l_{t+1} \phi_{lz,t+1} \right) \right)$$

We can again replace the multipliers by the labor wedges, and note that

$$\frac{w_t}{w_{z,t} \phi_{l,t}\left(l_t\right)} \phi_{lz,t} = \frac{\varepsilon_{\phi z,t}}{\varepsilon_{wz,t}} \rho^{\phi}_{lz,t}$$

to obtain formula (2.32).

**Proof of Corollary (3) :**

The derivation of the time evolution of the labor wedge follows Farhi and Werning (2013). Take any weighting function $\pi\left(\theta\right) > 0$ and let $\Pi\left(\theta\right)$ denote a primitive of $\pi\left(\theta\right)/\theta$. Starting from the expression of the optimal labor wedge in (2.25), multiply both sides of the expression

158

by $\pi(\theta) > 0$. Integrating by parts, yields:

$$\int \frac{\tau_{Lt}(\theta^t)}{(1 - \tau_{Lt}(\theta^t))} f^t(\theta_t|\theta_{t-1}) \frac{\theta_t}{\varepsilon_{w\theta,t}} \frac{\varepsilon_t^c}{1 + \varepsilon_t^u} \frac{1}{u_t'(c(\theta^t))} \frac{\pi(\theta_t)}{\theta_t} d\theta_t = \int \mu(\theta^t) \frac{\pi(\theta_t)}{\theta_t} d\theta_t$$

$$= -\int \dot{\mu}(\theta^t) \Pi(\theta_t) d\theta_t$$

$$= \int \left( \frac{1}{u_t'(c(\theta^t))} - \frac{R\beta}{u_{t-1}'(c(\theta^{t-1}))} + \frac{R\beta\mu(\theta^{t-1})}{f(\theta_{t-1}|\theta_{t-2})} \frac{\partial f^t(\theta_t|\theta_{t-1})}{\partial \theta_{t-1}} \frac{1}{f^t(\theta_t|\theta_{t-1})} \right) f^t(\theta_t|\theta_{t-1}) \Pi(\theta_t) d\theta_t$$

$$= \int \left( \frac{1}{u_t'} - \frac{R\beta}{u_{t-1}'} + \frac{R\beta\tau_{Lt-1}}{(1 - \tau_{Lt-1})} \frac{1}{u_{t-1}'} \frac{\theta_{t-1}}{\varepsilon_{w\theta,t-1}} \frac{\varepsilon_{t-1}^c}{1 + \varepsilon_{t-1}^u} \frac{\partial f^t(\theta_t|\theta_{t-1})}{\partial \theta_{t-1}} \frac{1}{f^t(\theta_t|\theta_{t-1})} \right) f^t(\theta_t|\theta_{t-1}) \Pi(\theta_t) d\theta_t d\theta_t$$

where the third line uses the expression for $\lambda_-$ and $\gamma_-$ from respectively (2.54) and (2.55) evaluated at $t-1$. The 4th line uses the optimal wedge from (2.25) at time $t-1$ to substitute for the multiplier $\mu(\theta^{t-1})$. Using the inverse Euler Equation, yields a general formula for any stochastic process and weighting function:

$$E_{t-1}\left( \frac{\tau_{Lt}(\theta^t)}{(1 - \tau_{Lt}(\theta^t))} \frac{\theta_t}{\varepsilon_{w\theta,t}} \frac{\varepsilon_t^c}{1 + \varepsilon_t^u} \frac{u_{t-1}'(c(\theta^{t-1}))}{u_t'(c(\theta^t))} \frac{\pi(\theta_t)}{\theta_t} \right) \tag{2.56}$$

$$= Cov\left( \frac{u_{t-1}'(c(\theta^{t-1}))}{u_t'(c_t(\theta^t))}, \Pi(\theta_t) \right) + \frac{R\beta\tau_{Lt-1}}{(1 - \tau_{Lt-1})} \frac{\theta_{t-1}}{\varepsilon_{w\theta,t-1}} \frac{\varepsilon_{t-1}^c}{1 + \varepsilon_{t-1}^u} \int \frac{\partial f^t(\theta_t|\theta_{t-1})}{\partial \theta_{t-1}} \Pi(\theta_t) d\theta_t$$

For the particular weighting function $\pi(\theta_t) = 1$ (with $\Pi(\theta_t) = log(\theta_t)$), and with the AR(1) process assumed for $\log(\theta_t)$, the formula becomes as in (2.29).

**Proof of Corollary (4):**

The net wedge on human capital can be rewritten similarly as in the proof of Proposition 3, using the same weighting function $\pi(\theta) = 1$:

$$\int \frac{t_{st}}{u_t'(c_t(\theta^t)) \varepsilon_{w\theta,t}} \frac{1}{(1 - \rho_{\theta s,t})} f^t(\theta_t|\theta_{t-1})$$

$$= \int \left( \frac{1}{u_t'(c_t(\theta^t))} - \frac{R\beta}{u_{t-1}'(c(\theta^{t-1}))} \right) f^t(\theta_t|\theta_{t-1}) \log(\theta_t) d\theta_t + R\beta \frac{\mu(\theta^{t-1})}{f(\theta_{t-1}|\theta_{t-2})} \int \frac{\partial f^t(\theta_t|\theta_{t-1})}{\partial \theta_{t-1}} \log(\theta_t) d\theta_t$$

where the second line uses the expression for $\lambda_-$ and $\gamma_-$ from respectively (2.54) and (2.55) evaluated at $t-1$. Using the optimal wedge from (2.20) at time $t-1$ to substitute for the multiplier $\mu(\theta^{t-1})$, the boundary conditions for $\mu$ (and the resulting Inverse Euler Equation), and the log AR(1) process for $\theta_t$, formula (2.30) in the text is obtained.

## 2.A.2 Implementation

**A sequential reformulation of the recursive allocations:**

Any solution to the recursive social planner's problem can be mapped into a solution which depends on the full past history, using a recursive construction. To see this, denote the solutions to the recursive problem at time $t$, for each realized type $\theta_t$, as a function of all state variables by:

$$v_t^* \left( v_{t-1}, \Delta_{t-1}, s_{t-1}, \theta_{t-1}, \theta_t \right), \Delta_t^* \left( v_{t-1}, \Delta_{t-1}, s_{t-1}, \theta_{t-1}, \theta_t \right), \omega_t^* \left( v_{t-1}, \Delta_{t-1}, s_{t-1}, \theta_{t-1}, \theta_t \right),$$

$$y_t^* \left( v_{t-1}, \Delta_{t-1}, s_{t-1}, \theta_{t-1}, \theta_t \right), s_t^* \left( v_{t-1}, \Delta_{t-1}, s_{t-1}, \theta_{t-1}, \theta_t \right), c_t^* \left( v_{t-1}, \Delta_{t-1}, s_{t-1}, \theta_{t-1}, \theta_t \right)$$

and the solutions to the planner's sequential problem by $\left\{ x_t^* \left( \theta^t \right) \right\} = \left\{ y_t^* \left( \theta^t \right), s_t^* \left( \theta^t \right), c_t^* \left( \theta^t \right) \right\}$ The dependence on initial promised utilities in period 1, due to initial heterogeneity, is dropped for notational convenience; they can be just carried as an additional conditioning variable. This allocation gives rise to a sequence of utilities for the agent, generated recursively by:

$$\omega_t^* \left( \theta^t \right) = u_t \left( c_t^* \left( \theta^t \right) \right) - \phi_t \left( \frac{y_t^* \left( \theta^t \right)}{w_t \left( \theta_t, s_t^* \left( \theta^t \right) \right)} \right) + \beta \int \omega_{t+1}^* \left( \theta^t, \theta_{t+1} \right) f^{t+1} \left( \theta_{t+1} | \theta_t \right) d\theta_{t+1}$$

$$v_t^* \left( \theta^t \right) = \int \omega_{t+1}^* \left( \theta^t, \theta_{t+1} \right) f^{t+1} \left( \theta_{t+1} | \theta_t \right) d\theta_{t+1}$$

$$\Delta_t^* \left( \theta^t \right) = \int \omega_{t+1}^* \left( \theta^t, \theta_{t+1} \right) \frac{\partial}{\partial \theta_t} f^{t+1} \left( \theta_{t+1} | \theta_t \right) d\theta_{t+1}$$

To initialize the allocations, set $\omega_1^* \left( \theta_1 \right) = \omega_1^* \left( v_0, \Delta_0, s_0, \theta_0, \theta_1 \right) = \underline{U} \left( \theta_1 \right)$ (if there is initial heterogeneity in $\theta_1$, with $v_0$ arbitrary in that case, $\Delta_0 = 0$), $y_1^* \left( \theta_1 \right) = y_1^* \left( v_0, \Delta_0, s_0, \theta_0, \theta_1 \right)$, $s_1^* \left( \theta_1 \right) = s_1^* \left( v_0, \Delta_0, s_0, \theta_0, \theta_1 \right)$, and construct iteratively the full, history dependent allocation for all histories $\theta^t$ using the difference equations:

$$\omega_t^* \left( \theta^t \right) = \omega_t^* \left( v_{t-1}^* \left( \theta^{t-1} \right), \Delta_{t-1}^* \left( \theta^{t-1} \right), s_{t-1}^* \left( \theta^{t-1} \right), \theta_{t-1}, \theta_t \right)$$

where $v_{t-1}^*$ and $\Delta_{t-1}^*$ are written as functions of $\omega_{t-1}$, i.e.,

$$v_{t-1}^* \left( \theta^{t-1} \right) = \frac{1}{\beta} \left[ \omega_{t-1}^* \left( \theta^{t-1} \right) - u_{t-1} \left( c_{t-1}^* \left( \theta^{t-1} \right) \right) + \phi_{t-1} \left( \frac{y_{t-1}^* \left( \theta^{t-1} \right)}{w_{t-1} \left( \theta_{t-1}, s_{t-1}^* \left( \theta^{t-1} \right) \right)} \right) \right]$$

$$\Delta_{t-1}^* \left( \theta^{t-1} \right) = \frac{1}{\beta} \left[ \frac{\partial \omega_{t-1}^* \left( \theta^{t-1} \right)}{\partial \theta_{t-1}} - \phi_{t-1}'(y_{t-1}^* \left( \theta^{t-1} \right) / w_{t-1} \left( s_{t-1}^* \left( \theta^{t-1} \right), \theta_{t-1} \right)) \frac{y_{t-1}^* \left( \theta^{t-1} \right)}{w_{t-1} \left( s_{t-1}^* \left( \theta^{t-1} \right), \theta_{t-1} \right)^2} \frac{\partial w_{t-1}}{\partial \theta_{t-1}} \right]$$

Construct also:

$$
\begin{aligned}
y_t^* \left( \theta^t \right) &= y_t^* \left( v_{t-1}^* \left( \theta^{t-1} \right), \Delta_{t-1}^* \left( \theta^{t-1} \right), s_{t-1}^* \left( \theta^{t-1} \right), \theta_{t-1} \right) \\
s_t^* \left( \theta^t \right) &= s_t^* \left( v_{t-1}^* \left( \theta^{t-1} \right), \Delta_{t-1}^* \left( \theta^{t-1} \right), s_{t-1}^* \left( \theta^{t-1} \right), \theta_{t-1} \right) \\
c_t^* \left( \theta^t \right) &= c_t^* \left( v_{t-1}^* \left( \theta^{t-1} \right), \Delta_{t-1}^* \left( \theta^{t-1} \right), s_{t-1}^* \left( \theta^{t-1} \right), \theta_{t-1} \right)
\end{aligned}
$$

Using these sequential optimal policies in the planner's problem we can rewrite the costs as a function of the history, i.e., $K_t \left( \theta^{t-1} \right) = K_t \left( v_{t-1}^* \left( \theta^{t-1} \right), s_{t-1}^* \left( \theta^{t-1} \right), \theta_{t-1} \right)$. or, if agents also have heterogeneous initial wealth levels, $K_t \left( b_0, \theta^{t-1} \right)$. Note that in the planner's problem (i.e., in the direct revelation mechanism) initial wealth holdings are irrelevant since the planner can fully observe and allocate consumption. Since agents can borrow at the same rate as the government, without loss of generality, agents can do all the borrowing and saving on their own. The implicit wealth levels are then defined recursively as: $\frac{1}{R} b_t^* \left( b_0, \theta^t \right) = y_t^* \left( b_0, \theta^t \right) - c_t^* \left( b_0, \theta^t \right) - M \left( e_t^* \left( b_0, \theta^t \right) \right) + b_{t-1}^* \left( b_0, \theta^t \right)$.

**General Deductibility Scheme**

With nonlinear cost, the general expression for the deferred deductibility scheme is:

$$
\begin{aligned}
-\frac{\partial T_t}{\partial e_t} &= \sum_{j=1}^{T-t} \beta^{j-1} E_t \left( \frac{u_{t+j-1}'}{u_t'} \frac{\partial T_{t+j-1}}{\partial y_{t+j-1}} \left( M_{t+j-1}' - \frac{1}{R} M_{t+j}' \right) \right) + \beta^{T-t} E_t \left( \frac{u_T'}{u_t'} \left( \frac{\partial T_T}{\partial y_T} \right) \right) \quad (2.57) \\
&\quad - \sum_{j=1}^{T-t} \beta^j E_t \left( \frac{u_{t+j}'}{u_t'} \left( \left( 1 - \xi_{M',t+j}' \right) E_{t+j-1} \left( M_{t+j}' \right) \frac{\partial T_{t+j}}{\partial b_{t+j-1}} - \frac{\partial T_{t+j}}{\partial s_{t+j-1}} \right) \right)
\end{aligned}
$$

The first set of terms capture the deferred deductibility from the income tax base. Because the marginal cost is no longer constant, the deduction in period $t + j$ occurs at the dynamic marginal cost effective in that period $(M_{t+j}' - \frac{1}{R} M_{t+j+1}')$, not at the "historic" marginal cost faced by the agent at the time of the purchase $M_t'$, i.e., a purchase of $\Delta e$ at time $t$ is deducted as $(M_{t+j}' - \frac{1}{R} M_{t+j+1}') \Delta e$ from $y_{t+j}$ at $t + j$. Otherwise, there would be arbitrage possibilities.[75] Similarly to the text, the "no-arbitrage" term takes into account the differential tax increases

---

[75] Note that with linear cost, as in the main text, this is just $(1 - \beta)$ with $\beta = \frac{1}{R}$ for all $t < T$, and 1 for $t = T$.

from physical capital versus human capital, except that now the nonlinear, risk adjusted cost $\left(1 - \xi'_{M',t+j}\right) E_{t+j-1} \left(M'_{t+j}\right)$ enters the picture. The deduction is risk adjusted, as witnessed by the insurance factors, $\xi'_{M',t+1} \equiv -Cov\left(\frac{\beta u'_{t+1}}{u'_t} - 1, M'_{t+1}\right) / \left(E_t\left(\frac{\beta u'_{t+1}}{u'_t} - 1\right) E_t\left(M'_{t+1}\right)\right)$, defined in the text.

As stated in the text, formula (2.40) can be rewritten in terms of the risk-adjusted, dynamic cost:

$$-\frac{\partial T_t}{\partial e_t} + \beta E_t \left(\frac{u'_{t+1}}{u'_t}\left(\frac{\partial T_{t+1}}{\partial e_{t+1}} - \frac{\partial T_{t+1}}{\partial s_t}\right)\right) = \frac{\partial T_t}{\partial y_t} M'^d_t - \left(1 - \frac{\partial T_t}{\partial y_t}\right)\frac{1}{R}\left(1 - \xi'_{M',t+1}\right) E_t\left(\beta R\frac{u'_{t+1}}{u'_t}\frac{\partial T_{t+1}}{\partial b_t}\right) E_t\left(M'_{t+1}\right)$$
(2.58)

Solving this relation forward, yields the analogous to (2.57) with $M'^d_t$.

