

Essays in Macroeconomics with Heterogeneity

A THESIS

**SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA**

BY

SEBASTIAN DYRDA

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY**

José Víctor Ríos-Rull

September, 2015

© SEBASTIAN DYRDA 2015
ALL RIGHTS RESERVED

Acknowledgements

I own deep debt of gratitude to my thesis advisor José Víctor Ríos-Rull and V.V. Chari for their patience, guidance, support and teaching me how to do research in economics. They have been leading examples of scholars for me.

I am also indebted to Christopher Phelan and Jonathan Heathcote for their help and encouragement.

I would like to thank my coauthor and friend Marcelo Pedroni for countless discussions and hard work on our paper.

Generous financial support during my time at the University of Minnesota has been provided by the Leonid Hurwicz Fellowship, Graduate Research Partners Program Fellowship, Hella Mears Graduate Fellowship and Kauffman Dissertation Fellowship from Ewing Marion Kauffman Foundation.

I am deeply grateful to my parents, who supported me enormously during the entire graduate program. The person who I own the most and who has contributed hugely to my doctorate is my wife Agnieszka. I would like to thank her for patience and love.

For my beloved wife Agnieszka.

Abstract

This thesis consist of two essays on macroeconomics with heterogeneity. First essay quantifies the importance of aggregate fluctuations in microeconomic uncertainty for the firm dynamics over the business cycle in an economy with frictional financial markets. To begin, it documents facts on asymmetric response across age and size groups of firms in the U.S. to the changes in aggregate economic conditions. I argue that age rather than size is a relevant margin for the cyclical employment dynamics; in particular total employment of young firms varies 2.6 times more relative to the old firms. Then I propose a theory that, contrary to the existing studies, generates endogenously a link between firm's age and size and its ability to obtain financing, and induces an asymmetric response to shocks. A key element of my theory is a financial friction, originated from the firm's private information and long-term, efficient lending contract between a risk averse entrepreneur and financial intermediary, which manifests itself as a borrowing constraint. I argue that, for any given expected return on project, young firms are more constrained in borrowing and they grow out of the constraint as they age up to the optimal, unconstrained size. Next I establish that, for any given age, firm's financing increases in line with the average return on a project. In times of high idiosyncratic uncertainty the financial contract calls for tightening of the borrowing constraint transmitting the initial impulse into a decline in demand for production inputs and further, including general equilibrium effects, into an economic downturn. This mechanism affects disproportionately young firms. Not only are they more constrained in borrowing but also they start smaller due to a reduced level of initial financing. A quantitative version of the model accounts for the fall of the aggregate output, employment and investment, decline of credit to GDP ratio and asymmetric employment dynamics of different groups of firms observed in the US data in recessions.

Second essay studies optimal taxation in an environment where heterogeneous households face uninsurable idiosyncratic risk. To do this, we formulate a Ramsey problem in a standard infinite horizon incomplete markets model. We solve numerically for the optimal

path of proportional capital and labor income taxes, (possibly negative) lump-sum transfers, and government debt. The solution maximizes welfare along the transition between an initial steady state, calibrated to replicate key features of the US economy, and an endogenously determined final steady state. We find that in the optimal (utilitarian) policy: (i) capital income taxes are front-loaded hitting the imposed upper bound of 100 percent for 33 years before decreasing to 45 percent in the long-run; (ii) labor income taxes are reduced to less than half of their initial level, from 28 percent to about 13 percent in the long-run; and (iii) the government accumulates assets over time reducing the debt-to-output ratio from 63 percent to -17 percent in the long-run. Relative to keeping fiscal instruments at their initial levels, this leads to an average welfare gain equivalent to a permanent 4.9 percent increase in consumption. Even though non-distortive lump-sum taxes are available, the optimal plan has positive capital and labor taxes. Such taxes reduce the proportions of uncertain and unequal labor and capital incomes in total income, increasing welfare by providing insurance and redistribution. We are able to quantify these welfare effects. We also show that calculating the entire transition path (as opposed to considering steady states only) is quantitatively important. Implementing the policy that maximizes welfare in steady state leads to a welfare *loss* of 6.4 percent once transitory effects are accounted for.

Contents

Acknowledgements	i
Abstract	iii
List of Tables	ix
List of Figures	x
1 Introduction	1
2 Optimal Fiscal Policy in the Heterogenous Agents Economy	3
2.1 Introduction	3
Related Literature	7
2.2 Mechanism: Two-Period Economy	10
2.2.1 Uncertainty economy	11
2.2.2 Inequality economy	14
2.2.3 Uncertainty and inequality	16
2.2.4 Relationship with infinite horizon problem	16
2.3 The Infinite-Horizon Model	18
2.3.1 The Ramsey Problem	20
2.3.2 Solution method	20
2.4 Calibration	22
2.4.1 Preferences and technology	22
2.4.2 Borrowing Constraints	24

2.4.3	Fiscal policy	24
2.4.4	Labor income process	25
2.4.5	Model performance	26
2.5	Main Results	27
2.5.1	Aggregates	28
2.5.2	Distributional Effects	28
2.5.3	Welfare decomposition	30
2.5.4	Fixed instruments	31
2.5.5	Transitory effects	33
2.6	Complete Market Economies	36
2.6.1	Economy 1: representative agent	38
2.6.2	Economy 2: add heterogeneity in initial assets	38
2.6.3	Economy 3: add heterogeneity in productivity levels	39
2.6.4	Economy 4: add heterogeneity in both	40
2.7	Robustness	41
2.7.1	Welfare function	42
2.7.2	Labor income process	44
2.8	Conclusion	45
3	Endogenous Borrowing Constraints in Heterogenous Firms Economy	46
3.1	Introduction	46
3.2	Business cycle facts on firm dynamics	55
3.2.1	Data	56
3.2.2	Age rather than size matters.	57
3.2.3	The role of the extensive margin	59
3.2.4	The Great Recession and it's aftermath	61
3.3	Model of firm dynamics with asymmetric response to uncertainty shocks	63
3.3.1	Environment	64
3.3.2	Firms	65
3.3.3	Credit market and Financial Intermediation	67
3.3.4	Workers	67
3.3.5	Dynamic lending contract	67

3.3.6	Aggregation	71
3.3.7	Recursive equilibrium	73
3.4	Theoretical results	73
3.4.1	Preliminaries	74
3.4.2	Borrowing constraint as a function of age.	75
3.4.3	Access to financing increases with average return	79
3.4.4	Existence of a stationary equilibrium	81
3.4.5	Policy functions, equilibrium and movements in uncertainty	82
3.5	Calibration and Quantitative results	84
3.5.1	Calibration	85
3.5.2	The Effects of Uncertainty Shocks	89
3.5.3	The Effects of Productivity Shocks	93
	References	95
	Appendix A.	104
A.1	Proofs for two-period economies	104
A.1.1	Uncertainty economy	104
A.1.2	Inequality economy	106
A.2	Proofs for complete market economies	106
A.2.1	Economy 2	110
A.2.2	Economy 3	112
A.2.3	Economy 4	112
A.3	Welfare decomposition	112
A.4	Algorithms	115
A.5	Tables and Figures	117
	Appendix B.	125
B.1	Data	125
B.1.1	BDS vs. CPS vs. Establishment survey	125
B.1.2	Age/Size distribution of firms and employment	127
B.1.3	Cross correlations	127

B.2 Proofs	129
B.2.1 Goods market clearing.	129
B.2.2 Proofs from theory section	130
B.3 Algorithms	145

List of Tables

2.1	Benchmark Model Economy: Target Statistics and Parameters	23
2.2	Proportion in favor of reform	29
2.3	Welfare decomposition	31
2.4	Welfare decomposition: Fixed instruments	32
2.5	Final Stationary Equilibrium: transitory effects	34
2.6	Welfare decomposition: transitory effects	34
2.7	Robustness: Welfare Function	43
2.8	Robustness: Labor income process	44
3.1	Employment volatility for different groups of firms.	58
3.2	Standard deviation of employment over age and size distribution.	59
3.3	Decomposition of employment variance for different group of firms	61
3.4	Number of firms in pre- and post-2007 period.	62
3.5	Exogenously Determined Parameters of the Baseline Economy	86
3.6	Preference and Technology Parameters and Associated Targets	86
3.7	Size/Age Distribution Parameters and Associated Targets	87
3.8	Effect of shocks on employment (% dev. from steady state)	93
A.1	Distribution of wealth	117
A.2	Income sources by quintiles of wealth	117
B.1	Cross correlations of cyclical components of employment and GDP.	126
B.2	Distribution of firms over age and size groups (%).	127
B.3	Distribution of employment over age and size groups (%).	127
B.4	Contemporaneous correlation of cyclical components of employment and GDP.	128

List of Figures

2.1	Optimal taxes in the presence of both uncertainty and inequality.	17
2.2	Income tax schedule	25
2.3	Optimal Fiscal Policy: Benchmark	27
2.4	Inequality measures	29
2.5	Optimal Taxes: Complete Market Economies	37
2.6	Optimal Fiscal Policy with 4 nodes	42
3.1	Cyclical component of employment for different group of firms.	60
3.2	Employment index for various groups of firms in pre-2007 and 2007-2009 recessions.	63
3.3	Life cycle of an average firm and role of the aggregate uncertainty.	81
3.4	Qualitative properties of the value and contract policy functions	84
3.5	Effects of Uncertainty Shock on macroeconomic outcomes.	92
A.1	Aggregates: Benchmark	118
A.2	Optimal Fiscal Policy: Fixed Capital Taxes	119
A.3	Optimal Fiscal Policy: Fixed Labor Taxes	120
A.4	Optimal Fiscal Policy: Lump-Sum Transfers to Output	121
A.5	Optimal Fiscal Policy: Fixed debt-to-output	122
A.6	Aggregates: Constant Policy	123
A.7	Optimal Fiscal Policy: Constant Policy	124
A.8	Economy 3: Θ_t	124
B.1	Aggregate employment: BDS vs. CPS vs. Establishment survey.	126
B.2	Cross correlations of cyclical components of employment and GDP.	128

Chapter 1

Introduction

Heterogeneity of individuals is ubiquitous feature of reality. However, for many years macroeconomic profession has been operating under the assumption of representative household and representative firm. This approach by construction limited the set of questions macroeconomists were able to address. Recently, macroeconomic models with heterogeneous agents have become a major tool used to address research questions. This thesis contributes to this stream of research. It presents two essays which use the environment with heterogeneity to address research questions. The first essay Optimal Fiscal Policy in the Heterogeneous Agents Economy starts with the fundamental question in public economics: to what extent should governments use fiscal policy instruments to provide redistribution and insurance? The main contribution of this essay is to provide a quantitative analysis and answer to this question using the model that replicates inequality and individual risk present in the US economy. The second essay Endogenous Borrowing Constraints in Heterogeneous Firms Economy starts with the key observation that age rather than size of the firm is a determinant of the cyclical employment dynamics. As a result employment of the young firms is 2.6. times more volatile relative to the employment of the old firms over the business cycle. Next, the essay proposes a theory of endogenous borrowing constraints and asymmetric firm dynamics over the business cycles, which builds on two ubiquitous features of the credit markets: relevance of past performance and long-term nature of financial arrangements. The innovative part of this essay is a quantitative theory of economic downturn and credit crunch that relies on the efficient, risk sharing financial

arrangements between firms and financial intermediaries facing the presence of the private information in the financial markets. The theme which unites the essays is the focus on the role of heterogeneity for the economic questions of interests. On the one hand, heterogeneity of agents in terms of wealth and productivity implies new insights for the shape of fiscal policy and leads to results contrasting starkly with well established ones in the representative agent literature. On the other hand, the empirical observation documented in the second essay calls immediately for the model with firm heterogeneity in terms of their age and size. The concept of a representative firm is inadequate to understand the driving forces behind business cycle employment fluctuations of various groups of firms shaping the aggregate employment dynamics.

Chapter 2

Optimal Fiscal Policy in the Heterogenous Agents Economy

2.1 Introduction

How and to what extent should governments tax capital and labor income if they care about individual income inequality and risk? We want to provide a quantitative answer to this question. We, therefore, need a model that is able to generate realistic levels of income inequality and uninsurable risk. Our approach in this paper is to numerically solve a Ramsey problem in a quantitative general equilibrium model with heterogenous agents and uninsurable idiosyncratic risk - from now on referred to as the standard incomplete markets (SIM) model¹.

The SIM model has been used extensively for positive analysis and been relatively successful at matching some basic facts about inequality and uncertainty². In this environment agents face uncertainty with respect to their individual labor productivity which they cannot directly insure against (only a risk-free asset is available). Depending on their productivity realizations they make different savings choices which leads to endogenous wealth inequality. As a result, on top of the usual concern about not distorting agents

¹This type of model was originally developed and analyzed by [Bewley \(1986\)](#), [Imrohoroglu \(1989\)](#), [Huggett \(1993\)](#), and [Aiyagari \(1994\)](#).

²Our calibration strategy is similar to the ones in [Domeij and Heathcote \(2004\)](#) and [Castañeda, Díaz-Giménez, and Ríos-Rull \(2003\)](#).

decisions, a (utilitarian) Ramsey planner has two additional objectives: to redistribute resources across agents, and to provide insurance against their idiosyncratic productivity risk.

The study of optimal fiscal policy in the SIM model has focused, so far, on the maximization of steady state welfare³. In contrast, we allow policy to be *time varying* and the welfare function to depend on the associated *transition* path. We calibrate the initial steady state to replicate several aspects of the US economy; in particular the fiscal policy, the distribution of wealth, and statistical properties of the individual labor income process. The final steady state is, then, endogenously determined by the path of fiscal policy. The Ramsey planner finances an exogenous stream of government expenditures with four instruments: proportional capital and labor income taxes, (possibly negative) lump-sum transfers, and government debt.

Labor and capital income taxes are distortive, however, they can be used to provide insurance and redistribution. The only uncertainty that agents face, in our environment, is with respect to their labor productivities⁴. Hence, labor income is the only risky part of the agents' income. By taxing labor income and rebating the extra revenue via lump-sum, the planner can reduce the proportion of the agents' income that is uncertain and effectively provide insurance. On the other hand, capital income is particularly unequal, since the inequality of individual asset holdings is high, and by taxing capital the planner can reduce the proportion of unequal income in total income and, this way, provide redistribution. The effect of government debt is more subtle. Increasing government debt the government crowds out capital which affects prices indirectly, in particular reducing wages and increasing interest rates which leads to a less uncertain but more unequal distribution of income. The optimal fiscal policy weighs all these effects against each other.

We find that capital income taxes should be front-loaded hitting the imposed upper bound of 100 percent for 33 years then decreases to 45 percent in the long-run. Labor

³See, for instance, [Aiyagari and McGrattan \(1998\)](#), [Conesa, Kitao, and Krueger \(2009a\)](#), and [Nakajima \(2010\)](#).

⁴[Panousi and Reis \(2012\)](#) and [Evans \(2014\)](#) focus instead on investment risk. One justification for our focus on labor income risk is the fact that it is a bigger share of the total income for most agents in the economy. The bottom 80 percent in the distribution of net worth have a share of labor income above 77 percent, in the 2007 SCF.

income taxes are reduced to less than half of their initial level, from 28 percent to about 13 percent in the long-run. The ratio of lump-sum transfers to output is reduced to about a half of its initial level of 8 percent and the government accumulates assets over time; the debt-to-output ratio decreases from 63 percent to -17 percent in the long-run. Relative to keeping fiscal instruments at their initial levels, this leads to a welfare gain equivalent to a permanent 4.9 percent increase in consumption.

Unlike the Ramsey problem solved for representative-agent economies, in this paper we do not rule out lump-sum taxation. The optimal levels of distortive taxation are, therefore, derived rather than imposed. Even though lump-sum taxes are available, the planner chooses to tax both capital and labor income at positive rates, rebating the associated revenue via lump-sum *transfers*. Relative to a system that obtains all revenue via lump-sum taxes, such a tax system changes the composition of agents' after-tax income, reducing the proportions associated with uncertain and unequal labor and capital incomes and increasing the proportion of certain and equal transfer income; providing insurance and redistribution. To clarify this point and to understand exactly how the optimal policy reacts to changes in uncertainty and inequality we provide an analytic characterization of the solution to the Ramsey problem in a simple two-period version of the SIM model.

We decompose the average welfare gains of 4.9 percent associated with implementing the optimal policy into three parts: (i) 3.7 percent come from the more efficient allocation of aggregate resources due to the reduction of the distortions of agents' decisions; (ii) 4.9 percent come from redistribution - the reduction in ex-ante inequality; and (iii) -3.7 percent come from the reduction in insurance - there is more uncertainty about individual consumption and labor streams under the optimal policy. The optimal policy implies an overall increase of capital taxes and a reduction of labor taxes. The net effect on the distortions of agents' savings and labor supply decisions is positive. The higher capital taxes decrease the proportion of the agents' income associated with the highly unequal asset income and lead to the redistributive gains. Finally, a lower labor income tax leads to a higher proportion of the agents' income to come from the uncertain labor income, thus the negative insurance effect.

We show that disregarding transitory welfare effects can be severely misleading. To

make this point we compute the stationary fiscal policy that maximizes welfare in the final steady state, which leads to a 9.8 percent greater steady state welfare than the initial steady state. However, once transitory effects are considered, implementing this policy leads to a welfare *loss* of 6.4 percent relative to keeping the initial fiscal policy. Relative to the fiscal policy that maximizes welfare over transition it leads to a welfare loss of 11.3 percent.

In order to illustrate the role of market incompleteness in our findings, we develop the following build-up. We start from the representative agent economy and sequentially introduce heterogeneity in initial assets; different (but constant and certain) individual productivity levels; and, finally, uninsurable idiosyncratic productivity risk which adds up to the SIM model. At each intermediate step, building on the work of [Werning \(2007\)](#), we analytically characterize and then numerically compute the optimal fiscal policy over transition identifying the effect of adding each feature. In particular, we show that the planner will choose to keep capital taxes at the upper bound in the initial periods if there is asset heterogeneity, before reducing it to zero. Productivity heterogeneity rationalizes positive (and virtually constant) labor taxes. The key qualitative difference of the solution once uninsurable idiosyncratic productivity risk is introduced is that long-run capital income taxes are set to a positive level. Rationales for this result already exist in the literature and are discussed in the next section. To our knowledge, however, the level of the optimal long-run capital taxes in the SIM model had not been obtained before.

Finally, we present robustness exercises with respect to the welfare function and the calibration of the labor income process. Our benchmark results are for the utilitarian welfare function which implies a particular social choice with respect to the equality versus efficiency trade-off. We introduce a parameter in the welfare function that allows for different choices, in particular for the planner to completely ignore equality concerns. The long-run levels of capital and labor taxes are surprisingly resistant to changes in this parameter. What does change significantly, however, is how long the capital tax is maintained at the upper bound; the more the planner “cares” about inequality the more years it keeps those taxes at the upper bound. With respect to different calibrations of the labor income process, the magnitudes of the taxes are affected, but the qualitative features are maintained.

Related Literature

This paper is related to several strands of literature. First, it is related to the literature on the steady state optimal fiscal policy in the SIM model. In an influential paper, [Conesa, Kitao, and Krueger \(2009a\)](#) solve for the tax system that maximizes steady state welfare in an overlapping generations SIM model. Their result includes an optimal long-run capital income tax of 36 percent. It is important to note that though this result is similar to ours the reasons behind it are different. They diagnose that their optimal capital tax level follows from the planner's inability to condition taxes on age, and the fact that a positive capital tax can mimic age-conditioned taxes in a welfare improving way (see [Erosa and Gervais \(2002\)](#)). This mechanism is not present in our analysis since we abstract from life-cycle issues.

[Aiyagari \(1995\)](#) and [Chamley \(2001\)](#) provide rationales for positive long-run capital taxes in environments similar to ours. [Aiyagari \(1995\)](#)'s logic depends on the planner choosing the path of government expenditure (appearing separably in the agent's utility function). The associated Euler equation implies the modified golden rule level of capital which can only be achieved by taxing savings; the planner does not have precautionary motives while the agents do. In our environment positive long-run capital taxes are preserved with exogenous governmental spending. [Chamley \(2001\)](#) shows, in a partial equilibrium version of the SIM model, that enough periods in the future every agent has the same probability of being in each of the possible individual (asset/productivity) states. It is, therefore, Pareto improving to transfer from the consumption-rich to the consumption-poor in the long-run. If the correlation of asset holdings with consumption is positive, this transfer can be achieved by a positive capital tax rebated via lump-sum. In short, an agent's asset level in the long-run is a good proxy for how lucky she has been; hence, taxing it is a good way to provide insurance in the long-run. In recent work, [Dávila, Hong, Krusell, and Ríos-Rull \(2012\)](#) solve the problem of a planner that is restricted to satisfy agents' budget constraints, but is allowed to choose the savings of each agent. If the consumption-poor's share of labor income is higher than the average, increasing the aggregate capital stock relative to the undistorted equilibrium can improve welfare through its indirect effect on wages and interest rates. In our setup, the Ramsey planner taxes capital to affect after tax interest rates directly and achieves the same goal.

Another important work on fiscal policy in the SIM model is [Aiyagari and McGrattan \(1998\)](#), who search for the level of debt-to-output that maximizes steady state welfare. Interestingly, they find that the optimal level is very close to the pre recession level of around 67 percent. The fact that they abstract from the transitional dynamics makes the result even more remarkable: the government could choose its level of asset without having to finance it over time, it could, for instance choose to have enough assets to finance all its expenditures and yet it chooses to remain in debt. By holding debt, the government crowds out capital increasing interest rates and decreasing wages. This effectively provides insurance since the proportion of uncertain labor income out of total income is reduced. This benefit is what drives the choice of the government to hold debt. However, there is another effect associated with such a policy; it increases inequality (the proportion of the unequal asset income out of total income increases). This negative effect is not particularly important in [Aiyagari and McGrattan \(1998\)](#) because their calibration focuses on matching labor income processes which leads to an underestimation of wealth inequality. [Winter and Roehrs \(2014\)](#) replicate their experiment with a calibration that targets wealth inequality statistics and find the opposite result, i.e. the government chooses to hold high levels of assets. Our calibration procedure is closer to that of [Winter and Roehrs \(2014\)](#), which elucidates our result that the Ramsey planner chooses to accumulate assets over time.

[Heathcote, Storesletten, and Violante \(2014\)](#) and [Gottardi, Kajii, and Nakajima \(2014b\)](#) characterize the optimal fiscal policy in stylized versions of the SIM model. Their approaches lead to elegant and insightful closed-form solutions. The environment and Ramsey problem in [Gottardi, Kajii, and Nakajima \(2014b\)](#) is similar to ours except for the simplifications that yield tractability; i.e. exogenous labor supply, the absence of borrowing constraints, and i.i.d. shocks to human capital accumulation. [Heathcote, Storesletten, and Violante \(2014\)](#), on the other hand, focus on different, though related, questions. By abstracting from capital accumulation, they are able to retain tractability in a model with progressive taxation, partial insurance, endogenous government expenditure and skill choices (with imperfect substitution between skill types). This leads to several interesting dimensions that, in our paper, we abstract from. However, the simplifications in these

models do not allow them to match some aspects of the data which we find to be important for the determination of the optimal tax system. In particular, the model in [Heathcote, Storesletten, and Violante \(2014\)](#) implies no wealth inequality (wealth is zero for all agents). Our calibration strategy allows us to match the distribution of wealth in the US.

We also contribute to the literature highlighting the importance of transition for policy prescriptions in incomplete markets models. [Domeij and Heathcote \(2004\)](#) use the SIM model to evaluate the implementation of a zero capital income tax policy taking into account the transitional welfare effects. They conclude that such a reform would be detrimental to welfare due to its transitory effect on inequality. [Krueger and Ludwig \(2013\)](#), [Poschke, Kaymak, and Bakis \(2012\)](#), and [Winter and Roehrs \(2014\)](#) also conduct experiments in this spirit. [Acikgoz \(2013\)](#) claims that the optimal long-run fiscal policy is independent of initial conditions and the transition towards it. He, then, studies the properties of fiscal policy in the long-run, but is silent about the optimal transition path which is the focus of this paper.

There is an extensive literature that studies the Ramsey problem in complete market economies; see [Chari and Kehoe \(1999\)](#) for a survey. The most well known result for the deterministic subset of these economies is due to [Judd \(1985\)](#) and [Chamley \(1986\)](#); capital taxes should converge to zero in the long run. Among others, [Jones, Manuelli, and Rossi \(1997\)](#) and [Atkeson, Chari, and Kehoe \(1999\)](#), show this result is robust to a relaxation of a number of assumptions. As was described above we make an effort to relate our main results to the results in this literature.

The New Dynamic Public Finance literature takes an alternative approach to answer our initial question. It focuses on the design of a mechanism that would allow the planner to extract information about the agents' unobservable productivities efficiently. It assumes tax instruments are unrestricted and in this sense it dominates the Ramsey approach in terms of generality, since the latter ignores the information extraction problem⁵ and imposes ad-hoc linearity restrictions on the tax system. One of the main results stemming

⁵The Ramsey planner is also unable to observe productivity levels, it is not allowed to condition taxes on them.

from this literature is the inverse Euler equation; see [Goloso, Kocherlakota, and Tsyvinski \(2003\)](#). [Farhi and Werning \(2012\)](#) show that starting from the allocations from the steady state of an undistorted SIM model and applying perturbations to implement the inverse Euler equation leads to small welfare gains, of the order of 0.2 percent. Moreover, it is difficult to solve the private information problem in dynamic economies with persistent shocks. [Farhi and Werning \(2013\)](#) and [Troshkin, Tsyvinski, and Goloso \(2010\)](#) have made advancements in this direction in partial equilibrium settings and find that restrictions to linear taxes lead to small welfare losses. Our view is that, even if only as a benchmark to more elaborate tax systems, it is useful to understand the properties of a simpler optimal linear tax system in a quantitative general equilibrium environment.

The rest of the paper is organized as follows. Section [2.2](#) illustrates the main mechanism behind our results in a two-period economy. Section [2.3](#) describes the infinite horizon model, sets up the Ramsey problem and discusses our solution technique. Section [2.4](#) describes the calibration. Section [2.5](#) presents the main results of the paper. Section [2.6](#) presents the build-up from the complete market economy results to our main results. Section [2.7](#) provides results for alternative welfare functions and calibrations and Section [2.8](#) concludes.

2.2 Mechanism: Two-Period Economy

In the SIM model, there are two dimensions of heterogeneity: productivity and wealth. Agents have different levels of productivity which follow an exogenous random process. In addition, markets are incomplete and only a risk-free asset exists. Therefore, the idiosyncratic productivity risk cannot be diversified away. It follows that the history of shocks, affects the amount of wealth accumulated by each agent and there is an endogenously determined distribution of wealth.

In a two-period economy, it is possible to evaluate how each dimension of heterogeneity affects the optimal tax system. Since there is no previous history of shocks the initial wealth inequality can be set exogenously. In this section, we characterize, under some assumptions about preferences, the optimal tax system when the government has access to

linear labor and capital income taxes, and (possibly negative) lump-sum transfers. First, we assume agents have the same level of wealth but face an idiosyncratic productivity shock - we call this the *uncertainty economy*. Then, we shut down uncertainty and introduce ex-ante wealth inequality - this is referred to as the *inequality economy*. Next we consider the case in which there is uncertainty *and* inequality and discuss the relationship with the infinite horizon problem.

2.2.1 Uncertainty economy

Consider an economy with a measure one of ex-ante identical agents who live for two periods. Suppose they have time-additive, von Neumann-Morgenstern utility functions. Denote the period utility function by $u(c, n)$ where c and n are the levels of consumption and labor supplied. Assume u satisfies the usual conditions and denote the discount factor by β . In the first period each agent is endowed with ω units of the consumption good which can be either consumed or invested into a risk-free asset, a , and supplies \bar{n} units of labor inelastically.

In period 2, consumers receive income from the asset they saved in period 1 and from labor. Labor is supplied endogenously by each agent in period 2 and the individual labor productivity, e , is random and can take two values: e_L with probability π and $e_H > e_L$ with probability $1 - \pi$, with the normalization $\pi e_L + (1 - \pi) e_H = 1$. Due to the independence of shocks across consumers a law of large numbers operates so that in period 2 the fraction of agents with e_L is π and with e_H is $(1 - \pi)$. Letting n_i be the labor supply of an agent with productivity e_i , it follows that the aggregate labor supply is $N = \pi e_L n_L + (1 - \pi) e_H n_H$.

The planner needs to finance an expenditure of G in period 2. It has three instruments available: labor and capital income taxes, τ^n and τ^k , and lump-sum transfers T which can be positive or negative. Let w be the wage rate and r the interest rate. The total period 2 income of an agent with productivity e_i is, therefore, $(1 - \tau^n) w e_i n_i + (1 + (1 - \tau^k) r) a + T$. In period 2, output is produced using capital, K , and labor and a constant-returns-to-scale neoclassical production function $f(K, N)$. We assume that $f(\cdot)$ is net of depreciation.

Definition 1. A *tax distorted competitive equilibrium* is a vector $(K, n_L, n_H, r, w; \tau^n, \tau^k, T)$ such that

1. (K, n_L, n_H) solves

$$\max_{a, n_L, n_H} u(\omega - a, \bar{n}) + \beta E[u(c_i, n_i)] \quad s.t. \quad c_i = (1 - \tau^n) w e_i n_i + \left(1 + (1 - \tau^k) r\right) a + T;$$

2. $r = f_K(K, N)$, $w = f_N(K, N)$, where $N = \pi e_L n_L + (1 - \pi) e_H n_H$;

3. and, $\tau^n w N + \tau^k r K = G + T$.

The Ramsey problem is to choose τ^n , τ^k , and T to maximize welfare. Since agents are ex-ante identical there is no ambiguity about which welfare function to use, it is the expected utility of the agents. If there is no risk, i.e. $e_L = e_H$, the agents are also ex-post identical and the usual representative agent result applies: since negative lump-sum transfers are available, it is optimal to obtain all revenue via this undistortive instrument and set $\tau^n = \tau^k = 0$.

In order to provide a sharp characterization of the optimal tax system we make the following assumption discussed below⁶.

Assumption 1. *No income effects on labor supply and constant Frisch elasticity, κ , i.e.*

$$u_{cn} - u_{cc} \frac{u_n}{u_c} = 0, \quad \text{and} \quad \frac{u_{cc} u_n}{n(u_{cc} u_{nn} - u_{cn}^2)} = \kappa.$$

We pursue a variational approach. Suppose $(K, n_L, n_H, r, w; \tau^n, \tau^k, T)$ is a tax distorted equilibrium⁷. We consider a small variation on the tax system $(d\tau^n, d\tau^k, dT)$, such that all the equilibrium conditions are satisfied. Then, evaluate the effect of such a variation on welfare, taking as given the optimal decision rules of the agents. Using this method we establish the following proposition (derivations and proofs are in Appendix A.1).

⁶In a similar two-period environment, [Gottardi et al. \(2014a\)](#) characterize the solution to Ramsey problem without Assumption A. However, they impose an alternative assumption about endogenous variables which are satisfied under Assumption A. Further, this assumption allows us to provide a sharper characterization of the optimal tax system (besides the signs of taxes we also characterize the levels).

⁷Since the equilibrium does not exist for $\tau^n \geq 1$ or $\tau^k \geq (1+r)/r$, we impose the restrictions that $\tau^n < 1$ and $\tau^k < (1+r)/r$.

Proposition 1. *In the uncertainty economy, if u satisfies Assumption A, then, the optimal tax system is such that $\tau^k = 0$,*

$$\tau^n = \frac{(\nu - 1) \pi(1 - \pi) (e_H n_H - e_L n_L)}{(\nu - 1) \pi(1 - \pi) (e_H n_H - e_L n_L) + \kappa N (\pi \nu + (1 - \pi))} > 0, \quad (2.1)$$

where $\nu \equiv \frac{u_c(c_L, n_L)}{u_c(c_H, n_H)}$, and $T < 0$ balances the budget.

Notice that the planner could choose to finance G with T but chooses a positive distortive labor income tax instead. The revenue from labor taxation is rebated via lump-sum transfers and the proportion of the agents' income that comes from the uncertain labor income is reduced. Hence, this tax system effectively provides insurance to the agents. Why not provide full insurance by taxing away all the labor income? This is exactly what would happen if labor were supplied inelastically. In fact, notice that in this case $\kappa = 0$ and equation (2.1) implies $\tau^n = 1$. However, with an endogenous labor supply the planner has to balance two objectives: minimize distortions to agents' decisions and provide insurance. This balance is explicit in equation (2.1) seeing as a higher κ implies a lower τ^n . That is, the more responsive labor supply is to changes in labor taxes the more distortive these taxes are and the planner chooses a lower labor tax. In the limit, if $\kappa \rightarrow \infty$ it will be optimal to set $\tau^n = 0$.

With income effects on labor supply, distortions of the savings decision would spill over to the labor supply decision and vice-versa. Thus, it could be optimal, for instance, to choose τ^k so as to mitigate the distortion imposed by a positive τ^n . This complex relationship complicates the analysis considerably. Assumption 1 unties this relationship and as a result it is optimal to set $\tau^k = 0$.

Next, suppose that $e_L = 1 - \epsilon^{unc}/\pi$ and $e_H = 1 + \epsilon^{unc}/(1 - \pi)$, so that ϵ^{unc} is a mean preserving spread on the productivity levels. It is easy to see that if $\epsilon^{unc} = 0$ equation (2.1) implies that $\tau^n = 0$. The effect of an increase in ϵ^{unc} on the optimal τ^n is not as obvious since the right hand side of equation (2.1) contains endogenous variables. An application of the implicit function theorem, however, clarifies that as long as $\partial\nu/\partial\epsilon^{unc} > 0$ and $\partial\nu/\partial\tau^n < 0$, it follows that $\partial\tau^n/\partial\epsilon^{unc} > 0$, i.e. the optimal labor income tax is increasing in the level of risk in the economy. Under standard calibrations, the equilibrium ratio of

marginal utilities, ν , is in fact increasing in the level of risk ($\partial\nu/\partial\epsilon^{unc} > 0$) and decreasing in the labor income tax ($\partial\nu/\partial\tau^n < 0$), as an example see section 2.2.3.

2.2.2 Inequality economy

Consider the environment described above only without uncertainty and with initial wealth inequality. That is, suppose the productivity levels do not vary between agents, i.e. $e_L = e_H = 1$, and that ω can take two values: ω_L for a proportion p of the agents and $\omega_H > \omega_L$ for the rest, with $\bar{\omega} \equiv p\omega_L + (1-p)\omega_H$.

Definition 2. A *tax distorted competitive equilibrium* is $(a_L, a_H, n_L, n_H, r, w; \tau^n, \tau^k, T)$ such that

1. For $i \in \{L, H\}$, (a_i, n_i) solves

$$\max_{a_i, n_i} u(\omega_i - a_i, \bar{n}) + \beta u(c_i, n_i), \quad \text{s.t. } c_i = (1 - \tau^n) w n_i + \left(1 + (1 - \tau^k) r\right) a_i + T;$$

2. $r = f_K(K, N)$, $w = f_N(K, N)$, where $K = pa_L + (1-p)a_H$ and $N = pn_L + (1-p)n_H$;

3. and, $\tau^n w N + \tau^k r K = G + T$.

In this economy the concept of optimality is no longer unambiguous. Since agents are different ex-ante, a decision must be made with respect to the social welfare function. In what follows, by optimal we mean the one that maximizes $W \equiv pU_L + (1-p)U_H$; the utilitarian welfare function. The following proposition follows.

Proposition 2. In the inequality economy, if u satisfies Assumption A and has CARA or GHH as in equation (2.4), then the optimal tax system is such that $\tau^n = 0$,

$$\tau^k = \frac{\left(\frac{1+r}{r}\right) (\nu - 1) p(1-p) (\omega_H - \omega_L)}{(\nu - 1) p(1-p) (\omega_H - \omega_L) + \frac{\rho}{\psi} (p\nu + (1-p))} > 0, \quad (2.2)$$

where $\rho \equiv \frac{2+(1-\tau^k)r}{2+r}$ for CARA, $\rho \equiv \frac{1+\beta^{-\frac{1}{\sigma}}(1+(1-\tau^k)r)^{\frac{\sigma-1}{\sigma}}}{1+r+\beta^{\frac{1}{\sigma}}(1+(1-\tau^k)r)^{\frac{1}{\sigma}}}$ for GHH, and ψ is the level of absolute risk aversion⁸. $T < 0$ balances the budget.

⁸The level of absolute risk aversion is endogenous in the GHH case.

The planner chooses a positive capital income tax which distorts savings decisions but allows for redistribution between agents. The ex-ante wealth inequality is exogenously given. However, agents with different wealth levels in the first period will save different amounts and have different asset levels in the second period. This endogenously generated asset inequality is the one the tax system is able to affect. A positive capital tax rebated via lump-sum transfers directly reduces the proportion of the agents' income that will be dependent on unequal asset income achieving the desired redistribution which implies a reduction of consumption inequality. A related result was established in [Dávila et al. \(2012\)](#). They show that the competitive equilibrium allocation in the SIM model is constrained inefficient. That is, the incomplete market structure itself induces outcomes that could be improved upon if consumers merely acted differently; if they used the same set of markets but departed from purely self-interested optimization. The constrained inefficiency results from a pecuniary externality. The savings and labor supply decisions of the agents affects the wage and interest rates and, therefore, the uncertainty and inequality in the economy. These effects are not internalized by the agents and inefficiency follows. Notice that the planner's problem in their environment is significantly different from the Ramsey problem described here. There the planner affects allocations directly and prices indirectly whereas the Ramsey planner affects (after tax) prices directly and allocations indirectly. In the inequality economy, for instance, [Dávila et al. \(2012\)](#) show that there is underaccumulation of capital. A higher level of capital would decrease interest rates and increase wages, reducing inequality. A naive extrapolation of this logic would suggest that capital taxes should be negative so as to encourage savings. This logic, however, does not take into account the more relevant direct effect of the tax system on after tax prices. Proposition 2 shows that the opposite is true: capital taxes should be positive.

One of the key elements of equation (2.2) is the inverse of the coefficient of absolute risk aversion, $1/\psi$, which is proportional to the agents' intertemporal elasticity of substitution. This elasticity indicates the responsiveness of savings to changes in τ^k . Hence, the higher this elasticity is the lower is the optimal τ^k , since providing redistribution becomes more costly. The $\tau^n = 0$ result is again associated with Assumption 1.

Assuming that $\omega_L = 1 - \epsilon^{ine}/p$ and $\omega_H = 1 - \epsilon^{ine}/(1-p)$. The effect of an increase

in ϵ^{ine} on the optimal τ^k can again be found by applying the implicit function theorem on equation (2.2). It follows that, if $\partial\nu/\partial\epsilon^{ine} > 0$ and $\partial\nu/\partial\tau^k < 0$, then $\partial\tau^k/\partial\epsilon^{ine} > 0$; the optimal capital income tax is increasing in the level of inequality in the economy. Under the assumptions of Proposition 2 it is possible to show that this will always be the case.

2.2.3 Uncertainty and inequality

If both uncertainty and inequality are present, the optimal tax system has to balance three objectives: minimize distortions, provide insurance and redistribution. A reasonable conjecture is that under Assumption 1 the optimal tax system will be a convex combination of the ones in Propositions 1 and 2, that is, positive labor and capital income taxes with magnitudes associated with the levels of uncertainty and inequality in the economy. A more subtle extrapolation of the results above points to another interesting prediction associated with Assumption 1: the capital (labor) income taxes should be invariant with respect to the level of uncertainty (inequality). We corroborate these conjectures with a numerical example the results of which are in Figure 2.1⁹.

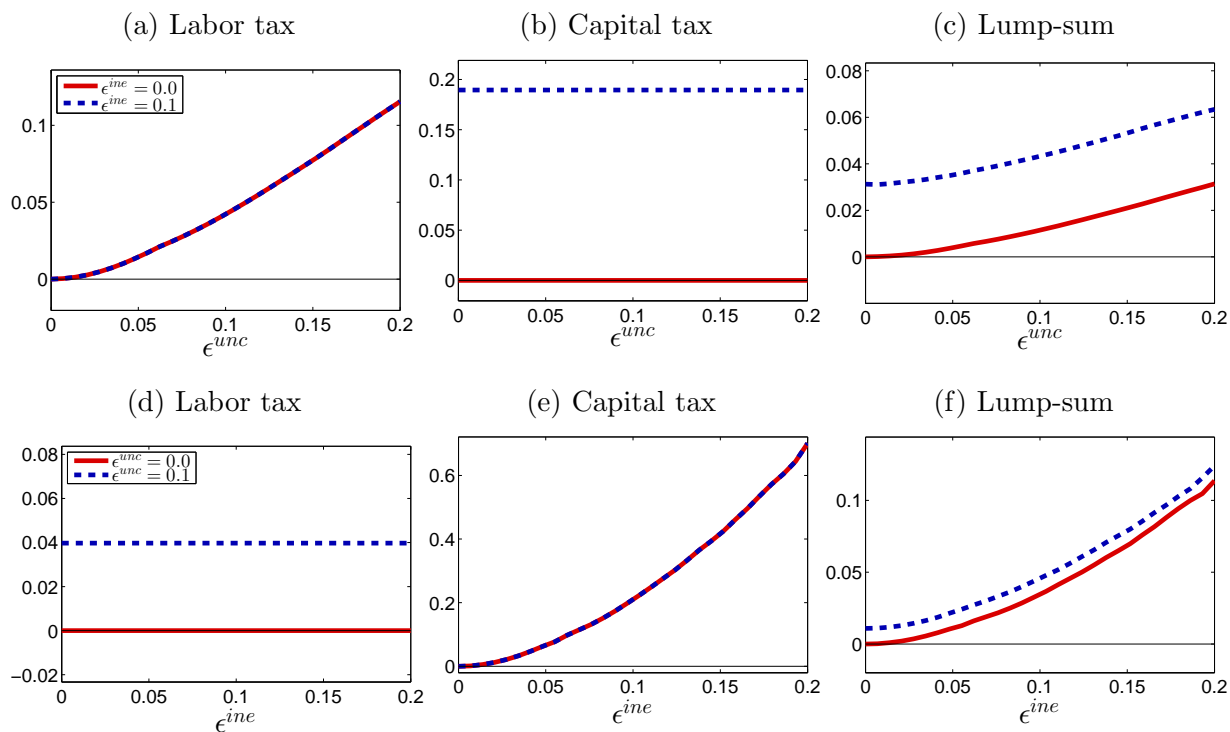
The first row of Figure 2.1 shows the optimal tax system with the level of uncertainty (embodied by the parameter ϵ^{unc}) in the x -axis with two levels of inequality: $\epsilon^{ine} = 0$ (solid line) and $\epsilon^{ine} = 0.1$ (dashed line). The solid lines corroborate Proposition 1. The comparison between the dashed and the solid lines corroborates the conjectures made above. The labor tax is increasing with the level of uncertainty and independent on the level of inequality whereas capital taxes increase with the level of inequality and are independent on level of risk. The second row of Figure 2.1 shows the results for the analogous experiment with ϵ^{ine} on the x -axis and $\epsilon^{unc} = 0$ (solid) and $\epsilon^{unc} = 0.1$ (dashed).

