Other People's Money

Essays on Capital Market Frictions

Mario R.C. Bersem

Other People's Money

Essays on Capital Market Frictions

Inauguraldissertation zur Erlangung des akademischen
Grades eines Doktors der Wirtschaftswissenschaften der
Universität Mannheim

durch

MARIO RONALD CHRISTIAN BERSEM

Dekan	Prof. Dr. Martin Peitz
Referenten	Prof. Dr. Ernst-Ludwig von Thadden Prof. Enrico Perotti, PhD Enrique Schroth, PhD
Mündliche Prüfung	13er Juli 2012

Universität Mannheim

Contents

Chapter 1

Introduction

Finance studies the exchange of money for repayment promises—finance is about using *other people's money*. In canonical complete-markets models, financial exchange is no different from the exchange of money for a loaf of bread, say, and the first welfare theorem of economics pertains.[1]

In reality, financial trades can be more problematic: they are trades over time and carry an element of risk, as unforeseen contingencies may arise. Furthermore, financial exchange causes *principal-agent* problems, as borrowers cannot be perfectly controlled by financiers. Capital markets, in other words, are not complete contingent markets.

It follows that canonical models need to be modified by recognizing capital market frictions such as those driven by imperfect information, imperfect enforcement and political intervention. This dissertation investigates capital market frictions across three themes.

The first theme of this dissertation is **sovereign debt**, the topic of chapter 2. Recent experience in the EU shows that enforcement of repayment promises is complex when the borrower is a state. Furthermore, governments are better informed about their repayment capacity than creditors are. Chapter 2 argues that enforcement and information frictions explain why sovereign borrowers issue simple debt contracts that frequently lead to debt crises and debt renegotiations. Such contracts are optimal because they save on costly audits by creditors.

[1] Specifically, if complete contingent contracts are available and markets are free (i.e. in an Arrow-Debreu world), the resulting market allocation is efficient.

The second theme of this dissertation concerns pensions, the topic of chapter 3. It is often argued that **collective pension funds**, such as found in the Netherlands, can enhance the welfare of their participants. Chapter 3 highlights one rationale for pension funds based on credit constraints (i.e. a capital market friction). Chapter 3 next explores an agency problem (i.e. a second friction) that may limit pension funds' ability to increase welfare.

The third theme of this dissertation concerns **political intervention in capital markets**, the topic of chapter 4. Financial liberalization and expanded access to capital are historically seen as signs of greater freedom. Yet many democratic countries choose to restrain the resource allocation called for by a free capital market. Chapter 4 argues that democracies may choose to resist free capital markets depending on demographical context, the concentration of wealth, and the rate of technological progress. In effect, democracies favor income stability over economic growth when the population is older, when the concentration of wealth is uneven, and when the rate of technological progress is high. The rest of this introduction gives summaries of the three main chapters, and describes their methodology, before concluding with avenues for future research based on the current work.

1.1 Chapter Summaries

Each of the chapters 2-4, is a stand-alone contribution and can be read independently of the other chapters. Each chapter starts with an introduction, then presents a model and results, before ending with a conclusion. Proofs are generally relegated to the appendix, unless they are short. In the following, I will try to give a brief, non-technical summary of each chapter.

1.1.1 Sovereign Debt

Chapter 2 is based on Bersem (2012) and focuses on sovereign debt. Debt is a financial contract in which a borrower receives some money and agrees to pay it back at a later date. When the borrower does not repay, creditors obtain certain rights vis-à-vis the borrower's assets, e.g., they may obtain the right to seize and sell the collateral that the borrower posted. Such rights facilitate financial trades and allow the proper functioning of financial markets. Indeed, credit markets are largest in countries where creditor and investor rights are strongest (La Porta et al., 1997;1998).

Sovereign debt is the debt contracted by sovereign borrowers, i.e., states like Greece. Creditor rights are typically difficult to impose on sovereign borrowers. No court can force an unwilling sovereign debtor to repay. Rather, the repayment of sovereign debt is a political question, decided upon by governments based on economic and political considerations—creditors of Greece were duly reminded of this simple fact in the March 2012 default.[2]

Still, sovereign debt markets are huge: there was more than $34 trillion of outstanding sovereign debt in 2009; thereby, sovereign debt accounted for about 40% of the value of global bond markets (Source: Bank of International Settlements, BIS).[3] The literature on sovereign debt has spelled out, both theoretically and empirically, how positive repayment can be sustained, even in the absence of a court. Reinhart and Rogoff (2009) summarize this literature as, 'concerns over future access to capital markets, maintaining trade, and possibly broader international relations all support debt flows.' Chapter 2 builds on the sovereign debt literature by taking as given that positive repayment can be sustained; the chapter seeks to extend the sovereign debt literature by asking an obvious complementary question: what is the form of the optimal borrowing agreement?[4]

Sovereign borrowers can be described as borrowers who can repudiate their repayment obligations at any time, i.e., repay nothing; and, whose exact repayment capacity is hard to observe for outsiders, as it depends on the specific political economic calculus of the government in office. The main question of chapter 2 is how to extend credit to such a borrower. Common sense—and knowledge of history—suggest that creditors need to be careful.[5]

Chapter 2 shows that, given enforcement and information constraints, the op-

[2]This enforcement problem is known, in the literature, as the *willingness-to-pay problem* (Eaton and Gersovitz, 1981).

[3]Greece alone had €350 billion of debt, before its March 2012 default—the biggest sovereign default in history.

[4]Optimal in the sense of minimizing agency costs.

[5]Reinhart and Rogoff (2009) survey eight centuries of financial crises, among which there are many sovereign debt crises.

timal borrowing agreement is a contract that specifies (i) a fixed payment, or *face value*, in high income states, and (ii) a default if the sovereign's willingness-to-pay falls short of the face value, where (iia) default is partial rather than complete, and (iib) the default repayment depends on the power that creditors have to punish repudiation. This result explains three salient facts of sovereign borrowing. First, the result explains why sovereign borrowers, choose to issue simple debt instruments instead of more contingent contracts, as Shleifer (2003) and others have argued they should (cf. Borensztein and Mauro, 2004). Such contracts are not optimal, because the auditing requirements would be prohibitively costly, i.e., the agency costs are too high for such contracts.

Furthermore, as Reinhart and Rogoff (2009) document, some countries default at very low debt-to-GDP levels; other countries continue to repay their debt at very high debt-to-GDP levels.[6] Chapter 2 explains this empirical finding by pointing to four factors that determine the government's repayment decision: (i) the available budget, (ii) the economic costs of repudiation (i.e. creditor power), (iii) the political costs of default, and (iv) creditor coordination costs.

Finally, conditional on default, there is a wide dispersion on how much creditors recover from the sovereign borrower (Sturzenegger and Zettelmeyer, 2006). Chapter 2 explains this empirical finding by pointing to creditor power as the relevant determinant of recovery. The most powerful creditor is the International Monetary Fund (IMF); historically, the IMF takes priority over all other creditors. It follows that, even if rates on IMF loans are lower than on other loans, IMF lending is not concessionary: the IMF simply expects to be repaid with higher probability, and to recover more in case of a default.

Chapter 2 shows that an increase in the costs of repudiation, be they political or economic costs, lowers the interest rate on sovereign debt through a commitment effect: higher costs of repudiation commit the sovereign to repay the debt at face value in more states of the world; thus, reducing sovereign risk.

[6] A similar picture emerges for other measures of a sovereign's ability to pay.

1.1.2 Collective Pension Funds

Chapter 3 is based on Bersem and Hollanders (2012) and focuses on collective pension funds. A pension is a payment stream that people receive upon retirement, i.e., when they leave the labor force. Rather than leaving it to individuals to save for their retirement, most advanced economies have pension systems in which individuals are required to participate. Common to such systems is that people contribute in their active working years, which entitles them to a pension benefit upon retirement.[7] But there is considerable variation between countries in how the pension system operates, how it is financed, and how pension benefits are determined.[8]

For example, some countries, like Germany, operate pay-as-you-go (PAYG) pension systems, where the active working population pays for the current retirees. Other countries, like the Netherlands, operate additional prefunded schemes, in which people *save* through pension contributions and receive a pension benefit that is set according to the pension contract. Roughly, one can distinguish two types of pension contracts: *defined-contribution* (DC) type contracts, where the pension benefit depends explicitly on investment returns (e.g. the famous 401(k) plans in the U.S.), and *defined-benefit* (DB) type contracts, where the pension is set according to a formula that may depend on average pay, years of employment, age at retirement, and other factors (e.g. the second pillar in the Netherlands).