## Proof of Proposition (24)

In the first step, we construct the history-independent savings tax, in the spirit of Werning (2011), with added human capital. Consider an incentive compatible allocation expressed as a function of the full history $c_t\left(\theta^t\right)$, $y_t\left(\theta^t\right)$, $e_t\left(\theta^t\right)$, $s_t\left(\theta^t\right)$, and its associated continuation utility $\omega_t\left(\theta^t\right)$, and suppose it is implemented as the outcome of a direct revelation mechanism with no savings. Allow agents to save any desired amount, with the restriction $b_T \geq 0$ (end of period $T$ asset level). Consider a general tax function $T_{Kt}\left(b_t, r^t\right)$ as a function of savings and the history of reports up to period $t$. Given the report of the agent up to period $t-1$, and the report of the current shock, the planner assigns $c_t\left(r^{t-1}, r_t\right)$, $y_t\left(r^{t-1}, r_t\right)$, $e_t\left(r^{t-1}, r_t\right)$ and the agent chooses savings $b_t$. Let $V_t\left(b_{t-1}, r^{t-1}, \theta_t\right)$ be the continuation value of an agent with beginning of period savings $b_{t-1}$, a history of reports $r^{t-1}$, and a realized shock $\theta_t$. The agent's problem is:

$$V_t\left(b_{t-1}, r^{t-1}, \theta_t\right) = \max_{r_t, b_t} \tilde{V}_t\left(b_{t-1}, b_t, r^{t-1}, r_t, \theta_t\right)$$

with $\tilde{V}_t\left(b_{t-1}, b_t, r^{t-1}, r_t, \theta_t\right)$ defined as the value from saving an amount $b_t$ and reporting $r_t$:

$$\tilde{V}_t\left(b_{t-1}, b_t, r^{t-1}, r_t, \theta_t\right) = \left\{\begin{array}{c} u_t\left(c_t\left(r^{t-1}, r_t\right) + b_{t-1} - \left(\frac{1}{R}b_t + T_{Kt}\left(b_t, r^{t-1}, r_t\right)\right)\right) - \phi_t\left(\frac{y\left(r^{t-1}, r_t\right)}{w_t\left(\theta_t, s_t\left(r^{t-1}, r_t\right)\right)}\right) \\ + \beta E\left(V_{t+1}\left(b_t, r^t, \theta_{t+1}\right)|\theta_t\right) \end{array}\right\}$$

In period $T - 1$, define for each type realization $\theta_{T-1}$, current asset level $b_{T-2}$, history of reports $r^{T-1}$, and savings levels $b_{T-1}$ a fictitious tax $T^\theta_{Kt}$ which makes an agent just indifferent

between saving $b_{T-1}$ and saving zero.

$$\omega_{T-1}\left(\theta^{T-1}\right) = u_{T-1}\left(c_{T-1}\left(r^{T-1}\right) + b_{T-2} - \frac{1}{R}b_{T-1} - T^{\theta}_{KT-1}\left(b_{T-1}, r^{T-1}, \theta_{T-1}\right)\right)$$

$$-\phi_{T-1}\left(\frac{y_{T-1}\left(r^{T-1}\right)}{w_{T-1}\left(\theta_{T-1}, s_{T-1}\left(r^{T-1}\right)\right)}\right) + \beta E\left(V_T\left(b_{T-1}, r^{T-1}, \theta_T\right)|\theta_{T-1}\right)$$

Taking the sup over all types $\theta_{T-1}$ yields a history-dependent, but type-independent savings tax $T^r_K\left(b_{T-1}, r^{T-1}\right) = \sup_{\theta_{T-1}} T^{\theta}_{KT-1}\left(b_{T-1}, r^{T-1}, \theta_{T-1}\right)$.

By induction, suppose that in period $t$ the agent is faced with a continuation value function $V_{t+1}\left(b_t, r^t, \theta_{t+1}\right)$. Define the tax function for period $t$ as $T_{Kt}\left(b_t, r^t\right) = \sup_{\theta_t} T^{\theta}_{Kt}\left(b_t, r^t, \theta_t\right)$ with $T^{\theta}_{Kt}\left(b_t, r^t, \theta_t\right)$ to equate:

$$\omega_t\left(\theta^t\right) = u_t\left(c_t\left(r^t\right) + b_{t-1} - \frac{1}{R}b_t - T^{\theta}_{Kt}\left(b_t, r^t, \theta_t\right)\right) - \phi_t\left(\frac{y_t\left(r^t\right)}{w_t\left(\theta_t, s_t\left(r^t\right)\right)}\right) + \beta E\left(V_{t+1}\left(b_t, r^t, \theta_{t+1}\right)|\theta_t\right)$$

Work backwards to define the tax functions in this fashion for all periods. The sequence of tax functions $\left\{T_{Kt}\left(b_t, r^t\right)\right\}_{t=1}^{T-1}$ thus defined implement zero savings each period for all sequences of reports. Next, take the supremum over all histories of reports $r^t$ to obtain a history independent tax which implements zero savings.

$$T_K\left(b\right) = \sup_{r^t} T_{Kt}\left(b, r^t\right)$$

In the second step, we construct the loan and repayment schedule that mimics the direct mechanism above. Note that $L^t$ can directly be mapped into a history of human capital and education levels $e^t$ (recall that $s_0 = 0$), using that $e_t = M_t^{-1}\left(L_t\right)$ Hence, $\left(L^{t-1}, y^{t-1}\right)$ and $\left(e^{t-1}, y^{t-1}\right)$ will be used interchangeably. First, set implicit finite (but potentially very large) upper and lower limits on asset holdings, $\bar{b} > 0$ and $\underline{b} < 0$. This can be done either by extending the proposed savings tax so that for $b_t > \bar{b}$ and $b_t < \underline{b}$, $\forall t$, it is confiscatory (e.g., tax away all wealth and imposes a large penalty on borrowing) or by directly setting borrowing and saving limits. Let $B \equiv \left[\underline{b}, \bar{b}\right]$.

In each period, all allocations which can arise as outcomes in the optimum are made afford-

able:

$$D_t\left(L^{t-1}, y^{t-1}, e_t^*\left(\theta^{t-1}, \theta\right), y_t^*\left(\theta^{t-1}, \theta\right)\right) + T_Y\left(y_t^*\left(\theta^{t-1}, \theta\right)\right) = y_t^*\left(\theta^{t-1}, \theta\right) - c_t^*\left(\theta^{t-1}, \theta\right) \quad (2.59)$$

$$L_t\left(e_t^*\left(\theta^{t-1}, \theta\right)\right) = M_t\left(e_t^*\left(\theta^{t-1}, \theta\right)\right) \quad (2.60)$$

for all $\left(L^{t-1}, y^{t-1}\right)$ such that $\theta^{t-1} \in \Theta^{t-1}\left(\left\{M_1^{-1}\left(L_1\right), ..., M_{t-1}^{-1}\left(L_{t-1}\right)\right\}, y^{t-1}\right) \neq \emptyset$ and all $\theta \in \Theta$.

To mimic the direct revelation mechanism, we need to exclude pairs $(e_t, y_t)$ which would not be assigned to any type $\theta_t$ in the social planner's problem after a history $\theta^{t-1}$, and, consequently, exclude histories $\left(y^{t-1}, e^{t-1}\right)$, which do not correspond to any history $\theta^{t-1}$. Call "non-allowed" a choice which is not assigned in the social planner's problem for *any* type $\theta_t$ after history $\theta^{t-1}$, i.e., such that $(e_t, y_t) \notin Q_{e,y}^{t-1}\left(\theta^{t-1}\right)$. The repayments at non-allowed levels have to be sufficiently dissuasive to make them strictly dominated by allowed choices. Then the history of reports would exactly be tracked by $r^{t-1} = \theta^{t-1} \in Q_{e,y}^{t-1}\left(y^{t-1}, e^{t-1}\right)$, and the agent's problem becomes equivalent to making a report $r_t \in \Theta$ in each period, by choosing a pair $(e_t, y_t)$ designed for some $\theta_t$ after $\theta^{t-1}$. We know that in this case, the previously constructed history-independent savings tax would enforce zero savings.

There are several ways to rule out non-allowed allocations, and the goal here is just to provide a possible one, which is to simply set the repayment prohibitively high, so that irrespective of savings, it is never optimal to chose such allocations. For instance, after a history $\theta^{t-1} \in \Theta^{T-1}\left(e^{T-1}, y^{T-1}\right)$ and for any choice $(e_t, y_t) \notin Q_{ey}^t\left(\theta^{t-1}\right)$, set

$$D_t\left(L^{t-1}, y^{t-1}, e_t, y_t\right) + T_Y\left(y_t\right) > \left[\bar{b} - \underline{b} + y_t\right]$$

i.e., the repayment plus income tax at least confiscate income and impose an additional penalty such that all wealth is confiscated and agents can never borrow sufficiently to retain positive consumption.[76] This leaves the agent with zero consumption, and will never be chosen. More generally, arbitrarily large repayments can be set. The second and less draconian way is to take

---

[76] To extend the repayment scheme's domain to histories $L^{t-1}, y^{t-1}$ for which $\Theta^{T-1}\left(e^{T-1}, y^{T-1}\right) = \emptyset$, set for all $e_t$, $y_t$ after such histories:

$$D_t\left(L^{t-1}, y^{t-1}, e_t, y_t\right) + T_Y\left(y_t\right) > \bar{b} - \underline{b} + y_t$$

the envelope of the repayment schedules which, after each history, for any possible beginning of period wealth and optimal savings choice, would make the agent just indifferent between any non-allowed allocation and his optimal allocation.[77] Whatever the method chosen, once the non-allowed choices are ruled out, each period, after every history, the agent faces a problem equivalent in outcomes to the direct revelation mechanism, i.e., he faces only allocations which are also available to him in the social planner's problem after that history. Accordingly, the savings tax ensures that he will find it optimal not to save. By temporal incentive compatibility, he will chose the allocation designed for him.

**Proof of Proposition (25) :**

The proof is an extended and modified version of the proof of Proposition in Albanesi and Sleet (2006), adding human capital. With iid shocks, the recursive formulation of the relaxed program no longer requires the states $\Delta$ and $\theta_{t-1}$:

$$K\left(v, s_-, t\right) = \min_{\left(c(\theta), l(\theta), \omega(\theta), s(\theta), v(\theta)\right)} \int \left[c\left(\theta\right) + M_t\left(s\left(\theta\right) - s_-\right) - w\left(\theta, s\left(\theta\right)\right) l\left(\theta\right) + \frac{1}{R} K\left(v\left(\theta\right), s\left(\theta\right), t+1\right)\right] f\left(\theta\right) d\theta$$

subject to:

$$\omega\left(\theta\right) = u_t\left(c\left(\theta\right)\right) - \phi_t\left(l\left(\theta\right)\right) + \beta v\left(\theta\right)$$

$$\dot{\omega}\left(\theta\right) = \frac{w_{\theta,t}}{w_t} l\left(\theta\right) \phi_{l,t}\left(l\left(\theta\right)\right)$$

$$v = \int \omega\left(\theta\right) f^t\left(\theta|\theta_-\right) d\theta$$

Denote by $K^{-1,s}$ and $K^{-1,v}$ the partial inverse functions of $K\left(v, s, t\right)$ with respect to its arguments $s$ and $v$ respectively. Define the set

$$Q_{e,y}^t\left(b_-, s_-\right) = \left\{e, y : e = e_t^*\left(v, s_-, \theta_t\right), y = y_t^*\left(v, s_-, \theta_t\right) \text{ for some } \theta_t \in \Theta, \text{ with } v = K^{-1,v}\left(b_-, s_-\right)\right\}$$

to be the set of output levels $y_t$ and education levels $e_t$ which are available to an agent with promised utility $v$ and previous human capital level $s_-$ in the social planner's problem. The

---

[77]A more sophisticated implementation, which smooths the repayment schedule to make it differentiable is currently explored by the author. That implementation involves adding wealth as a conditioning variable in the repayment function. See as well the next implementation below.

value function of the agent who starts a period with wealth $b_{t-1}$ and human capital $s_{t-1}$ is denoted by $V_t(b_{t-1}, s_{t-1})$, as in the main text.

In each period $t$, the agent's problem can be split into two stages, because of the separability between consumption and labor. In stage 1, he chooses labor supply $l_t$ (equivalently, output $y_t$) and human capital expenses $e_t$. He pays a tax $T_t(b_{t-1}, s_{t-1}, y_t, e_t)$ and is left with a total resource amount $b_t^m = y_t - T_t(b_{t-1}, s_{t-1}, y_t, e_t) - M(e_t) + b_{t-1}$. In stage 2, he chooses consumption and next-period bond holdings, $b_t$ to maximize $V_t^m(b_t^m, s_{t-1} + e_t)$, the intermediate value function from resource level $b_t^m$, defined as:

$$V_t^m(b_t^m, s_t) = \max_{c_t, b_t} (u_t(c_t) + \beta V_{t+1}(b_t, s_t))$$
$$s.t. \quad : \quad b_t^m = c_t + \frac{1}{R} b_t$$

with $V_T^m(b_T^m, s_T) = u_T(b_T^m)$. Denote the market outcomes by $\hat{b}_t(b_t^m, s_t)$ and $\hat{c}_t(b_t^m, s_t)$. In stage 1, the problem of the agent is:

$$V_t(b_{t-1}, s_{t-1}) = \max_{\{y_t(\theta), e_t(\theta), b_t^m(\theta)\}} \int_\Theta (-\phi_t(y_t(\theta)/w_t(\theta, s_{t-1} + e_t(\theta))) + V_t(b_t^m(\theta), s_t(\theta))) f(\theta) d\theta$$
$$s.t. \quad : \quad b_t^m = y_t - T_t(b_{t-1}, s_{t-1}, y_t, e_t) - M(e_t) + b_{t-1}$$

Let the market outcomes be denoted by $\hat{y}_t(b_{t-1}, s_{t-1}, \theta_t)$, $\hat{e}_t(b_{t-1}, s_{t-1}, \theta_t)$, and $\hat{b}_t^m(b_{t-1}, s_{t-1}, \theta_t)$ for each type $\theta_t$.

In each period, the planner solves a two-stage problem as well. In the first stage, he allocates human capital expenses $e_t$, output $y_t$, and an intermediate promised utility $v_t^m$. In the second stage, he allocates consumption $c_t$, and continuation utility $v_t$. In stage 2, given an intermediate promised utility $v_t^m$, and an acquired human capital level $s_{t-1} + e_t = s_t$, the planner solves the program with intermediate continuation cost function $K^m(v_t^m, s_t, t)$:

$$K^m(v_t^m, s_t, t) = \min_{c_t, v_t} \left( c_t + \frac{1}{R} K(v_t, s_t, t+1) \right)$$
$$s.t. \quad : \quad u_t(c_t) + \beta v_t = v_t^m$$

with $K(v_T, s_T, T+1) \equiv 0$. Denote the solutions to this problem by $c_t^*(v_t^m, s_t)$ and $v_t(v_t^m, s_t)$.

In stage 1, the problem of the planner is hence:

$$K\left(v_{t-1}, s_{t-1}, t\right) = \min_{\{v_t^m(\theta), e_t(\theta), y_t(\theta)\}} \int_\Theta \left(M\left(e_t\left(\theta\right)\right) - y_t\left(\theta\right) + K^m\left(v_t^m\left(\theta\right), s_{t-1} + e_t\left(\theta\right), t\right)\right) f\left(\theta\right) d\theta$$

$$s.t. \quad : \quad v_t^m\left(\theta\right) - \phi_t\left(y_t\left(\theta\right) / w_t\left(\theta, s_{t-1} + e_t\left(\theta\right)\right)\right)$$

$$\geq \quad v_t^m\left(\theta'\right) - \phi_t\left(y_t\left(\theta'\right) / w_t\left(\theta, s_{t-1} + e_t\left(\theta'\right)\right)\right) \quad \forall \theta, \theta'$$

$$\int_\Theta \left(v_t^m\left(\theta\right) - \phi_t\left(y_t\left(\theta\right) / w_t\left(\theta, s_{t-1} + e_t\left(\theta\right)\right)\right)\right) f\left(\theta\right) d\theta = v_{t-1}$$

Denote the solutions to the planner problem by $v_t^{m*}\left(v_{t-1}, s_{t-1}, \theta_t\right)$, $e_t^*\left(v_{t-1}, s_{t-1}, \theta_t\right)$, and $y_t^*\left(v_{t-1}, s_{t-1}, \theta_t\right)$.

In stage 2, the problem of an agent who has acquired human capital $s_t$ and who starts with intermediate wealth $b_t^m = K^m\left(v_t^m, s_t, t\right)$ is exactly the dual of the planner's problem who has promised utility $v_t^m = \left(K^m\right)^{-1, v}\left(b_t^m, s_t, t\right)$ for an agent with human capital $s_t$ (where $\left(K^m\right)^{-1, v}$ is the partial inverse of $K^m$ with respect to its first argument), so that the consumption choice of the agent and planner will coincide if mapped appropriately, i.e., $\hat{c}_t\left(K^m\left(v_t^m, s_t, t\right), s_t\right) = c_t^*\left(v_t^m, s_t\right)$. Furthermore, $\hat{b}_t\left(K^m\left(v_t^m, s_t, t\right), s_t\right) = K\left(v_t\left(v_t^m, s_t\right), s_t, t+1\right)$. To see this, suppose instead that there was another pair $\left(\tilde{c}, \tilde{b}\right) \neq \left(c^*, K\right)$ such that $\tilde{c} + \frac{1}{R}\tilde{b} = b_t^m$, but which yields higher utility: $u_t\left(\tilde{c}\right) + \beta V_{t+1}\left(\tilde{b}, s_t\right) > v_t^{m*} = \left(K^m\right)^{-1, v}\left(b_t^m, s_t, t\right)$ Note that the choice of $s_t$, already made in the previous stage is now fixed. Under the assumption that $V_{t+1}\left(., s_t\right)$ is increasing and continuous in its first argument, and given that $u_t\left(c\right)$ is increasing and continuous in $c$, there is also a pair $\left(\tilde{c}', \tilde{b}'\right)$ such that $\tilde{c} \leq c^*$, $\tilde{b} \leq K$ with one or both of these inequalities strict and such that $u_t\left(\tilde{c}\right) + \beta V_{t+1}\left(\tilde{b}, s_t\right) = v_t^{m*}$. But then $\left(\tilde{c}', \tilde{b}'\right)$ is better than $\left(c^*, K\right)$ in the planner's problem and hence, $\left(c^*, K\right)$ could not have been optimal, a contradiction.