2.2.4 Relationship with infinite horizon problem

The two-period examples are useful to understand the key trade-offs faced by the Ramsey planner, since they allow for the exogenous setting of the levels of uncertainty (ex-post risk)

⁹We use GHH preferences which satisfy Assumption 1. The most relevant interpretation of this two-period economy is that each period corresponds to half of the working life of a person. Accordingly, we set $\beta = 0.95^{20}$ and $\delta = 1 - 0.9^{20}$. Other parameters are set to satisfy the usual targets: $\sigma = 2$, $\kappa = 0.72$, $\chi = 6$, $\bar{n} = 0.3$, $\omega = 3.5$, $\pi = p = 0.5$, and $f(K, N) = K^\alpha N^{1-\alpha} - \delta K$ with $\alpha = 0.36$. G is set to 0, but any other feasible level would just shift the lump-sum transfers correspondingly.

Figure 2.1: Optimal taxes in the presence of both uncertainty and inequality.



and inequality (ex-ante risk). In the infinite horizon version of the SIM model, however, these concepts are inevitably intertwined. The characterization of the optimal tax system, therefore, becomes considerably more complex. Labor income taxes affect not only the level of uncertainty through the mechanism described above, but also the labor income inequality and the distribution of assets over time. An agent's asset level at a particular period depends not only on its initial value, but on the history of shocks this agent has experienced. Therefore, capital income taxation affects not only the ex-ante risk faced by the agent but also the ex-post. Nevertheless, these results are useful to understand some of the key features of the optimal fiscal policy in the infinite horizon model as will become clear in what follows.

2.3 The Infinite-Horizon Model

Time is discrete and infinite, indexed by t . There is a continuum of agents with standard preferences $E_0 [\sum_t \beta^t u(c_t, n_t)]$ where c_t and n_t denote consumption and labor supplied in period t and u satisfies the usual conditions. Individual labor productivity, $e \in E$ where $E \equiv \{e_1, \dots, e_L\}$, are i.i.d. across agents and follow a Markov process governed by Γ , a transition matrix¹⁰. Agents can only accumulate a risk-free asset, a . Let $A \equiv [\underline{a}, \infty)$ be the set of possible values for a and $S \equiv E \times A$. Individual agents are indexed by the pair $(e, a) \in S$. Given a sequence of prices $\{r_t, w_t\}_{t=0}^\infty$, labor income $\{\tau_t^n\}_{t=0}^\infty$, (positive) capital income $\{\tau_t^k\}_{t=0}^\infty$, and lump-sum transfers $\{T_t\}_{t=0}^\infty$, each household, at time t , chooses $c_t(a, e)$, $n_t(a, e)$, and $a_{t+1}(a, e)$ to solve

$$v_t(a, e) = \max u(c_t(a, e), n_t(a, e)) + \beta \sum_{e_{t+1} \in E} v_{t+1}(a_{t+1}(a, e), e_{t+1}) \Gamma_{e, e_{t+1}}$$

subject to

$$(1 + \tau^c)c_t(a, e) + a_{t+1}(a, e) = (1 - \tau_t^n)w_t e n_t(a, e) + (1 + (1 - I_{\{a \geq 0\}} \tau_t^k)r_t)a + T_t$$

$$a_{t+1}(a, e) \geq \underline{a}.$$

Note that value and policy functions are indexed by time, because policies $\{\tau_t^k, \tau_t^n, T_t\}_{t=0}^\infty$ and aggregate prices $\{r_t, w_t\}_{t=0}^\infty$ are time-varying. The consumption tax, τ^c , is a parameter¹¹. Let $\{\lambda_t\}$ be a sequence of probability measures over the Borel sets \mathcal{S} of S with λ_0 given. Since the path for taxes is known, there will be a deterministic path for prices and for $\{\lambda_t\}_{t=0}^\infty$. Hence, we do not need to keep track of the distribution as an additional state; time is a sufficient statistic.

¹⁰A law of large numbers operates so that the probability distribution over E at any date t is represented by a vector $p_t \in \mathbb{R}^L$ such that given an initial distribution p_0 , $p_t = p_0 \Gamma^t$. In our exercise we make sure that Γ is such that there exists a unique $p^* = \lim_{t \rightarrow \infty} p_t$. We normalize $\sum_i p_i^* e_i = 1$.

¹¹We could potentially allow consumption taxes to also be chosen by the Ramsey planner and it is not without loss of generality that we impose this restriction. There are two reasons for this choice. The first is practical, we are already using the limit of the computational power available to us, and allowing for one more choice variable would increase it substantially. Second, for the US in particular capital and labor income taxes are chosen by the Federal Government while consumption taxes are chosen by the states, so this Ramsey problem can be understood as the one relevant for a Federal Government that takes consumption taxes as given. We need to add τ^c as a parameter for calibration purposes.

Competitive firms own a constant-returns-to-scale technology $f(\cdot)$ that uses capital, K_t , and efficient units of labor, N_t , to produce output each period ($f(\cdot)$ denotes output net of depreciation - δ denotes the capital depreciation rate). A representative firm exists that solves the usual static problem. The government needs to finance an exogenous constant stream of expenditure, G , and lump-sum transfers with taxes on consumption, labor income, and (positive) capital income. It can also issue debt $\{B_{t+1}\}$ and, thus, has the following intertemporal budget constraint

$$G + r_t B_t = B_{t+1} - B_t + \tau^c C_t + \tau_t^n w_t N_t + \tau_t^k r_t \hat{A}_t - T_t, \quad (2.3)$$

where C_t is aggregate consumption and \hat{A}_t is the tax base for the capital income tax.

Definition 3. Given an initial distribution λ_0 and a policy $\pi \equiv \{\tau_t^k, \tau_t^n, T_t\}_{t=0}^\infty$, a **competitive equilibrium** is a sequence of value functions $\{v_t\}_{t=0}^\infty$, an allocation $X \equiv \{c_t, n_t, a_{t+1}, K_t, N_t, B_t\}_{t=0}^\infty$, a price system $P \equiv \{r_t, w_t\}_{t=0}^\infty$, and a sequence of distributions $\{\lambda_t\}_{t=0}^\infty$, such that for all t :

1. Given P and π , $c_t(a, e)$, $n_t(a, e)$, and $a_{t+1}(a, e)$ solve the household's problem and $v_t(a, e)$ is the respective value function;
2. Factor prices are set competitively,

$$r_t = f_K(K_t, N_t), \quad w_t = f_N(K_t, N_t);$$

3. The probability measure λ_t satisfies

$$\lambda_{t+1} = \int_S Q_t((a, e), \mathcal{S}) d\lambda_t, \quad \forall \mathcal{S} \in \mathcal{S}$$

where Q_t is the transition probability measure;

4. The government budget constraint, (2.3), holds and debt is bounded;
5. Markets clear,

$$C_t + G_t + K_{t+1} - K_t = f(K_t, N_t), \quad K_t + B_t = \int_{A \times E} a_t(a, e) d\lambda_t.$$

2.3.1 The Ramsey Problem

We now turn to the problem of choosing the optimal tax policy in the economy described above. We assume that, in period 0, the government announces and commits to a sequence of future taxes $\{\tau_t^k, \tau_t^n, T_t\}_{t=1}^\infty$, taking period 0 taxes as given. We need the following definitions:

Definition 4. Given λ_0 , for every policy π **equilibrium allocation rules** $X(\pi)$ and **equilibrium price rules** $P(\pi)$ are such that π , $X(\pi)$, $P(\pi)$ and corresponding $\{v_t\}_{t=0}^\infty$ and $\{\lambda_t\}_{t=0}^\infty$ constitute a competitive equilibrium.

Definition 5. Given λ_0 , τ_0^k , τ_0^n , T_0 and a welfare function $W(\pi)$, the **Ramsey problem** is to $\max_\pi W(\pi)$ such that $X(\pi)$ and $P(\pi)$ are equilibrium allocation and price rules.

In our benchmark experiments we assume that the Ramsey planner maximizes the utilitarian welfare function: the ex-ante expected lifetime utility of a newborn agent who has its initial state, (a, e) , chosen at random from the initial stationary distribution λ_0 . The planner's objective is thus given by

$$W(\pi) = \int_S E_0 \sum_{t=0}^{\infty} \beta^t u(c_t(a, e|\pi), n_t(a, e|\pi)) d\lambda_0.$$

In Section 2.7 we provide results for alternative welfare functions.

2.3.2 Solution method

We solve this problem numerically. Given an initial stationary equilibrium, for any policy π we can compute the transition to a new stationary equilibrium consistent with the policy¹² and calculate welfare $W(\pi)$. We then search for the policy π that maximizes $W(\pi)$. This is, however, a daunting task since it involves searching in the space of infinite sequences. In order to make it computationally feasible we impose the following ad-hoc constraints: that each path $\{\tau_t^k, \tau_t^n, T_t\}_{t=1}^\infty$ be smooth over time and become constant after a finite amount of periods. We denote the set of policies that satisfy these properties by Π_R . These conditions are restrictive, but they allow the problem to be solved and are flexible enough to characterize some of the key features of the optimal paths of taxes.

¹²As long as the taxes become constant at some point.

The statement about the ad-hoc constraints must be qualified. It is well known from the existing solutions to the Ramsey problem in complete markets economies that capital taxes should be front-loaded. We obtain similar results in Section 2.6. Hence, in defining the set Π_R we take this under consideration. That is, we allow capital taxes to hit the imposed upper bound of 100 percent for the first t^* periods, where a model period is equivalent to one calendar year. Importantly, t^* is endogenously chosen and is allowed to be zero, so the fact that the solution displays a capital tax at the upper bound for a positive amount of periods is not an assumption but a result. Other than this, we assume that the paths for $\{\tau_t^k\}_{t=t^*+1}^\infty$ and $\{\tau_t^n, T_t\}_{t=1}^\infty$ follow splines with nodes set at exogenously selected periods. The placement of the nodes is arbitrary, we started with a small number of them and sequentially added more until the solution converged. In the main experiment the planner was allowed to choose 17 variables in total: t^* , $\tau_{t^*+1}^k$, τ_{45}^k , τ_{60}^k , τ_{100}^k , τ_1^n , τ_{15}^n , $\tau_{t^*+1}^n$, τ_{45}^k , τ_{60}^k , τ_{100}^k , T_1 , T_{15} , T_{t^*+1} , T_{45} , T_{60} , and T_{100} . In the intermediate periods the paths follow a spline function and after the final period they become constant at the last level. The choice of the periods 1, 15, 45, 60, and 100, were a result of the fact that for experiments with less nodes, the optimal t^* was always close to 30, hence we placed the nodes at the same distance from each other except for the last ones which are supposed to capture the long run levels¹³.

Solving the problem described above is a particularly hard computational task. Effectively we are maximizing $W(\pi)$ on the domain $\pi \in \Pi_R$, where each element of Π_R can be defined by a vector with a finite number of elements (the nodes described above). We know very little about its properties; it is a multivariate function with potentially many kinks, irregularities and multiple local optima¹⁴. Thus, we need a powerful and thorough procedure to make sure we find the global optimum. We use a global optimization algorithm that randomly draws a very large number of policies in Π_R and computes the transition between the exogenously given initial stationary equilibrium and a final stationary equilibrium that depends on the policy. Then, we compute welfare $W(\pi)$ for each of those policies and select those that yield the highest levels of welfare. These selected policies are then clustered, similar policies placed in the same cluster. For each cluster we run an

¹³If the solver chooses t^* close to one of these predetermined nodes the algorithm replaces that node for $t = 30$. For instance, if $t^* = 43$ the periods became 1, 15, 30, $t^* + 1$, 60, and 100.

¹⁴See [Güvenen \(2011\)](#) for a discussion of how to deal with such problems.

efficient derivative free local optimizer. The whole procedure is repeated depending on how many local optima have been found and a Bayesian stopping rule is used to figure out if enough global procedures have been run. A more detailed description of the algorithm can be found in Appendix A.4¹⁵.

2.4 Calibration

We calibrate the initial stationary equilibrium of the model economy to replicate key properties of the US economy relevant for the shape of the optimal fiscal policy. Table 2.1 summarizes our parameters choices together with the targets we use to discipline their values and their model counterparts. We use data from the NIPA tables for the period between 1995 and 2007¹⁶ and from the 2007 Survey of Consumer Finances (SCF).

2.4.1 Preferences and technology

We assume GHH preferences¹⁷ with period utility given by

$$u(c, n) = \frac{1}{1 - \sigma} \left(c - \chi \frac{n^{1 + \frac{1}{\kappa}}}{1 + \frac{1}{\kappa}} \right)^{1 - \sigma}, \quad (2.4)$$

where σ is the coefficient of relative risk aversion, κ is the Frisch elasticity of labor supply and χ is the weight on the disutility of labor. These preferences exhibit no wealth effects on labor supply, which is consistent with microeconomic evidence showing these effects are in fact small¹⁸. Further, they imply that aggregate labor supply is independent of the distribution of wealth which is convenient for computing out of steady state allocations in our main experiment. We set the intertemporal elasticity of substitution to 0.5; the number frequently used in the literature (e.g. [Dávila et al. \(2012\)](#) and [Conesa et al. \(2009a\)](#)). For

¹⁵The algorithm was parallelized for multiple cores. For each global iteration, we drew 131,072 policies and computed the transition and welfare for each of them. The number of transitions run for each cluster is endogenously determined by the local solver, on average it amounted to around 150 transitions to find each local maximum. A total of 8 global iterations were needed. We performed our analysis on the Itasca cluster at the Minnesota Supercomputing Institute using 1024 cores.

¹⁶We choose this time period to be consistent with the one used to pin down fiscal policy parameters which we take from [Trabandt and Uhlig \(2011\)](#).

¹⁷See [Greenwood et al. \(1988\)](#).

¹⁸See [Holtz-Eakin et al. \(1993\)](#), [Imbens et al. \(2001\)](#) and [Chetty et al. \(2012\)](#) for details.

Table 2.1: Benchmark Model Economy: Target Statistics and Parameters

Statistic	Target	Model	Parameter	Value
Preferences and Technology				
Intertemporal elasticity of substitution	0.50	0.50	σ	2.00*
Frisch elasticity	0.72	0.72	ν	0.72*
Average hours worked	0.30	0.30	χ	4.12
Capital to output	2.72	2.71	β	0.97
Capital income share	0.38	0.38	α	0.38*
Investment to output	0.27	0.27	δ	0.10
Borrowing Constraint				
Households with negative wealth (%)	18.6	19.1	\underline{a}	-0.04
Fiscal Policy				
Capital income tax (%)	36.0	36.0	τ_k	0.36*
Labor income tax (%)	28.0	28.0	τ_n	0.28*
Consumption tax (%)	5.0	5.0	τ_c	0.05*
Transfer to output (%)	8.0	8.0	T	0.08
debt-to-output (%)	63.0	63.0	G	0.15
Labor Productivity Process				
Wealth Gini index	0.82	0.81	e_1/e_2	0.62
Percentage of wealth in 1st quintile	-0.2	-0.2	e_3/e_2	3.89
Percentage of wealth in 4th quintile	11.2	10.2	Γ_{11}	0.94
Percentage of wealth in 5th quintile	83.4	83.4	Γ_{12}	0.05
Percentage of wealth in top 5%	60.3	60.8	Γ_{21}	0.01
Correlation btw wealth and labor income	0.29	0.29	Γ_{22}	0.92
Autocorrelation of labor income	0.90	0.90	Γ_{31}	0.01
Standard Deviation of labor income	0.20	0.20	Γ_{32}	0.04

Notes: Parameter values marked with (*) were set exogenously, all the others were endogenously and jointly determined.

the Frisch elasticity, κ , we rely on estimates from [Heathcote et al. \(2010a\)](#) and use 0.72. This value is intended to capture both the intensive and the extensive margins of labor supply adjustment together with the typical existence of two earners within a household.

It is also close to 0.82, the number reported by [Chetty et al. \(2011\)](#) in their meta-analysis of estimates for the Frisch elasticity using micro data. The value for χ is chosen so that average hours worked equals 0.3 of total available time endowment¹⁹. To pin down the discount factor, β , we target a capital to output ratio of 2.72, and the depreciation rate, δ , is set to match an investment to output ratio of 27 percent²⁰.

The aggregate technology is given by a Cobb-Douglas production function $Y = AK^\alpha N^{1-\alpha} + (1 - \delta)K$ with capital share equal to α . The total factor productivity A is set to normalize output per capita, Y , to 1. The capital share parameter, α , is set to its empirical counterpart of 0.38.

2.4.2 Borrowing Constraints

We discipline the borrowing constraint \underline{a} using the percentage of households in debt (negative net worth). We target 18.6 percent following the findings of [Wolff \(2011\)](#) based on the 2007 SCF.

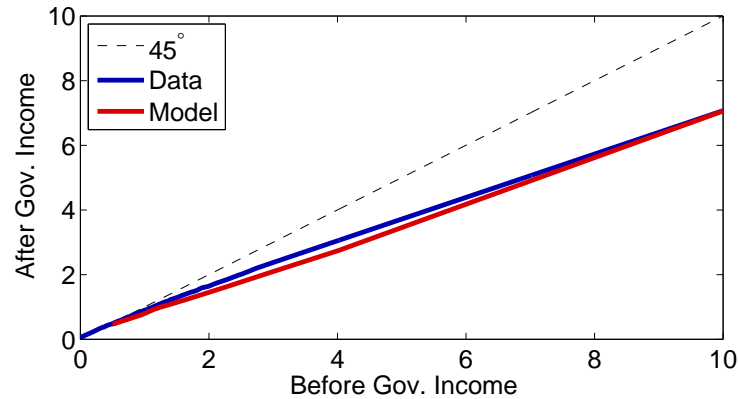
2.4.3 Fiscal policy

In order to set the tax rates in the initial stationary equilibrium we use the effective average tax rates computed by [Trabandt and Uhlig \(2011\)](#) from 1995 to 2007 and average them. The lump-sum transfers to output ratio is set to 8 percent and we discipline the government expenditure by imposing a debt to output ratio of 63 percent also following [Trabandt and Uhlig \(2011\)](#). The latter is close to the numbers used in the literature (e.g. [Aiyagari and McGrattan \(1998\)](#), [Domeij and Heathcote \(2004\)](#) or [Winter and Roehrs \(2014\)](#)). The calibrated value implies a government expenditure to output ratio of 15 percent, the data counterpart for the relevant period is approximately 18 percent. Further, we also approximate well the actual income tax schedule as can be seen in [Figure 2.2](#).

¹⁹It is understood that in any general equilibrium model all parameters affect all equilibrium objects. For the presentation purposes, we associate a parameter with the variable it affects quantitatively most.

²⁰Capital is defined as nonresidential and residential private fixed assets and purchases of consumer durables. Investment is defined in a consistent way.

Figure 2.2: Income tax schedule



Notes: The data was generously supplied by [Heathcote et al. \(2014\)](#) who used PSID and the TAXSIM program to compute it. The axis units are income relative to the mean.

2.4.4 Labor income process

The individual labor productivity levels e and transition probabilities in matrix Γ are chosen to match the US wealth distribution, statistical properties of the estimated labor income process and the correlation between wealth and labor income. There are three levels of labor productivity in our model. Since we normalize the average productivity to one we are left with two degrees of freedom. The transition matrix is 3×3 . The fact that it is a probability matrix implies its rows add up to one, therefore we are left with an additional six degrees of freedom. Thus, we end up with eight parameters to choose

It is common to use the Tauchen method when calibrating the Markov process for productivities. This method imposes symmetry of the Markov matrix which further reduces the number of free parameters. Following [Castañeda et al. \(2003\)](#) we do not impose symmetry which allows us to target at the same time statistics from the labor income process and the individual wealth distribution.

To match the wealth distribution we target shares of wealth owned by the first, fourth and fifth quintile, the share of wealth owned by individuals in the top 5 percent and the Gini index. The targets are taken from the 2007 Survey of Consumer Finances²¹. We also

²¹For a general overview of this data see [Díaz-Giménez et al. \(2011\)](#).

target properties of individual labor income estimated as the AR(1) process, namely its autocorrelation and its standard deviation²². According to [Domeij and Heathcote \(2004\)](#), existing studies estimate the first order autocorrelation of (log) labor income to lie between 0.88 and 0.96 and the standard deviation (of the innovation term in the continuous representation) of 0.12 and 0.25. We calibrate the productivity process so that the Markov matrix and vector e imply an autocorrelation of (log) labor income of 0.9 and a standard deviation of 0.2²³ (in [Section 2.7](#) we provide robustness results with respect to these choices). Finally, we target the correlation between wealth and labor income which is 0.29 in the 2007 SCF data. This way we discipline to some extent the labor income distribution using the wealth distribution that we match accurately. The resulting productivity vector, transition matrix and stationary distribution of productivities, λ_e^* , are

$$e = \begin{bmatrix} 0.79 \\ 1.27 \\ 4.94 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} .956 & .043 & .001 \\ .071 & .929 & .000 \\ .012 & .051 & .937 \end{bmatrix}, \quad \text{and} \quad \lambda_e^* = \begin{bmatrix} .616 \\ .377 \\ .007 \end{bmatrix}.$$

2.4.5 Model performance

[Table A.1](#) presents statistics about the wealth and labor income distributions. We target five of the wealth distribution statistics, so it is not surprising that we match that distribution quite well. [Table A.2](#) presents another crucial dimension along which our model is consistent with the data: income sources over the quintiles of wealth. The composition of income, specially of the consumption-poor agents, plays an important role in the determination of the optimal fiscal policy. The fraction of uncertain labor income determines the strength of the insurance motive and the fraction of the unequal asset income affects the redistributive motive. Our calibration delivers, without targeting, a good approximation of the income composition. Finally, we also match the consumption Gini which remained fairly constant around 0.27 in the period from 1995 to 2007 (see [Krueger and Perri \(2006\)](#)).

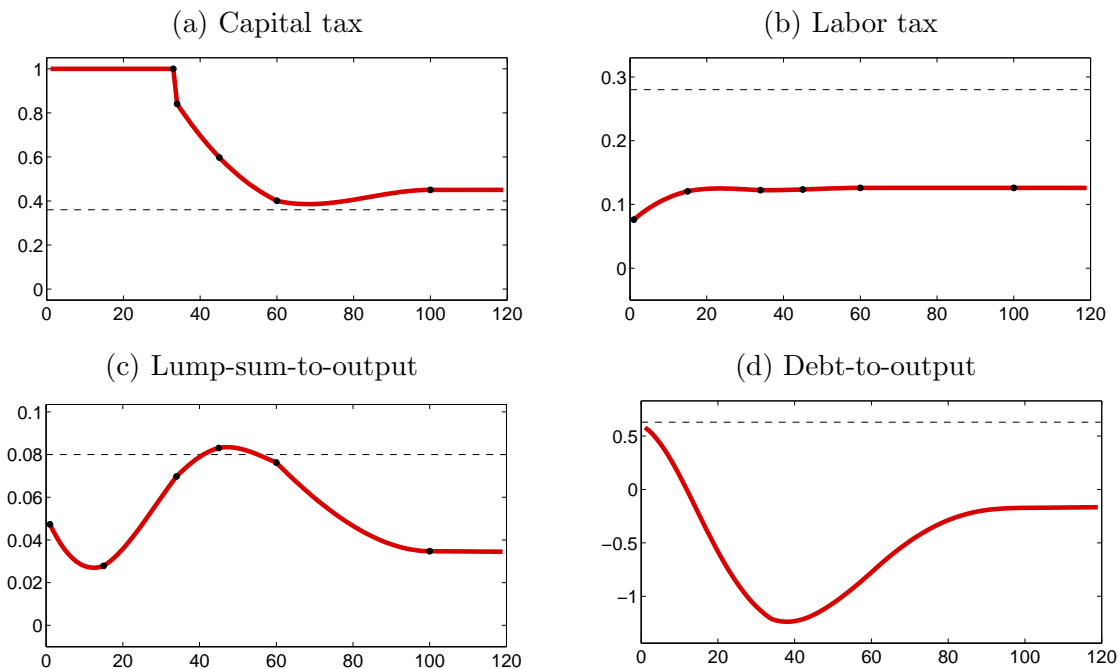
²²Including transitory shocks would allow a better match to the labor income process. However, these types of shocks can, for the most part, be privately insured against (see [Guvenen and Smith \(2013\)](#)) so we chose to abstract from them to keep the model parsimonious.

²³We follow [Nakajima \(2012\)](#) in choosing these targets. The targets are associated with labor income, wen , which includes the endogenous variables w and n . Therefore, to calibrate the parameters governing the individual productivity process, the model must be solved repeatedly until the targets are satisfied.

2.5 Main Results

The optimal paths for the fiscal policy instruments are portrayed in Figure 2.3. Capital taxes should be front-loaded hitting the upper bound for 33 initial periods then decrease to 45 percent in the long-run. Labor income taxes are substantially reduced to less than half of its initial level, from 28 percent to about 13 percent in the long-run. The ratio of lump-sum transfers to output decreases initially to about 3 percent, then increases back to its initial level of 8 percent before it starts converging to its final level of 3.5 percent. The government accumulates assets in the initial periods of high capital taxes reaching a level of debt-to-output of about -125 percent, which then converges to a final level of -17 percent. Relative to keeping fiscal instruments at their initial levels, this leads to a welfare gain equivalent to a permanent 4.9 percent increase in consumption.

Figure 2.3: Optimal Fiscal Policy: Benchmark



Notes: Dashed line: initial stationary equilibrium; Solid line: optimal transition; The black dots are the choice variables: the spline nodes and t^* , the point at which the capital tax leaves the upper bound.

2.5.1 Aggregates

The aggregates associated with the implementation of the optimal policy are shown in Figure A.1. The capital level initially decreases by about 8 percent in the first 13 years, but then increases towards a final level 20 percent higher than the initial steady state. The increase might be surprising at a first glance given the higher capital taxes. First notice that, even if capital income taxes were set to 100 percent forever, there would still be precautionary incentives for the agents with relatively high productivity to save: if they receive a negative shock they can then consume their savings. The decrease in government debt also contributes substantially to this increase - an effect we explain further below in Section 2.5.4. Most importantly though, the level of aggregate labor increases by about 15 percent immediately after the policy change following the reduction in labor taxes, increasing the marginal productivity of capital.

The higher levels of capital and labor lead to higher levels of output and consumption, which increases by 15 and 20 percent respectively over the transition. The concomitant increase in average consumption and labor has ambiguous effects on the welfare of the average agent. Hence, we also plot in Figure A.1f what we call the average consumption-labor composite, defined below in equation (2.5), which is the more relevant measure for welfare. On impact the labor-consumption composite increases by 13 percent as the higher consumption levels (due to the initial reduction in savings) more than compensate for the higher supply of labor. It then decreases for some periods following the reduction in output and the increasing savings. In the long-run it returns to a level about 13 percent higher than the one in the initial steady state.

2.5.2 Distributional Effects

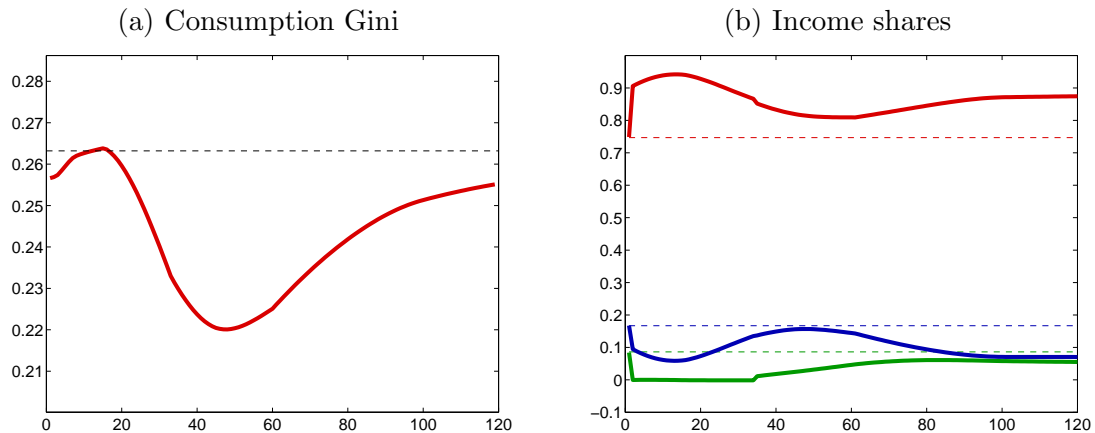
Movements in the levels do not provide a full picture of what results from the implementation of the optimal fiscal policy. It is also important to understand its effects on inequality and on the risk faced by the agents. Figure 2.4a plots the evolution of the Gini index for consumption²⁴. Notice that, though it takes some time for the reduction to start, the

²⁴Since labor supply is proportional to productivity levels, the inequality of hours is unaffected by the policy, it is in fact determined exogenously. Hence, here we can focus on consumption inequality.

consumption Gini is significantly reduced over the transition reaching a low about 16 percent lower than the initial level. As will become clear below, this reduction in inequality is behind most of the welfare gains associated with the optimal policy. Not surprisingly, such a change would be supported by most agents in the economy with the exception of the highly productive and, therefore, wealthier ones - see Table 2.2.

Figure 2.4b displays the evolution of the shares of labor, capital and transfer income out of total income. Importantly, notice that the share of labor income is significantly increased under the optimal policy. Since all the risk faced by agents in the SIM model is associated with their labor income, it turns out that they face more risk after the policy is implemented. This has an obvious negative effect on welfare which is, however, outweighed by the gains associated with the higher levels of consumption and the reduction in inequality it provides. The next sections will clarify some of these issues.

Figure 2.4: Inequality measures



Notes (a) and (b): Dashed lines: initial stationary equilibrium; Solid lines: optimal transition. Notes (b): Red lines: labor income share; Blue lines: transfer income share; Green lines: asset income share

Table 2.2: Proportion in favor of reform

$e = L$	$e = M$	$e = H$	All
99.6	98.3	3.7	99.5

2.5.3 Welfare decomposition

Here we present a result that will be particularly helpful for understanding the properties of the optimal fiscal policy. First, let x_t be the individual consumption-labor composite (the term inside the utility function 2.4), that is

$$x_t \equiv c_t - \chi \frac{n_t^{1+\frac{1}{\kappa}}}{1 + \frac{1}{\kappa}}, \quad (2.5)$$

and X_t denote its aggregate. The utilitarian welfare function can increase for three reasons. First, it will increase if the utility of the average agent, $U(\{X_t\})$, increases; we call this the *level effect*. Reductions in distortive taxes will achieve this goal by allocating resources more efficiently²⁵. Second, since agents are risk averse, it increases if the uncertainty about individual paths $\{x_t\}_{t=0}^{\infty}$ is reduced; we call this the *insurance effect*. By redistributing from the (ex-post) lucky to the (ex-post) unlucky, a tax reform can reduce the uncertainty faced by the agents. Finally, it will increase if the inequality across the certainty equivalents of the individual paths $\{x_t\}_{t=0}^{\infty}$, for agents with different initial (asset/productivity) states, is reduced; we call this the *redistribution effect*. By redistributing from the rich (ex-ante lucky) to the poor (ex-ante unlucky), the tax reform reduces the inequality between agents. In Appendix A.3 we give precise definitions for each of these effects and show how it is possible to measure them. Then, letting Δ be the average welfare gain, Δ_L the gains associated with the level effect, Δ_I with the insurance effect, and Δ_R with the redistribution effect, we prove the following proposition.

Proposition 3. *If preferences are GHH as in (2.4), then*

$$1 + \Delta = (1 + \Delta_L)(1 + \Delta_I)(1 + \Delta_R).$$

Hence, it is possible to decompose the average welfare gains into the components described above²⁶. The results for this decomposition for our main results are in Table 2.3.

²⁵This is the only relevant effect in a representative agent economy.

²⁶The welfare gains described above are in terms of consumption-labor composite units. The decomposition does not hold exactly in terms of consumption units. To keep our results comparable with others, we report the average welfare gains in terms of consumption units and normalize the numbers for Δ_L , Δ_I , and Δ_R accordingly.

Most of the welfare gains implied by the implementation of the optimal fiscal policy come from the reduction in ex-ante inequality (redistribution effect). The also substantial welfare gains associated with the reduction in distortions (level effect) is almost exactly offset by welfare losses due to the increase in uncertainty (insurance effect).

Table 2.3: Welfare decomposition

Average welfare gain	Level effect	Insurance effect	Redistribution effect
Δ	Δ_L	Δ_I	Δ_R
4.9	3.7	-3.7	4.9

2.5.4 Fixed instruments

In order to understand the role played by each instrument in the optimal fiscal policy, we ran experiments in which we hold each of them fixed and optimize only with respect to the others. Figures A.2, A.3, A.4, and A.5 display the solutions and Table 2.4 the welfare decomposition for each of these experiments.

Capital taxes

It is clear from the welfare decomposition in Table 2.4 that the path of capital taxes plays a crucial role in the redistributive gains associated with the unrestricted optimal policy. Restricting capital taxes to their initial level brings the redistribution effect from 4.9 percent to -0.2 percent. In line with the result in Proposition 2, the increase in capital taxes especially in the initial years leads to a strong redistribution effect as the proportion of unequal asset income is reduced (actually brought to zero in the first 33 years). Relative to the optimal policy, the restriction on capital taxes also leads to higher labor taxes (which explains the better insurance effect) and a lower accumulation of assets by the government.

Labor taxes

Fixing labor taxes at their initial level is particularly detrimental to the level effect. In the optimal policy labor taxes are reduced substantially and the labor supply distortions

reduced accordingly. The redistributive gains are virtually unaffected whereas the insurance effect is improved, which is consistent with the result in Proposition 1 since the restriction implies higher labor taxes. The fact that the insurance effect is still negative might be surprising though. What is behind this effect is the role played by the accumulation of assets by the government which we explain below.

Lump-sum transfers

Restricting lump-sum transfers to its initial level doesn't affect the results as much as the other restrictions; the average welfare gains are reduced from 4.9 percent to 4.4 percent. Most of the losses come from the reduction in the level effect. The restriction leads to a higher overall level of transfers and, therefore, higher labor taxes relative to the unrestricted optimal policy whereas capital taxes are virtually unaffected. This leads to an overall higher level of distortions which explains the lower level effect.

Table 2.4: Welfare decomposition: Fixed instruments

	Δ	Δ_L	Δ_I	Δ_R
Fixed capital taxes	1.0	3.7	-2.5	-0.2
Fixed labor taxes	3.3	0.0	-1.6	4.8
Fixed lump-sum	4.4	1.8	-2.5	5.1
Fixed debt	4.0	3.8	-3.2	3.2
Benchmark	4.9	3.7	-3.7	4.9

Government debt

In the absence of borrowing constraints an increase in government debt is innocuous, in response agents simply adjust their savings one-to-one and the Ricardian equivalence holds. In the SIM model, however, agents face borrowing constraints (which are binding for some of them). The Ricardian equivalence breaks down and in response to an increase in government debt aggregate savings increase by less than one-to-one. Since the asset

market must clear (i.e. $A_t = K_t + B_t$), it follows that capital must decrease as a result. Hence, increases in government debt crowd out capital while decreases crowd in capital²⁷.

In order to understand why the government accumulates assets in the optimal policy it is important to look at its effect on equilibrium prices²⁸. A lower amount of government debt leads to a higher level of capital which reduces interest rates and increases wages. Hence, besides the positive level effect associated with the higher levels of capital such a policy also affects the insurance and redistribution effects. It effectively reduces the proportion of the agents' income associated with the unequal asset income and increases the proportion associated with uncertain labor income. The result is a positive redistribution effect and a negative insurance effect. Thus, when government debt-to-output is held fixed the redistributive gains are reduced from 4.9 percent to 3.2 percent while the insurance loss is reduced from -3.7 percent to -3.2 percent. This also clarifies why the planner chooses to accumulate assets when the instrument is not restricted: the welfare gains associated with the resulting redistribution outweigh the losses from the increased uncertainty.

2.5.5 Transitory effects

In this section we first compute the optimal fiscal policy ignoring transitory welfare effects. A comparison with our benchmark results allows us to measure the importance of accounting for these transitory effects. If the difference was small this would be a validation of experiments of this kind performed in the literature. It turns out, however, that the results are remarkably different. A better option, is to solve for the optimal policy with constant instruments accounting for transitory welfare effects. The welfare loss associated with holding the instruments constant, however, is still significant. The results are summarized in Tables 2.5 and 2.6.

²⁷See Aiyagari and McGrattan (1998) and Winter and Roehrs (2014) for an extensive discussion on this issue.

²⁸The fact that the government accumulates assets does not imply that it becomes the owner of part of the capital stock. Agents own the capital, but on average owe the government (in the form of IOU contracts) more than the value of their capital holdings.

Table 2.5: Final Stationary Equilibrium: transitory effects

	τ^h	τ^k	T/Y	B/Y	K	H	r	w
Initial equilibrium	28.0	36.0	8.0	63.0	1.65	0.33	4.1	1.14
Stat. equil.	18.0	-	3.7	- 326.1	4.01	0.44	0.0	1.45
Stat. equil. fixed debt	4.7	-5.2	-5.4	63.0	2.84	0.43	1.9	1.26
Constant policy	7.6	73.7	3.5	49.8	1.31	0.36	7.1	1.01
Benchmark	12.6	45.1	3.5	-16.9	2.00	0.38	3.7	1.16

Notes: The values of τ^h , τ^k , T/Y , B/Y , and r are in percentage points.

Table 2.6: Welfare decomposition: transitory effects

	Δ	Δ_L	Δ_I	Δ_R
Stat. equil.	24.7	19.6	-4.6	9.3
Stat. equil. fixed debt	9.8	18.8	-5.2	-2.6
Constant policy	3.3	3.4	-3.0	3.0
Benchmark	4.9	3.7	-3.7	4.9

Stationary equilibrium policy

Here the planner chooses stationary levels of all four fiscal policy instruments to maximize welfare in the final steady state. In particular, the planner can choose any level of government debt without incurring in the transitional costs associated with it. It chooses a debt-to-output ratio of -326 percent. At this level the amount of capital that is crowded in is close to the golden rule level, that is, such that interest rates (net of depreciation) equal to zero. Thus, taxing capital income in this scenario has no relevant effect and we actually find multiple solutions with different levels of capital taxes which is why we do not display that number in Table 2.5. The average welfare gains associated with this policy

are of 24.7 percent, that is, agents would be willing to pay this percentage of their consumption in order to be born in the stationary equilibrium of an economy that has this policy instead of the initial stationary equilibrium. However, these welfare gains ignore the transitory effects, it is as if the economy jumped immediately to a new steady state in which the government has a large amount of assets without incurring in the costs associated with accumulating it.

A more reasonable experiment, which is closer to the one studied by [Conesa et al. \(2009a\)](#), is to restrict the level of debt-to-output ratio to remain at its initial level. When this is the case, the planner reduces labor taxes and capital taxes substantially obtaining most of the necessary revenue via lump-sum taxes. This has detrimental insurance and redistribution effects, but the associated level effect more than makes up for it. The policy leads to a welfare gain of 9.8 percent relative to the initial steady state when transitory effects are ignored. However, once transitory effects are considered, implementing this policy leads to a welfare *loss* of 6.4 percent. Hence, ignoring transitory effects can be severely misleading. Importantly, the transitory distributional effects of the policy and the costs associated with the accumulation of capital (or assets by the government) are ignored.

Transition with constant policy

Here we consider the problem of finding the *constant* optimal fiscal policy that maximizes the same welfare function we use in our benchmark experiment, in which transitory effects are accounted for. We present a comparison with the benchmark results in Figures [A.7](#) and [A.6](#). The level of capital taxes is close to average between the upper bound of 100 percent and the final capital tax in the benchmark experiment. Labor taxes are reduced from a long-run level of 12.6 percent to 7.6 percent and lump-sum transfers converge much faster to the final level of 3.5 percent. The main difference in the fiscal policy instruments is the fact that with a constant policy the government is not able to accumulate assets via higher initial capital taxes. The debt-to-output ratio remains close to the initial level²⁹. As a result of the higher long-run capital tax and relatively higher debt-to-output ratio, capital decreases by about 20 percent in the long-run whereas it *increases* by approximately the

²⁹We do not restrict debt-to-output ratio to be constant in this experiment.

