Optimal Policy and the Coexistence of Markets and Governments

by

Jenny Simon

Submitted to the Department of Economics
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2011

ARCHIVES
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
NOV 2 2 2011
LIBRARIES

Author .....
Department of Economics
August 15, 2011

Certified by .....
Daron Acemoglu
Elizabeth and James Killian Professor of Economics
Thesis Supervisor

Certified by .....
Iván Werning
Professor of Economics
Thesis Supervisor

Accepted by .....
Michael Greenstone
3M Professor of Environmental Economics
Chairman, Department Committee on Graduate Theses

© Jenny Simon, 2011. All rights reserved.
The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Optimal Policy and the Coexistence of Markets and Governments

by

Jenny Simon

Submitted to the Department of Economics on August 15, 2011, in partial fulfillment of the requirements to the degree of Doctor of Philosophy in Economics.

Abstract

This thesis explores three aspects of the coexistence of governments and markets from an optimal policy point of view.

In chapter 1, I study how the presence of financial markets shapes the government's ability to redistribute. Individuals do not constrain consumption to equal their net-of-tax income every period, but instead use markets to allocate their resources over time. This restricts the set of policy instruments available to the government. At the same time, however, markets enable agents to enter long-term consumption commitments. Changing these contracts is costly. These potential default costs mitigate the government's ex-post incentives to renege on the promised tax schedule, and therefore provide a commitment device for the government. In that sense, financial markets may facilitate rather than hinder redistribution.

In chapter 2, I present a rationale for corporate income taxes to discriminate between debt and equity financing. For risk-averse entrepreneurs, equity generates more surplus than debt, because it provides financing *and* insurance. A government seeking to extract surplus from entrepreneurs would naturally tax equity-generated income more than debt-generated income. Moreover, in the presence of private information, the government can use taxes to discriminate between different types of entrepreneurs. This degree of freedom allows a manipulation of the relevant incentive constraints, and an increase in overall efficiency. The optimal non-linear tax schedule to achieve the desired discrimination is isomorphic to one that taxes debt-generated income at a lower rate than equity-generated income.

In chapter 3, I explore how fast people adapt to institutional change. I study the differential reaction of former East and West Germans to a series of health care reforms. Along with the decrease in coverage under the public health insurance, former East Germans were significantly less likely to sign complementary insurance contracts in the private market. I show that the differential uptake rates of additional insurance are consistent with a model in which agents learn over time that institutions have changed and they are now responsible for optimizing their coverage. Thus, I provide evidence for the existence of a substantial transition period in the individual adaptation to new institutions.

Thesis Supervisor: Daron Acemoglu
Title: Elizabeth and James Killian Professor of Economics

Thesis Supervisor: Iván Werning
Title: Professor of Economics

Acknowledgments

I am deeply indebted to my advisor Professor Daron Acemoglu. Both his invaluable big picture advice and his willingness to dig through all the details of various drafts of my papers have been essential for my thesis. Many times I left his office in awe over the brilliance and clarity of his advice. Even when I had made little progress, he would find words of encouragement. I especially appreciated his guidance, honesty, and support during the process of finding a job; I owe my success to him.

I would also like to thank my secondary advisor Professor Iván Werning. On many occasions, he provided great insights that made me see the potential of my projects more clearly. I am grateful for his thoughtful, often challenging feedback that shaped how I define and approach research questions. His work has been and will continue to be the main inspiration for my own research.

Professor Mike Golosov, Professor James Poterba, and Professor Peter Diamond also had a great influence on my professional development. I am grateful for the training I received from them, and for their valuable advice at various stages of my studies. Moreover, I would like to thank my undergraduate advisor Professor Harald Uhlig. Without his encouragement and support I would not have been able to attend MIT.

My classmates have contributed immensely to my personal experience at MIT. Joaquin Blaum, Maya Eden, Luigi Iovino, Michael Peters, and Anna Zabai have been a source of friendship as well as good advice. Moreover, I would in particular like to thank Danielle Li and Florian Scheuer for their helpful feedback on my research.

I acknowledge very generous financial support throughout the Ph.D. program from the Peters-Beer-Stiftung. Their generosity has made it possible for me to attend this program. In my first year, I also received support from the World Economy Laboratory and the German Academic Exchange Service.

I dedicate this thesis to my parents, Gudrun and Gerd Simon. I am grateful beyond words for their unconditional love throughout all my life and especially the years at MIT. Knowing that no matter what happened I could always return to my family in Berlin has given me the strength to pursue this degree. Especially in tough times, they always knew how to put things in perspective. The laughs we shared, their encouragement to follow my dreams, but also their gentle reminders of what really matters in life have been essential for me to finish this program. Their ability to take life as it comes and make the best of it will always be an inspiration for me.

I also thank my sister Julia Simon, my grandparents Brigitte and Horst Simon, Elfriede and Fred Schneider, and my grand-aunt Irene Trampe, as well as my parents in law Sabine and Peter Huse. They all have been cheering from the sidelines and given me a sense of belonging even from afar.

Last, but certainly not least, I would like to thank Tim Huse. He has been with me all the steps of the way. Tim has selflessly given up his life in Berlin to move to the United States with me. He has weathered the hard times of finding a job in a foreign country to provide for our little family. In all stages of my work he served as a valuable critic, but more importantly has provided inexhaustible love and support. Without him, I would not have been able to succeed in this endeavor. Tim, I love you, and I look forward to the next chapter of our lives together.

Jenny Simon

Cambridge, Massachusetts

To my parents.

Contents

1 Financial Markets as a Commitment Device for the Government 13

1.1 Introduction 14

1.2 Endowment Economy with Private Information 20

1.3 Government with Information Constraints 21

1.3.1 Government with commitment 25

1.3.2 Government without commitment 27

1.4 Individual Access to a Financial Market 30

1.5 Commitment Through the Financial Market 38

1.5.1 Government with commitment 38

1.5.2 Government without commitment 40

1.6 Discussion 47

References 52

1.7 Appendix 54

1.7.1 Proof of Lemma 3 54

1.7.2 Proof of Lemma 4 58

1.7.3 Proof of Proposition 3 60

1.7.4 Proof of Proposition 5 66

2 Optimal Debt Bias in Corporate Income Taxation 69

2.1 Introduction 70

2.2 Setup 76

2.3 Equilibrium without Taxes 79

2.3.1 Observable Types 80
2.3.2 Unobservable Types 81
2.4 Equilibrium with Taxes 83
2.4.1 Implementing the Optimal Non-linear Tax Schedule 86
2.4.2 Continuum of Types with CARA Preferences 89
2.5 Discussion 91
References 94
2.6 Appendix 97
2.6.1 Proof of Lemma 11 97
2.6.2 Proof of Lemma 12 99
2.6.3 Proof of Lemma 6 100
2.6.4 Proof of Lemma 1 101

3 Adapting to Capitalism:
Private Health Insurance Uptake Among Former East Germans 103
3.1 Introduction 104
3.2 Health Insurance in Germany 108
3.2.1 Health Care Reforms 110
3.3 A Simple Model With Exogenous Learning 111
3.3.1 Heterogeneous Beliefs about Institutions 112
3.3.2 Comparing Two Populations 113
3.4 Data 115
3.5 Results 118
3.5.1 Age and Cohort Effects 123
3.5.2 Risk Taking and Risk Aversion 127
3.5.3 Preferences For a Larger Welfare State 129
3.5.4 Aggregate Effects: Regional Differences 131
3.6 Discussion 134
References 136
3.7 Appendix 138

List of Tables

3.1 Basic Regression 119
3.2 Basic Regression with East-Year Interactions 121
3.3 Changing East-Year Interactions - Test Statistics 122
3.4 Age Regression 125
3.5 Basic Regression For Different Age Groups 126
3.6 Risk Regression 128
3.7 Risk Regression 130
3.8 Preference Regression 131
3.9 Regression with State Controls 133
3.10 Summary Statistics 139
3.11 Basic Regression - Linear Probability Model 140
3.12 Basic Regression with East-Year Interactions - continued from table 3.2 141
3.13 Age Regression - continued from table 3.4 142
3.14 Basic Regression For Different Age Groups - continued from table 3.5 143
3.15 Regression by cohort - continued from table 3.14 144
3.16 Preference Regression - continued from table 3.8 145
3.17 Risk Regression - continued from tables 3.6 and 3.7 146
3.18 Regression with State Controls- continued from table 3.9 147

Chapter 1

Financial Markets as a Commitment Device for the Government

How does the presence of financial markets shape the government's ability to implement social redistribution? Individuals do not constrain consumption to equal their net-of-tax income every period, but instead use financial markets to allocate their resources over time. Thus, optimal redistributive policy ought to take agents' involvement in financial markets into account. From an allocative point of view, it restricts the set of policy instruments available to the government. At the same time, however, financial markets enable agents to borrow against their promised income and enter long-term consumption commitments. At any point in time, changing these contracts is costly. These potential default costs mitigate the government's ex-post incentives to renege on the promised tax schedule, and therefore provide a commitment device for the government. I show that whenever default costs are positive, the government is able to commit to a schedule that only pools some agents of similar type together. In that sense, financial markets may facilitate rather than hinder redistribution.

1.1 Introduction

In the presence of private information, the ability of a government to implement social redistribution depends crucially on its power to commit to future policy. This paper identifies a mechanism by which the existence of markets, and in particular financial markets, may enhance a government's ability to commit, and thus facilitate redistribution across society. Financial markets provide an opportunity for agents to enter long-term consumption commitments, and to borrow against their expected future income. Changing consumption plans and defaulting on loans is costly. A government choosing policy sequentially has to take the continuation value of agents' contracts into account: Deviation from previous announcements may lead to costly default. I show that any such default costs alter the government's ex-post incentives to renege on the promised transfer scheme and thereby effectively provide a commitment device for the government.

In market economies, individuals typically do not constrain their consumption to equal net-of-tax income every period. Instead, they use financial markets to allocate resources over time, allowing them to make long-term consumption plans. For instance, the ability to take out a mortgage enables individuals to live in a house that reflects their life-time income rather than in a rental unit that reflects their present disposable income every period. At any point in time, when re-optimizing redistributive policy, the government needs to take agents' consumption commitments into account. If agents end up with less net income than they expected, they have to adjust their consumption plans downward. Whenever agents have entered long-term commitments, such adjustments are costly. For example, defaulting on a mortgage may trigger costs of very diverse nature: A foreclosed house often does not sell for the same amount as it was worth to the original owner. Administration of defaults is costly. But also non-pecuniary losses may occur: When agents have to move out of the house they grew attached to, they may suffer further disutility. Chetty and Szeidl (2007) report that nearly 65% of the average US household's budget is allocated to

consumption commitments that cannot be adjusted costlessly. Such commitments would not be possible without access to financial markets. In this paper, I show that the ability of agents to enter consumption commitments, rather than just the plain opportunity to allocate resources over time, is what makes the existence of financial markets valuable for the government. Precisely because agents enter long-term commitments that cannot costlessly be adjusted, the government may gain the ability to commit to not changing the promised transfer scheme. Therefore, markets may in fact facilitate rather than hinder redistribution. The paper thus provides a new perspective on the concept of a social market economy, where markets play a crucial role for redistributive policy.

I consider a deterministic two-period endowment economy, where agents receive heterogeneous endowments every period. Income types are persistent and are private information. Individual income increases over time, so that every agent would like to smooth consumption by transferring resources from period 2 to period 1. Moreover, a benevolent government would like to redistribute across agents. In a seminal contribution, Mirrlees (1971) established that when the individual ability to generate income is private information, optimal redistributive policy needs to trade off allocative efficiency against information rent extraction. In the presented environment, a benevolent government with an exogenous commitment device can use a fully separating transfer scheme to implement this optimal trade-off between efficiency and equity: Analogous to Mirrlees' (1971) "efficiency at the top" result, at the constrained efficient allocation, only agents of the highest type receive perfectly smooth consumption. The government uses the degree of consumption smoothing as an incentive for agents to reveal their income type truthfully and to contribute to social redistribution. When agents are able to use a financial market to borrow against their future income, *every* agent gains the opportunity to smooth his income perfectly over time. The government *with* ex-ante commitment thus finds itself unable to implement the constrained efficient allocation, where it could use the degree of smoothing as an incentive for truthful revelation. Consequently, less redistribution is implemented. In

this situation, it is irrelevant whether or not agents enter long-term consumption commitments. It is only the allocative aspect of agents contracting in the financial market that constrains the set of policy instruments available to the government. When the government has an exogenous commitment device, the existence of markets may thus hinder redistribution.[1]

On the contrary, I argue that a government that *cannot* commit ex-ante to future policy may gain from the existence of a financial market. Roberts (1984) was the first to show that lack of commitment in dynamic taxation settings with private information may lead to a government not being able to implement any social redistribution. Analogously, in the economy presented here, when policy can be chosen sequentially over time, the government has an incentive to use the information about agents gathered in the past to achieve a better redistributive outcome ex-post. Agents anticipate this behavior. The resulting time-consistency constraint leads to the severely inefficient outcome of no social redistribution as well as almost no smoothing of individual consumption over time.[2] In this case, the government may in fact gain from agents' involvement in a financial market. When agents pledge their income in a private contract and enter a long-term consumption commitments, deviating from past policy announcements may not be optimal for the government anymore, despite its desire to redistribute. Depending on how many agents would have to default on their debt, the associated cost may be too high to justify any welfare gains from additional redistribution. I show that whenever default is costly, the government is effectively able to commit at least to a partially separating transfer schedule. Here, it is the consumption commitment characteristic of agents' contracts rather than the allocative aspect that has a favorable effect for the government's ability to effectively commit to future policy. The presence of financial markets and agents' involvement therein enables the implementation of social redistribution.

[1] Many authors have considered environments in which agents cannot only contract with a principal, but also in anonymous outside markets that make it harder to extract information from the agents truthfully. See for example Hammond (1987) for a general treatment or Golosov and Tsyvinsky (2007) for a more recent example from the dynamic public finance literature.

[2] See for example Golosov et al. (2006) for a derivation of this phenomenon in a general setting.

As the main result, I derive a simple condition that links the size of the default costs and the concavity of the utility function to the degree of separation a government is able to commit to. Intuitively, this condition equalizes the marginal benefit from additional redistribution toward the low end of the type distribution to the marginal cost due to default. The larger the default costs are, given the concavity of utility, the more separation and so the more social redistribution is possible. Conversely, for fixed default costs, a more concave utility function makes less separation possible, because it increases the ex-post welfare gain from redistribution across agents. The financial market effectively provides the government with a device for limited commitment, i.e. with the power to commit to not exploit a limited amount of information.

Moreover, the derived results provide insights into the form of partial separation that emerges at the optimum. If the government is allowed to randomize transfers between seemingly identical agents, the optimal allocation is such that agents are perfectly separated below a cutoff type and completely pooled above. The government collects perfect information about agents up to a threshold type chosen according to how much separation it can commit to, and simply pools together all higher income types. On the contrary, if the government is constrained to comply with horizontal equity, agents are optimally pooled together in groups throughout the type distribution. The resulting transfer schedule then has a "tax bracket" structure. The government collects coarse information over the complete type distribution, higher default costs allow for more and smaller brackets, so that more detailed information is collected.

In summary, this paper identifies a mechanism by which the economic environment a government operates in can provide a potential commitment device that does not rely on reputational considerations or political constraints and works in a finite horizon. Thereby, my results help reconcile the observation of policies that are suggestive of governments being able to commit, even though there is no apparent commitment device. In that sense, this paper provides a new perspective on the concept of a social

market economy, where the presence of well-functioning markets plays a crucial role for social redistribution. In addition, the fact that the degree of commitment can vary with the default costs introduces a rationale for why tax policy might not use *all* available information, but rather just condition on *coarse* private information.

Related Literature

This paper contributes to a recently growing literature on the interaction and co-existence of markets and governments. One branch of this literature attempts to identify circumstances under which markets outperform benevolent governments. Netzer and Scheuer (2010) consider a setup in which time-inconsistency arises because of an adverse selection problem. They show that markets can outperform benevolent governments even when they face the same adverse selection problem, because they provide greater incentives to exert effort. In their model, markets endogenously generate a form of commitment to refrain from full insurance and pooling. The key characteristic of markets enabling them to implement separating equilibria is competitiveness: agents are free to switch their insurance provider, just like firms can renege on the insurance contracts. This two-sided lack of commitment is what distinguishes the markets they consider from a benevolent government.

Acemoglu, Golosov, and Tsyvinski (2008a, 2008b) also compare the efficiency of markets and governments in a setting without commitment. They explore the impact of political economy constraints on optimal redistributive policy. To that end, they consider infinitely repeated games with equilibria that crucially rely on reputation effects - a channel completely abstracted from in this paper.

In contrast to these contributions, I am not comparing the performance of markets and governments, but rather ask how the presence and functioning of markets influences the government's ability to implement redistributive policy. Some characteristics of markets have been considered in the literature: Scheuer (2010) explores the impact of incomplete credit markets on optimal entrepreneurial taxation. He finds

that a market friction which gives rise to cross-subsidization between different types of potential entrepreneurs may induce inefficient entry at both ends of the skill distribution, which in turn promotes an additional corrective role for type-differential, redistributive taxation, even when the government originally has no redistributive objective. Unlike in Scheuer's paper, I consider a market that is not incomplete in that sense, and instead is able to provide credit without cross-subsidization.

Bisin and Rampini (2006) study a setup similar to the one considered here, but focus on the allocative role of *anonymous* markets. They find that allowing agents access to such markets is beneficial in a world where the government has no commitment, because it allows them to allocate resources over time without revealing any information, thereby increasing efficiency. However, the government's commitment problem is unchanged, no social redistribution can be implemented. In contrast, I analyze a market that does not act as a "tax haven" by enabling agents to hide information from the government. The crucial characteristic of private contracts I consider is that they constitute consumption commitments that cannot costlessly be changed. This increases the government's commitment power, enabling it to implement some social insurance.

The paper is organized as follows: I start by formulating a basic two period endowment economy with private information about income types in section 1.2. In section 1.3, I derive the constrained efficient allocations for governments with and without commitment when only the government has access to a borrowing technology as benchmarks for the following analysis. Section 1.4 extends the setup by introducing a financial market that allows agents to borrow against their future income. I derive the constraints that agents' involvement in a financial market imposes on the planning problem. Here I discuss in detail the characteristics of the financial market that lead to it functioning as a potential commitment device for the government. Section 1.5 analyzes the efficient allocations under this additional constraint for governments that can or cannot commit to future policy ex-ante. The main point of the paper is

derived: due to default costs, a non-commitment government is able to implement an allocation with partial separation, and thus can provide some social redistribution. Finally, section 1.6 concludes.

1.2 Endowment Economy with Private Information

The model economy lasts for 2 periods (t=1,2) and is inhabited by a continuum of agents of unit mass. Agents derive utility from a single consumption good according to

$$U = \sum_{t=1}^{2} u(c_t).$$

Utility is time-separable, and the per period utility function $u(\cdot)$ is strictly increasing, concave, and $\lim_{c \to 0} u'(c) = \infty$. I also assume that u displays constant elasticity of intertemporal substitution. To simplify the following analysis, I assume that agents do not discount between periods.

Agents receive heterogeneous income at the beginning of each period. Their income types, denoted θ, are perfectly persistent over time and are private information. Across the population θ is continuously distributed over a support $\Theta = [\underline{\theta}, \bar{\theta}]$, $F(\theta)$ denotes its cdf, which I assume to be continuously differentiable. Further, I assume that $\underline{\theta} \geq \frac{1}{2}\bar{\theta}$.[3] Apart from income heterogeneity, agents are identical.

Per period income is deterministic, and increases over time and across types. In particular, I assume that it is $t\theta$. Consequently, agents would like to smooth consumption over time and consume a constant fraction $\frac{3}{2}\theta$ of their overall income in each period.

[3] I make this assumption to rule out cases off the equilibrium path where agents and banks collude against the government. The impact of this assumption is discussed in section 1.4.

Consider the problem of a benevolent social planner with a utilitarian objective and equal Pareto weights on all agents. He chooses an allocation $\{c_t(\theta)\}_{t,\Theta}$ that assigns a consumption level to each type $\theta \in \Theta$, for each period $t = 1, 2$.

$$\begin{aligned} &\max_{\{c_t(\theta)\}_{t,\Theta}} \int_\Theta \left(\sum_{t=1}^{2} u\left(c_t(\theta)\right) \right) dF(\theta) \\ &\text{s.t.} \int_\Theta \left(\sum_{t=1}^{2} c_t(\theta) \right) dF(\theta) \leq 3 \int_\Theta \theta dF(\theta) \end{aligned} \tag{1.1}$$

The aggregate feasibility constraint reflects the assumption that there exists a technology to costlessly transfer resources between periods. The optimal allocation solving this problem is described as follows:

Lemma 1 (First-Best Allocation)

At the first-best allocation there is full social redistribution and perfect smoothing of consumption over time. All agents consume a constant fraction $c_1(\theta) = c_2(\theta) = c = \frac{3}{2} \int_\Theta \theta dF(\theta)$ *of the economy's total endowment in each period.*

Proof: The first order condition with respect to any agent's consumption in either period satisfies

$$u'(c_t(\theta)) - \lambda = 0 \quad \forall t, \theta$$

where λ is the Lagrange multiplier on the aggregate feasibility constraint. Thus, $c_t(\theta) = c_{t'}(\theta') \quad \forall t, t', \theta, \theta'$. □

1.3 Government with Information Constraints

Suppose a benevolent government can borrow and save at the risk free gross interest rate $R = 1$, i.e. it can costlessly transfer resources between periods. To implement the desired allocation, it would like to institute a schedule of type specific transfers $\{T_1(\theta), T_2(\theta)\}$. However, it faces private information constraints: When conditioning the allocation on income types, the government must rely on information reported by the agents. This turns the setup into a policy game between agents, choosing which

type to report, and the government, choosing the transfers to implement.

To analyze this game formally, consider first the timing of action:

1. Agents learn their income type θ.
2. The government announces a schedule $\{T_1, T_2\}$.
3. Period 1:
   a) Agents receive their first period endowment θ and send a report σ.
   b) The government implements transfers $\{T_1(\sigma)\}$.
4. Period 2:
   a) Agents receive their second period endowment 2θ.
   b) The government implements transfers $\{\hat{T}_2(\sigma)\}$, possibly different from the schedule announced before, depending on its commitment power.

As usual in such setups with private information, it is crucial whether or not the government can commit to not exploit the revealed information at a later point in time, i.e. whether or not it can commit to not changing the announced allocation to the disadvantage of some agents after information has been revealed. I assume that the government can always commit to the announced schedule at least *within* period 1, i.e. it will always implement transfers in period 1 according to the original announcement. The potential commitment problem that is the subject of this paper arises between periods 1 and 2. At that point the government has learned information about the agents' types. If it is not committed to the announcement made earlier, it may decide to implement transfers $\{\hat{T}_2(\sigma)\}$ that differ from the initial announcement $\{T_2(\sigma)\}$. Whether or not a government can commit is public information, and agents take it into account when choosing which type to report.

Let $\sigma : \Theta \mapsto \Sigma$ denote an agent's reporting strategy, a function that maps from the set of possible realizations of income types Θ to a set of possible reports Σ. For

future reference, let σ^* denote the direct truth-telling strategy where agents simply reveal their type truthfully, i.e. $\sigma^*(\theta) = \theta$. The utility obtained from any reporting strategy σ, given the government's transfers $\{T_1, T_2, \hat{T}_2\}$ is

$$U(T(\sigma)|\theta) = u(\theta + T_1(\sigma)) + u(2\theta + \hat{T}_2(\sigma)) \tag{1.2}$$

For truth-telling to be optimal for an agent of type θ, it must be that

$$U(T(\sigma^*)|\theta) \geq U(T(\sigma)|\theta) \quad \forall\sigma \tag{1.3}$$

The government's strategy involves choosing a set of transfer schedules $T = \{T_1, T_2, \hat{T}_2\}_M$. When the government has commitment, $\{T_2\}$ and $\{\hat{T}_2\}$ are exogenously constrained to be equal.

Definition 1

A (perfect Bayesian) equilibrium in the game between agents and the government is given by strategies σ^e and T^e and a belief system $\mathbb{B}$, such that σ^e and T^e are best responses to each other, given $\mathbb{B}$, and beliefs are derived from Bayesian updating[4].

To analyze the equilibrium of this game, I employ a general mechanism design approach (as e.g. in Bester and Strausz (2001) and Skreta (2007, 2010)) where a fictitious mechanism designer is in charge of choosing strategy sets for the agents (the set of possible reports Σ) and for the government (a set of possible transfer schedules T). While abstract, this approach has a number of advantages: The fictitious planner is always able to commit. The Revelation Principle then allows attention to be restricted to direct revealing mechanisms, i.e. agents' strategy set can without loss of generality be restricted to the set of possible types Θ. The optimal mechanism simply has to satisfy incentive compatibility for truth-telling (1.3). This is true even when the government (a player in this game) has no commitment, because the fictitious planner can restrict the government's strategy set as well: in particular, he can decide how much of the information agents report is revealed to the government. For-

[4]In the following analysis there will be no need to explicitly derive or condition on these beliefs.

mally, this amounts to the optimal mechanism specifying an *information revelation rule* $m : \Theta \mapsto M$ that maps from agents' reports to some set of possible messages the government observes. The government is then restricted to choose transfers T that condition only on these messages. The function m could be such that no information is revealed (i.e. m is constant), full information is revealed (i.e. m is the identity function), but could also allow for any form of partial information revelation (i.e. m is constant over some subset of Θ so that some agents are pooled together). Thus, this setup allows me to explicitly study situations where the government has limited commitment in the sense that it can commit not to exploit a limited amount of information. The main focus of the analysis in this paper will be on the optimal form of the information revelation rule m as a proxy for the commitment power of the government and the characteristics of the resulting allocation.

It is useful to think about the economic interpretation of the information revelation rule: In reality, when taxes and transfers are conditional on private information, the government must decide how people report this information. For example, the first step to implementing an income tax is to design a tax return form that people use to report their income. The government, knowing how much information it can commit not to exploit in the future, can choose an institutional design that asks agents only for *coarse* information. The tax return, for example, could only ask for an agent's approximate income, or an income bracket. The function m can be interpreted as this institution.

Moreover, note that since agents know whether or not the government has commitment, they correctly anticipate the government's incentive to re-optimize policy in the second period, and so condition their reporting strategy on $\{\hat{T}_2\}$ rather than on the announced $\{T_2\}$. Thus, there is no need to specify both of them separately. To summarize, the problem of the fictitious planner is to design a mechanism $\Gamma = (m, \{T_1(m), T_2(m)\})$ that satisfies incentive compatibility for all agents:

$$U(T_1(m(\theta)), T_2(m(\theta))|\theta) \geq U(T_1(m(\hat{\theta})), T_2(m(\hat{\theta}))|\theta) \quad \forall \theta, \hat{\theta}$$

It is without loss of generality to restrict attention to deterministic mechanisms[5] so that the set of possible messages is $M = \Theta$, and to assume that $m : \Theta \mapsto \Theta$ is weakly increasing. Moreover, I normalize m such that

$$m(\hat{\theta}) = \hat{\theta} \quad \text{for } \hat{\theta} = min\{\theta : m(\theta) = m\}^6$$

In the remainder of this section I will derive the optimal mechanisms when the government can or cannot commit, and agents do not participate in a financial market. These will serve as benchmarks. Section 1.4 then proceeds by deriving any additional constraints on the mechanism design problem that arise when agents can borrow individually in a financial market. Section 1.5 analyzes the resulting change in the optimal use of information and the implemented allocation.

1.3.1 Government with commitment

When the government has commitment, the optimal mechanism solves the following problem:

[5]Due to the CRRA assumption, non-degenerate stochastic mechanisms are suboptimal. Since the objective function is concave, introducing risk could only improve matters if some incentive constraints were relaxed. Making payoffs for lower type agents riskier does indeed relax higher types' incentive constraints. However, since CRRA implies decreasing absolute risk aversion, the loss for the low types from facing such risk is always higher than the gain in terms of relaxing incentive constraints for higher types. See for example Fudenberg and Tirole (1991).

[6]This just means that when some types are pooled together, the message sent to the government is normalized to be equal to the lowest type in that group.

$$\max_{\Gamma} \int_{\Theta} \sum_{t=1}^{2} u(t\theta + T_t(m(\theta)))dF(\theta) \tag{1.4}$$

$$\text{s.t.} \int_{\Theta} [T_1(m(\theta)) + T_2(m(\theta))]dF(\theta) \leq 0 \tag{1.5}$$

$$\theta \in \arg\max_{\hat{\theta}} \sum_{t=1}^{2} u(t\theta + T_t(m(\hat{\theta}))) \quad \forall \theta, \hat{\theta} \in \Theta \tag{1.6}$$

$$m : \Theta \mapsto \Theta \tag{1.7}$$

It maximizes a utilitarian welfare function (1.4) with equal Pareto weights on every agent, subject to aggregate feasibility (1.5) and incentive compatibility (1.6), choosing the information revelation rule m and the transfer schedule $\{T_t(\theta)\}$ for $t = 1, 2$ optimally.