For the first stage, consider an agent with initial wealth and human capital levels $b_{t-1}$ and $s_{t-1}$. First, map the allocations from the social planner's problem to allocations defined on the state space $\left(b_{t-1}, s_{t-1}, \theta_t\right)$:

$$y_t^*\left(b_{t-1}, s_{t-1}, \theta_t\right) = y_t^*\left(K^{-1, v}\left(b_{t-1}, s_{t-1}, t\right), s_{t-1}, \theta_t\right)$$

$$e_t^*\left(b_{t-1}, s_{t-1}, \theta_t\right) = e_t^*\left(K^{-1, v}\left(b_{t-1}, s_{t-1}, t\right), s_{t-1}, \theta_t\right)$$

$$b_t^{*m}\left(b_{t-1}, s_{t-1}, \theta_t\right) = K_t^m\left(v_t^{m*}\left(v_{t-1}, s_{t-1}, \theta_t\right), s_{t-1} + e_t^*\left(v_{t-1}, s_{t-1}, \theta_t\right), t\right)$$

167

Then, set the tax level such that for all $y, e \in Q_{e,y}^t (b_{t-1}, s_{t-1})$,

$$T_t (b_{t-1}, s_{t-1}, y_t^* (b_{t-1}, s_{t-1}, \theta_t), e_t^* (b_{t-1}, s_{t-1}, \theta_t))$$

$$= y_t^* (b_{t-1}, s_{t-1}, \theta_t) - b_t^{*m} (b_{t-1}, s_{t-1}, \theta_t) + b_{t-1} - M_t (e_t^* (b_{t-1}, s_{t-1}, \theta_t))$$

(that is, make all allocations consistent with an allocation in the planner's problem just affordable).

To extend the definition of the tax function to the full domain of allocations, even those which would not arise in the social planner's problem, three steps are taken. First, to rule out wealth levels $b_{t-1}$ not observed along the equilibrium path, i.e., such that

$$b_{t-1} \neq K \left( v_{t-1}^* \left( \theta^{t-1} \right), s_{t-1}^* \left( \theta^{t-1} \right), t \right)$$

for any $\theta^{t-1}$, set the borrowing limits to be $\underline{b_{t-1}} = \min_{v,s} K(v, s, t)$ where the min is taken over the possible values of $v$ and $s$ at time $t$ in the planner's program. Second, if $b_{t-1} = K \left( v_{t-1}^* \left( \tilde{\theta}^{t-1} \right), s_{t-1}^* \left( \tilde{\theta}^{t-1} \right), t \right)$ for some $\tilde{\theta}^{t-1}$ but $s_{t-1} \neq s_{t-1}^* \left( \tilde{\theta}^{t-1} \right)$ (that is, the levels of wealth and human capital from the past period are not mutually consistent), set the tax $T_t$ such that, for all $e_t, y_t$:

$$T_t (b_{t-1}, s_{t-1}, e_t, y_t) = y_t + \max \{ b_{t-1}, 0 \} - \min \{ \underline{b_t}, 0 \}$$

Finally, if the agent is on the equilibrium path with $b_{t-1}$ and $s_{t-1}$, and letting $v_{t-1} = K^{-1,v} (b_{t-1}, s_{t-1}, t)$, set $T_t$ such that for all pairs $y_t, e_t \notin Q_{e,y}^t (b_{t-1}, s_{t-1})$, and all $\theta_t$

$$-\phi_t \left( \frac{y_t}{w_t (\theta_t, s_{t-1} + e_t)} \right) + \beta V_t^m (b_{t-1} + y_t - T_t (b_{t-1}, s_{t-1}, y_t, e_t) - M_t (e_t), s_{t-1} + e_t)$$

$$< -\phi_t \left( \frac{y_t^* (v_{t-1}, s_{t-1}, \theta_t)}{w_t (\theta_t, s_{t-1} + e_t^* (v_{t-1}, s_{t-1}, \theta_t))} \right) + v_t^{m*} (v_{t-1}, s_{t-1}, \theta_t)$$

In the first stage, given the dissuasive taxes on choices which never arise in the planner's problem, the agent can either choose the full allocation ($y_t$ and $e_t$) destined for some type $\theta_t$ (that could be his own true type), which will then lead him to choose also the continuation wealth optimal for that same type, or he could choose $y_t$ optimal for some type $\theta_t^1$ but $e_t$

optimal for some type $\theta_t^2$. This will however leave him with lower value given the taxes on the off-equilibrium paths. Hence, if a type deviates, he must deviate to the full allocation of another type. By the temporal incentive compatibility constraint, he will choose not to do so.

### 2.A.3 Unobservable Human Capital

**Proof of Proposition (26) :**

A recursive formulation of the problem with unobservable human capital is:

$$K\left(v,\Delta,\Delta^s,\theta_-,s_-,t\right) = \min \int \left[ \begin{array}{c} c^a\left(\theta\right) - w_t\left(\theta,s\left(\theta\right)\right)l\left(\theta\right) \\ +\frac{1}{R}K\left(v\left(\theta\right),\Delta\left(\theta\right),\Delta^s\left(\theta\right),\theta,s\left(\theta\right),t+1\right) \end{array} \right] f^t\left(\theta|\theta_-\right)d\theta \quad (2.61)$$

$$\dot{\omega}\left(\theta\right) = \frac{w_{\theta,t}}{w_t}l\left(\theta\right)\phi_t'\left(l\left(\theta\right)\right) + \beta\Delta\left(\theta\right)$$

$$\omega\left(\theta\right) = u_t\left(c^a\left(\theta\right) - M_t\left(s\left(\theta\right) - s_-\right)\right) - \phi_t\left(l\left(\theta\right)\right) + \beta v\left(\theta\right)$$

$$v = \int \omega\left(\theta\right)f^t\left(\theta|\theta_-\right)d\theta$$

$$\Delta = \int \omega\left(\theta\right)\frac{\partial f^t\left(\theta|\theta_-\right)}{\partial\theta_-}\left(\theta|\theta_-\right)d\theta$$

$$M_t'\left(s\left(\theta\right) - s_-\right)u_t'\left(c^a\left(\theta\right) - M_t\left(s\left(\theta\right) - s_-\right)\right) = \frac{w_{s,t}}{w_t}l\left(\theta\right)\phi_t'\left(l\left(\theta\right)\right) - \beta\Delta^s\left(\theta\right)$$

$$\Delta^s = -\int \left(\frac{w_{s,t}}{w_t}l\left(\theta\right)\phi_t'\left(l\left(\theta\right)\right) - \beta\Delta^s\left(\theta\right)\right)f^t\left(\theta|\theta_-\right)d\theta$$

where the maximization is over $\left(c^a\left(\theta\right),l\left(\theta\right),s\left(\theta\right),\omega\left(\theta\right),v\left(\theta\right),\Delta\left(\theta\right),\Delta^s\left(\theta\right)\right)$, with $\Delta^s\left(\theta\right)$ as defined in $(2.45)$.

Let $\phi\left(l\right)$ be the disutility of labor, $\phi'\left(l\right)$ and $\phi''\left(l\right)$ its first and second order partials. The function $\tilde{c}^a\left(l,\omega - \beta v,s,s_-,\theta\right)$ defines assigned resources as a function of labor $l$, current period utility (rewritten again as: $\tilde{u} = \omega - \beta v$), current and beginning of period human capital levels $s$ and $s_-$, and the current realization of the type. Then, with the definition $c\left(\theta\right) = c^a\left(\theta\right) - M_t\left(s\left(\theta\right) - s_-\right)$, the constraint $\omega\left(\theta\right) = u_t\left(c\left(\theta\right)\right) - \phi_t\left(l\left(\theta\right),z\left(\theta\right) - z_-\right) + \beta v\left(\theta\right)$ becomes redundant, and the choice variables are $\left(l\left(\theta\right),s\left(\theta\right),\omega\left(\theta\right),v\left(\theta\right),\Delta\left(\theta\right),\Delta^s\left(\theta\right)\right)$. Let the multipliers on the constraints in program (2.61) be (in the order of the constraints): $\mu\left(\theta\right)$,

$\lambda_-$, $\gamma_-$, $\gamma^E(\theta)$, $\gamma^S_-$. The corresponding Hamiltonian is:

$$\left( \tilde{c}^a\left(l(\theta),\omega(\theta)-\beta v(\theta),s(\theta),s_-,\theta\right) - w_t(\theta,s(\theta))\,l(\theta) + \frac{1}{R}K\left(v(\theta),\Delta(\theta),\Delta^s(\theta),\theta,s(\theta),t+1\right) \right) f^t(\theta|\theta_-)$$

$$+\lambda_-\left[v-\omega(\theta)f^t(\theta|\theta_-)\right] + \gamma_-\left[\Delta-\omega(\theta)\frac{\partial f^t(\theta|\theta_-)}{\partial\theta_-}\right] + \mu(\theta)\left[\frac{w_{\theta,t}}{w_t}l(\theta)\phi_t'(l(\theta)) + \beta\Delta(\theta)\right]$$

$$-\gamma^S_-\left[\left(\frac{w_{s,t}}{w_t}l(\theta)\phi'(l(\theta)) - \beta\Delta^s(\theta)\right)f^t(\theta|\theta_-) + \Delta^s\right]$$

$$-\gamma^E(\theta)\left[M_t'(s(\theta)-s_-)u_t' - \frac{w_{s,t}}{w_t}l(\theta)\phi'(l(\theta)) + \beta\Delta^s(\theta)\right]$$

with the boundary conditions:

$$\lim_{\theta\to\bar{\theta}}\mu(\theta) = \lim_{\theta\to\underline{\theta}}\mu(\theta) = 0$$

From the FOC of the agents, along the optimum, and all other choice variables of the planner held constant:

$$\frac{d\tilde{c}^a}{dl} = \frac{\phi'(l)}{u'(c)} = w(1-\tau_L), \qquad \frac{d\tilde{c}^a}{ds_t} = M_t'(s_t-s_{t-1}), \qquad \frac{d\tilde{c}^a}{d\omega} = \frac{1}{u_t'(c)}$$

Using the three envelope conditions: $\frac{\partial K(v(\theta),\Delta(\theta),\Delta^s(\theta),\theta,s(\theta),t+1)}{\partial v(\theta)} = \lambda(\theta)$, $\frac{\partial K(v(\theta),\Delta(\theta),\Delta^s(\theta),\theta,s(\theta),t+1)}{\partial\Delta(\theta)} = \gamma(\theta)$, and $\frac{\partial K(v(\theta),\Delta(\theta),\Delta^s(\theta),\theta,s(\theta),t+1)}{\partial\Delta^s(\theta)} = -\gamma^s(\theta)$, the FOCs for $\omega(\theta)$, $v(\theta)$, $\Delta(\theta)$, and $\Delta^s(\theta)$ can be rewritten respectively as:

$$[\omega(\theta)]: -\frac{1}{u_t'(c)}f^t(\theta|\theta_-) + \frac{\gamma^E(\theta)}{f^t(\theta|\theta_-)}M_t'(e_t)\frac{u_t''(c)}{u_t'(c)}f^t(\theta|\theta_-) + \lambda\,f^t(\theta|\theta_-) + \gamma_-\frac{\partial f^t(\theta|\theta_-)}{\partial\theta_-} = \dot{\mu}(\theta)$$

$$[v(\theta)]\quad:\quad \frac{1}{u_t'(c_t)} = \frac{1}{\beta R}\lambda(\theta) + \frac{\gamma^E(\theta)}{f^t(\theta|\theta_-)}\frac{u_t''}{u_t'}M_t'(e_t) \tag{2.62}$$

$$[\Delta(\theta)]\quad:\quad -\frac{\gamma(\theta)}{R\beta} = \frac{\mu(\theta)}{f^t(\theta|\theta_-)}$$

$$[\Delta^s(\theta)]\quad:\quad -\beta\frac{\gamma^E(\theta)}{f^t(\theta|\theta_-)} + \beta\gamma^S_- = \frac{1}{R}\gamma^S(\theta)$$

Let $\tilde{\gamma}^E(\theta) = \frac{\gamma^E(\theta)}{f^t(\theta|\theta_-)}$ and $\theta_{-2}$ again denote the shock two periods back. Hence, the multiplier $\mu(\theta)$ solves:

$$\mu\left(\theta\right) = \int_{0}^{\bar{\theta}} \left( \frac{1}{u_t'\left(c\right)} \left(1 - \tilde{\gamma}^E\left(\theta_s\right) M_t'\left(e\left(\theta_s\right)\right) u_t''\left(c\right)\right) - \lambda_- + \frac{R\beta}{f\left(\theta_-|\theta_{-2}\right)} \mu_- \frac{\partial f^t\left(\theta_s|\theta_-\right)}{\partial \theta_-} \frac{1}{f\left(\theta_s|\theta_-\right)} \right) f\left(\theta_s|\theta_-\right) d\theta_s$$

Using the boundary conditions:

$$\lambda_- = \int_{\underline{\theta}}^{\bar{\theta}} \left( \frac{1}{u_t'\left(c_t\right)} \left(1 - \tilde{\gamma}^E\left(\theta_m\right) M_t'\left(e\left(\theta_m\right)\right) u_t''\left(c_t\right)\right) f\left(\theta_m|\theta_-\right) d\theta_m \right)$$

This equation, together with equation (2.62), lagged by one period, shows that the standard Inverse Euler no longer holds and is instead as in formula (2.49). The FOCs for labor and human capital are (with $\theta'$ and $s'$ being the next period's type and human capital).

$$[l\left(\theta\right)] : \frac{\tau_L\left(\theta\right)}{1 - \tau_L\left(\theta\right)} = \frac{\mu\left(\theta\right)}{f^t\left(\theta|\theta_-\right)} \frac{w_{\theta,t}}{w_t} u'\left(c_t\right) \frac{1 + \varepsilon^u}{\varepsilon^c} + u'\left(c_t\right) \left(\frac{w_{s,t}}{w_t}\right) \left(\frac{\gamma^E\left(\theta\right)}{f^t\left(\theta|\theta_-\right)} - \gamma_-^S\right) \frac{1 + \varepsilon^u}{\varepsilon^c} - \frac{\gamma^E\left(\theta\right)}{f^t\left(\theta|\theta_-\right)} M_t' u_t''$$

$$[s\left(\theta\right)] \quad : \quad \left(-M_t'\left(s\left(\theta\right) - s_-\right) + w_{s,t} l\left(\theta\right)\right) f^t\left(\theta|\theta_-\right) + \frac{1}{R} \int \left(M_{t+1}'\left(s'\left(\theta'\right) - s\left(\theta\right)\right)\right) f^{t+1}\left(\theta'|\theta\right) d\theta' f^t\left(\theta|\theta_-\right)$$

$$+\mu\left(\theta\right) \frac{l\left(\theta\right) \phi_t'\left(l\left(\theta\right)\right)}{w_t} \frac{w_{\theta,t} w_{s,t}}{w_t} \left(1 - \rho_{\theta s,t}\right) + \left(\gamma^E\left(\theta\right) - f^t\left(\theta|\theta_-\right) \gamma_-^S\right) l\left(\theta\right) \phi'\left(l\left(\theta\right)\right) \left(\frac{w_{s,t}}{w_t}\right)^2 \left(1 - \rho_{ss,t}\right)$$

$$+\gamma^E\left(\theta\right) M_t''\left(s\left(\theta\right) - s_-\right) u_t' - \frac{1}{R} \int \gamma^E\left(\theta'\right) M_{t+1}''\left(s'\left(\theta'\right) - s\left(\theta\right)\right) u_{t+1}'\left(c\left(\theta'\right)\right) f^{t+1}\left(\theta'|\theta\right) d\theta' f^t\left(\theta|\theta_-\right)$$

$$= \quad 0$$

where $\rho_{ss,t} = \frac{w_{ss,t} w_t}{w_{s,t}}$. Using the definition of the net wedge in (2.19) with $\tau_{St}$ set identically to 0 (because of the agent's Euler equation), the FOC for $s\left(\theta\right)$ allows to derive:

$$\left(\frac{\gamma^E\left(\theta\right)}{f^t\left(\theta|\theta_-\right)} - \gamma_-^S\right) = \frac{t_{st} w_t}{u'\left(c_t\right) w_{s,t}\left(1 - \rho_{ss,t}\right)} - \frac{\mu\left(\theta\right)}{f^t\left(\theta|\theta_-\right)} \frac{w_{\theta,t}}{w_{s,t}} \frac{\left(1 - \rho_{\theta s,t}\right)}{\left(1 - \rho_{ss,t}\right)} + \frac{w_t \left(\frac{1}{R} E\left(\tilde{\gamma}_{t+1}^E M_{t+1}'' u_{t+1}'\right) - \tilde{\gamma}_t^E M_t'' u_t'\right)}{u_t' w_{s,t}\left(1 - \rho_{ss,t}\right) w_{s,t} l_t\left(1 - \tau_{Lt}\right)}$$

Using this expression into the FOC for $l\left(\theta\right)$ yields the optimal tax formula.