Prefunded DB pension schemes run into trouble when they are *underfunded*, i.e., when pension liabilities, which are fixed, exceed pension assets, which may fluctuate. Such pension shortfalls are an inherent risk—and recurring feature—of funded DB pension schemes.[9] After the 2008 credit crisis, more than half of the pension funds in the Netherlands were underfunded. The decline in pension wealth led to controversy over who should pay to restore the solvency of

[7]This seems an obvious requirement, but note that the first generation in a pay-as-you-go (PAYG) pension system receives a pension without having paid contributions.

[8]World Bank (1994) gives a useful categorization of pension systems into three pillars: a *state pension*, aimed at poverty reduction, and financed through taxes; an *occupational pension*, aimed to maintain the standard of living, and prefunded; and a *private pension*, allowing for individual supplements, also prefunded.

[9]For example, General Motor's defined-benefit pension plans reported a shortfall of $35 billion in 2011; this exceeded GM's market value.

the pension system—Dutch regulations required a return to solvency within 5 years.[10] In recognition of the intrinsic tensions in DB pension systems with mismatch risk, DB pension schemes are being replaced with DC pension schemes, cf. Goudzwaard et al. (2009). This then leads to the question how such DC schemes perform. Chapter 3 explores one rationale for a DC pension scheme that has received little attention in the literature: that pension funds exist to lift credit constraints and implement the optimal optimal life-cycle investment strategy of participating generations.[11]

The literature on modern life-cycle investment theory shows that individuals' optimal investment strategy depends on their age: the young—who have human capital as well as financial capital—should invest their financial capital in a riskier manner than retirees—who have only financial capital (Bodie, Merton, and Samuelson, 1992). With their human capital, the young are naturally hedged against stock market risk. Typically, these models require the young to take a leveraged position in the stock market, i.e., to borrow and invest the proceeds in the stock market. If the young face credit constraints, this strategy is not feasible and pension funds have a role to play: pension funds can implement the young's preferred investment strategy by extending them credit.[12] There are two reasons why pension funds are better placed than private sector lenders to extend credit to the young: (i) participation is mandatory, reducing the adverse selection problem; and (ii) pension funds have access to a tax on human wealth, which allows them to enforce repayment. In effect, the pension fund helps to secure the human capital of participants as collateral (Bovenberg et al., 2007).

Chapter 3 shows that implementation of optimal investment strategies can in fact be achieved by a DC pension scheme, where participants pay contributions into a *generational account* (e.g. yearly), and pension funds invest these contributions on behalf of participants. The ability of pension funds to increase welfare is dependent on the assumption that pension funds can collateralize the human cap-

[10]By cutting entitlements, pension funds decrease their liabilities; by raising contributions—while keeping entitlements fixed—pension funds increase their assets.

[11]The notable exception is Bovenberg et al. (2007) ; Teulings and de Vries (2006) mention but do not pursue this rationale for pension funds.

[12]Bovenberg et al. (2007) provide a review of different rationales for pension funds; the review discusses alleviation of borrowing constraints as one possible rationale.

ital of participants. Recent experience in the Netherlands shows that this assumption canot be taken for granted, as it may prove impossible to raise contributions after a low stock market outcome—and in particular in a recession. It follows that the DC pension scheme we describe may run into the same problems as a DB pension scheme with mismatch risk, and that the ex-ante optimal risk level at the pension fund cannot be separated from the ex post contribution policy.

Indeed, one suggestive interpretation of the distributional conflict witnessed in the Netherlands is that pension funds—which, de facto, run a combination of DB and DC pension contracts—took risk on behalf of the young, assuming that contributions could be raised in case of a pension shortfall. When this proved infeasible, a controversy over who should pay for the shortfall was the result. This *ex post* distributional conflict leads to an *ex ante* governance conflict at the pension fund: older participants wish to limit risk taking such that they are repaid in every contingency.

1.1.3 Political Intervention in Capital Markets

Chapter 4 is based on Bersem, Perotti, and von Thadden (2012) and focuses on political intervention in capital markets. Capital markets, also denoted financial markets, are markets where banks and other financial intermediaries trade financial securities; they are markets through which the financial sector moves money from point A, where it is, to point B, where it is needed. In the ideal type of the free capital market, capital moves naturally towards its most profitable use and, by allowing the financing of new ventures, keeps alive the process of *creative destruction*, whereby old firms and organizations are constantly challenged and replaced by new ones (Schumpeter, 1942). In reality, however, the functioning of capital markets varies greatly between countries: it is much harder to obtain financing for new, daring, disruptive ideas in some countries than it is in others.

(Wurgler, 2000) has shown that industries with better growth prospects are able to invest more in countries that are more financially developed; these are also the countries in which declining sectors shrink faster.[13] Rajan and Zingales

[13]Wurgler (2000) finds, furthermore, that (i) a high degree of minority investor protection, and (ii) a lesser extent of state ownership in the economy are both associated with a better allocation of capital; evidence suggestive of the importance of political influence on capital markets.

(1998) show that industries that are more reliant on external finance, grow faster in countries where financial markets are more developed. Indeed, this was already observed by Walter Bagehot in his famous book *Lombard Street: A Description of the Money Market*:[14]

> "Political economists say that capital sets towards the most profitable trades, and that it rapidly leaves the less profitable non-paying trades. But in ordinary countries this is a slow process."

But what can account for these differences? As Rajan and Zingales (2003) point out, markets rely on the political goodwill for their infrastructure (which includes, e.g., the rule of law and how it is enforced).

The main contribution of chapter 4 is to explain why unrestricted capital mobility may be opposed in democracies as a result of the wealth and age distribution in a country's population. The starting insight of the chapter is that labor is less mobile than capital: while capital can easily be redeployed—think of land or real estate—it is hard to retrain workers once they have acquired specific skills. The result is a political conflict between generations (old vs. young); and within generations (workers vs. the capitalists). Capitalists favor the reallocation of capital to its most productive use. Old workers, who have outdated skills, resist the reallocation of capital to newer sectors, as this leads to a fall in their productivity and wages; old workers seek a political alliance to restrict capital mobility. Chapter 4 identifies young workers as the decisive class in society. What makes them pivotal is not their number—they are a minority just as every other voter class—but the fact that their preferences are the least extreme.

The preferences of the young worker depend crucially on the voting process: If capital market frictions can be repealed in the future, young workers will not favor them; but, if capital market frictions are permanent, young workers form an alliance with old workers to restrict capital mobility. The intuition is that young workers trade off lower wages when young against a job guarantee (i.e. higher wages) when old. Young voters prefer to restrict capital mobility more if technological obsolescence is high, as this increases the wage drop when old.

[14]Bagehot was an English banker and editor to *the Economist* newspaper. Lombard Street (1873) describes the world of finance, and the role of central banks, in common language.

Chapter 4 shows that opposition against capital markets can be sustained if frictions are hard to reverse. There are clear examples of institutional frictions in capital markets which are hard to reverse. Bankruptcy law, for example, defines specific conditions to the assignment of assets from declining sectors. While in some countries—such as the United Kingdom—bankruptcy law is designed to protect financial interests, in others—like France and Italy—it explicitly instructs the liquidator to reassign capital in a manner which protects employment.

1.2 Methodology

This dissertation studies financial contracting across different themes. The literature in finance has long recognized how the neoclassic paradigm of enforceable complete contracts needs to be modified by recognizing frictions such as those driven by imperfect information, imperfect enforcement, and political intervention. This dissertation fits into this research agenda.

Each chapter of this dissertation presents one economic model, a set of ideas about some specific aspect of capital markets. They cover the design of sovereign debt, pension finance, and the process of capital reallocation. They are motivated by empirical observations which are hard to square within the neoclassical paradigm.

The presentation of the models in this dissertation is formal: built on a specific set of pertinent assumptions, each chapter derives rigorously a specific set of implications. In choosing the approach for each chapter, the guiding force was the nature of the question at hand. With Occam's razor in mind, I've tried to make the models as simple as possible, but not simpler. It is important to state as a reminder that economic models are an abstraction, like a map, but that the map is not the territory.[15]

1.2.1 Overview

Chapter 2 (*Incentive-Compatible Sovereign Debt*) seeks to make a positive contribution by explaining why sovereign borrowers issue very simple debt contracts.