Lemma 2 (Information Revelation with Commitment)

If the government can fully commit, the optimal information revelation rule is such that complete information about types is revealed: $m(\theta) = \theta$ for all $\theta \in \Theta$.

Proof: Since the government has full commitment, Lemma 2 follows directly from the Revelation Principle[7]. □

When the government is able to commit to not changing the announced transfer schedule after information is revealed, it is optimal to implement a fully separating allocation. The resulting constrained efficient allocation has the following characteristics:

Lemma 3 (Optimal Allocation with Commitment)

At the optimal allocation with commitment:

(i) There is partial social redistribution - total consumption is increasing in type, but less steeply than under autarky:

$$0 < \frac{\partial(c_1(\theta) + c_2(\theta))}{\partial\theta} < 3$$

[7]See for example Myerson (1979) and Harris and Townsend (1981).

(ii) The degree of smoothness of consumption increases with type, only the highest type smooths consumption perfectly:

$$c_1(\bar{\theta}) = c_2(\bar{\theta})$$

$$c_1(\theta) < c_2(\theta) \quad \& \quad \frac{\partial \frac{c_1(\theta)}{c_2(\theta)}}{\partial \theta} > 0 \quad \forall \theta < \bar{\theta}$$

Proof: See appendix 1.7.1.

The setup resembles the traditional static Mirrlees (1971) model, where the desire to smooth consumption efficiently over time corresponds to the optimal labor/leisure choice in Mirrlees' setup. The optimal allocation depicts the classic trade-off between allocative efficiency and informational rent extraction under adverse selection. Even though both forms of redistribution (across the population as well as across time) are in the government's interest, the private information constraints introduce a trade-off between the two. Since the elasticity of intertemporal substitution is constant, all types are willing to give up the same fraction of their total income for smoothing consumption over time. In absolute terms, agents with higher income types would pay more for consumption smoothing than lower income types. The government uses the degree of smoothness as an incentive for higher types to reveal themselves and agree to higher contributions to social redistribution - the ability to do so crucially depends on the government being able to commit to the allocation ex-ante. Perfect consumption smoothing for the highest type is analog to Mirrlees' (1971) "efficiency at the top" result, non-perfect smoothing for all other types refers to the distortion of efficiency for all types other than the highest.

1.3.2 Government without commitment

If policy is chosen sequentially and the government cannot commit to a schedule ex-ante, the before stated optimization problem becomes even more constrained. The optimal allocation can be found by solving the above planning problem subject to an

additional commitment constraint.[8] It must be clear that when types are revealed, the government does not have an incentive to renege on the promised allocation at a later point in time. The problem is the same as above (equations (1.4) through (1.7)), with the following additional constraint:

$$\begin{aligned} \{T_2(m(\theta))\} \in \arg \max_{\{\hat{T}_2(m(\theta))\}} & \int_\Theta u(2\theta + \hat{T}_2(m(\theta)))dF(\theta) \\ \text{s.t.} & \int_\Theta [T_1(m(\theta)) + \hat{T}_2(m(\theta))]dF(\theta) \leq 0 \end{aligned} \tag{1.8}$$

This constraint requires that in period 2, the government won't change the promised transfer schedule based on information it learned in period 1. Since types are persistent, this amounts to maximizing second period welfare, only constrained by feasibility.

Lemma 4 (Information Revelation without Commitment)
If the government cannot commit, the optimal information revelation rule is such that no information about types is revealed: $m(\theta) = \underline{\theta}$ for all θ.

Proof: See appendix 1.7.2.

When the government cannot commit to not exploit information about types in period 2, it is not optimal to implement any separation at all. All agents will pool with the lowest type, no information about types is revealed.

The argument of the proof is as follows. Because of the commitment constraint (1.8), the government loses the ability to offer any separation in period 2 consumption: Since the necessity to provide incentives for agents to reveal their type truthfully vanishes after the first period, the government would always change the announced allocation when provided with the opportunity to do so. Such deviation from the ex-ante optimal contract, though, is not beneficial for all agents. The government offering above

[8] While before agents were moving after the government, lack of commitment introduces a second stage to the game, where only the government can move again. The commitment constraint essentially imposes subgame-perfection on the equilibrium, as e.g. in Kydland and Prescott (1977).

mean type agents a worse allocation after learning their true income is known as the *ratchet effect*[9]. Agents anticipate this, so incentives for truthful revelation need to be provided through transfers in period 1. However, to achieve any separation in types, the incentive payments would have to be so high, that redistribution would go from the bottom to the top of the income distribution - inequality would rise compared to autarky. Thus, complete pooling is the optimal choice of information revelation. Consequently, no redistribution across agents (i.e. social insurance) and almost no redistribution across time (i.e. consumption smoothing) will be implemented:

Lemma 5 (Optimal Allocation without Commitment)
At the optimal allocation without commitment:

(i) There is no social redistribution - agents consume their total endowment, total consumption increases in type as under autarky:

$$c_1(\theta) + c_2(\theta) = 3\theta \quad \forall\theta$$

(ii) Only one type θ^* smooths consumption perfectly:

$$c_1(\theta^*) = c_2(\theta^*) = \frac{3}{2}\theta^*$$
$$c_1(\theta) = \theta + \frac{1}{2}\theta^* \neq c_2(\theta) = 2\theta - \frac{1}{2}\theta^* \quad \forall\theta \neq \theta^*$$

Proof: When no information about types is revealed, the only instrument to increase welfare is to hand out non-differential transfers. The government will choose these optimally to smooth consumption for on particular type θ^*. All other agents therefore smooth only the part of income equal to that of type θ^* and consume their remaining income on the spot. □

In this economy, the government's lack of commitment has dramatic implications. Not only is the government unable to implement any social redistribution, the re-

[9]The insight that the only incentive compatible sequence of spot contracts is one without dynamic insurance is due to Townsend (1982).

sulting allocation is also very inefficient: Even though transferring resources across time is costless, this technology remains almost unused, because it would require the revelation of private information. Roberts' (1984) insight applies in this economy.

1.4 Individual Access to a Financial Market

Suppose now that agents have access to a financial market in which they can save and borrow at interest rate $R = 1$, i.e. they can use the same technology to transfer resources over time that is available to the government. Naturally, agents would use this opportunity to smooth consumption over time. Neither allocation analyzed in section 1.3 had full smoothing for all agents, so access to such a financial market likely imposes a binding constraint on the optimal mechanism. The purpose of this section is to derive the constraints that stem from the contracts agents may write in such a financial market.

Bisin and Rampini (2006) first showed, that in a setup similar to the one presented here, financial markets that can be used *anonymously* may be a beneficial constraint for a government without commitment. Such markets improve efficiency in the allocation of resources over time without disclosing information about types to the government. This leads to an increase in welfare. The government's commitment problem, however, remains unchanged. Still, no social redistribution would be possible.

I emphasize a different mechanism: Agents use the financial market to smooth consumption over time. To do that, they pledge future income in private contracts that resemble long-term consumption commitments (e.g. mortgages). When a government changes the announced allocation, these contracts may have to be renegotiated or even be defaulted on - a process that is usually costly. This introduces a cost to deviating from the announced allocation ex-post. In fact, it will turn out that agents' involvement in a financial market may essentially provide the government with a form

of limited commitment. The presence of markets may thus facilitate rather than hinder redistribution across agents. To derive this insight formally, I will lay out the critical characteristics of the market environment and derive the structure of private contracts that emerge in the presented economy. These contracts will be treated as constraints to the mechanism design problem. Section 1.5 proceeds with analyzing their impact on the optimal revelation of information and the resulting allocation when the government can or cannot commit through an exogenous commitment device.

Assumptions about the financial market

The market consists of many banks that have access to unlimited outside funding. Agents and banks can write contracts

$$[(h_1(\sigma), h_2(\sigma)), (b_1(\sigma), b_2(\sigma))] \quad \text{with } h_t, b_t \geq 0$$

where the bank agrees to provide h_t units of consumption in period t for a payment of b_t by the agent who announced type σ. This structure of contracts is very general. In what follows I will discuss which of these four determinants of contracts are of relevance to the results.

I make the following assumptions that shape the type of debt contracts signed by agents in this economy.

(A1) Competition between banks ensures that they make zero profits. The gross interest rate agents face is $R = 1$.

(A2) Banks can always enforce their contracts with the individual agents. This enforcement power is never revoked, or in other words, the government is always able to commit not to shut down the market. However, banks have the first take on *net-of-tax* income only. They do not possess any power over the government to enforce bailouts.

(A3) The market is not completely anonymous as in Bisin and Rampini (2006). The government can observe the contracts agents sign in the financial market up to the precision with which it observes agents' type announcements. This means that agents who are pooled together by the information revelation rule m are still able to write differential contracts without revealing any additional information to the government. However, the government observes if the contracts are feasible given the announced type. Thus, the financial market does not act as a "tax haven". Agents who reported a lower income than they actually have are not able to secretly smooth their consumption. On the other hand, all agents who reported their type truthfully are able to smooth consumption perfectly without revealing their precise type. This restriction on observability of transactions in the financial market is introduced for expositional convenience, and I will point out its impact on the results. The main results are unchanged if contracts were completely observable.

(A4) The government is not able to restrict access to the financial market or punish agents' market involvement except when it reveals that they were lying about their type. This amounts to assuming that the government cannot announce a transfer schedule T ex-ante that conditions payments on whether agents will contract in the financial market. It does not exclude the possibility that a government without commitment implements differential transfers ex-post depending on the contracts that were signed.

Consider the following modified timing of events:

1. Agents learn their income type θ.
2. A mechanism $\Gamma = (m, \{T_1(m), T_2(m)\})$ is announced.
3. Period 1:
   a) Agents receive their first period endowment θ and send a report $\sigma(\theta)$.
   b) The government observes messages $m(\sigma)$ and implements transfers $\{T_1(m)\}$.

c) Agents may contract in the financial market, first period payments b_1 and h_1 are executed

4. Period 2:

a) Agents receive their second period endowment 2θ.

b) The government implements transfers $\{\hat{T}_2(m)\}$, possibly different from the schedule announced before, depending on its commitment power. It takes the contracts agents signed into account.

c) Second period payments b_2 and h_2 are executed

Assumption (A2) is reflected in the fact that transactions in the financial market always take place after the government implemented its transfers. Together with assumption (A4), this implies that the government cannot levy a tax on h_t directly.[10] It can only influence net income and thereby bound the possible debt payments b_t.

A government that can choose policy sequentially will take the contracts agents signed into account when re-optimizing the transfer schedule in period 2. This is the key argument: The continuation value of agents' contracts may be such that the government finds it not optimal to renege on the promised allocation and so effectively gains commitment. In case an agent of (announced) type σ defaults on his loan, i.e. in case he cannot pay the amount $b_2(\sigma)$ agreed upon, his contract is renegotiated to $[\hat{h}_2(\sigma), \hat{b}_2(\sigma)]$. It is without loss of generality to assume that the bank cuts the contracted payment h_2 to zero:

$$\hat{h}_2(\sigma) = 0 \tag{1.9}$$

Yet, the bank remains in power to collect any outstanding balance $d_1(\sigma) = h_1(\sigma) - b_1(\sigma)$ from period 1. Such renegotiation, however, comes at a cost: The bank does not value h_2 at the same rate the agent does. From the bank's point of view, saving

[10] In reality, Austria is one of few exceptions to this assumption. The Austrian government levies a tax of currently 0.8% of the loaned amount in any debt contract on the debitor.

the second period payment is worth only $h_2 - H_B$. There are several interpretations for this cost: First, re-allocating resources to a new project is costly for a bank. Renegotiating contracts may also require costly administration. Second, notice that h_2 can be interpreted as collateral on the loan. It is only natural to assume that the bank might not be able to resell the asset for the same value it had for the particular agent. Banks, however, cannot make losses when agents default. I assume that they remain in power to collect the difference H_B from the defaulting agent's net-of-tax income. Therefore, the defaulting agent cannot consume before repaying

$$\hat{b}_2(\sigma) = min\{d_1(\sigma) + H_B, 2\theta + \hat{T}_2(m(\sigma))\} \tag{1.10}$$

If his net-of-tax income is less than $d_1 + H_B$, he would have to consume zero. Moreover, an agent who has to default on his loan may suffer an additional utility loss H_A that the government also has to take into account when reneging on its promised transfer schedule.

The default costs $H_A + H_B = H$ summarize the key characteristic of contracts in private financial markets that I want to emphasize in this paper: Contracts are not easily reversible, nor is it costless to renegotiate or default. The costs may be of very diverse nature: On the one hand, one might think of pure resource costs for administering the renegotiation. Re-allocating funds from one loan to another is also costly. A bank might not be able to resell an asset for the same value agreed upon previously with the now defaulting agent. Such costs are summarized by H_B. On the other hand, agents might suffer a loss in utility when they have to default on their loan in addition to the resource costs of the bank. They may have made life plans contingent on this loan that require further costly alteration. For example, they might have grown attached to their house, which they financed with a mortgage, and lose utility when they have to move. Such costs are summarized by H_A.

In short, agents who have access to financial markets may tie up their resources in a

contract whose continuation value needs to be taken into account when redistributing across the population. Redistributing across agents more than initially announced becomes costly ex-post, and so alters the government's optimization problem.

Structure of debt contracts

Banks set borrowing limits per announced type σ that reflect total net-of-tax income. Assumption (A1) implies that all contracts will be such that

$$h_1(\sigma) + h_2(\sigma) = b_1(\sigma) + b_2(\sigma) \tag{1.11}$$

Conditioning the contracts on agents' reports σ rather than their types θ eludes to the fact that banks must rely on the information agents reveal about themselves. In the first period all agents are net-borrowers. This opens up the possibility that an agent reports a much higher type to take advantage of a high borrowing limit and plans a sure default. To avoid such adverse selection, banks would like to verify that agents are at least of the type they claimed. While banks cannot verify an agent's type directly, notice that they can offer contracts that require a *down payment* of

$$b_1(\sigma) = \sigma + T_1(m(\sigma)) \tag{1.12}$$

This proof of solvency acts as a screening device, i.e. it signals to the bank that the agent is indeed at least of the type he claimed he was.[11] Competition then ensures that each agent can find a bank offering a contract with a borrowing limit that reflects the exact net income of the type he announced. Banks cannot gain by offering contracts that don't require down payments, since only agents who misreported their type would sort into those.

[11] This setup allows me to abstract from any additional adverse selection problem the financial market may face. Scheuer (2010) considers the impact of a financial market with adverse selection on optimal policy. Since I assumed $\underline{\theta} \geq \frac{1}{2}\bar{\theta}$, banks could never gain from letting agents borrow more than their type renders feasible. Without that assumption one could imagine a case where agents borrow much more than they can repay, forcing the government to bail them out. With that assumption, any agent would always be able to pay even the highest types tax in period 2, the default loss would then have to be absorbed by the bank.

Notice that because of these down payments, default is only possible if the government deviates from the announced transfer schedule. For the remainder of the analysis, it is enough to keep track of the net debt obligation each agent carries over to period 2:

$$d_1(\sigma) = h_1(\sigma) - b_1(\sigma)$$

Agents in this economy use the market to borrow against their second period income, to smooth consumption and consume more in period 1 than they are endowed with. Contracts will thus typically have $h_1 > 0$. How much agents can borrow depends on the type they reported. Banks will set borrowing limits that reflect total net income:

$$h_1(\sigma) + h_2(\sigma) \leq 3(\sigma) + T_1(m(\sigma)) + T_2(m(\sigma))$$

Even though the market and information structure impose some constraints, agents can still choose between a variety of contracts. Given a schedule of transfers and his report σ, an agent chooses to contract in the financial market to maximize his life-time utility:

$$\max_{h_t, b_t} \sum_{t=1}^{2} u(t\theta + T_t(m(\sigma)) + h_t(\sigma) - b_t(\sigma))$$

$$\text{s.t. } h_1(\sigma) + h_2(\sigma) = b_1(\sigma) + b_2(\sigma)$$

$$b_1(\sigma) \in \{0, \sigma + T_1(m(\sigma))\}$$

$$b_2(\sigma) \begin{cases} \leq 2\sigma + T_2(m(\sigma)) & \text{if } b_1(\sigma) > 0 \\ = 0 & \text{if } b_1(\sigma) = 0 \end{cases}$$

Whenever transfers are such that the agent's net income is not smooth over time, the

optimal solution to this problem is such that

$$\begin{aligned} b_1 &= \sigma + T_1(m(\sigma)) \\ h_1 &= \frac{1}{2}(3\sigma + T_1(m(\sigma)) + T_2(m(\sigma))) \\ \rightarrow d_1 = h_1 - b_1 = &= \frac{1}{2}(\sigma - T_1(m(\sigma)) + T_2(m(\sigma)) \end{aligned} \tag{1.13}$$

Optimal contracts are not uniquely pinned down. In the second period, payments could be as low as

$$b_2(\sigma) = d_1(\sigma) \quad \text{and} \quad h_2(\sigma) = 0 \tag{1.14}$$

(in this case, the agent would simply repay his outstanding debt), or as high as

$$b_2(\sigma) = 2\sigma + T_2(m(\sigma)) \quad \text{and} \quad h_2 = h_1 = \frac{1}{2}(3\sigma + T_1(m(\sigma)) + T_2(m(\sigma))) \tag{1.15}$$

I refer to these possibilities as *net* or *gross* contracts respectively. All contracts in between these extremes leave the agent with the same consumption allocation. However, when signing a gross contract, agents enter a consumption commitment for period 2 beyond the repayment of their net balance. While both types of contracts serve to allocate resources over time and to smooth consumption, a gross contract does that in the form of a long-term commitment. In other words, a gross contract may be interpreted as a mortgage, where the agent constrains his consumption of housing to a particular house not only in period 1, but also in period 2 with the help of a financial contract. If the government has ex-ante commitment, agents are indifferent between signing net and gross contracts. However, if the government has no ex-ante commitment, it will turn out to be individually optimal for agents to sign gross contracts.

The financial market provides agents with the opportunity to smooth their consumption perfectly over time. Moreover, agents may sign contracts that resemble long-term consumption commitments (e.g. mortgages). While this additional characteristic of private contracts in a financial market is irrelevant for the government with commit-

ment, it will turn out to be crucial in determining the de facto commitment power of a government that chooses policy sequentially.

1.5 Commitment Through the Financial Market

When agents have access to a financial market, the private contracts described by equations (1.13) through (1.15) constrain the choice of the optimal mechanism $\Gamma = (m, \{T_1(m), T_2(m)\})$.

1.5.1 Government with commitment

When the government can commit to a schedule $\{T_1(m), T_2(m)\}$ ex-ante, the optimal mechanism now solves problem (1.4) subject to feasibility (1.5) and a set of *modified* incentive compatibility constraints:

$$\theta \in \arg\max_{\hat{\theta}} \sum_{t=1}^{2} u(t\theta + T_t(m(\hat{\theta})) + h_t(\hat{\theta}) - b_t(\hat{\theta})) \quad \forall \theta, \hat{\theta} \in \Theta \tag{1.16}$$

taking the contracts agents sign as given. With access to the financial market, all agents who revealed their true type are able to perfectly smooth consumption themselves.

Lemma 6 (Information Revelation with Commitment)
If the government can fully commit, the optimal information revelation rule is such that complete information about types is revealed: $m(\theta) = \theta$ for all $\theta \in \Theta$.

Proof: Since the government has full commitment, Lemma 6 follows directly from the Revelation Principle.[12] □

[12] If contracts were perfectly observable, i.e. the restriction in assumption (A3) would not apply, the simple revelation principle would not apply. Then, agents who are pooled together by the function m would also pool on the financial market so to not reveal any additional information. Thus, by pooling agents together, the government could influence how agents can use the financial market, and implement allocations with non-smooth consumption for pooled agents. However, even in that case it turns out to be optimal for the government to choose full separation over any partial pooling arrangement, so the result is unchanged.

The government still implements a fully separating allocation. However, the set of *modified* incentive constraints (1.16) implies that it cannot achieve the allocation outlined in Lemma 3. Due to the fact that it cannot use the smoothing of consumption as an incentive anymore, less redistribution across the population is implemented at the optimum.

Lemma 7 (Optimal Allocation with Commitment and Financial Market) *At the optimal allocation with commitment, when agents have access to a financial market:*

(i) There is partial social insurance, but less than without the financial market: Total consumption is increasing in type, more than in Lemma 3, but less than under autarky:

$$0 < \frac{\partial(c_1^N(\theta) + c_2^N(\theta))}{\partial\theta} < \frac{\partial(c_1(\theta) + c_2(\theta))}{\partial\theta} < 3$$

where N denotes the allocation derived in Lemma 3 without the market.

(ii) All agents smooth consumption perfectly over time: $c_1(\theta) = c_2(\theta)\ \forall\theta$.

Proof: Given how agents contract in the financial market, the government essentially faces the extra constraints

$$\theta + T_1(\theta) = 2\theta + T_2(\theta) \quad \forall\theta$$

Except for the highest type $\bar{\theta}$, this constraint changes the allocation the government would have liked to implement. Incentives for truthful revelation can now only be given by higher total consumption, using the degree of smoothness as incentive is not an option anymore. Redistributing across agents thus becomes more expensive, less social insurance is possible. □

Notice that the equilibrium is not unique: Since the government can commit to the allocation, all agents will be indifferent as to whether or not they borrow in the financial market, as long as they receive smooth net-of-tax income. Because of the

ex-ante commitment, it is irrelevant for the government, whether agents have entered long-term consumption commitments. It is only the allocative aspect of agents' contracts that impacts the governments ability to redistribute across society. The implemented consumption allocation, however, is the same in all equilibria. I mark the resulting allocation with superscript c for future reference.

This leads to the following proposition:

Proposition 1 (Government with Commitment and Financial Market) *A government that can commit to an allocation ex-ante does not benefit from agents having access to a financial market.*

Proof: Without individual access to financial markets, the government still had the technology to provide perfectly smooth consumption for all agents. Yet, it optimally chose not to do so. Thus, the extra constraint reduces overall welfare, the allocation is inferior to the allocation of Lemma 3 from the government's point of view. Assumption (A4) implies that the government cannot deter agents from using the market by announcing punishments for doing to.□

The presence of markets reduces the set of policy instruments available to the government. When agents can costlessly take care of individual consumption smoothing, the resulting allocation does not allow the government to implement the desired trade-off between redistribution over time and redistribution across the population. Instead, perfect smoothing over time, but less social insurance will be realized.

1.5.2 Government without commitment

For a government that cannot commit to a second period schedule $\{T_2(m)\}$ ex-ante, individual access to a financial market interferes with its optimization problem in the same way as if it had commitment. But there is an additional effect: the lack of commitment constraint (1.8) is modified as well. When deciding about redistributive policy after information has been revealed, the government has to take into account

the continuation value of the debt contracts agents hold. Redistributing away from an agent who pledged all his income results in a costly default. Thus, for a government without commitment, the optimal mechanism solves the same problem as above, but subject also to the *modified* commitment constraint

$$
\begin{aligned}
\{T_2(m)\} \in \arg\max_{\{\hat{T}_2\}} \int_\Theta & \qquad (1.17)\\
u(2\theta + \hat{T}_2(m(\theta)) + h_2(\theta) - b_2(\theta)) \quad & \mathbb{I}\{2\theta + \hat{T}_2(m(\theta)) \geq b_2(\theta)\} \quad + \\
u(2\theta + \hat{T}_2(m(\theta)) - (d_1(\theta) + H_B) - H_A) \quad & \mathbb{I}\{2\theta + \hat{T}_2(m(\theta)) < b_2(\theta)\} \quad dF(\theta) \\
\text{s.t.} \int_\Theta T_1(m(\theta)) + \hat{T}_2(m(\theta)) dF(\theta) \leq 0 &
\end{aligned}
$$

taking the contracts agents can sign as given. Notice that both forms of default costs enter in the same way into the consideration: They both determine the continuation value of the contract the agent signed, and the government has to take potential losses that may result from default into account.

Suppose all agents sign a *gross* debt contract, i.e. $b_2(\theta) = 2\theta + T_2(m(\theta))$ for all types. In this case, any deviation from the previously announced allocation will lead to default. Given the enforcement power of the financial market, agents who are forced to default will have to be at least provided with a payment that covers $d_1 + H$ to avoid zero consumption. The above problem de-facto reduces to redistributing based on what agents planned to consume in the second period, h_2:

$$
\begin{aligned}
&\max_{\{\hat{x}_2(\theta)\}} \int_\Theta u(\hat{x}_2(m(\theta))) dF(\theta) \\
&\text{s.t.} \int_\Theta \hat{x}_2(m(\theta)) dF(\theta) \leq \int_\Theta [h_2(\theta) - H\mathbb{I}\{\hat{x}_2(m(\theta)) < h_2(\theta)\}] dF(\theta)
\end{aligned}
\qquad (1.18)
$$

This formulation of the government's problem at the beginning of period 2 nicely depicts the main point: the government is still free to redistribute, but doing so is costly. The default cost $H = H_A + H_B$ conceptually enters only on the resource side

of the feasibility constraint.[13]

From here it is also immediate that a large enough H would prevent any deviation from the promised schedule:

Proposition 2 (Limit Case: Full Commitment)
When all agents pledge their complete income in the financial market and if $H \geq \frac{1}{2}(3\bar{\theta} + T_1^c(\bar{\theta}) + T_2^c(\bar{\theta}))$, *the government can implement the same allocation as if it had full commitment.*

Proof: Suppose the government had promised the full commitment schedule $\{T_2^c(\theta)\}_\Theta$. The highest type $\bar{\theta}$ will accordingly plan to consume $h_2^c(\bar{\theta}) = \frac{1}{2}(3\bar{\theta} + T_1^c(\bar{\theta}) + T_2^c(\bar{\theta}))$ in period 2. In case he has to default, the default cost is more than what he actually planned to consume. The government would gain no resources for redistribution from letting even the highest type default, and thus would never attempt any redistribution ex-post. It can therefore implement the same allocation as if it had full commitment ex-ante. □

The result of Proposition 2 should be understood as a limit result: If default costs are so high that they leave no value after renegotiation, it obviously serves as a device for full commitment. Such high default costs are unrealistic, they could be interpreted as not offering default as an option. It is interesting, however, that a finite default cost would be enough to induce full commitment. The more relevant case, though, is one where default costs are too low to offer full commitment:

Proposition 3 (Information Revelation: Limited Commitment)
If the government has no ex-ante commitment, but all agents pledge their complete income in the financial market and the default costs H *are positive, the optimal information revelation rule is such that it pools agents above a cutoff type* $\tilde{\theta}$ *together*

[13]Farhi and Werning (2008) consider a government that faces an exogenous cost of reform. The analysis in the present paper can be interpreted as providing one possible micro-foundation for such a reform cost and showing how it leads to limited commitment.

but separates all other types:

$$m(\theta) = \theta \qquad \forall \theta \leq \tilde{\theta}$$
$$m(\theta) = \tilde{\theta} \qquad \forall \theta > \tilde{\theta}$$

The cutoff $\tilde{\theta}$ and transfers T must be such that

$$u'(h_2(\underline{\theta}))(h_2(\tilde{\theta}) - h_2(\underline{\theta}) - H) = u(h_2(\tilde{\theta})) - u(h_2(\underline{\theta})) \tag{1.19}$$

where $h_2(\theta) = \frac{1}{2}(\theta + T_1(m(\theta)) + T_2(m(\theta)))$.

Proof: See appendix 1.7.3.

The proposition states that a government without commitment gains the power to commit to a partially separating allocation if agents hold gross financial contracts, as long as they face positive default costs. How much separation is possible, or in other words how many types at the top of the distribution will pool, depends on the default cost H and on the concavity of the utility function. The intuition for the constraint is simple: the marginal benefit from deviating from the promised allocation (on the left hand side) is measured by the marginal utility of the lowest type (since he is the one distributed toward) times the amount of resources available for redistribution. The marginal cost of such deviation (on the right hand side) is the utility loss of the highest type: his consumption is equalized with that of the lowest type.

More separation of types, i.e. a higher $\tilde{\theta}$, leads to a larger differentiation in period 2 consumption $h_2(\tilde{\theta}) - h_2(\underline{\theta})$ (due to the incentive constraints in the ex-ante optimization problem), and in turn to a tightening of the constraint. A higher default cost H on the other hand relaxes the constraint, so that more separation is possible. In fact, when $H \geq h_2^c(\bar{\theta})$, the government will be able to commit to the same allocation as the full commitment government (Proposition 2). Proposition 3 states, however, that any positive default cost, even a very small one, allows the government at least some

commitment. Notice also that the degree of possible separation is negatively linked to the concavity of the utility function. If $u(\cdot)$ is more concave, the ex-post gain from redistribution increases.[14] In order to be able to withstand the higher temptation to let a fraction of agents at the top default, more agents have to be pooled together, the cutoff $\tilde{\theta}$ has to be lower.