**Proof of Proposition (27):**

If training time is unobservable, but expenses are observable, another endogenous state variable is introduced, analogous to $\Delta^s$ in the unobservable $s_t$ case:

$$\Delta^z\left(\theta\right) = -E_t\left(\phi_{z,t+1}\right)$$

As human capital expenses are observable, the agent's consumption is again directly controlled by the planner. Define the expenditure function indirectly as a function of the other choice variables: $\tilde{c}_t\left(l\left(\theta\right),\omega\left(\theta\right)-\beta v\left(\theta\right),z\left(\theta\right),z_-,\theta\right)$. The recursive program is (with multipliers in brackets after each corresponding constraint):

$$K\left(v,\Delta,\Delta^z,\theta_-,s_-,z_-,t\right) \tag{2.63}$$

$$= \min_{\{l,\omega,v,s,z,\Delta,\Delta^z\}} \int \left[ \begin{array}{c} \tilde{c}_t\left(l\left(\theta\right),\omega\left(\theta\right)-\beta v\left(\theta\right),z\left(\theta\right),z_-,\theta\right)+M\left(s\left(\theta\right)-s_-\right) \\ -w\left(\theta,s\left(\theta\right),z\left(\theta\right)\right)l\left(\theta\right)+\frac{1}{R}K\left(v\left(\theta\right),\Delta\left(\theta\right),\Delta^z\left(\theta\right),\theta,s\left(\theta\right),z\left(\theta\right),t+1\right) \end{array} \right] f^t\left(\theta|\theta_-\right)d\theta$$

$$\omega\left(\theta\right)=u_t\left(c\left(\theta\right)\right)-\phi_t\left(l\left(\theta\right),z\left(\theta\right)-z_-\right)+\beta v\left(\theta\right)$$

$$\dot{\omega}\left(\theta\right)=\frac{w_{\theta,t}}{w_t}l\left(\theta\right)\phi_{l,t}\left(l\left(\theta\right),z\left(\theta\right)-z_-\right)+\beta\Delta\left(\theta\right) \qquad \left[\mu\left(\theta\right)\right]$$

$$v=\int\omega\left(\theta\right)f^t\left(\theta|\theta_-\right)d\theta \qquad \left[\lambda_-\right]$$

$$\Delta=\int\omega\left(\theta\right)\frac{\partial f^t\left(\theta|\theta_-\right)}{\partial\theta_-}d\theta \qquad \left[\gamma_-\right]$$

$$\Delta^z=-\int\left(\frac{w_{z,t}l\left(\theta\right)}{w_t}\phi_{l,t}-\beta\Delta^z\left(\theta\right)\right)f^t\left(\theta|\theta_-\right)d\theta \qquad \left[\gamma_-^Z\right]$$

$$\phi_{z,t}=w_{z,t}l\left(\theta\right)\frac{1}{w_t}\phi_{l,t}-\beta\Delta^z\left(\theta\right)......\left[\gamma^E\left(\theta\right)\right]$$

With the Hamiltonian:

$$\left(\tilde{c}\left(l\left(\theta\right),\omega\left(\theta\right)-\beta v\left(\theta\right),z\left(\theta\right),z_-,\theta\right)-w_t\left(\theta,s\left(\theta\right),z\left(\theta\right)\right)l\left(\theta\right)\right)f^t\left(\theta|\theta_-\right)$$

$$+\frac{1}{R}K\left(v\left(\theta\right),\Delta\left(\theta\right),\Delta^z\left(\theta\right),\theta,s\left(\theta\right),z\left(\theta\right),t+1\right)f^t\left(\theta|\theta_-\right)$$

$$+\lambda_-\left[v-\omega\left(\theta\right)f^t\left(\theta|\theta_-\right)\right]+\gamma_-\left[\Delta-\omega\left(\theta\right)\frac{\partial f^t\left(\theta|\theta_-\right)}{\partial\theta_-}\right]+\mu\left(\theta\right)\left[\frac{w_{\theta,t}}{w_t}l\left(\theta\right)\phi_{l,t}+\beta\Delta\left(\theta\right)\right]$$

$$-\gamma^E\left(\theta\right)\left[w_{z,t}l\left(\theta\right)\frac{1}{w_t}\phi_{l,t}-\beta\Delta^z\left(\theta\right)-\phi_{z,t}\right]-\gamma_-^Z\left[\left(w_{z,t}l\left(\theta\right)\frac{1}{w_t}\phi_{l,t}-\beta\Delta^z\left(\theta^l\right)\right)f^t\left(\theta|\theta_-\right)+\Delta^z\right]$$

Taking the FOC with respect to $z_t$, $l_t$ and $s_t$ (without yet making assumption 6) so as to

have the general formulas: (recall that $\theta'$ and $s', z'$ denote the subsequent period's choices):

$$
[z(\theta)] \quad : \quad -\frac{\phi_{z,t}}{u_t'(c(\theta))} + w_{z,t}l(\theta) - \frac{\gamma^E(\theta)}{f^t(\theta|\theta_-)}\phi_{zz,t}
$$

$$
-\frac{\mu(\theta)}{f^t(\theta|\theta_-)}\left(l(\theta)\,\phi_{l,t}\frac{1}{w_t}w_{\theta z,t} - l(\theta)\,\phi_{l,t}w_{\theta,t}w_{z,t}\frac{1}{w_t^2} + \frac{w_{\theta,t}}{w_t}l(\theta)\,\phi_{lz,t}\right)
$$

$$
+\frac{1}{R}\int\left(\frac{\phi_{z,t+1}}{u_{t+1}'(c(\theta'))}\right)f^{t+1}(\theta'|\theta) + \frac{1}{R}\int\gamma^E(\theta')\,\phi_{zz,t+1}
$$

$$
+\frac{1}{R}\int\left(\frac{\mu(\theta')}{f^{t+1}(\theta'|\theta)}\frac{w_{\theta,t+1}}{w_{t+1}}l(\theta')\,\phi_{lz,t+1}\right)f^{t+1}(\theta'|\theta)
$$

$$
-\frac{1}{R}\int\left(\gamma^Z(\theta) + \frac{\gamma^E(\theta')}{f^{t+1}(\theta'|\theta)}\right)\left(w_{z,t+1}l(\theta')\frac{1}{w_{t+1}}\phi_{lz,t+1}\right)f^{t+1}(\theta'|\theta)
$$

$$
= \quad -\left(\gamma_-^Z + \frac{\gamma^E(\theta)}{f^t(\theta|\theta_-)}\right)\left(w_{zz,t}l(\theta)\frac{1}{w_t}\phi_{l,t} - (w_{z,t})^2 l(\theta)\frac{1}{w_t^2}\phi_{l,t} + w_{z,t}l(\theta)\frac{1}{w_t}\phi_{lz,t}\right)
$$

$$
[l(\theta)] : \left(\frac{\phi_{l,t}}{u_t'(c(\theta))} - w_t\right) + \frac{\mu(\theta)}{f^t(\theta|\theta_-)}\frac{w_{\theta,t}}{w_t}\phi_{l,t}\frac{1+\varepsilon_t^u}{\varepsilon_t^c} - \left(\gamma_-^Z + \frac{\gamma^E(\theta)}{f^t(\theta|\theta_-)}\right)\frac{w_{z,t}}{w_t}\frac{1+\varepsilon_t^u}{\varepsilon_t^c}\phi_{l,t} + \frac{\gamma^E(\theta)}{f^t(\theta|\theta_-)}\phi_{zl,t} = 0
$$

$$
[s(\theta)] \quad : \quad M_t'\left(s(\theta) - s_-\right) - l(\theta)\,w_{s,t} - \frac{1}{R}\int M'\left(s'(\theta') - s(\theta)\right)f(\theta'|\theta)\,d\theta =
$$

$$
\frac{\mu(\theta)}{f^t(\theta|\theta_-)}l(\theta)\,\phi_{l,t}\frac{1}{w_t^2}w_{\theta,t}w_{s,t}\left(1 - \rho_{\theta s,t}\right) + \left(\gamma_-^Z(\theta_-) + \frac{\gamma^E(\theta)}{f^t(\theta|\theta_-)}\right)w_{z,t}w_{s,t}\frac{1}{w_t^2}l(\theta)\,\phi_{l,t}\left(\rho_{zs,t} - 1\right)
$$

The FOC with respect to $\Delta^z$ yields:

$$
[\Delta^z(\theta)] : \frac{1}{\beta R}\gamma^Z(\theta) = \gamma_-^Z + \frac{\gamma^E(\theta)}{f^t(\theta|\theta_-)} \tag{2.64}
$$

Suppose now that assumption 6 holds. From the definition of the wedge $\tau_{Zt}$ and the net wedge $t_{zt}$, as well as the formula for the optimal $t_{zt}^*$ in (2.32), and the relation in (2.64), the FOC for $z_t$ allows to write:

$$
\left(\gamma_{t-1}^Z + \frac{\gamma_t^E(\theta_t)}{f^t(\theta_t|\theta_{t-1})}\right) = -\frac{[-t_{zt} + t_{zt}^*]}{u'(c_t)\frac{w_{z,t}}{w_t}\left(\rho_{zz,t} - 1\right)} + \frac{\left[\frac{\gamma_t^E(\theta_t)}{f^t(\theta_t|\theta_{t-1})}\phi_{zz,t} - \frac{1}{R}\int\gamma_{t+1}^E(\theta_{t+1})\,\phi_{zz,t+1}\right]}{w_{z,t}l_t u_t'(c_t)\left(1 - \tau_{Lt}\right)\frac{w_{z,t}}{w_t}\left(\rho_{zz,t} - 1\right)} \tag{2.65}
$$

Combining the FOCs for $l$ and $s$ yields:

$$[l_t] : \left( \frac{\tau_{Lt}}{1 - \tau_{Lt}} - \frac{\tau_{Lt}^*}{1 - \tau_{Lt}^*} \right) \frac{w_t}{w_{z,t}} \frac{\varepsilon_t^c}{1 + \varepsilon_t^u} = - \left( \gamma_{t-1}^Z f^t \left( \theta_t | \theta_{t-1} \right) + \gamma_t^E \left( \theta_t \right) \right) = \frac{\left( t_{st} - t_{st}^* \right)}{u' \left( c_t \right) \left( \rho_{zs,t} - 1 \right)} \frac{w_t}{w_{z,t}}$$

or, canceling terms, we obtain the expression in proposition 27. For completeness, the full expressions for $\tau_{Lt}$ and $t_{st}$ are:

$$\left( t_{st} - t_{st}^* \right) = \frac{\left( \rho_{zs,t} - 1 \right)}{\left( \rho_{zz,t} - 1 \right)} \left( t_{zt}^* - t_{zt} \right) + \frac{\left( \rho_{zs,t} - 1 \right)}{\left( \rho_{zz,t} - 1 \right)} \frac{\left[ \frac{\gamma_t^E \left( \theta_t \right)}{f^t \left( \theta_t | \theta_{t-1} \right)} \phi_{zz,t} - \frac{1}{R} \int \gamma_{t+1}^E \left( \theta_{t+1} \right) \phi_{zz,t+1} \right]}{w_{z,t} l_t \left( 1 - \tau_{Lt} \right)}$$

$$\left[ \frac{\tau_{Lt}}{\left( 1 - \tau_{Lt} \right)} - \frac{\tau_{Lt}^*}{\left( 1 - \tau_{Lt}^* \right)} \right] = \frac{w_t}{w_{z,t}} \frac{1 + \varepsilon_t^u}{\varepsilon_t^c} \frac{1}{\left( \rho_{zz,t} - 1 \right)} \left( t_{zt}^* - t_{zt} \right)$$

$$- \frac{1 + \varepsilon_t^u}{\varepsilon_t^c} \frac{1}{\left( \rho_{zz,t} - 1 \right)} \frac{\left[ \frac{\gamma_t^E \left( \theta_t \right)}{f^t \left( \theta_t | \theta_{t-1} \right)} \phi_{zz,t} - \frac{1}{R} \int \gamma_{t+1}^E \left( \theta_{t+1} \right) \phi_{zz,t+1} \right]}{w_{z,t} l_t \left( 1 - \tau_{Lt} \right)}$$

# Chapter 3

# Optimal Income, Education and Bequest Taxes in an Intergenerational Model

## 3.1 Introduction

Investing in the education of their children is a key concern for many parents. Parents often make it a priority to ensure that their offspring can attend the best schools that they can afford, and to save for college. From primary school to college, education expenses can be a major financial burden on families. Yet, parents can also transfer resources to their children through financial bequests. There is a three-way interaction between education subsidies, bequest taxes, and income taxes. Income taxes confiscate some of the returns to education for children, but also redistribute resources towards low income parents, which facilitates their investments in their offspring's education. Bequest taxes affect the choice between transferring resources through education purchases or through financial bequests. Investments in education in turn directly impact the available income and bequest tax bases and are major determinants of the pre-tax income distribution. Finally, bequests affect the incentives to work and, hence, the revenues from income taxes.

These interactions naturally lead to thinking about the optimal balance between income

taxes, education subsidies, and bequest taxes. Yet, the vast majority of optimal tax research assumes that education choices are exogenous to tax policy and that bequest and education decisions are delinked. Therefore, this chapter addresses the following questions. First and foremost, can we express the optimal formulas for education subsidies in terms of estimable statistics? These statistics should ideally be robust to heterogeneity in preferences and primitives. It is known that income tax formulas can typically be reformulated in terms of behavioral elasticities that capture the efficiency costs of taxation and distributional parameters that capture the redistributive value of taxation (Piketty and Saez 2013a). It would be valuable to be able to do the same for education subsidies without having to resort to strong assumptions about the underlying primitives and heterogeneities of agents. Second, how should the income tax account for education investments and bequests? Should parents' education expenses for their children be tax deductible? A strong result by Bovenberg and Jacobs (2005) states that income taxes and education subsidies are "Siamese Twins," i.e., they should be set equal to each other. In a static model, this corresponds to full deductibility of education expenses. Does this still hold in a generalized dynamic model with risk? At the very least, do income taxes and education subsidies comove? Finally, once we step into the realm of dynamic intergenerational models, the burden of estimating the relevant elasticities to tax policy increases. Hence, are there alternative specifications of the optimal formulas that can specifically be targeted to estimates that are easier to obtain? This leads us to reformulate optimal policies using "reform-specific" elasticities, as already introduced by Piketty and Saez (2013b).

Specifically, this chapter jointly determines the optimal linear income tax, bequest tax, and education subsidy in a dynamic intergenerational model. People live for one period during which they have children, supply labor, and then die. The setup is as in the standard dynastic Barro-Becker model: each generation cares about the expected discounted utility of all future generations. Parents can choose to purchase education for their children and also to leave them financial bequests. Education investments are risky and are subject to idiosyncratic shocks. The wage of each individual is a function of the endogenous education investment and a stochastic "ability." As in the standard Mirrlees (1971) income taxation model, ability is a comprehensive measure of the exogenous component of productivity and can capture, for instance, labor market or health shocks. People can also have idiosyncratic preference shocks that affect their tastes

for work and consumption. The government's goal is to maximize the expected utility of all dynasties as of the first generation. I briefly consider a simplified one generation model before moving to the intergenerational model.

The rationale to limit the set of tools to linear taxes and subsidies is to obtain clear formulas despite heterogeneous productivity shocks and preferences. In Chapter 2, I focused on the unrestricted optimal dynamic mechanism, the implementation of which requires history dependent, nonlinear tools.

The dynamic intergenerational model in this chapter highlights the different forces at play. First, even in a static version of the model, the Siamese Twins result for income taxes and education subsidies does not always hold and it is not optimal to make education expenses fully tax deductible,.unless a very special condition is satisfied. Namely the relative marginal revenue effect and the relative distributional effect of education subsidies and income taxes should be equal. Even more, because of the fully general wage and utility functions, it is not even the case that the education subsidy and the income tax should comove. In the dynamic model, bequest taxes can comove positively or negatively with education subsidies depending on whether bequests and education transfers are substitutes or complements for parents overall. A crucial determinant of education subsidies and bequest taxes are their distributional incidences, i.e., how concentrated education expenses and bequests are among high marginal utility agents. The optimal formulas can equivalently be derived in terms of reform-specific elasticities that capture the same trade-off between the efficiency and distributional impact of each instrument.

### 3.1.1 Related Literature

This chapter is related to the large literature on human capital, starting with Becker (1964), Ben-Porath (1967), and Heckman (1976a,b). A growing number of papers document the importance of human capital as a determinant of earnings (Card, 1995, Goldin and Katz, 2008, Acemoglu and Autor, 2011), and the financial and other factors shaping individuals' decisions to acquire education (Lochner and Monge-Naranjo, 2011, Altonji, Blom and Meghir, 2012, Avery et al., 2013).

It is also related to the optimal income taxation literature since Mirrlees (1971), its linear counterpart (Sheshinski, 1972), and most of all to the elasticities approach proposed by Saez

(2001) and explained in Piketty and Saez (2013a).