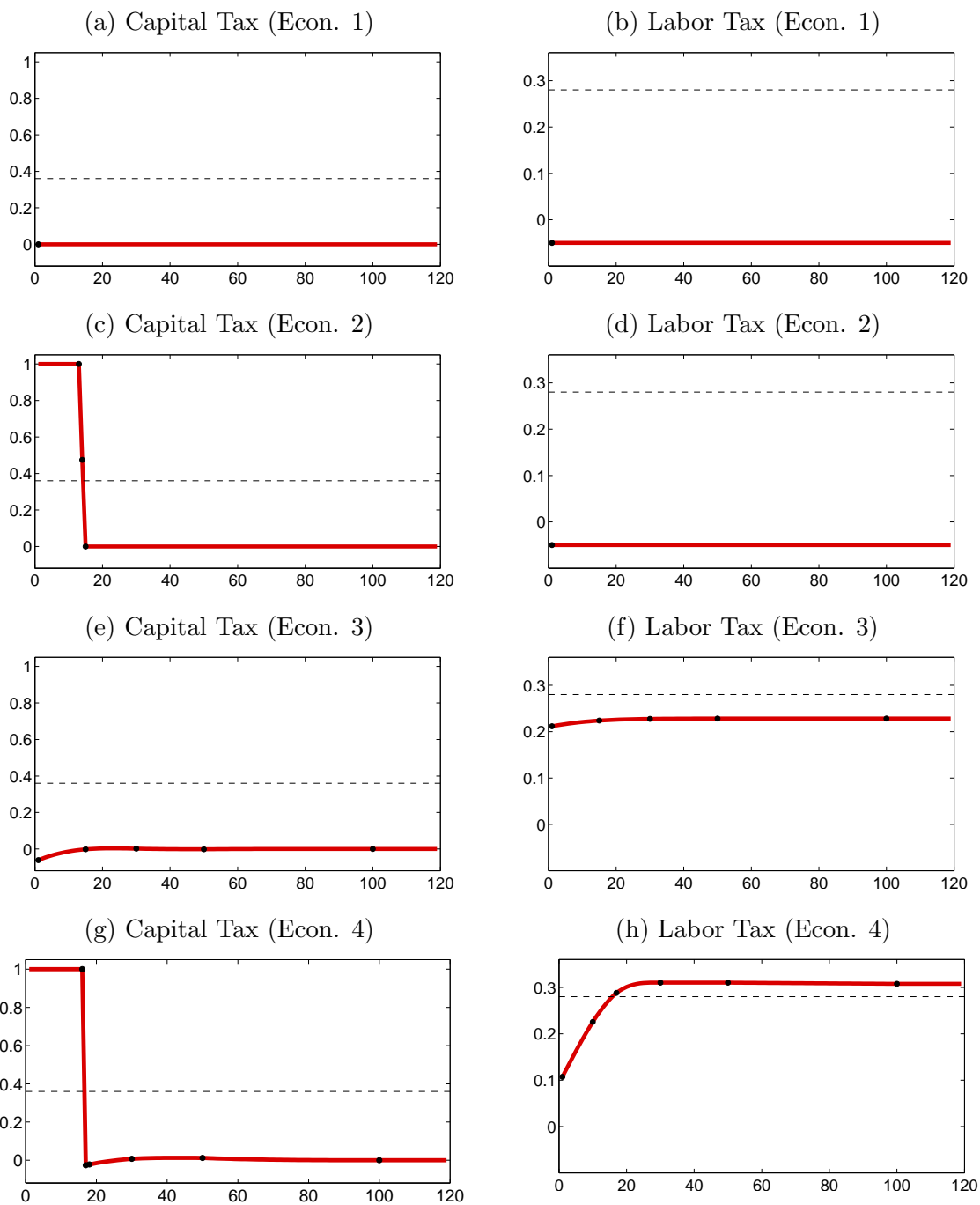
same amount in the benchmark experiment. The associated higher interest rates and lower wages lead to the reduction in the redistributive gains and reduces the insurance losses associated with the lower labor tax. This policy leads to an average welfare gain of 3.3 percent whereas the time varying policy increases welfare by 4.9 percent. That is, the restriction to constant policies leads a welfare loss of 1.6 percent.

2.6 Complete Market Economies

To our knowledge, this paper is the first to solve the Ramsey problem in the SIM environment. In order to provide further insight and relate it to other results in the literature, we provide a build up to our benchmark result. First, we start from the representative agent economy (Economy 1) and introduce heterogeneity only in initial assets (Economy 2), heterogeneity only in individual productivity levels (constant and certain) (Economy 3), and heterogeneity both in initial assets and in individual productivity levels (Economy 4). Introducing idiosyncratic productivity shocks and borrowing constraints brings us back to the SIM model. At each step, we analyze the optimal fiscal policy identifying the effect of each feature.

In what follows we examine the optimal fiscal policy in Economies 1-4. Their formal environments can be quickly described by starting from the SIM environment delineated above. Economy 4 is the SIM economy with transition matrix, Γ , set to the identity matrix, and borrowing constraints replaced by no-Ponzi conditions. Then, we obtain Economy 3 by setting initial asset levels to its average, Economy 2 by setting the productivity levels to its average, $e = 1$, and Economy 1 by equalizing both initial assets and levels of productivity. Figure 2.5 contains the numerical results.

Figure 2.5: Optimal Taxes: Complete Market Economies



Notes: Dashed line: initial taxes; Solid line: optimal taxes.

2.6.1 Economy 1: representative agent

To avoid a trivial solution, the usual Ramsey problem in the representative agent economy does not consider lump-sum transfers to be an available instrument. Since in this paper we do, the solution is, in fact, very simple. It is optimal to obtain all revenue via lump-sum taxes and set capital and labor income taxes so as not to distort any of the agent's decisions. This amounts to $\tau_t^k = 0$ and $\tau_t^n = -\tau^c$ for all $t \geq 1$. Since consumption taxes are exogenously set to a constant level, zero capital taxes leaves savings decisions undistorted and labor taxes equal to minus the consumption tax ensures labor supply decisions are not distorted as well. In this setup the Ricardian equivalence holds so that the path for lump-sum taxes and debt are indeterminate: there is no lesson to be learned from this model about the timing of lump-sum taxes or the path of government debt. This will also be the case in Economies 2, 3 and 4.

2.6.2 Economy 2: add heterogeneity in initial assets

Introducing heterogeneity in the initial level of assets we can diagnose the effect of this particular feature on the Ramsey policies by comparing it to the representative agent ones. We extend the procedure introduced by [Werning \(2007\)](#)³⁰ to characterize the optimal policies for this and the next two economies. We describe them in a proposition leaving the proof to [Appendix A.2](#).

Proposition 4. *There exists a finite integer $t^* \geq 1$ such that the optimal³¹ tax system is given by $\tau_t^k = 1$ for $1 \leq t < t^*$ and $\tau_t^k = 0$ for all $t > t^*$; and $\tau_t^n = -\tau^c$ for all $t \geq 1$.*

Once again, there is no reason to distort labor decisions since labor income is certain and the same for all agents. However, the paths for capital taxes and lump-sum transfer do differ from the representative agent ones. [Proposition 2](#) provides a rationale for taxing capital in this case; since agents have different initial asset levels, capital taxes can be used to provide redistribution. This fact together with the fact that capital taxes are zero in

³⁰[Werning \(2007\)](#) solves for separable and balance growth path utility functions. Besides solving for GHH preferences we also impose the upper bound on capital income taxes and remove the possibility of time zero taxation.

³¹All propositions in this section are valid for any set of welfare weights, not only the Utilitarian ones. The associated numerical results do assume a Utilitarian welfare function though.

the long-run determine the optimal path for capital taxes³². Capital taxes are positive and front-loaded, hitting the upper bound in the initial periods subsequently being driven to zero. The extra revenue obtained via capital taxation is redistributed via lump-sum transfers (or a reduction in lump-sum taxes relative to the representative agent level). It is important to reemphasize that since lump-sum transfers are an unrestricted instrument, there is no reason to tax capital in the initial periods other than for redistributive motives.

In order to have a sense of the magnitudes of t^* and the increase in lump-sum transfers, we apply the same procedure to the one we used to solve for the optimal tax system in the benchmark economy. All we need to do is choose the initial distribution of assets. The stationary distribution of assets in this economy is indeterminate³³, hence, we can choose any one we want. To keep the results comparable we choose the initial stationary distribution from the benchmark experiment³⁴.

2.6.3 Economy 3: add heterogeneity in productivity levels

It turns out that the Ramsey policies for this economy are a bit more complex. Let Φ , Ψ , and Ω^n be constants (defined in Appendix A.2) and define

$$\Theta_t \equiv \frac{C_t}{\Omega^n \chi \frac{\kappa}{1+\kappa} N_t^{\frac{1+\kappa}{\kappa}}} - 1.$$

The following proposition can be established.

Proposition 5. *Assuming capital taxes are bounded only by the positivity of gross interest rates, the optimal labor tax, τ_t^n , can be written as a function of Θ_t given by*

$$\tau_t^n(\Theta_t) = \frac{(1 + \tau^c) \Psi \Theta_t}{\Phi \Theta_t + \Psi(\sigma + \Theta_t)} - \tau^c, \quad \text{for } t \geq 1, \quad (2.6)$$

³²[Straub and Werning \(2014\)](#) show that capital taxes can be positive in environments similar to this. The reason why their logic does not apply here is the fact that the planner has lump-sum taxes as an available instrument. In particular, the proof of Proposition 4 does not impose convergence of any Lagrange multipliers.

³³For the preferences chosen above, consumption is linear on, and labor supply is independent of the individual asset level. It follows that the equilibrium levels of aggregates are independent of the asset distribution and equal to the representative agent ones (see [Chatterjee \(1994\)](#)). In a steady state, $\beta(1 + (1 - \tau^k)r) = 1$ and, therefore, every agent will keep its asset level constant.

³⁴In fact, a rescaling of it since the steady state aggregate level of assets is different when there is no idiosyncratic risk (since there is no precautionary savings).

with sensitivity

$$\Theta_t \frac{d\tau_t^n(\Theta_t)}{d\Theta_t} = \frac{\sigma(\tau_t^n(\Theta_t) + \tau^c)^2}{(1 + \tau^c)\Theta_t}. \quad (2.7)$$

It is optimal to set the capital-income tax rate according to

$$\frac{R_{t+1}}{R_{t+1}^*} = \frac{\tau_t^n + \tau^c}{\tau_{t+1}^n + \tau^c} \frac{1 - \tau_{t+1}^n}{1 - \tau_t^n}, \quad \text{for } t \geq 1. \quad (2.8)$$

Since labor income is unequal, there is a reason to tax it, in order to provide redistribution. Optimal labor taxes are not constant over time since they depend on Θ_t . If they were constant, however, equation (2.8) would imply $\tau_t^k = 0$ for all $t \geq 2$. Thus, capital taxes will fluctuate around zero to the extent that labor taxes vary over time. We disregard the upper bound on capital taxes, $\tau_{t+1}^k \leq 1$, because it would complicate the result even further and in a non-interesting way. It could be that the bound is violated if the variation of Θ_t between t and $t+1$ is large enough. However, as discussed below, quantitatively this is unlikely.

To obtain a numerical solution we set the productivity levels to the ones in the benchmark economy and apply the same procedure. To have a sense of the magnitude of the sensitivity of τ_t^n to Θ_t we plug the initial stationary equilibrium numbers ($\tau^n = 0.221$, $\tau^c = 0.046$, $\sigma = 2$, and $\Theta \approx 2$) into equation (2.7). This implies a sensitivity of 0.06, i.e. a 1 percent increase in Θ_t changes the tax rate by 0.06 of a percentage point, from 0.221 to 0.2209. We can then calculate the path of Θ_t , which we plot in Figure A.8. Notice that the volatility of Θ_t over time is unsubstantial. It follows that the optimal labor taxes are virtually constant and capital taxes virtually zero.

In any case, the fact that capital is taxed at all seems to be inconsistent with the logic put forward so far. It is not, when labor taxes vary over time they distort the savings decision, capital taxes are then set to “undo” this distortion. The analogous is not the case in Economy 2 because of the absence of income effects on labor supply; distortions of the savings decision do not affect the labor supply.

2.6.4 Economy 4: add heterogeneity in both

The result for this economy is a combination of the last two.

Proposition 6. *There exists a finite integer $t^* \geq 1$ such that the optimal tax system is given by $\tau_t^k = 1$ for $1 \leq t < t^*$, τ_t^k follows equation (2.8) for $t > t^*$; τ_t^n evolves according to equation (2.8) for $1 \leq t < t^*$; and τ_t^n is determined by equation (2.6) for all $t \geq t^*$.*

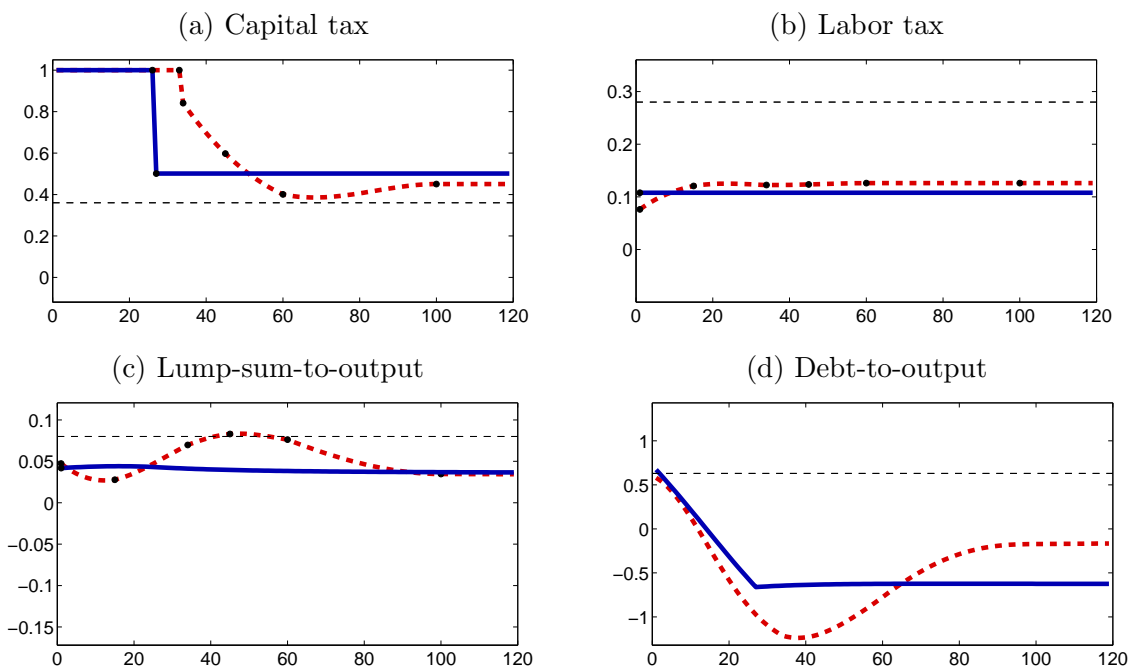
Optimal capital taxes are very similar to Economy 2 and for the same reasons. Labor taxes are determined by the same equation as in Economy 3 for $t \geq t^*$. In initial period, $1 \leq t < t^*$, while capital taxes are at the upper bound, $R_t = 1 < R_t^*$ and, therefore, equation (2.8) implies that labor taxes should be increasing. Lump-sum transfers are higher than the in Economies 2 and 3 since they are used to redistribute the capital *and* labor tax revenue.³⁵

2.7 Robustness

Figure 2.6 shows that the solution with 4 nodes ($t^*, \tau_{t^*+1}^k, \tau_1^n$, and T_1) produces a reasonable approximation for the benchmark solution, at least with respect to its basic features. In this section, we make use of this fact, and present results for alternative welfare functions and for different calibrations of the labor income process using these 4 nodes.

³⁵Bhandari et al. (2013) solve recursively for Ramsey policies in an economy similar to Economy 4 with aggregate risk.

Figure 2.6: Optimal Fiscal Policy with 4 nodes



Notes: Dashed thin line: initial stationary equilibrium; Dashed thick line: optimal transition with 17 nodes (benchmark); Solid line: optimal transition with 4 nodes.

2.7.1 Welfare function

All the results presented so far used the same social welfare function: the utilitarian one, which places equal Pareto weights on each agent. This implies a particular social preference with respect to the equality versus efficiency trade-off. Here we consider different welfare functions that rationalize different preferences about this trade-off. With this in mind we propose the following function

$$W^{\hat{\sigma}} = \left(\int \bar{x}(a_0, e_0)^{1-\hat{\sigma}} d\lambda_0 \right)^{\frac{1}{1-\hat{\sigma}}},$$

where λ_0 is the initial distribution of individual states (a_0, e_0) , \bar{x} denotes the individual certainty equivalents of labor-consumption composite (given a particular initial state (a_0, e_0)), and, following [Benabou \(2002\)](#), we call $\hat{\sigma}$ the planner's degree of inequality aversion. First notice that if $\hat{\sigma} = \sigma$ (the agents' degree of risk aversion), maximizing $W^{\hat{\sigma}}$ is equivalent to

maximizing the utilitarian welfare function³⁶. If $\hat{\sigma} = 0$, then maximizing W^0 is equivalent to maximizing $(1 + \Delta_L)(1 + \Delta_I)$, that is, the planner has no redistributive concerns and focuses instead in the reduction of distortions and the provision of insurance³⁷. Finally, as $\hat{\sigma} \rightarrow \infty$ the welfare function approaches $W^\infty = \min(\bar{x}(a_0, e_0))$. Hence, by choosing different levels for $\hat{\sigma}$ we can place different weights on the equality versus efficiency trade-off, from the extreme of completely ignoring equality ($\hat{\sigma} = 0$), passing through the utilitarian welfare function ($\hat{\sigma} = \sigma$), and in the limit reaching the Rawlsian welfare function ($\hat{\sigma} \rightarrow \infty$). Table 2.7 displays the results for different levels of $\hat{\sigma}$.

Table 2.7: Robustness: Welfare Function

	t^*	τ^k	τ^n	T/Y	B/Y	Δ	Δ_L	Δ_I	Δ_R
$\hat{\sigma} = 0$	0	34.7	12.2	0.0	79.8	0.58	5.32	-2.74	-1.80
$\hat{\sigma} = 1$	19	49.9	10.1	2.9	-36.4	4.56	3.73	-3.83	4.81
$\hat{\sigma} = 2^*$	26	49.7	10.8	3.6	-62.5	4.64	2.97	-3.84	5.68
$\hat{\sigma} = 3$	29	49.8	10.4	3.5	-76.8	4.64	2.90	-4.01	5.94
$\hat{\sigma} = 4$	30	48.9	11.5	4.1	-76.0	4.61	2.52	-3.78	6.05
$\hat{\sigma} = 5$	32	49.2	11.3	4.0	-84.2	4.59	2.45	-3.88	6.21

Notes: (*) When $\hat{\sigma} = 2 = \sigma$ the welfare function is utilitarian, this is the solution plotted in Figure 2.6. The values for T/Y and B/Y are the ones from the final steady state. For the welfare decomposition we use the utilitarian welfare function for comparability.

When $\hat{\sigma} = 0$ the planner has no redistributive motive and, accordingly, $t^* = 0$ which is consistent with the results displayed above, in particular in Section 2.6. The benchmark result that capital taxes should be held fixed at the upper bound for the initial periods is inherently linked to the redistributive motive of the planner. It follows that higher $\hat{\sigma}$ imply higher t^* 's (lower lump-sum-to-output ratios and higher debt-to-output ratios). Otherwise, overall, specially for $\hat{\sigma} \geq 1$, the results do not change significantly with changes in $\hat{\sigma}$. In particular, the final levels of capital and labor taxes are remarkably similar.

³⁶Notice that $(\int \bar{x}(a_0, e_0)^{1-\sigma} d\lambda_0)^{\frac{1}{1-\sigma}}$ is a monotonic transformation of $\int \frac{\bar{x}(a_0, e_0)}{1-\sigma}^{1-\sigma} d\lambda_0$, which is equivalent to the utilitarian welfare function.

³⁷This result can be established following a similar procedure to the one used in proof of Proposition 3. The online appendix contains the proof.

2.7.2 Labor income process

The labor income process (summarized by the Markov matrix, Γ , and the vector of productivity levels, e) is a key determinant of the amount of uncertainty and inequality faced by agents in the economy. These parameters are a discrete approximation for a continuous process for labor income, $li_t \equiv we_t n_t$, that is

$$\log(li_{t+1}) = \rho \log(li_t) + \varepsilon, \quad \text{where } \varepsilon \sim N(0, \sigma_\varepsilon^2).$$

In our benchmark calibration we target $\rho = 0.9$ and $\sigma_\varepsilon = 0.2$. Given the importance of these choices for our results and the lack of consensus in the literature about them (see Section 2.4.4 for a discussion), we provide here the results for alternative numbers for ρ and σ_ε . For each of these we recalibrate the economy modifying only the corresponding target, Table 2.8 contains the results.

Table 2.8: Robustness: Labor income process

	t^*	τ^k	τ^n	T/Y	B/Y	Δ	Δ_L	Δ_I	Δ_R
$\rho = 0.85$	24	34.8	4.8	0.0	-100.2	5.43	4.81	-3.72	4.48
$\rho = 0.95$	21	42.8	11.5	3.7	-49.5	3.91	3.63	-3.35	3.74
$\sigma_\varepsilon = 0.15$	28	28.1	4.9	0.1	-126.3	5.64	4.59	-4.09	5.31
$\sigma_\varepsilon = 0.25$	34	57.8	11.6	4.7	-75.9	4.52	2.51	-4.29	6.52
Benchmark	26	49.7	10.8	3.6	-62.5	4.64	2.97	-3.84	5.68

Notes: The values for T/Y and B/Y are the ones from the final steady state.

As one would expect, the magnitudes of the results do change considerably given changes in these important parameters. However, reassuringly, the qualitative features of the fiscal policy instruments and of where the welfare gains come from is not substantially affected.

2.8 Conclusion

In this paper we quantitatively characterize the solution to the Ramsey problem in the standard incomplete market model. We find that even though the planner has the ability to obtain all revenue via undistortive lump-sum taxes, it chooses instead to tax capital income heavily and labor income to a lesser extent. Moreover, we show that it is beneficial for the government to accumulate assets over time. With a welfare decomposition we diagnose that, relative to the current US tax system, this policy leads to an overall reduction of the distortions of agent's decisions, to a substantial amount of redistribution and to a reduction in the amount of insurance provided by the government. Importantly, we also show that disregarding the transitory dynamics and focusing only on steady states can lead to severely misleading results.

Finally, we do not view our results as a final answer to our initial question: to what extent should governments use fiscal policy instruments to provide redistribution and insurance? Instead, we understand it as a contribution to the debate. The model we use abstracts from important aspects of reality, as any useful model must, and we miss some important dimensions. For instance, in the model studied above an agent's productivity is entirely a matter of luck, it would be interesting to understand the effects of allowing for human capital accumulation. We also assume the government has the ability to fully commit to future policies, relaxing this assumption could lead to interesting insights.

Chapter 3

Endogenous Borrowing Constraints in Heterogenous Firms Economy

3.1 Introduction

Disruptions in the financial markets have been viewed to play an important role in shaping aggregate fluctuations and firm dynamics over the business cycle. In particular, a conventional belief is that small firms are more sensitive to the business cycle due to a limited access to credit markets¹. In this paper I challenge this belief. I begin with documenting a key observation that age of the firm rather than its size is a determinant of the cyclical employment volatility. I look at this observation through the lens of theory of endogenously frictional financial markets and propose a model of firm dynamics reflecting two ubiquitous features of the credit market: relevance of past performance and long-term nature of financial arrangements. I apply this theory to study, in a quantitative macroeconomic model, an impact of the aggregate shocks to microeconomic uncertainty on macroeconomic aggregates and on employment dynamics across various groups of firms.

¹One of the numerous examples of such belief is the speech of Ben Bernanke at the Federal Reserve Meeting Series: "[Addressing the Financing Needs of Small Businesses](#)".

My main empirical finding is that age of the firm rather than its size is a determinant of the asymmetric response of firms to changes in the aggregate economic conditions. The standard deviation of the cyclical component of employment of young firms is 2.6 times larger relative to the standard deviation of the cyclical component of employment of the old firms, whereas the small and large firms differ with this regards at most by 30 percent. Groups of firms that are similar along certain dimensions (share in total employment, average number of employees) but differ in age exhibit entirely different employment dynamics at the business cycle frequency. I argue that these differences are not driven by the entry of new firms to the group of young ones. Relative volatilities remain almost unchanged after I remove start-ups from the set of young firms. I further validate my main finding by looking at employment dynamics inside different size groups of firms. I found that, conditional on size, cyclical employment volatility declines with age. Whereas controlling for age, the cyclical employment volatility increases with size. Thus, the least volatile group is small, old firms, which constitute a sizeable fraction of the US businesses measured both in terms of number of firms and share of total employment. The existence of this group challenges the conventional belief about small firms being particularly sensitive to the aggregate conditions. I document using the most recent observations from the Business Dynamics Statistics data that the asymmetry of employment dynamics between young and old firms was particularly strong in the 2007-2009 recession and following recovery. Young firms reduced their employment stock by 24.2 percent between the beginning of recession and the last observation in 2012, accounting largely for jobless recovery. This is in a sharp contrast with the old firms which by 2012 fully recovered to their pre-recession employment levels.

Motivated by these facts I develop a theory of economic downturns and asymmetric cyclical employment dynamics across firms. I propose a general equilibrium model of firm dynamics with endogenously frictional financial markets. Key ingredients of my theory are existence of private information on the side of the firm and an efficient, long-term, lending arrangement between an individual firm and financial intermediary. Financial friction originates from these two ingredients and manifests itself as an endogenous borrowing constraint. A key contribution of my theory is an endogenously generated link between

firm's age and size and its ability to obtain financing. Each firm is run by an entrepreneur who seeks to maximize the stream of profits from the investment project. Every period a return on a project is subject to privately observed idiosyncratic shock, which I interpret as demand shock as it occurs after the production takes place. Moreover, firms differ in terms of average demand and therefore in the expected return on the project. The firm's operation is financed through the loan from the financial intermediaries. The combination of the history of idiosyncratic shocks, which can be interpreted in the model as proxy for performance, together with the differences in the mean returns on projects lead to rich age/size distribution of firms in equilibrium. To begin I show that, for any given expected return, an optimal contract imposes an endogenous borrowing constraint on a new firm, i.e. it is unable to obtain a level of financing it would achieve under full information. I provide sufficient conditions under which as firm ages the incentive problem vanishes and firm moves towards unconstrained level of financing. Then, I characterize the structure of consumptions and payments, both conditional on the realizations of the idiosyncratic shocks, over the firm's life cycle. Further, I show that an increasing relationship between firm average demand and access to financing which holds in the full information economy is preserved in the environment with private information. As a result firms with larger optimal scale of operation (size) are able to borrow more relative to firms with lower average demand even though they still may be constrained relative to their own efficient financing level. Finally, I show the existence of a stationary distribution of firms in my environment and further the existence of a recursive, stationary equilibrium.

In the quantitative part of the paper I exploit the dependence of the access to financing on firm's size and age to investigate the effects of the aggregate fluctuations in microeconomic uncertainty for individual contract policy, distribution of firms and hence macroeconomic aggregates by studying transitional dynamics of the economy. My findings in this part are twofold: (i) an increase of microeconomic uncertainty triggers recession even though contracts are complete (ii) a recession is characterized by an asymmetric response of employment across different groups of firms. I find that for a realistically calibrated economy an unanticipated increase in microeconomic uncertainty, disciplined by the data on cross-sectional distribution of firm level TFP over the last four recessions, reduces

aggregate output by 0.71 percent and aggregate employment by 0.61 percent, causing a significant recession. Moreover, an economic downturn in my model, in line with the data, is characterized by a fall of credit to GDP ratio, drop of investment and labor productivity. Furthermore, employment stock of young firms falls 4.1 times more relative to the employment of the old accounting for 51 percent of the average difference in the pre-2007 recessions. At the same time small firms reduce employment by 23 percent less relative to the large ones, in line with the data. These results are in a stark contrast with two natural benchmarks I consider: aggregate shock to microeconomic uncertainty operating in the economy with the full information and aggregate productivity shock as a source of fluctuations. In a frictionless economy an aggregate shock to micro uncertainty has absolutely no effects, i.e. an economy remains in the initial equilibrium. The reason is that with full information only the expected return on the firm's project matters for the lending, consumptions and payments and there are no incentive considerations. Therefore efficient level of lending can be sustained every period in line with perfect insurance in terms of consumption. Since the aggregate shock to micro uncertainty is mean preserving it has no effects on allocations and prices. If instead productivity shock drives the fluctuations the economy falls into recession regardless of the presence of the informational friction. However, the economic downturn in this case is characterized by a symmetric fall of employment, investment and output across all firms and constant credit to GDP ratio, counter to the data. It is due to the homogeneity result I establish in a theoretical part. Since the aggregate productivity shock implies a symmetric reduction of the expected return to the project and the contract policy functions are monotonous with respect to the return, the shock affects all firms the same way regardless of age and size. Therefore my quantitative results indicate it is a combination of fluctuations in microeconomic uncertainty and private information that are crucial to account *jointly* for a decline of main macroeconomic aggregates and asymmetric response of employment across various groups of firms observed in the US recessions.

To shed more light on the economics of my model consider first an individual contracting problem between entrepreneur (firm) and financial intermediary. Entrepreneur, who has access to a decreasing returns to scale technology, is risk averse and lender (financial

intermediary) is risk neutral. In the initial period of operation firm draws a type determining its average demand and hence an average return on the project. Every period demand is subject to idiosyncratic shock, that is privately observed by the entrepreneur. Before the shock is realized production inputs need to be paid and production takes place. Since entrepreneur (firm) has no wealth at the beginning of the operation it enters into a mutually beneficial, efficient, long-term lending relationship with the financial intermediary that allows for financing the production process every period.

Absent informational friction, within considered environment, the lender would completely insure the borrower and would provide a statically efficient level of financing in every period, which equalizes marginal benefit from additional investment with its marginal cost. Entrepreneur would receive, independent of the realization of the demand shock, a constant stream of consumption. Moreover, the realization of the shock would have no effect on the continuation of the contract and every firm would start operating at its own optimal scale. Private information in the financial market paired with an efficient, dynamic contract between entrepreneur and lender introduces a tradeoff between production efficiency, providing insurance and maintaining proper intertemporal incentives. Financial intermediary no longer provides efficient level of financing to the firm, thus informational asymmetry generates financial friction which manifests itself as an endogenous borrowing constraint. After demand is realized and observed by entrepreneur, an efficient arrangement imposes revealing true realizations through the combination of payments to the financial intermediary and continuation utilities that are contingent on the realization of the idiosyncratic shock. Following low demand realization financial intermediary requires low repayment, but also delivers low continuation value. After high realization of demand shock intermediary requires high repayment, but also delivers high continuation value for the entrepreneur. This way financial contract provides some insurance against idiosyncratic risk, albeit imperfect. Such patterns of lending, payments and continuation values together with equilibrium interest rate level induce firm is growing with age towards its optimal size determined by initially drawn type and production technology. Conditional on receiving a long enough sequence of high demand realizations, which can be thought of a proxy for good performance, the endogenous borrowing constraint relaxes with age and

firm has more access to borrowing.

A combination of lending dynamics over the firm's lifetime and initial type of the firm leads to a non-degenerate age/size distribution of firms in equilibrium of my model. A key element disciplining my quantitative exercise is to match the data counterpart of this distribution. I make sure in my model in line with the data most of the firms are small but not necessarily young. Moreover, I target employment shares among age and size groups. Matching joint age/size distribution of firms and employment is crucial for the analysis of the effects of uncertainty shocks. As idiosyncratic uncertainty increases dispersion of demand realizations rises, which induces larger incentives for the entrepreneur to misreport his type and consume the additional output. In order to separate the types (provide incentives) and prevent misreporting the continuation utilities need to be spread away more, which is costly for the financial intermediary. Thus, facing tradeoffs introduced by private information intermediary balances this cost with reduction of insurance and more importantly with tightening of the endogenous borrowing constraint. In a calibrated version of the model it is the young firms, irrespective of their initially drawn type, that are constrained in borrowing, so they reduce their demand for labor and capital the most as constraint gets tighter. Old firms, again regardless of size, are on average less constrained in borrowing, in particular there exists a fraction of them which already achieved their optimal size. Thus, due to this composition effect young firms as a group reduce demand for labor inputs more relative to the old ones, regardless of the size determined by the initially drawn type, over the deterministic transition following the uncertainty shock. Moreover, in my quantitative experiment more firms which drew a low average demand (small firms) are unconstrained relative to the group with high average demand. As a result employment of small firms in my economy is less responsive relative to the employment of the large ones, in line with the data. Reduction of demand for labor input leads to a decline in wage rate. This downward pressure tends to raise capital and labor demand, as well as output of unconstrained firms. This group consists almost entirely of old firms, irrespective of their size. Thus general equilibrium effect counters the initial effect of the uncertainty shock. In my quantitative exercise the initial impulse dominates the decline in wage rate and economy falls into a recession. Economic downturn in my model resembles

actual recessions observed in the US data. Output, employment and investment falls. More importantly credit to GDP declines. Finally, an economy in recession exhibits asymmetric employment patterns across various groups of firms, in particular young firms are more volatile than old ones and small vary less than large ones.

Related literature

This paper contributes to several strands of literature. First, it corresponds to the empirical literature on firm dynamics over the business cycle. Thus far most of this literature has focused on the role of firm size and the cycle ([Gertler and Gilchrist \(1994\)](#), [Chari, Christiano, and Kehoe \(2008\)](#), [Moscarini and Postel-Vinay \(2012\)](#)). The main conclusion from the existing studies is that large firms are more responsive to the NBER recessions, whereas small firms tend to respond more to credit market tightening. Recently, [Fort, Haltiwanger, Jarmin, and Miranda \(2013\)](#) explored the role of local housing market and aggregate financial conditions for the dynamics of the net growth rate in the business cycle context. They find innovation to the state-specific cyclical indicator associated with a downturn (e.g., a rise in the state unemployment rate) reduces the differential in the net job creation rate between young/small and large/mature businesses and that the effect persists for a number of years. In other words net growth rate of young/small businesses falls more in contractions than does the net growth rate of large/mature businesses. They interpret this as evidence that young/small businesses are more vulnerable to business cycle shocks. Also they find that a decline in housing prices in the state yields a further reduction in the differential in the net job creation rate between young/small and large/mature businesses. The approach in this paper is somehow different. Firstly, my focus is on the employment stock, rather than the growth rates, of different group of firms. I extract cyclical components of the employment time series among different groups and report their properties. Unlike [Fort et al. \(2013\)](#), I compare the cyclical dynamics among groups of firms that share certain characteristics (share in total employment, average number of employees) but differ in age and argue the latter is a key determinant of asymmetric patterns of employment. Secondly, I document entry margin does not contribute to the observed differences, the fact that is not highlighted by [Fort et al. \(2013\)](#). Finally, I provide a decomposition quantifying the contribution of the extensive margin to the employment

volatility of size and age groups. I view my analysis and findings as complementary to those by [Fort et al. \(2013\)](#).

Secondly, this paper contributes to the large literature on financial frictions. They are viewed to play central role in a propagation of aggregate fluctuations and they have been extensively explored in the economic literature ([Bernanke and Gertler \(1989\)](#), [Carlstrom and Fuerst \(1997\)](#), [Bernanke, Gertler, and Gilchrist \(1999\)](#), [Kocherlakota \(2000\)](#) and also [Quadrini \(2011\)](#), [Brunnermeier, Eisenbach, and Sannikov \(2012\)](#) for recent surveys). A common assumption in this literature is that markets are exogenously incomplete and firms utilize one period debt contracts to overcome the incompleteness. When borrowers and lenders are able to form long-term lending relationship, which are contingent on all public information (complete contracts), the macroeconomic consequences might be quite different. Though, the role of long-term financial contracts in shaping dynamics of macroeconomic aggregates remains largely unexplored² and this paper provides some new insights to this issue. I propose an environment in which financial friction originates from the presence of the private information and efficient, long term financial contract between a firm and a financial intermediary and manifests itself as an endogenous borrowing constraint. Moreover, in a quantitative literature on financial frictions the severity of the distortion is determined by exogenous shocks (for example shocks to the value of the collateral like in [Zetlin-Jones and Shourideh \(2012\)](#) or [Buera and Shin \(2013\)](#)). I develop a model in which credit frictions fluctuate endogenously and I link these fluctuations to changes in microeconomic uncertainty, which are disciplined by the data on cross sectional distribution of firm level TFP and sales.

The latter links my paper to the strand of the literature on the fluctuations in idiosyncratic uncertainty. The main conclusion from this literature is that at the micro level, the recessions have been accompanied by large increases in the cross-section dispersion of TFP and sales ([Bachmann and Bayer \(2009\)](#), [Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry](#)

²Notable exceptions are [Cooley, Marimon, and Quadrini \(2004\)](#) and [Verani \(2013\)](#) who study the role of limited commitment and private information respectively in a general equilibrium, business cycle models subject to technology shock.

(2014), Bloom (2014)). Stock and Watson (2012) document role of the heightened uncertainty was particularly large in the recent recession. Bloom et al. (2014), motivated by these regularities, build a general equilibrium model with heterogeneous firms where fluctuations of micro uncertainty are the source of the aggregate shocks. In their model the micro uncertainty and real economic activity nexus operates through the existence of the adjustment costs preventing firms from actions in times of heightened uncertainty. Alternatively, uncertainty can also increase the probability of default, by expanding the size of the left-tail default outcomes, raising the default premium and the aggregate deadweight cost of bankruptcy. This role of uncertainty in raising borrowing costs can reduce micro and macro growth, as emphasized in papers on the impact of uncertainty in the presence of financial constraints (Arellano, Bai, and Kehoe (2012), Christiano, Motto, and Rostagno (2014), Gilchrist, Sim, and Zakrajsek (2014)). All these papers though hinge on the incomplete markets assumption and one period debt contracts. My main contribution relative to this strand of the literature is to propose and quantify a novel mechanism where uncertainty shocks, within a complete contract environment, are endogenously translated into movements in the borrowing constraints and further cause real effects. Moreover, my model generates asymmetric response to aggregate shocks across firms of different size and age, which this literature is silent about.

Finally this paper contributes to the dynamic contracting literature with private information. It relates to two branches of existing literature. The first branch studies the optimal consumption insurance among risk-averse agents when individual endowments or efforts are unobservable (Green (1987), Thomas and Worrall (1990), Atkeson and Lucas (1992), Atkeson and Lucas (1995)). Smith and Wang (2006) embody this insurance problem into a stationary recursive equilibrium and find in a comparative statics exercise that changes in microeconomic uncertainty have negligible effects on aggregates. The second branch assumes risk-neutral agents and studies the optimal investment schedule maximizing the resources generated by the firm (Quadrini (2004), Clementi and Hopenhayn (2006), DeMarzo and Fishman (2007)). The general equilibrium version of this environment subject to technology shocks is studied by Verani (2013). The current paper combines the

main features of these two branches and the contract solves the trade-off between the optimal consumption insurance - as entrepreneurs are risk averse - and the optimal investment schedule - as resources depend on investment. It embodies the contracting problem into a general equilibrium framework with heterogeneous firms and studies an impact of aggregate shocks to microeconomic uncertainty.

The rest of the paper is organized as follows. Section 3.2 presents facts on firm dynamics over the business cycle. Next, in Section 3.3 I present a dynamic model of firms with asymmetric response to uncertainty shocks. Further, in Section 3.4 I provide a theoretical results characterizing my environment, in particular an access to exogenous financing as a function of size and age. Then, in Section 3.5 I present calibration and quantitative results of my paper. Finally, Section 6 concludes.

3.2 Business cycle facts on firm dynamics

In this section I document facts about the firm dynamics over the business cycle among different group of firms. I provide evidence to support the following facts:

1. Standard deviation of employment of the young firms is 2.6 larger than the standard deviation of employment of the old firms at the business cycle frequency. The least volatile group of firms are small and old ones.
2. Movements in the number of firms (extensive margin) account for 34% of the aggregate employment variance. This contribution varies between different groups of firms: it is positive for young and negative for old.
3. Since 2007 employment of the young firms went down by 24.5% and in 2012 it was at the historically lowest level.

3.2.1 Data

Data source. The primary data source I use is the Business Dynamics Statistics (BDS). It contains data on employment and job flows of establishments and firms by their characteristics for practically total non-agricultural sector of the US economy. I use the data on a large cross section of firms from 1982 to 2012, to construct the detrended series. I argue BDS is a useful and reliable source of information, compared to the most commonly used like Current Population Survey and Establishment Survey, about the cyclical movements of the aggregate employment in the US and as such can be used to quantify the contributions of different groups of firms and margins into the aggregate employment fluctuations (Appendix B.1.1). While working with the BDS data some complications arise. Firstly, while the benefit of the BDS is its large coverage its the main drawback is that it is a cross-sectional data. Thus, one can only track the employment and job flows of the particular group of firms, without knowing which firms are growing and which are contracting. To observe the latter one would need panel data. Secondly, there have been significant low-frequency secular trends in the number of firms, age and size structure over the period that the data covers. To remove the systematic changes I detrend the data using a Hodrick-Prescott filter.

Definitions. To document business cycle regularities among different groups of firms I consider three ways of partitioning the total population of the firms in the BDS. First, I consider young vs. old firms division, where I define young firms to be five years old and less and old firms to be six years and older. Secondly, I consider the small vs. large firms partition, where I define small firms to have less than 20 employees and large firms to have 20 and more employees. The third way of partitioning the sample is to divide it into small firms with less than 100 employees and large firms which have 100 and more employees. I chose the employment cutoff in the second and third partition to create groups that are comparable with the young firms with regards to the share in the total employment and the average number of employees. In the following sections, I use these three definitions to illustrate my main empirical findings.