When $H = 0$, the case of no commitment is recovered: As long as the utility function is strictly concave, condition (1.19) is only satisfied when $h_2(\tilde{\theta}) = h_2(\underline{\theta})$, i.e. when there is no separation at all. As in in the benchmark case without a financial market, when the government is not able to implement any separation in the second period, it cannot provide enough incentives through first period transfers to implement any social redistribution from the top to the bottom of the distribution.

The default costs essentially serve as a commitment device for the government. With any positive default costs, the government can gain limited commitment: It can credibly commit to not exploit a limited amount of information. By pooling agents at the top of the distribution together, only limited information is revealed: For agents of type $\theta \leq \tilde{\theta}$ the true type is revealed, while for all agents of type $\theta > \tilde{\theta}$ the government only learns that they are part of the high income group, but not their exact type.

The following lemma summarizes the characteristics of the best allocation the government is able to commit to:

Lemma 8 (Optimal Allocation without Commitment and Financial Market) *At the optimal allocation without commitment, when agents have access to a financial market, the default cost is* $0 < H < h_2^c(\bar{\theta})$ *and the conditions in Proposition 3 are met:*

(i) There is partial social insurance - total consumption increases in type, but less

[14] Because I study the problem of a government with a utilitarian objective with equal Pareto weights on all agents, the concavity of the individual utility function also measures the potential welfare gain from redistribution. More generally, the form of the government's objective function is the crucial characteristic to determine the optimal cutoff $\tilde{\theta}$.

steeply than under autarky:

$$0 < \frac{\partial(c_1(\theta) + c_2(\theta))}{\partial\theta} < 3$$

(ii) All agents smooth consumption perfectly over time: $c_1(\theta) = c_2(\theta)\ \forall\theta$.

Proof: All agents who report their type truthfully are able to use the financial market for perfect consumption smoothing. This provides incentives for agents to reveal themselves. As in the Lemma 3, the government uses the information gained about types to implement some social insurance. □

The government, even though per se not able to commit to an allocation ex-ante, is able to implement some redistribution across agents, i.e. it can provide some level of social insurance. This leads directly to the following proposition:

Proposition 4 (Government without Commitment and Financial Market) *If default costs H are positive, a government that cannot commit to an allocation ex-ante always benefits from agents having access to financial markets.*

Proof: Proposition 3 establishes that with positive default costs H the government is able to commit to a partially separating allocation that provides some social insurance - an improvement over the pooling allocation in Lemma 5 without financial markets. However, contrary to the case in which the government has commitment ex-ante, here the commitment power hinges critically on agents actually borrowing in the financial market. It remains to be shown that agents will indeed sign *gross* contracts in the financial market. Since agents are small, non-strategic players in this policy game, it would be a stretch to assume that they coordinate on signing such contracts in order to provide a commitment device for the government. Notice, however, that once a schedule is announced and types have been reported, i.e. at the stage of choosing a debt contract, it is individually optimal to sign a gross contract and enter a consumption commitment for period 2:

At the beginning of period 2, the last stage of the game, the government without an exogenous commitment device chooses the transfer schedule to solve problem (1.17), knowing what contract each agent has signed. The government would like to equalize consumption as much as possible. If an agents has signed a contract such that

$$b_2(\theta) < m(\theta) + T_2(m(\theta))$$

it is costless and thus optimal for the government to redistribute the unpledged portion of promised after-tax income away from that agent, and redistribute it toward lower types. This leaves the agents worse off than if he had signed a gross contract. However, regardless of whether or not all other agents signed gross contracts, the government will not find it optimal to let those who did default. Thus, it is a dominant strategy for agents to sign such gross contracts. □

The role of the financial market is to give agents the opportunity to pledge their expected resources in debt contracts which in turn influences the government's ability to commit at least to a partially separating allocation. Therefore, it also facilitates redistribution across the population. This mechanism is the central insight of this paper. The crucial characteristic of contracts in the financial market is not that agents are free to allocate resources. It is that in order to allocate resources, agents are able to enter consumption commitments that cannot costlessly be changed.

Additional Assumption: Horizontal Equity

The key insight of this paper is that the prospect of agents defaulting on their loans enables the government to commit to some separation even after it learned agents' private information. In the previous section I derived the specific form of separation at the optimal allocation: Agents below a certain cutoff are perfectly separated, while agents above that cutoff pool. Thus, in this specific setup, the commitment comes from the fact that agents at the top of the income distribution would default on their loans, if the government decided to redistribute across the population more than it

announced ex-ante.

In this section, I derive the form of partial pooling at the optimal allocation under the additional assumption of horizontal equity. It means that the government is bound to treat equal agents equally, i.e. it cannot randomize transfers between seemingly equal agents. While there is nothing in the model that necessitates this assumption, it has some realistic appeal.

Proposition 5 (Information Revelation under Horizontal Equity)
If the government has no ex-ante commitment, all agents pledge their complete income in the financial market and the default cost is $0 < H < h_2^c(\bar{\theta})$*, the optimal information revelation rule is such that it pools agents into finitely many groups throughout the type distribution. Any* $H > 0$ *allows separation into at least two groups.*

Proof: See appendix 1.7.4.

If the government is bound to horizontal equity, it is not optimal anymore to pool agents just at the top of the distribution. In fact, since the government cannot let just a fraction of any pooled group of agents default, less pooling at the top is necessary, more information about the highest types can be revealed and used to provide social insurance.

At the optimum, agents will be pooled throughout the distribution in groups of varying size - a structure that can be interpreted as tax brackets. The effective commitment power of the government then stems from the concern of agents defaulting on their loan throughout the distribution, not just at the top.

1.6 Discussion

This paper uncovers a mechanism by which the presence of a financial market may effectively provide the government with a (limited) commitment device, thereby enabling the implementation of commitment-type policies. It thus helps reconcile the

observation of policies that are suggestive of governments being able to commit, even though there is no apparent commitment device. Moreover, the model provides a rational for why governments do not implement policies contingent on complete information: When they have no ex-ante commitment power, a reasonable default cost provides them with *limited* commitment, i.e. with the power to commit to not exploit a limited amount of information.

In the presence of private information, the ability of a government to implement social redistribution crucially depends on its power to commit to future policy. In reality, there is little reason to believe that governments possess some exogenous commitment device. Instead, commitment must stem from the environment the government operates in. The literature has focused on political economy constraints as mechanisms for commitment. In contrast to that, the presented paper highlights the fact that also the *economic* environment might enhance the commitment power of the government. In that sense, the paper establishes a theoretical foundation for what can be referred to as a social market economy, where the presence of well functioning competitive markets that allow agents to enter consumption commitments plays a crucial role for social redistribution.

People typically do not just rely on the government and simply consume their net-of-tax income every period. Instead, they use private financial markets to allocate their resources over time. Redistributive policy ought to take that into account. To address the question how the presence of a financial market shapes the government's ability to implement redistributive policy, I studied a standard Mirrlees framework. In the presented economy agents receive heterogeneous income, and a benevolent government might attempt two forms of redistribution: Smoothing of individual consumption over time and social redistribution. Private information about income types, however, introduces a trade-off between the two.

If the government can commit to future policy ex-ante, it is able to implement a

fully separating allocation. In this case, agents having access to a financial market reduces the set of policy instruments available to the government: It loses the ability to provide incentives for truthful revelation through the degree of consumption smoothness. Agents can use the financial market to smooth consumption by themselves. The government is deprived of the power to discriminate along this dimension, extracting private information from the agents gets harder. Consequently, a government with full commitment cannot gain from agents' involvement in financial markets. It finds itself unable to implement the constrained optimal trade-off between allocative efficiency and equity. In such a setup, financial markets hinder redistribution across the population.

However, if the government cannot commit to an allocation ex-ante and is thus unable to implement social redistribution, it might gain from agents' involvement in financial markets. In fact, it might gain the power to commit at least to a partially separating allocation, making some social redistribution possible. The reason is that in order to smooth their income over time, agents pledge future income in the financial market in contracts that induce long-term consumption commitments. Such private contracts are typically not easily reversible. A surprise redistribution, after agents have revealed their type and signed individual debt contracts, will lead to some agents having to default on their debt. Such default, however, is costly. These default costs mitigate the government's desire to exploit information and implement full social insurance ex-post.

The costs may be of very diverse nature: On the one hand, one might think of pure resource costs for administering the default on a loan. For banks, re-allocating funds from one loan to another is also costly. A bank might not be able to resell an asset for the same value agreed upon previously with the now defaulting agent. On the other hand, agents might suffer a loss in utility when they have to default on their loan in addition to the resource costs of the bank. They may have made life plans contingent on this loan that require further costly alteration. They might, for example, have

grown attached to their house, which they financed with a mortgage, and lose utility when they have to move. I argued that any such costs alter the government's ex-post optimization problem, since it has to take the continuation value of agents' contracts into account. Agents' involvement in a financial market thus effectively provides a commitment device for the government: Even though it has the ability to re-optimize its policy over time, it does not find it useful to do so at any point.

How much commitment is possible depends on the size of the default costs. I derived a simple condition that links the size of the default costs and the concavity of the utility function to the degree of separation a government is able to commit to. I show that whenever the default costs are positive, some separation can persist after information is revealed: The government will optimally pool agents at the top of the type distribution together and separate all other types perfectly. This allows for some social insurance to be implemented and is thus a strict improvement on the no-commitment equilibrium with complete pooling and no social insurance.

The intuition for the constraint is simple: it equates the marginal benefit from deviating from the promised allocation (as measured by the marginal utility of the lowest type who would be distributed toward times the amount of resources available for redistribution) with the marginal cost of such deviation (the utility loss of the highest type who would have to default). For a given functional form of utility, the higher the default costs, the more separation can be implemented. Conversely, for given default costs, the more concave the utility function is, the higher would be the ex-post welfare gain from redistribution. The government would be more tempted to deviate from announced policy ex-post, and so is able to commit only to less separation ex-ante.

The particular form of optimal information revelation changes when the government is restricted to horizontal equity, i.e. if it cannot randomize transfers between seemingly equal agents. This constraint exogenously mitigates the commitment problem, and while the allocation with pooling above a threshold is still feasible, it is not op-

timal anymore. Instead, it turns out optimal for the government to pool agents into finitely many groups throughout the type distribution. The resulting transfer schedule then has a "tax bracket" structure. The government collects coarse information about agents' types over the complete distribution. The higher the default costs are, the more and the smaller these brackets can be, so that finer information can be collected.

References

Acemoglu, D., Golosov, M. and Tsyvinski, A.: 2008a, Markets versus governments, *Journal of Monetary Economics* **55**, 159–189.

Acemoglu, D., Golosov, M. and Tsyvinski, A.: 2008b, Political economy of mechanisms, *Econometrica* **76**, 619–641.

Acemoglu, D., Golosov, M. and Tsyvinski, A.: 2010, Dynamic mirrlees taxation under political economy constraints, *Review of Economic Studies* **77**(3), 841–881.

Bester, H. and Strausz, R.: 2001, Contracting with imperfect commitment and the revelation principle: The single agent case, *Econometrica* **69**, 1077–1098.

Bisin, A. and Rampini, A.: 2006, Markets as beneficial constraints on the government, *Journal of Public Economics* **90**, 601–629.

Chetty, R. and Szeidl, A.: 2007, Consumption commitments and risk preferences, *Quarterly Journal of Economics* **122**(2), 831–877.

Farhi, E. and Werning, I.: 2008, The political economy of nonlinear capital taxation, *Working Paper* .

Fudenberg, D. and Tirole, J.: 1991, *Game Theory*, MIT Press, Cambridge.

Golosov, M., Tsyvinski, A. and Werning, I.: 2006, New dynamic public finance: A user's guide, *NBER Macroeconomic Annual* .

Golosov, M. and Tsyvinsky, A.: 2007, Optimal taxation with endogenous insurance markets, *Quarterly Journal of Economics* **122**(2), 487–534.

Hammond, P.: 1987, Markets as constraints: multilateral incentive compatibility in continuum economies, *Review of Economic Studies* **54**, 399–412.

Harris, M. and Townsend, R. M.: 1981, Resource allocation under asymmetric information, *Econometrica* **49**, 33–69.

Kydland, F. E. and Prescott, E. C.: 1977, Rules rather than discretion: the inconsistency of optimal plans, *Journal of Political Economy* **85**, 473–91.

Mirrlees, J. A.: 1971, An exploration in the theory of optimum income taxation, *review of Economic Studies* **38**, 175–208.

Myerson, R.: 1979, Incentive compatibility and the bargaining problem, *Econometrica* **47**, 61–73.

Netzer, N. and Scheuer, F.: 2010, Competitive markets without commitment, *Working Paper* .

Roberts, K.: 1984, The theoretical limits to redistribution, *Review of Economic Studies* **51**, 177–195.

Scheuer, F.: 2010, Entrepreneurial taxation, occupational choice, and credit market frictions, *Working Paper* .

Skreta, V.: 2007, Sequential optimal mechanisms, *Review of Economic Studies* **73**, 1085–1111.

Skreta, V.: 2010, Optimal auction design under non-commitment, *Working Paper* .

Townsend, R. M.: 1982, Optimal multiperiod contracts and the gain from enduring relationships under private information, *Journal of Political Economy* **90**(6), 1166–86.

1.7 Appendix

1.7.1 Proof of Lemma 3

First, notice that the first-best allocation is not incentive compatible: From an agent's point of view, his consumption allocation $x_1 = x_2 = x$ is fixed, no matter what type he reports. He then chooses to report type $\hat{\theta}$ to solve

$$\max_{\hat{\theta}} u(x + (\theta - \hat{\theta})) + u(x + 2(\theta - \hat{\theta}))$$

Since utility is time-separable and per period utility is strictly increasing, first and second period consumption are not perfect complements. Thus, every type has an incentive to hide income from the government, thereby receiving the same allocation of consumption as under truth-telling x plus the extra hidden income $t(\theta - \hat{\theta})$. Each agent then optimally chooses tho report the lowest possible type $\underline{\theta}$. Full social insurance *and* perfect smoothing cannot be implemented.

Consider next the allocation with perfect smoothing over time for all types and no redistribution across agents, i.e. $x_1(\theta) = x_2(\theta) = \frac{3}{2}\theta$. This allocation is incentive compatible: The agent solves

$$\max_{\hat{\theta}} u(\theta + \frac{1}{2}\hat{\theta}) + u(2\theta - \frac{1}{2}\hat{\theta})$$

Optimality requires

$$\frac{\partial}{\partial \hat{\theta}} = \frac{1}{2}(u'(\theta + \frac{1}{2}\hat{\theta}) - u'(2\theta - \frac{1}{2}\hat{\theta})) = 0 \tag{1.20}$$

$$u'(\theta + \frac{1}{2}\hat{\theta}) = u'(2\theta - \frac{1}{2}\hat{\theta}) \tag{1.21}$$

$$\rightarrow \hat{\theta} = \theta \tag{1.22}$$

The last step follows because $u(\cdot)$ is strictly concave. At this allocation, per period consumption $x_t(\theta) = t\theta + T_t(\theta)$ increases with slope $\frac{3}{2}$. The concavity of the utility

function implies that it is strictly optimal for all agents report the true type. This means that the incentive constraints are not binding for any type. Thus, there is room for welfare increasing redistribution across agents. It follows directly that total consumption will be increasing less than under autarky, i.e.

$$\frac{\partial(x_1(\theta) + x_2(\theta))}{\partial\theta} < 3$$

Next, I will derive the properties of the optimal allocation that result from such redistribution. Redistributing across agents from top to bottom requires that the sum of transfers $T_1(\theta) + T_2(\theta)$ should be decreasing in type, i.e.

$$\frac{\partial(T_1(\theta) + T_2(\theta))}{\partial\theta} = T_1'(\theta) + T_2'(\theta) < 0 \leftrightarrow -\frac{T_2'(\theta)}{T_1'(\theta)} > 1 \tag{1.23}$$

for all types $\theta < \bar{\theta}$. Just at the highest type, the contribution to the social redistribution system need not be increasing, i.e. $T_1(\bar{\theta}) = -T_2(\bar{\theta})$.

When agents choose which type to report, they solve

$$\max_{\hat{\theta}} u(\theta + T_1(\hat{\theta})) + u(2\theta + T_2(\hat{\theta}))$$

A necessary condition for incentive compatibility thus is that the first order condition of this problem be zero at $\hat{\theta} = \theta$:

$$\frac{u'(\theta + T_1(\theta))}{u'(2\theta + T_2(\theta))} = \frac{-T_2'(\theta)}{T_1'(\theta)} \tag{1.24}$$

First, notice that (1.23) together with (1.24) and concavity of $u(\cdot)$ implies that $x_1(\theta) < x_2(\theta)$ for all types $\theta < \bar{\theta}$, but $x_1(\bar{\theta}) = x_2(\bar{\theta})$. That is, perfect smoothing for the highest type is optimal and smoothing is distorted for all other types.

For (1.24) to also be sufficient for incentive compatibility, it must be the case that

the second order condition for optimality is also satisfied at $\hat{\theta} = \theta$

$$u''(\theta+T_1(\theta))(T_1')^2+u'(\theta+T_1(\theta))T_1''+u''(2\theta+T_2(\theta))(T_2')^2+u'(2\theta+T_2(\hat{\theta}))T_2'' < 0 \quad (1.25)$$

Further differentiating (1.24) yields

$$u''(\theta+T_1(\theta))T_1'x_1'+u'(\theta+T_1(\theta))T_1''+u''(2\theta+T_2(\theta))T_2'x_2'+u'(2\theta+T_2(\hat{\theta}))T_2'' = 0 \quad (1.26)$$

where $x_t(\theta) = t\theta + T_t(\theta)$ and so $x_t'(\theta) = t + T_t'(\theta)$.

Combining (1.25) and (1.26) gives the following monotonicity requirement

$$u''(\theta + T_1(\theta))T_1'x_1' + u''(2\theta + T_2(\theta))T_2'x_2' > u''(\theta + T_1(\theta))(T_1')^2 + u''(2\theta + T_2(\theta))(T_2')^2 \quad (1.27)$$

which simplifies to

$$u''(\theta + T_1(\theta))T_1' + 2u''(2\theta + T_2(\theta))T_2' > 0 \quad (1.28)$$

A sufficient condition for this to hold is that

$$2 > \frac{u''(x_1)}{u''(x_2)} \quad (1.29)$$

which due to CRRA implies

$$x_1 > \frac{1}{2}x_2 \quad (1.30)$$

Autarky implies $x_1 = \frac{1}{2}x_2$, so that this condition is met when smoothness of consumption is increased for all agents. Thus, the full set of IC constraints can be replaced by the local incentive constraints (1.24) and the requirement that $x_1 > \frac{1}{2}x_2$.

The government's problem then is to solve

$$\max_{\{T_1,T_2\}} \int_{\underline{\theta}}^{\bar{\theta}} u(\theta + T_1(\theta)) + u(2\theta + T_2(\theta))$$

$$\text{s.t.} \int_{\Theta} T_1(\theta) + T_2(\theta) \leq 0$$

$$u'(\theta + T_1(\theta))T_1'(\theta) + u'(2\theta + T_2(\theta))T_2'(\theta) = 0 \quad \forall\theta$$

The first order conditions to this problem yield the following optimality condition:

$$u'(\theta + T_1(\theta)) - u'(2\theta + T_2(\theta)) = \gamma(\theta)(u''(\theta + T_1(\theta))T_1'(\theta) + u''(2\theta + T_2(\theta))T_2'(\theta))$$

where $\gamma(\theta)$ are the Lagrange multipliers on the incentive compatibility constraints. From this condition it follows that when $x_1(\theta) < x_2(\theta)$

$$u''(\theta + T_1(\theta))T_1'(\theta) + u''(2\theta + T_2(\theta))T_2'(\theta) < 0 \tag{1.31}$$

CRRA implies that

$$\frac{x_2}{x_1} = \frac{u''(x_1)}{u''(x_2)}\frac{u'(x_2)}{u'(x_1)} \tag{1.32}$$

so that

$$x_1(\theta) < x_2(\theta) \rightarrow \frac{x_2(\theta)}{x_1(\theta)} > \frac{u'(x_2(\theta))}{u'(x_1(\theta))} \tag{1.33}$$

Moreover, note that

$$\frac{x_2(\theta)}{x_1(\theta)} = \frac{2\theta + T_2(\theta)}{\theta + T_1(\theta)} \leftrightarrow \frac{x_2'(\theta)}{x_1'(\theta)} = \frac{2 + T_2'(\theta)}{1 + T_1'(\theta)} \tag{1.34}$$

We would like to show that the degree of smoothness as measured by the ratio $\frac{x_1}{x_2}$ is increasing in type, i.e

$$\frac{\partial \frac{x_1(\theta)}{x_2(\theta)}}{\partial\theta} = \frac{x_1'(\theta)x_2(\theta) - x_1(\theta)x_2'(\theta)}{(x_2(\theta))^2} > 0 \leftrightarrow \frac{x_2(\theta)}{x_1(\theta)} > \frac{x_2'(\theta)}{x_1'(\theta)} \tag{1.35}$$

Combining optimality (1.31), CRRA (1.33), and (1.34) with (1.24) and (1.28) implies that (1.35) holds, and thus the degree of consumption smoothness increases with type. This concludes the proof.

1.7.2 Proof of Lemma 4

First, suppose the information revelation rule was such that all information reported by the agents would be revealed to the government, i.e. $m(\theta) = \theta$ for all types θ. Constraint (1.8) implies that if the government possesses any information about types at the beginning of the second period, it will exploit it so to equalize consumption as much as possible. To see that, consider the first order conditions of the government's problem (1.8) at t=2:

$$u'(2\theta + \hat{T}_2(\theta)) - \lambda = 0 \quad \forall\theta$$

These conditions imply that the government will choose $\{\hat{T}_2\}$ so to equalize consumption across all agents, $x_2(\theta) = x_2\ \forall\theta$. From the agent's point of view then the consumption allocation in period 2 is fixed, and he solves:

$$\max_{\hat{\theta}} u(x_1(\hat{\theta}) + \theta - \hat{\theta}) + u(x_2 + 2(\theta - \hat{\theta}))$$

For truth-telling to be optimal, it is necessary that the first and second order conditions are satisfied at $\hat{\theta} = \theta$, i.e. $\forall\theta$:

$$(x_1'(\theta) - 1)u'(x_1(\theta)) - 2u'(x_2) = 0 \tag{1.36}$$

$$(x_1'(\theta) - 1)^2 u''(x_1(\theta)) + x_1''(\theta)u'(x_1(\theta)) + 4u''(x_2) < 0 \tag{1.37}$$

Further differentiating (1.36) yields

$$x_1'(\theta)(x_1'(\theta) - 1)u''(x_1(\theta)) + x_1''(\theta)u'(x_1(\theta)) = 0$$

which reduces (1.37) to

$$-(x_1'(\theta)-1)u''(x_1(\theta))+4u''(x_2)<0 \tag{1.38}$$

This, together with (1.36) implies that for the allocation to be incentive compatible, it must be such that $\forall\theta$

$$\begin{aligned} -\frac{u''(x_1(\theta))}{u'(x_1(\theta))} &< -2\frac{u''(x_2)}{u'(x_2)} \\ \leftrightarrow -\frac{u''(x_1(\theta))}{u'(x_1(\theta))}x_1(\theta)x_2 &< -2\frac{u''(x_2)}{u'(x_2)}x_2x_1(\theta) \\ \leftrightarrow x_2\frac{1}{\epsilon} &< 2x_1(\theta)\frac{1}{\epsilon} \end{aligned}$$

where ϵ is the elasticity of intertemporal substitution, which is constant by assumption. Thus, it must be true for all types that

$$x_1(\theta)>\frac{1}{2}x_2 \tag{1.39}$$

Moreover, (1.36) can be rearranged as

$$x_1'(\theta)=2\frac{u'(x_2)}{u'(x_1(\theta))}+1$$

This differential equation determines the shape of the consumption schedule in period 1. Two properties are important: $x_1(\theta)$ is *increasing* in type, with a slope strictly larger than 1, and with *increasing slope*. The lowest type, $\underline{\theta}$ will receive the lowest period 1 consumption. To relax incentive constraints for the higher types, it is optimal to start from the lowest possible $x_1(\underline{\theta})$. A lower bound is $x_1(\underline{\theta})=\frac{1}{2}x_2$. What is x_2?

$$x_2=2\int_{\underline{\theta}}^{\bar{\theta}}\theta dF(\theta)-\int_{\underline{\theta}}^{\bar{\theta}}x_1(\theta)dF(\theta) \tag{1.40}$$

The second summand cannot be solved without further assumptions on the utility function. But we can use a conservative lower bound to see what the government

would at most be able to achieve with a fully separating allocation. To that end, suppose we ignore that $x_1(\theta)$ has to be increasing with *increasing* slope, and rather assume that it will increase with *constant* slope $x_1'(\underline{\theta}) \approx 2$. This is not a bad approximation, since constant elasticity of intertemporal substitution implies $u'''(\cdot) < 0$ and so $\frac{u'(x_2)}{u'(\frac{1}{2}x_2)} \geq \frac{1}{2}$ is not a terrible assumption. This lower bound allows to compute an upper bound on x_2:

$$x_2 \leq 2 \int_{\underline{\theta}}^{\bar{\theta}} \theta dF(\theta) - \int_{\underline{\theta}}^{\bar{\theta}} \frac{1}{2}x_2 + 2(\theta - \underline{\theta})dF(\theta) \tag{1.41}$$

$$\leftrightarrow x_2 \leq \frac{4}{3}\underline{\theta} \tag{1.42}$$

This leaves the lowest type at best with the consumption allocation $[\frac{2}{3}\underline{\theta}, \frac{4}{3}\underline{\theta}]$. Notice that this means he is distributed *away* from in the aggregate and also doesn't gain any smoothness. This cannot be optimal from a social welfare point of view. It means that the only separating allocation that can be implemented is one that increases inequality and lowers welfare compared to autarky, and thus it is not optimal.

Notice that the argument does not change when the government learns only partial information about types. Since the second period allocation is fixed, providing incentives for any separation through first period transfers is so costly that it is not optimal to do so. Thus, the optimal information revelation rule is one where no information is revealed, i.e.

$$m(\theta) = \underline{\theta} \quad \forall \theta$$

This concludes the proof.

1.7.3 Proof of Proposition 3

The proposition states that a government without commitment is able to implement an allocation with at least partial separation, if all agents pledge their complete income in the financial market. Strictly positive default costs alter the ex-post problem

of the government: it might even after the revelation of information not have an incentive to redistribute fully, because this would lead to costly default by agents who are redistributed away from.

The proof proceeds as follows: In a first step I will establish the optimal form of the information revelation rule. It turns out to be optimal that agents with income types above a threshold $\tilde{\theta}$ are pooled together, while all agents below the cutoff are completely separated. The second step derives the optimal cutoff type, dependent on the the size of the default costs H and the concavity of the utility function.

First notice that the following Lemma holds:

Lemma 9

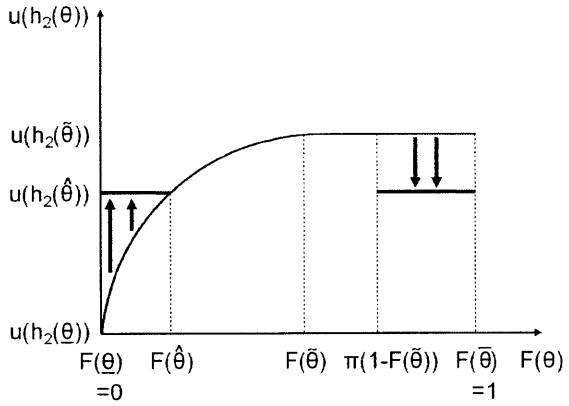
If the government wants to default, it will always default on the highest (observed) type first. Even if the density of highest types is large (e.g. due to pooling at the top), it prefers to randomize and default on some of the highest types rather than on lower types.

Proof: Since ex-ante incentive compatibility implies that the promised allocation in $t = 2$ is increasing in type, the gains from redistributing ex-post are highest when letting the highest types default. The default costs H, on the other hand, are constant per default. □

Suppose agents above some cutoff $\tilde{\theta}$ are pooled together. Even if it is not optimal to default on all of them, it might still be profitable for the government to default only on a fraction π of them. The reason is that neither the gained resources nor the gain in welfare from redistributing these resources are are linear in π. The resources saved are optimally distributed toward the lowest types. Thus, the gain is the highest for the first redistributed dollars and decreases thereafter.

Let $\hat{\theta}$ denote the type below which agents get extra resources when the government lets a fraction π of agents above the cutoff $\tilde{\theta}$ default. The following graph clarifies

notation:



The resources gained for redistribution are

$$\pi(1 - F(\tilde{\theta}))(h_2(\tilde{\theta}) - H - x_2(\hat{\theta})) \tag{1.43}$$

This is because the types who are forced to default will receive the same allocation $x_2(\hat{\theta})$ as the agents at the bottom of the distribution who are distributed toward. $\hat{\theta}$ is a function of the resources gained, and so a function of π so that the gain is generally not linear in π.