A long-standing question has been the optimal trade-off between labor and physical capital taxation in dynamic models (Atkinson and Sandmo, 1980). A stark finding has been that capital taxation is infinitely distortive in the long-run and, hence, steady-state capital taxes are optimally zero (Chamley, 1986, Judd, 1985). Piketty and Saez (2013b) shift the focus from physical capital to bequest taxation in an intergenerational model, and highlight that with uncertainty and distributional concerns, the optimal bequest tax is not zero. Farhi and Werning (2010) study nonlinear bequest and income taxation in a dynamic Mirrleesian framework, and find that a progressive bequest subsidy is optimal. Instead, with more general altruistic preferences, both taxes and subsidies could be optimal (Farhi and Werning, 2013). In this chapter, I shift the focus to an additional way of transferring resources to the next generation, namely education investments by parents for their children.

Several studies have considered the optimal treatment of physical versus human capital in dynamic macro models with representative agents (Jones, Manuelli and Rossi, 1993, Judd, 1999). The difference to the current chapter is that I allow for stochastic heterogeneity in both preferences and earnings ability.

A series of papers have considered optimal taxation jointly with education policies. The benchmark result that will be one of the focuses of this chapter is by Bovenberg and Jacobs (2005). Using a static taxation model, the authors find that education subsidies and income taxes are "Siamese Twins," and should always be set equal to each other, which is equivalent to making education expenses fully tax deductible. Benabou (2002) jointly analyses taxes and education in a dynastic Ramsey model. Krueger and Ludwig (2013) build an overlapping generations general equilibrium model, in which "education" is a binary decision that occurs exclusively before entry into the labor market. The main difference to these papers is, again, that I do not need to impose restrictions on the underlying heterogeneity or preferences in order to obtain estimable formulas for the optimal policies.

Several recent dynamic optimal tax papers look at the impact of taxation on human capital in the spirit of the New Dynamic Public Finance approach that considers general, unrestricted mechanisms (Bohacek and Kapicka, 2008, Kapicka, 2013a, Anderberg, 2009, Grochulski and Piskorski, 2010, Findeisen and Sachs, 2013, Kapicka and Neira, 2013, and Chapter 2 of this

178

Thesis). By contrast, this chapter tries to analyze restricted, linear instruments, that are functions of estimable elasticities and distributional characteristics.

The rest of the chapter is organized as follows. Section 3.2 presents the dynamic intergenerational model and the government's tools. Section 3.3 first solves for the optimal policies in a benchmark one generation model with no bequests. Section 3.4 solves for the optimal linear policies in the full dynastic model, using both standard and reform-specific elasticities. Section 3.5 concludes.

## 3.2 Intergenerational Model of Human Capital Investment

### 3.2.1 Preferences and Dynastic Utility

This section starts with the setup of an intergenerational model of human capital investment. The economy consists of agents who live for one period. Agents are born, have one single child each, and then die. Total population is hence constant and normalized to 1, so that average per capita and aggregate variables are the same. Denote a generic agent from generation $t$ and dynasty $i$ by $ti$.

Parents are the ones who purchase education for their children. Parents in generation $t$ can buy education $s_{t+1}$ for their child of generation $t+1$ at a linear cost $s_{t+1}$. In turn, generation $t$ also receives the human capital that their own parents from generation $t-1$ purchased for them. Human capital completely depreciates between generations. The first generation of dynasty $i$ at time 1 has an exogenously given distribution of human capital $s_{1i}$. This setup mirrors the fact that most investments in human capital occur before and during college and that parents account for a large share of these expenses. Indeed, parents typically bear the full burden of the bill for primary and secondary schooling. Beyond this, parents covered around 40% of total college expenses in 2012 – down from 50% before the financial crisis (Sallie Mae, 2012). Around 60% of students receive some help from their parents for college (Haider and McGarry, 2012). In their study of education choices in the kibbutzim, Abramitzky and Lavy (2013) show that investments in schooling by parents are quite responsive to redistributive policies.

Agents receive their human capital from their parents before they start working and consuming. The wage rate $w_{ti}$ of any agent is determined by his stock of human capital and his

stochastic ability $\theta_{ti}$:

$$w_{ti}(s) \equiv w(s, \theta_{ti})$$

$w$ is strictly increasing in education and ability and concave in education. Ability $\theta_{ti}$ is drawn from a stationary, ergodic distribution that allows for correlation between generations. There are several possible interpretations for $\theta$. For example, with a separable wage form $w_{ti}(s) = \theta_{ti} + h(s)$, for some function $h$, $\theta$ resembles a stochastic version of productivity from the static Mirrlees (1971) model. With a wage such as $w_{ti}(s) = \theta_{ti}h(s)$, $\theta$ is perhaps more naturally interpreted as the stochastic return to human capital. To keep with the tradition in the literature, $\theta$ will be called "ability" throughout. Ability to earn income can be stochastic for several reasons, among which are health shocks, individual labor market idiosyncrasies or luck.

If an agent works $l_{ti} \geq 0$ hours at a wage rate $w_{ti}$, he earns gross income $y_{ti} = w_{ti}l_{ti}$. His utility is given by:

$$u_{ti}(c, y, s) \equiv u\left(c, \frac{y}{w(s, \theta_{ti})}; \eta_{ti}\right)$$

where $\eta_{ti}$ is an idiosyncratic preference shock. $u$ is increasing in consumption $c$, decreasing in output $y$, and increasing in human capital $s$.

In addition to financing their education, parents can also leave financial bequests to their children. Bequests left by generation $t$ are denoted by $b_{t+1i}$ and earn a generational gross rate of interest $R$. Thus, generation $t$ inherits a pre-tax bequest of $Rb_{ti}$ from their parents. The initial generation 1 has an exogenously given distribution of bequests $b_{1i}$.

The expected utility of a dynasty $i$, $U_{1i}$ as perceived from period 1 and discounted by a generational discount factor $\beta$, is given by:

$$U_{1i} = E_1\left(\sum_{t=1}^{\infty} \beta^{t-1} u_{ti}(c_{ti}, y_{ti}, s_{ti})\right) \tag{3.1}$$

where $E_t$ is the expectation with respect to the realizations of $\theta$ and $\eta$ for all $i$ as of time $t$. Or, recursively:

$$U_{ti} = u_{ti} + \beta U_{t+1i}$$

180

### 3.2.2 Taxes and Budget Constraints

Government policies consist of a linear labor income tax $\tau_{Lt}$, a linear human capital subsidy $\tau_{St}$, a linear tax on the capitalized bequests $\tau_{Bt}$, and a lump-sum demogrant $G$. Hence, the budget constraint of person $ti$ is:

$$c_{ti} + b_{t+1i} + (1 - \tau_{St}) s_{t+1i} = Rb_{ti} (1 - \tau_{Bt}) + w_{ti} (s_{ti}) l_{ti} (1 - \tau_{Lt}) + G_t$$

One important simplification here is that, because each generation only lives for one period, the parental bequests and education investments deliver their income at the same moment in life of the beneficiaries (their children). In a more realistic overlapping generations model with multiple periods of life, this would not be the case; bequests will in general be received later in the beneficiary's life than education. This would not matter if there were perfect certainty and no credit constraints. Parents have consistent preferences with their children and can make the optimal trade-off between education and bequests without delay. However, when there is uncertainty in the child's ability and earning potential, being able to decide on bequests later in life, when the earning potential has been revealed, might make a difference. Indeed, parents may then try to target bequests to children who have received low ability and low income shocks so as to smooth consumption and provide insurance across generations.

### 3.2.3 Equilibrium

Aggregate consumption, human capital investments, bequests received, and output are denoted respectively by $c_t$, $s_{t+1}$, $b_t$ and $y_t$ for generation $t$. I assume that the stochastic processes for $\theta$ and $\eta$ are ergodic, so that, at constant policies (i.e., constant linear tax rates and demogrant), there is a unique ergodic steady-state equilibrium which is independent of the initial distribution of bequests and human capital (see also Piketty and Saez, 2013b). Hence, if tax policies $(\tau_{Lt}, \tau_{St}, \tau_{Bt}, G_t)$ converge to constant levels $(\tau_L, \tau_S, \tau_B, G)$ in the long run, then human capital $s_{t+1}$, output $y_t$, and bequests $b_t$ also converge to their steady state levels and depend on the steady state tax policies.

It is assumed that the government sets the policies so as to satisfy a period-by-period budget

constraint:[1]

$$G_t = \tau_{Lt} y_t + \tau_B R b_t - \tau_{St} s_{t+1} \tag{3.2}$$

## 3.3 Optimal Static Taxes and Subsidies

Suppose first that there are no intergenerational considerations, that each person invests in their own human capital, and that the distribution of shocks $\eta$ and $\theta$ is iid over time. This leads to a sequence of identical static problems. In this case, preferences are simply given by $U_i = u_i (c_i, y_i, s_i)$ and the budget constraint is:

$$c_i + (1 - \tau_S) s_i = w_i (s_i) l_i (1 - \tau_L) + G$$

Social welfare is a weighted sum of individual utilities, with $g_i$ the weight on individual $i$:

$$SWF = \int_i g_i u_i (c_i, y_i, s_i) \, di$$

The government needs to satisfy his single period budget constraint:

$$G = \tau_L y - \tau_S s$$

where $y$ and $s$ are aggregate output and aggregate human capital.

**The optimal labor tax in the presence of human capital:** Suppose that the government sets the optimal income tax rate $\tau_L$, taking the education subsidy as given. Denote the individual elasticity of output by $\varepsilon_{yi} \equiv d\log(y_i) / d\log(1 - \tau_L)$, the individual elasticity of education to the income tax by $\varepsilon_{si}^y \equiv d\log(s_i) / d\log(1 - \tau_L)$ and the aggregate weighted elasticities of, respectively, output and human capital by:

$$\varepsilon_Y \equiv \frac{d\log y}{d\log(1 - \tau_L)} = \int_i \varepsilon_{yi} \frac{y_i}{y} di$$

---

[1]Piketty and Saez (2013b) show that, if debt were allowed and the government was optimizing the economy-wide capital accumulation, then a modified golden rule would hold, with $\beta R = 1$ and the optimal formulas would be unaffected. Despite the fact that their model does not contain human capital, their result carries over unchanged.

$$\varepsilon_S^Y \equiv \frac{d\log s}{d\log(1-\tau_L)} = \int_i \varepsilon_{si}^y \frac{s_i}{s} di$$

Using the individuals' optimization and the envelope condition, a straightforward maximization implies that the change in welfare from a change in the linear tax rate $d\tau_L$ is given by:

$$\left(\int_i g_i u_{c,i} di\right)\left(y + \frac{\tau_L}{1-\tau_L}\int_i \varepsilon_{yi} y_i - \frac{\tau_S}{1-\tau_L}\int_i \varepsilon_{si}^y s_i\right) d\tau_L - \int_i g_i u_{c,i} y_i di d\tau_L = 0$$

Define the distributional characteristic of income to be

$$\bar{y} \equiv \frac{\int_i g_i u_{c,i} y_i di}{y \int_i g_i u_{c,i} di}$$

Hence, $\bar{y}$ is higher if income is correlated positively with the marginal utility of income.

**Proposition 28** *The optimal static income tax at any given human capital subsidy is:*

$$\tau_L^* = \frac{1 - \bar{y} - \tau_S \frac{s}{y} \varepsilon_S^Y}{1 - \bar{y} - \varepsilon_Y}$$

The expression for $\tau_L^*$ captures the typical trade-off between redistribution as measured by $(1 - \bar{y})$ and efficiency as measured by $\varepsilon_Y$. In addition, there is a term involving the education subsidy in the numerator. Indeed, if education choices respond to the income tax, there is a type of fiscal externality to the education subsidy base. Suppose that education choices respond negatively to the income tax ($\varepsilon_S^Y > 0$) and that there is a positive education subsidy in place. Then, the income tax is naturally lower than if education choices were insensitive to income taxes and the more so the higher the subsidy to education is. This is because a higher education subsidy is a measure of how much incentives are provided for education and it is costly to counteract them with the distortive income tax. If the education subsidy was zero, there would be no such fiscal externality, and the income tax would be set according to the standard formula. Note however, in the presence of human capital, the elasticity of income $\varepsilon_Y$ is a composite of the elasticity of labor supply and the elasticity of the wage (driven by the elasticity of human capital) to the income tax. Nevertheless, all that matters is the full elasticity of taxable income, observed in the data.

**The optimal education subsidy at any given income tax:** We can symmetrically

derive the optimal education subsidy. Denote the individual elasticity of human capital by $\varepsilon_{si} \equiv d\log(s_i)/d\log(\tau_S - 1)$, the individual elasticity of income to the education subsidy by $\varepsilon_{yi}^s \equiv d\log(y_i)/d\log(\tau_S - 1)$ and the aggregate weighted elasticities by:

$$\varepsilon_S \equiv \frac{d\log s}{d\log(\tau_S - 1)} = \int_i \varepsilon_{si} \frac{s_i}{s} di$$

$$\varepsilon_Y^S \equiv \frac{d\log y}{d\log(\tau_S - 1)} = \int_i \varepsilon_{yi}^s \frac{y_i}{y} di$$

Note that for a subsidy $\tau_S < 1$, $\varepsilon_{si} < 0$ since $ds_i/d\tau_s > 0$. Similarly to the definition of $\bar{y}$ above, define the distributional characteristic of education to be:

$$\bar{s} \equiv \frac{\int_i g_i u_{c,i} s_i di}{s \int_i g_i u_{c,i} di}$$

The higher $\bar{s}$ and the more education is concentrated among high social welfare weight agents. In the standard utilitarian framework, $g_i u_{c,i} = u_{c,i}$ is just the marginal utility of income, so that higher consumption agents have lower social marginal welfare weights. Hence, if education is concentrated among high income agents, then $\bar{s}$ is small. However, as Saez and Stantcheva (2013) show for the optimal income tax rate, the same formula derived below will hold if one replaces $g_i u_{c,i}$ by generalized social welfare weights.

Again, the change in welfare from a change in the linear education subsidy $d\tau_S$ is given by:

$$\left( \int_i g_i u_{c,i} di \right) \left( -s - \frac{\tau_S}{\tau_S - 1} \int_i \varepsilon_{si} s_i + \tau_L \int_i \varepsilon_{yi}^s y_i \right) d\tau_S + \int_i g_i u_{c,i} s_i di d\tau_S = 0$$

**Proposition 29** *The static optimal human capital subsidy for a given labor tax $\tau_L$ is given by:*

$$\tau_S^* = \frac{1 - \bar{s} + \frac{y}{s} \varepsilon_Y^S \tau_L}{1 - \bar{s} + \varepsilon_S} \tag{3.3}$$

This formula looks very similar to a standard income tax formula, despite education being a productive input and not a final consumption good. Hence, there is no "production efficiency" result a la Diamond and Mirrlees (1971). In general, the optimal subsidy will not be zero because of the redistributive effect of education $1 - \bar{s}$ and a finite elasticity $\varepsilon_S$. The income tax appears in the numerator because of the fiscal spillover: if output responds positively to

education subsidies, then a higher education subsidy has an additional positive effect on revenues raised, which is stronger the higher the income tax rate is.[2]

**Full Optimum: optimizing jointly the income tax and education subsidy.** At the full optimum, with both $\tau_L$ and $\tau_S$ optimally set, the labor and human capital taxes are given by:

$$\tau_S^* = \frac{(1-\bar{s})\left(1-\bar{y}-\varepsilon_Y\right) + \frac{y}{s}\varepsilon_Y^S\left(1-\bar{y}\right)}{\left[(1-\bar{s}+\varepsilon_S)\left(1-\bar{y}-\varepsilon_Y\right) + \varepsilon_Y^S\varepsilon_S^Y\right]}$$

$$\tau_L^* = \frac{(1-\bar{y})\left(1-\bar{s}+\varepsilon_S\right) - \frac{s}{y}\varepsilon_S^Y\left(1-\bar{s}\right)}{\left[(1-\bar{s}+\varepsilon_S)\left(1-\bar{y}-\varepsilon_Y\right) + \varepsilon_Y^S\varepsilon_S^Y\right]}$$

The full optimum is discussed next in relation to Bas and Bovenberg's "Siamese Twins" result, stating that the linear income tax and the linear education subsidies should be set equal to each other.

### 3.3.1 Should Education Expenses be Fully Tax Deductible?

A natural benchmark is the full deductibility of education expenses, i.e., $\tau_S^* = \tau_L^*$. Bas and Bovenberg find that full deductibility of human capital expenses is optimal with a special form of the earnings function that guarantees that all agents benefit equally at the margin, in proportional terms, from human capital investments. In this generalized setup, where both the wage and the utility function are unrestricted, we can infer a similar result but based on the observed elasticities and redistributive effects:

**Corollary 7** *The education subsidy should optimally be set equal to the income tax rate if and only if:*

$$\frac{\left(\frac{y}{s}\varepsilon_Y^S - \varepsilon_S\right)}{\left(\frac{s}{y}\varepsilon_S^Y - \varepsilon_Y\right)} = \frac{(1-\bar{s})}{(1-\bar{y})} \tag{3.4}$$

*for $1-\bar{s} \neq 0$ and $1-\bar{y} \neq 0$.[3]*

The left hand side in expression (3.4) is the ratio of the efficiency cost of the education subsidy and the efficiency cost of the income tax. The right hand side, is the ratio of their

---

[2] Note that because $\tau_S$ is defined as a subsidy and because the elasticity is defined with respect to $\tau_S - 1$, which is negative for $\tau_S < 1$, the denominator $1-\bar{s}+\varepsilon_S$ is typically negative. If output responds positively to education subsidies, then $\varepsilon_Y^S < 0$.