[15] John Kay - 'The Map is not the Territory: An Essay on the State of Economics.'

After all, national economies and thus fiscal capacity have specific risk exposures, which may be best hedged in their financing. Yet this is almost never the case. We pursue an explanation driven by the institutional constraints imposed by sovereignty, which limits direct contractual enforcement, and the superior information held by governments over private parties. We show how the optimal government debt contract resolves these constraints in a simple form.

Chapter 3 (*Collective Pension Funds*) seeks to make a normative contribution by showing what features of pension funds may enhance welfare. Pensions can improve intergenerational risk sharing by increasing the ability of the young to sustain capital investment. They could not achieve the same privately, as adverse selection and moral hazard limits their ability to borrow against their human capital to invest in risky, high return assets. The analysis is built on a simple modeling of the underlying tension between generations in the process.

Chapter 4 (*Sand in the Wheels of Capitalism*) seeks to explain why democratic societies may oppose free capital markets. Financial liberalization and expanded access to capital for new enterprises are historically seen as signs of greater freedom. Yet many democratic countries choose to restrain or contain the process of free resource allocation called for by prices set on a free capital markets. The literature has shown that such choices may affect growth and the rate of innovation. The model here shows that the redistributive effect of a technological shock creates a strong political demand to limit capital reallocation away from obsolete sectors. It shows that income stability may be chosen above economic growth when certain conditions allow the creation of persistent frictions to capital reallocation. So the model offers an endogenous explanation for the existence of avoidable financial frictions.

In the following I discuss the methodology of each chapter in more detail, before concluding with avenues for future research.

Chapter 2 (Incentive-Compatible Sovereign Debt)

Chapter 2 uses financial contract theory to model the interaction between a sovereign borrower and potential financiers (Hart, 2001). Specifically, I use a version of the costly state audit model, which goes back to Townsend (1979) and Gale and

Hellwig (1985).

The basic setup is familiar from corporate finance models: an agent seeks financing from a group of financiers. There are gains from trade. The question is whether they can be realized, and if so with what contract.

Two frictions complicate the trade: (i) there is asymmetric information that can be resolved only at a cost; and (ii) there is no court to enforce repayment by the agent. The second assumption is made to capture that the agent is sovereign, rather than a corporation. Other assumptions particular to this sovereign finance model: the cost of audit is borne by the agent, future income is an endowment.

The optimal contract I derive saves on audit costs and implements the second-best allocation. (The first-best cannot be achieved due to prevailing enforcement and information frictions.) Chapter 2 argues that this contract matches some empirical facts of sovereign borrowing.

Chapter 3 (Collective Pension Funds)

Chapter 3 models the interaction between individuals of different age using an overlapping-generations (OLG) model, which goes back to **?**.

Chapter 3 focuses on optimal investment and, by fixing savings, abstracts from individuals' labor supply and consumption decision, as in Gollier (2008). There is a capital market to which all individuals have access. The problem reduces to choosing an optimal investment strategy. The question is whether this strategy can be implemented.

Chapter 3 introduces a friction: future savings cannot be used as collateral, which implies that the young are credit constrained. The assumption captures that human capital does not collateralize loans for adverse selection and moral hazard reasons (Constantinides et al., 2002). A pension fund can improve on the market allocation if (i) participation is mandatory; and (ii) it has access to a tax on human wealth. In effect, the pension fund works as a commitment device for the young to pledge their human capital as collateral to older generations.

Credibility of the commitment device is crucial for the result. Chapter 3 argues that when this assumption becomes problematic—in light of recent experience in the Netherlands—it reduces the scope for pension funds to increase welfare.

Chapter 4 (Sand in the Wheels of Capitalism)

Chapter 4 models the economic and political interaction between different generations with an OLG model that is extended with a simple majority vote (Persson and Tabellini, 2000). Redistributive effects of policy on labor and capital returns are at the heart of political economy explanations for the structure of the economy.

Chapter 4 assumes vintage human capital, which means that the labor market is segmented. This assumption captures the difficulty of retraining workers once they have specific skills. A realistic second type of heterogeneity arise because the capital is largely owned by a subset of the population (the *capitalists*). The capital market reallocates capital across firms, operating under conditions set by political decisions.

The economic rationale for capital reallocation is that new sectors, where the young work, are more productive than older sectors. The old wish to block capital reallocation as it reduces their wages. The political conflict exists as long as capital and labor are complementary factors of production; as long as human capital is less mobile than physical capital; and as long as human capital risk cannot be fully insured.

In each period, a vote takes place: individuals can choose a capital market friction, which slows down the subsequent reallocation of capital in the economy. The question is whether such a friction will be chosen under majority rule.

In the Chapter we study different specifications of the political model, both in terms of possible voting strategies (open-loop, subgame perfect, and markov perfect) and in terms of the persistence of the chosen policy (persistent vs. reversible). Our aim is to understand by what type of policies capital market frictions can be sustained. We show that capital market frictions are politically sustained when redistributive effects are strong, only and only if persistent frictions can be established.

As usual in political economy models, one may wonder why economically suboptimal outcomes cannot be resolved by bargaining. Shouldn't it be possible to compensate the old workers for allowing capital reallocation? Prohibiting this type of efficient bargaining is the hold-up problem associated with the relinquishment of power, a core issue in political economy and the source of much inefficiency (Acemoglu, 2003).

1.3 Future Research

Chapter 2 (*Incentive-Compatible Sovereign Debt*) is the starting point of an extensive theoretical and empirical research agenda in sovereign debt. Empirically, the cross-sectional implications of the model must be subjected to rigorous statistical testing. Do shifts in political power or ultimate holdings of sovereign debt lead to the secondary price responses that the model predicts? Recent events in the European Union suggest that they do, as shifts in political power were consistently followed by secondary market responses. Theoretically, the framework must be extended to develop a fully dynamic model of sovereign debt; a model that endogenizes the cost of repudiation and allows the study of repayment and refinancing decisions in one unifying framework.

Chapter 3 (*Collective Pension Funds*) explores one rationale for prefunded pension funds; future work must include others. In particular, I've abstracted from intergenerational risk-sharing between non-overlapping generations. It is an important open question whether such risk-sharing is best achieved via government debt and tax policies, as in **?**, or via pension funds, as in Gollier (2008). The current framework can be extended to study both in a dynamic OLG setting and compare, by calibration, the performance of different pension schemes.

Chapter 4 (*Sand in the Wheels of Capitalism*) predicts that opposition to free capital markets is strongest in democracies with narrow capital market participation, and with older populations. Broad capital market participation is found in some democracies, in particular those with funded pension schemes. Empirical tests must show if capital reallocation is less restricted in democracies with fully funded pension systems. Ageing populations form another testing ground for our theory. Empirical tests must show if capital reallocation is more restricted in democracies with older populations. Finally, as noted, capital market frictions cannot be bargained away. By broadening capital market participation, however, capitalists could change the young worker's preference. This is a possible extension of the theory reminiscent of Rajan (2010), who argues persuasively that credit expansion has historically been used to assuage the concerns of a group that is left behind.

References

ACEMOGLU, D. (2003): "Why not a political Coase theorem? Social conflict, commitment, and politics," *Journal of Comparative Economics*, 31, 620–652.

ARROW, K. J., AND G. DEBREU (1954): "Existence of an Equilibrium for a Competitive Economy," *Econometrica*, 22(3).

BAGEHOT, W. (1873): *Lombard Street: A Description of the Money Market*. King & Co, London.

BERSEM, M. (2012): "Incentive-Compatible Sovereign Debt," Discussion paper, University of Mannheim.

BERSEM, M., AND D. HOLLANDERS (2012): "Prefunded Pension Schemes: a Conflict over Risk and Contributions," Discussion paper, University of Mannheim.

BERSEM, M., E. PEROTTI, AND E.-L. VON THADDEN (2012): "Sand in the Wheels of Capitalism: on the Political Economy of Capital Market Frictions," Discussion paper, University of Mannheim.

BODIE, Z., R. MERTON, AND W. SAMUELSON (1992): "Labor supply flexibility and portfolio choice in a life cycle model," *Journal of Economic Dynamics and Control*, 16(3-4), 427–449.

BORENSZTEIN, E., AND P. MAURO (2004): "The case for GDP-indexed bonds," *Economic Policy*, 19(38), 165–216.

EATON, J., AND M. GERSOVITZ (1981): "Debt with Potential Repudiation: Theoretical and Empirical Analysis," *Review of Economic Studies*, 48(2), 289–309.