The redistribution causes a loss of utility for the defaulting agents at the top, which for the same argument is nonlinear in π:

$$\pi(1 - F(\tilde{\theta}))(u(h_2(\tilde{\theta})) - u(x_2(\hat{\theta}))) \tag{1.44}$$

The gain in welfare comes from the utility gain for the types at the low end of the distribution, below $\hat{\theta}$:

$$F(\hat{\theta})u(x_2(\hat{\theta})) - \int_{\underline{\theta}}^{\hat{\theta}} h_2(\theta)dF(\theta) \tag{1.45}$$

The new consumption level $x_2(\hat{\theta})$ is derived by distributing resources equally between the defaulting high type group and the low type group:

$$x_2(\hat{\theta}) = \frac{\int_{\underline{\theta}}^{\hat{\theta}} h_2(\theta)dF(\theta) - \pi(1 - F(\tilde{\theta}))(h_2(\tilde{\theta}) - H)}{F(\hat{\theta}) + \pi(1 - F(\tilde{\theta}))} \tag{1.46}$$

The net gain from letting a fraction π of agents in the pooled group at the top of the distribution default thus is nonlinear in π. This makes it possible that the government might optimally choose to randomize between seemingly equal agents instead of defaulting on all of them.

Since all agents can smooth their consumption perfectly regardless of whether they are pooled together with other types, the government will choose as much separation as possible to gain as much information as it can commit not to exploit. Thus, it will choose to separate all agents below the cutoff. This establishes the optimal form of the information revelation rule:

$$\begin{aligned} m(\theta) &= \theta & \forall\theta \leq \tilde{\theta} \\ m(\theta) &= \tilde{\theta} & \forall\theta > \tilde{\theta} \end{aligned}$$

All agents below the cutoff are asked for precise information, all types above the cutoff can truthfully only report the same income.

The second step of the proof involves finding the optimal pooling cutoff $\tilde{\theta}$ so that the government will not find it optimal ex-post to let even a few of the pooled agents default, and so consequently does not find it optimal to let anyone default.

Given a promised allocation with pooling at the top, the government will choose

the optimal fraction π of default on the pooled group of agents according to

$$\max_{\pi}[F(\hat{\theta}) + \pi(1 - F(\tilde{\theta}))]u(h_2(\hat{\theta})) - \int_{\underline{\theta}}^{\hat{\theta}} u(h_2(\theta))dF(\theta) - \pi(1 - F(\tilde{\theta}))u(h_2(\tilde{\theta}))$$

$$\text{s.t. } h_2(\hat{\theta}) = \frac{\left[\int_{\underline{\theta}}^{\hat{\theta}}(h_2(\theta))dF(\theta) + \pi(1 - F(\tilde{\theta}))(h_2(\tilde{\theta}) - H)\right]}{(F(\hat{\theta}) + \pi(1 - F(\tilde{\theta})))} \tag{1.47}$$

$$0 \leq \pi \leq 1 \tag{1.48}$$

The problem states that the government maximizes the welfare gain from defaulting on a fraction π of the pooled agents subject to how many agents can be provided with higher consumption depending on the resources saved due to not paying out the promised high income to the high types. As introduced earlier, I denote with $\hat{\theta}$ the cutoff below which agents are better off after the redistribution. The government wants to distribute the saved resources to the low types such that it makes optimal use of the highest marginal utilities of more consumption. As a result, all agents up to type $\hat{\theta}$ will get the same consumption as the type $\hat{\theta}$ was promised ex-ante, i.e $x_2(\hat{\theta}) = h_2(\hat{\theta})$. Of course the government will choose to provide the same level of consumption to the agents who were just forced to default. The cutoff $\hat{\theta}$ is obviously endogenous to the choice of π - the constraint (1.47) defines the optimal cutoff implicitly.

The first order condition to this optimization problem, disregarding constraint (1.48) for the moment, is:

$$[F(\hat{\theta}) + \pi(1 - F(\tilde{\theta}))]u'(h_2(\hat{\theta}))\frac{\partial h_2(\hat{\theta})}{\partial \pi} - (1 - F(\tilde{\theta}))[u(h_2(\tilde{\theta})) - u(h_2(\hat{\theta}))] = 0 \tag{1.49}$$

where

$$\frac{\partial h_2(\hat{\theta})}{\partial \pi} = \frac{(1-F(\tilde{\theta}))}{[F(\hat{\theta})+\pi(1-F(\tilde{\theta}))]^2}((h_2(\tilde{\theta})-H)[F(\hat{\theta})+\pi(1-F(\tilde{\theta}))] \\ -[\int_{\underline{\theta}}^{\hat{\theta}} h_2(\theta)dF(\theta)+\pi(1-F(\tilde{\theta}))(h_2(\tilde{\theta}-H))]) \tag{1.50}$$

Note that since $\hat{\theta}$ is always chosen optimally depending on π, by the Envelope Theorem the derivative of $\hat{\theta}$ with respect to π need not be taken into account. First note that since

$$\int_{\underline{\theta}}^{\hat{\theta}} h_2(\theta)dF(\theta) = F(\hat{\theta})h_2(\hat{\theta}) - \pi(1-F(\tilde{\theta}))(h_2(\tilde{\theta}) - h_2(\hat{\theta}) - H)$$

we can rewrite

$$\frac{\partial h_2(\hat{\theta})}{\partial \pi} = \frac{(1-F(\tilde{\theta}))(h_2(\tilde{\theta}) - h_2(\hat{\theta}) - H)}{F(\hat{\theta})+\pi(1-F(\tilde{\theta}))} > 0 \tag{1.51}$$

and so the first derivative simplifies to:

$$\frac{d}{d\pi} = (1-F(\tilde{\theta}))\left[(h_2(\tilde{\theta}) - h_2(\hat{\theta}) - H) - (u(\tilde{\theta}) - u(h_2(\hat{\theta})))\right] \tag{1.52}$$

The second order condition to this problem is always negative:

$$\frac{d^2}{d\pi^2} = (1-F(\tilde{\theta}))u''(h_2(\hat{\theta}))(h_2(\tilde{\theta}) - h_2(\hat{\theta}) - H)\frac{\partial h_2(\hat{\theta})}{\partial \pi} < 0$$

Thus, there is only one optimal default probability π^*. Next, I will derive a condition under which the government will find it optimal to choose $\pi^* = 0$. For $\pi^* = 0$ to be optimal, we need the first derivative (1.52) to be less or equal to zero at $\pi = 0$. Less than zero makes $\pi = 0$ optimal because of the non-negativity constraint (1.48) disregarded before. Setting $\pi = 0$ leads to $\hat{\theta} = \underline{\theta}$. Then evaluating (1.52) at $\pi = 0$, gives the following final condition:

$$u'(h_2(\underline{\theta}))(h_2(\tilde{\theta}) - h_2(\underline{\theta}) - H) \leq u(h_2(\tilde{\theta})) - u(h_2(\underline{\theta})) \tag{1.53}$$

Given H and the functional form of $u(\cdot)$, the government can commit to a schedule $\{h_2(\theta)\}_\Theta$ that pools agents above $\tilde{\theta}$ and satisfies constraint (1.53). In fact, should this condition not bind, less agents can be pooled together, which is preferable for the government. Thus it will always choose $\tilde{\theta}$ such that the condition holds with equality.

It remains to be shown that for any positive H some separation is possible, i.e. there exists a $\tilde{\theta} > \underline{\theta}$ such that condition (1.53) is satisfied. Notice that when there is no separation ($\tilde{\theta} = \underline{\theta}$) and $H > 0$, the condition is always slack:

$$u'(h_2(\underline{\theta}))(-H) < 0 \tag{1.54}$$

Thus, there is room for separation until the constraint binds, as long as $H > 0$. This concludes the proof of Proposition 3.

1.7.4 Proof of Proposition 5

This proof proceeds by analyzing how the additional constraint of horizontal equity changes the optimization problem in Proof 1.7.3.

First, notice that Lemma 9 (in Proof 1.7.3) does not hold anymore. Since the government cannot randomize transfers, once cannot conclude that in case any default is profitable ex-post, it is optimal to let the highest types default. If enough agents at the top are pooled together, it might be that defaulting on all of them is suboptimal, while defaulting on some of them might have given a welfare improvement, but is not allowed anymore. In that case, the government might decide to default on lower type agents, simply because they are not so many. In what follows, I will show when that can happen, and how the optimal pooling is chosen so to prevent the desirability of any default ex-post.

Starting from an allocation where all agents are perfectly separated, it is still optimal

to default on the highest types first (unless H is high enough to give full commitment). Thus, to be able to commit, some agents at the top need to be pooled. The cutoff $\tilde{\tilde{\theta}}$ has to be such that

$$[F(\hat{\theta}) + (1 - F(\tilde{\tilde{\theta}}))]u(h_2(\hat{\theta})) \leq \int_{\underline{\theta}}^{\hat{\theta}} u(h_2(\theta))dF(\theta) + (1 - F(\tilde{\tilde{\theta}}))u(h_2(\tilde{\tilde{\theta}})) \tag{1.55}$$

$$\text{where } h_2(\hat{\theta}) = \frac{\int_{\underline{\theta}}^{\hat{\theta}}(h_2(\theta))dF(\theta) + (1 - F(\tilde{\tilde{\theta}}))(h_2(\tilde{\tilde{\theta}}) - H)}{(F(\hat{\theta}) + (1 - F(\tilde{\tilde{\theta}})))} \tag{1.56}$$

This ensures that the gain from defaulting on *all* agents above $\tilde{\tilde{\theta}}$ (left side of equation (1.55)) is smaller than the associated loss (right side of equation (1.55)). Notice that if the cutoff $\tilde{\tilde{\theta}}$ is chosen such that the government is exactly indifferent (i.e. equation (1.55) holds with equality), then $\tilde{\tilde{\theta}}$ is larger than the cutoff $\tilde{\theta}$ derived by solving the problem without the additional constraint of horizontal equity (problem (1.47) in proof 1.7.3). The reason is simply that without the horizontal equity requirement, the government had to be deterred from defaulting with any positive probability - a stronger requirement than indifference for default probability $\pi = 1$.

Next, suppose $\tilde{\tilde{\theta}}$ is chosen as high as possible, such that (1.55) binds, and all types below this cutoff are perfectly separated. Then, the government will always find it optimal to let a few agents just below the cutoff default. In other words, one can always find a positive ϵ, so that

$$[F(\hat{\theta}) + (F(\tilde{\tilde{\theta}}) - F(\tilde{\tilde{\theta}} - \epsilon))]u(h_2(\hat{\theta})) > \int_{\underline{\theta}}^{\hat{\theta}} u(h_2(\theta))dF(\theta) + \int_{\tilde{\tilde{\theta}}-\epsilon}^{\tilde{\tilde{\theta}}} u(h_2(\theta))dF(\theta) \tag{1.57}$$

$$\text{where } h_2(\hat{\theta}) = \frac{\int_{\underline{\theta}}^{\hat{\theta}}(h_2(\theta))dF(\theta) + \int_{\tilde{\tilde{\theta}}-\epsilon}^{\tilde{\tilde{\theta}}} (h_2(\theta) - H)dF(\theta)}{(F(\hat{\theta}) + (F(\tilde{\tilde{\theta}}) - F(\tilde{\tilde{\theta}} - \epsilon)))} \tag{1.58}$$

To prevent default on types below the cutoff $\tilde{\tilde{\theta}}$, the government can either choose $\tilde{\tilde{\theta}} = \tilde{\theta}$ (then since default is not even optimal with probability zero, so it cannot be optimal for any lower type, since less resources would be gained) or it can pool agents below $\tilde{\tilde{\theta}}$ so that it won't find it optimal to default on them. In other words, the government is still able to implement the same allocation with pooling at the top only as before. However, it may find it optimal to let fewer agents pool at the top, but implement pooling throughout the distribution. In fact, it is easy to verify that whenever agents in some range are perfectly separated, it is optimal to pool them together and in return achieve more separation higher up the type distribution. The reason is simply that more precise information about agents with higher income types is more valuable, since more redistribution across agents can be achieved.

Thus, at the optimal allocation, agents will be pooled into finitely many groups throughout the type distribution. As the default costs increase, the number of groups increases. Only when default costs are high enough to give full commitment, perfect separation (i.e. "pooling" into infinitely many groups) can be sustained. Since any positive default cost allowed some separation in the case without horizontal equity, any positive default costs always allows separation into at least 2 groups of agents with the horizontal equity requirement. This concludes the proof of Proposition 5.

Chapter 2

Optimal Debt Bias in Corporate Income Taxation

I present a rationale for a government to discriminate between debt and equity financing when taxing corporate income. For risk-averse entrepreneurs, equity generates more surplus than debt, because it provides financing *and* insurance. A government seeking to extract surplus from entrepreneurs would naturally tax equity-generated income more than debt-generated income. I also establish a less obvious reason for why the government might want to extract surplus from entrepreneurs: It is well understood that when the quality of projects is unobservable to investors, risk-averse entrepreneurs with higher return projects retain a larger share of equity to signal their type (Leland and Pyle (1977)). I show that in such an adverse selection setting, while competitive investors are constrained to offer actuarially fair terms, the government can use taxes to discriminate between types. This degree of freedom allows a manipulation of the relevant incentive constraints so that a lower level of debt suffices for separation, and an increase in overall efficiency can be obtained. Since entrepreneurs separate along their debt-to-equity ratios, the optimal non-linear tax schedule to achieve the desired discrimination is isomorphic to one that taxes debt-generated income at a lower rate than equity-generated income.

2.1 Introduction

Many tax codes do not treat debt and equity financing equally. While interest payments for a loan can often be deducted from the corporate income tax base, dividends to equity holders are taxed as profits on the firm side (and then often again as capital income on the investor's side). This constitutes a discrimination in favor of debt financing which is widely believed to be suboptimal. In this paper, I present a reason for why a government might optimally choose to discriminate between debt and equity financing. When provided by a competitive financial market, debt and equity financing generate different levels of surplus for the entrepreneur: While both help the entrepreneur to realize the *implementation gain* of his project, equity also provides *insurance*. A government aiming at extracting the surplus from entrepreneurs has thus no reason to tax income generated by different means of financing at the same rate.

The rationale for a government to aim at extracting surplus, though, is less obvious: Risk-averse entrepreneurs whose projects differ in expected returns, typically differ in their willingness to pay for insurance of the same risk. Yet, if financial markets are competitive and insurance can be obtained at actuarially fair terms, every such entrepreneur prefers equity over debt. Perfect insurance is obtained only when the complete ownership of the project is sold to an investor. However, if the characteristics of projects are unobservable to investors ex-ante, catering to the different entrepreneurs becomes a problem of screening types. Leland and Pyle (1977) have shown that in this case entrepreneurs with higher return projects will retain a larger share of their project to signal their type to investors.

How does a tax discrimination between debt and equity influence the outcome of this signaling game? Much along the lines of Spence (1973), using more debt to finance a new investment is a costly, yet socially inefficient signal. It is wasteful, because the risk-averse party needs to take more risk than if information were symmetric. Then,

subsidizing the wasteful signal might simply lead to an increased use of debt, but most likely not to a difference in the outcome of the signaling game[1]. However, this logic only applies if debt receives an absolute subsidy. The typical corporate income tax schedule instead entails only a *relative* subsidy of debt over equity.

In this paper, I show that the relative discrimination between debt and equity can instead lead to a more efficient outcome. Adverse selection necessitates the use of debt as a signal. The level of debt necessary for separation is thus dictated by the incentive constraints of a screening problem. Competition between investors does not allow for any further price discrimination between types - everyone receives his equilibrium amount of insurance at actuarially fair terms. The government, however, is not restricted to the prices dictated by competition. Instead, it can use differential taxation to implement a form of price discrimination between types that resembles the behavior of a monopolist. This degree of freedom allows additional manipulation of the incentive constraints and can so lead to a *lower* overall use of debt. Such an increase in efficiency requires to extract surplus from the entrepreneurs. While the equilibrium with taxes can be more efficient than the competitive equilibrium, it overall distributes away from the entrepreneurs and toward other parts of the economy (not modeled here). Studies comparing monopolistic versus competitive insurance markets come to similar conclusions (e.g. Stiglitz (1977), Dahlby (1987), de Feo and Hindriks (2009)).

The optimal discrimination between types generally calls for a non-linear tax on cor-

[1]There are two caveats: First, if the entrepreneurs' preferences exhibit *increasing* absolute risk aversion, a linear subsidy on debt might draw in more lower types, thereby weakening the signal. At the shares of retained equity that the high types are able to afford when trading off their risk aversion against the combined incentives for using debt (the signaling value and the tax incentives), investors might not be able to infer their types anymore, as the same tax subsidy might have led to lower types retaining the same shares of their projects. A separating equilibrium might then fail to exist. Second, the tax schedule would have to include a no-arbitrage condition, so that types can take out at most so much debt to bridge the gap in financing that arose due to a lower equity issue. With a tax subsidy, the highest types, who would be willing to keep almost all the ownership of their project will run into that constraint. They cannot increase the share of debt any higher to separate themselves from lower types, some pooling at the top might result.

porate income. I show that since entrepreneurs separate along their debt-to-equity ratios, the optimal tax schedule is isomorphic to one that taxes the fraction of income generated with equity differently from that generated with debt financing. In such a schedule, each type faces a higher tax rate on equity-financed income than on debt-financed income, because the surplus generated by equity financing is larger (implementation *and* insurance gains). In some cases, even a linear schedule with separate tax rates on the fraction of income generated by debt versus equity and a debt bias, approximates the optimal non-linear tax schedule that maximizes revenue.

Related Literature

The modern discussion about what determines the capital structure of corporations started with the seminal contribution of Modigliani and Miller (1958). They state that in an efficient market and in absence of taxes, bankruptcy and agency costs, and asymmetric information, the value of the firm is invariant with respect to its capital structure. Relaxing any of these conditions in turn and analyzing the resulting optimal capital structure has since been a major focus of the corporate finance literature. Harris and Raviv (1991) provide a comprehensive overview.

Many tax codes favor debt financing over equity. For corporations, interest payments on loans are to a large extend deductible from the tax base, while dividend returns to equity holders are not. In the US, Graham (2000) estimates the tax benefits for debt to amount to 9.7% of firm value. de Mooij (2011) reports that the cost of equity-financed investment was higher than that of debt-financed investment in 2007 for firms in the US, Japan, and the EU-27. Bradley et al. (1984) survey the large literature that has investigated the effects of this discrimination between means of financing. The general argument is that firms choose the optimal level of debt trading off tax incentives against the potential costs of financial distress resulting from debt financing.

Whether firms do indeed base their choice of capital structure on the tax incen-

tives for debt has long been questioned[2]. However, recent studies do find evidence for the hypothesis that "the desirability of debt finance at the margin increases with the firm's effective marginal tax rate on deductible interest" (MacKie-Mason (1990), p.1482). The behavioral response of firms to the "debt bias" has been quantified most recently by Gorden and Lee (2001, 2007). They estimate that shifting from the average tax distortion to no tax distortion would reduce debt-to-capital ratios by 0.022, implying that an additional 2.2% of capital would be financed with equity rather than debt (Gordon and Lee (2007)).

These studies focus on the optimal decision of financing from the firm's point of view, taking the debt bias in corporate income tax schedules as given. None of them ask why it would be in the government's interest to implement such a discrimination. Despite the arguments against this discrimination, it has persisted over time, and might have even become larger (de Mooij (2011)). Yet, the general consensus is that the discrimination between debt and equity should be eliminated (as for example argued by Auerbach et al. (2010) in the Mirrlees Review).

In contrast to the static tax incentive versus bankruptcy costs trade-off view, a large stream of the corporate finance literature following Modigliani and Miller (1958) has considered asymmetric information between entrepreneurs and financiers and its effect on firms' capital structure. A firm's choice of the means of financing, so the main argument, might contain information about its underlying value. The signal conveyed to investors and the associated costs in terms of firm valuation are what determines the decision to issue new equity or debt.

In a seminal contribution, Myers and Majluf (1984) have argued that information asymmetries can explain why the stock price of a company typically declines after new equity is issued. If investors cannot observe the underlying value of a firm (e.g.

[2]MacKie-Mason (1990) briefly surveys Titman and Wessels (1988), Fischer et al. (1989), Ang and Peterson (1986), Long and Malitz (1985), Bradley et al. (1984), and Marsh (1982) as "studies that fail to find plausible or significant tax effects" (p.1471).

the quality of a new project for which the firm seeks financing), they will factor in some probability that the entrepreneur or manager of the firm is behaving opportunistically, issuing shares when they are overvalued. Issuing debt does not affect the price of financing in the same way - the value of debt, especially safe or highly rated debt, is largely unaffected by a change in the stock price. A number of studies have tested the implications of this theory empirically, and have found evidence for stock price decreases related to new equity issues (see Dierkens (1991) for a summary). Myers (1984) therefore proposes a "pecking order" of corporate finance. According to this theory, firms finance new projects with internal fund first, and only after exhausting them seek outside capital, first in the form of safe debt and only when absolutely necessary through new equity. These papers consider firms as risk-neutral agents. One important difference between debt and equity, however, is the amount of risk the entrepreneur can shift to the investor.

On the investment side of the financing transaction, Gordon and Bradford (1980) have considered a model with risk-averse investors. They conclude that an investor's portfolio mix of stocks and bonds will depend on the tax rates he faces as well as his risk-aversion. When tax rates are not linear and absolute risk aversion is not perfectly constant, the individually optimal mix will depend on the overall size of the portfolio.

Leland and Pyle (1977) instead consider a setup with risk-averse entrepreneurs that is most closely related to the one presented here. In this case, everything else equal, debt is a less attractive means to raise capital, because the entrepreneur has to bear more risk compared to selling shares of his firm. The authors then show that retaining a higher stake in their firm can serve as a signal of confidence to investors, so that entrepreneurs with higher quality projects would be willing to retain higher shares of equity (and instead issue debt to meet their financing needs) simply to separate themselves from lower type entrepreneurs. Entrepreneurs weigh the benefit from receiving more favorable terms of financing (due to a higher market valuation of their firm) against the cost of having to bear more risk. This finding parallels those of the

classic insurance literature pioneered by Rothschild and Stiglitz (1976), as well as the literature on education as signaling device initiated by Spence (1973). Accordingly, the results of the presented paper relate to findings of the insurance literature that compares monopolistic and competitive provision, first introduced by Stiglitz (1977) and Dahlby (1987), and more recently generalized by Chade and Schlee (2011) and de Feo and Hindriks (2009).

Recently, other authors have considered the optimal taxation point of view to corporate income taxation in setups with asymmetric information. The main concern of this literature has been an inefficient entry of entrepreneurs in models of occupational choice. Gathak et al. (2007) show that if entrepreneurs differ along only one dimension, a lump-sum tax on entrepreneurs can correct an excessive entry of low type entrepreneurs. Scheuer (2011) investigates a model where entrepreneurs differ along two dimensions. Then, credit market imperfections lead to the government optimally intervening with a nonlinear subsidy on entrepreneurial profits, due to an inefficient entry of entrepreneurial types at both ends of the ability distribution. None of these papers derives an optimal tax schedule that involves a discrimination between debt and equity. However, contrary to this paper, they all consider risk-neutral entrepreneurs, and therefore disregard the potential of tax discrimination the dimension of risk aversion.

The rest of the paper is structured as follows: In section 2.2, I set up a model reminiscent of Leland and Pyle (1977) where risk-averse entrepreneurs with heterogeneous projects seek financing from competitive investors. Asymmetric information about the quality of projects results in an adverse selection problem. To illustrate my point, I employ a setup with two types of entrepreneurs. Section 2.3 describes the competitive equilibrium in this economy. Section 2.4 then analyzes a government's opportunities to intervene. It is first shown that the revenue maximizing tax schedule implements a price discrimination between types and generally leads to a different level of debt for the high return type entrepreneur than in the competitive equilib-

rium. Section 2.4.1 then proceeds to show that a tax schedule with a debt bias can be designed to collect the same revenue as the optimal tax schedule. Finally, section 2.4.2 analyzes a specific example in which even a linear tax schedule with different rates on debt versus equity generated income approximates the optimal non-linear schedule. Section 2.5 concludes.

2.2 Setup

Consider an economy populated by entrepreneurs who seek financing for their projects and investors who compete to provide the funds.

Entrepreneurs

There exists a continuum of entrepreneurs of size one. Each entrepreneur owns the idea for a project, but has no initial wealth to cover the required setup costs I to implement his project. Entrepreneurs are risk-averse, their utility function u is increasing, strictly concave, differentiable and exhibits non-increasing absolute risk aversion (NIARA).

Entrepreneurs are of two different types, indicated by index $i \in \{L, H\}$, β and $(1-\beta)$ are the respective shares in the population. Types differ with respect to the return their project can generate. In particular, I assume that an implemented project produces a gross return of

$$Y(\theta_i, E) = I + \theta_i + E \tag{2.1}$$

Here, θ_i is the *individual* mean return (net of the setup costs I) that differs with the type of the entrepreneur, and is known to him even before the implementation of the project. E represents an *aggregate* shock, its realization is unknown to everybody at the time of interaction between entrepreneurs and investors.

I assume that

(A1) $E \in \{\epsilon, -\epsilon\}$, with $\mathbb{E}[E] = 0$.

(A2) ϵ is small compared to θ_i: $0 < \epsilon << \theta_L < \theta_H$.

Assumption (A1) implies that a risk-neutral agent would disregard the aggregate shock in his optimization of ex-ante utility. However, since entrepreneurs are risk-averse, they do take the aggregate risk into account when deciding whether to implement their project. Assumption (A2) implies that the initial setup costs I are always recovered. Every project has a positive return.

While without any initial wealth, entrepreneurs do have outside options, denoted ψ_i. For example, one might think about entrepreneurs being able to implement their project elsewhere in the world, or to simply remain in the labor force of the economy's productive sector. I assume that

(A3) $\psi_i < C_i$, where C_i is the certainty equivalent defined by $u(C_i) = \mathbb{E}[u(\theta_i + E)]$;

(A4) $C_H > \theta_L$.

Assumption (A3) ensures that entrepreneurs are willing to implement their projects if they are offered financing at sufficiently good terms. Assumption (A4) puts a joint restriction on $\theta_L - \theta_H$ (the spread of the mean returns), the aggregate risk E and the concavity of the entrepreneurs' utility function. It essentially places a lower bound on the spread of safe outside options. With this assumption, I am restricting attention to cases where an equilibrium always exists.

Investors

The financial market consists of a large number of risk-neutral investors, each with unlimited funds. They can either invest at the safe gross interest rate normalized to

(A5) $R = 1$

or finance projects.

Investors can offer financing contracts to entrepreneurs. A financial contract is a pair $(x, T) \in (0, 1) \times \mathbb{R}$. I denote with $x \in (0, 1)$ the share of the project that remains in the ownership of the entrepreneur. Thus, $x = 0$ corresponds to an equity contract where complete ownership of the project is transferred to the investor, while $x = 1$ denotes a pure debt contract where the entrepreneur remains the owner of his project. x is thus a measure of the degree of insurance the entrepreneur purchases (where a smaller x corresponds to more insurance).

In any case, the contract specifies a fixed payment T to the entrepreneur, which is net of the setup cost I. It is without loss of generality to consider only contracts where T is not a function of the realization of the aggregate shock E[3]. It is useful to think of T as reflecting the price the entrepreneur pays for a financing service with a degree of insurance x. Since *realized* output Y_i perfectly reveals all private information, contracts must be restricted to not be contingent on it. Otherwise, a simple penalty for lying about the true type would easily circumvent the adverse selection problem that is at the heart of this study. Realized payoffs do depend on realized output for all players holding parts of the ownership rights.

Timing

The strategic interaction considered is the following:

1. Entrepreneurs learn their type.

2. Investors offer a set of contracts, denoted $\mathbb{X} = \{(x, T) \in (0, 1) \times \mathbb{R}\}$. They are subsequently committed to honor the terms of the offers.

3. Entrepreneurs choose whether to implement their projects (let $\xi_i \in \{0, 1\}$ represent that decision), and if so which contract to accept. They can accept only one

[3]In principle, a debt contract could specify a T that depends on the realization of E. This would be the case if entrepreneurs are asked to pay an interest rate on their debt that is so high that they might not be able to afford it in the bad state. However, competition between investors will ensure that interest rates on debt contracts are at least equal to the safe rate $R = 1$. Assumption (A2) then ensures that limited liability is never a binding constraint.

contract[4]. Accordingly, payments T_i are made and the projects implemented.

4. Aggregate uncertainty is realized and returns are distributed according to the contracted ownership of shares.

The concept of competition between investors is similar to that introduced by Rothschild and Stiglitz (1976). As is well know, in this setup an equilibrium might not exist. Assumption (A4) does ensure, however, that an equilibrium always exists[5].

2.3 Equilibrium without Taxes

In the last active stage, each entrepreneur makes an implementation decision $\xi_i \in \{0,1\}$ and decides which contract to sign, taking the set of offered contracts as given. He solves:

$$\begin{aligned} \max_{\xi,(x,T)} \quad & \xi\, \mathbb{E}[u(x(\theta+E)+T)] + (1-\xi)\, u(\psi) \\ \text{s.t.} \quad & \xi \in \{0,1\} \quad \text{and} \quad (x,T) \in \mathbb{X} \end{aligned} \tag{2.2}$$

The financial market in this economy is competitive. Thus, even though investors are maximizing their expected profits, in equilibrium they will make zero profits in expectation. Entrepreneurs on the other hand are risk-averse, how much of the aggregate risk they have to bear plays a decisive role in their implementation and financing decision.