3.2.2 Age rather than size matters.

I start by examining the differences in volatility of the employment stock between certain groups of firms at the business cycle frequency. Table 3.1 summarizes my findings. Young firms in the US economy account for 16.0% of the aggregate employment and hire on average 8.1 employees. Old firms account for 84.0% of the aggregate employment and hire on average 31.6 employees. The first group of small firms, with less than 20 employees, accounts for 19.7% of total employment, a number comparable to the share of the young firms, and the average number of employees is 4.9 in this group. The second group, firms with less than 100 employees, accounts for 37.7% of total employment and the average number of employees is 8.4 in this group, which is a similar figure to the average for young firms. Third column of the table reports the standard deviation of the logged, HP-filtered time series of employment time series for all three divisions of the BDS sample. It illustrates the main point of this section, i.e. age rather than size is a determinant of the asymmetric response of employment across different groups of firms to the cyclical changes in the aggregate economic conditions. The standard deviation of employment of young firms is 2.6 times larger relative to the standard deviation of the old firms. It is also 3.2 times larger relative to the small firms with less than 20 employees (2.5 times larger for the second definition of the small firms). The groups of firms that are similar along certain dimension (share in total employment, average number of employees) exhibit entirely different employment dynamics at the business cycle frequency. These differences are not driven by the entry of new firms to the group of young ones. Relative standard deviations remain almost unchanged after I restrict the definition of young firms to those at age between 1 and 5 years.

The fact that age is a major determinant of employment volatility can be further validated by looking at the employment dynamics inside the size groups of firms. Table 3.2 presents volatility of employment across subgroups in the age/size distribution of the population of firms³. Left panel documents standard deviations of employment for the size threshold of 20 employees, whereas the right panel for the size threshold of 100 employees. Regardless of the size threshold employment volatility declines with age and increases

³In the Appendix B.1.2 I provide companion tables documenting firms distribution, employment shares and volatilities after restricting definition of young firms to those between 1 and 5 years old for presented age/size categories.

Table 3.1: Employment volatility for different groups of firms.

	Share of total employment	Average number of employees	Standard deviation of employment
All firms	100	21.8	1.47
Young (0-5)	16.0	8.1	3.20
Old (5+)	84.0	31.6	1.25
Young, no entry (1-5)	13.0	8.8	3.32
Small (0-19)	19.7	4.9	1.04
Large (20+)	80.3	149.2	1.64
Small (0-99)	37.7	8.4	1.31
Large (100+)	62.3	695.4	1.67

Notes: Employment series are logged and HP filtered with parameter $\lambda = 6.25$. Share of total employment and average number of employees are average values. Annual data, 1982-2012. Source: Business Dynamics Statistics (BDS).

with size. Put differently even inside size groups young firms are more volatile than old ones in terms of employment. This robust finding strengthens the main message of the empirical section of my paper: age of the firm is the relevant margin if one seeks for the volatility differences. Table 3.2 also documents an existence of a group of small old firms that are the least volatile of all groups. For the size threshold of 20 employees this groups accounts for 49.9% of all firms (for 100 employees threshold the number is 56%) and for 11.9% of the total employment (25.1% for 100 employees threshold). The existence of such groups of firms challenges the conventional view that small firms are those who reduce their employment the most in recessions and highlights the key role of firm's age for response to changes in the aggregate economic conditions.

Table 3.1 and Table 3.2 document differences between firms of various size and age in terms of employment volatility. They are however silent on the comovement over the cycle of these groups and GDP. Figure 3.1 presents the time series of logged, filtered employment used to compute the standard deviations in Table 3.1. Apart from illustrating the main point about the role of age rather than size in determining the asymmetric cyclical behavior it reveals two additional features of the data. Firstly, the employment time series for all groups of firms are positively correlated with each other with the correlation coefficient ranging

Table 3.2: Standard deviation of employment over age and size distribution.

	Small (0-19)	Large (20+)	All sizes		Small (0-99)	Large (100+)	All sizes
Young (0-5)	1.96	4.93	3.20	Young (0-5)	2.37	7.66	3.20
Old (6+)	0.85	1.40	1.25	Old (6+)	1.02	1.50	1.25
All ages	1.04	1.64	1.47	All ages	1.31	1.67	1.47

Notes: Employment series are logged and HP filtered with parameter $\lambda = 6.25$. Source: Own calculations. Business Dynamics Statistics, 1982-2012.

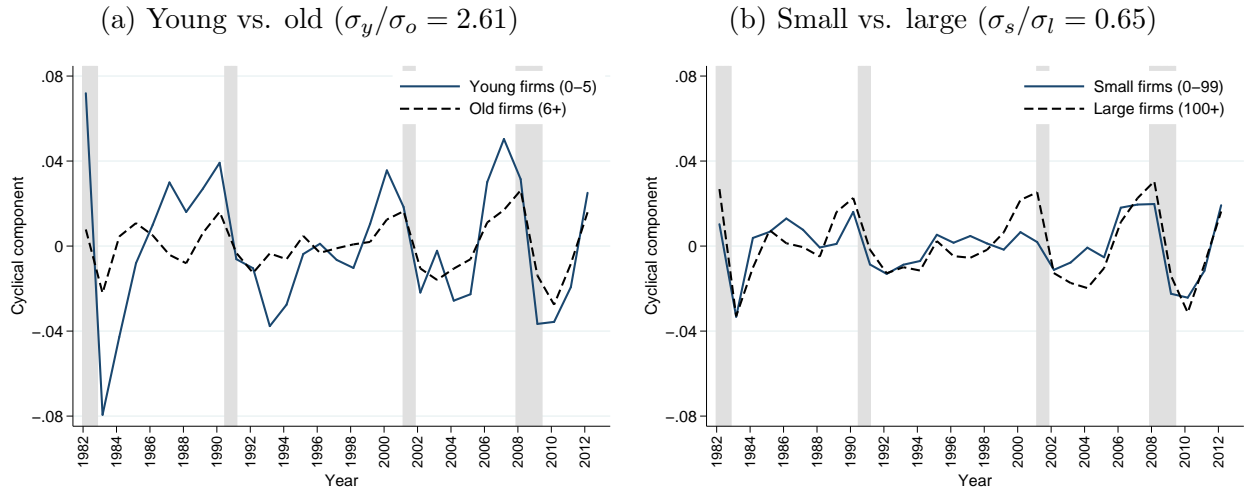
between 0.71 and 0.99. Secondly, all groups exhibit positive contemporaneous correlation with the cyclical component of GDP with the correlation coefficient varying between 0.45 and 0.67. I report all contemporaneous correlations in Table B.4 in the Appendix B.1.3. In fact, also the phase shifts of employment and GDP for different groups of firms look alike, which I document in Figure B.2 in Appendix B.1.3. The bottom line is that the magnitude of the response to the changes in the aggregate conditions is a factor differentiating the young from the old firms (also the young from small), while the timing of the response is similar across different groups of firms.

3.2.3 The role of the extensive margin

Large differences between various groups of firms are driven by the movements in employment per firm (intensive margin) as well as by the movements in the number of firms by itself (extensive margin). In this section I document that contribution of each margin differs largely across age groups but is almost identical for size groups. One way to quantify the role of extensive margin for the changes in employment over the business cycle is to consider the following simple decomposition. Let employment within a group of firms j be denoted by E_j and the number of firms by F_j . Then the following identity holds

$$E_j = \frac{E_j}{F_j} \times F_j$$

Figure 3.1: Cyclical component of employment for different group of firms.



Notes: Shaded areas are NBER recessions. Employment series are logged and HP filtered with parameter $\lambda = 6.25$. Source: Own calculations. Business Dynamics Statistics, 1982-2012

This decomposition says one can write the employment as a product of employment per firm and the number of firms within group j . Then taking logs and variances yields

$$V(\log(E_j)) = V\left(\log\left(\frac{E_j}{F_j}\right)\right) + V(\log(F_j)) + 2Cov\left(\log\left(\frac{E_j}{F_j}\right), \log(F_j)\right)$$

Table 3.3 reports the result of this decomposition across different groups of firms. Movements in the extensive margin account for 34.1% of total employment variance (all firms) at the business cycle frequency. This contribution however differs significantly as one looks at the first partition of the sample i.e. young vs. old firms. An extensive margin accounts for 55.7% of the young firms employment variance. Thus, more than half of the cyclical employment movements in this group of firms is due to the entry and exit. This is in stark contrast with the group of old firms, where the extensive margin contributes negatively to the cyclical movements of employment and it dampens its variance. The movements in the extensive margin for the old firms reduces by 28.7% the variance of the employment per firm⁴. Such difference suggests that the role extensive margin is another factor that

⁴It is important to highlight this decomposition does not attribute life time employment patterns to the

distinguishes young and old firms. As I report in Table 3.3 this is not the case for the small vs. large firms partition. The role of extensive margin is virtually the same for both groups. The movements in the number of firms and covariance term account for more than half of the cyclical employment movements of these two groups of firms.

Table 3.3: Decomposition of employment variance for different group of firms

	All firms	Young (0-5)	Old (5+)	Small (0-19)	Large (20+)
$V\left(\log\left(\frac{E_j}{F_j}\right)\right)$	65.9	44.3	128.8	45.6	42.3
	Contribution of the extensive margin				
	34.1	55.7	-28.7	54.3	57.6
$V(\log(F_j))$	27.8	28.6	49.5	55.9	68.2
$2Cov\left(\log\left(\frac{E_j}{F_j}\right), \log(F_j)\right)$	6.3	27.1	-78.5	-1.6	-10.6

Notes: Series are logged and HP filtered with parameter $\lambda = 6.25$. Annual data, 1982-2012. Source: Own calculations. Business Dynamics Statistics (BDS).

3.2.4 The Great Recession and it's aftermath

The 2007-2009 recession was extraordinary, relative to previous recessions, not only in terms of depth and length when measured with the standard macroeconomic aggregates, but also in terms of the response across different group of firms. Two empirical findings that differentiate the recent downturn and following recovery from the previous business cycle episodes are: (i) between 2007 and 2012 young firms reduced employment by 24.2 percent, whereas in previous episodes young firms recovered to their initial employment level after 5 years (ii) in pre-2007 downturns large firms reduced their employment more relative to the small ones. During the last recession this pattern was reversed - small firms reduced employment more. Figure 3.2 illustrates these facts.

variance of employment. Since the data set is a repeated cross section of the firms thus employment time series tracks employment over time within a particular group with the use of the same definition over the whole time period. Therefore the variance of employment can be thought of as the cyclical movements of the employment life cycle patterns.

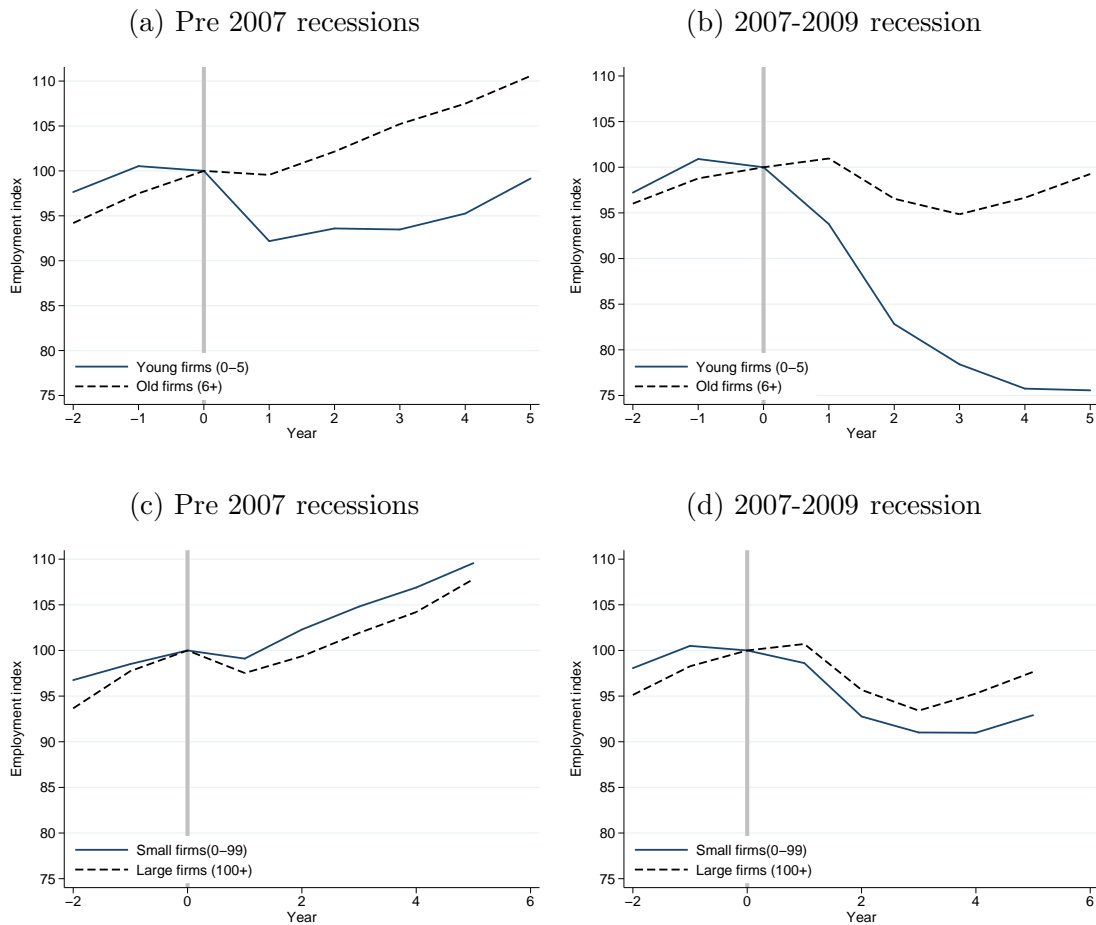
The asymmetric employment patterns of young and old firms between 2007 and 2012 consists of two phases. Firstly, the drop in employment of young firms during the actual recession (2007-2009) time was larger than in any pre-2007 recessions, which reflected the depth of the recent downturn. Secondly, and more importantly, since 2009 young firms did not recover and reduced the employment even more, so that the cumulative employment fall amounts to 24.5 percent by 2012. As a result in March 2012 the employment level of the young firms was on the historically lowest level since 1982, the first data point in the BDS for which the employment of young firms can be computed. Table 3.4 sheds more light on the sources of this un-precedential fall. It reports the number of young firms, entering firms (firms of age 0) and old firms in 2007, 2011 and pre-2007 period. Note, that the sum of young and old firms amounts to the total number of firms in the BDS. The number of old firms is growing since the sample of firms in the BDS is getting older, i.e. there is a secular trend due to change in the demographic structure of the data set. The number of young firm however is not subject to this issue and it reflects more accurately the role of extensive margin in the Great Recession. In 2011 the number of young firms in the US economy was the lowest since 1982. To put the size of changes into perspective, note the US economy lost 389.1 thousands of young firms between 2007 and 2011, which amounts to 18.5% of their number before the Great Recession. To large extent this change was driven by the historically small number of new businesses. In 2010 and 2011, the inflows of newly born firm to the US economy were two lowest since 1982.

Table 3.4: Number of firms in pre- and post-2007 period.

	Young firms	Entering firms	Old firms
Lowest level before 2007	1,748.5	411.8	1,801.1
2007	2,109.5	529.2	3,189.0
2011	1,720.4	409.0	3,260.3

Source: Own calculations. Business Dynamic Statistics (BDS).

Figure 3.2: Employment index for various groups of firms in pre-2007 and 2007-2009 recessions.



Notes: Employment normalized at the year of NBER recession start to 100. Pre-2007 recessions are the downturns starting in 1981, 1990, and 2001. Source: Business Dynamic Statistics (BDS).

3.3 Model of firm dynamics with asymmetric response to uncertainty shocks

In this section I develop a dynamic model with heterogeneous firms and define a recursive stationary equilibrium. Each firm is run by an entrepreneur. Firms have private information about their demand and enter into a long-term lending relationship in order to

finance their operation. These two features generate endogenously a borrowing constraint that is binding for a fraction of firms in the economy and are crucial to account for asymmetric employment patterns over business cycles documented in the data. One can view my model as incorporating *jointly* features of model of a firm dynamics with variable investment (Clementi and Hopenhayn (2006), DeMarzo and Fishman (2007)) and dynamic insurance problem (Thomas and Worrall (1990), Atkeson and Lucas (1992)) under private information into a general equilibrium framework suited for a quantitative work.

3.3.1 Environment

Time is discrete, lasts forever and is indexed by $t = 0, 1, \dots$. There are three types of agents in the economy: a large number of homogenous workers, a large number of firms (entrepreneurs), and a large number of financial intermediaries (lenders). There is a single consumption good in the economy. Each firm draws a permanent type s from the finite set of types S in the first period of operation, which determines the average demand of the firm. The source of idiosyncratic uncertainty is a shock to the demand of the firm, $\theta_s \in \Theta_S$. Let θ_{st} be the realization of this shock at any time t for the firm of type s and denote the individual history by $\theta_s^t = (\theta_{sj}, \theta_{sj+1}, \dots, \theta_{st})$ of the firm starting to operate at time $j \leq t$.

Timing. The timing of the events within a period can be summarized as follows:

1. New firms are born and draw type s .
2. Financial intermediaries lend resources to the firms.
3. Production inputs are hired and production takes place.
4. Idiosyncratic shock θ_{st} is realized.
5. Consumption and payments take place. Firms exit exogenously.

Information. Financial intermediaries observe the amount of lending to the firms and payments received from the firms. Moreover, the initially drawn type s is public information. Intermediaries can not observe the realization of the idiosyncratic shock θ_{st} and entrepreneur's consumption.

3.3.2 Firms

Preferences. I assume the ownership of the firms is concentrated, i.e. a firm is associated with an entrepreneur. Each firm faces a time-invariant probability $\zeta < 1$ of surviving into the next period. The total measure of firms in the economy is equal to one. Newly born firm draws a type $s \in S$ according to the probability distribution Γ , that assigns a probability Γ_s to each type. Type determines an expected demand and hence expected revenue from the project. Firm of type s that starts operating in period j , values a stochastic sequence of consumption good $\{c(\theta_s^t)\}_{t=j}^{\infty}$ through the lens of the entrepreneur's preferences, i.e.

$$\sum_{t=j}^{\infty} \sum_{\theta_s^t} (\beta\zeta)^{t-j} \Pr(\theta_s^t) U(c(\theta_s^t))$$

where the period utility function U is strictly increasing, strictly concave, and satisfies standards conditions, $\Pr(\theta_s^t)$ is a probability of a particular history θ_s^t and $\beta < 1$ is a discount factor.

Technology. Each firm of type s operates a long-lived project that produces output in each period of life of the firm. Every period project requires capital input, k , and labor input, n , which need to be purchased in advance, i.e. before the production is sold and the idiosyncratic shock is realized. Each entrepreneur is born without wealth. Therefore every period in order to finance the project he must borrow resources l in the credit market to operate a project. Firm has an access to a decreasing returns to scale production technology $f : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ which satisfies standard conditions listed in Assumption ... Denote γ to be a degree of returns to scale. Technology transforms the capital and labor inputs into a consumption good. Produced output is subject to a idiosyncratic demand shock $\theta_{st} \in \Theta_s = \{\theta_{s1}, \dots, \theta_{sN}\}$ with $N < \infty$, with the fixed probability distribution Π_s that assigns positive probability $\pi(\theta_s)$ to all θ_{st} values and can potentially depend on initially drawn type $s \in S$. Without loss of generality, let $\theta_m < \theta_n$ if $m < n$. Demand shock θ is i.i.d. over time. Thus every period the following feasibility constraint has to hold

$$c(\theta_s^t) + m(\theta_s^t) \leq \theta_{st}^{1-\gamma} f(k(\theta_s^{t-1}), n(\theta_s^{t-1})) \quad \forall \theta_s^t, \theta_s^{t-1} \quad (3.1)$$

where m denotes the repayment from the firm to the financial intermediary. From 3.1 it becomes clear that given that financial intermediary can not observe consumption and demand shock it is unable to figure out the actual realization of the demand. Denote w_t to be the wage rate, r_t to be the interest rate and δ be a depreciation of capital. Firm operates under the constraint that both wage bill $w_t n(\theta_s^{t-1})$ and cost of renting capital $(r_t + \delta) k(\theta_s^{t-1})$ are covered by the loan l , i.e.

$$w_t n(\theta_s^{t-1}) + (r_t + \delta) k(\theta_s^{t-1}) \leq l(\theta_s^{t-1}) \quad \forall \theta_s^{t-1} \quad (3.2)$$

I assume that the primitives of the entrepreneur's satisfy the following assumptions

Assumption 2. $U : (0, \infty) \rightarrow \mathbb{R}$ is C^2 , strictly increasing, strictly concave function and it satisfies $\lim_{c \rightarrow 0} U'(c) = \infty$, $\lim_{c \rightarrow \infty} U'(c) = 0$ and $\sup U(c) < \infty$. The production technology $f : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$ is C^2 , strictly increasing, has decreasing returns to scale of degree γ .

Firm's problem. Entrepreneur seeks to maximize the life time utility from consumption subject to the technological constraints (3.1) and (3.2) and given the loan l and repayment m . Thus, the problem of the firm is

$$\max_{c(\cdot), n(\cdot), k(\cdot)} \sum_{t=j}^{\infty} \sum_{\theta_s^t} (\beta \zeta)^{t-j} \Pr(\theta_s^t) U(c(\theta_s^t)) \quad (3.3)$$

subject to

$$\begin{aligned} c(\theta_s^t) + m(\theta_s^t) &\leq \theta_{st}^{1-\gamma} f(k(\theta_s^{t-1}), n(\theta_s^{t-1})) && \forall \theta_{st}, \theta_s^{t-1} \\ w_t n(\theta_s^{t-1}) + (r_t + \delta) k(\theta_s^{t-1}) &\leq l(\theta_s^{t-1}) && \forall \theta_s^{t-1} \end{aligned}$$

To economize on notation let the maximized output given the loan be $F(l)$, i.e.

$$\begin{aligned} F(l(\theta_s^{t-1})) &\equiv \max_{k(\cdot), n(\cdot)} f(k(\theta_s^{t-1}), n(\theta_s^{t-1})) && (3.4) \\ &\text{subject to} \\ &w_t n(\theta_s^{t-1}) + (r_t + \delta) k(\theta_s^{t-1}) \leq l(\theta_s^{t-1}) \end{aligned}$$

Then, as I argue in the Appendix ... if the dynamic contract specifies lending and repayments the feasibility constraint pins down the consumption of the entrepreneur, which solves the problem (3.3).

3.3.3 Credit market and Financial Intermediation

Financial intermediaries (lenders) arise as institutions participating in the long-term credit market in which they provide funds to firms in the exchange for payments. They are risk neutral and value a stream of consumption good. They discount future with the inverse of the real interest rate $\frac{1}{(1+r)}$. A project of an individual firm is long-lived and its returns are private information, therefore it is optimal for financial intermediaries and firms to enter the long-term dynamic, lending relationships. A detailed specification of the contract will be discussed in the next section. There is free entry into the financial intermediation industry. Thus, in equilibrium all financial intermediaries make zero profits and hence their ownership is immaterial. As a result it is without the loss of generality to consider a single, representative financial intermediary, which is what I do for the rest of the paper. At any point in time the representative intermediary holds a portfolio of contracts with a large number of firms of different types s and histories θ_s^t .

3.3.4 Workers

Workers are hand to mouth and do not participate in the asset market. In each period, they decide how much to work and how much to consume. They maximize the utility $U^w : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ of consumption goods and labor $\{c_t^w, h_t\}_{t=0}^\infty$ subject to the budget constraint, i.e. they solve

$$\max_{c_t^w, h_t} U^w(c_t^w, h_t) \quad \text{s.t.} \quad c_t^w = w_t h_t \quad (3.5)$$

where w_t is the wage rate.

3.3.5 Dynamic lending contract

Every firm that starts operating at any period j receives an offer from the financial intermediary. The offer consist of a contract menu whose terms can be contingent on all public information. I assume that both financial intermediaries and firms are fully committed to

the contract. Hence, no party is allowed to leave the contract in any ex-post state of the world. Then the contract is defined as follows.

Definition 6. A *dynamic contract* is a vector $\mathbf{x}_s \equiv \{l(\theta_s^{t-1}), c(\theta_s^t), m(\theta_s^t)\}_{t=j}^\infty$ specifying for a each firm of type $s \in S$ an amount of lending $l : \Theta_s^{t-1} \rightarrow \mathbb{R}_+$, entrepreneurs consumption $c : \Theta_s^t \rightarrow \mathbb{R}_+$, transfer to the financial intermediary $m : \Theta_s^t \rightarrow \mathbb{R}$.

At every t the contract specifies the amount of lending from the financial intermediary to the firm l , the entrepreneur's consumption c and transfers from the firm to the financial intermediary m . The latter two are contingent on the realization of the demand shock θ_{st} . Feasibility imposes a technological restriction on the contract space, i.e. the sum of entrepreneur's consumption and payments to the financial intermediary can not exceed the return on project. Below I provide the definition of feasible contract.

Definition 7. A dynamic contract \mathbf{x}_s is *feasible* if $\forall t \geq j$ and $\forall \theta_s^{t-1} \in \Theta_s^{t-1}, \forall \theta_{st}$

$$c(\theta_s^t) + m(\theta_s^t) \leq \theta_{st}^{1-\gamma} F(l(\theta_s^{t-1})) \quad (\text{BC})$$

Consider now the restriction on the contract space imposed by the fact that θ is privately observed. By invoking the Revelation Principle, I can without the loss of generality restrict the message space to the set Θ_s for all $s \in S$, i.e. consider only the direct revelation mechanism in which the firm reports it's true type. Define the continuation utility for the entrepreneur associated with the contract \mathbf{x}_s after history θ_s^t (according to the truth telling) as

$$v(\theta_s^t) \equiv \sum_{n=1}^{\infty} \sum_{\theta_s^{t+n}} (\beta\zeta)^{n-1} \Pr(\theta_s^{t+n} | \theta_s^t) U(c(\theta_s^{t+n}))$$

which is useful to define an incentive compatible contract.

Definition 8. A dynamic contract \mathbf{x}_s is *incentive compatible* if it satisfies the following incentive compatibility constraint $\forall t \geq j$ and $\forall \theta_s^{t-1} \in \Theta_s^{t-1}, \forall \theta_{st}, \theta'$:

$$\begin{aligned} U\left(\theta_{st}^{1-\gamma} F(l(\theta_s^{t-1})) - m(\theta_s^{t-1}, \theta_{st})\right) + \beta\zeta v(\theta_s^{t-1}, \theta_{st}) &\geq \\ U\left(\theta_{st}^{1-\gamma} F(l(\theta_s^{t-1})) - m(\theta_s^{t-1}, \theta')\right) + \beta\zeta v(\theta_s^{t-1}, \theta') &\end{aligned} \quad (\text{IC})$$

Incentive compatibility requires that actual realization of the demand θ_{st} is more or equally profitable to report relative to any other realization θ' . On the top of the feasibility and incentive compatibility the contract has to deliver in period j at least the initial promised utility of the entrepreneur, $v_s^0 \in [v_{\min}, v_{\max}]$. This is summarized by the following participation constraint

$$\sum_{t=j}^{\infty} \sum_{\theta_s^t} (\beta\zeta)^t \Pr(\theta_s^t) U(c(\theta_s^t)) \geq v_s^0 \quad (PC)$$

The financial intermediary seeks to maximize the net present value of the payments from the firm subject to the feasibility constraint, incentive compatibility constraint and participation constraint. Thus, the optimal dynamic lending contract solves

$$J(v_s^0) = \max_{\mathbf{x}_s} \sum_{t=j}^{\infty} \sum_{\theta_s^t} \left(\frac{\zeta}{1+r} \right)^{t-j} \Pr(\theta_s^t) [m(\theta_s^t) - l(\theta_s^{t-1})] \quad (3.6)$$

subject to

(BC), (IC) and (PC).

Define the set $\mathcal{I} \equiv \{v \mid \exists \mathbf{x} \text{ s.t. } (BC), (IC) \text{ and } (PC) \text{ holds}\}$ of utility values that can be generated by feasible and incentive compatible contracts. For any initial utility v_s the optimal dynamic lending contract solves problem (3.6) i.e. it maximizes the value obtained by the financial intermediary among all the feasible, incentive compatible contracts.

Randomization. The constraint set in problem (3.6) is not necessarily convex because of the presence of a concave function, $U\left(\theta_{st}^{1-\gamma} F(l(\theta_s^{t-1})) - m(\theta_s^{t-1}, \theta')\right)$, on the right hand side of the incentive compatibility constraint. Thus, randomization may be optimal. It is possible to rule out randomization as part of the optimal contract by making an additional assumption following Aguiar, Amador, and Gopinath (2009).

Assumption 3. Let $C : [U(0), U(\infty)] \rightarrow \mathbb{R}$ and $C = U^{-1}$, $H = F^{-1}$ and $u(\theta) = U(c(\theta))$ and $\underline{u} = U((\theta_i + \theta_j) F(l) + C(U(c_j)))$ and . Define a function

$$G(\underline{u}, u) = -H\left(\frac{C(\underline{u}) - C(u(\theta_i))}{\theta_i - \theta_j}\right) + \frac{C(\underline{u}) - C(u(\theta_j))}{\theta_i - \theta_j}$$

where $\theta_i > \theta_j$. G is concave.

This assumption allows me to transform the constraint set into the linear in u and \underline{u} (see Appendix B.2.2 for the proof). In the quantitative part of the paper I check ex-post for every solution of the individual contract whether the optimal contract satisfies the assumption.

Recursive formulation. Following arguments and techniques by [Atkeson and Lucas \(1992\)](#) one can show that the dynamic contracting problem (3.6) admits a recursive formulation using the entrepreneur's continuation utility $v_s \equiv v(\theta_s^t)$, as a state variable. It solves the following recursive problem for $v_s \in [v_{\min}, v_{\max}]$

$$B_s(v_s) = \max_{l, m(\theta_s), v'(\theta_s)} \left\{ -l + \sum_{\theta_s \in \Theta_s} \pi(\theta_s) \left[m(\theta_s) + \frac{\zeta}{(1+r)} B_s(v'(\theta_s)) \right] \right\} \quad (3.7)$$

subject to

$$v_s = \sum_{\theta_s \in \Theta_s} \pi(\theta_s) [U(\theta_s^{1-\gamma} F(l) - m(\theta_s)) + \beta \zeta v'(\theta_s)]$$

$$U(\theta_s^{1-\gamma} F(l) - m(\theta_s)) + \beta \zeta v'(\theta_s) \geq U(\theta_s^{1-\gamma} F(l) - m(\theta')) + \beta \zeta v'(\theta') \quad \forall \theta_s, \theta'$$

where $B_s(v_s)$ is the maximal discounted value of net payments that the financial intermediary can attain subject to the constraint that the recursive contract delivers a value v_s to the firm (promise keeping constraint) and recursive version of the incentive compatibility constraint. Using the fact that $c(\theta_s) = \theta_s^{1-\gamma} F(l) - m(\theta_s)$ the problem (3.7) can be rewritten as

$$B_s(v_s) = \max_{l, c(\theta_s), v'(\theta_s)} \left\{ -l + \sum_{\theta_s \in \Theta_s} \pi(\theta_s) \left[\theta_s^{1-\gamma} F(l) - c(\theta_s) + \frac{\zeta}{1+r} B_s(v'(\theta_s)) \right] \right\} \quad (3.8)$$

subject to

$$v_s = \sum_{\theta_s \in \Theta_s} \pi(\theta_s) [U(c(\theta_s)) + \beta \zeta v'(\theta_s)]$$

$$U(c(\theta_s)) + \beta \zeta v'(\theta_s) \geq U((\theta_s^{1-\gamma} - \theta_s'^{1-\gamma}) F(l) + c(\theta')) + \beta \zeta v'(\theta') \quad \forall \theta, \theta'$$

It becomes clear from this formulation of the contracting problem that incentive considerations affect lending given the presence of $(\theta_s^{1-\gamma} - \theta_s'^{1-\gamma}) F(l)$ term in the incentive compatibility constraint. Also, the dispersion of demand outcomes (*ceteris paribus*) affects the tightness of the incentive compatibility constraint. These two features of the dynamic contract are going to play crucial role in the propagation of the aggregate shock.

3.3.6 Aggregation

At any period the financial intermediary holds the portfolio of lending contracts with a large number of firms of different types $s \in S$. This portfolio can be summarized by the probability distribution over the space of the continuation utilities. Let $V = [v_{min}, v_{max}]$ and let $(V, \mathcal{B}(V))$ be a measurable space of promised utilities, where $\mathcal{B}(V)$ denotes the Borel set. Define a measure $\mu_s : \mathcal{B}(V) \rightarrow [0, 1]$ over the space of continuation utilities for firms of type s . The type s is fixed, thus the updating operator for each s is defined as follows

$$T\mu_s(V) = \int_V Q(v_s, V) d\mu_s(v_s) \quad \forall \mathcal{A} \in \mathcal{B}(V)$$

where $Q(v_s, \mathcal{A}) : V \times \mathcal{B}(V) \rightarrow \mathbb{R}$ is a transition function defined as

$$Q(v_s, \mathcal{A}) = \begin{cases} \zeta \sum_{\theta_s \in \Theta_s} \pi(\theta_s) \mathbb{I}\{v'(v_s, \theta_s) \in \mathcal{A}\} & \text{for } \forall \mathcal{A} \in \mathcal{B}(V) \\ (1 - \zeta) & \text{for } \mathcal{A} = v_s^0 \end{cases}$$

Aggregate lending in the economy is then defined by

$$L = \sum_{s \in S} \Gamma_s \int_V l(v_s) d\mu_s(v_s) = \sum_{s \in S} \Gamma_s \int_V (wn(v_s) + (r + \delta)k(v_s)) d\mu_s(v_s)$$

whereas the aggregate payments received by the financial intermediary are

$$P = \sum_{s \in S} \Gamma_s \int_V \pi(\theta_s) m(v_s, \theta_s) d\mu_s(v_s)$$

Note that by the law of large numbers the fraction of firms of type s that received shock θ_s is exactly $\pi(\theta_s)$, so there is no uncertainty about the size of the aggregate payments for the financial intermediary. As a result, the asset holdings of the financial intermediary

evolve according to

$$A' = (1 + r)A + (P - L)$$

i.e. the stock of assets in the portfolio of financial intermediary depends on the assets inherited from the previous period $(1 + r)A$ and the aggregate net aggregate payments $(P - L)$. In the stationary equilibrium this equation becomes $rA + P - L = 0$. To complete the description of the financial intermediary I have to pin down the initial promised utility for the entrepreneur who starts to operate a firm of type s , v_s^0 . Lending market is perfectly competitive thus in equilibrium the financial intermediary earns zero profits. Then the v_s^0 is determined by the following free entry condition

$$B_s(v_s^0) = 0 \tag{3.9}$$

Before I state the formal definition of the equilibrium I define the rest of aggregate variables in the economy. The aggregate output is given by

$$Y = \sum_{s \in S} \Gamma_s \int_V \pi(\theta_s) [\theta_s^{1-\gamma} F(l_s(v_s))] d\mu_s(v_s)$$

the aggregate consumption of the entrepreneurs is

$$C^e = Y - P$$

the aggregate labor input is

$$N = \sum_{s \in S} \Gamma_s \int_V n_s(v_s) d\mu_s(v_s)$$

and the aggregate capital input is

$$K = \sum_{s \in S} \Gamma_s \int_V k_s(v_s) d\mu_s(v_s)$$

The markets clearings are $K = A$ and $N = h$ for the capital and labor market respectively. Then by the Walras law the market clearing for the consumption good becomes (see the

Appendix B.2.1 for derivation)

$$Y = C^e + C^w + K' - (1 - \delta) K \quad (3.10)$$

i.e. the aggregate output is divided into consumption of the entrepreneurs, consumption of workers and investment.

3.3.7 Recursive equilibrium

The aggregate state of the economy in any period can be summarized by S -dimensional vector of distributions over continuation utilities $\mu = (\mu_1, \dots, \mu_S)$. Thus a stationary, recursive equilibrium is defined as follows:

Definition 9. A *stationary recursive equilibrium* consists of: (i) an allocation of the household $\{c^w, h\}$ (ii) a contract policy $\{l(v_s), m(v_s, \theta_s), c(v_s, \theta_s)\}_{s \in S}$ (iii) an allocation of the firm $\{n(v_s), k(v_s)\}_{s \in S}$ (iv) prices $\{r, w\}$ (v) initial promised utility value v_s^0 (vi) the measure μ_s over the space of promised utility, such that :

1. Given $\{w\}$, an allocation of the workers $\{c^w, h\}$ solves the (3.5).
2. Given $\{r, w\}$, Contract policy $\{l(v_s), m(v_s, \theta_s), c(v_s, \theta_s)\}_{s \in S}$ solves the problem (3.8).
3. Given $\{r, w\}$, an allocation of the firm $\{n(v_s), k(v_s)\}_{s \in S}$ solves the problem (3.4).
4. Markets clear: $A = K$ and $N = h$.
5. The initial promised utilities $\{v_s^0\}_{s \in S}$ solves the problem (3.9).
6. The measures $\{\mu_s\}_{s \in S}$ are stationary.

3.4 Theoretical results

In this section I describe theoretical results that shed light on the properties of the dynamic lending contract and their implications for access to the financing as a function of firm's size and age. Further, I provide an analysis of how aggregate shocks to microeconomic uncertainty affect properties of the contract policy functions.

3.4.1 Preliminaries

I start with the simplification of the constraint set of the problem 3.7. Denote the incentive compatibility constraint for all $s \in S$ and $m, n \in N$ by

$$C_{n,m}^s \equiv U(\theta_{sn}^{1-\gamma} F(l) - m(\theta_{sn})) + \beta \zeta v'(\theta_{sn}) - [U(\theta_{sm}^{1-\gamma} F(l) - m(\theta_{sm})) + \beta \zeta v'(\theta_{sm})]$$

where n is the actual demand state and m is the reported demand state. The lemma below allows me to consider only the local constraints. Satisfying local downward and upward constraints implies all the global constraints are also satisfied. This result is standard in dynamic contracting environment with private information either in case of risk neutral agents and optimal investment or in case of risk-averse agents with unobservable endowments. It holds also in my environment, which embodies both features.

Lemma 4. *If the local downward constraints $C_{m,m-1}^s \geq 0$ and upward constraint $C_{m,m+1}^s \geq 0$ hold for each $m \in N$, then the global constraints $C_{m,n}^s \geq 0$ holds $\forall m, n \in N$.*

Proof. Appendix B.2.2 □

Another standard result can also be established in my model. Financial intermediary imposes a weakly lower payment on an entrepreneur reporting a lower demand in exchange for a lower future continuation utility. Entrepreneur reporting higher demand shock returns a weakly higher payment in exchange for higher continuation utility in the future. This way financial intermediary is able to provide partial insurance to risk averse entrepreneurs. In the Proposition 7 I further characterize payments and consumptions of entrepreneurs in my environment and compare it to the case with risk neutral entrepreneurs with variable investment.

Lemma 5. *An incentive compatible contract policy satisfies $m(\theta_{sn}) \geq m(\theta_{sn-1})$ and $v'(\theta_{sn}) \geq v'(\theta_{sn-1})$ for $\theta_{sn} > \theta_{sn-1}$ and for all $s \in S$.*

Proof. Appendix B.2.2 □

In the next lemma I establish useful properties of the value function $B_s(v)$.

Lemma 6. (i) Under Assumptions 2 and 3 for every $s \in S$ value function $B_s : [v_{\min}, v_{\max}] \rightarrow \mathbb{R}$ is strictly concave and maximizers $v'(\theta_s), m(\theta_s) : [v_{\min}, v_{\max}] \rightarrow \mathbb{R}$ and $c(\theta_s), l(v_s) : [v_{\min}, v_{\max}] \rightarrow \mathbb{R}_+$ are continuous, single-valued functions.

(ii) Under Assumption 2, the value function B_s is differentiable.

Proof. Appendix B.2.2 □

Under Assumption 2, strict concavity follows from $G(\underline{u}, u)$ being concave and constraint set to be convex. As a result by standard dynamic programming arguments I can argue that contract policy functions are continuous and single valued functions. This, paired with the differentiability (which holds independently on Assumption 3) of the value function is useful for theoretical characterization of a dynamic contract as well as for a computational algorithm I use in the quantitative part of the paper (see Appendix 12).

3.4.2 Borrowing constraint as a function of age.

To illustrate the role of private information in my model, and in particular it's role in access to exogenous financing, it is useful to introduce a full information benchmark first. Consider a relaxed version of the contracting problem 3.8 in which the incentive compatibility constraint is dropped. With no incentive considerations, firm's project will be financed to maximize the flow of the profits of the financial intermediaries i.e. up to the point where marginal cost of lending an additional unit of resources is equalized with the expected marginal benefit of it. Hence, lending would be determined by the solution to the following problem

$$\max_l \mathbb{E} [\theta_s^{1-\gamma}] F(l) - l \tag{3.11}$$

This leads to the following definition of a static efficiency.

Definition 10. A statically efficient level of lending, l_s^* , is determined by

$$\mathbb{E} [\theta_s^{1-\gamma}] F'(l_s^*) = 1.$$

Moreover define the value of the financial after realization of the demand shock

Definition 11. Let the value of the discounted stream of profits for the financial intermediary after realization of the shock be denoted by $g(v_s, \theta_s) = l(v_s) - m(v_s, \theta_s) + \frac{\zeta}{1+r} B(v'(v_s, \theta_s))$.

The next proposition establishes three properties of the dynamic contract policies that I will later use to characterize an access to financing as a function of firm's age.

Proposition 7. For all $s \in S$ a contract policy is such that:

- (i) The contract policy is dynamic: $\forall v_s \in [v_{\min}, v_{\max}]$, $m(v_s, \theta_{si}) > m(v_s, \theta_{sj})$, $c(v_s, \theta_{si}) > c(v_s, \theta_{sj})$, and $v'(v_s, \theta_{si}) > v'(v_s, \theta_{sj})$ for $\theta_{si} > \theta_{sj}$.
- (ii) There are distortions in lending. There exists $v^* \in [v_{\min}, v_{\max}]$ such that $l(v) < l^*$ for all $v \in [v_{\min}, v^*]$ and $l(v) = l^*$ for all $v \in [v^*, v_{\max}]$.
- (iii) There is a coinsurance. For $\forall v \in [v_{\min}, v_{\max}]$, $g(v, \theta_i) > g(v, \theta_j)$.