[3] *If $1 = \bar{s}$ and $1 = \bar{y}$, then $\tau_S^* = \tau_L^* = 0$, a degenerate case.*

redistributive effects. The optimal human capital subsidy should be set equal to the optimal income tax if and only if the relative efficiency cost is equal to the relative redistributive effect of human capital versus output. On the other hand, if the redistributive effect of human capital is disproportionately large relative to its efficiency cost, then it will be optimal to set $\tau_S^* > \tau_L^*$ and to subsidize education expenses beyond just making them tax deductible.

It is easy to check that the Bovenberg and Jacobs (2005) setting with a multiplicatively separable wage $w = \theta s$, isoelastic, separable, and quasilinear utility $u_i(c, y, s) = c - \frac{1}{\gamma}\left(\frac{y}{\theta_i s}\right)^\gamma$ implies that for any welfare weights, $\bar{y} = \bar{s}$, $\varepsilon_Y^S = \gamma$, $\varepsilon_Y = 1 - \gamma$, $\varepsilon_S^Y = -\gamma$, $\varepsilon_S = \gamma - 1$, so that the equality (3.4) indeed holds.

In Chapter 2, it was shown that the relative redistributive effects of education and output are determined by how complementary ability $\theta$ and human capital are in the wage function. If the marginal wage benefit of human capital is proportionately higher for higher ability agents, then $(1 - \bar{s})$ will be large relative to $(1 - \bar{y})$ and $\tau_S^* < \tau_L^*$, so that education expenses are only partially tax deductible. In addition, the redistributive effects can also be driven by the idiosyncratic preference shock $\eta$.

It is not true more generally that even a relaxed version of the Siamese Twins result must hold. I.e., it is not necessarily the case that the labor tax and the education subsidy should comove positively. From expression (3.3), we see that if output is increasing in human capital (so that $\varepsilon_Y^S < 0$ for any $\tau_S < 1$), the human capital subsidy is increasing in the labor tax rate $\tau_L$. Indeed, a higher subsidy stimulates output and allows to raise more revenues from the income tax base the higher $\tau_L$ is. However, the effect of human capital on output is the result of two effects, namely, the effect of human capital on the wage and the effect of the wage on labor supply. Because the utility function is fully general, substitution and income effects of the wage rate on labor supply can occur in many different configurations. It is hence not guaranteed that $\tau_S$ and $\tau_L$ will always comove.

### 3.3.2 Unobservable Human Capital

It might be that at least part of education expenses are not perfectly observable by the government, or that their reporting can be manipulated by parents, or that there are political constraints on education subsidies. In either of these cases, human capital investments should

formally be treated as not observable and hence, $\tau_S \equiv 0$. The optimal income tax will have to be adjusted differently, so as to compensate for the lack of a tool to directly control human capital investments. The optimal linear tax $\tau_L^{*,u}$ with unobservable human capital is given by:

$$\tau_L^{*,u} = \frac{1 - \bar{y}}{1 - \bar{y} - \varepsilon_Y}$$

The difference to the observable human capital case in formula (3.3) is the lack of term $-\tau_S \frac{s}{y} \varepsilon_S^Y$ in the numerator. If the substitution effect dominates and education responds negatively to income taxes ($\varepsilon_S^Y < 0$), then if $\tau_S^*$ would have been positive were education observable, the optimal income tax needs to be set lower when human capital is not observable. This is intuitive: with observable human capital, the government would like to subsidize human capital. Yet, if the subsidy on human capital is constrained to be zero, incentives for human capital can only be indirectly provided through a lower income tax.

## 3.4 Optimal Linear Policies in a Dynastic Model

We now turn back to the intergenerational model presented in Section 3.2. With different generations, there are at least two possible criteria for a government setting optimal policies. The first is to maximize the expected welfare of today's generation. The second criterion would be to maximize the steady state utility of dynasties and is not pursued here. We are more interested in reforms that will be evaluated by agents living today, rather than in the steady state, and which take into account the transition to the new steady state.

### 3.4.1 A Dynamic Reform Approach

Social welfare is equal to the expected discounted utility of a dynasty at time 0 (i.e., before today's generation's uncertainty is realized).

$$SWF_0 = \max E_0 \sum_{t=1}^{\infty} \beta^{t-1} \left[ u_{ti} \left( (1 - \tau_{Lt}) y_{ti} - s_{t+1i} (1 - \tau_{St}) + R (1 - \tau_{Bt}) b_{ti} - b_{t+1i} + G_t, y_{ti}, s_{ti} \right) \right]$$

subject to

$$G_t = \tau_{Lt} y_t + \tau_{Bt} R b_t - \tau_{St} s_{t+1}$$

Consider a small reform $d\tau_{St} = d\tau_S$ that affects subsidy rates for all generations after some $T$. There is perfect foresight about the time and magnitude of the reform, and, hence, the dynasty will have anticipatory effects even before time $T$. At the optimal $\tau_S$, the change in social welfare from the reform must be zero. From the envelope theorem linked to the agents' first-order conditions $dSWF_0$ is:

$$
\begin{aligned}
dSWF_0 &= \sum_{t \geq T} \beta^{t-1} E\left(s_{t+1i} u_{c,ti}\right) d\tau_S - \sum_{t \geq T} \beta^{t-1} E\left(u_{c,ti}\right) s_{t+1} d\tau_S \\
&\quad + \sum_{t < T} \beta^{t-1} E\left(u_{c,ti}\right) \left(-\tau_{St} ds_{t+1} + \tau_{Lt} dy_t + \tau_{Bt} R db_t\right) \\
&\quad + \sum_{t \geq T} \beta^{t-1} E\left(u_{c,ti}\right) \left(-\tau_{St} ds_{t+1} + \tau_{Lt} dy_t + \tau_{Bt} R db_t\right)
\end{aligned}
$$

The first term is the direct welfare effect of the reform. By the envelope theorem, it is equal to the weighted reduction in consumption from the subsidy change. This is true even if parents make the choices instead of their children, as the dynastic setup ensures that the preferences of the next generation are perfectly internalized by the current generation. If parents did not put the same weight on the utility impact of education for their children as children themselves did, there would then be an additional welfare effect on the children from the parents' change in investment decisions. This would occur for instance if children's education directly entered parental utility: $u_{ti}\left(c_{ti}, y_{ti}, s_{ti}, s_{t+1i}\right)$. There would then be double counting of the effect of any education increase, both on parents and on children.

The second term is the mechanical revenue effect driven by the loss in tax revenue from the higher subsidy, at constant individual choices. These two effects only take place after the reform. The last two terms are the behavioral responses, before and after the reform respectively. The third term is the anticipatory effect.

Define the elasticities of aggregate human capital $s_{t+1}$, output $y_t$, and bequests $b_t$ to a small

change in the education subsidy $d\tau_S$ (for all $t > T$) to be:

$$\varepsilon_{St+1} = \frac{ds_{t+1}}{d(\tau_S - 1)}\frac{(\tau_{St} - 1)}{s_{t+1}}, \qquad \varepsilon_{Yt}^S = \frac{dy_t}{d(\tau_S - 1)}\frac{(\tau_{St} - 1)}{y_t}, \qquad \varepsilon_{Bt}^S = \frac{db_t}{d(\tau_S - 1)}\frac{(\tau_{St} - 1)}{b_t}$$

Because of the ergodic stationary steady state, for $t > T$ after the reform, each of these elasticities converges to the corresponding long-run elasticity when $t \to \infty$. Before the reform, the elasticities correspond to anticipatory elasticities. Even though generations can in principle start reacting a long time before the reform, the responses are attenuated the further away in the future the reform is. It is convenient, hence, to a first approximation, to assume that the anticipatory effects of the reform only start once the steady state paths of all variables have been reached (i.e., $T$ is large enough, so that aggregate variables have converged even before then).

On the steady state path, human capital, output, and bequests are constant. Hence, we can divide through by $E(u_{c,ti})s_{t+1}$, constant in the steady state, where the expectation is taken over all $i$:

$$dSWF_0 = \sum_{t \geq 1}^{\infty} \beta^{t-1}\frac{1}{s_{t+1}}\left(-\tau_{St}\varepsilon_{St+1}\frac{s_{t+1}}{\tau_S - 1} + \tau_{Lt}\varepsilon_{Yt}^S\frac{y_t}{(\tau_S - 1)} + \tau_{Bt}\varepsilon_{Bt}^S\frac{Rb_t}{(\tau_S - 1)}\right)$$
$$+ \sum_{t \geq T}\beta^{t-1}\frac{E(s_{t+1i}u_{c,ti})}{E(u_{c,ti})s_{t+1}} - \sum_{t \geq T}\beta^{t-1}$$

Let the elasticities $\varepsilon_S'$, $\varepsilon_Y^{S\prime}$ and $\varepsilon_B^{S\prime}$ be the long-run elasticities of the present discounted value of each corresponding tax base, i.e.:

$$\varepsilon_S' \equiv (1-\beta)\sum_{t \geq 1}\beta^{t-1-T}\varepsilon_{St+1}, \qquad \varepsilon_Y^{S\prime} \equiv (1-\beta)\sum_{t \geq 1}\beta^{t-1-T}\varepsilon_{Yt}^S, \qquad \varepsilon_B^{S\prime} \equiv (1-\beta)\sum_{t \geq 1}\beta^{t-1-T}\varepsilon_{Bt}^S$$

$$(3.5)$$

For instance, $\varepsilon_S'$ is the elasticity of the present discounted value of the education subsidy base with respect to a distant subsidy change. It is the sum of the discounted average of the standard post-reform elasticities and of the discounted average of the anticipatory elasticities. The same is true for the cross-elasticities of output and bequests to a distant subsidy change, $\varepsilon_Y^{S\prime}$ and $\varepsilon_B^{S\prime}$.

As before, define the redistributive values of output, human capital and bequests to be:

$$\bar{y} \equiv \frac{E\left(u_{c,ti}y_{ti}\right)}{E\left(u_{c,ti}\right)y_t}, \qquad \bar{s} \equiv \frac{E\left(u_{c,ti}s_{t+1i}\right)}{E\left(u_{c,ti}\right)s_{t+1}}, \qquad \bar{b} \equiv \frac{E\left(u_{c,ti}b_{ti}\right)}{E\left(u_{c,ti}\right)b_t} \qquad (3.6)$$

To reiterate, each of these factors measures the strength of the covariance between the corresponding variable and the marginal utility. The larger $\bar{y}$, $\bar{s}$, or $\bar{b}$ are, the more output, human capital, and bequests are concentrated among those with high marginal utilities of consumption.

If $\tau_S$ is optimally set, then it has to be that at $\tau_S$, the change in welfare $dSWF_0$ is zero. Suppose that $\tau_L$ and $\tau_B$ are just kept at their constant levels throughout. Then, we can rearrange the expression to obtain the optimal education subsidy $\tau_S^*$.

**Proposition 30** *The optimal human capital subsidy, for any $\tau_L$ and $\tau_B$, is given by:*

$$\tau_S^* = \frac{1 - \bar{s} + \varepsilon_Y^{S\prime}\tau_L\frac{y}{s} + \varepsilon_B^{S\prime}\tau_B R\frac{b}{e}}{1 - \bar{s} + \varepsilon_S'} \qquad (3.7)$$

*with $\bar{s}$ the distributional characteristic of education as defined in (3.6), and $\varepsilon_S'$, $\varepsilon_Y^{S\prime}$, and $\varepsilon_B^{S\prime}$, respectively, the long-run elasticities of the discounted human capital, income, and bequest tax bases as defined in (3.5).*

Several features of optimal formula (3.7) are worth noting. First, the typical inverse elasticity effect is apparent: the subsidy is smaller when $\varepsilon_S'$, which is negative for any $\tau_S < 1$, is larger. It is also possible that the optimal subsidy is actually a tax. This is most likely to occur if the distributional value of education $\bar{s}$ is small, i.e., mostly high consumption agents acquire education.

**Tax deductibility of education expenses:** In the dynamic model, full tax deductibility of education expenses from both the taxable income and the bequest tax bases would require that $\tau_S = \tau_L + \tau_B$. This would be akin to a dynamic version of the Siamese Twin result that also includes the bequest tax. It is clear that full deductibility is in general not optimal. A weaker result would be that education subsidies, income taxes, and bequest taxes comove positively with each other.

As in the static formula above, the revenue effect from the income tax appears in the numerator: if income is increasing in the human capital subsidy, the subsidy comoves with the

190

income tax rate and the more so if output is more elastic to the human capital subsidy, i.e., if $\varepsilon_Y^{S\prime}$ is larger in absolute value. However, the effect can go in the other direction. Because of income effects and the non-separable utility function, a higher human capital subsidy can also reduce the need for parents – and the children who receive a larger human capital stock from them – to work.

In the dynamic model, the same sort of effect also occurs with the bequest tax, again with an ambiguous direction. When bequest taxes are higher, the human capital subsidy should be reduced if a lower human capital subsidy encourages more bequests through a strong substitution effect ($\varepsilon_S^{B\prime} > 0$). However, a higher education subsidy also has an income effect that might increase bequests, which are a normal good. Depending on which effect dominates, both the income tax and the education subsidy on the one hand, and the bequest tax and the education subsidy on the other hand could be comoving or not.

**Evaluating reforms to education subsidies:** The aggregate elasticities and the distributional parameters are of course endogenous to the taxes and subsidies, and the formula is merely an implicit formula. However, it is especially useful for evaluating reforms around the current status quo. Indeed, the formula holds for any bequest and income taxes $\tau_B$ and $\tau_L$, so that the right-hand side could be evaluated at the current tax and subsidy levels. If the implied $\tau_S^*$ is above the current $\tau_S$, a reform that decreases $\tau_S$ would improve social welfare.

**The distributional characteristic of education matters:** The optimal subsidy is higher when those with high social welfare weights, that is, those with low consumption, also have high education expenses. The redistributive value of education is closely linked to how complementary education and ability are in the wage function, as explained in subsection 3.3.1. In addition, it can also be driven by idiosyncratic preferences for human capital and work, as captured by $\eta$. However, it will not only depend on the technological and preference primitives, but also on the institutional setup of the education system. Suppose that the government provides basic education for all children through a public school system, which is free, and, by extension, through public universities. If parents choose to acquire additional education for their children, they can do so at their own expense, for example in private schools or universities, or through private tutoring lessons. If the marginal subsidy applies to those additional

191

education expenses, it is likely that only high income agents incur them, so the approximation $\bar{s} = 0$ might be reasonable. On the other hand, if there is no free-of-charge public education system, subsidies on marginal education expenses would have to be larger, as even low income agents would need to purchase education.

**Generalized Social Welfare Weights**: Saez and Stantcheva (2013) show that it is possible to replace the standard social welfare weights, equal to $u_{c,ti}/E\left(u_{c,ti}\right)$ here, by generalized social welfare weights $g_{ti}$ that directly place a social marginal value on an additional dollar transferred to person $i$.[4] At one extreme, it could be that children who receive no education from their parents should be compensated. In the limit, $\bar{s} = 0$ and $\tau_S$ is set to the revenue-maximizing, Rawlsian subsidy rate:

$$\tau_S^{Rawls} = \frac{1 + \varepsilon_Y^{S\prime}\tau_L\frac{y}{e} + \varepsilon_B^{S\prime}\tau_B R_e^b}{1 + \varepsilon_S'}$$

On the other hand, it could be that society places a lot of value on parents who invest in their children's human capital. This would lead to a very large subsidy if $\bar{s} \gg 1$. Again, the reasonable social weight to place on people at different education levels depends on the institutional setting and on who exactly acquires education. If everyone expects to have a fair shot at receiving education from their parents and/or leaving education to their kids, then social preferences would naturally be such that $\bar{s} = 1$ and the optimal subsidy is only driven by efficiency considerations, trading off the differential revenue impacts of the education subsidy, income tax, and bequest tax:

$$\tau_S^{Efficiency} = \frac{\varepsilon_Y^{S\prime}\tau_L\frac{y}{s} + \varepsilon_S^{B\prime}\tau_B R_e^b}{\varepsilon_S'}$$

**Heterogeneous Altruism for Parents**: In this model, parents are the key agents that invest in children's education. But some parents might care more than others about their children and the future generations of their dynasty. This can be captured by heterogeneous generational discount rates, $\beta_{ti}$. In this case, formula (3.7) nevertheless applies, but replacing

---

[4]The small difficulty here is that one may weight differently parents who purchase a lot of education for their children and children who simply inherit a lot of education from their parents (akin to bequests).

$\bar{s}$ by:

$$\bar{s} = \lim_{T} \frac{\sum_{t \geq T} E\left(\beta_{1i}...\beta_{ti}u_{c,ti}s_{ti}\right)}{\sum_{t \geq T} E\left(\beta_{1i}...\beta_{ti}u_{c,ti}\right)s_t}$$

More altruistic parents have higher discount factors and, all else equal, invest more in the education of their children. Hence, social welfare becomes automatically concentrated on altruistic dynasties, which care about their children, and simultaneously invest more in education. This implies that $\bar{s} > 1$. This is quite unappealing as, eventually, society puts no weight on descendants from selfish parents. An alternative social criterion that does not suffer from this shortcoming would be the steady state welfare of the dynasty, instead of the welfare as of the first generation.