Definition 2

An equilibrium is a set of contracts $\mathbb{X} = \{(x_i, T_i) \in (0,1) \times \mathbb{R}, i = L, H\}$ *and an implementation decision* ξ_i *of each entrepreneur such that:*

(i) entrepreneurs maximize their expected utility;

(ii) investors make zero expected profits.

[4]This is a short cut to assuming that trades are observable.

[5]Classic (e.g. Riley (1979)) as well as more recent studies (e.g. Dubey and Geanakoplos (2002)) have modified the Rothschild-Stiglitz concept of competitive equilibrium to deal with the non-existence of equilibrium problem. The general consensus is that there always exists a separating equilibrium.

2.3.1 Observable Types

In a first-best world, investors are able to observe an entrepreneur's type and can offer type-specific contracts. Since investors are competitive, the optimal contract for a type i entrepreneur maximizes his expected utility, subject a zero expected profit constraint for the investor. The implementation decision is made taking into account the entrepreneur's outside option ψ_i:

$$\begin{aligned} \max_{\xi_i,(x_i,T_i)} \quad & \xi_i \, \mathbb{E}[u(x_i(\theta_i + E) + T_i)] + (1 - \xi_i) \, u(\psi_i) \\ \text{s.t.} \quad & (1 - x_i)\theta_i - T_i \geq 0 \end{aligned} \tag{2.3}$$

Lemma 10

If types are observable, in equilibrium all projects are implemented in the economy. All entrepreneurs obtain full insurance at actuarially fair terms: $\xi_i = 1$ *and* $(x_i^*, T_i^*) = (0, \theta_i) \quad \forall i$.

Proof: Suppose that $\xi_i = 1$. Then, the optimal contract solves

$$\max_{(x_i,T_i)} \quad \frac{1}{2}[u(x_i(\theta_i + \epsilon) + T_i) + u(x_i(\theta_i - \epsilon) + T_i)]$$

$$\text{s.t.} \quad \frac{1}{2}[u(x_i(\theta_i + \epsilon) + T_i) + u(x_i(\theta_i - \epsilon) + T_i)] \geq u(\psi_i) \tag{2.4}$$

$$(1 - x_i)\theta_i - T_i \geq 0 \tag{2.5}$$

The first-order conditions for this optimization are:

$$[x_i] \quad \frac{1}{2}(1 + \mu)\left[u'(x_i(\theta_i + \epsilon) + T_i)(\theta_i + \epsilon) + u'(x_i(\theta_i - \epsilon) + T_i)(\theta_i - \epsilon)\right] = \lambda\theta_i \tag{2.6}$$

$$[T_i] \quad \frac{1}{2}(1 + \mu)\left[u'(x_i(\theta_i + \epsilon) + T_i) + u'(x_i(\theta_i - \epsilon) + T_i)\right] = \lambda \tag{2.7}$$

where μ and λ are the Lagrange multipliers associated with the individual rationality constraint of the entrepreneur (2.4) and the zero expected profit condition for the

investor (2.5) respectively. These necessary conditions for optimality require

$$\theta_i = \frac{u'(x_i(\theta_i+\epsilon)+T_i)(\theta_i+\epsilon) + u'(x_i(\theta_i-\epsilon)+T_i)(\theta_i-\epsilon)}{u'(x_i(\theta_i+\epsilon)+T_i) + u'(x_i(\theta_i-\epsilon)+T_i)} \tag{2.8}$$

which implies $x_i^* = 0$. The zero expected profit constraint (2.5) determines $T_i^* = \theta_i$. A contract $(x_i = 0, T = \theta_i)$ satisfies the individual rationality constraint of entrepreneur i, and so he optimally chooses $\xi_i = 1$. □

Borch's Rule[6] of optimal risk sharing implies that the risk-neutral party (here the investor) should bear all the risk. Each entrepreneurial type receives full insurance. Competition between investors implies that in equilibrium entrepreneurs will be offered actuarially fair insurance. Being offered such favorable terms, all entrepreneurs optimally decide to implement their projects, and sell their firms to an investor in an equity contract.

Financial contracts serve two purposes: They provide financing to set up the project as well as insurance against the aggregate risk. In a first-best world, one could think of separate markets for these two tasks. Each entrepreneur would then issue safe debt contract to finance the setup costs I and sign a separate insurance contract. The first-best allocation in such a world would be equivalent to the presented setup.

2.3.2 Unobservable Types

Suppose now that investors in the economy cannot observe an entrepreneur's type, and so are uncertain about the mean expected return of their investment when offering to buy a share of equity. The first-best set of contracts can not be an equilibrium anymore: Since $T_L^* < T_H^*$, every entrepreneur would claim to have a high return idea to maximize the price he can fetch from selling the ownership rights to his project. The investor's zero expected profit condition would then be violated, he would make certain losses. Recall that while *realized* returns are perfectly informative about the

[6]See Borch (1962).

entrepreneur's type, investors are by assumption precluded from offering contracts with payments T contingent on realized returns.

The equilibrium implementation decisions and set of contracts in this adverse selection problem are the solution to a standard screening problem. Since preferences exhibit NIARA, they satisfy the single-crossing property. By the Revelation Principle[7] it suffices to design contracts that are incentive compatible for each type of entrepreneur to choose the contract that is meant for him.

$$\max_{\xi_i, \mathbb{X}} \quad \begin{aligned} &\beta \quad \{\xi_L \, \mathbb{E}[u(x_L(\theta_L + E) + T_L)] + (1 - \xi_L)\, u(\psi_L)\} + \\ &(1 - \beta)\, \{\xi_H \, \mathbb{E}[u(x_H(\theta_H + E) + T_H)] + (1 - \xi_H)\, u(\psi_H)\} \end{aligned}$$

$$\text{s.t.} \quad \xi_L \, \mathbb{E}[u(x_L(\theta_L + E) + T_L)] \geq \xi_L \, \mathbb{E}[u(x_H(\theta_L + E) + T_H)] \tag{2.9}$$

$$\xi_H \, \mathbb{E}[u(x_H(\theta_H + E) + T_H)] \geq \xi_H \, \mathbb{E}[u(x_L(\theta_H + E) + T_L)] \tag{2.10}$$

$$\xi_L \beta[(1 - x_L)\theta_L] + \xi_H(1 - \beta)[(1 - x_H)\theta_H] \geq \xi_L \beta \, T_L + \xi_H \beta \, T_H \tag{2.11}$$

The optimization is now further constrained by (2.9) and (2.10), the incentive compatibility constraints for entrepreneurial types L and H respectively. Investors breaking even in expectation is ensured by constraint (2.11). With the added complication of an information asymmetry between entrepreneurs and investors, not all entrepreneurs obtain full insurance in equilibrium:

Lemma 11

If types are unobservable, in the unique equilibrium all projects are implemented: $\xi_i = 1$ *for* $i = L, H$. *In particular,*

(i) type L entrepreneurs obtain full insurance at actuarially fair terms: $(x_L, T_L) = (0, \theta_L)$;

[7]See for example Myerson (1979) and Harris and Townsend (1981).

(ii) type H entrepreneurs obtain partial insurance at actuarially fair terms:
$(x_H, T_H) = (0 < x_H \leq 1, (1 - x_H)\theta_H)$.

Proof: See appendix 2.6.1.

Even though investors are risk-neutral, they cannot provide full insurance to all entrepreneurs. Due to asymmetric information, full insurance for all types could only be granted in a pooling equilibrium, where type H entrepreneurs receive less than actuarially fair insurance and subsidize type L entrepreneurs. Just as in the canonical Rothschild-Stiglitz model, such a pooling contract is not an equilibrium: A contract that offers a little less insurance and attracts only high types, without subsidizing low types, is a profitable deviation. This type of cream-skimming by investors rules out a pooling equilibrium.

Just as in the insurance literature, in the separting equilibrium, higher types (i.e. lower risks) receive less insurance. This property of the equilibrium parallels the findings of Leland and Pyle (1977): Higher type entrepreneurs retain a higher share of their projects to signal their type to the market.

Competition between investors drives their expected profits to zero. All the surplus generated from implementing projects and providing (partial) insurance accrue to the entrepreneurs.

2.4 Equilibrium with Taxes

Suppose the government of the economy is aiming at extracting the surplus entrepreneurs make. One might think of entrepreneurs as being foreign to the economy, and free to set up their projects anywhere in the world. A government might then want to ensure that some (or all) of the surplus generated through the use of the economy's financial market are recovered. More generally, one might also think about a government trying to redistribute away from entrepreneurs to other parts of the econ-

omy (not modeled in the presented setup).

I assume that the government announces a tax schedule τ at the beginning of time and is subsequently committed to it. Taxes can in principle depend on types, and on the specifics of the financing contract (x, T), and are always collected from the entrepreneur. Investors and entrepreneurs take the announced taxes as given in their decisions. I denote the problem of a type i entrepreneur with $\mathbb{U}_i$:

$$\begin{aligned} \max_{\xi,(x,T)} \quad & \xi\, \mathbb{E}[u(x(\theta+E)+T-\tau(x,T,\theta))] + (1-\xi)u(\psi) \\ \text{s.t.} \quad & \xi \in \{0,1\} \quad \text{and} \quad (x,T) \in \mathbb{X} \end{aligned}$$

As before, the optimal financing contracts are found as a solution to the screening problem (2.9)-(2.11), where investors take the effect of the tax schedule on the entrepreneurs' decision problem into account. I denote the investor's problem by $\mathbb{P}$. The government then solves:

$$\begin{aligned} \max_{\{\tau(x,T,\theta)\}} \quad & \beta\, \xi_L \tau(x_L, T_L, \theta_L) + (1-\beta)\, \xi_H \tau(x_H, T_H, \theta_H) \\ \text{s.t.} \quad & \xi_i \in \operatorname{argmax} \mathbb{U}_i \quad \forall i & (2.12) \\ & \mathbb{X} \in \operatorname{argmax} \mathbb{P} & (2.13) \end{aligned}$$

The government maximizes tax revenue and only cares about the entrepreneurs' utility insofar as it would like their participation constraints to be satisfied.

Definition 3

An equilibrium with taxes is a set of contracts $\mathbb{X} = \{(x_i, T_i) \in (0,1) \times \mathbb{R}, i = L, H\}$, *an implementation decision* ξ_i *of each entrepreneur, and a tax schedule* $\{\tau(x_i, T_i, \theta_i)\}$ *such that:*

(i) entrepreneurs maximize their expected utility, taking taxes as given;

(ii) investors make zero expected profits;

(iii) the government maximizes tax revenue.

The equilibrium is comparable in structure to the competitive equilibrium. However, the government will tax entrepreneurial surplus.

Lemma 12

When types are unobservable, in the equilibrium with taxes all projects are implemented: $\xi_i = 1$ *for* $i = L, H$. *In particular,*

(i) type L entrepreneurs obtain full insurance and investors pay actuarially fair terms:
$(x_L, T_L) = (0, \theta_L)$;

(ii) type H entrepreneurs obtain partial insurance and investors pay actuarially fair terms:
$(x_H, T_H) = (0 < x_H \leq 1, (1 - x_H)\theta_H)$;

(iii) type H entrepreneurs will be taxed so that they are indifferent between implementing their projects or their outside option.

Proof: See appendix 2.6.2.

This equilibrium structure is analogous to the one in Stiglitz (1977), who analyzes a monopoly insurance problem. Indeed, the government's objective to maximize tax revenue coincides with that of a monopolist investor. As in Stiglitz (1977), contracts in the equilibrium with taxes (denoted with superscript G) will generally differ from those in the competitive equilibrium without taxes (denoted with superscript C).

Proposition 6

Generically, $\mathbb{X}^G \neq \mathbb{X}^C$.

Proof: See appendix 2.6.3.

In the equilibrium without taxes, it follows from competition between investors that debt earns no interest beyond $R = 1$ and all insurance is sold at actuarially fair terms. Insurance coverage for the high types is determined solely by the incentive constraints of the low types. Competition between investors leaves no room for price

discrimination between types of entrepreneurs. The government however doesn't face competition. Using taxes, it can implement effective prices that differ from those consistent with competition. It essentially acts like a monopolist, who is able to charge differential mark-ups. This degree of freedom allows the government to manipulate the incentive constraint of the low types such that a different level of insurance for the high type emerges in equilibrium.

Corollary 1

When $\tau_H > \tau_L$, type H entrepreneurs receive more insurance than in the competitive equilibrium ($x_H^G < x_H^C$).

Proof: See appendix 2.6.4.

The possibility of an increase in efficiency again parallels findings in the insurance literature. For insurance markets with adverse selection, it has been shown by Dahlby (1987) that coverage for the low risk types is higher when purchased from a monopolist insurer, rather than in a competitive market. More generally, de Feo and Hindriks (2009) show that monopolists are often more efficient at providing insurance under adverse selection than a competitive market.

It should be noted that while the equilibrium with taxation can be more efficient than the competitive equilibrium, the necessary discrimination has distributional consequences. The government's objective is to maximize the revenue extracted from the entrepreneurs. One might interpret this as a re-distributional objective away from entrepreneurs and toward other parts of the economy (not modeled in this paper).

2.4.1 Implementing the Optimal Non-linear Tax Schedule

Lemma 12 stated that the government is able to extract all the surplus from the entrepreneurs. Generally, a non-linear tax schedule will be optimal to achieve that objective. It will satisfy the individual rationality constraints for both types with equality, so that no surplus is left to the entrepreneurs:

$$u(\theta_L - \tau_L) = u(\psi_L) \tag{2.14}$$

$$\mathbb{E}[u(\theta_H - \tau_H + x_H^G E)] = u(\psi_H) \tag{2.15}$$

Proposition 7

The optimal tax schedule $\{\tau_L, \tau_H\}$ *is isomorphic to one that taxes the fraction of income generated by equity at a higher rate than the fraction of income generated by debt.*

Proof: Define R_i as the absolute risk premium type i would be willing to pay to avoid the risk he would be exposed to from holding all shares of his project:

$$\mathbb{E}[u(\theta_i - \tau_i + E)] = u(\theta_i - \tau_i - R_i) \tag{2.16}$$

Since the aggregate risk is small (by assumption (A2)), the absolute risk premium can be approximated by

$$R_i \approx \frac{1}{2} r_i Var(E) \tag{2.17}$$

where $r_i = -\frac{u''}{u'}$ is the Arrow-Pratt measure of absolute risk aversion for type i evaluated at $\theta_i - \tau_i$. Using this definition, the optimal non-linear tax schedule satisfies

$$T_i = \theta_i - \psi_i - (x_i^G)^2 R_i \quad \text{for } i = L, H \tag{2.18}$$

By assumption (A3), each entrepreneur's outside option is $\psi_i < C_i$. It can always be written as $\psi_i = \delta_i \theta_i - R_i$. Thus, the surplus to be taxed away is

$$T_i = (1 - \delta_i)\theta_i + (1 - (x_i^G)^2) R_i \quad \text{for } i = L, H \tag{2.19}$$

The first summand, $(1 - \delta_i)\theta_i$, represents the *implementation gain*, i.e. the surplus generated only from obtaining financing for the setup costs and so being able to implement the project. The rest, $(1 - (x_i^G)^2)R_i$, represents the *insurance gain*, i.e. the additional surplus generated from receiving insurance. From this intuition, it is

clear that the same tax revenue τ_i can be generated by taxing the fractions of income generated by debt or equity at different rates:

$$\begin{aligned} x_i^G\theta_i\tau_i^D + (1-x_i^G)\theta_i\tau_i^E &= \tau_i \\ &= (1-\delta_i)\theta_i + (1-(x_i^G)^2)R_i \\ &= x_i^G(1-\delta_i)\theta_i + (1-x_i^G)(1-\delta_i)\theta_i + (1-(x_i^G)^2)R_i \end{aligned}$$

Here τ_i^D is the tax *rate* applied to the share of income retained by the entrepreneur. It taxes the surplus generated with debt, i.e. only a fraction of the implementation gain. τ_i^E is the tax *rate* applied to the fraction of income sold as equity to the investors. It taxes the surplus generated by equity, which consists of both a proportional fraction of the implementation gain as well as the insurance gain. Notice that

$$\tau_i^D = 1-\delta_i \tag{2.20}$$

$$\tau_i^E = 1-\delta_i + (1-x_i^G)\frac{R_i}{\theta_i(1-x_i^G)} \tag{2.21}$$

$$\rightarrow \tau_i^D < \tau_i^E \tag{2.22}$$

Thus, the optimal tax schedule τ_i is isomorphic to one that taxes the fraction of income generated by equity at a higher rate than income generated by debt. □

Proposition 7 states that a tax schedule with a debt bias might indeed be optimal, given the government's objective of extracting all surplus from the entrepreneurs. In reality, the debt bias in a typical corporate income tax schedule takes a particular form: It allows the costs of debt to be deducted from the tax base, whereas payments to equity holders are (to a large extend) considered taxable profits. In the model, such a debt bias would occur when a corporate income tax rate $\bar{\tau}_i$ would be levied on $\theta_i + T_i$, i.e. on all of the entrepreneurs generated income before paying out any equity holders. The entrepreneur thus earns $(\theta + T_i)(1-\bar{\tau}_i) - (1-x_i^G)\theta_i$, where competition determines $T_i = (1-x_i^G)\theta_i(1-\bar{\tau}_i)$. This results in a double taxation of the fraction

of income generated by equity financing:

$$(\theta + T_i)(1 - \bar{\tau}_i) - (1 - x_i^G)\theta_i = x_i^G\theta_i(1 - \bar{\tau}_i) + (1 - x_i^G)\theta_i(1 - \bar{\tau}_i)^2 \quad (2.23)$$

$$= x_i^G\theta_i(1 - \tau_i^D) + (1 - x_i^G)\theta_i(1 - \tau_i^E) \quad (2.24)$$

with $\tau_i^D < \tau_i^E$[8]. Thus, this particular form of a debt bias is nothing different than taxing the fraction of income generated with equity at a higher rate than the fraction generated with debt. The debt bias observed in many tax codes might be optimal, at least in structure, to maximize revenue generated from corporate income taxation.

2.4.2 Continuum of Types with CARA Preferences

So far, I have shown that the optimal tax schedule to extract all surplus generated for the entrepreneurs can be implemented using separate tax rates for fractions of income corresponding to the share of ownership retained by the entrepreneur or sold to an investor. The marginal tax rates on these fractions of income differ, because with the means of financing, the surplus generated differs: While both debt and equity help the entrepreneur to realize the implementation gain, only equity provides insurance against the aggregate risk, and so generates additional surplus. Thus, the tax rate on debt financed income will be lower than that on equity financed income for each type. However, generally, these tax rates still depend on types. So the question arises why the government would postulate a tax schedule with two separate tax rates per type when it could also just announce one tax payment τ_i per type that incorporates all surplus generated.

In what follows, I show that in some cases, a simple *linear* tax schedule for debt-financed and equity-financed income can approximate the optimal non-linear corporate tax. Since the fractions of income financed with equity vary with type, the

[8]In reality, there are many rules of what exactly can or cannot be deducted from the tax base, so that τ^E will never be exactly $2\tau^D - (\tau^D)^2$ as suggested by this simple example.

effective tax entrepreneurs face is still non-linear:

$$x_i\theta_i(1-\tau^D)+(1-x_i)\theta_i(1-\tau_E)=\theta_i(1-\tau_i) \tag{2.25}$$

$$\text{where} \quad \tau_i=\tau^E+x_i(\tau^D-\tau^E) \tag{2.26}$$

Suppose there exist a continuum of different types of entrepreneurs, index by $\theta \in [\underline{\theta}, \bar{\theta}]$, $0<\underline{\theta}<\bar{\theta}<1$, with a density function $f(\theta)$ that is strictly positive, continuous and differentiable everywhere, and satisfies a monotone likelihood ration property. Stiglitz (1977) showed that under these conditions, a fully separating equilibrium exists and is such that the lowest type $\underline{\theta}$ receives full insurance, whereas all other types receive only partial insurance coverage decreasing with type. Chade and Schlee (2011) more recently extended the analysis to provide conditions for full separation under more general type distributions.

Moreover, suppose that entrepreneurial preferences exhibit constant absolute risk aversion, and that each entrepreneur's outside option is a constant fraction of the certainty equivalent his project generates, i.e. it can be expressed as $\psi_i=\delta\theta_i-R$. Again, R is the absolute risk premium, which is now constant for all agents. Under these assumptions, the implementation gain of an entrepreneur is proportional to his type, while the insurance gain depends only the degree of insurance coverage. Further, consider a population of entrepreneurs with $\bar{\theta} \leq 2\underline{\theta}$.

From equation (2.18), we know that the optimal tax schedule collects revenue

$$\tau(\theta)=(1-\delta)\theta+(1-x(\theta)^2)R \tag{2.27}$$

from a type θ entrepreneur. The revenue collected with linear tax rates on debt and equity financed income is

$$x(\theta)\theta\tau^D+(1-x(\theta))\theta\tau^E \tag{2.28}$$

The tax rates τ^D and τ^E are pinned down by the implementation gain and the insurance gain from full insurance. A type that signs an equilibrium contract without any insurance ($x(\theta) = 1, T(\theta) = 0$), faces only the tax rate τ^D, which extracts the full implementation gain:

$$\theta\tau^D = (1-\delta)\theta \quad \rightarrow \quad \tau^D = 1 - \delta \tag{2.29}$$

The lowest type receives full insurance, so he realizes the implementation gain and the full insurance gain, and faces only tax rate τ^E, so that

$$\underline{\theta}\tau^E = (1-\delta)\underline{\theta} + R \quad \rightarrow \quad \tau^E = 1 - \delta + \frac{R}{\underline{\theta}} > \tau^D \tag{2.30}$$

With these tax rates, the surplus extracted from any type θ with a contract ($0 < x(\theta) < 1, T(\theta)$) generates tax revenue that is approximately the same as the optimal tax revenue in equation (2.27):

$$\begin{aligned} x(\theta)\theta\tau^D + (1-x(\theta))\theta\tau^E &\approx (1-\delta)\theta + (1-x(\theta)^2)R \\ \theta(\tau^E + x(\theta)(\tau^D - \tau^E)) &\approx (1-\delta)\theta + (1-x(\theta)^2)R \\ \theta(1-\delta+(1-x(\theta))R) &\approx (1-\delta)\theta + (1-x(\theta)^2)R \\ \frac{\theta}{\underline{\theta}} &\approx (1+x(\theta)) \end{aligned} \tag{2.31}$$

When the spread of mean returns is small, a linear schedule of tax rates $\{\tau^D, \tau^E\}$ applied to the fractions of income generated by debt and equity financing respectively approximates the optimal non-linear tax schedule.

2.5 Discussion

Many governments discriminate between debt and equity financing when taxing corporate income. Conventional wisdom however suggests that the means of financing should be treated equivalently. I present a rationale for why a government might choose to discriminate between debt and equity: Debt and equity financing generate

different levels of surplus for the entrepreneur. While both help him to realize the *implementation gain* of his project, equity also provides *insurance*. A government aiming at extracting the surplus from the entrepreneurs thus has no reason to tax income generated by different means of financing at the same rate.

A difference in surplus generated by equity versus debt only occurs when entrepreneurs are risk-averse. Yet, if insurance can be obtained at actuarially fair terms (as is the case in competitive financial markets), every entrepreneur prefers equity over debt. It is then due to an adverse selection problem that different types of entrepreneurs choose different debt-to-equity ratios: The associated screening problem results in an equilibrium that separates entrepreneurs using the share of retained earnings as a screening device. The level of debt necessary for separation is solely determined by the incentive constraints. Competition between investors does not allow for any further price discrimination between types. The government, however, can introduce taxes such that different types effectively face different mark-ups over actuarially fair insurance terms. This additional opportunity for discrimination can relax incentive constraints and lead to a more efficient outcome, with a higher overall degree of insurance.

To implement the optimal discrimination scheme between types, the government can make use of the fact that separation occurs along the debt-to-equity ratio. A differential taxation of income generated with debt versus income generated with equity financing is one way to achieve optimal discrimination. This mechanism provides another less obvious justification for a debt bias in corporate income taxation.

It should be noted that while the equilibrium with taxation can be more efficient than the competitive equilibrium, the necessary discrimination has distributional consequences. I analyzed a government whose objective is to maximize the revenue extracted from the entrepreneurs. One might interpret this as a re-distributional objective away from entrepreneurs and toward other parts of the economy (not modeled

in this paper).

Alternatively, one might consider an economy that would like to attract foreign entrepreneurs to set up their firms in the country. In search for an opportunity to finance their projects, entrepreneurs can decide where in the world to set up their firm. They make this decision merely based on expected utility maximization, taking into account any uncertainty they might face, and optimizing over the terms of financing they are offered by investors in different countries. If investors in the economy are competitive, they might well be able to attract foreign entrepreneurs. However, all the surplus generated accrues to the entrepreneur, i.e. outside the economy. The government might then try to regain some of that surplus to distribute among members of the economy.

References

Ang, J. and Peterson, D.: 1986, Optimal debt versus debt capacity: A disequilibrium model of corporate debt behavior, *in* A. Chen (ed.), *Research in Finance*, University of Chicago Press.

Auerbach, A., Devreux, M. and Simpson, H.: 2010, Taxing corporate income, *in* J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles and J. Poterba (eds), *Dimensions of Tax Design: the Mirrlees Review*, Oxford University Press.

Borch, K. H.: 1962, Equilibrium in a reinsurance market, *Econometrica* **30**, 424–44.

Bradley, M., Jarrell, G. and Kim, E.: 1984, On the existence of an optimal capital structure: Theory and evidence, *Journal of Finance* **39**, 857–878.

Chade, H. and Schlee, E.: 2011, Optimal insurance with adverse selection, *mimeo*, Arizona State University.

Dahlby, B.: 1987, Monopoly versus competition in an insurance market with adverse selection, *Journal of Risk and Insurance* **54**, 325–331.

de Feo, G. and Hindriks, J.: 2009, Harmful competition in the insurance markets, *Technical report*, University of Strathclyde, Glasgow.

de Mooij, R.: 2011, Tax biases to debt finance: Assessing the problem, finding solutions, *Imf staff discussion note*, International Monetary Fund.

Dierkens, N.: 1991, Information asymmetry and equity issues, *Journal of Financial and Quantitative Analysis* **26**, 181–199.

Dubey, P. and Geanakoplos, J.: 2002, Competitive pooling: Rothschild-stiglitz reconsidered, *Quarterly Journal of Economics* **117**, 1529–1570.

Fischer, E., Heinkel, R. and Zechner, J.: 1989, Dynamic capital structure choice: Theory and tests, *Journal of Finance* **44**, 19–40.

Gathak, M., Morelli, M. and Sjöström, T.: 2007, Entrepreneurial talent, occupational choice, and trickle up policies, *Journal of Economic Theory* **137**, 27–48.

Gordon, R. and Bradford, D.: 1980, Taxation and the stock market valuation of capital gains and dividends: Theory and empirical results, *Journal of Public Economics* **14**, 109–136.

Gordon, R. and Lee, Y.: 2001, Do taxes affect corporate debt policy? evidence from u.s. corporate tax return data, *Journal of Public Economics* **82**, 195–224.

Gordon, R. and Lee, Y.: 2007, Interest rates, taxes and corporate financial policies, *National Tax Journal* **60**, 65–84.

Graham, J.: 2000, How big are the tax benefits of debt?, *Journal of Finance* **55**, 1901–1941.

Harris, M. and Raviv, A.: 1991, The theory of capital structure, *Journal of Finance* **46**, 297–356.

Harris, M. and Townsend, R. M.: 1981, Resource allocation under asymmetric information, *Econometrica* **49**, 33–69.

Leland, H. and Pyle, D.: 1977, Information asymmetries, financial structure, and financial intermediation, *Journal of Finance* **32**, 371–388.

Long, M. and Malitz, I.: 1985, Investment patterns and financial leverage, *in* B. Friedman (ed.), *Corporate Capital Structures in the United States*, University of Chicago Press.

MacKie-Mason, J.: 1990, Do taxes affect corporate financing decisions?, *Journal of Finance* **45**, 1471–1493.

Marsh, P.: 1982, The choice between equity and debt: An empirical study, *Journal of Finance* **37**, 121–144.

Modigliani, F. and Miller, M.: 1958, The cost of capital, corporation finance, and the theory of investment, *American Economic Review* **48**, 261–297.

Myers, S.: 1984, The capital structure puzzle, *Journal of Finance* **39**, 575–592.

Myers, S. and Majluf, N.: 1984, Corporate financing and investment decisions when firms have information that investors do not have, *Journal of Financial Economics* **13**, 187–222.

Myerson, R.: 1979, Incentive compatibility and the bargaining problem, *Econometrica* **47**, 61–73.

Riley, J.: 1979, Informational equilibrium, *Econometrica* **47**, 331–359.