GALE, D., AND M. HELLWIG (1985): "Incentive-Compatible Debt Contracts: The One-Period Problem," *Review of Economic Studies*, 52, 647–663.

GOLLIER, C. (2008): "Intergenerational risk-sharing and risk-taking of a pension fund," *Journal of Public Economics*, 92, 1463 – 1485.

HART, O. (1995): *Firms, Contracts, and Financial Structure*. Oxford University Press (USA).

HART, O. D. (2001): "Financial Contracting," *Journal of Economic Literature*, 39(4), 1079–1100.

PERSSON, T., AND G. TABELLINI (2000): *Political Economics: Explaining Economic Policy*. The MIT Press, Cambridge, Massachusetts.

PORTA, R. L., F. LOPEZ-DE-SILANES, A. SHLEIFER, AND R. W. VISHNY (1997): "Legal Determinants of External Finance," *The Journal of Finance*, 52(3), pp. 1131–1150.

PORTA, R. L., F. LOPEZ-DE SILANES, A. SHLEIFER, AND R. W. VISHNY (1998): "Law and Finance," *Journal of Political Economy*, 106(6), pp. 1113–1155.

RAJAN, R. (2010): *Fault Lines: How Hidden Fractures Still Threaten the World Economy*. Princeton University Press.

RAJAN, R. G., AND L. ZINGALES (1998): "Financial Dependence and Growth," *American Economic Review*, 88(3), 559–586.

——— (2003): *Saving Capitalism from the Capitalists*. Crown Business, New York.

REINHART, C. M., AND K. S. ROGOFF (2009): *This Time is Different: Eight Centuries of Financial Folly*. Princeton University Press, Princeton and Oxford.

SCHUMPETER, J. A. (1942): *Capitalism, Socialism, and Democracy*. New York: Harper Brothers.

SHLEIFER, A. (2003): "Will the Sovereign Debt Market Survive?," *American Economic Review Papers and Proceedings*, 93(2), 85–90.

STURZENEGGER, F., AND J. ZETTELMEYER (2006): *Debt Defaults and Lessons from a Decade of Crises*. MIT Press, Cambridge.

TEULINGS, C., AND C. DE VRIES (2006): "Generational Accounting, Solidarity and Pension Losses," *de Economist*, 146, 63–83.

TOWNSEND, R. M. (1979): "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, 21, 265–293.

WURGLER (2000): "Financial Markets and the Allocation of capital," *Journal of Financial Economics*, 58, 187–214.

Chapter 2

Incentive-Compatible Sovereign Debt[1]

Abstract. In a model of sovereign borrowing and lending—a model with asymmetric information, costly state disclosure, and no court to enforce repayment—I show that a sovereign borrower optimally issues a contract that specifies (i) a fixed payment, or face value, in high income states, and (ii) a default if the sovereign's willingness-to-pay falls short of the face value, where (iii) default is partial rather than complete, and (iv) the default repayment depends on the power that creditors have to punish repudiation. The result explains why sovereign borrowers issue simple debt instruments instead of more contingent contracts. An increase in the costs of repudiation lowers the interest rate on sovereign debt through a commitment effect: higher costs of repudiation, both political an economic, commit the sovereign to repay the debt at face value in more states of the world; thus, reducing sovereign risk.

[1] I am indebted to Ernst-Ludwig von Thadden, Enrico Perotti, and Enrique Schroth for their advice. I thank Adriano Rampini, Andras Niedermayer, André Stenzel, Arnoud Boot, Ernst Maug, John Moore, József Sákovics, Knut Heen, Klaus Adam, Malin Arve, Mike Burkart, Nicolas Schutz, Petra Loerke, Philipp Zahn, and seminar audiences at the Stockholm School of Economics, the University of Vienna, the University of Edinburgh, and Copenhagen Business School for comments.

2.1 Introduction

This chapter analyses credit extension to a borrower in the absence of a court. Creditors can punish *repudiation* by the borrower, but they cannot seize any of the borrower's income. The borrower is also better informed as to her income. Creditors can get informed if the borrower agrees to a *public audit*, which is costly. The problem occurs naturally in sovereign borrowing. Sovereign debt contracts have been shown to be hard to enforce, not least because governments have private information, and are reluctant to subject to a public audit of their books by the International Monetary Fund (IMF), or fellow member states. In this chapter, I adopt a simple model of sovereign borrowing and lending to answer the following question: if a government seeks to finance an expenditure today, but receives income only in the future, *what is the optimal financial contract* the government can offer to international creditors?

In an Arrow-Debreu world, with complete contingent contracts, the question is easily answered: the optimal contract is either indeterminate or the optimal contract doesn't exist–depending on whether expected income exceeds the expenditure. The market for sovereign finance, by contrast, is plagued by at least two frictions. First, there is an enforcement friction: there is little collateral, and seizure of sovereign assets is complicated.[2] It follows that a sovereign borrower can repudiate any contract she has entered and repay zero to her creditors; this is known as the *willingness-to-pay problem.* Why a sovereign borrower ever chooses repayment over repudiation, in the absence of a court, is a central question of the sovereign debt literature. The sovereign pays for two reasons in this chapter. First, the sovereign pays because she is concerned with the economic costs of repudiation. In line with the literature on the willingness-to-pay problem, I assume that repudiation is economically costly.[3] Second, and novel to my

[2]Domestic courts are subject to the laws of the sovereign and, therefore, cannot be used to force the sovereign to repay. As for outside courts, there are few assets located abroad, and those that are located abroad are often protected by sovereign immunity, cf. Sturzenegger and Zettelmeyer (cf. 2006).

[3]In a seminal contribution, Eaton and Gersovitz (1981) show that reputational concerns can sustain positive repayment by the sovereign; other papers that show how positive repayment can be sustained, in the absence of a court, include Grossman and Van Huyck (1988), Bulow and Rogoff (1989a,b), Worrall (1990), Fernandez and Rosenthal (1990), Atkeson (1991), Cole and

model, the sovereign pays because she is concerned with the political costs of repudiation: the sovereign is forced to resign if she repudiates.[4]

Aside from the enforcement friction, there is an information friction: the sovereign has private information about her income, or *ability-to-pay*. The sovereign can disclose her true ability-to-pay to the creditors, but this is costly: the government has to invite an outside auditor, like the IMF, and dislikes the increased scrutiny and interference that follow a public audit. Building on the enforcement friction and the information friction, I propose a new theory of sovereign debt to explain why sovereign borrowers issue simple debt instruments instead of more contingent contracts

The optimal contract specifies (i) a fixed payment, or *face value*, in unaudited states; (ii) a payment equal to the creditor punishment threat in audited states; and (iii) an audit if and only if the sovereign's willingness-to-pay falls short of the face value The intuition for the optimal contract is that it economises on costly auditing, which is what the costly state verification literature has emphasised (cf. Townsend, 1979; Gale and Hellwig, 1985). Compared to the familiar standard debt contract, there is still a fixed payment in high-income states, i.e., the optimal contract is still a *debt contract*. But the optimal contract specifies partial repayment for audited states, rather than full repayment (or *maximum recovery*); in addition, the usual budget constraint is replaced by a willingness-to-pay constraint.

It is natural to interpret a public audit as a *sovereign default* episode. Examples abound: Russia defaulted in 1998, Pakistan in 1999, Argentina in 2001. Indeed, default episodes are politically costly to governments who are likely to lose office, as Borensztein and Panizza (2009) document. Default episodes also involve a transfer of information: more information comes available through, e.g., IMF reports and increased press coverage.[5]

Kehoe (1998), Kletzer and Wright (2000), Sandleris (2008), Guembel and Sussman (2009), and Hellwig and Lorenzoni (2009). Eaton and Fernandez (1995) and Reinhart and Rogoff (2009) survey the literature.

[4]The political cost is best thought of as a punishment by the electorate for the economic hardship suffered. Indeed, governments are often forced to resign in the wake of a default, as Argentina's government was in 2001. See Blustein (2005) for a detailed account.

[5]The 2010 debt crisis in Greece serves as a case in point: details on its tax-collection system, and the size of its public sector entitlements, became widely known after the EU-IMF-ECB bailout.