**The Optimal Income Tax:** The optimal linear income tax $\tau_L^*$ can be similarly derived for given human capital subsidy and bequest tax:

$$\tau_L^* = \frac{1 - \bar{y} + \varepsilon_S^{Y\prime}\frac{s}{y}\tau_S - \varepsilon_B^{Y\prime}\tau_B R\frac{b}{y}}{1 - \bar{y} + \varepsilon_Y'} \tag{3.8}$$

with $\varepsilon_S^{Y\prime}$, $\varepsilon_B^{Y\prime}$, and $\varepsilon_Y'$ the long-run elasticities of the discounted present value of, respectively $s$, $b$, and $y$ to a reform in $\tau_L$ taking place after time $T$. These elasticities are again the composites of the anticipatory elasticities and the post-reform elasticities. The typical trade-off between the redistributive and insurance benefit of taxation (in $\bar{y}$) and the efficiency cost of taxation (in $\varepsilon_Y'$) is present. The fiscal spillovers to the education subsidy base and the bequest tax base enter as well.

**The Optimal Bequest Tax:** Finally, the optimal bequest tax is, at given $\tau_S$ and $\tau_L$:

$$\tau_B^* = \frac{1 - \bar{b} + \varepsilon_S^{B\prime}\frac{s}{b}\tau_S - \varepsilon_Y^{B\prime}\tau_L\frac{y}{b}}{1 - \bar{b} + \varepsilon_B'}$$

The bequest tax is generically not zero. There are two different sets of reasons for this. First, if education subsidies and income taxes are already set, and the government needs to maximize the bequest tax holding the former fixed, there are spillover effects on revenue through the education subsidy and the income tax bases. Bequest taxes indirectly influence education purchase and

work decisions. Second, even in the absence of income taxes and education subsidies, the bequest tax is not zero as long as $\bar{b} \neq 1$. We would have $\bar{b} = 1$ if either utility were linear, or if everyone left the same bequest amount due to homogeneous, quasilinear preferences.

This result fits well with Piketty and Saez (2013b), except that they do not consider education as an alternative way for parents to transfer resources to their children. Farhi and Werning (2013) find that a wide range of positive or negative bequest taxes can be optimal depending on the social objective.

### 3.4.2 Reform Elasticities

There is another way to determine optimal tax and subsidy rates, which might prove to be more convenient in some situations. Indeed, the shortcoming of (3.7) is that it relies on the (endogenous) cross-elasticity of output or wealth to education subsidies, $\varepsilon_Y^{S\prime}$ and $\varepsilon_B^{S\prime}$, at a given $\tau_L$ and $\tau_B$. Sometimes, all that is observed in the data is the total response of any given variable to a reform. A full reform, however, is rarely just an isolated shift in one tax or subsidy rate, and often a combination of changes in several tax tools, with revenue-neutrality or not. In any country, there might have been specific reforms already implemented, which can serve as natural experiments to estimate elasticities. For each reform, one can derive the implied optimal education subsidy as a function of the "reform elasticities," i.e., the full responses observed during that reform. It is then not crucial to know what the cross-elasticities were. The analysis can be performed for different types of reforms, and is illustrated below for a change in education subsidies financed by an increase in income taxes.

Suppose again that the government maximizes the full dynasty's expected welfare subject to its per period budget constraint. Consider a small revenue-neutral reform $d\tau_{St} = d\tau_S$ for $t > T$, and a corresponding series of income tax reforms $d\tau_{Lt}$ to maintain budget balance, around constant $\tau_S$ and $\tau_L$. The bequest tax $\tau_B$ is left unchanged. $T$ is again large enough for all variables to have converged to their steady state paths. At an optimum, the change in social welfare from this reform must be zero. Using the envelope theorem from the agents' first-order conditions:

$$dSWF_0 = \sum_{t>T} \beta^{t-1} E\left(s_{t+1i} u_{c,ti}\right) d\tau_S - \sum_{t\geq 1} \beta^{t-1} E\left(y_{ti} u_{c,ti}\right) d\tau_{Lt} = 0$$

Define the elasticities in each period to be:

$$\varepsilon_{Bt} \equiv \frac{db_t}{d(\tau_S - 1)}\frac{\tau_{St} - 1}{b_t}|_G, \quad \varepsilon_{St+1} \equiv \frac{ds_{t+1}}{d(\tau_S - 1)}\frac{\tau_{St} - 1}{s_{t+1}}|_G, \quad \varepsilon_{Yt} \equiv \frac{dy_t}{d(1 - \tau_L)}\frac{1 - \tau_{Lt}}{y_t}|_G$$

where $db_t$, $ds_{t+1}$, and $dy_t$ are the responses of, respectively, aggregate savings, aggregate human capital, and aggregate output to the *full* reform $(d\tau_S, d\tau_{Lt})$ at constant revenue $G$. Note that these elasticities capture the joint total effects of the simultaneous changes in $\tau_S$ and $\tau_L$ on $b_t$, $s_{t+1}$ and $y_t$. Hence, they are composites, for each variable, of own-tax and cross-tax effects.[5] By contrast, the formula in (3.7) isolated the pure effect of the education subsidy on all other variables.

Exactly as above, due to anticipatory effects with forward-looking life-cycle agents, the reaction of these variables to the reform may start even before the reform period $T$. Assuming that $T$ is large enough, the anticipatory reactions only start once all three choice variables have reached their steady state paths.

For the budget to remain balanced in all periods, we need the income tax to adjust such that, for $t > T$:

$$\left(-\varepsilon_{St+1}\frac{\tau_S}{\tau_S - 1} + R\frac{b_t}{s_{t+1}}\varepsilon_{Bt}\frac{\tau_B}{\tau_S - 1} - 1\right)s_{t+1}d\tau_S = -\left(1 - \varepsilon_{Yt}\frac{\tau_L}{1 - \tau_L}\right)y_t d\tau_{Lt}$$

and for $t \le T$ :

$$\left(-\varepsilon_{St+1}\frac{\tau_S}{\tau_S - 1} + R\frac{b_t}{s_{t+1}}\varepsilon_{Bt}\frac{\tau_B}{\tau_S - 1}\right)s_{t+1}d\tau_S = -\left(1 - \varepsilon_{Yt}\frac{\tau_L}{1 - \tau_L}\right)y_t d\tau_{Lt}$$

Substituting for these tax changes in $dSWF_0$, dividing by $s_{t+1}E(u_{c,ti})$ (constant in the steady state) yields:

$$dSWF_0 = \bar{s}\frac{\beta^{T-1}}{1 - \beta} + \bar{y}\left(\sum_{t<T}\beta^{t-1}\frac{\left(-\varepsilon_{St+1}\frac{\tau_S}{\tau_S-1} + R\frac{b_t}{s_{t+1}}\varepsilon_{Bt}\frac{\tau_B}{\tau_S-1}\right)}{\left(1 - \varepsilon_{Yt}\frac{\tau_L}{1-\tau_L}\right)} + \sum_{t\ge T}\beta^{t-1}\frac{\left(-\varepsilon_{St+1}\frac{\tau_S}{(\tau_S-1)} + R\frac{b_t}{s_{t+1}}\varepsilon_{Bt}\frac{\tau_B}{\tau_S-1} - 1\right)}{\left(1 - \varepsilon_{Yt}\frac{\tau_L}{1-\tau_L}\right)}\right)$$

Let $\varepsilon'_S$ and $\varepsilon'_B$ be the total long-run response of the discounted human capital, bequest and

---

[5] In particular, the normalization by $\tau_{St} - 1$ of $\varepsilon_{Bt}$ is arbitrary. We could have normalized by $(1 - \tau_L)$ instead.

output bases to the reform, as in subsection 3.4.1. Define $\varepsilon'_Y$ as the composite elasticity of output that ensures that the following equality holds:

$$\frac{\left(-\varepsilon'_S \frac{\tau_S}{(\tau_S-1)} + R\frac{b}{e}\varepsilon'_B \frac{\tau_B}{\tau_S-1} - 1\right)}{\left(1 - \varepsilon'_Y \frac{\tau_L}{1-\tau_L}\right)}$$

$$= (1-\beta)\sum_{t<0}\beta^{t-T}\frac{\left(-\varepsilon_{St+1}\frac{\tau_S}{\tau_S-1} + R\frac{b_t}{s_{t+1}}\varepsilon_{Bt}\frac{\tau_B}{\tau_S-1}\right)}{\left(1 - \varepsilon_{Yt}\frac{\tau_L}{1-\tau_L}\right)} + (1-\beta)\sum_{t\geq T}\beta^{t}\frac{\left(-\varepsilon_{St+1}\frac{\tau_S}{\tau_S-1} + R\frac{b_t}{s_{t+1}}\varepsilon_{Bt}\frac{\tau_B}{\tau_S-1} - 1\right)}{\left(1 - \varepsilon_{Yt}\frac{\tau_L}{1-\tau_L}\right)}$$

Rearranging the expression for $dSWF_0 = 0$ at constant $\tau_S$ and $\tau_L$, yields the optimal human capital subsidy, formulated according to this specific reform.

**Proposition 31** *The optimal human capital subsidy that maximizes the expected welfare of the dynasty, for any $\tau_L$ and $\tau_B$, is given by:*

$$\tau_S^* = \frac{1 - \frac{\bar{s}}{\bar{y}}\left(1 - \varepsilon'_Y \frac{\tau_L}{1-\tau_L}\right) + R\frac{b}{s}\varepsilon'_B\tau_B}{1 - \frac{\bar{s}}{\bar{y}}\left(1 - \varepsilon'_Y \frac{\tau_L}{1-\tau_L}\right) + \varepsilon'_S} \tag{3.9}$$

*where the long-run elasticities $\varepsilon'_B$, $\varepsilon'_Y$ and $\varepsilon'_S$ are estimated from a revenue neutral reform that changes $\tau_S$ and adjusts $\tau_L$ to maintain budget balance and $\bar{s}$ and $\bar{y}$ are the distributional factors of human capital and income, as defined in (3.6).*

The elasticities in formula (3.9) are reform-specific, i.e., they measure the total impact on the aggregate variables $b_t$, $s_{t+1}$, and $y_t$ of changing $\tau_S$ while adjusting $\tau_L$ to maintain revenue-neutrality. Hence, this formulation is most useful when there have been past reforms resembling exactly this one, so that the elasticities can be estimated in the data. Conversely, an analog of formula (3.9) can easily be derived again for other reforms, with suitably redefined elasticities. The advantages of formulation (3.9) are that, if such suitable reforms already exist, the elasticities are readily available without having to separately estimate all cross-tax effects. Again, it is not necessary to assume that either $\tau_L$ or $\tau_B$ are optimally set in the economy.

The difference between formulas (3.7) and (3.9) is that, in the latter, the income tax is adjusted in step with the education subsidy to ensure budget balance. Hence, there is no direct revenue effect, but any change in the education subsidy is mirrored by a change in the

income tax. This leads to two changes. First, the redistributive effect $\bar{s}$ is replaced by the comprehensive redistributive effect $\frac{\bar{s}}{\bar{y}} \left(1 - \varepsilon'_Y \frac{\tau_L}{1-\tau_L}\right)$ that takes into account that although a human capital subsidy may mostly benefit high social value agents, it will also lead to a rise in the income tax that will partially offset this benefit. The relevant parameter is the relative distributional value $\bar{s}/\bar{y}$. Second, as both $\tau_L$ and $\tau_S$ are adjusted to maintain budget balance, there is no additional revenue effect $\varepsilon_Y^{S'}\tau_L \frac{y}{s}$ from the income tax base in the numerator as in (3.7).

It bears repeating that, with perfect estimation tools that would allow us to uncover all cross elasticities and all reform-specific elasticities, at any tax levels, in a world where reforms were cleanly isolated with no concurrent changes in other policies, (3.7) and (3.9) would yield the same answer. They are merely two ways to approach the same question, taking into account the empirical burden of estimating different elasticities.

### 3.4.3 Unobservable Human Capital

As in subsection 3.3.2, if human capital is unobservable, the optimal income tax will adjust, so as to compensate for the lack of a tool to directly control human capital investments. Supposing that bequests remain observable, the optimal linear tax is now:

$$\tau_L^{*,unobs} = \frac{1 - \bar{y} - \frac{b}{y}\varepsilon_B^{Y'}\tau_B}{1 - \bar{y} + \varepsilon'_Y}$$

Exactly as in the static case, the difference to the observable human capital case in formula (3.7) is the lack of term $-\tau_S \frac{s}{y}\varepsilon_S^{Y'}$ in the numerator, with $\varepsilon_S^{Y'}$ the long-run elasticity of the discounted education base to the retention rate. If the substitution effect dominates, so that the present discounted value of the education base responds negatively to income taxes ($\varepsilon_S^{Y'} < 0$), and if $\tau_S^*$ would have been positive were education observable, the optimal income tax needs to be set lower when human capital is not observable.

## 3.5 Conclusion

This chapter studies optimal dynamic linear income and bequest taxes and education subsidies in a dynastic intergenerational model. Parents invest in the education of their children and can

197

also leave financial bequests. Each generation is subject to idiosyncratic ability and preference shocks. The government aims to provide redistribution and insurance to maximize the expected discounted welfare of the dynastics, from the point of view of today's generation. I derive formulas for the optimal taxes and subsidies as functions of estimable behavioral elasticities and redistributive factors.

In a generalized model with risk, non-separable utility, and a general interaction between education and ability in the wage function, it is in general not optimal to make education expenses fully tax deductible. Instead, education subsidies, bequest taxes and income taxes can comove positively or negatively. Because of the distributional value of education and its finite elasticity to education subsidies, the optimal subsidy on education is generically non-zero, even if education is a production input.

The same model can be used to study more general investments by parents in the human capital of their children, notably health expenses. Finally, with non-complete depreciation of human capital between generations, the model can be reformulated as a lifecycle model of a single agent and address the issue of optimal linear lifecycle taxation and subsidies for human capital.

This theoretical research points to two important empirical explorations. First, how do parents value the educational achievements of their children? The dynastic setup is very stringent in the sense that all generations are in perfect agreement with each other. Yet, there may be imperfectly altruistic preferences in reality. Second, it has not yet been convincingly documented how strongly parents react to bequest taxation and education subsidies when choosing between alternative ways of transferring resources to their children.

# Bibliography

Abraham, A. and N. Pavoni. 2008. "Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending," *Review of Economic Dynamics* 11, 781-803.

Acemoglu, D. and D. Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings," in O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics,* Vol. 4.

Acemoglu, D. and J. Pischke 1998. "Why do Firms Train? Theory and Evidence", *Quarterly Journal of Economics,* 113(1), 79-119

Akerlof, G. 1976. "The Economics of Caste and of the Rat Race and Other Woeful Tales", *Quarterly Journal of Economics,* 90(4), 599-61

Albanesi, S. and C. Sleet 2006. "Dynamic Optimal Taxation with Private Information," *Review of Economic Studies,* 73(1), 1-30

Alesina, A., Glaeser, E., and B. Sacerdote 2005. "Work and Leisure in the United States and Europe: Why so Different?", *NBER Macroeconomics Annual 20,* M. Gertler and K. Rogoff, eds.

Altonji, J. and C. Paxson 1988. "Labor Supply Preferences, Hours Constraints, and Hours-Wage Trade-offs," *Journal of Labor Economics,* 6(2), 254-276

Altonji, J., E. Blom and C. Meghir. 2012. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annual Review of Economics,* Annual Reviews, vol. 4(1), 185-223.

Anderberg, D. 2009. "Optimal Policy and the Risk Properties of Human Capital Reconsidered," *Journal of Public Economics* 93(9-10), 1017-1026.

Autor, D.. 2001. "Why do Temporary Help Firms Provide Free General Skills Training?", *Quarterly Journal of Economics,* 116(4), 1409-1448

Angrist, J., and A. Krueger 1991. "Does compulsory school attendance affect schooling and earnings," *Quarterly Journal of Economics* 106, 979-1041.

Ashenfelter, O. and A. Krueger 1994. "Estimates of the economic return to schooling for a new sample of twins", *American Economic Review* 84:1157-1173.

Ashenfelter, O. and C. Rouse 1998. "Income, schooling, and ability: Evidence from a new sample of identical twins," *Quarterly Journal of Economics* 113, 253-284.

Atkeson, A. and R. Lucas 1995. "Efficiency and Equality in a Simple Model of Efficient Unemployment Insurance," *Journal of Economic Theory* 66, 64-88.

Atkinson, A. and J. Stiglitz 1976. "The Design of Tax Structure: Direct versus Indirect Taxation," *Journal of Public Economics* 6, 55-75.

Autor, D.H., L. Katz and A. Krueger. 1998. "Computing inequality: have computers changed the labor market?," *Quarterly Journal of Economics* 113, 1169-1213.