Rothschild, M. and Stiglitz, J.: 1976, Equilibrium in competitive insurance markets: An essay in the economics of imperfect information, *Quarterly Journal of Economics* **90**, 629–649.

Scheuer, F.: 2011, Entrepreneurial taxation, occupational choice, and credit market frictions, *mimeo*, Stanford University.

Spence, M.: 1973, Job market signaling, *Quarterly Journal of Economics* **87**, 355–374.

Stiglitz, J.: 1977, Monopoly, non-linear pricing and imperfect information: The insurance market, *Review of Economic Studies* **44**, 407–430.

Titman, S. and Wessels, R.: 1988, The determinants of capital structure choice, *Journal of Finance* **43**, 1–19.

2.6 Appendix

2.6.1 Proof of Lemma 11

First, I show that all projects are always implemented: Assumption (A2) implies that every project has a positive return. Competition between investors and assumption (A5) then imply that every entrepreneur can always issue (safe) debt at the gross interest rate $R = 1$. Thus, assumption (A3) implies that every project generates at least a positive implementation gain for the entrepreneur, even if he cannot obtain any insurance.

Second, I establish that the only equilibrium is a separating equilibrium. Given that all projects are implemented, the screening problem (2.9) through (2.11) can be rewritten as

$$\max_{\mathbb{X}} \quad \beta \quad \mathbb{E}[u(x_L(\theta_L + E) + T_L)] + (1 - \beta)\mathbb{E}[u(x_H(\theta_H + E) + T_H)]$$

$$\text{s.t.} \quad \mathbb{E}[u(x_L(\theta_L + E) + T_L)] \geq u(\psi_L) \tag{2.32}$$

$$\mathbb{E}[u(x_H(\theta_H + E) + T_H)] \geq u(\psi_H) \tag{2.33}$$

$$\mathbb{E}[u(x_L(\theta_L + E) + T_L)] \geq \mathbb{E}[u(x_H(\theta_L + E) + T_H)] \tag{2.34}$$

$$\mathbb{E}[u(x_H(\theta_H + E) + T_H)] \geq \mathbb{E}[u(x_L(\theta_H + E) + T_L)] \tag{2.35}$$

$$\beta[(1 - x_L)\theta_L] + (1 - \beta)[(1 - x_H)\theta_H] \geq \beta T_L + \beta T_H \tag{2.36}$$

This is a standard screening problem where maximization of entrepreneurial surplus is subject to individual rationality and incentive compatibility constraints, as well as

a zero profit condition for investors. As in the classic Rothschild and Stiglitz (1976) setup, a pooling contract cannot be an equilibrium. Due to competition between investors, the only candidate pooling contract would offer full insurance at *average* actuarially fair terms:

$$(x, T) = (0, \beta\theta_L + (1 - \beta)\theta_H) \tag{2.37}$$

High types would obtain full insurance but subsidize low types. A profitable deviation is possible. There exists a contract (x', T') that offer less than full insurance and satisfies:

$$\mathbb{E}[u(x'\theta_H + T' + x'E)] > u(\beta\theta_L + (1 - \beta)\theta_H) \tag{2.38}$$

$$\mathbb{E}[u(x'\theta_L + T' + x'E)] < u(\beta\theta_L + (1 - \beta)\theta_H) \tag{2.39}$$

Only high types would take up this contract, the investor could offer T' so that the first condition binds, and make a profit. Low types would stick with the pooling contract, which then makes certain losses. Thus, the pooling contract cannot be an equilibrium.

Next, it is shown that the only separating equilibrium must be such that type L entrepreneurs obtain full insurance, and type H entrepreneurs only partial insurance. In a separating equilibrium, the zero profit condition must hold for each type separately, so that T_i is pinned down by

$$T_i = (1 - x_i)\theta_i \tag{2.40}$$

Competition implies that one type will get full insurance. Otherwise a profit could be made by offering full insurance to one type. However, contracts $(x_H = 0, T_H = \theta_H)$ (full insurance) and $(x_L > 0, T_L = (1 - x_L)\theta_L)$ (partial insurance) can never satisfy type L's incentive compatibility constraint (2.34):

$$\mathbb{E}[u(x_L\theta_L + T_L + x_L E)] = \mathbb{E}[u(\theta_L + x_L E)] < u(\theta_L) < u(\theta_H) \tag{2.41}$$

Thus, a separating equilibrium must be such that $(x_L, T_L) = (0, \theta_L)$, i.e. type L receives full insurance. Since type H entrepreneurs are also risk-averse, they strictly prefer higher levels of insurance if offered at actuarially fair terms. The terms are ensured by competition, so that the unique separating equilibrium is that with the highest possible level of insurance for type H. It is pinned down by the incentive constraint of type L (2.34):

$$u(\theta_L) = \mathbb{E}[u(x_H\theta_L + (1 - x_H)\theta_H) + x_H E] \tag{2.42}$$

It remains to be shown that such a separating equilibrium always exists. In Rothschild and Stiglitz (1976), a separating equilibrium might fail to exist, if the terms offered to type H are so bad that they would prefer to pool with type L. Here, the worst terms that could possibly be offered to type H entrepreneurs would be no insurance, i.e. $(x_H, T_H) = (1, 0)$. By assumption (A4), a type H entrepreneur would still prefer that contract over the full insurance contract offered to type L $(x_L, T_L) = (0, \theta_L)$:

$$\mathbb{E}[u(\theta_H + E)] = u(C_H) > u(\theta_L) \tag{2.43}$$

This concludes the proof.

2.6.2 Proof of Lemma 12

The proof is analogous to that for lemma (11). It remains true that most surplus is generated when one type gets full insurance and the other as much as possible, given incentive compatibility constraints. One must simply note that the government cannot increase revenue by implementing a tax such that type H entrepreneurs get full insurance and type L entrepreneurs get only partial insurance. To make such contracts incentive compatible, the government would have to pay a subsidy to type L entrepreneurs without generating more revenue. Point (*iii*) is straight forward to see: the government's objective is to maximize revenue, so that leaving any surplus to type H would be wasteful.

2.6.3 Proof of Lemma 6

In the competitive equilibrium, type H's contract $(x_H^C, T_H^C) = (x_H^C > 0, (1 - x_H^C)\theta_H)$ was such that type L's incentive constraint (2.9) binds:

$$\mathbb{E}[u(x_H^C\theta_L + (1 - x_H^C)\theta_H + x_H^C E)] = u(\theta_L) \tag{2.44}$$

Define $R_i^{\mathbb{X}}$ as the absolute risk premium type i would be willing to pay to avoid the risk he would be exposed to from holding all shares of his project at contracts $\mathbb{X}$:

$$\mathbb{E}[u(\theta_i + E)] = u(\theta_i - R_i^{\mathbb{X}}) \tag{2.45}$$

Since the aggregate risk is small (by assumption (A2)), the absolute risk premium can be approximated as

$$R_i^{\mathbb{X}} \approx \frac{1}{2} r_i^{\mathbb{X}} Var(E) \tag{2.46}$$

where $r_i^{\mathbb{X}} = -\frac{u''}{u'}$ is the Arrow-Pratt measure of absolute risk aversion for type i evaluated at the contract $\mathbb{X}$. Using this definition, condition (2.44) can be rewritten as:

$$u(x_H^C\theta_L + (1 - x_H^C)\theta_H - (x_H^C)^2 R_L^C) = u(\theta_L) \tag{2.47}$$

[9] Thus, x_H^C solves

$$(1 - x_H^C)(\theta_H - \theta_L) - (x_H^C)^2 R_L^C = 0 \tag{2.48}$$

In the equilibrium with taxes, the incentive constraint is:

$$\mathbb{E}[u(x_H^G\theta_L + (1 - x_H^G)\theta_H - \tau_H + x_H^G E)] = u(\theta_L - \tau_L) \tag{2.49}$$

and x_H^G analogously solves

$$(1 - x_H^G)(\theta_H - \theta_L) - (x_H^C)^2 R_L^G = \tau_H - \tau_L \tag{2.50}$$

$$(1 - x_H^G)(\theta_H - \theta_L) - (x_H^C)^2 R_L^C = \tau_H - \tau_L + (x_H^G)^2 (R_L^G - R_L^C) \tag{2.51}$$

[9] Using that $Var(xE) = x^2 Var(E)$.

Thus, generally $x_H^G \neq x_H^C$, i.e. the equilibrium sets of contracts differ.

However, there might exist one specific parameterization such that $x_H^G = x_H^C$. Suppose preferences are CARA, so that the absolute risk premium R is constant. If the surplus taxed from both types is exactly the same, then the set of contracts is unchanged. The surplus consists of implementation and insurance gain. The insurance gain with CARA is proportional to the level of insurance obtained. The implementation gain would have to be such that it exactly offsets the difference in insurance. While this is possible, a tiny difference in outside options would already yield a different set of contracts.

2.6.4 Proof of Lemma 1

In the competitive equilibrium, x_H^C solves (2.48):

$$(1 - x_H^C)(\theta_H - \theta_L) - (x_H^C)^2 R_L^C = 0 \tag{2.52}$$

whereas in the equilibrium with taxes, x_H^G solves (2.51):

$$(1 - x_H^G)(\theta_H - \theta_L) - (x_H^C)^2 R_L^C = \tau_H - \tau_L + (x_H^G)^2 (R_L^G - R_L^C) \tag{2.53}$$

Since preferences are NIARA, $R_L^G \geq R_L^C$. Thus, when $\tau_H > \tau_L$, the right hand side of (2.53) is positive, so that $x_H^G < x_H^C$.

Chapter 3

Adapting to Capitalism: Private Health Insurance Uptake Among Former East Germans

I study the differential reaction of former East and West Germans to a series of health care reforms that started in 1997. Along with the gradual decrease in coverage under the public health insurance system, former East Germans were significantly less likely to sign complementary health insurance contracts in the private market. I show that the differential uptake rates of additional private insurance after the reforms are consistent with a model in which agents optimize their individual insurance status only if they are aware of the organizational form of the health care system (or more generally the welfare state), and in which East Germans are initially less likely to have the correct beliefs, but learn over time that institutions have changed and they are now responsible for optimizing their insurance coverage. While it is widely recognized that the development of new institutions in transition economies takes time, people's adjustment to them has received little attention. This study provides evidence for the existence of a substantial transition period in the individual adaptation to new institutions.

3.1 Introduction

When transitioning from socialism to a free-market economy, do people adapt to the new circumstances immediately? Undoubtedly, major shifts in the political system do not escape people's notice. They often follow extended demonstrations, spectacular coups, or even violent uprising. However, the institutional changes that go along with this transition, although fundamental, may not be apparent immediately. For instance, an all-encompassing welfare state with an extreme level of redistribution is a core idea of socialism. Confronted with a capitalist system, do people understand their new individual responsibility immediately, or do they adapt their decision making over time? Economic transition consists of both the development of new institutions and people's adaptation to them. While it is now widely recognized that the former takes time, the latter has received little attention so far. Using data on uptake rates of private health insurance of former East and West Germans, I find that the individual adjustment period can be substantial.

The case of Germany presents a unique opportunity to study the question. The reunification of the socialist East with the capitalist West in 1990 came after 40 years of separation[1]. While in the past two decades many Eastern European countries have started to transition from communism to western-oriented democracies, two characteristics of the German case make it especially suitable to study people's reaction to institutional change. First, the influence of socialism can be interpreted as an exogenous shock. The division of Germany was not a choice of the German people, but imposed by the Allied Forces, and the new border determined by where the forces were standing at the end of World War II. During the time of separation, migration

[1]Political separation of the two German states lasted for 40 years. After World War II, Germany had been divided by the allied forces into four occupation zones, i.e. a British zone in the North West, a French zone in the South West, a U.S. zone in the South, and a Soviet zone in the East of Germany. Due to the aggravating political situation among the Soviet Union and the three Western powers, the American and British zone were merged in 1946 and were joined by the French zone in 1947. In May 1949, the Federal Republic of Germany (FRG) was founded on this territory, while in October 1949 the Eastern zone became the German Democratic Republic (GDR). However, only the erection of the Berlin Wall in August 1961 completed the physical separation.

was minimal[2]. Reunification then came rather surprisingly. The large protest that lead to the Fall of the Berlin Wall in November 1989 had started only two months earlier. Reunification of the two German states was finalized a mere year later. Former West Germans thus constitute a credible control group for former East Germans. Second, with the Unification Treaty (1990), East Germany implemented the political and economic system of the FRG in its pre-existing form. New institutions did therefore not need to be developed; they were already well-functioning and rapidly imposed onto the East German population. Any observed adaptation process can thus be interpreted as people adjusting to the new institutions rather than the parallel development of these institutions.

I study the differential reaction of former East and West Germans to a series of health care reforms that started in 1997. Before 1990, both German health care systems had provided almost universal coverage. Their organization however differed: While in East Germany all health care provision had been state owned and health care free to citizens, West Germany has had a market for health services, and a public health insurance - funded out of payroll taxes - had provided extensive coverage to the vast majority of the population. With reunification, the market-based system was implemented in East Germany. However, since coverage remained de facto the same, the institutional change may not have been immediately apparent to East Germans at the time. It was not until 1997 that more than the organizational details of the health insurance system changed. I find that following the reform shocks, along with the decrease in coverage under the public health insurance system, former East Germans were significantly less likely to sign complementary health insurance contracts in the private market than former West Germans. Such differences could be driven by demographic factors, differences in risk attitudes, or aggregate economic effects that differ between the Eastern and Western parts of Germany. However, the different

[2]Until the fall of the Wall in 1989, migration was minimal in either direction due to violent military border protection in the GDR coupled with a rigorous restriction of the number of GDR citizens officially allowed to travel (Küsters and Hofmann (1998)), and very little migration from prospering West Germany to the East (Münz and Ulrich (1997)).

reaction to the decrease in coverage could also be attributed to East Germans being less aware of the fact that they are now responsible for their own insurance status, and are able to buy private insurance. After living in a socialist regime, they only adapt over time to the capitalist institutions of the unified Germany. The goal of this paper is to isolate this effect, while controlling for the others.

In particular, I show that the uptake rates of additional private insurance after the reforms are consistent with a model in which agents optimize their individual insurance status only if they are aware of the organizational form of the health care system (or more generally the welfare state). East Germans are initially less likely to have the correct beliefs, but learn over time that institutions have changed and they are now responsible for optimizing their insurance coverage. Moreover, I show that older East Germans, i.e. those who have lived with the socialist institutions longer, are even less likely to have correct beliefs than their younger equivalents. Thus, this study provides evidence for a transition period in people's adaptation to new institutions.

Related Literature

Since the fall of Communism in Central and Eastern Europe, the interest in what determines the speed and success of the transition process has been strong. Two decades after the transition started, varying experiences of the reforming economies inspire a growing literature to highlight the impact of institutions on the transition path. Murrell (2008) provides an overview. Investigations on how institutions develop in economies following major political change have considered many factors that might determine the success of the transition, for example international assistance (Cochrane (2007)). Yet, little notice has been taken of how people's reaction influences this process. Arguably, the success and speed of transition hinge critically on how well the people living in transition economies are able to adapt to institutional change. This paper aims at providing evidence for the existence of a substantial transition period in this adaptation process.

Only a few studies have analyzed behavioral differences between former East and West Germans. Bucher-Koenen and Lusardi (2011) report that in 2009, financial literacy was lower among people living in the eastern states of Germany, a fact that they interpret as former citizens of the GDR having not yet caught up to West Germans in terms of financial education. The authors then link financial literacy to retirement planning decisions. Sauter (2009) analyzes participation rates in security markets among East Germans. He finds habit persistence to be a strong explanatory factor for the low participation. Fuchs-Schündeln (2008) focuses on the differential savings behavior of East and West Germans after reunification and finds that they are consistent with a life-cycle consumption model with precautionary savings. Differential behavior could also stem from differences in preferences. Alesina and Fuchs-Schündeln (2007) analyze the effect of Communism on an individual's taste for public social policy. They find that former East Germans are more likely to favor a high state responsibility, pointing to a possible feedback effect of political regimes on policy preferences. From a different perspective, sociologists have developed a literature on social capital formation in the transition economies of Eastern Europe. The emphasis here has been on the establishment of informal networks, and the formation of trust after the centrally planned and controlled systems broke down (see Keefer and Knack (2005) for a survey).

The strategic management literature has explored how firms adapt to a new market-oriented economy. Apart from the challenges of privatization (Uhlenbruck and de Castro (2000)), and organizational restructuring (Filatotchev et al. (2000)), also a learning process has been documented: Kriauciunas and Kale (2006) find that while the so-called imprinting effect of the socialist environment adversely affects firms' ability to change their operating knowledge, firms that search for new knowledge from distant sources (i.e. non-socialist countries) are able to successfully change their knowledge to meet the demands of the new market-oriented economy. Murrell (2005) presents related literature, Djankov and Murrell (2002) survey empirical evidence.

The remainder of the paper is organized as follows: Section 3.2 provides an institutional background about the German health insurance system before and after 1990. Section 3.3 outlines a simple model of exogenous learning to frame the subsequent data analysis. In Section 3.4 I describe the data set used; section 3.5 reports the results of the empirical analysis. Specific emphasis is placed on differences across age groups (section 3.5.1), risk taking and risk aversion (section 3.5.2) as well as the role of preferences for a bigger welfare state (section 3.5.3). Last, I analyze potential regional aggregate effects (section 3.5.4). Section 3.6 concludes.

3.2 Health Insurance in Germany

Public health insurance has a long tradition in Germany. Introduced in 1883, and initially for workers of certain industries only (miners, guilds, factory workers), it represents one of the first national social insurance systems in the world. In the following decades, public health insurance coverage was gradually extended to cover larger parts of the population (10% in 1885, 51% in 1925). Its core elements of being mandatory, pay-as-you-go, financed by both employers and employees, and being managed by so-called sickness funds persist to date (Busse and Riesberg (2004)).

After the country's political separation that followed the end of World War II, along with the divergent political systems, the health care system developed quite differently in the two new states. In the FRG, the national health insurance system was continued in a market-based format in which health insurance was mandatory but could be obtained through the public system or from private insurers. Public health insurance was the overall dominant form, with 83% of the population being covered by 1960 and 88% by 1987. Financing was organized through equal contributions from employers and employees (6% of income in 1950; 12.6% in 1987). Co-payments for benefits were only nominal (WHO (2000)). Publicly insured individuals could also purchase complementary private coverage for select cases, but only a very small frac-

tion of the population actually did[3]. Only for state employees and self-employed, often wealthier, individuals private insurance was required.

In the GDR, the social insurance was maintained in principle with a health insurance system that was universally mandatory. Nearly 100% coverage rate was provided by only two large managing bodies ("sickness funds"): One for employees, workers, and their families (covering 89% of population), and one for professionals, members of agricultural cooperatives, artists and self-employed and their families (covering 11% of population). De facto, however, the importance of the social insurance system was only very limited as the majority of health care providers and facilities were state employed and owned, so that health care was free of cost to citizens and supplemental health insurance was not needed (Busse and Riesberg (2004)).

With the German reunification, the GDR introduced the health insurance law and system of the FRG (Article 21, Unification Treaty (1990)), integrating 17 million former East German citizens into the existing system of the FRG. Health insurance continued to be mandatory in the unified Germany, with public health insurance being the predominant form of provision[4]. The extensive benefits of the public health insurance included coverage of almost all health care as well as benefit payments to compensate for salary loss during recovery[5].

[3]According to the Association of Private Health Insurers, less than 3% of people had any sort of private health insurance contract, including travel insurance.

[4]88% of the German population was covered under the public health insurance system in 1997 and in 2003. 10% of the population were covered by private health insurance, including nearly 4% civil servants with free governmental care and complementary private insurance. 2% of the population were covered by other governmental plans (e.g. military, social welfare) and another 0.2% (mainly self-employed) have no prepaid coverage for health care. Individuals who buy full, i.e. not only complementary, private health insurance opt out of public health insurance system and its financing (WHO (2000)).

[5]Coverage benefits are described in the Social Code Book V, and generally include prevention of disease and health promotion at the workplace, screening for disease, treatment of disease (including inpatient and outpatient care, dental care, medication, medical devices, rehabilitation, etc.), and emergency care. Benefit payments are managed by sickness funds, paying their employed insured individuals 70% of the last gross salary (max. 90% of net salary) for from week 7 up to week 78 of certified illness, while employers continue to pay 100% of the salary during the first 6 weeks of sickness (WHO (2000)).

3.2.1 Health Care Reforms

Even before 1990, both health care systems had suffered from financial problems. In the GDR, severe under-financing, personnel shortages, and lack of modern medical equipment and supplies led to the erosion of quality of care in the 1970s and 80s. Public health indicators fell behind Western standards, e.g. regarding infant mortality and life expectancy, so that in 1989 a national health conference decided extensive health care reform - but with the fall of the Berlin Wall the GDR ceased to exist that year. In the FRG, an era of cost-containment had already started in 1977 with the introduction of the Health Insurance Cost-Containment Act to ensure stability in contribution rates, and aimed at increasing technical and allocative efficiency (Busse and Riesberg (2004)).

After the health insurance system of the FRG was implemented in the former GDR, aggravating demographic trends, price increases in medical supplies, increasing wages, and a trend of high income individuals to opt for private insurance further increased the cost pressure on the public health insurance system. As a result, a long series of reforms set in. While initially focused on cost-containment through increased efficiency, measures shifted towards cutting benefits, and increasing co-payments as well as contribution rates in the mid 1990s. Three reform acts in 1997 increased the co-payments for medication, hospital stays, and dentures. The reimbursement for glasses were eliminated, and allowances for preventive and rehabilitative care substantially decreased. From 1998 to 2000, under a new government, some of these measures were revoked, but re-introduced and expanded in smaller reforms in 2002 and 2003, and in the Public Health Insurance Modernization Act of 2004 (Steffen (2011)).

As a result, an increasing number of publicly insured individuals sought complementary private health insurance for benefits previously covered by the public health insurance. While in 1990 less than 3% had additional private health insurance, this

number rose to 9.3% in 2002[6]. According to the Association of Private Health Insurers (2004), the most commonly bought private insurance policies cover dental benefits, specialty medication, procedures by chief physicians, and hospital accommodation in private rooms. However, the trend towards additional coverage is not uniform across the German population. This paper documents that former East Germans are less likely to sign a complementary health insurance contract than West Germans.

3.3 A Simple Model With Exogenous Learning

Suppose an economy is populated by a continuum of agents who potentially differ along many dimensions. Let $x_{i,t}$ denote a vector of individual characteristics that contains all information about agent i's socio-economic situation in period t, his risk preferences, as well as any other attributes that may influence his choice of insuring against health risks. The government of this economy influences an agent's insurance status through public provision of health insurance, financed by taxes. Thus, the public health insurance system is part of a solidarily financed welfare state. I denote the government's choice of public health insurance contracts in period t with $\mathbb{G}_t$.

The private market offers a set of insurance contracts $\mathbb{I}_t$ that complement public health insurance. I assume that private insurers do not offer an actuarially fair contract for each individual agent, so that some agents in the population are better off not signing any complementary health insurance. Various kinds of market restrictions or imperfections may cause this incompleteness. For example, insurers might not be allowed to discriminate along all possible dimensions, even if the relevant characteristics of an agent are verifiable. In Germany, private health insurance carriers are not allowed to discriminate individuals based on their origin (e.g. former East vs. former West), and are also prohibited from extensive health screenings (e.g. for

[6]Estimates by the Association of Private Health Insurers (2004). The official German micro census does not make it mandatory to answer questions about private insurance contracts. Private insurers do publish summary statistics of the number of insurance contracts signed. However, these are an overestimate of people with additional health insurance, since they include double counting of people with more than one contract and also include e.g. private travel health insurance contracts.

cancer) before approving applicants to private health insurance (Busse and Riesberg (2004)). Bolton and Dewatripont (2005) discuss Akerlof's 1970 incomplete contracts setting in an insurance framework. When principals are constrained to offer only one insurance contract to a heterogeneous group of agents, some low risk agents might prefer not to be insured over pooling in an insurance contract where they subsidize higher risk types.

3.3.1 Heterogeneous Beliefs about Institutions

Agents have heterogeneous beliefs about the institutional goal of public health insurance. Some agents correctly believe that the scope of the welfare state reflects political and aggregate economic constraints. They are aware that the specific coverage under the public health insurance might not necessarily be enough for them individually. Other agents wrongly believe that the welfare state is all-encompassing and health related costs are completely nationalized as under a socialist regime. These agents think that public health insurance provides the necessary coverage for all agents and that there is no need for them to consider complementing their individual health insurance with a private policy.

Let $b_{i,t} \in \{0,1\}$ denote agent i's belief in period t. In period $t = 0$ a fraction α of the population has the wrong belief ($b_{i,0} = 0$) about the scope of the welfare state. Every period a fraction x of the population learns the true type of the government. This exogenous learning shock is idiosyncratic and independent of agents' individual characteristics $x_{i,t}$ and beliefs $b_{i,t}$. If an agent with the correct belief $b_{i,t} = 1$ receives the shock, nothing changes. An agent with the wrong belief, however, changes his views and thus possibly his behavior, now realizing that he should optimize his insurance status by himself.

Agents can perfectly observe $\mathbb{G}_t$ and $\mathbb{I}_t$. They choose to sign a contract in the private insurance market if it maximizes their expected utility, given their beliefs. Let $y_i \in \{0,1\}$ denote agent i's decision to sign a complementary health insurance

contract with a private insurer. Then

$$y_{i,t} = 1 \quad \text{if} \quad w_{i,t} = b_{i,t} * y^*_{i,t} > 0 \tag{3.1}$$

$$b_{i,t} \in \{0, 1\} \tag{3.2}$$

$$y^*_{i,t} = \mathbb{E}\left[U(y_{i,t} = 1) \mid x_{i,t}, \mathbb{G}_t, \mathbb{I}_t\right] - \mathbb{E}\left[U(y_{i,t} = 0) \mid x_{i,t}, \mathbb{G}_t, \mathbb{I}_t\right] \tag{3.3}$$

3.3.2 Comparing Two Populations

The purpose of this paper is to compare two populations, former East and West Germans (denoted with superscript $j = E, W$ respectively). They now live in the same economy and share the government as well as the private insurance market, so that $\mathbb{G}_t$ and $\mathbb{I}_t$ are the same for both populations. The two populations possibly differ in demographic make up. Most importantly, however, different shares of these populations initially have the wrong beliefs about the welfare state. In particular, I assume that among the former East Germans the wrong beliefs are initially more prevalent, i.e. $\alpha^E > \alpha^W$.

From an aggregate point of view, a fraction Y^j_t of agents in each population signs a complementary health insurance contract. Given $\mathbb{G}_t$ and $\mathbb{I}_t$, these aggregate insurance levels depend on the agents' beliefs as well as the demographic characteristics of the respective population:

$$Y^j_t = Y^{j*}_t * (1 - \alpha^j + \alpha^j x \sum_{n=0}^{t-1} (1 - x)^n) \tag{3.4}$$

Since the populations of former East and West Germans potentially have different demographic characteristics, there is no reason to believe that $Y^E = Y^W$ in any period. The *optimal* levels Y^{j*}_t, however, are unobservable, so that observed differences

in the *actual* aggregate levels of complementary health insurance Y_t^j do not allow conclusions about differences in beliefs.

The probability for any individual agent to take up complementary health insurance is

$$Pr(y_{i,t} = 1) = Pr(b_{i,t} = 1) * Pr(y^*_{i,t} > 0) \tag{3.5}$$

Being East or West German only influences the probability of having the right beliefs about the welfare state. By setup, the probability of getting insurance once the agent realizes he has to optimize, is independent from him being a member of either population. Thus, a regression with individual level data, that includes a dummy for being East German, as well as controls for individual characteristics $x_{i,t}$, yields a coefficient

$$\beta_t^E = Pr(y_{i,t} = 1 \,|\, x_{i,t}, East) - Pr(y_{i,t} = 1 \,|\, x_{i,t}, West) \tag{3.6}$$

Since only the decision for or against additional insurance, $y_{i,t}$, is observable, and the latent variable $y^*_{i,t}$ is not, β^E is only identified off individuals who would sign an insurance if they had the right beliefs, i.e.

$$\begin{aligned}
\beta_t^E &= Pr(y_{i,t} = 1 \,|\, y^*_{i,t} > 0, East) - Pr(y_{i,t} = 1 \,|\, y^*_{i,t} > 0, West) && (3.7)\\
&= Pr(b_{i,t} = 1 \,|\, East) - Pr(b_{i,t} = 1 \,|\, West) && (3.8)\\
&= (\alpha^W - \alpha^E)(1 - x \sum_{n=0}^{t-1} (1-x)^n) && (3.9)
\end{aligned}$$

Since $\alpha^W < \alpha^E$, and $0 < x < 1$:

$$\beta_t^E < 0 \quad \text{and} \quad \beta_t^E < \beta_{t+1}^E \quad \forall t \tag{3.10}$$

East Germans in this model are less likely to buy additional health insurance at any

point in time because of their lower probability to have the correct beliefs. However, as they learn over time, this effect decreases, and converges zero. This simple model is intended to frame the following data analysis. While it cannot be concluded that this particular structure is the true underlying process, the data does provide evidence for the presence of a learning or adaptation process.