Proof. Appendix B.2.2 □

Part (i) states that optimal contract policy is dynamic, i.e. it uses a variation in the continuation utility and payments to the financial intermediary to provide intertemporal incentives and partial insurance. An efficient arrangement imposes revealing true realizations through the combination of payments to the financial intermediary and continuation utilities that are contingent on the realization of the idiosyncratic shock. Following low demand realization financial intermediary requires low repayment, but also delivers low continuation value. After high realization of demand shock intermediary requires high repayment, but also delivers high continuation value for the entrepreneur. This way financial contract provides some insurance against idiosyncratic risk, albeit imperfect. The entrepreneur who received high realization of the demand shock consumes strictly more relative to the one with lower realization. Part (ii) states that for any there exists a point v^* in the continuation utility domain such that providing statically efficient level of financing is feasible and incentive compatible. For any point left to v^* the informational friction implies an existence of the endogenous borrowing constraint, tightness of which can be measured by $(l^* - l(v))$, a difference between efficient level of financing and

level implied by the optimal contract with private information. The existence of this constraint is induced by the presence of key tradeoffs between production efficiency, providing insurance and maintaining proper intertemporal incentives. I discuss them in details in the next section. The existence of the endogenous borrowing constraint is crucial for the economic mechanism driving an asymmetric response of firms to the uncertainty shocks. As I argue in Section 3.4.5 shocks to microeconomic uncertainty induce movements in endogenous borrowing constraint, in particular increase in uncertainty tightens the constraint. Finally, part (iii) states that there is a coinsurance since following a low demand shock net payments for financial intermediary are lower relative to the payments after high demand shock. Thus, there is some relief in the amount of repayments when the marginal utility of consumption for entrepreneur is high. At the same time financial intermediary benefits from high realizations by extracting larger payments from the firm.

Given the existence of the v^* the question arises whether firms on average grow towards the unconstrained levels. There are two counteracting forces and a general theoretical characterization is not possible⁵. Firstly, due to the private information full intertemporal risk sharing is not achievable, and the optimal, dynamic contract induces that on average an entrepreneur borrows against his future income. As a result, this force pushes the entrepreneur's expected utility downwards over time. In particular, whenever $(1+r) = \beta^{-1}$ then $v > \mathbb{E}[v]$ as argued by [Green \(1987\)](#), [Thomas and Worrall \(1990\)](#) or [Atkeson and Lucas \(1992\)](#). I label this tendency of continuation utility to fall over the lifetime as an incentive effect. Secondly, the rate $(1+r)$ at which financial intermediary discounts future cash flows matters for the evolution of the continuation utilities. As interest rate increases financial intermediary is less patient and it is optimal to receive payments from the firm earlier. Other things equal increase in the interest rate pushes the expected continuation value up. The way to see it is to inspect the condition derived from the necessary and sufficient first order conditions (see [Appendix B.2](#))

$$B'(v) = \frac{1}{\beta(1+r)} \mathbb{E}[B'(v)]$$

⁵This is contrary to the environment with risk neutral entrepreneurs as in [Clementi and Hopenhayn \(2006\)](#) where one can show the firm's equity is a submartingale with two absorbing states, where one of them is unconstrained level of financing.

as interest rate increases the $\frac{1}{\beta(1+r)}$ falls the absolute value (given the utility function I impose in Assumption 7 $B'(v)$ is negative over the entire domain of v) of $\mathbb{E}[B'(v)]$ has to increase. Given that $\lim_{v \rightarrow v_{\min}} B'(v) = 0$ and $\lim_{v \rightarrow v_{\max}} B'(v) = \infty$ and monotonicity of the value function $B(v)$ it has to be that on average v increases. I label this force as the interest rate effect. Therefore a net effect of the two forces determines whether on average firms move towards v^* and hence the unconstrained level of financing.⁶ In the quantitative section I make sure that firms on average are growing in the economy. This property of the model as well as the presence of the endogenous borrowing constraint is illustrated in Figure 3.3.

Key tensions. Absent informational friction, within considered environment, the lender would completely insure the borrower and would provide a statically efficient level of financing in every period, which equalizes expected revenue from additional investment with its marginal cost. Entrepreneur would receive, independent on the realization of the demand shock, a constant stream of consumption. Moreover, the realization of the shock would have no effect on the continuation of the contract. The presence of the private information introduces a tension between production efficiency, providing insurance and maintaining incentives for the financial intermediary. To illustrate this trade-offs skip the dependence of the problem on $s \in S$ and let $N = 2$ with $\Theta = \{\theta_H, \theta_L\}$ and consider the following variation: for some $\varepsilon \in \mathbb{R}$ sufficiently close to zero, decrease v'_H and v'_L by ε/β and increase c_H and c_L such that $U(c_H)$ and $U(c_L)$ increase by ε . Then the feasible, incentive compatible lending under considered deviation increases by $\varepsilon_l = \frac{\varepsilon}{F'(l)}\Omega_1$ where

$$\Omega_1 = \frac{1}{(\theta_H^{1-\gamma} - \theta_L^{1-\gamma})} \left(\frac{1}{U' \left((\theta_H^{1-\gamma} - \theta_L^{1-\gamma}) F(l(v)) + c_L \right)} - \frac{1}{U'(c_L)} \right) > 0$$

⁶Somehow less restrictive requirement in my environment is the presence of any positive mass of unconstrained firms in equilibrium. This can happen even if the sufficient condition for average firm to grow, i.e. $\mathbb{E}[(v)] > v$ is not satisfied. Conditional on on survival and receiving long enough sequence of high demand shocks a firm can reach unconstrained level of financing, even if on average firms are not growing in the economy

This variation leaves the value of the intermediary unchanged, i.e. $\frac{\Delta B}{\varepsilon} = 0$, where

$$\frac{\Delta B}{\varepsilon} \approx \frac{\Omega_1 (\mathbb{E} [\theta^{1-\gamma}] F'(l) - 1)}{F'(l)} - \left[\frac{\pi_H}{U'(c_H)} + \frac{\pi_L}{U'(c_L)} \right] - \frac{1}{\beta(1+r)} [\pi_L B'(v'_L) + \pi_H B'(v'_H)] \quad (3.12)$$

By strict concavity of value function B and the fact that $v'_L < v'_H$ the third term in the equation above is positive. It reflects the benefit from reduction in continuation values. As resources are transferred from future to current period the promise keeping constraint requires an increase of both consumptions. This is costly for the intermediary and is reflected by the second, negative term in the equation 3.12 above. The first term reflects the incentive effects on lending. Note that absent informational frictions $\mathbb{E} [\theta^{1-\gamma}] F'(l) = 1$ and the first term vanishes. However, if only the incentive compatibility constraint is binding transferring resources from future to current periods optimal contract implies the intermediary lends more to an entrepreneur and thus its profit increases, given that $\mathbb{E} [\theta^{1-\gamma}] F'(l) - 1 > 0$ whenever the constraint is strictly binding.

3.4.3 Access to financing increases with average return

The access to exogenous financing grows with age as argued in the previous section. In this section I argue that the type s which determines the expected return on a project matters for the level of financing in my environment, firms are able to borrow more if their expected return on project (terminal size) is larger. Absent any information friction it follows directly from the problem (3.11) that lending increases with expected return on a project. Though it is not straightforward to see that this property extends to the case with private information. To argue that it does consider the following assumption on the structure of uncertainty, which I will later also use in the quantitative section.

Assumption 7. Assume $U(c) = \frac{c^{1-\rho}}{1-\rho}$ with $\rho > 1$. Let $N = 2$ with $\Theta = \{\theta_{sL}, \theta_{sH}\}$ for all $s \in S$ with $\pi_L = 1 - \pi_H$ and let

$$\theta_{sH} = \left(\bar{\theta}_s + \frac{\sigma}{\pi_H} \right)^{\frac{1}{1-\gamma}}, \quad \theta_{sL} = \left(\bar{\theta}_s - \frac{\sigma}{\pi_L} \right)^{\frac{1}{1-\gamma}}$$

implying $\mathbb{E} [\theta_s^{1-\gamma}] = \bar{\theta}_s$ and $std(\theta_s^{1-\gamma}) = \frac{\sigma}{\sqrt{\pi_L \pi_H}}$.

Thus, the type s the firm draws at the beginning of operation affects only the average

return on a project, but not the standard deviation of the idiosyncratic shock. Given this assumption I establish the following proposition.

Proposition 8. *Under Assumption 7 the optimal contract policies $x = \{c(\theta_s^t), m(\theta_s^t), l^*(\theta_s^{t-1})\}$ are functions homogenous of degree one in histories of idiosyncratic shocks.*

Proof. Appendix B.2.2 □

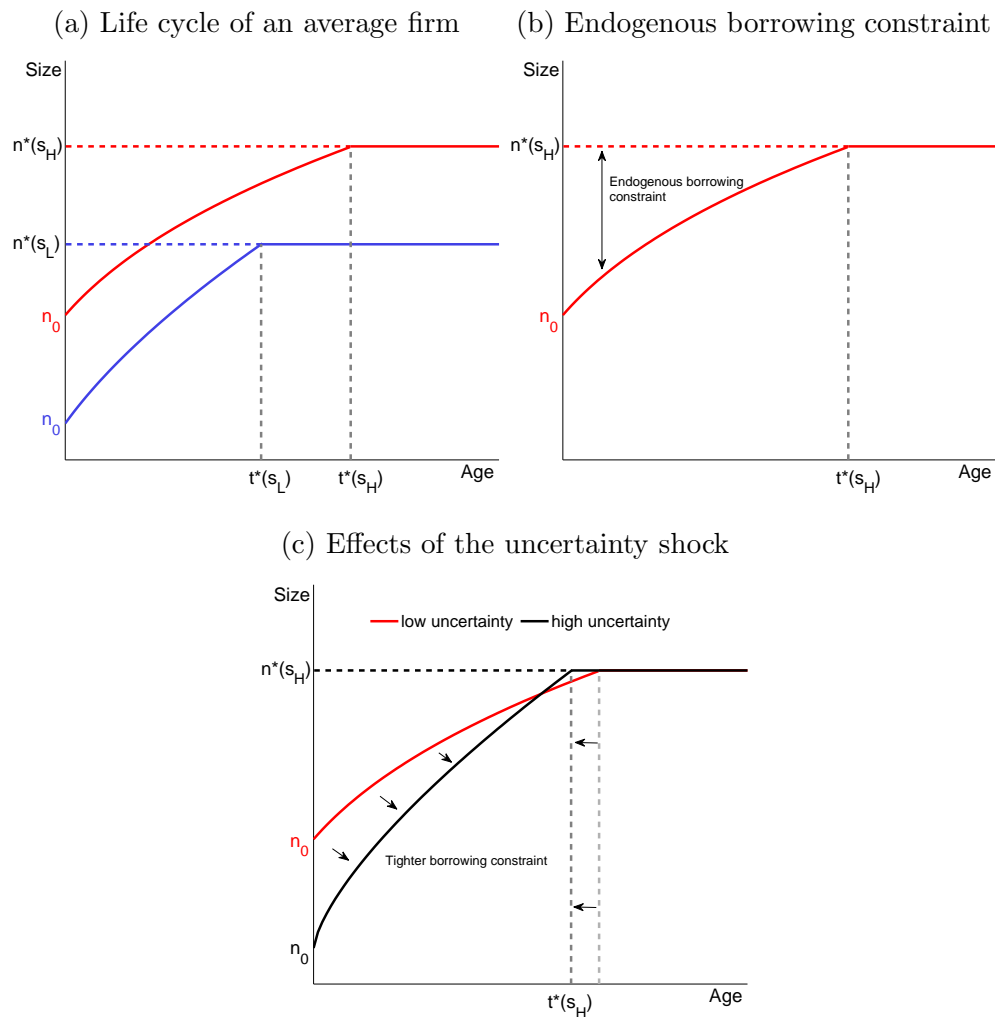
The relevance of this proposition can be illustrated by considering the existence of only one type s . Suppose x^* is a solution to the optimal contracting problem (3.6). Then let a history of shocks be scaled by a factor of $\lambda > 0$, i.e. $\Theta = \{\lambda\theta_L, \lambda\theta_{sH}\}$, implying that $\mathbb{E}[\theta_s^{1-\gamma}] = \lambda^{1-\gamma}\bar{\theta}$. The solution to the optimal contracting problem then, according to the proposition above, is λx^* , i.e. it is scaled. Given the Assumption 7 the following corollary holds

Corollary 8. *Under Assumption 7, amount of lending $l(v)$ increases at a constant rate as a function of the average returns on a project.*

Proof. Appendix B.2.2 □

Hence, the larger the average return on a project the more resources are lent to the firm. This property is illustrated in Figure 3.3. Also the initial value of the continuation utility v_0 , pinned down by the free entry condition 3.9, increases with firm's average return on a project. The fact that financing increases with expected return is an intuitive feature of my environment, if a project is on average more profitable for the entrepreneur efficiency dictates for the financial intermediary to invest more resources in it to maximize profits. The presence of private information distorts the efficient allocation of resources although the monotonicity of investment with respect to the average return on project is preserved.

Figure 3.3: Life cycle of an average firm and role of the aggregate uncertainty.



Notes: Size is measured in terms of employment n . s_H stands for high initial technology draw, s_L stands for low initial technology draw.

3.4.4 Existence of a stationary equilibrium

So far my theoretical characterization has been limited to an individual contract properties. One of the contributions of my paper is to embody dynamic, contracting problem into a general equilibrium model with a large number of firms, which further is used to

conduct a quantitative analysis. Though, before that I make sure that recursive stationary equilibrium defined in Section 3.3 actually exists. I establish and prove the existence in the following proposition.

Proposition 9. *Stationary recursive competitive equilibrium exists.*

Proof. Appendix B.2.2 □

3.4.5 Policy functions, equilibrium and movements in uncertainty

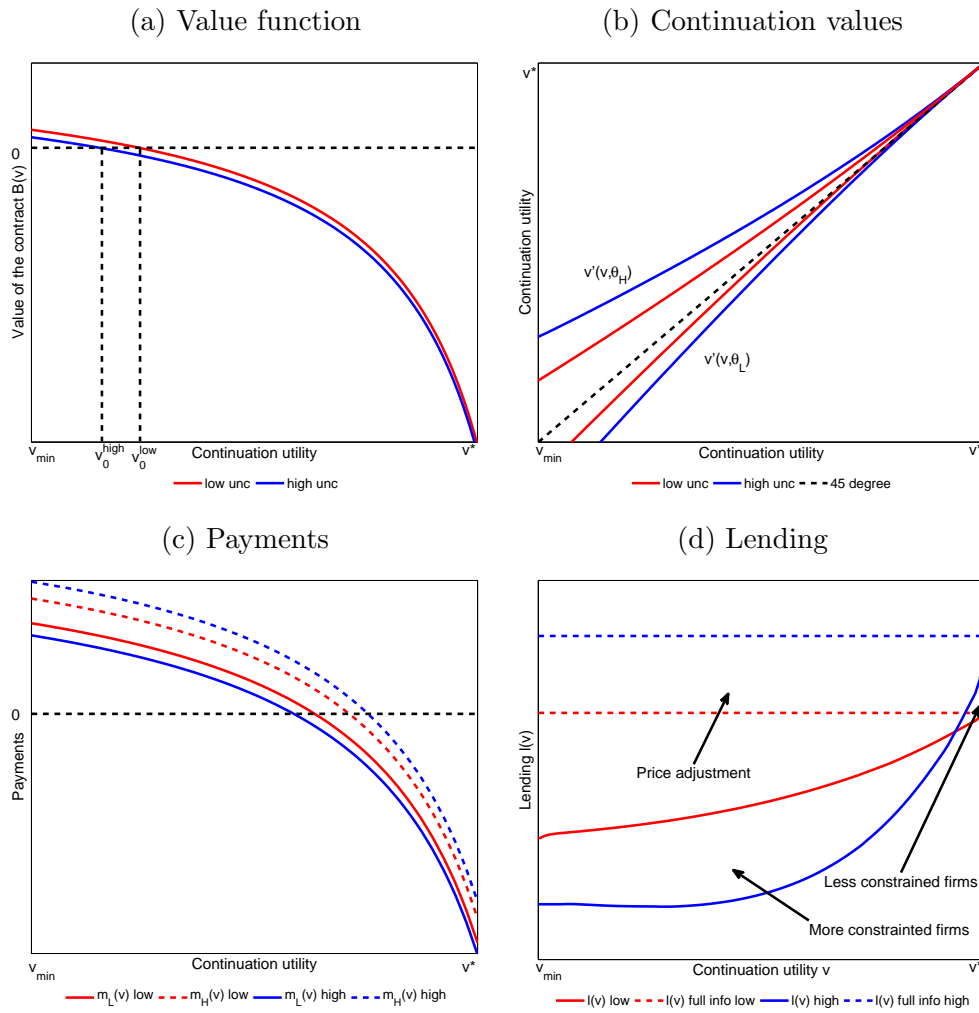
In previous sections I characterized the optimal contract policy functions. In this section I illustrate and discuss the qualitative properties of the policy functions and illustrate the effect of an increase in microeconomic uncertainty on the equilibrium of my model. Figure 3.4 illustrates value of the contract, continuation values, payments to the financial intermediary and lending in two scenarios: (i) low microeconomic uncertainty equilibrium (ii) high microeconomic uncertainty equilibrium. Start with the low uncertainty equilibrium. Consider a newly born firm. It's initial value v_0^{low} (initial utility of entrepreneur) is pinned down by the zero profit condition which is illustrated in panel (a) of Figure 3.4 by a point at which value of the contract crosses zero. Then, the evolution of the continuation utility is governed by the policy functions depicted in panel (b). Following high demand realization an optimal contract delivers high continuation value $v'_H(v) > v$, whereas after low realization contract delivers low continuation value $v'_L(v) < v$. As I argue in Section 3.4.2 I make sure the average firm's continuation utility increases with age, which is reflected by asymmetry between $v'_H(v) - v$ and $v'_L(v) - v$. Average payments from the firm to the financial intermediary falls with the continuation value. Thus, at the beginning of the operation firm makes positive payments to the financial intermediary and as it ages the payments become negative, i.e. it uses resources deposited for consumption. Finally, panel (d) presents lending over the domain of continuation utilities. Dashed red line depicts the full-information benchmark, in which firm receives statically efficient level of financing. The informationally constrained case (solid red line) level of lending increases with promised utility of entrepreneur⁷. Given, that on average promised utility of the firm is increasing

⁷I was unable to show any monotonicity result with regards to lending function. In most of my numerical simulations lending was strictly increasing albeit I also encountered the presence of decreasing part close to v_{min} . The reason for existence of such part can be as follows. It happens at the region of the state space

with time in my economy as firm ages it grows out of the constraint. Now consider, under Assumption 7, a comparative statics exercise with respect to the level of microeconomic uncertainty parameter σ . The response of the economy is illustrated with the blue lines in Figure 3.4. Facing higher risk in order to provide proper intertemporal incentives financial intermediary has to spread away more the continuation utilities, which given the strict concavity of the value function is costly. As a result due to tradeoffs illustrated by equation 3.12 amounts of insurance and lending have to decline. This is reflected firstly by larger spread in payments and more importantly by tightening of the borrowing constraint faced by the firms. The general equilibrium effect of falling wages and interest rates induces an increase in the full information level of financing by increasing the marginal benefit of investment. The net effect of these two forces matters. As a result firm, dependent on the position of the firm in the state space, can either be more constrained in borrowing or less constrained in borrowing. The latter case though holds for just a small interval of the domain of promised utilities (see the arrows in panel (d)). In particular firm is, on average, more constrained when young and less constrained when very old. This constitutes the first channel through which my model generates an asymmetric response to uncertainty shocks. Secondly, note that in the high uncertainty equilibrium the initial continuation value for the entrepreneur is lower ($v_0^{high} < v_0^{low}$) inducing smaller size of new entrants in times of heightened uncertainty. This is another channel contributing to asymmetric response across young and old firms. The effects of the uncertainty shocks for the employment over the life cycle of an average firm is illustrated in Figure 3.3.

where value of the contract flattens, i.e. it is relatively cheaper to spread away the continuation values rather than distort so much the efficiency margin. Taking this argument to the limit from the properties of $B(v)$ we have $\lim_{v \rightarrow v_{\min}} B''(v) = 0$ i.e. cost of spreading promised utilities goes to zero at the left endpoint, so a financial intermediary has room for not distorting so much the efficiency margin.

Figure 3.4: Qualitative properties of the value and contract policy functions



3.5 Calibration and Quantitative results

In this section I present the calibration of the model and undertake quantitative exercise illustrating the role of shocks to microeconomic uncertainty. I calibrate the economy to match steady state moments of the model and their counterparts in the data. I then

conduct impulse response analysis and discuss the macroeconomic impact of aggregate shock to microeconomic uncertainty shock. Next, I compare it's effects to those arising from the movements in aggregate productivity.

3.5.1 Calibration

A period in my model is a quarter. To calibrate the model, I partition my model parameters into those which I fix in advance following estimates typically found in the macroeconomic literature and those which I calibrate by matching model implied moments with the data. First, I set the functional forms for preferences and technology. I assume GHH preferences of workers with period utility given by

$$U^w(c, h) = \frac{1}{1 - \rho} \left(c - \psi \frac{h^{1+\frac{1}{\nu}}}{1 + \frac{1}{\nu}} \right)^{1-\rho}.$$

Moreover, I assume entrepreneurs have does not supply labor and hence their period utility is given by

$$U(c) = \frac{c^{1-\rho}}{1 - \rho}$$

Technology of production is a decreasing returns to scale function of capital and labor $f(k, n) = k^\alpha n^\gamma$. As for parameters set in advance, I set the intertemporal elasticity of substitution $1/\rho$ to 0.5; the number frequently used in the literature (e.g. [Davila et al. \(2012\)](#) and [Conesa et al. \(2009b\)](#)). For the Frisch elasticity, ν , I rely on estimates from [Heathcote et al. \(2010b\)](#) and use 0.72. This value is intended to capture both the intensive and the extensive margins of labor supply adjustment together with the typical existence of two earners within a household. It is also close to 0.82, the number reported by [Chetty et al. \(2011\)](#) in their meta-analysis of estimates for the Frisch elasticity using micro data. Share of capital, α , is disciplined by the long run properties of the US time series. Parameters exogenously imposed on the model are summarized in [Table 3.5](#).

The value for ψ is chosen so that average hours worked equals 0.3 of total available time endowment. Depreciation rate, δ , is set to match the investment to output ratio of 0.24. The discount factor of entrepreneurs is set so that the interest rate in the initial stationary

Table 3.5: Exogenously Determined Parameters of the Baseline Economy

Parameter	Value
Inverse of IES, ρ	2.0
Frisch elasticity of labor, ν	2.0
Share of capital, α	0.35
Share of labor, η	0.60

equilibrium is equal to 4% annually (equivalently to 0.099 quarterly). Note that implied product is $(1+r)\beta > 1$ implying under my calibration firms on average are growing, since $v < \mathbb{E}(v)$. The survival probability ζ controls directly the average age of firms in the economy and therefore it is disciplined by the percentage of young firms in the data equal to 0.41. I use a constant A to normalize output in the economy to 1.

Table 3.6: Preference and Technology Parameters and Associated Targets

Parameter	Value	Target	Value	Model
A	0.97	Aggregate output	1.0	1.0
δ	0.06	Investment-to-Output ratio	0.24	0.24
β	0.99	Average growth rate of firms	0.02	0.02
ζ	0.92	Share of young firms in total	0.41	0.41
ψ	2.2	Average hours worked	0.30	0.30

Size-Age Distribution. A key and novel aspect of the calibration of my model is to match jointly the age and size distribution of the firms in the US economy. Firstly, I assume firm can be of one of three fixed types, i.e. $I = \{1, 2, 3\}$, with expected returns on project $\bar{\theta}_1 < \bar{\theta}_2 < \bar{\theta}_3$. Moreover, let Γ_s be a probability of drawing a type $s \in S$. Table 3.7 summarizes the choices of parameters and associated targets which determine the size/age distribution of firms in the model. Business Dynamics Statistics provides data for firms of different size, where size is measured in terms of number of employees and classified in 12 bins. In my calibration I associate firms with less than 20 employees to be of type 1, firms with 20 to 99 employees to be of type 2 and firms with more than 100 employees to be of type 3. I use the fraction of firms within each of three groups to discipline Γ_s . Note,

88% of firms in the US economy is classified in group 1, 10% in groups 2 and only 2% in group three. In terms of employment share, the third group of firms account for 69% of total employment, the second group accounts for 15% and the first group for 16%. I use the former two numbers to discipline the $\bar{\theta}_2$ and $\bar{\theta}_3$, so that the models exactly matches employment shares from the data. I impose the idiosyncratic uncertainty to evolve in line with Assumption 7, i.e. $N = 2$ and

$$\begin{aligned}\theta_{sL}^{1-\gamma} &= \bar{\theta}_s - \frac{\sigma_s}{\pi_L} \\ \theta_{sH}^{1-\gamma} &= \bar{\theta}_s + \frac{\sigma_s}{\pi_H}\end{aligned}$$

Thus, the standard deviation of the idiosyncratic shock, which is privately observed in my model, is controlled by σ and π_H (given that $\pi_L = 1 - \pi_H$). I use these two parameters to reflect the shares of employment of young firms across size groups. My preferred definition of small firms that I use for the rest of the quantitative part and which I relate to it's data counterpart is firms with less than 100 employees. Thus, I target the share of employment of young firms in employment of small and large firms defined this way. Ability of the model to replicate *joint* age/size distribution of employment is a crucial feature for the quantitative part of the paper.

Table 3.7: Size/Age Distribution Parameters and Associated Targets

Parameter	Value	Target	Value	Model
$\bar{\theta}_2$	1.13	Employment share of group 2	0.18	0.18
$\bar{\theta}_3$	1.57	Employment share of group 3	0.62	0.59
Γ_1	0.88	Firm's share of group 1	0.88	0.88
Γ_2	0.10	Firm's share of group 2	0.10	0.10
Γ_3	0.02	Firm's share of group 3	0.02	0.02
σ_1	0.34	Employment share of young in group 1	0.33	0.35
σ_2	0.31	Average growth rate for group 2*	0.012	0.014
σ_3	0.27	Average growth rate for group 3*	0.006	0.007

Notes: Group 1 has 1-20 employees; Group 2 has 20-99 employees; Group 3 has 100+ employees. Young firms are less 0-5 and old 6+. Source: BDS

*Numbers based on calculations by Haltiwanger et. al. (2013)

Transitional dynamics: a stand on existing contracts. In my quantitative exercise I compute a transitional dynamics (under the perfect foresight) between steady states after an increase in standard deviation of the demand shock θ . In my environment such exercise requires taking an explicit stand on what happens with existing contracts after the economy is hit with the uncertainty shock. Denote by $\mu_{s,0}^*(v_{s,0}; \sigma_l)$ a stationary distribution over the state space associated with a low uncertainty. Denote value functions associated with low and high uncertainty by $B_s(v_s; \sigma_l)$ and $B_s(v_s; \sigma_h)$ respectively. An evolution of the distribution over deterministic transition is determined by the following law of motion

$$\mu_{s,t+1}(v_{s,t+1}) = T\mu_{s,t}(v_{s,t}) = \int_V Q(v_{s,t}, v_{s,t+1}) \mu_{s,t}(v_{s,t})$$

where $\mu_{s,1}(v_{s,1}; \sigma_h)$ is first period distribution from which the iteration is started after the economy is hit with the shock in this period. In terms of what happens with existing contracts two possible cases are possible: (i) all the risk is absorbed by the financial intermediary (ii) all the risk is absorbed by an entrepreneur. For (i) we have $v_{s1} = v_{s0}$ i.e. there is no change in the value of the continuation utility for any entrepreneur and hence

$$B_s(v_{s1}; \sigma_h) \neq B_s(v_{s0}; \sigma_l)$$

given that an idiosyncratic risk is greater in economy with σ_h . This implies

$$\mu_{s,1}(v_{s,1}; \sigma_h) = \mu_s^*(v_s; \sigma_l)$$

so the economy starts with the distribution inherited from the low uncertainty economy. For case (ii) we have

$$B_s(v_s; \sigma_l) = B_s(\hat{v}_s; \sigma_h)$$

for some $\hat{v}_s \neq v_s$ and hence

$$\mu_{s,1}(v_{s,1}; \sigma_h) = \mu_s^*(\hat{v}_s; \sigma_l)$$

which implies a shift in distribution over continuation utilities. In what follows I compute transitional dynamics for case (i) in which the financial intermediary absorbs the risk

induced by movements in micro uncertainty.

3.5.2 The Effects of Uncertainty Shocks

In this section I present the main quantitative result of my paper. I evaluate the implications of the aggregate shock to microeconomic uncertainty faced by firms. I model the uncertainty shock as an increase in standard deviation of the demand shock θ . I assume that the shock declines upon impact and this decline decays over time. I fix the decay of the impulse so that the shock has a half-life of 4 quarters. I calibrate the initial impulse so that on impact, a standard deviation of θ increases by 60% a number comparable to the increase in the standard deviation of the firm-level TFP⁸ in the US during last four recession, as argued by Bloom (2014). I compute the impulse response path of the economy as it transitions back to the steady state (under perfect foresight). In the current version I fix the interest rate at the initial equilibrium level. Figure 3.5 presents impulse responses of output, employment, investment, credit to GDP ratio, consumption and labor productivity all as percentage deviations from the steady state of the model. The impact on the employment among different groups of firms is summarized in Table 3.8. I draw four important conclusions from this impulse response exercise.

Firstly, a economy hit with an uncertainty shock falls into a recession. Impulse to the economy generates a 1.53% decline in output on impact. Decline in output is driven by the fact that constrained firms in the model are now more constrained and reduce their demand for both capital and labor. This is a direct consequence of tradeoffs faced by the financial intermediary in the dynamic contracting environment with private information. As idiosyncratic uncertainty increases maintaining intertemporal incentives becomes more expansive since, given the strict concavity of the value function, continuation values need to be more spread away. This lowers the value of the contract for the lender and thus resources devoted to provide insurance and in particular to make loans to firms fall, inducing tightening of the borrowing constraint. Importantly, employment falls by 1.35%, which is

⁸I interpret the idiosyncratic shock in my model as demand shock, since it occurs before the production takes place. However, equivalently, in line with arguments by Foster et al. (2012) this shock can be interpreted as firm-level TFP shock.

88% fall relative to the GDP, a number close to the data counterpart. Interestingly, reduction of employment and output is not driven by mis-allocation, measured productivity remains at its initial steady state level. This is contrary to the mechanism common in the literature on financial friction relying on the exogenously imposed collateral constraint (see for example [Zetlin-Jones and Shourideh \(2012\)](#)). There the collateral shock induces a greater degree of mis-allocation relative to the steady state outcome. Constrained firms in those models are the firms that experience positive shock to uninsurable, idiosyncratic productivity and these highly productive firms then grow more slowly along the impulse path than they would have in steady state. In my model mechanism is quite different. The presence of private information induces in times of high, microeconomic uncertainty firms will actually grow (or contract) faster (continuation utilities are more spread away), albeit the path towards unconstrained level is more steep which is reflected in a tighter borrowing constraint. The economic downturn is accompanied by a fall in credit to GDP ratio by roughly 0.3%. Contraction of credit is an ubiquitous feature of recessions in the US data, which my model is able to replicate. Finally, consumption of entrepreneurs increases whereas consumption of workers falls. An immediate drop of workers consumption reflects a combination of decline in wage and employment.

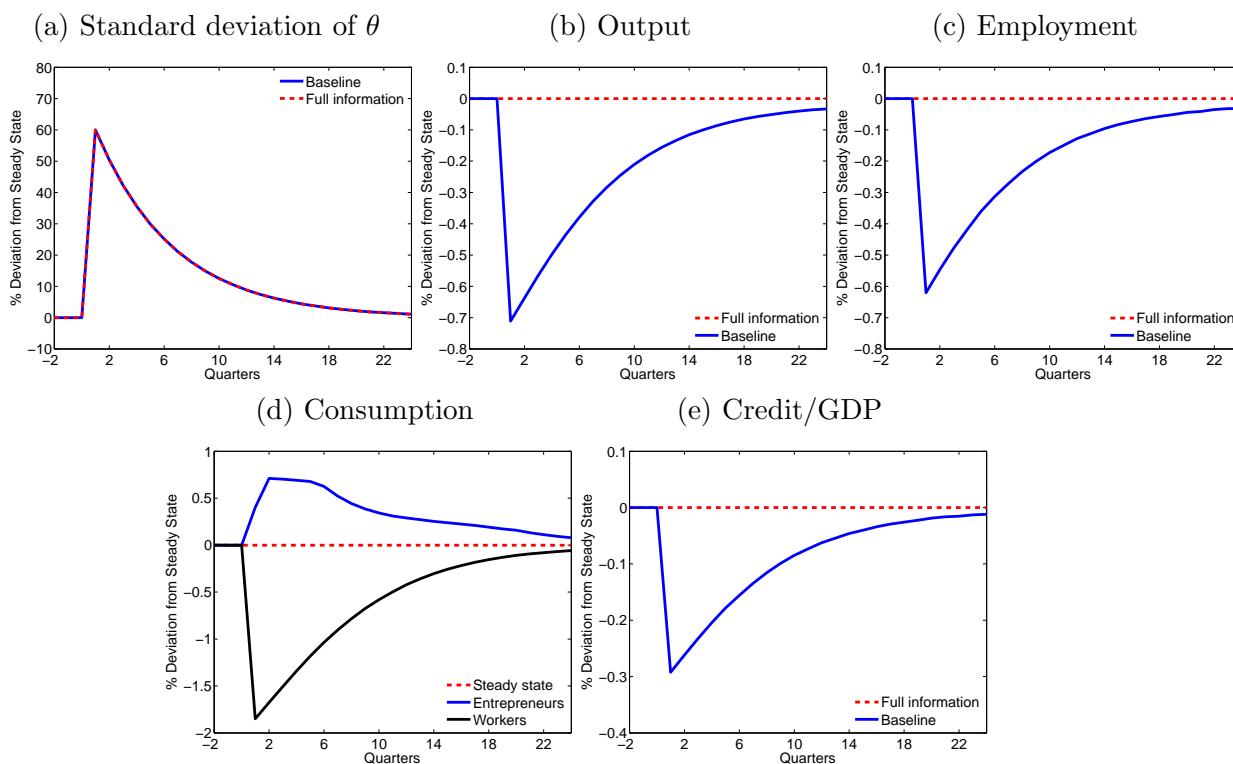
Secondly, the general equilibrium effect counters the initial impact of the uncertainty shock. A reduction in demand for labor by constrained firms imposes downward pressure on wages (interest rate is fixed in my transitional dynamics exercise). Thus the net effect of the two implies whether the macroeconomic aggregates fall or increase. Unconstrained firms who already receive the efficient level of financing facing lower wage rates increase their employment and investment. Thus the composition of firms between constrained (also the tightness of the constraint matters) and unconstrained ones in the economy is crucial for the behavior of the economic aggregates. In my environment this composition is endogenous. Although, I impose discipline on the distribution of firms by calibrating age/size firms and employment distribution observed in the data. Thus implicitly, through the discipline on fundamentals of the economy, my calibration procedure determines the fraction of the unconstrained firms in the economy. It turns out initial impulse dominates the general equilibrium effect and economy experiences recession, which is reflected by the

behavior of macroeconomic aggregates illustrated in Figure 3.5.

My third observation highlights the role of private information in my model as a propagation mechanism. A natural benchmark for baseline experiment is an economy with full information. Dotted lines in Figure 3.5 present impulse responses to the uncertainty shock in an economy with no informational frictions. Uncertainty shock in this economy has absolutely no effects on allocations and prices. The shock is mean preserving, thus it does not change the expected revenue of the project. With no informational friction statically efficient level is sustained and perfect insurance is provided by the financial intermediary at the same levels as in the initial steady state. Hence, shock has no effects on the economy (see also Table 3.8). The economy remains in the initial steady state. Therefore a difference between between solid lines and dotted lines illustrates a role of private information as a mechanism shaping macroeconomic aggregates over transition.

Finally, Table 3.8 compares the response to the uncertainty shock across different groups of firms. My key finding is that employment of young firms. My key finding is that employment of young firms responds 6.4 times more relative to the employment of the old firms. This roughly accounts for a 71.5% of differential response between young and old firms which was observed in the US economy (see Section 3.2.4). However, so far the model is nowhere close to reproduce a huge decline in employment of young firms following the recent recession. There are two channels through which uncertainty shocks triggers asymmetric response. Firstly, young firms are mostly financially constrained and thus since uncertainty shock tightens the endogenous borrowing constraint their employment, investment and therefore output is reduced more. Since the firms in my economy are growing over time and conditional on survival they eventually grow out of the endogenous constraint, a population of the old firms is concentrated more in a region where the financial friction is less severe, i.e. borrowing constraint is less strict or does not bind at all. Moreover, the general equilibrium effect counters the initial impact of the uncertainty shocks. For a fraction of unconstrained firms or those close to the unconstrained level the effect of falling wages is dominating and they expand by increasing employment and investment. Thus age composition, for each size type, of firms over the continuation utilities and hence over the

Figure 3.5: Effects of Uncertainty Shock on macroeconomic outcomes.



access to borrowing is a key feature determining an asymmetric response between young and old firms. Second channel through which uncertainty shocks induce larger reduction of the employment of young firms is reduction of the size of start-ups. The stream of financial intermediary's profit from an individual lending contract is lower in times of heightened uncertainty, therefore a free entry condition implies initial utility of entrepreneur, determining initial size, declines contributing to reduction of employment by young firms. This channel by definition is absent in the group of old firms. The model also generates asymmetric patterns of employment between small and large firms. Recall, I define in line with the data small firms to be less than 100 employees, which in my calibration are types one and two. As for the small vs. large firms margin there are two counteracting forces. Firstly, firms which drew the high average demand will start larger relative to the other start-ups,

since an optimal contract is more profitable and offers (due to zero profit condition) higher initial utility. This pushes the large star-ups closer to the unconstrained level. However, given that standard deviation of the uncertainty shock is common across firms large firms would require longer sequence of high demand shocks to grow out of the constraint. The first force contributes to lower variability of large firms employment relative to the small ones, whereas the second induces higher volatility since it keeps large firms longer in the constrained region. Which one dominates is a quantitative question. Given my calibration I found my model generates employment response of small firms which is 11 percent larger relative to the response of the large, accounting for 60.2% percent of difference in the data (see Table 3.8).

Table 3.8: Effect of shocks on employment (% dev. from steady state)

Age/Size	Data	Uncertainty shock (Baseline)	Uncertainty shock (Full-info)	Productivity shock (1%) (Baseline)
Total	-2.75	-1.35	0.00	-1.71
Young (0-5)	-10.15	-4.61	0.00	-1.71
Old (6+)	-1.19	-0.72	0.00	-1.71
Small (1-99)	-2.48	-1.26	0.00	-1.71
Large (100+)	-2.94	-1.40	0.00	-1.71

3.5.3 The Effects of Productivity Shocks

I now compare the effects of uncertainty shocks to the effects of aggregate productivity shock in my model. This is also a natural benchmark given a popularity of the real business cycle models. This exercise has two purposes. Firstly, it allows me to provide context for the magnitude of the effects of uncertainty shocks. Secondly it allows, within my environment, to disentangle the sources of aggregate fluctuations that are plausible to generate movements of the macroeconomic aggregates and asymmetric response across firms that are observed in recessions in the US. Here, I consider the transition dynamics that results from purely unanticipated decline in the aggregate productivity which slowly returns to

steady state. Similarly to the uncertainty shock I fix the half-life of the impulse to 4 quarters. I discipline the size of the productivity shock so that a measured productivity in the model falls on impact by one standard deviation of measured Solow residual in the United States. This implies, roughly, drop of measured productivity in the model by 1%. On impact employment falls by 1.71%. Hence, uncertainty shock that induces one standard deviation fall in the standard deviation of the microeconomic uncertainty leads to a recession of 79% the size of the one caused by the one standard deviation fall in the aggregate productivity. Employment series exhibit similar ratio. Although, counter to the uncertainty shock and to the data productivity shock does not induce a fall in credit to GDP ratio. Therefore, I consider uncertainty shock as an important contributor to the movements of the macroeconomic aggregates and plausible driving force of economic downturns.

This conclusion is further validated by inspecting response of employment across different groups of firms, which are illustrated in Figure 7 and presented Table 3.8. Economy hit with productivity shock exhibits a symmetrical response of employment across age and size groups of firms, which stands in stark contrast to the uncertainty shock case and more importantly contrary to the data. Aggregate productivity shock induces fall in the expected return on project of all firms, thus given the homogeneity result I prove in Section 3.4, the access to financing falls symmetric for all firms implying symmetric fall of employment. These findings validate the most important conclusion from the quantitative part of my paper i.e. it is a combination of fluctuations in microeconomic uncertainty paired with the private information friction that are essential to account *jointly* for phenomena observed during recessions in the US.

References

- ACIKGOZ, O. (2013): “Transitional Dynamics and Long-run Optimal Taxation Under Incomplete Markets,” MPRA Paper 50160, University Library of Munich, Germany.
- AIYAGARI, S. R. (1994): “Uninsured Idiosyncratic Risk and Aggregate Saving,” *The Quarterly Journal of Economics*, 109, 659–684.
- (1995): “Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting,” *Journal of Political Economy*, 103, 1158–1175.
- AIYAGARI, S. R. AND E. R. MCGRATTAN (1998): “The optimum quantity of debt,” *Journal of Monetary Economics*, 42, 447–469.
- ARELLANO, C., Y. BAI, AND P. J. KEHOE (2012): “Financial frictions and fluctuations in volatility,” Tech. rep.
- ATKESON, A., V. CHARI, AND P. J. KEHOE (1999): “Taxing capital income: a bad idea,” *Quarterly Review*, 3–17.
- ATKESON, A. AND J. LUCAS, ROBERT E (1992): “On Efficient Distribution with Private Information,” *Review of Economic Studies*, 59, 427–53.
- ATKESON, A. AND R. LUCAS (1995): “Efficiency and Equality in a Simple Model of Efficient Unemployment Insurance,” *Journal of Economic Theory*, 66, 64–88.
- BACHMANN, R. AND C. BAYER (2009): “The cross-section of firms over the business cycle: new facts and a DSGE exploration,” Tech. rep.