With the default interpretation, the characteristics of the optimal contract match some key facts of sovereign borrowing: first, the sovereign's default decision depends on her willingness-to-pay, rather than on her solvency;[6] second, default is partial rather than complete, and creditors get a haircut that depends on their power;[7] and third, countries issue plain bonds that promise a fixed payment.[8]

A further result is that an increase in the political cost of repudiation can increase welfare by alleviating the inefficiency due to the enforcement friction. High repudiation costs work as a commitment to repay in the absence of formal outside enforcement: the government is committed to repay the debt at face value in more states of the world.

In section 2.5, I extend the basic model to study the role of creditor coordination costs, along the lines of Bolton and Jeanne (2009). Clearly, creditor coordination is an important issue in sovereign debt renegotiations. If creditors cannot coordinate around a debt renegotiation, then such renegotiation breaks down, leading to a deadweight loss. Bolton and Jeanne (2007, 2009) take this observation to an extreme by assuming that sovereign debt can either be renegotiated at no cost, or not at all. I show that the debt contracts that are available to the sovereign *by assumption* in Bolton and Jeanne (2007, 2009), can be derived as optimal contracts: non-renegotiable debt is the optimal contract if an audit and creditor coordination are both costly; renegotiable debt is optimal if an audit is costly, but subsequent creditor coordination is costless

This chapter is related to theories of debt in the corporate finance literature, in particular to the costly state verification models pioneered by Townsend (1979) and Gale and Hellwig (1985).[9] My approach is new in combining the well-known

Greece, however, has not defaulted on its creditors at the time of writing. In light of the model in section 3.2, Greece has chosen to be audited, but has not yet made its repayment decision.

[6]Reinhart and Rogoff (2009) document a wide dispersion of debt-to-GDP levels at time of default, and argue that a sovereign's willingness-to-pay, rather than her ability-to-pay, determines the repayment decision.

[7]All creditors are not equal. As there is no court to enforce creditor priority, creditors can expect to be repaid according to the power they wield. The IMF, for example, is typically repaid in full, whereas private creditors receive a haircut.

[8]All parties involved understand that plain bonds are, implicitly, contingent on the state of the world. Still, the prevalence of plain bonds, instead of more explicitly contingent contracts, in sovereign borrowing is puzzling, as Borensztein and Mauro (2004) and Shleifer (2003) have argued.

[9]Other papers that study optimal contracting under *costly state verification* are Border and

costly state verification approach with an ex-post repayment decision, i.e. with the willingness-to-pay problem. A second innovation is that the audit cost is political: the sovereign dislikes to disclose its true ability-to-pay. The political reluctance to disclose, the political cost of repudiation, and the economic cost of repudiation drive the optimal contract design.

Conceptually, this chapter is close to Gale and Hellwig (1989) who consider a model of sovereign borrowing with asymmetric information and a willingness-to-pay problem. But Gale and Hellwig (1989) study the outcome of the ex-post debt renegotiation; they are not concerned with the ex-ante optimal contract design as I am here.[10]

Finally, this chapter is related to Bolton and Jeanne (2007, 2009) who argue that the sovereign debt market–left to itself–produces equilibria in which the sovereign debt structure is excessively hard to restructure. In both papers the sovereign, by assumption, can issue two types of debt: debt that is renegotiable (r-debt), and debt that is not renegotiable (n-debt). If the government is truly unable to repay, renegotiable debt allows for an efficient renegotiation of the debt burden, while non-renegotiable debt leads to a dead-weight loss. Still, the sovereign may choose to issue non-renegotiable debt because it offers some commitment value: n-debt strengthens the sovereign's repayment incentives in Bolton and Jeanne (2007), and n-debt cannot be diluted by subsequent debt issues in Bolton and Jeanne (2009). In this chapter, both renegotiable debt and non-renegotiable debt emerge as optimal contracts, see section 2.5.

Sobel (1987); Mookherjee and Png (1989); Krasa and Villamil (1994, 2000); Hvide and Leite (2010). Other papers that show the optimality of debt under some form of enforceability include Diamond (1984), Innes (1990), and Hart and Moore (1998).

[10]Specifically, Gale and Hellwig (1989) model debt renegotiation under asymmetric information as a signalling game: first the borrower decides how much to repay, then the creditors choose whether to accept the payment or punish the borrower and seize some output. As creditors can always use their punishment technology, the initial contract does not matter in Gale and Hellwig (1989). By contrast, in this chapter creditors can only punish the sovereign debtor if there is a breach of contract and output cannot be seized.

2.2 Model: A Simple Borrowing Problem

Consider a small open economy over two periods: the present ($t = 0$) and the future ($t = 1$). There is a single homogeneous good that can be consumed or invested. A sovereign government, or *sovereign*, seeks to finance a fixed government expenditure, $g > 0$, at time 0; the government expenditure benefits all residents in the economy equally.[11] As the sovereign has no funds at time 0, she seeks to raise the full amount from international creditors, in return for a promise to repay at date 1. A continuum of risk-neutral creditors provides funds at the prevailing opportunity cost of capital, normalised to 0. The sovereign seeks to borrow from a mass one subset of the creditors.

The sovereign's budget at date 1 is uncertain as of date 0. Budget uncertainty arises because future output is uncertain, as is the sovereign's ability to tax output, cut expenses, or generate income from other sources, e.g., from privatising state property, or undertaking structural reforms. The sovereign's budget, or *ability-to-pay*, is denoted by y, a random variable that takes values in an interval $T \subseteq \mathbb{R}_+$ and is distributed according to a cumulative distribution function $F(y)$.

Two frictions limit the efficiency of sovereign borrowing. The first friction arises from asymmetric information: while the sovereign observes y at no cost, outside creditors only observe y if the sovereign subjects to a public audit. If there is a public audit, the country comes under international public scrutiny, led by the IMF, and creditors learn about the sovereign's ability-to-pay. The public audit is costly to the sovereign: the sovereign faces interference with her policies and, as a result of increased transparency, possibly loses office.

The second friction arises from the lack of enforcement in the sovereign finance market: a sovereign borrower can repudiate any contract and repay 0. In line with the literature on the willingness-to-pay problem, I assume that repudiation is economically costly. This cost should be thought of as arising, either, from direct creditor sanctions, as in Sachs and Cohen (1982); or, from a loss of market access, as in Eaton and Gersovitz (1981).

The sovereign maximises the utility of the representative resident, and enjoys

[11] The government expenditure can be thought of as public consumption, as the expenditure does not raise future productivity of the economy. This assumption is not crucial for any of my results, but plausible in the context of sovereign borrowing, cf. also Bolton and Jeanne (2009).

a private benefit from holding office. The utility of the sovereign is given by

$$U_S = \chi_g V + c + S(y, \hat{y}) \tag{2.2.1}$$

where the first two terms capture the utility of the representative resident: χ_g is an indicator that equals 1 if the expenditure is financed; V represents the utility value the residents derive from the expenditure at date 0; and c is consumption at date 1, i.e. income net of any repayment to creditors, or punishment for repudiation. The third term, $S(y, \hat{y}) \geq 0$, is a non-pecuniary private of holding office, which is enjoyed depending on the sovereign's announcement and repayment decision at date 1. The sovereign enjoys the biggest private benefit if she repays without a public audit; she enjoys a smaller private benefit if she repays after a public audit; finally, she enjoys no private benefit if she repudiates–repudiation is costly, and so is a public audit. This is represented by a step-function

$$S(y, \hat{y}) = B\chi_{\{no\,audit, repayment\}}(y, \hat{y}) + b\chi_{\{audit, repayment\}}(y, \hat{y})$$

where χ is an indicator function, $B > 0$, and $0 < b \leq B$.

In autarky, residents consume y as it comes available and the sovereign enjoys her full private benefit from holding office; her expected utility at date 0 is

$$EU_S^{aut} = Ey + B \tag{2.2.2}$$

I assume that financing the government expenditure is efficient, $g < V$, and that the sovereign's expected income exceeds the expenditure, $Ey > g$. These assumptions ensure, first, that the sovereign *wants to* finance the expenditure expenditure and, second, that the sovereign *is able to* finance the expenditure if information is symmetric and enforcement is complete–i.e. in a first-best world. The sovereign's first-best expected utility is given by

$$EU_S^{FB} = V + Ey - g + B \tag{2.2.3}$$

Any contract that satisfies the budget constraint and, in expectation, pays out g to creditors implements the first-best allocation.