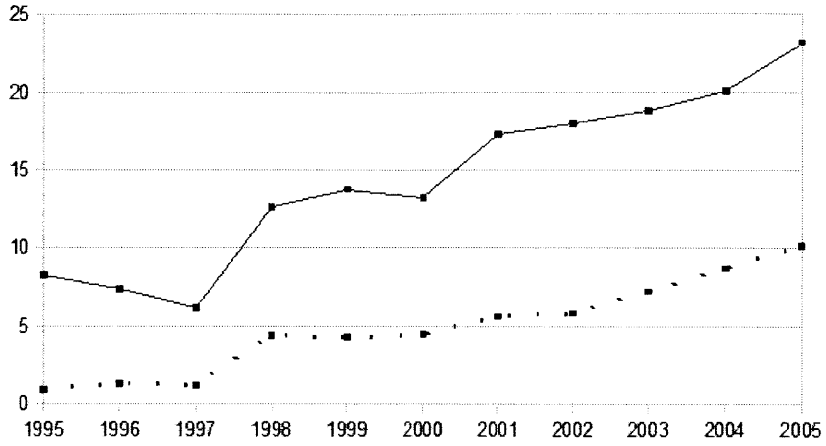
3.4 Data

I use data from the German Socio-Economic Panel (GSOEP (2007)). Since 1984, the German Institute of Economic Research "DIW" conducts a yearly survey of households and their members, representative of the population in the FRG. Since 1990, the study also represents the population in the "new" German states, the regions formerly belonging to the GDR. The GSOEP survey mainly covers basic information on population demographics, education, training, and qualification, labor market and occupational dynamics, earnings, income and social security, health and household production. Moreover, the survey has an emphasis on aspects of basic political orientation, preferences, values and satisfaction in life (Haisken-DeNew and Frick (2005)).

The original sample of the panel included only West-Germans. After 1990, a sample representative of the former GDR was added. In 2002, the sample was supplemented by a high-income group, so that since then the survey is representative of the German population without a truncation on income. The sample used in this paper includes only participants who were born in either part of Germany before reunification in 1990. Moreover, I restrict attention to people covered under the public health insurance system. The dependent variable is a dummy indicating whether the respondent has signed a complementary health insurance contract with a private insurer. This information is only available starting in 1995 up to 2005. The so restricted data set leaves a total of 126,346 observations for the analysis out of which 43,513 are responses of former East Germans, and 82,833 of former West Germans.

Figure 3-1 reports the fractions of East and West Germans with complementary

Figure 3-1: Population Shares with Complementary Health Insurance



This figure shows the shares of former East (dashed) and West (continuous) Germans in the sample who had a complementary health insurance contract in the respective year.

health insurance in the sample. Notice that the population shares with additional health insurance do not seem to converge over time. However, recall that the model only predicts aggregate convergence *conditional* on all characteristics of the population. The two German populations, though, are very different. The age structure differs between East and West Germans: East Germans are younger on average. Birth rates in East Germany were much higher than in West Germany. In 1980, the average number of births per woman was 1.9 in the GDR, but only 1.3 in the FRG (Pötzsch (2007)). Moreover, health care was worse in East Germany, resulting in lower life-expectancy, and so contributing to a younger population (Busse and Riesberg (2004)). An even more striking difference exists in terms of income. At the beginning of the sample, in 1995, former East Germans on average have 23% less income than West Germans. More importantly, this gap widens over time, and is 35% in 2005. Fuchs-Schündeln et al. (2010) have recently documented this divergence. There is also a difference in home ownership, which can be interpreted as a proxy for wealth: 31% (25%) less East Germans than West Germans own a home in 1995

(2005). Among East Germans, despite the age structure, both unemployment and retirement rates are higher - a trend that increased over time as well[7]. Given these differences, it is to be expected that the population share with additional insurance is higher (and possibly increasing much faster) among West Germans than among East Germans. On the other hand, East Germans are slightly more likely to have a high school and college degree, a factor that might work in the opposite direction. Yet, the education under the socialist regime likely did not increase knowledge about financial planning and private insurance. Moreover, the model implies that a reform shock has a stronger impact on the number of private insurance contracts the more people are actually optimizing (i.e. have the right beliefs). After the first health care reform in 1997, the increase in additional coverage is much stronger among the West Germans than East Germans. Following later reforms, the trends are much more equal (and in the last 5 years of the they are almost exactly the same). This pattern is consistent with the two populations converging to the same fraction of people with the correct beliefs over time. While these factors might explain why the insurance uptake rates have not converged, it is impossible to conclude from the aggregate numbers whether the proposed learning mechanism is at work. The remainder of the paper is thus concerned with an individual level analysis.

The main explanatory variables in the analysis are a dummy that takes on the value one if the respondent lived in East Germany before reunification, and interactions of the East dummy with year dummies. These variables capture the effect of being East German on the probability that the individual has complementary health insurance separately for each year the survey was taken, and so allow conclusions about the effect of being used to a different welfare state, and how it evolves over time. To control for each individual's socio-economic situation, I use three sets of controls in all regressions. The first set contains relevant socio-demographic information on age,

[7]In 1995 (2005), the unemployment rate was 8.1% (9.9%) in the West German states but 13.9% (18.7%) in the former East German territories (Statistik der Bundesagentur für Arbeit (2011)). After reunification, the German government promoted early retirement by not cutting retirement benefits too much, which led to a higher retirement rate among East Germans.

marital status, the individual's perceived health status, as well as his satisfaction with health. The second set includes information about the level of education and the current employment status of the respondent. The third set of controls covers income and wealth data: gross and net income as well as asset income on the household level, and a dummy variable for homeownership. Table 3.10 in the appendix reports summary statistics for all variables.

The GSOEP study periodically covers special topics: In 1996 and 2001, respondents were asked about their preferences regarding the scope of the welfare state. In particular, it was asked: "Who should be responsible for financial security in case of illness?" I generate a dummy variable indicating that the answer was "only the state" or "mostly the state", as opposed to "both the state and private forces", "mostly private forces", or "only private forces". In 2003, questions about risk taking behavior and risk aversion were included in the study. Respondents were asked to indicate on a scale from 0 to 10 how willing they were to take risks. Moreover, people were asked to indicate how much of a hypothetical lottery win they would be prepared to invest in a financially risky, yet lucrative investment. I use these special topic controls in sections 3.5.2 and 3.5.3.

3.5 Results

Table 3.1 reports the results from the baseline probit regression[8]. It includes the East dummy and the three sets of basic controls. The dependent variable is an indicator for whether the respondent has signed a complementary health insurance contract in the private market. As a robustness check, I also estimate the linear probability model (see table 3.11 in the appendix). While the OLS coefficients confirm the general results, naturally the size of the coefficients is very different in these two regressions.

[8]All tables report total coefficients. Most interpretations rest on the sign or relative size of coefficients. Marginal effects are only reported in the text, whenever the absolute size of an effect is of interest.

Table 3.1: Basic Regression

Dependent Variable: Complementary Health Insurance†	Coefficients (Standard Errors)		Coefficients (Standard Errors)
East†	-0.517***		
	(0.008)		
Age	-0.022*	High school degree†	0.334***
	(0.010)		(0.051)
Age squared $*10^{-3}$	0.591	Vocational training †	0.197***
	(0.338)		(0.059)
Age cubed $*10^{-5}$	-0.399	Completed college degree†	0.239*
	(0.259)		(0.119)
Female†	0.116***	Full or part time employed†	0.071*
	(0.033)		(0.036)
Married†	-0.083	Unemployed†	-0.076
	(0.052)		(0.050)
Married but separated†	0.231	Retired†	-0.019***
	(0.139)		(0.003)
Divorced†	0.070	Log (household income)	0.041***
	(0.038)		(0.007)
Widowed†	-0.148***	Log (household net income)	0.244***
	(0.023)		(0.017)
Health status	-0.012***	Log (household asset income)	0.055***
	(0.002)		(0.006)
Satisfaction with health	0.005	Homeowner†	0.034***
	(0.004)		(0.007)
Constant	-4.473***		
	(0.234)		
Observations	126,346		
Log likelihood	-41,341		

Probit regression. Omitted categories are male, single, intermediate schooling, not employed. The variable "Health status" ranges from 1 (very good) to 5 (bad). The variable "Health satisfaction" ranges from 0 (low) to 10 (high). Standard errors are clustered by East.
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

The most important explanatory variable is the East dummy. An East German respondent is significantly less likely to purchase complementary health insurance in the private market. To establish a learning effect, however, one needs to analyze how the effect of being from East Germany changes over time. Consider table 3.2, which reports the results of the same basic regression, with the set of explanatory variables

expanded by year dummies and East-year interactions. While the coefficient on East in table 3.1 measured the average effect of being East German over all eleven years included in the sample, it now corresponds to the same effect specifically in 1995. As in the basic regression, in 1995, East Germans were significantly less likely to buy additional health insurance. However, this effect vanishes over time. The coefficients on all East-year interactions are positive and significant. Thus, in every year, the East effect is less strong than in 1995. Table 3.3 reports the results of the one-sided hypothesis tests that the coefficients on the East-year interactions are indeed increasing over time. Except for 1999, 2001, and 2005, the East effect did shrink significantly in every year. Even for the three years it did not, the decrease was small. In any given year, the effect of having lived in the GDR on the probability to sign a complementary health insurance contract is the sum of the coefficients on East and the respective East-year interaction. It increases from -0.978 in 1995 to -0.434 in 2005, a fifty percent reduction within the eleven years of the sample.

The coefficients in table 3.2 measure the increase in the probability of signing a private insurance contract since the base year 1995. Recall from equation (3.5) that according to the learning model introduced, this probability is determined by two independent effects: It stems both from the change in the latent variable $y^*_{i,t}$, influenced by policy reforms, and the learning effect. The coefficients on the year dummies document that the likelihood for West Germans to buy additional health insurance also increases every year, starting in 1998, after the first big health care reform. However, notice that in each year, the east-year interaction coefficient is larger than the one on the year dummy. The probability of buying additional health insurance increases faster among East Germans than among West Germans. This is in line with the predictions of the model: While the change due to reforms is the same for all Germans, the learning effect is stronger for East Germans, simply because they started with a larger fraction of people with the wrong beliefs ($\alpha^E > \alpha^W$). While these results are evidence for a stronger learning effect among East Germans, one cannot conclude much about the speed of convergence. Reforms to the public health insurance system

Table 3.2: Basic Regression with East-Year Interactions

Dependent Variable: Complementary Health Insurance†	Coefficients (Standard Errors)		Coefficients (Standard Errors)
East†	-0.978***		
	(0.012)		
East * 1996†	0.230***	1996†	-0.095***
	(0.006)		(0.007)
East * 1997†	0.269***	1997†	-0.198***
	(0.004)		(0.013)
East * 1998†	0.470***	1998†	0.245***
	(0.001)		(0.001)
East * 1999†	0.403***	1999†	0.271***
	(0.008)		(0.012)
East * 2000†	0.458***	2000†	0.258***
	(0.008)		(0.014)
East * 2001†	0.444***	2001†	0.353***
	(0.015)		(0.023)
East * 2002†	0.460***	2002†	0.365***
	(0.012)		(0.026)
East * 2003†	0.512***	2003†	0.406***
	(0.014)		(0.027)
East * 2004†	0.548***	2004†	0.453***
	(0.008)		(0.028)
East * 2005†	0.544***	2005†	0.564***
	(0.007)		(0.032)
Observations	126,346		
Log likelihood	-40,414		

Probit regression. Omitted categories are 1995, the interaction of East and 1995, male, single, intermediate schooling, not employed. Standard errors are clustered by East. Table continued in the appendix (table 3.12).
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

change the baseline probability of any individual to seek additional coverage almost every year since 1997. Only if $\alpha^W = 0$, i.e. all West Germans are assumed to have the correct beliefs (so that their learning effect is zero), would the difference between East an West Germans in any given year correspond only to a learning effect among the former East German population.

The effects of the remaining controls are as expected. Among the socio-demographic controls shown in the first column of table 3.1, only the female and widowed dummies

Table 3.3: Changing East-Year Interactions - Test Statistics

H_0 :	Test statistic	p-value
0≥ East*1996	1376.47	0.000
East*1996≥ East*1997	308.93	0.000
East*1997≥ East*1998	5690.13	$1.87*e^{-69}$
East*1998≥ East*1999	92.98	1
East*1999≥ East*2000	$1.2*e^{5}$	0.000
East*2000≥ East*2001	4.54	0.9834
East*2001≥ East*2002	42.92	$2.85*e^{-11}$
East*2002≥East*2003	2194.76	0
East*2003≥ East*2004	37.56	$4.42*e^{-10}$
East*2004≥ East*2005	19.41	0.9999

This table reports χ^2 test statistics and p-values for the one-sided hypothesis tests that the East-year coefficients of the basic regression (table 3.2) are increasing over time.

are significant. Women are significantly more likely to buy additional health insurance. They might face higher health risks, at least for the types of circumstances not covered under the public health insurance system, or be more risk-averse than men (Borghans et al. (2009) document such gender differences in risk aversion.). Moreover, the better a respondent perceives his own health, the more likely he is to have a complementary health insurance[9]. The interpretation of this effect is complicated by a potential endogeneity problem: Health could be positively correlated with risk aversion, resulting in healthier people to be more likely to buy insurance. However, the reverse might also be true: People with more coverage could be healthier because they use the benefits of the insurance and see a doctor more often, or take advantage of the many preventive measures usually reimbursed and heavily advertised by insurance companies.

[9]The variable "Health status" ranges from 1 (very good) to 5 (bad).

Surprisingly, a respondent's age has only a small and barely significant effect on his likelihood to buy insurance. Since health risks typically increase with age, this coefficient should naturally be positive. However, one has to keep in mind that the age coefficient in this basic regression might capture a variety of effects: Older people might be more likely to need insurance, but also more prone to the wrong beliefs about the welfare state, especially if they lived in the GDR. Section 3.5.1 is aimed at decomposing these effects.

Higher educated respondents are more likely to buy health insurance. They might be better able to understand their own risk structure or the offers in the private market. Such an argument has been made for other areas of economic decision making. Lusardi and Mitchell (2009), for example, link the level of education obtained to financial literacy and document a positive causal effect from financial literacy to retirement planning efforts. All income and wealth variables have positive and significant coefficients. One might interpret this result along the same lines as the education effect: wealthier individuals tend to be more financially literate. It could also be the case that complementary health insurance is seen as a luxury good. A low-income individual might for example choose to neither spend any money on artificial dentition, nor on an insurance policy that would cover such costs. All severe health risks are covered under the public health insurance. Except for being retired, which has a negative effect on the likelihood to purchase additional insurance, the employment status of an individual seems to not play a significant role - the coefficient on full or part time employed is only significant at the 10% level.

3.5.1 Age and Cohort Effects

In terms of health risks, age is an important factor. The health care reforms have cut benefits for artificial dentition and glasses, and so have a particularly strong impact on older people. Moreover, the number of drugs regularly prescribed on average increases with age, so that a higher co-pay affects the older population more. Given the

higher risk they face, older people should be more likely to have additional coverage. Yet, age did not turn out to be a highly significant driving factor in the baseline regression. In this specification, however, age might absorb two opposing effects. While older people in general might be more likely to buy additional health insurance, they were also exposed to the respective political regime the longest. It might be the case that at least among former East Germans older people are more likely to have the wrong beliefs about the welfare state, and are therefore *less* likely to buy private insurance contracts than their younger countrymen. To disentangle these effects, I run a regression that includes an East-age interaction, as well as age-year and East-age-year interactions. To ease the interpretation, I do not include age squared or cubed as regressors. Table 3.4 reports the results.

Consider the second column of table 3.4. The effect of age among West Germans in any specific year is measured by the sum of the coefficients on age and the particular age-year interaction. Notice that for all years this sum is positive. Thus, older people are indeed more likely to buy insurance. However, this effects gets smaller over time. The coefficient on every age-year interaction is negative, indicating that age has less and less impact over time. Considering the reforms to the public health insurance system, this is only natural: ever decreasing coverage makes it increasingly necessary also for younger individuals to buy additional insurance, so that the age gap is shrinking over time.

The striking result of this regression is that among East Germans, older respondents are actually *less* likely to buy insurance. The obvious interpretation is that older East Germans have lived longer under the socialist regime and its all-encompassing welfare state, so that they are less likely to have the right beliefs than younger East Germans. This negative effect more than compensates the positive effect of age on the likelihood to buy additional insurance. The positive coefficients on the East-age-year interactions confirm this view: Over time, the probability of having the right belief converges across age groups, and the positive age effect becomes more dominant - a

Table 3.4: Age Regression

Dependent Variable: Complementary health insurance†	Coefficients (Standard errors)		Coefficients (Standard errors)
East†	0.0259		
	(0.023)		
East * Age	-0.026***	Age	0.006***
	(0.001)		(0.0002)
East * Age * 1996	-0.008***	Age * 1996	0.005***
	(0.0001)		(0.00003)
East * Age * 1997	0.007***	Age * 1997	0.005***
	(0.0003)		(0.000002)
East * Age * 1998	0.011***	Age * 1998	-0.003***
	(0.0003)		(0.00005)
East * Age * 1999	0.008***	Age * 1999	-0.0002***
	(0.0003)		(0.00002)
East * Age * 2000	0.006***	Age * 2000	-0.002***
	(0.0003)		(0.000006)
East * Age * 2001	0.006***	Age * 2001	-0.0009***
	(0.0002)		(0.00004)
East * Age * 2002	0.006***	Age * 2002	-0.001***
	(0.0003)		(0.0001)
East * Age * 2003	0.010***	Age * 2003	-0.001***
	(0.0003)		(0.0001)
East * Age * 2004	0.010***	Age * 2004	-0.002***
	(0.0003)		(0.0001)
East * Age * 2005	0.0129***	Age * 2005	-0.003***
	(0.0003)		(0.0002)
Observations	126,346		
Log likelihood	-40,139		

Probit regression. Omitted categories are 1995, the interaction of East and 1995, male, single, intermediate schooling, not employed. Standard errors are clustered by East. Table continued in the appendix (table 3.13).
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

pattern consistent with the exogenous learning model.

To further substantiate the claim that older East Germans are even less likely to have the right beliefs than younger East Germans, I look at separate regressions for different groups of cohorts. I divide the sample into 5 groups: those born between 1975 and 1989, those born between 1965 and 1974, those born between 1955 and 1964, those born between 1945 and 1954, and those born before 1945. East Germans

Table 3.5: Basic Regression For Different Age Groups

Dependent Variable: Complementary Health Insurance†	Born 1975-1989	Born 1965-1974	Born 1955-1964	Born 1945-1954	Born before 1945
East†	-0.332***	-0.579***	-0.808***	-1.195***	-1.594***
	(0.022)	(0.001)	(0.016)	(0.030)	(0.055)
East * 1996†		0.276***	0.204***	-0.016	0.003
		(0.001)	(0.007)	(0.013)	(0.010)
East * 1997†		0.383***	0.282***	0.153***	0.346***
		(0.002)	(0.019)	(0.017)	(0.006)
East * 1998†	0.179***	0.139***	0.359***	0.549***	0.761***
	(0.014)	(0.004)	(0.007)	(0.022)	(0.014)
East * 1999†	0.117***	0.159***	0.338***	0.561***	0.537***
	(0.016)	(0.003)	(0.017)	(0.024)	(0.003)
East * 2000†	0.272***	0.212***	0.334***	0.590***	0.508***
	(0.012)	(0.004)	(0.013)	(0.023)	(0.003)
East * 2001†	0.194***	0.152***	0.405***	0.495***	0.536***
	(0.006)	(0.006)	(0.023)	(0.040)	(0.008)
East * 2002†	0.227***	0.161***	0.418***	0.432***	0.556***
	(0.001)	(0.009)	(0.024)	(0.026)	(0.006)
East * 2003†	0.091***	0.299***	0.501***	0.567***	0.584***
	(0.004)	(0.005)	(0.026)	(0.031)	(0.006)
East * 2004†	0.184***	0.284***	0.443***	0.679***	0.646***
	(0.016)	(0.005)	(0.016)	(0.024)	(0.009)
East * 2005†	0.110***	0.316***	0.381***	0.698***	0.712***
	(0.014)	(0.00001)	(0.022)	(0.020)	(0.013)
Observations	17,368	22,149	26,882	18,859	41,088
Log likelihood	-5,510	-7,784	-9,262	-6,607	-10,476

Probit regression. Omitted categories are 1995, the interaction of East and 1995, male, single, intermediate schooling, not employed. Standard errors are clustered by East. Table continued in the appendix (table 3.14).
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

in the youngest group have spent only their childhood in a socialist regime. For them, the effect of being from East Germany should be the least pronounced. In the model, this corresponds to a smaller share of people with the wrong beliefs, i.e. a smaller α. However, the change over time should be slower for them. Since I assumed that the same fraction of the population receives the iid learning shock every period, there should be convergence in the share of the population with the correct beliefs.

Table 3.5 reports the results of the baseline regressions by age group. Notice that

for the youngest group, the sample does not include any respondents with complementary health insurance in 1995, 1996, or 1997 (when they were between 6 and 22 years old)[10]. Since the respective East-year interactions for this group were omitted, the remaining coefficients on the interaction terms do not directly compare to those of the older age groups. Focusing on the remaining 4 age groups, one can see a clear pattern consistent with the exogenous learning model: The effect of being from East Germany in 1995 is stronger the older the respondents are. This corresponds to a larger share α of people with the wrong beliefs initially. After the first big reforms (i.e. starting in 1998), the East effect decreases over time *within* each age group: The coefficients on the East-year interactions are positive and increasing from year to year. This learning effect, however, is stronger for older respondents. Comparing the interaction coefficients *across* different age groups for any given year, one can see the convergence pattern. The coefficients are smaller for younger respondents, older individuals learn at a faster pace.

3.5.2 Risk Taking and Risk Aversion

Living in an all-encompassing welfare state might influence people's risk taking behavior or even their risk-aversion. If East Germans are either less risk averse or simply take less risks, these differences in behavior or preferences would make them less likely to buy additional health insurance. To determine whether the differences in the probability of taking up private health insurance is due to differences in beliefs or preferences, I include measures of risk aversion and risk taking behavior in the analysis. In 2003, the survey included the following questions: "How would you rate your willingness to take risks on a scale from 0 (low) to 10 (high)?" and "What share of a lottery winning would you be prepared to invest in a financially risky, yet lucrative investment?"[11] The answer to the first question serves as a measure for a

[10]Moreover, none of the young respondents were widowed or retired during the time of the survey.

[11]The full question on the survey read: Please consider what you would do in the following situation: Imagine you had won 100,000 Euros in the lottery. Almost immediately after you collect the winnings, you receive the following financial offer from a reputable bank, the conditions of which are as follows: There is a chance to double the money within two years. It is equally possible that you could lose half of the amount invested. You have the opportunity to invest the amount, part of the

Table 3.6: Risk Regression

Dependent Variable: Complementary health insurance†	Coefficients (Standard errors)
East†	-0.782***
	(0.011)
Risk taking	0.03***
	(0.002)
East * Risk taking	0.03***
	(0.003)
Risk aversion	0.011***
	(0.001)
East * Risk aversion	0.037***
	(0.002)
Observations	4877
Log likelihood	-1,891

Probit regression. Omitted categories are male, single, intermediate schooling, not employed. Risk taking ranges from 0 (low) to 10 (high). Risk aversion ranges from 1 (low) to 6 (high). Standard errors are clustered by East. Table continued in the appendix (table 3.17).
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

respondent's risk taking behavior, while the second question provides an estimate of the respondent's risk aversion.

First, I analyze whether a respondent's willingness to take risks or his risk aversion influence his likelihood to buy additional health insurance, and how this differs between former East and West Germans. Table 3.6 reports the results of the baseline regression augmented by the risk variables as well as their interactions with the East dummy. Since this regression only includes observations from one year of the survey, the coefficients do not directly compare to the basic regression results. Notice however that the coefficient on the East dummy is still significantly negative. The main conclusion is not changed: East Germans are significantly less likely to have complementary health insurance, a difference that can be explained by them having the

amount or reject the offer. What share of your lottery winnings would you be prepared to invest in this financially risky, yet lucrative investment? The answer categories were: 100,000 Euros (1), 80,000 Euros (2), 60,000 Euros (3), 40,000 Euros (4), 20,000 Euros (5), or "Nothing, I would decline the offer" (6).

wrong beliefs about the welfare state with a higher probability than West Germans.

The coefficients on risk taking and risk aversion are as expected: The more willing a respondent is to take risks, the more likely he is to buy additional insurance. Most likely, insurance contracts can not control for these attitudes towards risk taking, so that this effect could be evidence for adverse selection or moral hazard. Naturally, the more risk averse a respondent reports he is, the more likely he is to have insurance. Interestingly, these effects are very similar for East and West Germans. For risk taking, the coefficients are almost exactly the same, while risk aversion is a little bit more influential among East Germans than it is among West Germans.

Second, I ask whether risk taking behavior or risk aversion is determined in part by which regime the respondent lived under before reunification. If that was the case, then the results in table 3.6 would have to be questioned. Consider table 3.7. It reports the results of ordinary least square regressions of the risk variables on the East dummy and the baseline controls (the full results are reported in the appendix). Former East Germans seem a little more willing to take risks than West Germans. If anything, this should make them *more* likely to buy additional insurance, but the effect is only significant at the 10% level. For risk aversion, the East dummy is not significant at all. These results make it safe to reject the hypothesis that the differences in the probability to take up complementary health insurance between East and West Germans is due to differences in risk taking behavior or risk aversion.

3.5.3 Preferences For a Larger Welfare State

Alesina and Fuchs-Schündeln (2007) use the same data set employed in this paper to document that living under a socialist regime influences preferences about public social policies. They show that former East Germans are more likely to prefer the state to be responsible for providing social services, insurance, and redistribution. Observing former East Germans to be less likely to seek additional insurance beyond the coverage of the public health insurance could be a consequence of or at least cor-

Table 3.7: Risk Regression

Dependent Variable:	Risk taking	Risk aversion
East†	0.121*	0.245
	(0.009)	(0.023)
Observations	4960	4962
R squared	0.0397	0.0914

Omitted categories are male, single, intermediate schooling, not employed. Standard errors are clustered by East. Risk taking ranges from 0 (low) to 10 (high). Risk aversion ranges from 1 (low) to 6 (high).Table continued in the appendix (table 3.17).
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

related with them having a stronger preference for state intervention. If, for example, an agent thinks that the contracts offered in the private market are unfair, he might have a stronger preference for the state to intervene *and* be less likely to buy private insurance.

To investigate this potential relationship, I include the same measure of preferences for a bigger welfare state that Alesina and Fuchs-Schündeln (2007) used in their analysis. In the surveys of 1996 and 2001, it was asked: "Who should be responsible for financial security in case of illness?" I generate a dummy variable indicating that the answer was "only the state" or "mostly the state", as opposed to "both the state and private forces", "mostly private forces", or "only private forces", and include it in the baseline regression. Table 3.8 reports the results. Again, since this regression only includes two of the years included in the baseline sample, the size of the coefficients is not necessarily comparable. Again, however, the effect of having lived under the socialist regime on the likelihood to complement insurance coverage with a private contract remains significantly negative. Even controlling for preferences about the welfare state, former East Germans are less likely to buy private health insurance.

Interestingly, the effect of the preferences for state responsibilities are quite different among former East and West Germans. While a stronger preference for state inter-

Table 3.8: Preference Regression

Dependent Variable: Complementary health insurance†	Coefficients (Standard errors)
East†	-0.727***
	(0.001)
Preference†	-0.208***
	(0.020)
East * Preference†	0.024
	(0.039)
Observations	21,836
Log likelihood	-6,180

Probit regression. Preference is a variable that contains the answer to the question of who should be responsible for the financial security in case of illness. It takes on the value 1 if the answer was "only the state" or "mostly the state", and 0 for "both state and private forces", "mostly private forces", or "only private forces". Omitted categories are 1996, the interaction of East and 1996, male, single, intermediate schooling, not employed. Standard errors are clustered by East. Table continued in the appendix (table 3.16).
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

vention makes respondents significantly less likely to purchase additional insurance among West Germans, there is no significant effect among the former East German population.

3.5.4 Aggregate Effects: Regional Differences

An agent's decision to obtain additional insurance coverage might be influenced by factors not captured as an individual characteristic, but rather inherent in his environment. Living in a big city, for example, might make it much easier to access the private insurance market. Insurance agencies are probably rare to find in more rural areas. Health services might be cheaper in some areas, which could make agents living there less likely to have a complementary health insurance. In short, the aggregate economic and demographic situation of the region an agent lives in might greatly

influence his likelihood to obtain additional insurance coverage in the private market. If former East Germans mostly live in areas that have a negative impact on the insurance decision, the observed differences might not be attributable to differences in beliefs.