Avery, C., M. Glickman, C. Hoxby, and A. Metrick. 2013. "A Revealed Preference Ranking of American Colleges and Universities," *Quarterly Journal of Economics* 128(1).

Baker, G., R. Gibbons, and K.J. Murphy, 1994. "Subjective Performance Measures in Optimal Incentive Contracts", *Quarterly Journal of Economics*, 109, 1125-56

Baron, D. and D. Besanko. 1984. "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, 1(3), 267-302.

Battaglini, M. and S. Coate. 2008. "Pareto Efficient Income Taxation with Stochastic Abilities," *Journal of Public Economics* 92(3-4), 844-868.

Battaglini, M. and R. Lamba. 2013. "Optimal Dynamic Contracting: the First Order Approach and Beyond," *Working Paper, Princeton University.*

Becker, G. 1964. *A Theoretical and Empirical Analysis, with Special Reference to Education*, The University of Chicago Press.

Benabou, R. 2002. "Tax and Education Policy in a Heterogeneous Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica*, 70(2), 481-517.

Ben-Porath, Y. 1967. "The Production of Human Capital and the Life Cycle of Earnings", *Journal of Political Economy*, Vol. 75, No. 4, pp. 352-365.

Bierbrauer, F. and P. C. Boyer 2010. "The Pareto-Frontier in a simple Mirrleesian model of income taxation", *Preprints of the Max Planck Institute for Research on Collective Goods*

Blomquist, Soren. 1982. "Should educational expenses be deductible?," *Journal of Public Economics* 18, 65-83.

Bohacek, R. and M. Kapicka. 2008. "Optimal Human Capital Policies," *Journal of Monetary Economics* 55(1), 1-16.

Boskin, Michael. 1977. "Notes on the Tax Treatment of Human Capital," *Proceedings of the Treasury Conference on Tax Policy*, U.S. Department of Treasury.

Bovenberg, L. and B. Jacobs. 2005. "Redistribution and Education Subsidies are Siamese Twins," *Journal of Public Economics*, 89(11-12), 2005-2035.

Bovenberg, L. and B. Jacobs. 2011. "Optimal Taxation of Human Capital and the Earnings Function," *Journal of Public Economic Theory*, 13, (6), 957-971.

Brody, E. 1994. "Paying back your country through income-contingent student loans," *San Diego Law Review*, 31, 449-518.

Card, D., 1995a. "Earnings, schooling and ability revisited" in S. Polachek (ed.) *Research in Labor Economics*, Vol. 14, JAI Press, Greenwich, CT, pp. 23-48.

Card, D., 1995b. "Using geographic variation in college proximity to estimate the return to schooling," in L. N. Christofides, E. K. Grant and R. Swidinsky (eds.), *Aspects of labor market behavior: essays in honour of John Vanderkamp*, University of Toronto Press, Toronto. pp. 201-222.

Caucutt, E, and K. Kumar. 2003. "Higher Education Subsidies and Heterogeneity: A Dynamic Analysis," *Journal of Economic Dynamics and Control*, 27, 1459-1502.

Chamley, C. 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54(3), pp. 607-622.

Chetty, R., 2009. "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance," *American Economic Journal: Economic Policy 1*, 31-52

Chetty, R. and E. Saez 2010. "Optimal Taxation and Social Insurance with Endogenous Private Insurance", *American Economic Journal: Economic Policy 2*, 85-114

Chetty, R., J. Friedman, T. Olsen, and L. Pistaferri 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records", *Quarterly Journal of Economics*, 126(2), 749-804

Chetty, R., J. Friedman, N. Hilger, D. Schanzenbach, D. Yagan. 2011. "How Does Your Kindergarden Classroom Affect Your Earnings? Evidence from Project STAR," *Quarterly Journal of Economics* 126(4), 1593-1660.

Chetty, R. 2012. "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 80(3), 969-1018.

Cole, H.L., and N. Kocherlakota. 2001. "Efficient allocations with hidden income and hidden storage," *Review of Economic Studies*, 68(3), 523-542.

Crocker, K. and A. Snow, 1985. "The Efficiency of Competitive Equilibria in Insurance Markets with Asymmetric Information", *Journal of Public Economics*, 26, 207-219

Crocker, K. and A. Snow, 1986. "The Efficiency Effects of Categorical Discrimination in the Insurance Industry", *Journal of Political Economy*, 94(2), 321-344

Cunha, F, J. Heckman and S. Navarro. 2005. "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economics Papers*, 57(2), 191.

Cunha, F, and J. Heckman. 2007. "Identifying and Estimating the Distributions of Ex Post and Ex

Ante Returns to Schooling," *Labour Economics*, 14(6), 870-893.

Cunha, F, and J. Heckman. 2008. "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, 43(4), 738-782.

Cunha, F, J. Heckman, L. Lochner, and D. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation," in E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, Vol. 1, Chapter 12, 697-812.

Da Costa, C. E., and L. J. Maestri. 2007. "The risk Properties of Human Capital and the Design of Government Policies," *European Economic Review* 51(3), 695-713.

Diamond, P.1998. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review* 88(1), 83-95.

Diamond, P. and J. Mirrlees. 1971. "Optimal Taxation and Public Production I: Production Efficiency," *American Economic Review* 61(1), 8-27.

Dickens, W. and Shelly Lundberg 1993. "Hours Restrictions and Labor Supply,", *International Economic Review*, 34(1), 169-92

Doepke, M., and R. Townsend. 2006. "Dynamic Mechanism Design with Hidden Income and Hidden Actions", *Journal of Economic Theory* 126(1), 235-285.

Dynarski, S. 2003. "Does Aid Matter? Measuring the Effects of Student Aid on College Attendance and Completion," *American Economic Review* 93(1), 279-288.

Farhi, E. and I. Werning 2013. "Insurance and Taxation over the Life Cycle," *Review of Economic Studies* 80, 596-635

Federal Reserve Bank of New York. 2013. *Quarterly Report on Household Debt and Credit*, May 2013

Feldstein, M 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax,", *Review of Economics and Statistics*, 81, 674-680

Fernandes, A. and C. Phelan. 2000. "A Recursive Formulation for Repeated Agency with History Dependence," *Journal of Economic Theory* 91(2), 223-247.

Findeisen, S. and D. Sachs 2012. "Education and Optimal Dynamic Taxation: The Role of Income-contingent Student Loans," *Working Paper, University of Konstanz*.

Fryer R. 2011. "Financial Incentives and Student Achievement: Evidence from Randomized Trials," *Quarterly Journal of Economics*, 126, 1755-1798.

Geanakoplos, J. and H. M. Polemarchakis 2008. "Pareto improving taxes,", *Journal of Mathematical Economics*, 44, 682-696

Gibbons, R. and L. Katz, 1991. "Layoffs and Lemons", *Journal of Labor Economics*, 9(4), 351-380

Gibbons, R and K. Murphy, 1992. "Optimal Incentive Contracts in the Presence of Career Concerns:

Theory and Evidence", *Journal of Political Economy*, 100, 468-505

Giertz, S., 2004. "Recent Literature on Taxable Income Elasticities", *US Congressional Budget Office Technical Paper*, 2004-16

Goldin, C. and L. Katz. 2008. "The Race Between Education and Technology," *Harvard University Press, Cambridge MA.*

Golosov, M, A. Tsyvinski and M. Troshkin. 2013. "Redistribution and Social Insurance," *Working paper, Princeton University.*

Golosov, M, A. Tsyvinski and I. Werning. 2006. "New Dynamic Public Finance: A User's Guide," *NBER Macroeconomic Annual*, MIT Press.

Golosov, M, and A. Tsyvinski. 2007. "Optimal Taxation with Endogenous Insurance Markets," *Quarterly Journal of Economics*, 122(2), 487-534.

Greenwald, B. and J. Stiglitz 1986. "Externalities in Economies with Imperfect Information and Incomplete Markets", *The Quarterly Journal of Economics*, 101(2), 229-264

Grochulski, B, and T. Piskorski. 2010. "Risky Human Capital and Deferred Capital Income Taxation," *Journal of Economic Theory*, 145(3), 908-943.

Guesnerie, R 1998. "A Contribution to the Pure Theory of Taxation," *Cambridge University Press*

Hader, S, and K. McGarry. 2012. "Parental Investments in College and Later Cash Transfers," *NBER working paper 18485.*

Heathcote, J., F. Perri and G. Violante. 2010. "Unequal we Stand: An Empirical Analysis of Economic Inequality in the United States, 1967-2006," *Review of Economic Dynamics,* 13, 15-51.

Heathcote, J., K. Storesletten and G. Violante. 2005. "Two Views of Inequality Over the Life Cycle," *Journal of the European Economic Association,* 3, 765-775.

Heckman, James. 1976a. "Estimates of a Human Capital Production Function Embedded in a Life-Cycle Model of Labor Supply," in N. Terleckyj (ed.), *Household Production and Consumption*, NBER, p. 225-264.

Heckman, James. 1976b. "A Life-Cycle Model of Earnings, Learning, and Consumption," *Journal of Political Economy*, 84(4), S11-S44.

Heckman, J., L. Lochner, and P. Todd. 2005. "Earnings Functions, Rates of Return, and Treatment Effects: The Mincer Equation and Beyond," *NBER Working Paper 11544.*

Hicks, J. 1970. "Elasticity of Substitution Again: Substitutes and Complements," *Oxford Economic Papers*, 25, 289-296.

Holmstrom, B.1999. "Managerial Incentive Problems: A Dynamic Perspective", *Review of Economic Studies*, 66(1), 169-182.

Hoxby, C. 1998. "Tax Incentives for Higher Education," in J. Poterba (ed.) Tax Policy and the

Economy, *MIT Press, Cambridge, Mass,* 49-82.

Hoy, M.1984. "The Impact of Imperfectly Categorizing Risks on Income Inequality and Social Welfare", *The Canadian Journal of Economics,* 17(3), 557-568.

Hoy, M.1989. "The Value of Screening Mechanisms under alternative insurance possibilities", *Journal of Public Economics,* 39, 177-206.

Huggett, M., G. Ventura and A. Yaron. 2011. "Sources of Lifetime Inequality," *American Economic Review,* 101, 2923-54.

Huggett, M. and G. Kaplan. 2011. "Human Capital Values and Returns: Bounds Implied by Earnings and Asset Returns Data," *Journal of Economic Theory,* 146(3), 897-919.

Jones, L., Manuelli, R. and P. Rossi. 1993. "Optimal taxation in Models of Endogenous Growth," *Journal of Political Economy,* 101(3), 485-517.

Judd, K.L. 1985. "Redistributive taxation in a simple perfect foresight model," *Journal of Public Economics,* 28(1), pp. 59-83.

Kapicka, M. and J. Neira 2013. "Optimal Taxation in a Life-Cycle Economy with Endogenous Human Capital Formation," *Working Paper, University of California, Santa Barbara.*

Kapicka, M. 2013a. "Optimal Mirrleesean Taxation with Unobservable Human Capital Formation," *Working Paper, University of California, Santa Barbara.*

Kapicka, M. 2013b. "Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-Order Approach," *Review of Economic Studies,* 80(3), 1027-1054.

Kaplow, L. 1994. "Human Capital Under and Ideal Income Tax," *Virginia Law Review,* 80(7), 1477-1514.

Kaplow, L. 2006. "On the desirability of commodity taxation even when income taxation is not optimal," *Journal of Public Economics,* 90, 1235-1250.

Karaivanov, A. and R. Townsend 2013. "Dynamic Financial Constraints: Distinguishing Mechanism Design from Exogenously Incomplete Regimes," *Working Paper, MIT.*

Kocherlakota, N. 2004. "Figuring out the impact of hidden savings on optimal unemployment insurance," *Review of Economic Dynamics,* 7, 541-554.

Kocherlakota, N. 2005. "Zero Expected Wealth taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica,* 73(5), 1587-1621.

Kremer, M. 2002. "Should taxes be independent of age?" *mimeo.*

Kugler, A. and G. Saint-Paul 2004. "How do Firing Costs affect Worker Flows in a World with Adverse Selection", *Journal of Labor Economics,* 22(3), 553-584.

Krueger, D. and F. Perri 2010. "Public versus Private Risk Sharing", *Journal of Economic Theory,* 146, 920-956.

**Krueger, D., and A. Ludwig.** 2013. "Optimal Capital and Progressive Labor Income Taxation with Endogenous Schooling Decisions and Intergenerational Transfers," *University of Pennsylvania Working Paper.*

**Laffont, J-J. and D. Martimort** 2001. *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press

**Landers, R., J. Rebitzer, and L. Taylor** 1996. "The Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms", *American Economic Review*, 86(3), 329-348.

**Laroque, G.** 2005. "Indirect taxation is superfluous under separability and taste homogeneity: A simple proof," *Economic Letters*, 87, 141-144.

**Lindbeck, A.,** 1995. "Welfare State Disincentives with Endogenous Habits and Norms," *The Scandinavian Journal of Economics*, 97(4), 477-494.

**Lochner, L. and A. Monge-Naranjo.** 2011. "The Nature of Credit Constraints and Human Capital," *American Economic Review*, 101, 2487-2529.

**Maldonado, D.** 2007. "Education Policies and Optimal Taxation," *International Tax and Public Finance*, 15(2), 131-143.

**Milgrom, R. and I. Segal.** 2002. "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70(2), 583-601.

**Mirrlees, J.** 1971. "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 38, 175-208.

**Mirrlees, J., S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba** 2011. "Tax by Design: The Mirrlees Review," *Oxford University Press.*

**Miyazaki, H.** 1977. "The Rat Race and Internal Labor Markets", *The Bell Journal of Economics*, 8(2), 394-418.

**Murphy, K., and F. Welch.** 1992. "The structure of Wages," *Quarterly Journal of Economics*, 107, 285-326.

**Chang Wei, C.** 2010. "What is the Price of College?" *Stats in Brief*, National Center for Education Statistics, US Department of Education (NCES 2011-175).

**OECD.** 2004. "Improving Skills for More and Better Jobs: Does Training Make a Difference?," *OECD Employment Outlook.*

**OECD.** 2013. "Education at a Glance, 2013," *OECD, Paris.*

**Palacios, M.** 2004. *Investing in Human Capital: A Capital Markets Approach to Higher Education Funding*, Cambridge: Cambridge University Press.

**Pavan, A., I. Segal and J. Toikka.** 2013. "Dynamic Mechanism Design: A Myersonian Approach," forthcoming *Econometrica.*

Prescott, E. C., and R. Townsend, 1984. "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard", *Econometrica*, 52(1), 21-45.

Rothschild, M.and J. Stiglitz, 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", *The Quarterly Journal of Economics*, 90(4), 629-649.

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies* 68(1), 205-229.

Saez, E., J. Slemrod, and S. Giertz, 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review", *Journal of Economic Literature*, 50:1, 3-50.

Saez, E. and S. Stantcheva, 2013. "Generalized Social Marginal Welfare Weights for Optimal Tax Theory", *NBER Working Paper 18835.*.

Samuelson, P. 1973. "Complementarity – An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory", *Journal of Economic Literature*, 12, (4), 1255-1289.

Scheuer, F., 2013a. "Entrepreneurial Taxation with Endogenous Entry", *forthcoming, American Economic Journal: Economic Policy*.

Scheuer, F., 2013b. "Adverse Selection in Credit Markets and Regressive Profit Taxation", *Journal of Economic Theory* 148, 1333-1360.

Schmalensee, R., 1984. "Imperfect Information and the Equitability of Competitive Prices", *The Quarterly Journal of Economics*, 99(3), 441-460.

Schultz, T. 1961. "Investment in Human Capital," *American Economic Review*, 51.

Slemrod, J. B., 1998. "Methodological Issues in Measuring and Interpreting Taxable Income Elasticities," *National Tax Journal*, 51(4), 773-88.

Slemrod, J. B., and S. Yitzhaki, 2002. "Tax avoidance, evasion, and administration," in *Handbook of Public Economics*, Alan Auerbach and Martin Feldstein eds. (Amsterdam: North-Holland).

Spear, S., and S. Srivastava 1987. "On Repeated Moral Hazard with Discounting", *Review of Economic Studies*, 54(4), 599-617.

Spence, M., 1978. "Product Differentiation and Performance in Insurance Markets", *Journal of Public Economics*, 10, 427-447

Storesletten, K., Telmer, C., and A. Yaron. 2004. "Consumption and Risk Sharing over the Life Cycle," *Journal of Monetary Economics* 51, 609-633.

Weinzierl, M. 2011. "The Surprising Power of Age-Dependent Taxes," *Review of Economic Studies*, 78, 1490-1518.

Werning, Ivan 2002 "Optimal Unemployment Insurance with Unobservable Savings," *Working Paper MIT*.

Werning, Ivan 2007a "Nonlinear Capital Taxation," *Working Paper MIT*.

**Werning, Ivan** 2007b "Optimal Fiscal Policy with Redistribution," *Quarterly Journal of Economics*, 122, 925-967.

**Werning, Ivan** 2007c "Pareto Efficient Income Taxation," *Working Paper MIT.*

**Wilson, C.,** 1977. "A Model of Insurance markets with Incomplete Information", *Journal of Economic Theory*, 12, 167-207