- BENABOU, R. (2002): “Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?” *Econometrica*, 70, 481–517.
- BERNANKE, B. AND M. GERTLER (1989): “Agency Costs, Net Worth, and Business Fluctuations,” *American Economic Review*, 79, 14–31.
- BERNANKE, B., M. GERTLER, AND S. GILCHRIST (1999): “The Financial Accelerator in a Quantitative Business Cycle Framework,” *Handbook of Macroeconomics*, edited by John B. Taylor and Michael Woodford, pp. 1341-1393, Amsterdam, New York and Oxford, Elsevier Science, North-Holland.
- BEWLEY, T. (1986): “Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers,” in *Contributions to Mathematical Economics in Honor of Gerard Debreu*, ed. by W. Hildenbrand and A. Mas-Collel, Amsterdam: North-Holland.
- BHANDARI, A., D. EVANS, M. GOLOSOV, AND T. J. SARGENT (2013): “Taxes, Debts, and Redistributions with Aggregate Shocks,” NBER Working Papers 19470, National Bureau of Economic Research, Inc.
- BLOOM, N. (2014): “Fluctuations In Uncertainty,” Working Papers 14-17, Center for Economic Studies, U.S. Census Bureau.
- BLOOM, N., M. FLOETOTTO, N. JAIMOVICH, I. SAPORTA-EKSTEN, AND S. J. TERRY (2014): “Really Uncertain Business Cycles,” Working Papers 14-18, Center for Economic Studies, U.S. Census Bureau.
- BOUARICHA, A. AND R. SCHNABEL (1994): “TENSOLVE: A Software Package for Solving Systems of Nonlinear Equations and Nonlinear Least Squares Problems Using Tensor Methods,” Tech. rep., Preprint MCS-P463-0894, Mathematics and Computer Science Division, Argonne National Laboratory.
- BRUNNERMEIER, M. K., T. M. EISENBACH, AND Y. SANNIKOV (2012): “Macroeconomics with Financial Frictions: A Survey,” Levine’s Working Paper Archive 78696900000000384, David K. Levine.

- BUERA, F. J. AND Y. SHIN (2013): “Financial Frictions and the Persistence of History: A Quantitative Exploration,” *Journal of Political Economy*, 121, 221 – 272.
- CARLSTROM, C. T. AND T. S. FUERST (1997): “Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis,” *American Economic Review*, 87, 893–910.
- CARROLL, C. D. (2006): “The method of endogenous gridpoints for solving dynamic stochastic optimization problems,” *Economics Letters*, 91, 312–320.
- CASTAÑEDA, A., J. DÍAZ-GIMÉNEZ, AND J.-V. RÍOS-RULL (2003): “Accounting for the U.S. Earnings and Wealth Inequality,” *Journal of Political Economy*, 111, 818–857.
- CHAMLEY, C. (1986): “Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives,” *Econometrica*, 54, 607–22.
- (2001): “Capital income taxation, wealth distribution and borrowing constraints,” *Journal of Public Economics*, 79, 55–69.
- CHARI, V. AND P. J. KEHOE (1999): “Optimal fiscal and monetary policy,” in *Handbook of Macroeconomics*, ed. by J. B. Taylor and M. Woodford, Elsevier, vol. 1 of *Handbook of Macroeconomics*, chap. 26, 1671–1745.
- CHARI, V. V., L. CHRISTIANO, AND P. KEHOE (2008): “On the Behavior of Small and Large Firms over the Business Cycle,” Working Paper, Research Department, Federal Reserve Bank of Minneapolis.
- CHATTERJEE, S. (1994): “Transitional dynamics and the distribution of wealth in a neoclassical growth model,” *Journal of Public Economics*, 54, 97–119.
- CHETTY, R., A. GUREN, D. MANOLI, AND A. WEBER (2011): “Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins,” *American Economic Review*, 101, 471–75.
- (2012): “Does Indivisible Labor Explain the Difference between Micro and Macro

- Elasticities? A Meta-Analysis of Extensive Margin Elasticities,” in *NBER Macroeconomics Annual 2012, Volume 27*, National Bureau of Economic Research, Inc, NBER Chapters, 1–56.
- CHRISTIANO, L. J., R. MOTTO, AND M. ROSTAGNO (2014): “Risk Shocks,” *American Economic Review*, 104, 27–65.
- CLEMENTI, G. L. AND H. HOPENHAYN (2006): “A Theory of Financing Constraints and Firm Dynamics,” *Quarterly Journal of Economics*, 121, 229–265.
- CONESA, J. C., S. KITAO, AND D. KRUEGER (2009a): “Taxing Capital? Not a Bad Idea after All!” *American Economic Review*, 99, 25–48.
- (2009b): “Taxing Capital? Not a Bad Idea after All!” *American Economic Review*, 99, 25–48.
- COOLEY, T., R. MARIMON, AND V. QUADRINI (2004): “Aggregate Consequences of Limited Contract Enforceability,” *Journal of Political Economy*, 112, 817–847.
- DÁVILA, J., J. H. HONG, P. KRUSELL, AND J.-V. RÍOS-RULL (2012): “Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks,” *Econometrica*, 80, 2431–2467.
- DAVILA, J., J. H. HONG, P. KRUSELL, AND J. V. R. RULL (2012): “Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks,” *Econometrica*, 80, 2431–2467.
- DEMARZO, P. M. AND M. J. FISHMAN (2007): “Agency and Optimal Investment Dynamics,” *Review of Financial Studies*, 20, 151–188.
- DÍAZ-GIMÉNEZ, J., A. GLOVER, AND J.-V. RÍOS-RULL (2011): “Facts on the distributions of earnings, income, and wealth in the United States: 2007 update,” *Quarterly Review*.
- DOMEIJ, D. AND J. HEATHCOTE (2004): “On The Distributional Effects Of Reducing Capital Taxes,” *International Economic Review*, 45, 523–554.

- EROSA, A. AND M. GERVAIS (2002): “Optimal Taxation in Life-Cycle Economies,” *Journal of Economic Theory*, 105, 338–369.
- EVANS, D. (2014): “Optimal Taxation with Persistent Idiosyncratic Investment Risk,” mimeo.
- FARHI, E. AND I. WERNING (2012): “Capital Taxation: Quantitative Explorations of the Inverse Euler Equation,” *Journal of Political Economy*, 120, 000 – 000.
- (2013): “Insurance and Taxation over the Life Cycle,” *Review of Economic Studies*, 80, 596–635.
- FLODEN, M. (2001): “The effectiveness of government debt and transfers as insurance,” *Journal of Monetary Economics*, 48, 81–108.
- FORT, T. C., J. HALTIWANGER, R. S. JARMIN, AND J. MIRANDA (2013): “How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size,” Working Papers 13-30, Center for Economic Studies, U.S. Census Bureau.
- FOSTER, L., J. HALTIWANGER, AND C. SYVERSON (2012): “The Slow Growth of New Plants: Learning about Demand?” Working Papers 12-06, Center for Economic Studies, U.S. Census Bureau.
- GERTLER, M. AND S. GILCHRIST (1994): “Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms,” *The Quarterly Journal of Economics*, 109, 309–40.
- GILCHRIST, S., J. W. SIM, AND E. ZAKRAJSEK (2014): “Uncertainty, Financial Frictions, and Investment Dynamics,” NBER Working Papers 20038, National Bureau of Economic Research, Inc.
- GOLOSOV, M., N. KOCHERLAKOTA, AND A. TSYVINSKI (2003): “Optimal Indirect and Capital Taxation,” *Review of Economic Studies*, 70, 569–587.
- GOTTARDI, P., A. KAJII, AND T. NAKAJIMA (2014a): “Constrained Inefficiency and Optimal Taxation with Uninsurable Risks,” KIER Working Papers 694, Kyoto University, Institute of Economic Research.

- (2014b): “Optimal taxation and debt with uninsurable risks to human capital accumulation,” CIGS Working Paper Series 14-007E, The Canon Institute for Global Studies.
- GREEN, E. J. (1987): “Lending and the Smoothing of Uninsurable Income,” *Contractual Arrangements for Intertemporal Trade*, 3–25.
- GREENWOOD, J., Z. HERCOWITZ, AND G. W. HUFFMAN (1988): “Investment, Capacity Utilization, and the Real Business Cycle,” *American Economic Review*, 78, 402–17.
- GUVENEN, F. (2011): “Macroeconomics With Heterogeneity: A Practical Guide,” NBER Working Papers 17622, National Bureau of Economic Research, Inc.
- GUVENEN, F. AND A. SMITH (2013): “Inferring labor income risk and partial insurance from economic choices,” Staff Report 485, Federal Reserve Bank of Minneapolis.
- HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2010a): “The Macroeconomic Implications of Rising Wage Inequality in the United States,” *Journal of Political Economy*, 118, 681–722.
- (2010b): “The Macroeconomic Implications of Rising Wage Inequality in the United States,” *Journal of Political Economy*, 118, 681–722.
- (2014): “Optimal Tax Progressivity: An Analytical Framework,” NBER Working Papers 19899, National Bureau of Economic Research, Inc.
- HOLTZ-EAKIN, D., D. JOULFAIAN, AND H. S. ROSEN (1993): “The Carnegie Conjecture: Some Empirical Evidence,” *The Quarterly Journal of Economics*, 108, pp. 413–435.
- HUGGETT, M. (1993): “The risk-free rate in heterogeneous-agent incomplete-insurance economies,” *Journal of Economic Dynamics and Control*, 17, 953–969.
- IMBENS, G. W., D. B. RUBIN, AND B. I. SACERDOTE (2001): “Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players,” *American Economic Review*, 91, 778–794.
- IMROHORUGLU, A. (1989): “Cost of Business Cycles with Indivisibilities and Liquidity Constraints,” *Journal of Political Economy*, 97, 1364–83.

- JONES, L. E., R. E. MANUELLI, AND P. E. ROSSI (1997): “On the Optimal Taxation of Capital Income,” *Journal of Economic Theory*, 73, 93–117.
- JUDD, K. L. (1985): “Redistributive taxation in a simple perfect foresight model,” *Journal of Public Economics*, 28, 59–83.
- KOCHERLAKOTA, N. (2000): “Creating Business Cycles through Borrowing Constraints,” *Federal Reserve Bank of Minneapolis Quarterly Review*, 24, 2–10.
- KRUEGER, D. AND A. LUDWIG (2013): “Optimal Progressive Taxation and Education Subsidies in a Model of Endogenous Human Capital Formation,” PIER Working Paper Archive 13-035, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.
- KRUEGER, D. AND F. PERRI (2006): “Does Income Inequality Lead to Consumption Inequality? Evidence and Theory,” *Review of Economic Studies*, 73, 163–193.
- KUCHERENKO, S. AND Y. SYTSKO (2005): “Application of Deterministic Low-Discrepancy Sequences in Global Optimization,” *Computational Optimization and Applications*, 30, 297–318, 10.1007/s10589-005-4615-1.
- MOSCARINI, G. AND F. POSTEL-VINAY (2012): “The Contribution of Large and Small Employers to Job Creation in Times of High and Low Unemployment,” *American Economic Review*, 102, 2509–39.
- NAKAJIMA, M. (2010): “Optimal capital income taxation with housing,” Working Papers 10-11, Federal Reserve Bank of Philadelphia.
- (2012): “Business Cycles In The Equilibrium Model Of Labor Market Search And Self-Insurance,” *International Economic Review*, 53, 399–432.
- PANOUSI, V. AND C. REIS (2012): “Optimal capital taxation with idiosyncratic investment risk,” Tech. rep.
- POSCHKE, M., B. KAYMAK, AND O. BAKIS (2012): “On the Optimality of Progressive Income Redistribution,” 2012 Meeting Papers 837, Society for Economic Dynamics.

- QUADRINI, V. (2004): "Investment and liquidation in renegotiation-proof contracts with moral hazard," *Journal of Monetary Economics*, 51, 713–751.
- (2011): "Financial frictions in macroeconomic fluctuations," *Economic Quarterly*, 209–254.
- RINNOOY KAN, A. H. G. AND G. T. TIMMER (1987): "Stochastic global optimization methods. part 1: clustering methods," *Math. Program.*, 39, 27–56.
- SMITH, A. J. AND C. WANG (2006): "Dynamic credit relationships in general equilibrium," *Journal of Monetary Economics*, 53, 847–877.
- STOCK, J. H. AND M. W. WATSON (2012): "Disentangling the Channels of the 2007-2009 Recession," NBER Working Papers 18094, National Bureau of Economic Research, Inc.
- STRAUB, L. AND I. WERNING (2014): "Positive Long Run Capital Taxation: Chamley-Judd Revisited," NBER Working Papers 20441, National Bureau of Economic Research, Inc.
- THOMAS, J. AND T. WORRALL (1990): "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-agent Problem," *Journal of Economic Theory*, 51, 367–390.
- TRABANDT, M. AND H. UHLIG (2011): "The Laffer curve revisited," *Journal of Monetary Economics*, 58, 305–327.
- TROSHKIN, M., A. TSYVINSKI, AND M. GOLOSOV (2010): "Optimal Dynamic Taxes," 2010 Meeting Papers 320, Society for Economic Dynamics.
- VERANI, S. (2013): "Aggregate Consequences of Dynamic Credit Relationships," Working Paper.
- WERNING, I. (2007): "Optimal Fiscal Policy with Redistribution," *The Quarterly Journal of Economics*, 122, 925–967.
- WINTER, C. AND S. ROEHRS (2014): "Reducing Government Debt in the Presence of Inequality," Tech. rep.

WOLFF, E. N. (2011): “Recent Trends in Household Wealth, 1983-2009: The Irresistible Rise of Household Debt,” *Review of Economics and Institutions*, 2.

ZETLIN-JONES, A. AND A. SHOURIDEH (2012): “External Financing and the Role of Financial Frictions over the Business Cycle: Measurement and Theory,” Tech. rep.

Appendix A

This appendix presents concise versions of the proofs. Extensive versions with more details are contained in a separate online appendix which can be found in our websites.

A.1 Proofs for two-period economies

A.1.1 Uncertainty economy

Define $\tau_R^k \equiv r\tau^k / (1 + r)$. Six equations determine a tax distorted equilibrium $(K, n_L, n_H, r, w; \tau^n, \tau_R^k, T)$ according to Definition 1: the first order conditions of the agent's problem (one intertemporal and two intratemporal), the first order conditions of the firm's problem

$$r = f_K(K, N), \quad \text{and} \quad w = f_N(K, N), \quad \text{where} \quad N = \pi e_L n_L + (1 - \pi) e_H n_H \quad (\text{A.1})$$

and the government's budget constraint. Using equation (A.1) to substitute out for r and w we are left with a system of four equations that any vector $(K, n_L, n_H, \tau^n, \tau_R^k, T)$ of equilibrium values must satisfy. The two degrees of freedom are a result of the fact that the planner has three instruments (τ^n, τ_R^k, T) that are restricted by one equation, the government's budget constraint. Defining welfare by

$$W \equiv u(\omega - K, \bar{n}) + \beta E \left[u \left((1 - \tau^n) f_N(K, N) e_i n_i + \left(1 - \tau_R^k \right) f_K(K, N) K + T \right), n_i \right]$$

and totally differentiating the four equilibrium equations together with this definition and making the appropriate simplifications using Assumption 1 we obtain the following equation

(the algebra is tedious and, therefore, suppressed¹):

$$dW = \Theta^n d\tau^n + \Theta^k d\tau_R^k,$$

where Θ^n and Θ^k are complicated functions of equilibrium variables².

Lemma 9. *Under Assumption 1, in equilibrium $n_H > n_L$ and $u_c(c_L, n_L) > u_c(c_H, n_H)$.*

The proof of this Lemma is contained in the online appendix.

Proof of Proposition 1. First notice that the optimal tax system must satisfy $\Theta^n = 0$ and $\Theta^k = 0$, otherwise there would exist variations in $(\tau^n, \tau_R^k) \in (-\infty, 1)^2$ that would increase welfare. $\Theta^k = 0$ simplifies to $\theta_1^k + \theta_2^k \tau^n + \theta_3^k \tau_R^k = 0$ where

$$\theta_1^k \equiv f_N f_{KN} N (V_c - U_c), \quad \theta_2^k \equiv f_N f_{KN} N ((1 + \kappa) U_c - V_c), \quad \text{and} \quad \theta_3^k \equiv f_K (f_N + f_{KN} K \kappa) U_c.$$

Solving this equation for τ_R^k , substituting it in $\Theta^n = 0$ and simplifying entails

$$V_c (1 - \tau^n) - U_c (1 - (1 + \kappa) \tau^n) = 0.$$

Solving for τ^n we obtain equation (2.1) and substituting it back in the equation for τ_R^k we obtain $\tau_R^k = 0$; and, therefore, $\tau^k = 0$. This is the only pair $(\tau^n, \tau_R^k) \in (-\infty, 1)^2$ that solves

¹Mathematica codes that compute all the algebraic steps are available upon request.

²Here are the exact formulas:

$$\begin{aligned} \Theta^k &\equiv \frac{f_K K U_c}{\Phi} \left\{ f_N f_{KN} N [(1 - \tau^n) (V_c - U_c) + \tau^n \kappa U_c] + \tau_R^k f_K (f_N + f_{KN} K \kappa) U_c \right\}, \\ \Theta^n &\equiv \frac{f_N N}{(1 - \tau^n) \Phi} \left\{ (1 - \tau_R^k) f_K^2 f_N K \left[(1 - \tau^n) (U_{cc} (U_c - V_c) + \tau_R^k (V_{cc} - U_{cc}) U_c) - (1 - \tau_R^k) \tau^n \kappa U_{cc} U_c \right] \right. \\ &\quad \left. + f_N [(1 - \tau^n) (V_c - U_c) + \tau^n \kappa U_c] \left[(1 - \tau_R^k) f_{KN} N U_c - K u_{cc}^0 \right] + (1 - \tau_R^k) \tau_R^k f_{KN} f_K K \kappa U_c^2 \right\}. \end{aligned}$$

where

$$\begin{aligned} U_c &\equiv \beta [\pi u_c(c_L, n_L) + (1 - \pi) u_c(c_H, n_H)], \quad U_{cc} \equiv \beta [\pi u_{cc}(c_L, n_L) + (1 - \pi) u_{cc}(c_H, n_H)], \\ V_c &\equiv \beta \left[\pi u_c(c_L, n_L) \frac{e_L n_L}{N} + (1 - \pi) u_c(c_H, n_H) \frac{e_H n_H}{N} \right], \quad V_{cc} \equiv \beta \left[\pi u_{cc}(c_L, n_L) \frac{e_L n_L}{N} + (1 - \pi) u_{cc}(c_H, n_H) \frac{e_H n_H}{N} \right], \\ \Phi &\equiv (1 - \tau_R^k) (f_K f_N f_{KN} K N ((1 - \tau^n) (V_{cc} - U_{cc}) + \tau^n \kappa U_{cc}) + (f_N + f_{KN} K \kappa) f_K^2 K U_{cc} - f_N f_{KN} N U_c) \\ &\quad + (f_N + f_{KN} K \kappa) K u_{cc}^0. \end{aligned}$$

the system $\Theta^n = 0$ and $\Theta^k = 0$. The fact that the optimal level of $\tau^n > 0$ follows from Lemma 9. \square

A.1.2 Inequality economy

The proof of Proposition 2 is entirely analogous and for that reason suppressed here. It can be found in the online appendix.

A.2 Proofs for complete market economies

The proofs follow straight-forwardly the approach introduced by [Werning \(2007\)](#). Hence, for details on the logic behind the procedure we refer the reader to that paper, here we focus mainly on the parts that comprise our value added. We depart from [Werning \(2007\)](#) in following ways: we use the GHH utility function (whereas he studies the separable and Cobb-Douglas cases), we do not allow the Ramsey planner to choose time zero policies and impose an upper bound of 1 for capital income taxes. These departures make the Ramsey planner's problem comparable to our benchmark experiment. The restriction on time zero policies is particularly important because it prevents the planner from confiscating the (potentially unequal) initial capital levels eliminating the corresponding redistribution motives.

Consider Economy 4 as described in Section 2.6. For simplicity, we assume that agents are divided into a finite number of types $i \in I$ of relative size π_i . Type i has an initial asset position of $a_{i,0}$ and a productivity level of e_i . Let p_t denote the price of the consumption good in period t in terms of period 0. Since markets are complete we can write down the present value budget constraint of the agent (remember that τ^c is a parameter),

$$\sum_{t=0}^{\infty} p_t ((1 + \tau^c) c_{i,t} + a_{i,t+1}) \leq \sum_{t=0}^{\infty} p_t ((1 - \tau_t^n) w_t e_i n_{i,t} + R_t a_{i,t} + T_t),$$

where $R_t \equiv 1 + (1 - \tau_t^k) r_t$. Rule out arbitrage opportunities by setting $p_t = R_{t+1} p_{t+1}$, and

define $T \equiv \sum_{t=0}^{\infty} p_t T_t$. Then, the budget constraint simplifies to

$$\sum_{t=0}^{\infty} p_t ((1 + \tau^c) c_{i,t} - (1 - \tau_t^n) w_t e_i n_{i,t}) \leq R_0 a_{i,0} + T. \quad (\text{A.2})$$

Similarly, the government's budget constraint simplifies to

$$R_0 B_0 + T + \sum_t p_t G = \sum_t p_t (\tau^c C_t + \tau_t^n w_t N_t + \tau_t^k r_t K_t). \quad (\text{A.3})$$

The resource constraint is given by

$$C_t + G + K_{t+1} = f(K_t, N_t), \quad \text{for all } t \geq 0. \quad (\text{A.4})$$

Definition 12. Given $\{a_{i,0}\}$, K_0 , B_0 and $(\tau_0^n, \tau_0^k, T_0)$, a competitive equilibrium is a policy $\{\tau_t^n, \tau_t^k, T_t\}_{t=1}^{\infty}$, a price system $\{p_t, w_t, r_t\}_{t=0}^{\infty}$, and an allocation $\{c_{i,t}, n_{i,t}, K_{t+1}\}_{t=0}^{\infty}$, such that: (i) agents choose $\{c_{i,t}, n_{i,t}\}_{t=0}^{\infty}$ to maximize utility subject to budget constraint (A.2) taking policies and prices (that satisfy $p_t = R_{t+1} p_{t+1}$) as given; (ii) firms maximize profits; (iii) the government's budget constraint (A.3) holds; and (iv) markets clear: the resource constraints (A.4) hold.

Given aggregate levels C_t and N_t , individual consumption and labor supply levels can be found by solving the following static subproblem

$$U^m(C_t, N_t; \varphi) \equiv \max_{c_{i,t}, n_{i,t}} \sum_i \pi_i \varphi_i u(c_{i,t}, n_{i,t}) \quad \text{s.t.} \quad \sum_i \pi_i c_{i,t} = C_t \quad \text{and} \quad \sum_i \pi_i e_i n_{i,t} = N_t \quad (\text{A.5})$$

where u is given by equation (2.4), for some vector $\varphi \equiv \{\varphi_i\}$ of market weights $\varphi_i \geq 0$. Let

$c_{i,t}^m(C_t, N_t; \varphi)$, and $n_{i,t}^m(C_t, N_t; \varphi)$ be the *argmax* of this problem. It can be shown that³

$$\begin{aligned} c_{i,t}^m(C_t, N_t; \varphi) &= \omega_i^c C_t + \chi \frac{\kappa}{1 + \kappa} \left((\omega_i^n)^{\frac{1+\kappa}{\kappa}} - \omega_i^c \Omega^n \right) (N_t)^{\frac{1+\kappa}{\kappa}} \\ n_{i,t}^m(C_t, N_t; \varphi) &= \omega_i^n e_i N_t \\ U^m(C_t, N_t; \varphi) &= \frac{\Omega^c}{1 - \sigma} \left(C_t - \Omega^n \chi \frac{\kappa}{1 + \kappa} (N_t)^{\frac{1+\kappa}{\kappa}} \right)^{1-\sigma} \end{aligned}$$

Then, implementability constraints can be written as

$$\begin{aligned} \sum_{t=0}^{\infty} \beta^t (U_C^m(C_t, N_t; \varphi) c_{i,t}^m(C_t, N_t; \varphi) + U_N^m(C_t, N_t; \varphi) n_{i,t}^m(C_t, N_t; \varphi)) \\ = U_C^m(C_0, N_0; \varphi) \left(\frac{R_0 a_{i,0} + T}{1 + \tau^c} \right) \quad \text{for all } i \in I \end{aligned} \quad (\text{A.6})$$

Proposition 10. *An aggregate allocation $\{C_t, N_t, K_{t+1}\}_{t=0}^{\infty}$ can be supported by a competitive equilibrium if and only if the resource constraints (A.4) hold and there exist market weights φ and a lump-sum tax T so that the implementability conditions (A.6) hold for all $i \in I$. Individual allocations can then be computed using functions $c_{i,t}^m$ and $n_{i,t}^m$, prices and taxes can be computed using the usual equilibrium conditions.*

The Ramsey problem is that of choosing policies $\{\tau_t^n, \tau_t^k, T_t\}_{t=1}^{\infty}$, taking $\{a_{i,0}\}$, K_0 , B_0 and $(\tau_0^n, \tau_0^k, T_0)$ as given, to maximize a weighted sum of the individual utilities,

$$\sum_{t=0}^{\infty} \beta^t \pi_i \lambda_i u(c_{i,t}, n_{i,t}), \quad (\text{A.7})$$

where $\{\lambda_i\}$ are the welfare weights normalized so that $\sum_i \pi_i \lambda_i = 1$ with $\lambda_i \geq 0$, subject to allocations and policies being a part of a competitive equilibrium and $\tau_t^k \leq 1$ for all $t \geq 1$.

First notice that in equilibrium it must be that $U_C^m(t) = \beta (1 + (1 - \tau_{t+1}^k) r_{t+1}) U_C^m(t+1)$,

³Where constants are defined as follows:

$$\omega_i^c \equiv \frac{(\varphi_i)^{\frac{1}{\sigma}}}{\sum_j \pi_j (\varphi_j)^{\frac{1}{\sigma}}}, \quad \omega_i^n \equiv \frac{(e_i)^\kappa}{\sum_j \pi_j (e_j)^{1+\kappa}}, \quad \Omega^c \equiv \left(\sum_i \pi_i (\varphi_i)^{\frac{1}{\sigma}} \right)^\sigma, \quad \text{and} \quad \Omega^n \equiv \left(\sum_j \pi_j (e_j)^{1+\kappa} \right)^{-\frac{1}{\kappa}}$$

so that

$$U_C^m(t) \geq \beta U_C^m(t+1), \quad (\text{A.8})$$

is equivalent to $\tau_{t+1}^k \leq 1$. Moreover, notice that τ_0^k and T_0 have not been substituted out in the implementability constraint. The fact that τ_0^n is given together with the equilibrium condition $(1 - \tau_0^n) w_0 = -U_N^m(0) / U_C^m(0)$ is equivalent to

$$N_0 = \bar{N}_0, \quad (\text{A.9})$$

where \bar{N}_0 is defined implicitly as a function of variables given to the Ramsey planner,

$$(1 - \tau_0^n) f_N(K_0, \bar{N}_0) = \Omega^n \chi(\bar{N}_0)^{\frac{1}{\kappa}}.$$

Finally, we can use Proposition 10 to rewrite the Ramsey problem as that of choosing $\{C_t, N_t\}_{t=0}^\infty$, T , and φ to maximize (A.7) subject to (A.4) for all $t \geq 0$, (A.6) for all $i \in I$ with multiplier μ_i , (A.8) for all $t \geq 0$ with multiplier η_t , and (A.9). Equivalently, we can write it as that of solving the following auxiliary problem

$$\max_{\{C_t, N_t\}_{t=0}^\infty, T, \varphi} \sum_{t=0}^{\infty} \beta^t W(C_t, N_t; \varphi, \mu, \lambda) - U_C^m(C_0, N_0; \varphi) \sum_{i \in I} \mu_i \left(\frac{R_0 a_{i,0} + T}{1 + \tau^c} \right),$$

subject to (A.4) for all $t \geq 0$, (A.8) for all $t \geq 0$, and (A.9), where

$$\begin{aligned} W(C_t, N_t; \varphi, \mu, \lambda) \equiv & \sum_i \pi_i \{ \lambda_i u(c_{i,t}^m(C_t, N_t; \varphi), n_{i,t}^m(C_t, N_t; \varphi)) \\ & + \mu_i (U_C^m(C_t, N_t; \varphi) c_{i,t}^m(C_t, N_t; \varphi) + U_N^m(C_t, N_t; \varphi) n_{i,t}^m(C_t, N_t; \varphi)) \}. \end{aligned}$$

With some algebra it can be shown that⁴

$$W(C_t, N_t; \varphi, \mu, \lambda) = \frac{1}{1-\sigma} \left(C_t - \Omega^n \chi \frac{\kappa}{1+\kappa} (N_t)^{\frac{1+\kappa}{\kappa}} \right)^{-\sigma} \left(\Phi C_t - (\Phi + (1-\sigma)\Psi) \Omega^n \chi \frac{\kappa}{1+\kappa} (N_t)^{\frac{1+\kappa}{\kappa}} \right) \quad (\text{A.10})$$

Define $R_t^* \equiv 1 + r_t$ and

$$\eta_{-1} \equiv \frac{R_0}{\beta(1+\tau^c)} \sum_i \pi_i \mu_i a_{i,0},$$

and first order conditions (for the following proofs we need only necessary conditions) together with equilibrium conditions imply the following equations⁵

$$\sum_i \pi_i \mu_i = 0 \quad (\text{A.11})$$

$$\frac{\tau_t^n + \tau^c}{1 + \tau^c} = \frac{\Psi \Theta_t}{\Phi \Theta_t + \Psi(\sigma + \Theta_t) + \Upsilon_t \sigma (\beta \eta_{t-1} - \eta_t)}, \quad \text{for } t \geq 1 \quad (\text{A.12})$$

$$\frac{R_{t+1}}{R_{t+1}^*} = \frac{\Phi \Theta_{t+1} + \Psi \sigma + \Upsilon_{t+1} \sigma (\beta \eta_t - \eta_{t+1})}{\Phi \Theta_t + \Psi \sigma + \Upsilon_t \sigma (\beta \eta_{t-1} - \eta_t)} \frac{\Theta_t}{\Theta_{t+1}}, \quad \text{for } t \geq 0 \quad (\text{A.13})$$

Notice that $\Upsilon_t > 0$ and $\Theta_t > 0$, for all $t \geq 0$.

A.2.1 Economy 2

Lemma 10. *If $e_i = 1$ for all $i \in I$, then $\Psi = 0$ and $\Phi > 0$.*

Proof. If $e_i = 1$ for all $i \in I$, then it follows from the definition of Ψ that

$$\Psi = \frac{\Omega^c \sum_j \pi_j \mu_j (e_j)^{1+\varepsilon}}{\varepsilon \sum_j \pi_j (e_j)^{1+\varepsilon}} = \frac{\Omega^c \sum_j \pi_j \mu_j}{\varepsilon \sum_j \pi_j} = 0$$

⁴Where constants are defined as follows:

$$\Phi \equiv \sum_j \pi_j \left(\frac{\lambda_j}{\varphi_j} - \sigma \mu_j \omega_j^c \right), \quad \text{and} \quad \Psi \equiv \frac{\Omega^c}{\kappa} \sum_j \pi_j \mu_j e_j \omega_j^n.$$

⁵Where $\Upsilon_t \equiv \Omega^c / \Omega^n \chi \frac{\kappa}{1+\kappa} (N_t)^{\frac{1+\kappa}{\kappa}}$.

where the last equality follows from equation (A.11). Next, notice that

$$u(c_{i,t}^m(C_t, N_t; \varphi), n_{i,t}^m(C_t, N_t; \varphi)) = \frac{(\omega_i^c)^{1-\sigma}}{1-\sigma} \left(C_t - \Omega^n \chi \frac{\kappa}{1+\kappa} (N_t)^{\frac{1+\kappa}{\kappa}} \right)^{1-\sigma}$$

and, therefore, the solution to the problem must satisfy $C_t > \Omega^n \chi \frac{\kappa}{1+\kappa} (N_t)^{\frac{1+\kappa}{\kappa}}$ for all $t \geq 0$. Otherwise, the objective function of the Ramsey problem would be $-\infty$. On the other hand, since $\Psi = 0$, it follows from equation (A.10) that

$$W(C_t, N_t; \varphi, \mu, \lambda) = \frac{\Phi}{1-\sigma} \left(C_t - \Omega^n \chi \frac{\kappa}{1+\kappa} (N_t)^{\frac{1+\kappa}{\kappa}} \right)^{1-\sigma}.$$

It follows that, if $\Phi \leq 0$, setting $C_0 = f(K_0, \bar{N}_0) - G$, and $C_t = N_t = 0$, for all $t \geq 1$ (so that $C_t = \Omega^n \chi \frac{\kappa}{1+\kappa} (N_t)^{\frac{1+\kappa}{\kappa}}$ for all $t \geq 1$) would maximize the objective function of the auxiliary problem while being feasible which is a contradiction. \square

Proof of Proposition 4. Using Lemma 10, from equation (A.12) it follows that

$$\tau_t^n = -\tau^c, \text{ for } t \geq 1.$$

Next, suppose $\eta_t = 0$, for all $t \geq 0$. Then, it follows from (A.13) that $\tau_1^k < 1$ if

$$-\frac{1}{\beta} \frac{\Phi \Theta_0}{\Upsilon_0 \sigma} \equiv P_1 < \eta_{-1} < M_1 \equiv \frac{1}{\beta} \frac{(R_1^* - 1) \Phi \Theta_0}{\Upsilon_0 \sigma},$$

and that $\tau_t^k = 0$ for $t \geq 2$. Hence, if $P_1 < \eta_{-1} < M_1$, the constraints will in fact never be binding. Now, suppose $\eta_t > 0$, for $t \leq t^* - 2$ and $\eta_t = 0$, for all $t \geq t^* - 1$, then it follows from (A.13) that $\tau_{t^*}^k < 1$ if

$$-\sum_{\tau=1}^{t^*} \frac{1}{\beta^\tau} \frac{\Phi \Theta_{\tau-1}}{\Upsilon_{\tau-1} \sigma} \equiv P_{t^*} < \eta_{-1} < M_{t^*} \equiv \sum_{\tau=1}^{t^*} \frac{1}{\beta^\tau} \frac{\left(\prod_{t=\tau}^{t^*} R_t^* - 1 \right) \Phi \Theta_{\tau-1}}{\Upsilon_{\tau-1} \sigma},$$

and that $\tau_t^k = 0$ for $t \geq t^* + 1$. The result follows from the fact that η_{-1} is finite, $\lim_{t \rightarrow \infty} P_t = -\infty$ and $\lim_{t \rightarrow \infty} M_t = \infty$. \square

A.2.2 Economy 3

Proof of Proposition 5. In this economy there is no heterogeneity in initial levels of asset, i.e. $a_{i,0} = a_0$ for all $i \in I$. Then it follows that

$$\eta_{-1} = \frac{R_0}{\beta(1+\tau^c)} \sum_i \pi_i \mu_i a_{i,0} = \frac{R_0}{\beta(1+\tau^c)} a_0 \sum_i \pi_i \mu_i = 0$$

where the last equality follows from equation (A.11). Since here we assume that τ_t^k does not have to be bounded by 1, it follows that $\eta_t = 0$ for all $t \geq 1$. Then, equation (2.6) follows directly from equation (A.12), (2.7) from its derivative with respect to Θ_t , and (2.8) from equations (A.12) and (A.13). \square

A.2.3 Economy 4

Proof of Proposition 6. Equation (2.8) can be established for all $t \geq 1$, by substituting (A.12) into (A.13). The existence of a t^* such that $\eta_t > 0$, for $t < t^* - 1$ and $\eta_t = 0$, for all $t \geq t^* - 1$, follows from a very similar logic to the one used in the proof of Proposition 4, which we suppress here⁶. Hence, for $t \geq t^*$ we can obtain τ_t^n by using (2.6), which follows from (A.12) with $\eta_t = 1$. For the same time period τ_t^k can then be found by using (2.8). Now, having $\tau_{t^*}^n$ we can use the fact that $\tau_t^k = 1$ and (2.8) moving backwards to obtain τ_t^n for $t < t^*$. \square

A.3 Welfare decomposition

Let $v(x) \equiv u(c, n)$ where x is the consumption-labor composite defined in Section 2.5.3 and u is defined in (2.4). Consider a policy reform. Denote by $x_t^R(a_0, e^t)$ the equilibrium consumption-labor composite path of an agent with initial assets a_0 and history of productivities e^t if the reform is implemented. Let $x_t^{NR}(a_0, e^t)$ be the equilibrium path in case there is no reform. The average welfare gain, Δ , that results from implementing

⁶With

$$P_{t^*} \equiv - \sum_{\tau=1}^{t^*} \frac{1}{\beta^\tau} \frac{\Phi \Theta_{\tau-1} + \Psi \sigma}{\Upsilon_{\tau-1} \sigma}, \quad \text{and} \quad M_{t^*} \equiv \sum_{\tau=1}^{t^*} \frac{1}{\beta^\tau} \frac{\left(\prod_{t=\tau}^{t^*} R_t^* - 1 \right) \Phi \Theta_{\tau-1} + \left(\frac{\Theta_{\tau-1}}{\Theta_{t^*}} \prod_{t=\tau}^{t^*} R_t^* - 1 \right) \Psi \sigma}{\Upsilon_{\tau-1} \sigma}$$

the reform is defined as the constant percentage increase to $x_t^{NR}(a_0, e^t)$ that equalizes the (utilitarian) welfare to the value associated with the reform, that is,

$$\int E_0 [U((1 + \Delta) \{x_t^{NR}(a_0, e^t)\})] d\lambda_0(a_0, e_0) = \int E_0 [U(\{x_t^R(a_0, e^t)\})] d\lambda_0(a_0, e_0), \quad (\text{A.14})$$

where λ_0 is the initial distribution over states (a_0, e_0) and $U(\{x_t(a_0, e^t)\}) \equiv \sum_{t=0}^{\infty} \beta^t v(x_t(a_0, e^t)) = \sum_{t=0}^{\infty} \beta^t u(c_t(a_0, e^t), n_t(a_0, e^t))$.

Define

$$X_t^j \equiv \int x_t^j(a_0, e^t) d\lambda_t^j(a_0, e^t), \quad \text{for } j = R, NR.$$

to be the average level of x at each t . Then, the level effect, Δ_L , is

$$U((1 + \Delta_L) \{X_t^{NR}\}) = U(\{X_t^R\}), \quad (\text{A.15})$$

In order to define the other two components we need some previous definitions. Let $\bar{x}^j(a_0, e_0)$ denote the individual consumption-labor certainty equivalent,

$$U(\{\bar{x}^j(a_0, e_0)\}) = E_0 \left[U(\{x_t^j(a_0, e^t)\}) \right], \quad \text{for } j = R, NR, \quad (\text{A.16})$$

(notice that $\bar{x}^j(a_0, e_0)$ can be chosen to be constant) and let \bar{X}^j be the aggregate consumption-labor certainty equivalent,

$$\bar{X}^j = \int \bar{x}^j(a_0, e_0) d\lambda(a_0, e_0), \quad \text{for } j = R, NR. \quad (\text{A.17})$$

The insurance effect, Δ_I , is defined by

$$1 + \Delta_I \equiv \frac{1 - p_{unc}^R}{1 - p_{unc}^{NR}}, \quad \text{where } U((1 - p_{unc}^j) \{X_t^j\}) = U(\{\bar{X}^j\}), \quad (\text{A.18})$$

and the redistribution effect, Δ_R , by

$$1 + \Delta_R \equiv \frac{1 - p_{ine}^R}{1 - p_{ine}^{NR}}, \quad \text{where } U((1 - p_{ine}^j) \{\bar{X}^j\}) = \int U(\{\bar{x}^j(a_0, e_0)\}) d\lambda(a_0, e_0). \quad (\text{A.19})$$

The following proposition holds⁷.

Proof of Proposition 3. First notice that $v(x) \equiv u(c, n)$ where u is the GHH utility function, defined in (2.4), satisfies the following regularity property: there exists a totally multiplicative function h : (i.e. $h(ab) = h(a)h(b)$, and $h(a/b) = h(a)/h(b)$) such that for any scalar α ,

$$v(\alpha x) = h(\alpha) v(x). \quad (\text{A.20})$$

Hence, suppressing the dependence on (a_0, e_0) , we obtain:

$$\begin{aligned} \int E_0 U(\{x_t^R\}) d\lambda_0^R &\stackrel{(\text{A.16})}{=} \int U(\{\bar{x}^R\}) d\lambda_0^R \stackrel{(\text{A.19})}{=} U((1 - p_{ine}^R) \{\bar{X}^R\}) \stackrel{(\text{A.20})}{=} h(1 - p_{ine}^R) U(\{\bar{X}^R\}) \\ &\stackrel{(\text{A.18})}{=} h(1 - p_{ine}^R) U((1 - p_{unc}^R) \{X_t^R\}) \stackrel{(\text{A.20})}{=} h((1 - p_{ine}^R)(1 - p_{unc}^R)) U(\{X_t^R\}) \\ &\stackrel{(\text{A.15})}{=} h((1 - p_{ine}^R)(1 - p_{unc}^R)) U((1 + \Delta_L) \{X_t^{NR}\}) \\ &\stackrel{(\text{A.20})}{=} h((1 + \Delta_L)(1 - p_{ine}^R)(1 - p_{unc}^R)) U(\{X_t^{NR}\}) \\ &\stackrel{(\text{A.20})}{=} h\left((1 + \Delta_L)(1 - p_{ine}^R) \frac{(1 - p_{unc}^R)}{(1 - p_{unc}^{NR})}\right) U((1 - p_{unc}^{NR}) \{X_t^{NR}\}) \\ &\stackrel{(\text{A.18})}{=} h((1 + \Delta_L)(1 + \Delta_I)(1 - p_{ine}^R)) U(\{\bar{X}^{NR}\}) \\ &\stackrel{(\text{A.20})}{=} h\left((1 + \Delta_L)(1 + \Delta_I) \frac{(1 - p_{ine}^R)}{(1 - p_{ine}^{NR})}\right) U((1 - p_{ine}^{NR}) \{\bar{X}^{NR}\}) \\ &\stackrel{(\text{A.19})}{=} h((1 + \Delta_L)(1 + \Delta_I)(1 + \Delta_R)) \int U(\{\bar{x}^{NR}\}) d\lambda_0^{NR} \\ &\stackrel{(\text{A.18})}{=} h((1 + \Delta_L)(1 + \Delta_I)(1 + \Delta_R)) \int E_0 U(\{x_t^{NR}\}) d\lambda_0^{NR} \\ &\stackrel{(\text{A.20})}{=} \int E_0 U((1 + \Delta_R)(1 + \Delta_I)(1 + \Delta_L) \{x_t^{NR}\}) d\lambda_0^{NR}. \end{aligned}$$

The result follows from the definition of Δ in equation (A.14). \square

⁷This result is similar to the one introduced by Benabou (2002) and used in Floden (2001).