With asymmetric information and no formal outside enforcement–i.e. in a second-best world–the interaction between the sovereign and international creditors is as follows. At date 0, *the financing stage*, the sovereign seeks to finance the expenditure by offering a contract to creditors. The contract determines (i) the sovereign's contractual payment obligation in each state; and (ii) what states are to be audited. Formally, a *contract* is defined as an array (O_1, I_d), where $O_1 = O_1(y)$ gives the date 1 contractual obligation as a function of the budget, and $I_d = I_d(y)$ is an indicator that equals 1 if there is an audit, 0 otherwise. At date 1, *the repayment stage*, the sequence of actions is as follows:

1. Nature chooses the state y, sovereign observes y ;
2. **Announcement:** Sovereign announces her ability-to-pay $\hat{y}$,
   (a) if $I_d(\hat{y}) = 1$, then creditors observe y, and the contractual obligation is $O_1(y)$;
   (b) if $I_d(\hat{y}) = 0$, then the contractual obligation is $O_1(\hat{y})$.
3. **Payment:** Sovereign makes a repayment decision $r \in \{0, 1\}$,
   (a) $r = 1$: she pays O_1, i.e. honours the contract, and the games ends, or
   (b) $r = 0$: she pays 0, i.e. repudiates, and creditors charge a punishment.

The contractual obligation of the sovereign is fully determined by her announcement at the repayment stage. If the sovereign announces a state for which the contract specifies an audit, then creditors observe the budget and the contractual obligation is set at $O_1(y)$. If instead the sovereign announces a state that remains unaudited, then based on the announcement the contractual obligation is set at $O_1(\hat{y})$. Finally, the sovereign makes her payment decision: she can either repudiate the contractual obligation she entered, or honour it. An outside arbitrator, with the same information as creditors, certifies whether the sovereign has honoured or repudiated her contractual obligation.[12] If the sovereign honours the contract,

[12]The IMF plays an important role in sovereign default episodes and subsequent debt renegotiations. For example, an IMF program is prerequisite to a renegotiation of any Paris club debt (i.e. debt owed to creditor nations). Furthermore, IMF data form the basis for debt renegotiations with

then investors have no further claim against her. If, instead, the sovereign repudiates the contract, there is a proportional output loss, γy, as in Sachs and Cohen (1982) and Bolton and Jeanne (2009). Creditors do not recover any payment if the sovereign repudiates; thus, the output loss represents a deadweight loss.[13] Note that the sovereign's repayment decision depends on the true budget, y, as well as on the contractual obligation, $O_1(\hat{y})$. The sovereign's payoff is summarised in the following table:

To conclude the section, consider the different entries in table 2.2.1. All entries but the lower-left (no audit, repayment), correspond to a sovereign default episode; and all default episodes are politically costly to the sovereign. Still, not all default episodes are equal. If there is a public audit and the sovereign subsequently repays, the outcome resembles a successful debt workout, or an excusable default as in Grossman and Van Huyck (1988). If the sovereign repudiates, there is an output loss of γy, but creditors do not recover any payment.[14]

the private sector; sometimes, the IMF even provides an explicit seal of approval for a proposed debt restructuring by sending *comfort letters* to creditors, cf. Díaz-Cassou, Erce, and Vázquez-Zamora (2008). While the IMF performs the role of an auditor, the IMF crucially cannot enforce payments or seize assets.

[13]The output loss is best thought of as arising from a loss of market access: as long as no settlement is reached with outside investors, the country is shut out off international markets (Bolton and Jeanne, 2009); the parameter $\gamma \leq 1$ captures the power of creditors to punish the sovereign for repudiation.

[14]Outright repudiation is rarely observed. An example is the refusal of Russia's Bolshevik government to repay Tsarist debts after the revolution in 1918.

	Repayment	Repudiation
Audit	$y - O_1(y)$	$y - \gamma y$
No Audit	$y - O_1(\hat{y}) + B_1$	$y - \gamma y$

Table 2.2.1: Sovereign's Payoff Matrix

2.3 Optimal contract

2.3.1 With repayment commitment

The optimal contract depends on whether the sovereign can commit to a future repayment strategy at the financing stage. As a benchmark, I derive the optimal contract assuming that the sovereign can commit, at date 0, to a future repayment strategy, at date 1. In the following, let y be the true state, while $\hat{y}$ denotes the announced state. Under commitment, the sovereign chooses *full-repayment*, i.e. a repayment strategy given by

$$r(y,\hat{y}) = \begin{cases} 1 & \text{if} \quad O_1(y) \leq y \quad \text{and} \quad I_d(\hat{y}) = 1, \\ 1 & \text{if} \quad O_1(\hat{y}) \leq y \quad \text{and} \quad I_d(\hat{y}) = 0, \\ 0 & \text{if} \quad \text{otherwise} \end{cases}$$

With full repayment, the sovereign pays any contractual obligation that respects the budget constraint. The remaining problem is to derive the optimal contract under a full-repayment commitment; a problem that is equivalent to a special case of Gale and Hellwig (1985), who consider a model of credit extension without any enforcement friction.

If there is an optimal contract, it takes the form of a *standard debt contract*. Three features define standard debt: (i) a fixed payment, or face value; (ii) an audit if and only if the sovereign's ability-to-pay falls short of debt's face value; and (iii) maximum recovery in case of an audit.[15] Formally, a contract (O_1, I_d) is said to be a *standard debt contract* if and only if

1. for some D, we have $O_1(y) = D$ if $I_d(y) = 0$;
2. $I_d(y) = 1$ if and only if $y < D$; and
3. $O_1(y) = y$ if $I_d(y) = 1$;

also see figure 2.3.1.

[15] As there is only a private cost of disclosure, maximum recovery implies that all income is transferred to creditors in case of disclosure. By contrast, the pecuniary costs of state observation in Gale and Hellwig (1985) imply that creditors recover only part of firm income in case of disclosure.

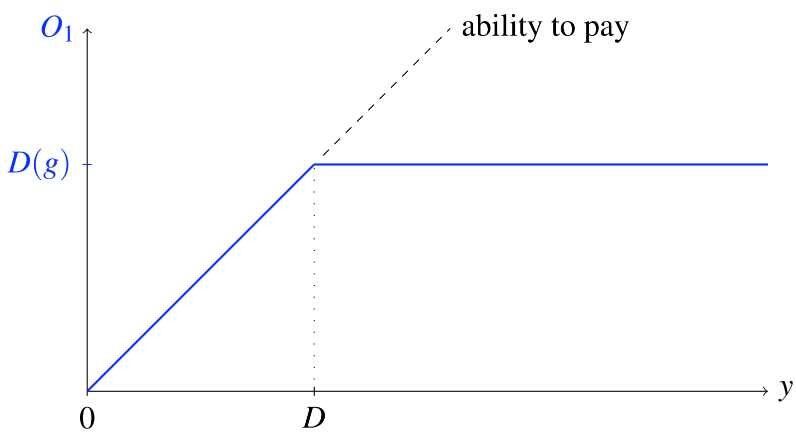

Figure 2.3.1: Payment of standard debt contract as a function of income.

2.3.2 Without repayment commitment

Without the ability to commit, the sovereign makes her repayment decision after the contractual obligation is set. To see which contractual obligations are repaid, and which are repudiated, consider the repayment stage at time 1. The optimal repayment strategy, $r(y,\hat{y})$, follows from comparing the sovereign's utility in case of repayment with her utility in case of repudiation; it is given by

$$r(y,\hat{y}) = \begin{cases} 1 & \text{if} \quad O_1(y) \leq \min\{\gamma y + b, y\} \quad \text{and} \quad I_d(\hat{y}) = 1, \\ 1 & \text{if} \quad O_1(\hat{y}) \leq \min\{\gamma y + B, y\} \quad \text{and} \quad I_d(\hat{y}) = 0, \\ 0 & \text{if} \qquad\qquad \text{otherwise} \end{cases} \tag{2.3.1}$$

The maximum contractual obligation that is repaid under the optimal repayment strategy, by definition, is the sovereign's *willingness-to-pay*; it depends both on the true state and the announced state, and is increasing in the costs of repudiation–the economic costγ, and the political cost B.

Repudiation leads to a deadweight loss, as creditors do not recover any payment if the sovereign repudiates. Contracts that are repaid almost surely at the repayment stage are called repudiation-proof. Formally, a contract, (O_1, I_d), is said to be *repudiation proof* if and only if:

$$\mathbb{P}\left(\{y \in T \mid \hat{y}(y) \text{ is such that } r(y, \hat{y}(y)) = 0\}\right) = 0$$

were $\hat{y}(y)$ is the sovereign's chosen announcement in state y under the given contract. Repudiation proofness is necessary for optimality, as the following proposition shows.