The GSOEP data set does allow me to identify in which state the respondent lived in each year of the survey[12]. I include dummies for all states as well as East-state interactions. The omitted category is the state Hamburg. Since 1991, Hamburg has consistently had the largest GDP per capita (Statistisches Bundesamt (2010)). It is also a city state with a dense population and urban infrastructure. The coefficients measure the difference between Hamburg and the respective state. Table 3.9 reports the results.

Consider first the coefficient on the East dummy. In Hamburg, former East Germans are significantly less likely to have additional health insurance than their fellow West German citizens. All coefficients on the East-state interactions are negative, while all coefficients on the state dummies are positive (with the exception of Berlin, where the East interaction is not significant, and the state dummy is negative). Two things follow from this. First, in every state, former East Germans are less likely to have complementary health insurance, confirming the results of the basic regression. Second, not surprisingly, where a respondent lives does have a significant effect on him purchasing insurance, since almost all coefficients are significant. Interestingly, though, this effect is opposite for former East and West Germans. While West Germans in almost all states are more likely to buy additional insurance than those living in Hamburg, for East Germans the opposite is true. This means that in every state, the effect of being East German is stronger than in among people living in Hamburg. One possible explanation could be that East Germans who migrated to Hamburg (an international harbor, and home to the headquarters of many large companies) are

[12]The dataset does not distinguish the states Rheinland-Pfalz and Saarland. Thus, there are only 15 states to control for, even though Germany has 16 Länder.

Table 3.9: Regression with State Controls

Dependent Variable: Complementary health insurance†	Coefficients (Standard errors)		Coefficients (Standard errors)
East†	-0.501***		
	(0.0478)		
East * Schleswig-Holstein†	-0.758***	Schleswig-Holstein†	0.001
	(0.033)		(0.009)
East * Niedersachsen†	-0.176***	Niedersachsen†	0.150***
	(0.021)		(0.023)
East * Bremen†	-0.163**	Bremen†	0.083***
	(0.061)		(0.001)
East * Nordrhein-Westfalen†	-0.653***	Nordrhein-Westfalen†	0.353***
	(0.037)		(0.018)
East * Hessen†	-0.485***	Hessen†	0.061***
	(0.033)		(0.01)
East * Rheinland-Pfalz, Saarland†	-0.277***	Rheinland-Pfalz, Saarland†	0.08**
	(0.027)		(0.029)
East * Baden-Würtemberg†	-0.338***	Baden-Würtemberg†	0.271***
	(0.0145)		(0.023)
East * Bayern†	-0.378***	Bayern†	0.274***
	(0.023)		(0.026)
East * Berlin†	0.008	Berlin†	-0.056***
	(0.069)		(0.005)
East * Mecklenburg-Vorpommern†	-1.082***	Mecklenburg-Vorpommern†	0.694***
	(0.079)		(0.039)
East * Brandenburg†	-0.371***	Brandenburg†	0.124***
	(0.041)		(0.019)
East * Sachsen-Anhalt†	-0.301**	Sachsen-Anhalt†	-0.019
	(0.104)		(0.053)
East * Thüringen†	-0.979***	Thüringen†	0.636***
	(0.086)		(0.053)
East * Sachsen†	-0.479***	Sachsen†	0.286***
	(0.087)		(0.047)
Observations	126,331		
Log likelihood	-40,126		

Probit regression. Omitted categories are Hamburg, the interaction of East and Hamburg, 1995, the interaction of East and 1995, male, single, intermediate schooling, not employed. Standard errors are clustered by East. Table continued in the appendix (table 3.18).
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

especially likely to have the correct beliefs about the welfare state. Either a selection effect (more pro-capitalism people moved to Hamburg) or a treatment effect (the environment made them learn faster) could be responsible for such a difference in the

average beliefs of East Germans in Hamburg and elsewhere.

3.6 Discussion

This paper analyzes the question whether people adapt to new institutions immediately, or learn only over time how to adjust their economic decisions. Germany presents a unique opportunity to study this question. Since at reunification the well-established economic and political system of West Germany was imposed onto East Germany, any observed adaptation process is rather due to people learning than to institutions developing. Moreover, former West Germans can serve as a meaningful control group. To identify a learning process among former East Germans, I analyze the economic decision of buying private health insurance.

In the spirit of a "reverse" difference-in-difference approach, where the treatment (socialism) occurred in the pre-period, and the shock (health care reform) is the same for treatment and control group, I study the differential reaction of former East and West Germans to a series of health care reforms that started in 1997. Along with the gradual decrease in coverage under the public health insurance system, former East Germans were significantly less likely to sign complementary health insurance contracts in the private market. I show that this difference can be interpreted as East Germans having not yet fully adapted to the new capitalist institutions. In particular, I show that the uptake rates of additional private insurance after the reforms are consistent with a model in which agents learn about institutions through an exogenous shock and optimize their individual insurance status only if they are aware of the organizational form of the health care system (or more generally the welfare state). East Germans are initially less likely to have the correct beliefs, but learn over time how institutions have changed and that they are now responsible for optimizing their insurance coverage.

An age decomposition of the regression analysis substantiated the convergence hy-

pothesis of the learning model. The effect of being from the East is more pronounced but vanishes faster among older Germans. This is consistent with older East Germans, who lived under the socialist regime longer, being less likely to have the correct beliefs about the welfare state initially, but receiving an exogenous learning shock with the same probability as their younger equivalents. The purpose of this exercise is not to claim that the simple exogenous learning model is the true underlying process of East Germans adapting to capitalism; most likely, the true learning process is more complex. Rather, this study provides evidence for the existence of a substantial transition period in people's adaptation to new institutions. Taking into account that people need time to adjust is critical for predicting the success and speed of an economy's transition from socialism to capitalism.

References

Akerlof, G.: 1970, The market for lemons, *Quarterly Journal of Economics* **84**, 488–500.

Alesina, A. and Fuchs-Schündeln, N.: 2007, Good bye lenin (or not?): The effect of communism on people's preferences, *American Economic Review* **97**.

Arbeitslosigkeit im Zeitverlauf: 2011, *Data*, Statistik der Bundesagentur für Arbeit. http://statistik.arbeitsagentur.de.

Bolton, P. and Dewatripont, M.: 2005, *Contract Theory*, The MIT Press.

Borghans, L., Golsteyn, B., Heckman, J. and Meijers, H.: 2009, Gender differences in risk aversion and ambiguity aversion, *Journal of the European Economic Association* **7**, 649–658.

Bucher-Koenen, T. and Lusardi, A.: 2011, Financial literacy and retirement planning in germany, *Working Paper 17110*, NBER.

Busse, R. and Riesberg, A.: 2004, *Health Care Systems in Transition - Germany*, European Observatory on Health Systems and Policies.

Cochrane, N.: 2007, Promoting sustainable market institutions in the transition economies: The role of international assistance, *104th seminar*, European Association of Agricultural Economists.

Die Private Krankenversicherung 2002/2003: 2004, Verband der privaten Krankenversicherer e.V.

Djankov, S. and Murrell, P.: 2002, Enterprise restructuring in transition: A quantitative survey, *Journal of Economic Literature* **15**, 739–792.

Filatotchev, I., Buck, T. and Zhukov, V.: 2000, Downsizing in privatized firms in russia, ukraine, and belarus, *Academy of Management Journal* **43**, 286–304.

Fuchs-Schündeln, N.: 2008, The response of household saving to the large shock of german reunification, *American Economic Review* **98**, 1798–1828.

Fuchs-Schündeln, N., Krüger, D. and Sommer, M.: 2010, Inequality trends for germany in the last two decades: A tale of two countries, *Review of Economic Dynamics* **13**, 103–132.

GSOEP: 2007, German socio-economic panel study, 1984-2006, *Data*, DIW Berlin. www.diw.de/soep.

Haisken-DeNew, J. P. and Frick, J. R.: 2005, Desktop companion to the german socio-economic panel (soep), *Manual*, DIW Berlin. Version 8.0.

Keefer, P. and Knack, S.: 2005, Social capital, social norms and the new institutional economics, *in* C. Menard and M. Shirley (eds), *Handbook of New Institutional Economics*, Springer Verlag, pp. 701–725.

Kriauciunas, A. and Kale, P.: 2006, The impact of socialist imprinting and search on resource change: A study of firms in lithuania, *Strategic Management Journal* **27**, 659–679.

Küsters, H. and Hofmann, D.: 1998, *Deutsche Einheit - Sonderedition aus den Akten des Bundeskanzleramtes 1989/90*, Oldenbourg Verlag, München.

Lusardi, A. and Mitchell, O.: 2009, How ordinary consumers make complex economic decisions: Financial literacy and retirement readiness, *Working Paper 15350*, NBER.

Münz, R. and Ulrich, R.: 1997, Changing patterns of german immigration, 1945-1994, *in* K. Bade and M. Weiner (eds), *Migration Past, Migration Future: Germany and the United States*, Berghahn Books, Providence/RI - Oxford, pp. 65–117.

Murrell, P.: 2005, Institutions and firms in transition economies, *in* C. Menard and M. Shirley (eds), *Handbook of New Institutional Economics*, Springer Verlag, pp. 667–699.

Murrell, P.: 2008, Transition and institutions, *in* L. Blume and S. Durlauf (eds), *The New Palgrave Dictionary of Economics*, Palgrave Macmillan.

Pötzsch, O.: 2007, Geburten in deutschland, *report*, Statistisches Bundesamt.

Sauter, N.: 2009, Tearing down the wall: (non-)participation and habit persistence in east german securities markets, *mimeo*, Munich Graduate School of Economics.

State Treaty between the FRG and the GDR Establishing a Monetary, Economic, and Social Union: 1990.

Steffen, J.: 2011, *Sozialpolitische Chronik. Krankenversicherung seit 1977*, Arbeitnehmerkammer Bremen.

Uhlenbruck, K. and de Castro, J.: 2000, Foreign acquisitions in central and eastern europe: Outcomes of privatization in transitional economies, *Academy of Management Journal* **43**, 381–402.

VGRdL: 2010, *Bruttoinlandsprodukt, Bruttowertschöpfung in den Ländern und Ost-West-Grossraumregionen Deutschlands 1991-2010*, Arbeitskreis "Volkswirtschaftliche Gesamtrechnungen der Länder" im Auftrag der Statistischen Ämter der 16 Bundesländer, des Statistischen Bundesamtes und des Bürgeramtes, Statistik und Wahlen.

WHO: 2000, *Health Care Systems in Transition*, World Health Organization.

3.7 Appendix

Table 3.10: Summary Statistics

	Total	West	East
Variable	Mean (Standard deviation)	Mean (Standard deviation)	Mean (Standard deviation)
East†	0.344		
	(0.475)		
Age	46.352	46.793	45.512
	(17.523)	(17.578)	(17.389)
Female†	0.538	0.541	0.533
	(0.499)	(0.498)	(0.499)
Married†	0.626	0.634	0.611
	(0.484)	(0.482)	(0.488)
Married but separated†	0.013	0.012	0.013
	(0.111)	(0.111)	(0.112)
Divorced†	0.054	0.050	0.061
	(0.226)	(0.219)	(0.238)
Widowed†	0.071	0.074	0.065
	(0.257)	(0.262)	(0.247)
Health status	2.619	2.607	2.640
	(0.935)	(0.949)	(0.909)
Satisfaction with health	6.608	6.717	6.401
	(2.192)	(2.209)	(2.146)
High school degree†	0.158	0.151	0.171
	(0.365)	(0.359)	(0.377)
Vocational training †	0.702	0.687	0.729
	(0.458)	(0.464)	(0.444)
Completed college degree†	0.158	0.118	0.234
	(0.364)	(0.322)	(0.423)
Full or part time employed†	0.499	0.500	0.497
	(0.500)	(0.500)	(0.500)
Unemployed†	0.089	0.054	0.156
	(0.285)	(0.225)	(0.363)
Retired†	0.261	0.257	0.269
	(0.439)	(0.437)	(0.444)
Household income	38506.300	42824.850	30285.340
	(38610.730)	(42331.280)	(28538.540)
Household net income	33754.600	36341.160	28830.710
	(23946.050)	(27128.080)	(15068.640)
Household asset income	2137.444	2806.130	864.508
	(12949.270)	(15771.740)	(3301.101)
Homeowner†	0.516	0.568	0.417
	(0.500)	(0.495)	(0.493)
Observations	126,346	82,833	43,513

Summary statistics for the explanatory variables included in all regression. Income variables in Euro. † Dummy variable.

Table 3.11: Basic Regression - Linear Probability Model

Dependent Variable: Complementary Health Insurance†	Coefficients (Standard Errors)		Coefficients (Standard Errors)
East†	-0.083**		
	(0.001)		
Age	-0.0038	High school degree†	0.074
	(0.004)		(0.028)
Age squared $*10^{-3}$	0.109	Vocational training †	0.031
	(0.114)		(0.017)
Age cubed $*10^{-5}$	-0.076	Completed college degree†	0.045
	(0.077)		(0.044)
Female†	0.02	Full or part time employed†	0.001
	(0.011)		(0.003)
Married†	-0.016	Unemployed†	0.056
	(0.009)		(0.008)
Married but separated†	0.049	Retired†	0.013
	(0.041)		(0.004)
Divorced†	0.014*	Log (household income)	0.006
	(0.001)		(0.003)
Widowed†	-0.02***	Log (household net income)	0.013
	(0.00001)		(0.003)
Health status	-0.003	Log (household asset income)	-0.007
	(0.001)		(0.007)
Satisfaction with health	0.001	Homeowner†	-0.011
	(0.0002)		(0.003)
Constant	-0.539		
	(0.134)		
Observations	126,346		

OLS regression. Omitted categories are male, single, intermediate schooling, not employed. The variable "Health status" ranges from 1 (very good) to 5 (bad). The variable "Health satisfaction" ranges from 0 (low) to 10 (high). Standard errors are clustered by East.
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

Table 3.12: Basic Regression with East-Year Interactions - continued from table 3.2

Dependent Variable: Complementary Health Insurance†	Coefficients (Standard Errors)		Coefficients (Standard Errors)
Age	-0.028***	High school degree†	0.322***
	(0.007)		(0.057)
Age squared $*10^{-3}$	0.650*	Vocational training†	0.182**
	(0.312)		(0.058)
Age cubed $*10^{-5}$	-0.430	Completed college degree†	0.253*
	(0.253)		(0.105)
Female†	0.108**	Log (household income)	0.050***
	(0.036)		(0.012)
Married†	-0.030	Log (household net income)	0.177***
	(0.022)		(0.035)
Married but separated†	0.246	Log (household asset income)	0.059***
	(0.134)		(0.005)
Divorced†	0.077	Homeowner†	0.011
	(0.055)		(0.019)
Widowed†	-0.103***	Full or part time employed†	0.084*
	(0.002)		(0.040)
Health status	-0.016***	Unemployed†	-0.084
	(0.002)		(0.054)
Satisfaction with health	0.002	Retired†	0.006
	(0.004)		(0.010)
Constant	-4.059***		
	(0.400)		
Observations	126,346		
Log likelihood	-40,414		

Probit regression. Omitted categories are 1995, the interaction of East and 1995, male, single, intermediate schooling, not employed. Standard errors are clustered by East.
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

Table 3.13: Age Regression - continued from table 3.4

Dependent Variable: Complementary Health Insurance†	Coefficients (Standard Errors)		Coefficients (Standard Errors)
East * 1996†	0.560***	1996†	-0.328***
	(0.005)		(0.004)
East * 1997†	0.089***	1997†	-0.423***
	(0.008)		(0.006)
East * 1998†	0.083***	1998†	0.363***
	(0.008)		(0.005)
East * 1999†	0.162***	1999†	0.272***
	(0.005)		(0.005)
East * 2000†	0.288***	2000†	0.353***
	(0.004)		(0.005)
East * 2001†	0.281***	2001†	0.385***
	(0.002)		(0.009)
East * 2002†	0.294***	2002†	0.416***
	(0.0005)		(0.008)
East * 2003†	0.230***	2003†	0.453***
	(0.002)		(0.006)
East * 2004†	0.253***	2004†	0.515***
	(0.003)		(0.004)
East * 2005†	0.140***	2005†	0.692***
	(0.003)		(0.0001)
Female†	0.103**	Vocational training †	0.191***
	(0.039)		(0.041)
Married†	-0.014***	Completed college degree†	0.301***
	(0.0001)		(0.063)
Married but separated†	0.256*	Log (household income)	0.037***
	(0.122)		(0.004)
Divorced†	0.095	Log (household net income)	0.200***
	(0.075)		(0.012)
Widowed†	-0.125***	Log (household asset income)	0.060***
	(0.024)		(0.006)
Health status	-0.017***	Homeowner†	-0.002
	(0.001)		(0.037)
Satisfaction with health	0.001	Full or part time employed†	0.077
	(0.001)		(0.049)
High school degree†	0.305***	Unemployed†	-0.098*
	(0.053)		(0.044)
Constant	-4.737***	Retired†	-0.001
	(0.240)		(0.003)
Observations	126,346		
Log likelihood	-40,139		

Probit regression. Omitted categories are 1995, the interaction of East and 1995, male, single, intermediate schooling, not employed. Standard errors are clustered by East.
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

Table 3.14: Basic Regression For Different Age Groups - continued from table 3.5

Dependent Variable: Complementary Health Insurance†	Born 1975-1989	Born 1965-1974	Born 1955-1964	Born 1945-1954	Born before 1945
1996†		0.032	-0.029	-0.115***	-0.069***
		(0.017)	(0.019)	(0.010)	(0.009)
1997†		-0.081**	-0.140***	-0.283***	-0.172***
		(0.030)	(0.034)	(0.013)	(0.013)
1998†	0.517***	0.534***	0.323***	0.096**	0.218***
	(0.015)	(0.041)	(0.034)	(0.031)	(0.002)
1999†	0.445***	0.579***	0.330***	0.132***	0.263***
	(0.006)	(0.056)	(0.051)	(0.029)	(0.006)
2000†	0.400***	0.650***	0.379***	0.097***	0.193***
	(0.001)	(0.073)	(0.056)	(0.028)	(0.011)
2001†	0.565***	0.751***	0.383***	0.254***	0.304***
	(0.008)	(0.089)	(0.065)	(0.014)	(0.019)
2002†	0.558***	0.790***	0.447***	0.211***	0.316***
	(0.015)	(0.105)	(0.067)	(0.005)	(0.025)
2003†	0.671***	0.808***	0.449***	0.295***	0.377***
	(0.018)	(0.118)	(0.069)	(0.0004)	(0.026)
2004†	0.671***	0.921***	0.512***	0.292***	0.426***
	(0.022)	(0.134)	(0.068)	(0.009)	(0.026)
2005†	0.797***	1.039***	0.635***	0.424***	0.511***
	(0.028)	(0.144)	(0.073)	(0.019)	(0.028)
Age	0.248	-0.355*	1.413**	-0.809	0.145**
	(0.306)	(0.142)	(0.449)	(0.627)	(0.052)
Age squared $*10^{-3}$	-12.58	11.83*	-36.05***	15.95	-1.757***
	(13.84)	(4.939)	(10.53)	(11.53)	(0.526)
Age cubed $*10^{-5}$	20.93	-13.37**	30.22***	-10.38	0.657***
	(19.70)	(5.135)	(8.228)	(6.957)	(0.166)
Female†	0.085**	0.086***	0.037	0.208***	0.167*
	(0.027)	(0.021)	(0.093)	(0.020)	(0.078)
Married†	-0.020	0.110***	-0.079	-0.007	-0.337***
	(0.029)	(0.025)	(0.116)	(0.004)	(0.008)
Married but separated†	0.197	0.070	0.319	0.436***	-0.090***
	(0.694)	(0.056)	(0.294)	(0.066)	(0.020)
Divorced†	-0.152	0.091	-0.004	0.297**	-0.161***
	(0.190)	(0.129)	(0.027)	(0.098)	(0.004)
Widowed†		0.606*	0.049	-0.108	-0.315***
		(0.306)	(0.205)	(0.060)	(0.015)
Health status	0.041***	0.030***	-0.010	0.001	-0.096***
	(0.0003)	(0.003)	(0.036)	(0.004)	(0.014)
Satisfaction with health	-0.004	0.008**	-0.002	-0.012	-0.002
	(0.004)	(0.003)	(0.011)	(0.014)	(0.008)
Observations	17,368	22,149	26,882	18,859	41,088
Log likelihood	-5,510	-7,784	-9,262	-6,607	-10,476

Probit regression. Omitted categories are 1995, the interaction of East and 1995, male, single, intermediate schooling, not employed. Standard errors are clustered by East. Table continued on the next page.

*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

Table 3.15: Regression by cohort - continued from table 3.14

Dependent Variable: Complementary Health Insurance†	Born 1975-1989	Born 1965-1974	Born 1955-1964	Born 1945-1954	Born before 1945
High school degree†	0.306***	0.217***	0.252***	0.449	0.425*
	(0.010)	(0.011)	(0.045)	(0.290)	(0.167)
Vocational training †	0.124***	0.017	0.152**	0.124	0.286***
	(0.022)	(0.021)	(0.053)	(0.089)	(0.076)
Completed college degree†	0.191	0.117	0.153***	0.298***	0.476***
	(0.127)	(0.066)	(0.040)	(0.063)	(0.095)
Log (household income)	0.089***	0.083	0.022	-0.038***	-0.003
	(0.004)	(0.046)	(0.027)	(0.009)	(0.010)
Log (household net income)	0.005*	-0.00002	0.316***	0.386***	0.395***
	(0.002)	(0.051)	(0.031)	(0.019)	(0.019)
Log (household asset income)	0.047***	0.054***	0.045***	0.046***	0.090***
	(0.006)	(0.010)	(0.008)	(0.013)	(0.009)
Homeowner†	-0.030	-0.060	-0.025	0.039	0.121***
	(0.063)	(0.048)	(0.030)	(0.047)	(0.011)
Full or part time employed†	0.097	0.121***	0.087*	0.151**	-0.007
	(0.105)	(0.014)	(0.036)	(0.056)	(0.014)
Unemployed†	-0.134	-0.117	-0.130***	-0.158***	-0.018
	(0.129)	(0.107)	(0.025)	(0.036)	(0.028)
Retired†		-0.276*	-0.164***	0.019	-0.066***
		(0.112)	(0.032)	(0.032)	(0.015)
Constant	-4.856*	0.355	-23.56***	8.062	-9.514***
	(2.153)	(1.272)	(6.801)	(11.02)	(1.578)
Observations	17,368	22,149	26,882	18,859	41,088
Log likelihood	-5,510	-7,784	-9,262	-6,607	-10,476

Probit regression. Omitted categories are 1995, the interaction of East and 1995, male, single, intermediate schooling, not employed. Standard errors are clustered by East.
*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

Table 3.16: Preference Regression - continued from table 3.8

Dependent Variable: Complementary Health Insurance†	Coefficients (Standard errors)		Coefficients (Standard errors)
2001†	0.452***	High school degree†	0.364***
	(0.013)		(0.042)
East * 2001†	0.198***	Vocational training†	0.180
	(0.009)		(0.093)
Age	-0.036***	Completed college degree†	0.275
	(0.007)		(0.149)
Age squared $*10^{-3}$	0.829***	Log (household income)	0.041*
	(0.075)		(0.017)
Age cubed $*10^{-5}$	-0.507**	Log (household net income)	0.230***
	(0.158)		(0.023)
Female†	0.089***	Log (household asset income)	0.052***
	(0.024)		(0.011)
Married†	0.002	Homeowner†	0.039***
	(0.029)		(0.003)
Married but separated†	0.340	Full or part time employed†	0.117
	(0.223)		(0.109)
Divorced†	0.097	Unemployed†	-0.102
	(0.064)		(0.088)
Widowed†	-0.069	Retired†	-0.037
	(0.080)		(0.023)
Health status	-0.049***	Constant	-4.391***
	(0.006)		(0.402)
Satisfaction with health	0.002	Observations	21836
	(0.013)	Log likelihood	-6,180

Probit regression. Preference is a variable that contains the answer to the question of who should be responsible for the financial security in case of illness. It takes on the value 1 if the answer was "only the state" or "mostly the state", and 0 for "both state and private forces", "mostly private forces", or "only private forces". Omitted categories are 1996, the interaction of East and 1996, male, single, intermediate schooling, not employed. Standard errors are clustered by East.

*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

Table 3.17: Risk Regression - continued from tables 3.6 and 3.7

Dependent Variable:	Complementary health insurance†	Risk taking	Risk aversion
Age	0.035	0.047	-0.187
	(0.035)	(0.012)	(0.054)
Age squared $*10^{-3}$	-0.880	-0.898	3.568
	(0.964)	(0.312)	(0.899)
Age cubed $*10^{-5}$	0.699	0.612	-2.335
	(0.772)	(0.204)	(0.525)
Female†	0.068***	0.213	-0.761*
	(0.007)	(0.087)	(0.057)
Married†	0.048	0.161	-0.410
	(0.052)	(0.05)	(0.087)
Married but separated†	0.548*	0.085	-0.16*
	(0.255)	(0.06)	(0.011)
Divorced†	0.29*	0.115	0.25*
	(0.12)	(0.048)	(0.01)
Widowed†	-0.114*	0.123	-0.212
	(0.052)	(0.025)	(0.284)
Number of kids under 16 in the household	-0.071*	-0.023	-0.082
	(0.029)	(0.024)	(0.059)
Number of adults in the household	-0.038	0.068	-0.041
	(0.028)	(0.027)	(0.035)
Health status	0.069***		
	(0.019)		
Satisfaction with health	0.017*		
	(0.007)		
High school degree†	0.188***	-0.152	0.112
	(0.041)	(0.035)	(0.036)
Vocational training†	0.16***	-0.009	0.091
	(0.026)	(0.02)	(0.251)
Completed college degree†	0.270***	-0.076	0.300
	(0.014)	(0.099)	(0.188)
Log (household income)	0.052	-0.010	-0.016
	(0.068)	(0.018)	(0.035)
Log (household net income)	0.127	-0.024	0.187
	(0.115)	(0.070)	(0.090)
Log (household asset income)	0.084***	-0.025	-0.023
	(0.004)	(0.008)	(0.014)
Homeowner†	-0.111	-0.016	0.060
	(0.084)	(0.009)	(0.075)
Full or part time employed†	0.049	-0.041	0.202
	(0.043)	(0.038)	(0.026)
Unemployed†	-0.251***	0.023	0.160
	(0.002)	(0.009)	(0.197)
Retired†	-0.279***	0.021	-0.364
	(0.048)	(0.038)	(0.188)
Constant	-4.282***	4.658*	6.745*
	(0.026)	(0.358)	(0.381)
Observations	4877	4960	4962
Log likelihood / R squared	-1.891	0.0397	0.0914

Probit regression in the first column. Omitted categories are male, single, intermediate schooling, not employed. Risk taking ranges from 0 (low) to 10 (high). Risk aversion ranges from 1 (low) to 6 (high).Standard errors are clustered by East.

*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

Table 3.18: Regression with State Controls- continued from table 3.9

Dependent Variable: Complementary health insurance†	Coefficients (Standard errors)		Coefficients (Standard errors)
East * 1996†	0.232***	1996†	-0.097***
	(0.006)		(0.007)
East * 1997†	0.268***	1997†	-0.201***
	(0.003)		(0.013)
East * 1998†	0.471***	1998†	0.246***
	(0.002)		(0.002)
East * 1999†	0.403***	1999†	0.269***
	(0.009)		(0.01)
East * 2000†	0.452***	2000†	0.259***
	(0.009)		(0.012)
East * 2001†	0.439***	2001†	0.354***
	(0.015)		(0.022)
East * 2002†	0.454***	2002†	0.367***
	(0.013)		(0.024)
East * 2003†	0.506***	2003†	0.405***
	(0.013)		(0.026)
East * 2004†	0.542***	2004†	0.455***
	(0.007)		(0.026)
East * 2005†	0.539***	2005†	0.566***
	(0.007)		(0.032)
Age	-0.029***	High school degree†	0.327***
	(0.007)		(0.063)
Age squared $*10^{-3}$	0.696*	Vocational training†	0.184**
	(0.321)		(0.058)
Age cubed $*10^{-5}$	-0.461	Completed college degree†	0.25*
	(0.26)		(0.108)
Female†	0.109**	Log (household income)	0.05***
	(0.037)		(0.011)
Married†	-0.029	Log (household net income)	0.173***
	(0.02)		(0.034)
Married but separated†	0.247	Log (household asset income)	0.059***
	(0.142)		(0.006)
Divorced†	0.086	Homeowner†	0.02***
	(0.054)		(0.003)
Widowed†	-0.099***	Full or part time employed†	0.089*
	(0.005)		(0.041)
Health status	-0.019***	Unemployed†	-0.077
	(0.002)		(0.053)
Satisfaction with health	0.001	Retired†	0.01
	(0.003)		(0.011)
Observations	126,331	Constant	-4.217***
Log likelihood	-40,126		(0.423)

Probit regression. Omitted categories are Hamburg, the interaction of East and Hamburg, 1995, the interaction of East and 1995, male, single, intermediate schooling, not employed. Standard errors are clustered by East.

*** Significant at 1%, ** significant at 5%, * significant at 10%. † Dummy variable.

Funny how time passes.