A.4 Algorithms

Here we describe the algorithms used to obtain our results.

Algorithm for computing the transition between steady states⁸

1. Solve for the initial stationary equilibrium.
2. Assume the economy converges to a new stationary equilibrium in \bar{t} periods and guess a sequence $K_2, \dots, K_{\bar{t}-1}$.
3. Solve for the new tax on labor such that given $K_2, \dots, K_{\bar{t}-1}$ and the path for the other taxes, government debt is unchanged between $\bar{t} - 1$ and \bar{t} . Compute the associated path for the government debt, $B_1, \dots, B_{\bar{t}-1}$ (for details see the Final Tax Computation section in the online appendix).
4. Solve for the final stationary equilibrium given final tax rates τ^k, τ^n, τ^c and T , and $B_{\bar{t}}$. Compute $K_{\bar{t}}$.
5. Solve for households savings decisions in transition.
6. Update the path of capital, i.e. take the initial stationary distribution over wealth and productivity and use the decision rules computed above to simulate the economy forward. Then, check for market clearing at each date and adjust $K_2, \dots, K_{\bar{t}-1}$ appropriately.
7. If the new sequence for capital is the close to the old, we have found the equilibrium path. Otherwise go back to step 5.
8. Increase \bar{t} until the solution stops changing.

Algorithm for global optimization⁹

1. Sample a large set X of points from a uniform distribution over the domain¹⁰.

⁸This is an extension of the procedure proposed by [Domeij and Heathcote \(2004\)](#). To solve for agent's decision rules we use the endogenous grid method (see [Carroll \(2006\)](#)).

⁹This procedure is described in more detail in [Kucherenko and Sytsko \(2005\)](#).

¹⁰We used pseudo-random numbers from a Sobol sequence which give more efficient results.

2. Evaluate the objective function for all points in X .
3. Select a reduced set X_r with the highest objective function values. Sort the elements of X_r into clusters and run a local¹¹ solver one time for each cluster¹².
4. Use a Bayesian stopping rule to determine whether or not the procedure should be repeated.

¹¹We used an open source local solver called BOBYQA.

¹²See [Rinnooy Kan and Timmer \(1987\)](#).

A.5 Tables and Figures

Table A.1: Distribution of wealth

	Bottom (%)		Quintiles					Top (%)	Gini
	0-5	1st	2nd	3rd	4th	5th	95-100		
Data	-0.1	-0.2	1.1	4.5	11.2	83.4	60.3	0.82	
Model	-0.1	-0.2	1.5	4.7	10.2	83.4	60.8	0.81	

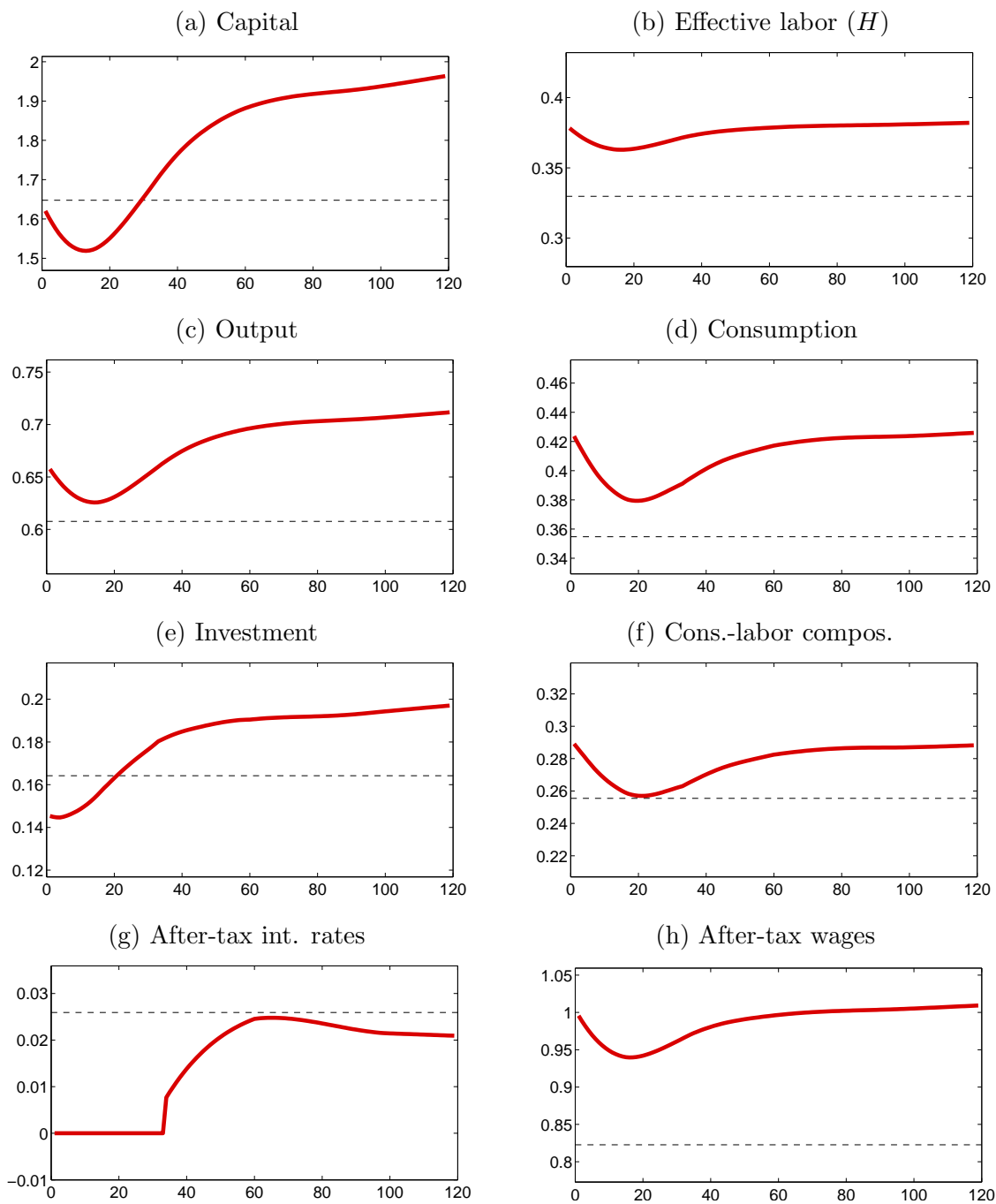
Notes: Data come from the 2007 Survey of the Consumer Finance.

Table A.2: Income sources by quintiles of wealth

Quintile	Model			Data		
	Labor	Asset	Transfer	Labor	Asset	Transfer
1st	83.7	-0.1	16.4	82.0	2.0	16.0
2nd	85.4	1.6	13.1	83.0	4.8	12.2
3rd	84.1	4.7	11.2	80.0	7.3	12.7
4th	81.4	8.6	10.0	77.6	10.3	12.1
5th	58.7	36.2	5.2	51.7	40.0	8.3

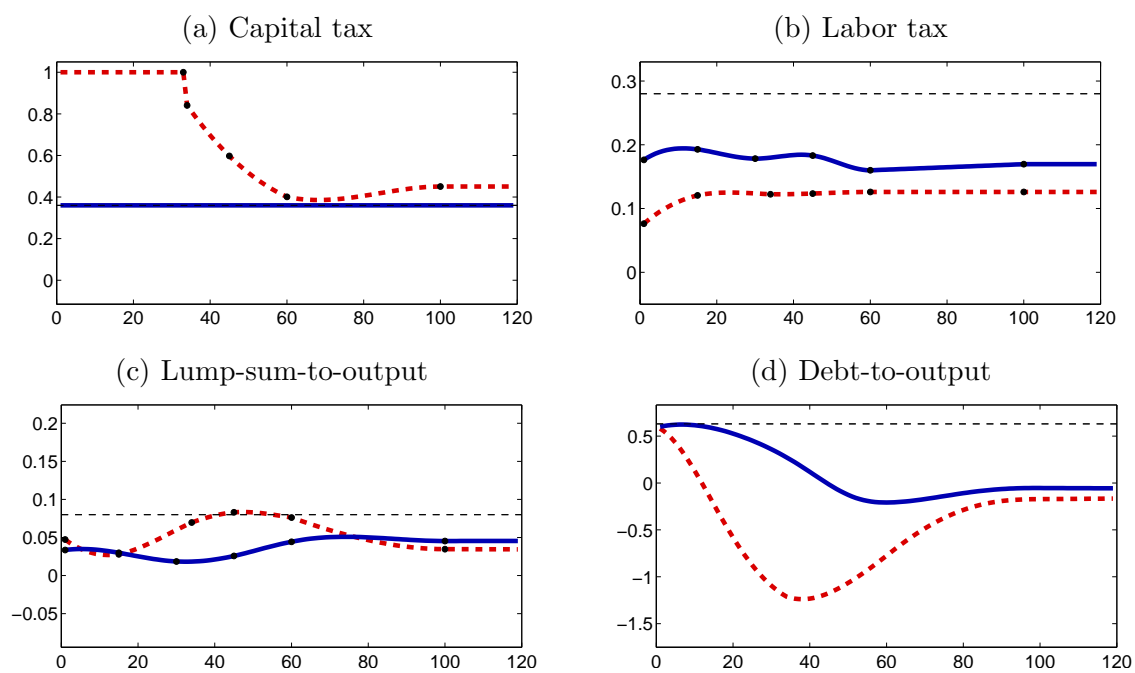
Notes: Table summarizes the pre-tax total income decomposition. We define the asset income as the sum of income from capital and business. Data come from the 2007 Survey of the Consumer Finance, the numbers are based on a summary by [Díaz-Giménez et al. \(2011\)](#).

Figure A.1: Aggregates: Benchmark



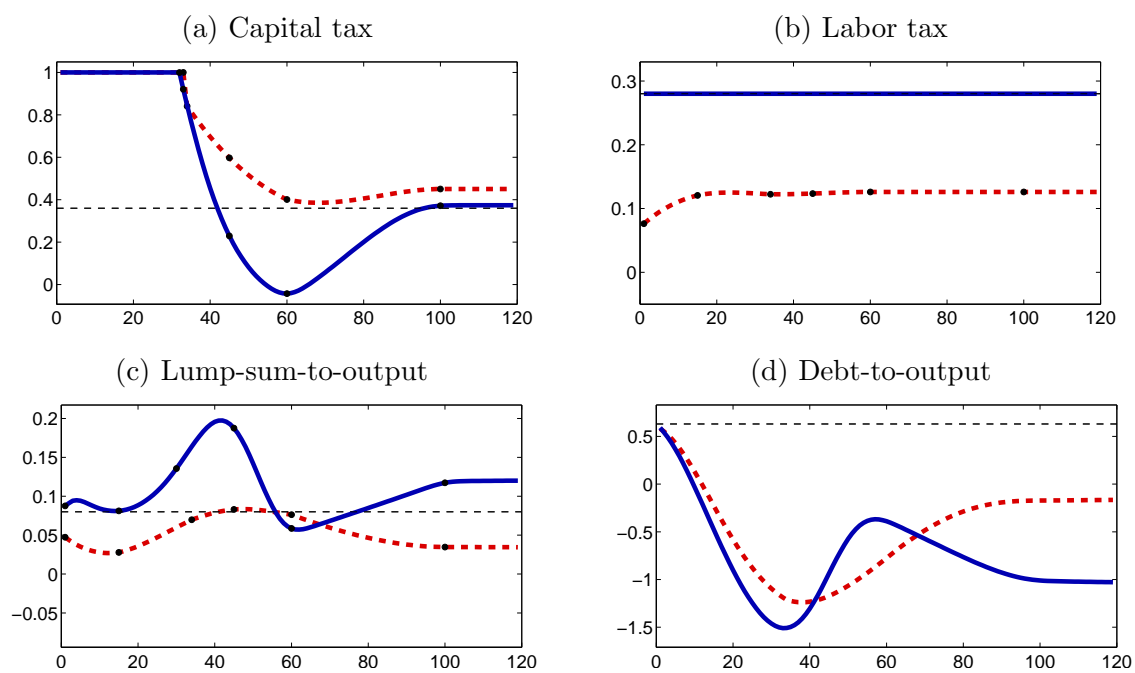
Notes: Dashed line: initial stationary equilibrium; Solid line: optimal transition.

Figure A.2: Optimal Fiscal Policy: Fixed Capital Taxes



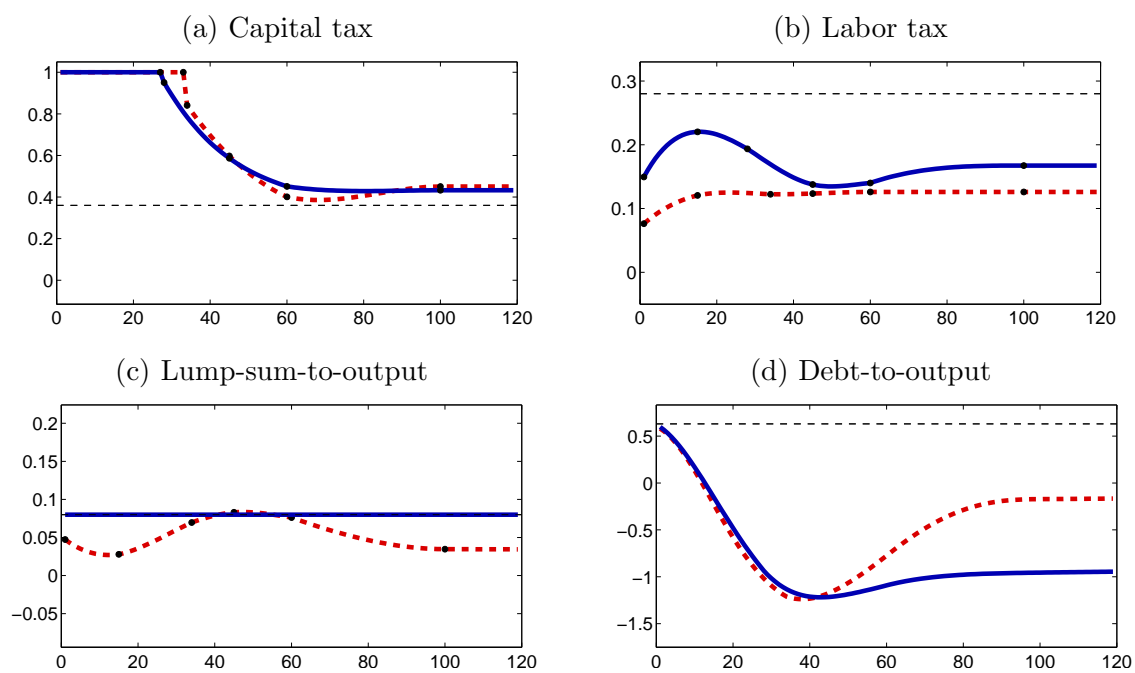
Notes: Dashed thin line: initial stationary equilibrium; Dashed thick line: optimal transition with unrestricted instruments (benchmark); Solid line: optimal transition with fixed capital taxes.

Figure A.3: Optimal Fiscal Policy: Fixed Labor Taxes



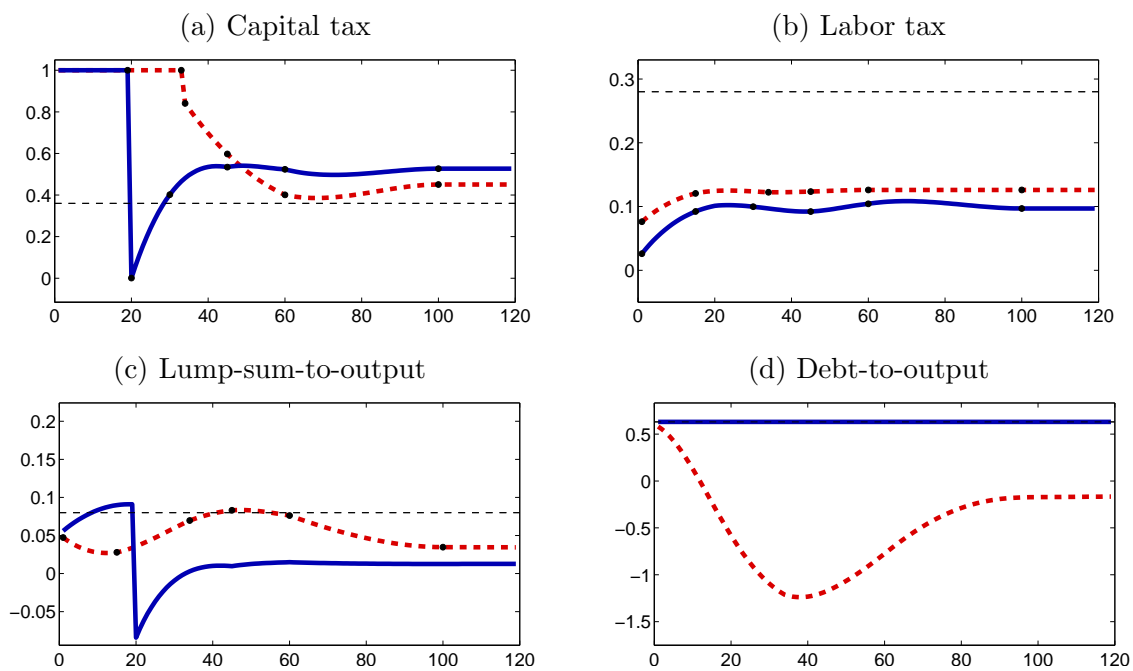
Notes: Dashed thin line: initial stationary equilibrium; Dashed thick line: optimal transition with unrestricted instruments (benchmark); Solid line: optimal transition with fixed labor taxes.

Figure A.4: Optimal Fiscal Policy: Lump-Sum Transfers to Output



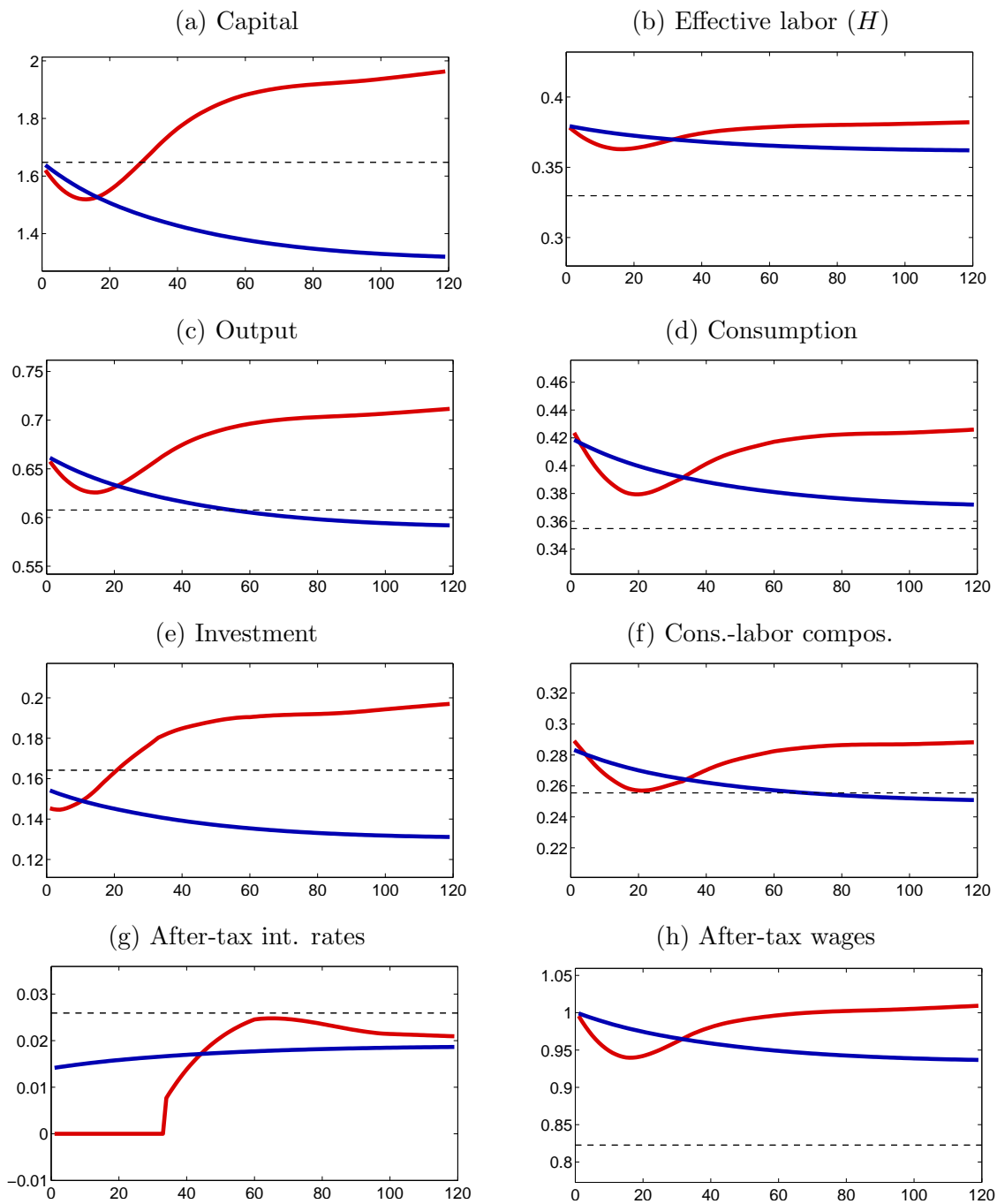
Notes: Dashed thin line: initial stationary equilibrium; Dashed thick line: optimal transition with unrestricted instruments (benchmark); Solid line: optimal transition with fixed lump-sum transfers to output ratio.

Figure A.5: Optimal Fiscal Policy: Fixed debt-to-output



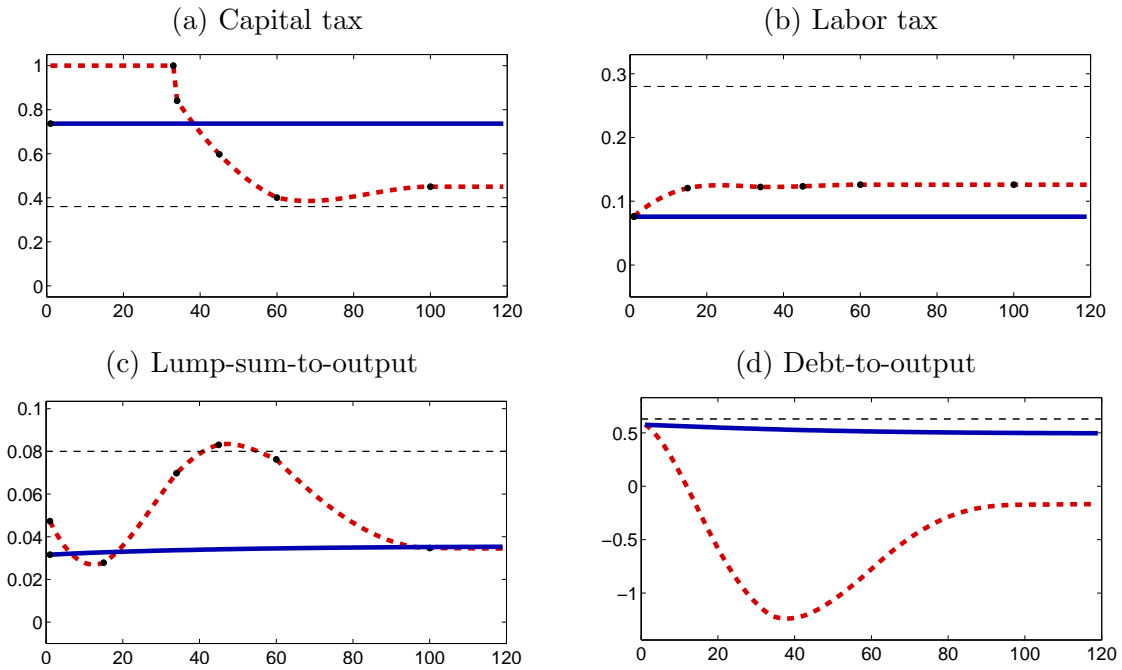
Notes: Dashed thin line: initial stationary equilibrium; Dashed thick line: optimal transition with unrestricted instruments (benchmark); Solid line: optimal transition with fixed debt-to-output ratio.

Figure A.6: Aggregates: Constant Policy

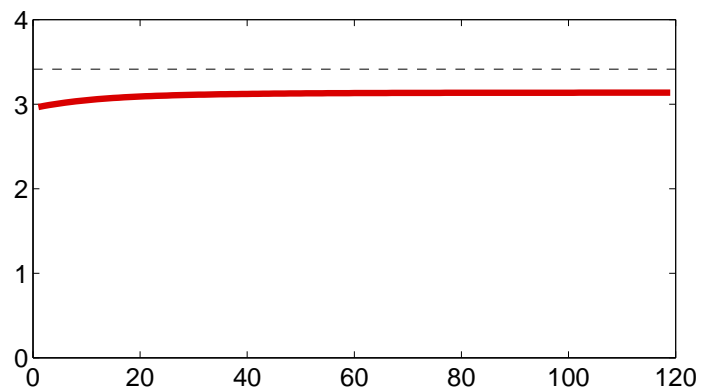


Notes: Dashed line: initial stationary equilibrium; Solid line: optimal transition.

Figure A.7: Optimal Fiscal Policy: Constant Policy



Notes: Dashed line: initial stationary equilibrium; Solid line: optimal transition; The black dots are the choice variables: the spline nodes and t^* , the point at which the capital tax leaves the upper bound.

Figure A.8: Economy 3: Θ_t 

Appendix B

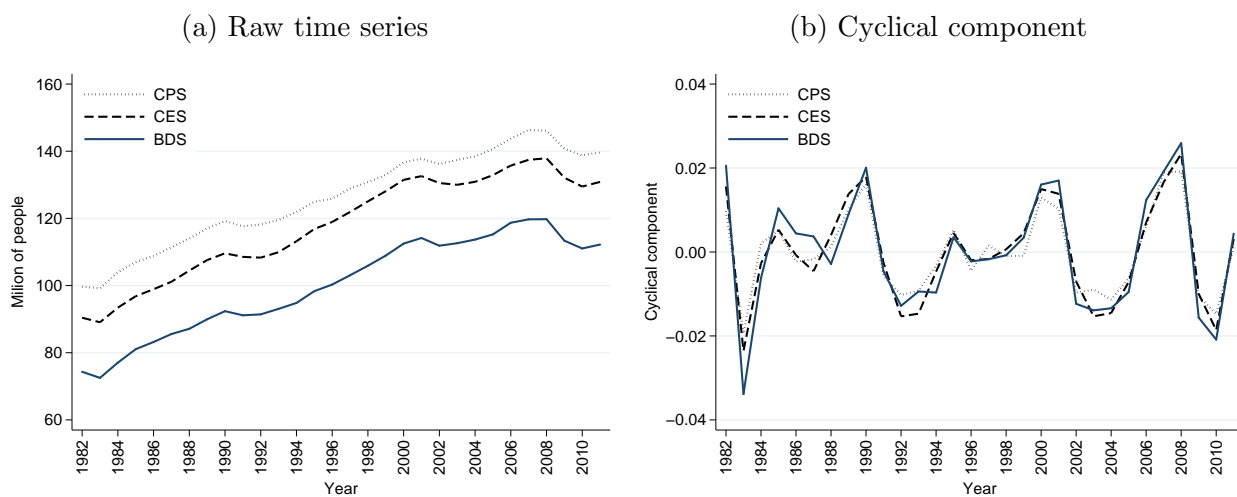
B.1 Data

In this part of Appendix I provide further details complementing facts documented in the empirical section of the paper.

B.1.1 BDS vs. CPS vs. Establishment survey

In the empirical section I argue Business Dynamics Statistics (BDS) data is a useful source of information about the movements of the aggregate employment in the US economy. In order to validate this claim I compare the cyclical properties of the total employment series from the BDS with the two most common sources of the data on employment, i.e. the Current Population Survey (CPS) and Establishment Survey. The average level of employment between 1982 and 2011 in the BDS data amount to 80% and 85% of the average, annual aggregate employment levels from the CPS and Establishment survey respectively. The left panel of Figure [B.1](#) presents the raw time series of all three data sets. Even though there exist significant differences in the coverage of these data sets and thus in the level of the aggregate employment, the cyclical properties are very similar. The right panel of Figure [B.1](#) plots the cyclical components of the three series and Table [B.1](#) summarizes the correlations of the cyclical components.

Figure B.1: Aggregate employment: BDS vs. CPS vs. Establishment survey.



Source: Own calculations, Business Dynamics Statistics, Current Population Survey, Establishment Survey, 1982-2011

Table B.1: Cross correlations of cyclical components of employment and GDP.

Data set	Lags										
	-5	-4	-3	-2	-1	0	1	2	3	4	5
CPS	0.07	-0.02	-0.11	-0.26	-0.14	0.18	0.43	-0.12	-0.19	0.05	0.23
Establishment survey	0.08	-0.02	-0.12	-0.21	-0.15	0.05	0.45	-0.05	-0.18	0.05	0.25
BDS	0.06	-0.03	-0.07	-0.23	-0.17	0.04	0.50	-0.01	-0.14	-0.03	0.11

Notes: All series are logged and HP filtered with parameter $\lambda = 6.25$. Annual data, 1982-2011.

B.1.2 Age/Size distribution of firms and employment

Table B.2 presents a distribution of the BDS population of firms over size and age groups.

Table B.2: Distribution of firms over age and size groups (%).

	Small (0-19)	Large (20+)	All sizes		Small (0-99)	Large (100+)	All sizes
Young (0-5)	39.7	2.6	42.3	Young (0-5)	42.0	0.3	42.3
Old (6+)	48.5	9.2	57.7	Old (6+)	55.9	1.8	57.7
All ages	88.2	11.8	100.0	All ages	97.9	2.1	100.0

Table B.3 presents a distribution of the BDS population of firms over size and age groups.

Table B.3: Distribution of employment over age and size groups (%).

	Small (0-19)	Large (20+)	All sizes		Small (0-99)	Large (100+)	All sizes
Young (0-5)	8.7	7.3	16.0	Young (0-5)	12.6	3.4	16.0
Old (6+)	11.0	73.0	84.0	Old (6+)	25.1	58.9	84.0
All ages	19.7	80.3	100.0	All ages	37.7	62.3	100.0

B.1.3 Cross correlations

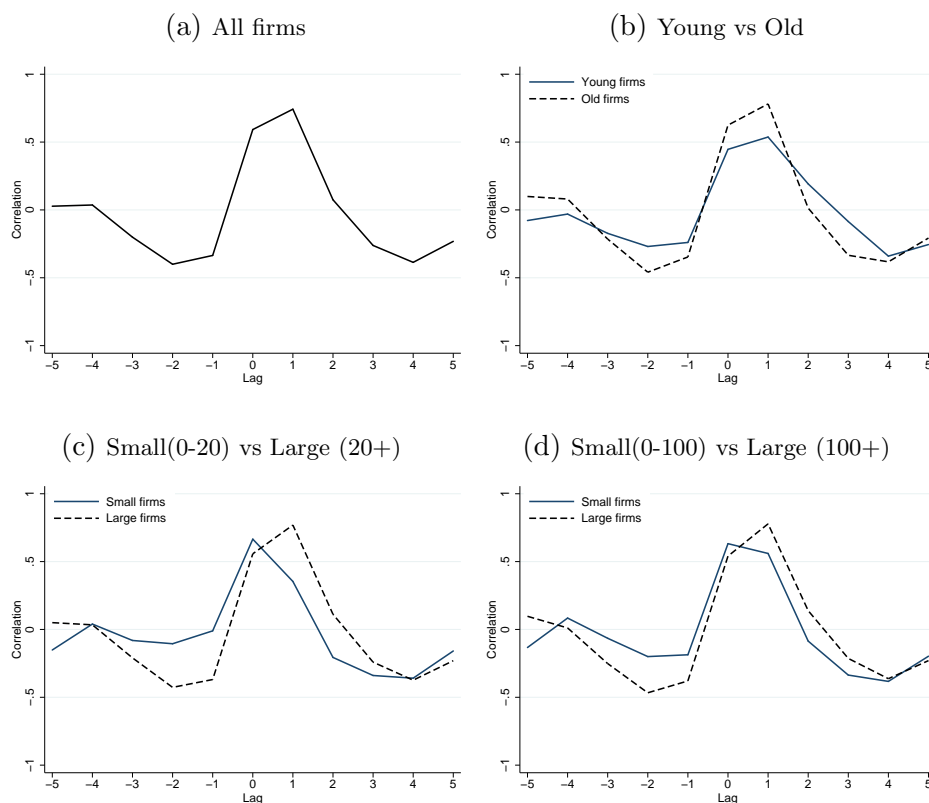
Table B.4 reports the contemporaneous correlations between the cyclical components of employment for different groups of firms and GDP. Figure B.2 illustrate the cross-correlations between employment and GDP for different groups of firms.

Table B.4: Contemporaneous correlation of cyclical components of employment and GDP.

	GDP	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GDP	1.00	0.59	0.45	0.62	0.67	0.56	0.63	0.54
(1) All firms		1.00	0.88	0.96	0.77	1.00	0.91	0.98
(2) Young (0-5)			1.00	0.70	0.68	0.87	0.80	0.86
(3) Old (5+)				1.00	0.73	0.95	0.87	0.94
(4) Small (0-20)					1.00	0.71	0.95	0.64
(5) Large (20+)						1.00	0.87	0.99
(6) Small (0-100)							1.00	0.81
(7) Large (100+)								1.00

Notes: All series are logged and HP filtered with parameter $\lambda = 6.25$. Annual data, 1982-2012. Source: Business Dynamics Statistics (BDS).

Figure B.2: Cross correlations of cyclical components of employment and GDP.



Source: Business Dynamics Statistics; Annual data, 1982-2012.

B.2 Proofs

This section presents the proofs of results given in the main text.

B.2.1 Goods market clearing.

Note that the market clearing in the consumption goods market can be derived as follows.

Define a consumption of entrepreneurs:

$$C^e = \sum_{s \in S} \Gamma_s \int_V \pi(\theta_s) [\theta_s^{1-\gamma} F(l_s(v_s))] d\mu_s(v_s) - \sum_{s \in S} \Gamma_s \int_V \pi(\theta_s) m(v_s, \theta_s) d\mu_s(v_s)$$

thus

$$C^e = Y - P. \tag{B.1}$$

The budget constraint of the financial intermediary implies

$$P = L + A' - (1 + r)A$$

and using the fact that

$$L = \sum_{s \in S} \Gamma_s \int_V (wn(v_s) + (r + \delta)k(v_s)) d\mu_s(v_s)$$

hence I get

$$P = w \sum_{s \in S} \Gamma_s \int_V n(v_s) d\mu_s(v_s) + (r + \delta) \sum_{s \in S} \Gamma_s \int_V k(v_s) d\mu_s(v_s) + A' - (1 + r)A$$

using market clearing for assets $K = A$ and (B.1) I arrive at

$$Y - C^e = w \sum_{s \in S} \Gamma_s \int_V n(v_s) d\mu_s(v_s) + (r + \delta) \sum_{s \in S} \Gamma_s \int_V k(v_s) d\mu_s(v_s) + K' - (1 + r)K$$

furthermore from the problem of the consumer I get

$$C^w = wh$$

and by definition of $N = \sum_{s \in S} \Gamma_s \int_V n_s(v_s) d\mu_s(v_s)$ and labor market clearing $N = h$ I get

$$Y - C^e = C^w + (r + \delta) \sum_{s \in S} \Gamma_s \int_V k(v_s) d\mu_s(v_s) + K' - (1 + r) K$$

furthermore using $K = \sum_{s \in S} \Gamma_s \int_V k_s(v_s) d\mu_s(v_s)$ I get

$$Y = C^w + C^e + K' - (1 - \delta) K$$

which is market clearing for consumption goods, see equation 3.10.

B.2.2 Proofs from theory section

In the Appendix I skip the dependence of the value functions and policy functions on the initial types s since all the results hold for any s . I also relabel the the demand shock so that $\theta = \theta^{1-\gamma}$ wherever the degree of returns to scale does not play any role to economize on notation. The contracting problem for an individual firm, incorporating a possibility of randomization, is given by

$$\widehat{B}(v) = \max_{l, c(\theta), v'(\theta)} \left\{ -l + \sum_{\theta \in \Theta} \pi(\theta) \left[\theta F(l) - c(\theta) + \frac{\zeta}{(1+r)} B(v'(\theta)) \right] \right\} \quad (\text{B.2})$$

subject to

$$v = \sum_{\theta \in \Theta} \pi(\theta) [U(c(\theta)) + \beta \zeta v'(\theta)]$$

$$U(c(\theta)) + \beta \zeta v'(\theta) \geq U((\theta - \theta') F(l) + c(\theta')) + \beta \zeta v'(\theta') \quad \forall \theta, \theta'$$

where

$$B(v) = \max_{\alpha \in [0,1], v_1, v_2} \alpha \widehat{B}(v_1) + (1 - \alpha) \widehat{B}(v_2) \quad (\text{B.3})$$

subject to

$$v = \alpha v_1 + (1 - \alpha) v_2$$

where without loss of generality I can restrict randomization to be between two points. In order to characterize the contract I use the formulation (B.2) – (B.3). Let the decision rules associated with (B.2) be $l(v) : [v_{\min}, v_{\max}] \rightarrow \mathbb{R}_+$, $m(v, \theta) : [v_{\min}, v_{\max}] \times \Theta \rightarrow \mathbb{R}$ and $v'(v, \theta) : [v_{\min}, v_{\max}] \times \Theta \rightarrow \mathbb{R}$. Moreover, let $\omega(v, \theta) \equiv U(\theta F(l(v) - m(v, \theta)))$. Denote by $V_{nr} \subset [v_{\min}, v_{\max}]$ be the (non-empty) region of the state space for which randomization is not optimal, i.e. $B(v) = \widehat{B}(v)$. Let $V_r \subset [v_{\min}, v_{\max}]$ be the (potentially empty) randomization region i.e. region of the state space for which it is optimal to randomize, i.e. $B(v) > \widehat{B}(v)$. Without the loss of generality the randomization is between two values in $[v_{\min}, v_{\max}]$.

It is instructive to establish certain properties of the value function $B(v)$. First, $B(v)$ can not exceed the value of the unconstrained first-best contract. Under the first-best the lending is pinned down by the following condition

$$l^* = (F')^{-1} \frac{1}{\mathbb{E}[\theta]}$$

moreover the unconstrained first-best contract collects $(\theta F(l^*) - c^*(v))$ where $c^*(v)$ solves the $v = \sum_{t=0}^{\infty} \frac{U(c^*)}{1-\beta\zeta}$. Thus

$$B_{\max} = \frac{1}{1-\beta\zeta} \left[-l^* + \sum_{\theta \in \Theta} \pi(\theta) [\theta F(l^*(v)) - c^*(v)] \right]$$

The value function $B(v)$ is strictly concave, which I show in the Lemma 6, and by assumptions on the utility I have $\lim_{c \rightarrow c_{\min}} U'(c) = \infty$. Hence, it becomes very cheap for the intermediary to increase a promised utility when the current promise is very low, that is $\lim_{v \rightarrow v_{\min}} B'(v) = 0$. On the other hand when promised utility is close to the upper bound, where the entrepreneur has a low marginal utility of additional consumption v_{\max} , i.e. $\lim_{v \rightarrow v_{\max}} B'(v) = -\infty$, so it becomes very expensive to to increase promised utility.

Proof of Lemma 4. Follows standard arguments see Chapter 19.5.2 in Sargent and Ljungqvist.