Proposition 2.3.1. *An optimal contract must be repudiation-proof.*

Proof. Consider an optimal contract, (O_1, I_d) and suppose it is not repudiation-proof. Then consider a new contract,$(\tilde{O}_1, \tilde{I_d})$, given by

$$\tilde{O}_1(y) = \begin{cases} \min\{\gamma y + b, y\} & \text{for } \{y \in T \mid \hat{y}(y) \text{ such that } I_d = 1 \text{ and } r = 0\} \\ \min\{\gamma y + B, y\} & \text{for } \{y \in T \mid \hat{y}(y) \text{ such that } I_d = 0 \text{ and } r = 0\} \\ O_1(y) & \text{for } \{y \in T \mid \hat{y}(y) \text{ such that } r = 1\} \end{cases}$$

and $\tilde{I_d} \equiv I_d$. Note that the new contract is repudiation-proof by construction, and leaves the sovereign with identical announcement incentives. It follows that, under the new contract, the sovereign receives the same payoff in all states, while creditors receive a higher expected payment, which contradicts the optimality of the initial contract. □

The proposition is intuitive: an optimal contract avoids the deadweight loss of repudiation by respecting the sovereign's willingness-to-pay constraints. Note that the standard debt contract is, in general, not repudiation-proof.[16]

At the announcement stage, the sovereign can lie about her income; she will if lying leads to a lower repayment. I check that the the contract, (O_1, I_d), is carried out as specified. If the announcement is such that the contract calls for an audit (i.e. $I_d(\hat{y}) = 1$), then creditors observe the true state, and the contractual obligation is set at $O_1(y)$; thus, I need only check that the sovereign has no incentive to falsely claim that her income is $\hat{y}$, with $I_d(\hat{y}) = 0$.

Let $W(y, \hat{y})$ denote the sovereign's date 1 payoff if her true income is y, while

[16]Under standard debt, the contractual obligation exceeds the sovereign's willingness-to-pay with positive probability, except for $\gamma = 1$.

she announces $\hat{y}$ for which $I_d(\hat{y}) = 0$, so

$$W(y,\hat{y}) := y - O_1(\hat{y}) + B$$

If the sovereign reveals the true state y, her date 1 utility is

$$y - O_1(y) + S(y,y)$$

A contract then is said to satisfy *truthful state revelation* if and only if: for any states y and $\hat{y}$ such that $I_d(\hat{y}) = 0$, we have $W(y,\hat{y}) \leq y - O_1(y) + S(y,y)$. Announcing a false state is unprofitable if a contract satisfies truthful state revelation; the structure imposed is given in the following lemma.

Proposition 2.3.2. *A contract (O_1, I_d) satisfies truthful state revelation if and only if there is a constant D such that (i) $O_1(y) = D$, whenever $I_d(y) = 0$; and (ii) for any y and $\hat{y}$ such that $I_d(\hat{y}) = 0$, $I_d(y) = 1$, we have $O_1(y) \leq O_1(\hat{y}) - (B - b)$.*

Proof. If (i) and (ii) hold, the sovereign cannot do better than truthfully reveal her income at the announcement stage. To see that (i) and (ii) are necessary, suppose $O_1(y)$ is not constant for unaudited states; then, the sovereign has an incentive to announce the unaudited state that results in the lowest repayment, contradicting truthful state revelation. Likewise, suppose condition (ii) is violated; then, there exists y and $\hat{y}$ with $I_d(\hat{y}) = 0$ and $I_d(y) = 1$, such that $O_1(y) > O_1(\hat{y}) - (B - b)$. This implies that the sovereign strictly prefers to announces $\hat{y}$ instead of y, again a contradiction. □

Condition (i) of the lemma ensures that the sovereign has no gain from announcing a different state if the actual realisation remains unaudited. Condition (ii) of the lemma ensures that the sovereign has no gain from announcing an unaudited state if the actual realisation calls for an audit. She may be tempted to do so to avert the loss of private benefit that is associated with an audit, i.e., to avert the loss of $(B - b)$.

An optimal contract, (O_1, I_d), solves

$$\max_{(O_1, I_d)} E\left(y + S(y,\hat{y})\right) - g$$

such that

$$EO_1(y) \geq g \tag{2.3.2}$$

and

$$\begin{aligned} O_1(y) &\leq \min\{\gamma y + B, y\} \quad \text{for } y \in \{x | I_d(x) = 0\} & (2.3.3)\\ O_1(y) &\leq \min\{\gamma y + b, y\} \quad \text{for } y \in \{x | I_d(x) = 1\} & (2.3.4) \end{aligned}$$

and there is a constant D such that

$$\begin{aligned} O_1(y) &= D \quad \text{for} y \in \{x | I_d(x) = 0\} & (2.3.5)\\ O_1(y) &\leq D - (B - b) \quad \text{for } y \in \{x | I_d(x) = 1\} & (2.3.6) \end{aligned}$$

An optimal contract maximises the sovereign's expected utility subject to the investor participation constraint and four incentive constraints: two repudiation proofness constraints, and two truthful revelation constraints. It is easy to show that the participation constraint of the investor must bind at an optimum, or $EO_1 = g$.[17] The maximisation problem reveals that the sovereign wishes to finance the government expenditure, while maximising the private benefit of holding office. To characterise the solution, I introduce a new type of contract: the sovereign debt contract. A contract is said to be a *sovereign debt contract* if and only if

(i) for some D, we have $O_1(y) = D$ if $I_d(y) = 0$;
(ii) $I_d(y) = 1$ if and only if $D > \min\{\gamma y + B, y\}$; and
(iii) $O_1(y) = \gamma y + b$ if $I_d(y) = 1$.

The sovereign debt contract specifies: (i) a fixed payment, or face value; (ii) disclosure if and only if the willingness-to-pay falls short of the face value; and (iii) a payment equal to the remaining creditor punishment threat in case of disclosure, also see figure 2.3.2.[18] Sovereign debt contracts are repudiation-proof, they satisfy truthful state revelation, and they are uniquely characterised by their face value. The following proposition shows that an optimal contract must be a

[17] If the participation constraint does not bind, then $P_1(y)$ can be decreased such that the participation constraint of the investor, and truthful state revelation, remain satisfied. The resulting increase in expected utility for the sovereign, contradicts optimality.

[18] In the graph shown, the political cost of auditing equals the political cost of repudiation, its upper bound.

sovereign debt contract.

Proposition 2.3.3. *Let* (O_1, I_d) *be an optimal contract, then* (O_1, I_d) is a sovereign debt contract.

Proof. See the Appendix. □

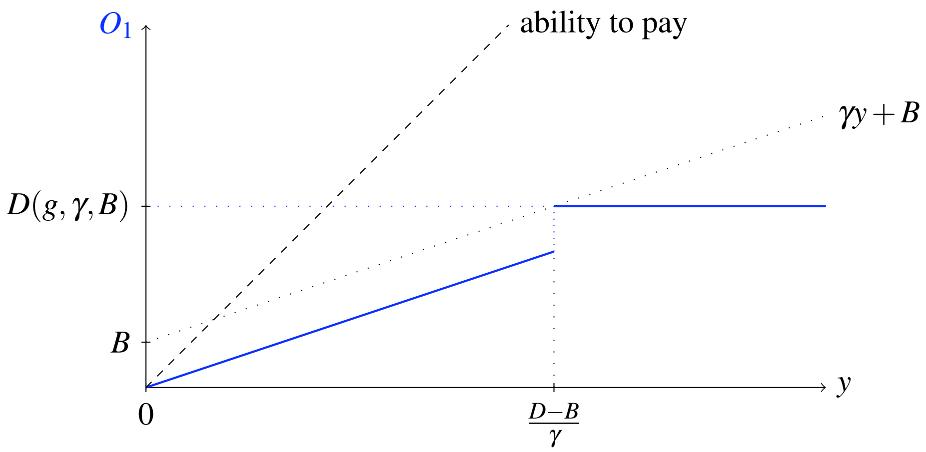

Figure 2.3.2: Payment of sovereign debt contract as a function of income

Intuitively, the sovereign debt contract is optimal because it (i) economises on the costs of auditing, and (ii) is never repudiated. While costly audits serve to establish the sovereign's ability-to-pay, repudiation leads to a pure deadweight loss. The sovereign debt contract is repaid at face value in high income states, where the sovereign's willingness-to-pay is also high; in low income states, where the sovereign's willingness-to-pay is also low, the sovereign debt contract specifies a public audit and an output-contingent payment.