□

Proof of Lemma 5. Add downward local constraint $C_{n,n-1} \geq 0$ and local upward constraint $C_{n-1,n} \geq 0$, i.e.

$$C_{n,n-1} \equiv U(\theta_n F(l) - m(\theta_n)) + \beta \zeta v'(\theta_n) - [U(\theta_n F(l) - m(\theta_{n-1})) + \beta \zeta v'(\theta_{n-1})] \geq 0$$

$$C_{n-1,n} \equiv U(\theta_{n-1} F(l) - m(\theta_{n-1})) + \beta \zeta v'(\theta_{n-1}) - [U(\theta_{n-1} F(l) - m(\theta_n)) + \beta \zeta v'(\theta_n)] \geq 0$$

to get

$$U(\theta_n F(l) - m(\theta_n)) - U(\theta_n F(l) - m(\theta_{n-1})) + U(\theta_{n-1} F(l) - m(\theta_{n-1})) - U(\theta_{n-1} F(l) - m(\theta_n)) \geq 0$$

therefore

$$U(\theta_n F(l) - m(\theta_n)) - U(\theta_n F(l) - m(\theta_{n-1})) \geq U(\theta_{n-1} F(l) - m(\theta_n)) - U(\theta_{n-1} F(l) - m(\theta_{n-1})) \quad (\text{B.4})$$

Note that a concavity of the $U(\cdot)$ implies

$$-U_m(\theta_n F(l) - m) < -U_m(\theta_{n-1} F(l) - m) \quad \forall m$$

where $U_m = \frac{\partial U}{\partial m}$. Suppose that the $m(\theta_n) < m(\theta_{n-1})$, hence

$$\begin{aligned} -\int_{m(\theta_n)}^{m(\theta_{n-1})} U_m(\theta_n F(l) - m) &< -\int_{m(\theta_n)}^{m(\theta_{n-1})} U_m(\theta_{n-1} F(l) - m) \\ U(\theta_n F(l) - m(\theta_{n-1})) - U(\theta_n F(l) - m(\theta_n)) &> U(\theta_{n-1} F(l) - m(\theta_{n-1})) - U(\theta_{n-1} F(l) - m(\theta_n)) \\ -U(\theta_n F(l) - m(\theta_n)) + U(\theta_n F(l) - m(\theta_{n-1})) &> -U(\theta_{n-1} F(l) - m(\theta_n)) + U(\theta_{n-1} F(l) - m(\theta_{n-1})) \end{aligned}$$

hence

$$U(\theta_n F(l) - m(\theta_n)) - U(\theta_n F(l) - m(\theta_{n-1})) < U(\theta_{n-1} F(l) - m(\theta_n)) - U(\theta_{n-1} F(l) - m(\theta_{n-1}))$$

which contradicts (B.4). Hence, it has to be that $m(\theta_n) \geq m(\theta_{n-1})$. Then from $C_{n,n-1}$ it is immediate that $v'(\theta_n) \geq v'(\theta_{n-1})$, which completes the proof. \square

Proof of Lemma 6. Under Assumptions 2 and 3 $B(v) = \widehat{B}(v)$. Rewrite the problem

(B.2) using the change of the variables. Instead of $(l, c(\theta), v'(\theta))$, consider choosing $(\underline{u}, u(\theta), v'(\theta))$ where $u = U(c)$ and $C : [U(0), U(\infty)] \rightarrow \mathbb{R}$ and $C = U^{-1}$. Also, let $H : [F(0), F(\infty)] \rightarrow [0, l^*]$ and $H = F^{-1}$. Then we have

$$c(\theta) = C(u(\theta))$$

and let

$$\underline{u} = U((\theta - \theta') F(l) + C(u(\theta')))$$

then

$$\begin{aligned} F(l) &= \frac{C(\underline{u}) - C(u(\theta'))}{(\theta - \theta')} \\ l &= H\left(\frac{C(\underline{u}) - C(u(\theta'))}{(\theta - \theta')}\right) \end{aligned}$$

and define

$$G(\underline{u}, u) = -H\left(\frac{C(\underline{u}) - C(u(\theta'))}{(\theta - \theta')}\right) + \frac{C(\underline{u}) - C(u(\theta'))}{(\theta - \theta')}$$

hence the problem can be rewritten as

$$\begin{aligned} B(v) &= \max_{\underline{u}, u(\theta), v'(\theta)} \left\{ G(\underline{u}, u) + \sum_{\theta \in \Theta} \pi(\theta) \left[-C(u(\theta)) + \frac{\zeta}{(1+r)} B(v'(\theta)) \right] \right\} \\ &\text{subject to} \\ v &= \sum_{\theta \in \Theta} \pi(\theta) [u(\theta) + \beta v'(\theta)] \\ u(\theta) + \beta v'(\theta) &\geq \underline{u} + \beta v'(\theta') \quad \forall \theta, \theta' \end{aligned}$$

The set of constraints is now linear in the choice variables, therefore convex. By Assumption 2 (strict concavity of the utility function), we have that $C = U^{-1}$ is strictly convex and therefore $-C$ is a strictly concave function. Under Assumption 2 (decreasing returns to scale in production) F is a strictly concave function, therefore H is strictly convex and $-H$ is strictly concave function. Thus, the function $G(\underline{u}, u) + \sum_{\theta \in \Theta} \pi(\theta) (-C(u(\theta)))$ is a strictly concave function. Then by Theorem 4.8 in SLP $B(v)$ is strictly concave and $(\underline{u}, u(\theta), v'(\theta))$ are continuous, single-valued functions. This completes a proof of part (i).

For part (ii), first note that for $v \in V_r$, B is linear, and thus it is differentiable. For $v \in V_{nr}$ I establish the differentiability by the application of the Beneviste, Sheinkman Theorem (see SLP Theorem 4.10). Let $x = (l, c(\theta), v'(\theta))$ be the solution that attains $B(v)$. Take any $v_0 \in V_{nr} \cap [v_{\min}, v_{\max}]$. Consider a neighborhood of v_0 , $D(v_0, \varepsilon) = (v_0 - \varepsilon, v_0 + \varepsilon)$ for some small $\varepsilon > 0$. Define $\hat{x} = (\hat{l}(v), \hat{c}(v, \theta), \hat{v}'(v, \theta))$ for all $v \in D(v_0, \varepsilon)$ such that

$$\begin{aligned}\hat{l}(v) &= l(v_0) + \frac{(v - v_0)}{F'(l(v_0))} \left[\frac{1}{(\theta - \theta')} \left(\frac{1}{U'((\theta - \theta') F(l(v_0)) + c(v_0, \theta'))} - \frac{1}{U'(c(v_0, \theta'))} \right) \right], \\ \hat{c}(v, \theta) &= c(v_0, \theta) + \frac{v - v_0}{U'(c(\theta))}, \\ \hat{v}'(v, \theta) &= v'(v_0, \theta).\end{aligned}$$

Denote to economize on notation that

$$y(v_0) = (\theta - \theta') F(l(v_0)) + c(v_0, \theta')$$

and note it satisfies the promise keeping constraint, i.e. for all $v \in D(v_0, \varepsilon)$

$$\begin{aligned}v &= \sum_{\theta \in \Theta} \pi(\theta) [u(\hat{c}(v, \theta)) + \beta \zeta \hat{v}'(v, \theta)] \\ &= \sum_{\theta \in \Theta} \pi(\theta) \left[u \left(c(v_0, \theta) + \frac{v - v_0}{U'(c(\theta))} \right) + \beta \zeta v'(v_0, \theta) \right] \\ &= \sum_{\theta \in \Theta} \pi(\theta) [u(c(v_0, \theta)) + \beta \zeta v'(v_0, \theta)] + (v - v_0)\end{aligned}$$

hence

$$v_0 = [u(c(v_0, \theta)) + \beta \zeta v'(v_0, \theta)].$$

Moreover, \hat{x} satisfies the incentive-compatibility constraints $\forall \theta > \theta'$. To see that start with the constraint at v_0 , adding $(v - v_0)$ both sides

$$U(c(v_0, \theta)) + (v - v_0) + \beta \zeta v'(v_0, \theta) = U((\theta - \theta') F(l(v_0)) + c(v_0, \theta)) + (v - v_0) + \beta \zeta v'(v_0, \theta)$$

thus the first two terms of the right hand side can be rearranged into

$$U \left((\theta - \theta') F \left(l(v_0) + \frac{(v - v_0)}{F'(l(v_0))} \left(\frac{1}{(\theta - \theta')} \left(\frac{1}{U'(y(v_0))} - \frac{1}{U'(c(v_0, \theta'))} \right) \right) \right) \right) + c(v_0, \theta) + \frac{v - v_0}{U'(c(v_0, \theta'))}$$

and analogously the first two terms of the left hand side into

$$U \left(c(v_0, \theta) + \frac{v - v_0}{U'(c(v_0, \theta'))} \right) + \beta \zeta v'(v_0, \theta)$$

then using definitions of $\hat{l}(v)$, $\hat{c}(v, \theta)$ and $\hat{v}'(v, \theta)$ and equalizing both sides I arrive at

$$U(\hat{c}(v, \theta)) + \beta \hat{v}'(v, \theta) = U \left((\theta - \theta') F(\hat{l}(v)) + \hat{c}(v, \theta') \right) + \beta \zeta \hat{v}'(v, \theta').$$

Next, define a function $\underline{B} : D(v_0, \varepsilon) \rightarrow \mathbb{R}$ as

$$\underline{B}(v) = -\hat{l}(v) + \sum_{\theta \in \Theta} \pi(\theta) \left[\theta F(\hat{l}(v)) - \hat{c}(v, \theta) + \frac{\zeta}{(1+r)} B(\hat{v}'(v, \theta)) \right]$$

note that for $v = v_0$ it is $\underline{B}(v_0) = B(v_0)$, since $\hat{l}(v_0) = l(v_0)$ and $\hat{c}(v_0, \theta) = c(v, \theta)$. Moreover by the fact that $x = (l, c(\theta), v'(\theta))$ is the solution that attains $B(v)$ and $\hat{x} = (\hat{l}(v), \hat{c}(v, \theta), \hat{v}'(v, \theta))$ is feasible and incentive compatible for all $v \in D(v_0, \varepsilon)$ I arrive at

$$\underline{B}(v) \leq B(v) \quad \forall v \in D(v_0, \varepsilon)$$

and $B'(v_0) = \underline{B}'(v)|_{v_0}$, where

$$\begin{aligned} \underline{B}'(v)|_{v_0} &= \frac{\sum_{\theta \in \Theta} \pi(\theta) \theta F'(l(v_0)) - 1}{F'(l(v_0))} \left[\frac{1}{(\theta - \theta')} \left(\frac{1}{U'(y(v_0))} - \frac{1}{U'(c(v_0, \theta'))} \right) \right] \\ &\quad - \sum_{\theta \in \Theta} \frac{\pi(\theta)}{U'(c(v_0, \theta))} \end{aligned}$$

which completes the proof. \square

The value function $B(v)$ is differentiable and under Assumption 3 it is strictly concave. In order to simplify notation let $c_i = c(v, \theta_i)$ and $v_i = v'(v, \theta_i)$ for $i = L, H$ for every

$s \in S$. Moreover let $\lambda_{IC}(v)$ and $\lambda_{PKC}(v)$ be the Lagrange multipliers on respectively incentive-compatibility and promise keeping constraints. I skip the dependence contract policy functions and multipliers on the continuation utilities v and types s to simplify the notation. Then, the necessary and sufficient conditions for the interior solution are:

$$l : F'(l) \sum_{i \in L, H} \pi_i \theta_i - 1 - \lambda_{IC} (\theta_H - \theta_L) F'(l) U'((\theta_H - \theta_L) F(l) + c_L) = 0 \quad (\text{B.5})$$

$$c_L : -\pi_L - \lambda_{PKC} \pi_L U'(c_L) - \lambda_{IC} U'((\theta_H - \theta_L) F(l) + c_L) = 0 \quad (\text{B.6})$$

$$c_H : -\pi_H - \lambda_{PKC} \pi_H U'(c_H) + \lambda_{IC} U'(c_H) = 0 \quad (\text{B.7})$$

$$v'_L : \frac{\pi_L}{(1+r)} B'(v'_L) - \lambda_{PKC} \beta \pi_L - \lambda_{IC} \beta = 0 \quad (\text{B.8})$$

$$v'_H : \frac{\pi_H}{(1+r)} B'(v'_H) - \lambda_{PKC} \beta \pi_H + \lambda_{IC} \beta = 0 \quad (\text{B.9})$$

and the envelope condition is

$$B'(v) = \lambda_{PKC} \quad (\text{B.10})$$

which after rearranging together with the promise keeping constraint and incentive compatibility constraint constitute the following set of equations determining allocation $\{c_L, c_H, v_L, v_H, l\}$

$$v - \pi_H [U(c_H) + \beta \zeta v'_H] - \pi_L [U(c_L) + \beta \zeta v'_L] = 0 \quad (\text{B.11})$$

$$U(c_H) + \beta \zeta v'_H - [U((\theta_H - \theta_L) F(l) + c_L) + \beta \zeta v'_L] = 0 \quad (\text{B.12})$$

$$-\pi_L - U'(c_L) \left[B'(v) \pi_L + \lambda_{IC} \frac{U'((\theta_H - \theta_L) F(l) + c_L)}{U'(c_L)} \right] = 0 \quad (\text{B.13})$$

$$-\pi_H - U'(c_H) [B'(v) \pi_H - \lambda_{IC}] = 0 \quad (\text{B.14})$$

$$\mathbb{E}[\theta] F'(l) - 1 - \lambda_{IC} (\theta_H - \theta_L) F'(l) U'((\theta_H - \theta_L) F(l) + c_L) = 0 \quad (\text{B.15})$$

where

$$\lambda_{IC} = \frac{\pi_H \pi_L}{\beta (1+r)} (B'(v'_L) - B'(v'_H)) \quad (\text{B.16})$$

Proof of Proposition 7. I start with establishing that incentive compatibility constraint

must bind at the optimal solution. Rewrite the problem (B.2) using the feasibility constraint $\theta_n F(l) - m_n = c_n$ for $n = L, H$ to replace consumption with payments

$$B(v) = \max_{l, m(\theta), v'(\theta)} \left\{ -l + \sum_{i=L, H} \pi_i \left[m_i + \frac{\zeta}{(1+r)} B(v'_i) \right] \right\} \quad (\text{B.17})$$

subject to

$$\begin{aligned} v &= \sum_{i=L, H} \pi_i [U(\theta_i F(l) - m_i) + \beta \zeta v'_i] \\ U(\theta_H F(l) - m_H) + \beta \zeta v'_H &\geq U(\theta_H F(l) - m_L) + \beta \zeta v'_L \quad \forall \theta, \theta' \end{aligned}$$

and suppose the incentive constraint is not binding at the solution $\{l, m(\theta), v'(\theta)\}$, i.e.

$$U(\theta_H F(l) - m_H) + \beta \zeta v'_H > U(\theta_H F(l) - m_L) + \beta \zeta v'_L \quad (\text{B.18})$$

By the Lemma 5, $m(\theta_H) \geq m(\theta_L)$ and $v'(\theta_H) \geq v'(\theta_L)$, thus it has to be that $v'_H(v) > v'_L(v)$. Consider now the following variation

$$\widehat{v}'_H(v) = v'_H(v) - \varepsilon$$

where ε is large enough to that (B.18) holds with equality. Now, let

$$\widehat{v}'_L(v) = v'_L(v) + \frac{\pi_H \varepsilon}{\pi_L}$$

then

$$\sum_{i=L, H} \pi_i \widehat{v}'_i(v) = \pi_H (v'_H(v) - \varepsilon) + \pi_L \left(v'_L(v) + \frac{\pi_H \varepsilon}{\pi_L} \right) = \sum_{i=L, H} \pi_i v'_i(v)$$

thus the variation satisfies the promise keeping constraint and is a mean preserving decrease in spread of the continuation values. Under Assumption 3, the value function $B(v)$ is strictly concave, hence $\sum_{i=L, H} \pi_i B(\widehat{v}'_i) > \sum_{i=L, H} \pi_i B(v'_i)$ and hence $B(v)$ increases contradicting the optimality. Hence, incentive constraint is binding at the solution to the problem B.2. Given Lemma 5 there are 4 possible cases: (1) $v'_H > v'_L, m_H > m_L$, (2) $v'_H = v'_L, m_H > m_L$, (3) $v'_H > v'_L, m_H = m_L$, (4) $v'_H = v'_L, m_H = m_L$. First note that case (2) violates the incentive compatibility constraint. Consider case (4) and note whenever

$m_H = m_L$ then $c_H > c_L$. Collapse first order conditions with respect to v'_L and v'_H into

$$\frac{\beta(1+r)}{\pi_H\pi_L}\lambda_{IC} = (B'(v'_L) - B'(v'_H))$$

then since $v'_H = v'_L$, if only solution is interior, it has to be that $\lambda_{IC} = 0$, which implies from collapsed (B.14) and (B.13)

$$\frac{1}{U'(c_L)} = \frac{1}{U'(c_H)}$$

hence $c_L = c_H$, a contradiction. Finally, consider case (3). Consider the following deviation

$$\begin{aligned}\hat{v}'_H(v) &= v'_H(v) - \frac{\varepsilon}{\beta\zeta}, & \hat{v}'_L(v) &= v'_L(v) \\ \hat{m}_H(v) &= m_H(v) + \frac{\varepsilon}{U'(c_H)}, & \hat{m}_L(v) &= m_L(v)\end{aligned}$$

which satisfies the promise keeping constraint and is incentive compatible. By strict concavity of the value function $B(v)$ there is $\sum_{i=L,H} \pi_i B(\hat{v}'_i) > \sum_{i=L,H} \pi_i B(v'_i)$ and also we have $\hat{m}_H(v) > m_H(v)$ implying together that $B(v)$ increases, contradicting optimality. Therefore under the optimal contract it has to be that $v'_H > v'_L, m_H > m_L$. Now, use again a collapsed conditions (B.8) and (B.9) together with the fact that $v'_H > v'_L$ to conclude that for any interior solution it has to be that $\lambda_{IC}(v) > 0$. Use modified conditions (B.6) and (B.7) respectively

$$\begin{aligned}\lambda_{PKC} &= -\frac{1}{U'(c_L)} - \lambda_{IC} \frac{U'((\theta_H - \theta_L)F(l) + c_L)}{\pi_L U'(c_L)} < -\frac{1}{U'(c_L)} - \frac{\lambda_{IC}}{\pi_L} < -\frac{1}{U'(c_L)} \\ \lambda_{PKC} &= -\frac{1}{U'(c_H)} + \frac{\lambda_{IC}}{\pi_H} > -\frac{1}{U'(c_H)}\end{aligned}$$

where the first inequality in the first line comes from the fact that for any interior solution $l > 0$ (see proof of part (ii)) and the second inequality is implied by $\lambda_{IC}(v) > 0$. Analogously the inequality in the second line is implied by $\lambda_{IC}(v) > 0$. Therefore combining the two conditions

$$\frac{1}{U'(c_H(v))} > -\lambda_{PKC}(v) > \frac{1}{U'(c_L(v))}$$

strict concavity of $B(v)$ implies $\lambda_{PKC}(v) < 0$ thus I obtain $c_H > c_L$. This completes a

proof of part (i). For part (ii) consider a problem of finding a minimum value of the continuation utility v^* such that the statically efficient lending l^* is both incentive compatible and feasible. The value v^* is a solution to the problem where the right hand side of the participation constraint is an objective function and the incentive and feasibility are the constraints, i.e.

$$\begin{aligned}
v^* &\equiv \min_{c_H, c_L, m_H, m_L, v_H, v_L} \{ \pi_H [U(c_H) + \zeta\beta v_H] + \pi_L [U(c_L) + \zeta\beta v_L] \} \\
&\text{subject to} \\
&U(c_H) + \beta\zeta v_H \geq U((\theta_H - \theta_L)F(l^*) + c_L) + \beta\zeta v_L \\
&c_H + m_H \leq \theta_H F(l^*) \\
&c_L + m_L \leq \theta_L F(l^*) \\
&v_H \geq v^*, v_L \geq v^*
\end{aligned}$$

and using the feasibility one can rule out the c_H, c_L from the problem and rewrite the objective as

$$\pi_H U(\theta_H F(l^*) - m_H) + \pi_L U(\theta_L F(l^*) - m_L) + \zeta\beta [\pi_H (v_H - v_L) + v_L]$$

further using binding incentive compatibility constraint objective is reduced to the problem

$$\begin{aligned}
v^* &\equiv \min_{m_H, m_L, v_L} \pi_L U(\theta_L F(l^*) - m_L) + \pi_H U(\theta_H F(l^*) - m_L) + \zeta\beta v_L \\
&v_L \geq v^*
\end{aligned}$$

Note that necessary condition is $v_L = v^*$ and v^* is a finite by Assumption 2 which completes the argument that such point exists. Now I argue that for any $v < v^*$ it has to be that $l < l^*$. Suppose now for contradiction that there exists $v \in [v_{\min}, v^*]$ such $l = l^*$. Then l is determined by

$$\sum_{i \in L, H} \pi_i \theta_i F'(l) - 1 = 0 \tag{B.19}$$

rewriting the first order condition (B.5) using (B.19) we obtain

$$0 = \lambda_{IC} U' \left((\theta_H - \theta_L) F \left(F'^{(-1)} \left(\frac{1}{\sum_{i \in L, H} \pi_i \theta_i} \right) \right) + c_L \right)$$

since for any finite c_L the equation above can not hold. By the fact that the solution is interior and incentive compatibility constraint is binding there is $\lambda_{IC} > 0$. Thus it has to be that $l < l^*$. \square

Proof of Proposition 8. The most convenient way to prove this result is with the use of the sequential formulation of the problem

$$J(v_0) = \max_{\mathbf{x}} \sum_{t=j}^{\infty} \sum_{\theta^t} \left(\frac{\zeta}{1+r} \right)^{t-j} \Pr(\theta^t) [m(\theta^t) - l(\theta^{t-1})] \quad (\text{B.20})$$

subject to

$$\begin{aligned} c(\theta^t) + m(\theta^t) &\leq \theta_n^{1-\gamma} F(l(\theta^{t-1})) \quad \forall t \geq j, \forall \theta^{t-1} \in \Theta^{t-1}, n = L, H \\ \sum_{t=j}^{\infty} \sum_{\theta^t} (\beta\zeta)^t \Pr(\theta^t) U(c(\theta^t)) &\geq v_0 \\ U(c(\theta^t)) + \beta\zeta v(\theta^t) &\geq \\ U\left(\left((\theta_H)^{1-\gamma} - (\lambda\theta_L)^{1-\gamma}\right) F(l(\theta^{t-1})) + c(\theta^{t-1}, \theta_L)\right) + \beta\zeta v(\theta^{t-1}, \theta_L), &\forall \theta^{t-1} \\ v_0 &\text{ given.} \end{aligned}$$

where

$$v(\theta^t) = \sum_{i=1}^{\infty} \sum_{\theta^{t+i}} (\beta\zeta)^{i-1} \Pr(\theta^{t+i} | \theta^t) U(c_i(\theta^{t+i}))$$

Let $x^* = \{c^*(\theta^t), m^*(\theta^t), l^*(\theta^{t-1})\}$ be the solution of the problem (B.20). The policy functions are the functions of a history of shocks, i.e. $c^* : \Theta^t \rightarrow \mathbb{R}_+$, $m^* : \Theta^t \rightarrow \mathbb{R}$ and $l^* : \Theta^{t-1} \rightarrow \mathbb{R}_+$, where $\Theta = \{\theta_L, \theta_H\}$. Consider any $\lambda > 0$, so that $\Theta = \{\lambda\theta_L, \lambda\theta_H\}$ then under Assumption 7

$$\lambda\theta_H = \lambda \left(\bar{\theta} + \frac{\sigma}{\pi_H} \right)^{\frac{1}{1-\gamma}}, \quad \lambda\theta_L = \lambda \left(\bar{\theta} - \frac{\sigma}{\pi_L} \right)^{\frac{1}{1-\gamma}}$$

and thus

$$(\lambda\theta_H)^{1-\gamma} = \lambda^{1-\gamma}\bar{\theta}\left(1 + \frac{\sigma}{\pi_H}\right), \quad (\lambda\theta_L)^{1-\gamma} = \lambda^{1-\gamma}\bar{\theta}\left(1 - \frac{\sigma}{\pi_L}\right)$$

therefore

$$\mathbb{E}\left[(\lambda\theta)^{1-\gamma}\right] = \lambda^{1-\gamma}\bar{\theta}, \quad Std\left((\lambda\theta)^{1-\gamma}\right) = \frac{\sigma}{\sqrt{\pi_L\pi_H}}$$

so that expected return on project increases by $\lambda^{1-\gamma}$ and the standard deviation remains unchanged. Immediately, we have $\lambda\theta^t = \{\lambda\theta_j, \dots, \lambda\theta_t\}$. Under Assumption 2 output of the firm is $F(l(\theta^{t-1})) = [l(\theta^{t-1})]^\gamma \Omega(r, w)$ where $\Omega(r, w)$ is a function of prices and exogenous parameters of the model only. Note that $\Pr(\lambda\theta^t) = \Pr(\theta^t)$ since probabilities of the particular history remain unchanged. Then, the contracting problem becomes

$$\hat{J}(\hat{v}_0) = \max_{\mathbf{x}} \sum_{t=j}^{\infty} \sum_{\theta^t} \left(\frac{\zeta}{1+r}\right)^{t-j} \Pr(\theta^t) m(\lambda\theta^t) - l(\lambda\theta^{t-1}) \quad (\text{B.21})$$

subject to

$$\begin{aligned} c(\lambda\theta^t) + m(\lambda\theta^t) &\leq (\lambda\theta_n)^{1-\gamma} (l(\lambda\theta^{t-1}))^\gamma \Omega(r, w) \quad \forall t \geq j, \forall \theta^{t-1} \in \Theta^{t-1}, n = L, H \\ \sum_{t=j}^{\infty} \sum_{\theta^t} (\beta\zeta)^t \Pr(\theta^t) U(c(\lambda\theta^t)) &\geq \hat{v}_0 \\ U(c(\lambda\theta^t)) + \beta\zeta v(\lambda\theta^t) &\geq \\ U\left(\left((\lambda\theta_H)^{1-\gamma} - (\lambda\theta_L)^{1-\gamma}\right) (l(\lambda\theta^{t-1}))^\gamma \Omega(r, w) + c(\lambda\theta^{t-1}, \lambda\theta_L)\right) &+ \beta\zeta v(\lambda\theta^{t-1}, \lambda\theta_L) \\ \hat{v}_0 &\text{ given.} \end{aligned}$$

where \hat{v}_0 is the initial promised utility pinned down by the free entry condition.

$$v(\lambda\theta^t) = \sum_{i=1}^{\infty} \sum_{\theta^{t+i}} (\beta\zeta)^{i-1} \Pr(\theta^{t+i}|\theta^t) U(c_t(\lambda\theta^{t+i}))$$

Guess the solution to the problem (B.21) has the form $\hat{x} = \lambda x^* = \{\lambda c^*(\theta^t), \lambda m^*(\theta^t), \lambda l^*(\theta^{t-1})\}$ and the initial promised utility is $\hat{v}_0 = \lambda^{1-\rho} v_0$. In what follows I will show this policy is feasible, satisfies the participation constraint and is incentive compatible. Next, I argue it maximizes the value of the lender and that \hat{v}_0 is consistent with zero profit condition.

Feasibility requires $\forall t \geq j, \forall \theta^{t-1} \in \Theta^{t-1}, n = L, H$

$$\lambda c^*(\theta^t) + \lambda m^*(\theta^t) \leq (\lambda \theta_n)^{1-\gamma} (\lambda l^*(\theta^{t-1}))^\gamma \Omega(r, w)$$

which clearly holds. For the participation constraint

$$\begin{aligned} \sum_{t=j}^{\infty} \sum_{\theta^t} (\beta \zeta)^t \Pr(\theta^t) U(\lambda c^*(\theta^t)) &\geq \hat{v}_0 \\ \sum_{t=j}^{\infty} \sum_{\theta^t} (\beta \zeta)^t \Pr(\theta^t) \lambda^{1-\rho} U(c^*(\theta^t)) &\geq \lambda^{1-\rho} v_0 \end{aligned}$$

where the second equation is due to the utility function form imposed in Assumption 7. Hence it is also satisfied. As for the incentive compatibility

$$\begin{aligned} U(\lambda c^*(\theta^t)) + \beta \zeta v(\lambda \theta^t) &\geq \\ U\left(\left((\lambda \theta_H)^{1-\gamma} - (\lambda \theta_L)^{1-\gamma}\right) (\lambda l^*(\theta^{t-1}))^\gamma \Omega(r, w) + \lambda c^*(\theta^{t-1}, \theta_L)\right) &+ \beta \zeta v(\lambda \theta^{t-1}, \lambda \theta_L) \end{aligned}$$

which given the properties of the utility function allows to factor out the $\lambda^{1-\rho}$ to get

$$\begin{aligned} \lambda^{1-\rho} U(c^*(\theta^t)) + \lambda^{1-\rho} \beta \zeta v(\theta^t) &\geq \\ \lambda^{1-\rho} U\left(\left((\theta_H)^{1-\gamma} - (\theta_L)^{1-\gamma}\right) (l^*(\theta^{t-1}))^\gamma \Omega(r, w) + c^*(\theta^{t-1}, \theta_L)\right) &+ \lambda^{1-\rho} \beta \zeta v(\theta^{t-1}, \theta_L) \end{aligned}$$

thus a proposed contract is incentive feasible. Thus, a contract policy \hat{x} is feasible, satisfies the participation constraint and is incentive compatible. Moreover, note that since x solves (B.20) we have

$$\begin{aligned} \hat{J}(\hat{v}_0) &= \max_{\mathbf{x}} \sum_{t=j}^{\infty} \sum_{\theta^t} \left(\frac{\zeta}{1+r}\right)^{t-j} \Pr(\theta^t) [\lambda m^*(\theta^t) - \lambda l^*(\theta^{t-1})] \\ &= \lambda \max_{\mathbf{x}} \sum_{t=j}^{\infty} \sum_{\theta^t} \left(\frac{\zeta}{1+r}\right)^{t-j} \Pr(\theta^t) [m^*(\theta^t) - l^*(\theta^{t-1})] \\ &= \lambda J(v_0) \end{aligned}$$

and thus

$$\widehat{J}(\lambda^{1-\rho}v_0) = \lambda J(v_0) \quad (\text{B.22})$$

therefore a proposed contract policy maximizes the value of the financial intermediary. Moreover, it follows that whenever $J(v_0) = 0$ then $\widehat{J}(\widehat{v}_0) = 0$, hence \widehat{v}_0 is consistent with zero profit condition. \square

Proof of Corollary 8. Consider any $b > 1$ and suppose that $\mathbb{E}[\widehat{\theta}^{1-\gamma}] = b\mathbb{E}[\theta^{1-\gamma}] = b\bar{\theta}$, then under Assumption 7 I have

$$b\bar{\theta} = b\mathbb{E}[\theta^{1-\gamma}] = b\left(\pi_H\theta_H^{1-\gamma} + \pi_L\theta_L^{1-\gamma}\right) = \left(\pi_H\left(b^{\frac{1}{1-\gamma}}\theta_H\right)^{1-\gamma} + \pi_L\left(b^{\frac{1}{1-\gamma}}\theta_L\right)^{1-\gamma}\right)$$

so it has to be

$$\widehat{\theta}_H = b^{\frac{1}{1-\gamma}}\theta_H, \quad \widehat{\theta}_L = b^{\frac{1}{1-\gamma}}\theta_L$$

then by Proposition 8 the optimal amount of lending is $l^*(\widehat{\theta}^{t-1}) = b^{\frac{1}{1-\gamma}}l(\theta^{t-1})$ thus it increases in line with the expected demand as desired. \square

Sketch of proof of Proposition 9. The proof is conducted in three steps. In the first step I show that there exists a stationary distribution of firms over the space of continuation utilities which can be attained in a finite number of periods starting from any initial distribution. In the second step I show the stationary distribution is continuous in prices. Finally, I define a continuous mapping Φ and apply Shauder Fixed-Point Theorem implying the mapping has at least one fixed point (stationary recursive equilibrium).

Step(1) Stationary distribution of firms exists. To prove the existence of the stationary distribution of firms I show that conditions in Theorem 12.12 in SLP are satisfied. Define the transtion function $Q(v_s, \mathcal{A}) : [v_{\min}, v_{\max}] \times \mathcal{B}([v_{\min}, v_{\max}]) \rightarrow \mathbb{R}$ as

$$Q(v_s, \mathcal{A}) = \begin{cases} \zeta \sum_{\theta_s \in \Theta_s} \pi(\theta_s) \mathbb{I}\{v'(v_s, \theta_s) \in \mathcal{A}\} \\ (1 - \zeta) \end{cases}$$

Firstly, I have to argue that transition function is monotone and satisfies Feller property. A transition function is monotone if for any bounded, increasing function f , the function Tf is also increasing. Consider any $v_s^1 > v_s^2$ and f increasing and note $v'(v_s^1, \theta_s) > v'(v_s^2, \theta_s)$, then

$$\int f(x) Q(v_s^1, dx) > \int f(x) Q(v_s^2, dx)$$

and hence $Q(v_s^1, \cdot) > Q(v_s^2, \cdot)$ by Exercise 12.11 in SLP, establishing monotonicity. For Feller property, note by Lemma 6 policy function $v'(v_s, \theta_s)$ is continuous for all $s \in S$ and therefore by Proposition 9.14 in SLP the transition function $Q(v_s, \mathcal{A})$ satisfies Feller property. I am left to show that Assumption 12.1 in SLP is satisfied. I have to argue that there exists a mixing point $v^* \in [v_{\min}, v_{\max}]$, a natural number $N \geq 1$ and $\varepsilon > 0$, such that $Q^N(v_{\min}, [v_s, v_{\max}]) \geq \varepsilon$ and $Q^N(v_{\max}, [v_s, v_{\min}]) \geq \varepsilon$. Given the exogenous probability of exit firm faces every period we have

$$Q(v_s, v_s^0) = (1 - \zeta) \text{ for all } v_s \in [v_{\min}, v_{\max}]$$

which clearly implies $v^* = v_s^0$ is a mixing point with $N = 1$ and $\varepsilon = (1 - \zeta)$.

Step(2) Stationary distribution is continuous in prices. In this part of the proof I apply the Theorem 12.13 in SLP.

Step(3) There exists a fixed point (equilibrium). Use the labor, capital and goods market clearing conditions to define a mapping $\Phi : \Omega \rightarrow \mathbb{R}^3$, where

$$\Phi(\omega) = \begin{bmatrix} \sum_{s \in S} \Gamma_s \int_V k_s(\omega) d\mu_s(\omega) - A(\omega) \\ \sum_{s \in S} \Gamma_s \int_V n_s(\omega) d\mu_s(\omega) - h \\ \sum_{s \in S} \Gamma_s \int_V \pi(\theta_s) \left[\theta_s^{1-\gamma} F(l_s(\omega)) \right] d\mu_s(v_s) - C^e(\omega) - C^w(\omega) - K(\omega) \end{bmatrix}$$

Note prices r and w have to be greater than zero and without loss of generality I assume they are bounded above by some arbitrarily large number. Thus, Ω compact and convex set. Then define a mapping

$$T(\omega) = \arg \max_{\omega \in \Omega} - \|\Phi(\omega)\|^2 \tag{B.23}$$

By arguments from Step 2 a solution to the problem [B.23](#) is continuous in ω and therefore a correspondence $T(\omega)$ is also continuous. Applying the Shauder Fixed-Point theorem (see SLP Theorem 17.4) establishes the result. \square

B.3 Algorithms

This section presents numerical algorithms used to compute the individual contracting problem, stationary recursive equilibrium and then transitional dynamics. Let me start with the description of the individual contract solution.

Algorithm 11. *The algorithm consists of the following steps*

1. *Discretize space $V = [v_{\min}, v_{\max}]$ using n_V points. I set $n_V = 400$.*
2. *For each $i \in N_V = \{1, \dots, n_V\}$ and each type $s \in S$ guess the derivative of the contract value function $B^0(v) = \{B_s^0(v_i)\}_{i=1, \dots, n_V, s \in S}$.*
3. *Update the guess and obtain $B^0(v)$ using the following procedure:*
 - (a) *Given the initial guess $B^0(v)$ use the necessary and sufficient conditions [B.11](#)-[B.15](#) to solve for $\{c(v_i, \theta_{sL}), c(v_i, \theta_{sH}), v'(v_i, \theta_{sL}), v'(v_i, \theta_{sH}), l(v_i)\}$ for each $i = 1, \dots, n_V, s \in S$. I use the nonlinear equation solver proposed by [Bouaricha and Schnabel \(1994\)](#). Also I use linear splines to approximate $B^0(v)$ at any proposed $v'(v_i, \theta_{sL})$ and $v'(v_i, \theta_{sH})$ by the solver.*
 - (b) *Update the guess using collapsed conditions [B.9](#) and [B.8](#) and envelope condition [B.10](#), i.e.*

$$B_s^1(v_i) = \frac{1}{\beta(1+r)} [\pi_L B_s^0(v'(v_i, \theta_{sL})) + \pi_H B_s^0(v'(v_i, \theta_{sH}))] \quad \forall s \in S, \forall i \in N_V$$

4. *Compare $B_1'(v)$ and $B_0'(v)$ and compute distance d :*

$$d = \max_{i \in \{1, 2, \dots, n_V\}} |B_{s1}'(v_i) - B_{s0}'(v_i)|$$

5. *If $d \leq \varepsilon_c$ then the stop. If $d > \varepsilon_c$ then update the derivative of the value function $B_0'(v) = B_1'(v)$ and go back to Step 3.*

6. Back out the value function $B(v)$ using the contract policy functions found in steps 3 and 4.

A recursive, stationary equilibrium can be computed by finding a pair of prices r and w that solve the pair of equations $G_1(r, w) = 0$ and $G_2(r, w) = 0$. I use a nonlinear equation solver to find them. The function G_1 corresponds to the labor market clearing $N = h$, while the function G_2 corresponds to the market-clearing in the asset market $K = A$. In order to evaluate these function I use the following algorithm.

Algorithm 12. *The algorithm consists of the following steps:*

1. *Guess the initial price vector $\{w_0, r_0\}$.*
2. *Given the prices solve for the optimal contract policies $\{c(v_i, \theta_{sL}), c(v_i, \theta_{sH}), v'(v_i, \theta_{sL}), v'(v_i, \theta_{sH}), l(v_i)$ for each $i = 1, \dots, n_V, s \in S$, and value of the contract $B_s(v_s)$ using Algorithm 11.*
3. *Given the value function obtained in Step 2, solve for the initial utility of entrepreneur using*

$$B_s(v_s^0) = 0$$

4. *Use Monte Carlo simulation to compute a stationary distribution of firms. Simulate a panel of firms (I use $N = 1,000,000$) for a large number of periods T using decision rules and initial utility computed in Steps 2 and 3 until the distribution of firms converges to a stationary one. Specifically, use a (pseudo)random number generator to generate a sequence $\{\theta_t^j\}_{j=1}^N$ and $\{u_t^j\}_{i=1}^N$ for each $t = 0, \dots, T$, where u_t^j are iid over time and firms and each u_t^j is uniformly distributed on $[0, 1]$ interval. Use these simulates sequences generate a simulated sequence of continuation utilities $\{v_t^j\}_{t=0}^T$ for each $j = 1, \dots, N$. In particular, use the following procedure*

- (a) *Set $v_0 = v_s^0$, and age of the firm of type s to be $a_s^j = 0$ and $t = 0$.*
- (b) *If $u_t^j \leq \zeta_t^j$ then firm survives and $v_{t+1} = v'(v_t, \theta_{st})$ and $a_s^j = t$, otherwise the firm is replaced by a new one and $v_{t+1} = v_0, a_s^j = 0$.*
- (c) *Increment t . Iterate on steps (a)-(b) until $t = T$.*

5. Given a sequence of $\{v_t^j\}_{t=0}^T$ compute policy functions using linear spline interpolation and contract policy functions solved on the grid in Step 2.
6. Compute aggregates defined in Section 3.3.6 using the empirical distribution from the simulation at time T (check if the distribution converged).
7. Evaluate $G_1(r, w)$ and $G_2(r, w)$, defined as follows

$$\begin{aligned} G_1(r, w) &= N(r, w) - h(w) \\ G_2(r, w) &= K(r, w) - (1/r)(L(r, w) - P(r, w)) \end{aligned}$$

I compute transitional dynamics in my main quantitative section using the following algorithm.

Algorithm 13. *The algorithm consists of the following steps:*

1. Fix a transition length T , uncertainty shock vector $\{\sigma_t\}_{t=1}^T$ and convergence criterion ε .
2. Solve for the initial stationary equilibrium using Algorithm 12. Denote the initial distribution by μ_0 , initial prices clearing the market by $\{w_0, r_0\}$.
3. Solve for the final stationary equilibrium with prices $\{w_T, r_0\}$. Denote the final distribution by μ_T .
4. Guess a sequence of aggregate labor stocks $\{\hat{H}_t\}_{t=1}^T$ of length T such that $\hat{H}_1 = H$ and $\hat{H}_T = H^{**}$.
5. Back out the vector of wages from household problem. Get from the final stationary equilibrium $B_T(v)$. Solve backward for the policy functions $\{c_t(v_i, \theta_{sL}), c_t(v_i, \theta_{sH}), v_t'(v_i, \theta_{sL}), v_t'(v_i, \theta_{sH})\}$ and further for value of the contract $\{B_t(v)\}_{t=1}^{T-1}$.
6. Compute a sequence of distributions forward using

$$\mu_{s,t+1}(v_{s,t+1}) = \int_V Q(v_{s,t}, v_{s,t+1}) d\mu_{s,t}(v_{s,t})$$

7. Compute the aggregate variables to check market clearings

$$\widehat{N}_t = \sum_s \int_V n_t(v_s) d\mu_{s,t}(v_{s,t})$$

8. Check whether $\max_{1 \leq t \leq T} |\widehat{H}_t - \widehat{N}_t| < \varepsilon$. If the criterion above is not satisfied at every t , update a new guess

$$\widehat{H}_t^{new} = (1 - \phi) \widehat{H}_t + \phi \widehat{N}_t$$

where $0 < \phi \leq 1$ and go back to Step 5.

Note this algorithm does not require running costly Monte Carlo experiment, which is required to track firm's age, simulation for every iteration. Once the market clearing price and policy functions are obtained through this algorithm I can run Monte Carlo simulation just once. To do so proceed as follows.

Algorithm 14. *The algorithm consists of the following steps:*

1. Use the population of firms to which an economy converged in Algorithm 12 as the starting one.
2. Simulate the economy over T periods using decision rules and prices computed using Algorithm 13 in a similar way as in Step 4 of Algorithm 12.
3. Compute the statistics conditional on size and age over transition.