Compared to standard debt, the fixed repayment feature is retained in the sovereign debt contract. The audit decision differs; in particular, there is an audit whenever the willingness-to-pay of the sovereign falls below a threshold. Finally, the sovereign debt contract does not specify maximum recovery for states that are audited. Rather, the amount that is recovered in audit states equals the punishment that creditors can inflict. For the boundary case, $\gamma = 1$, the sovereign debt contract coincides with standard debt. The intuition is that the willingness-to-pay problem poses no constraint if repudiation leads to a loss of the full budget. For the more plausible cases, with $\gamma < 1$, the payment of the sovereign debt contract is discontinuous at the audit threshold, i.e., at $y = \frac{D-B}{\gamma}$. The payment discontinuity ensures that the sovereign reveals her budget truthfully in all states.

The primitives of the contracting problem are (i) the private benefits of holding office, B and b (ii) the proportion of output that is lost if the sovereign repudiates, γ (iii) the government expenditure need g, and (iv) the distribution and support of income y, i.e. $F(y)$ and $T \subseteq \mathbb{R}_+$. To gain intuition for the existence problem, consider the contracting problem under symmetric information. With symmetric information, the only remaining friction is the willingness-to-pay problem and the scope for inefficiency is extreme: either the expenditure can be financed and the first-best is achieved, or there is no contract with which the expenditure can be financed. To see this, note that with symmetric information there is no need for costly auditing, and the sovereign can pledge a maximum of $\gamma y + B$ in each state, as long as the budget constraint is satisfied. Expected pledgeable income therefore equals

$$E\left(\min\{\gamma y + B, y\}\right) \tag{2.3.7}$$

If pledgeable income exceeds the expenditure (g), then the first-best can be achieved and the optimal contract is indeterminate; if the expenditure exceeds pledgeable income, then no contract allows the sovereign to finance the expenditure. The example shows that the primitives of the problem can be such that the sovereign is not able to finance the expenditure with any contract. In particular this is the case if creditors have little power to punish repudiation ($\gamma \ll 1$), or if the government expenditure is high ($g \gg 0$).

2.4 Sovereign Debt Contract

An individual investor who holds a sovereign debt contract, (O_1, I_d), expects a repayment of

$$EO_1 = \gamma \int_0^{D-B/\gamma} y f(y) dy + D \int_{D-B/\gamma}^{\infty} f(y) dy \tag{2.4.1}$$

Simple comparative statics show that

Proposition 2.4.1. *The expected repayment of a sovereign debt contract with a given face value D, is increasing in creditor power, γ, and in the private benefit of holding office, B.*

The proposition is intuitive. An increase in the economic cost of repudiation, γ, increases the sovereign's willingness-to-pay in all states. An increase in political cost of repudiation, B, also increases the sovereign's willingness-to-pay, but only in unaudited states.[19]

Consider the primary market for sovereign debt, i.e. the market at the date of issuance. The main question at date 0 is whether the sovereign can finance g. She can't if creditors have too little power or if the government expenditure is too high (cf. section 2.3.2). Proposition 2.4.1 then implies that an increase in B may lift the sovereign out of autarky; likewise, an increase in γ can leave the sovereign debtor better off. . For the primary market, the model predicts that the sovereign should find it easier to raise funds from powerful creditors, meaning that the sovereign pays a lower interest rate on a loan of given size. After the date of issuance, sovereign debt contracts trade in a secondary market, where their *market value* equals expected creditor repayment.[20]

Different events may move the secondary market price, or implied interest rate, of the sovereign debt contracts. Suppose, for example, that the government announces an audit at time 1, cf. timing of events in section 3.2. Then creditors will observe the state of the economy as information on the economy comes available. This may not be immediate. By contrast any market response will be immediate. As soon as the government announces the audit, the market value drops to

$$E(O_1|I_d = 1) = \gamma \int_0^{D-B/\gamma} yf(y)dy$$

[19]In disclosed states, the political audit cost is sunk and only the threat of creditor punishment deters the sovereign from repudiation.

[20]Formally, there is no secondary market in my model, as investors are homogeneous. Trade can easily be introduced, however, by assuming exogenous liquidity shocks. That the market value equals expected repayment follows from investor risk neutrality. It follows from investors' participation constraint that the market value at date 0 equals g.

where $f(y)$ is the probability density function. When creditors learn about y, the market value converges to γy.[21]

Another event of interest is a change of government before the repayment stage. Within the framework of the model, there are two channels through which a change in government can effect the market value of outstanding debt:

1. a change in the private benefit B (willingness-to-pay channel) ; or

2. a change in the distribution, F, and support ,T, of ability-to-pay y (competence channel) .

Suppose a new government takes office that is understood to be more competent than its predecessor as to collecting taxes, undertaking structural reforms, and privatising state property. Then the probability of a high income state increases and so does the market value of outstanding debt.[22] The opposite happens if a new government takes charge that is perceived to be less competent than its predecessor. Likewise, suppose a government takes over that is known to be highly committed to avoid a public audit (high B), then the market will view this favourably and the market value of outstanding debt increases. A new government that is perceived as less committed to pay the debt at face value (low B) leads to a decline in the market value of outstanding debt.

Finally, one may consider the impact of changes in γ and B in the secondary market. If changes take place before date 1, then the effect is given by proposition 2.4.1. Hence the market value of the sovereign debt contract increases with an increase in either γ or B.[23]

2.5 Alternative Repayment Game

I consider an alternative repayment game, in which a debt crisis is followed by a debt renegotiation, as in Bolton and Jeanne (2009). In the original formulation,

[21] Learning about y can, for example, be modelled as a narrowing of the support of y.

[22] Any market response must run through expectations of investors, as there is not yet a realisation of y.

[23] Note that both an increase in γ and an increase in B destroy truthful state revelation of the contract that was initially issued: the sovereign pays D even in states where the contract calls for an audit. Indeed, this is the reason why the market value increases.

there is no need to renegotiate the initial contract, as the contract specifies the course of action in each contingency. The alternative view, explored here, is that the contract cannot be explicitly conditioned on the state of the world, even if creditors observe that state. The repayment game, at time $t = 1$, is as follows:

1. Nature chooses the state y, sovereign observes y ;
2. **Announcement:** Sovereign announces her ability-to-pay $\hat{y}$,
   (a) if $I_d(\hat{y}) = 1$, investor observes y, and a debt renegotiation starts;
   (b) if $I_d(\hat{y}) = 0$, then contractual obligation is $O_1(\hat{y})$.
3. In case of a debt renegotiation, coordinated creditors make a repayment offer η. Otherwise, the contract binds both parties to $O_1(\hat{y})$.
4. **Payment:** Sovereign makes a repayment decision $r \in \{0, 1\}$,
   (a) she pays and the game ends ($r = 1$), or
   (b) she repudiates and creditors charge a punishment ($r = 0$).

If the sovereign announces a state that is not audited, then the contractual obligation, $O_1(\hat{y})$, is binding for all parties. If instead there is an audit, then creditors make a repayment offer η for which they are willing to swap the initial contract and lift repudiation sanctions; but, creditors can only make the offer if they manage to coordinate. Formally, there is a coordination cost c_R, incurred by creditors if they make an offer η. As the renegotiation surplus equals γy, creditors cannot be coordinated if the income realisation is too low, or $\gamma y < c_R$. In such states, no renegotiation takes place, creditors receive 0, and the sovereign suffers the economic and political cost of repudiation. If creditors can coordinate, i.e., $\gamma y \geq c_R$, then the creditor offer follows from solving the repayment game backwards along the public audit branch. Because the sovereign accepts any offer $\eta \leq \gamma y$, creditors set their offer at $\eta = \gamma y$, and receive a net payment of $\gamma y - c_R$.

I assume that creditors can either coordinate at no cost ($c_R = 0$); or creditors cannot coordinate at all ($c_R = \infty$). These assumptions are made to capture, in a stylised way, the difference between debt that is held by a handful of banks that find it easy to coordinate; and debt that is held by dispersed bondholders that cannot be coordinated, also see Bolton and Jeanne (2009).

If creditors can coordinate at no cost, then introducing the debt renegotiation is equivalent to setting the contractual obligation in audited states equal to $O_1(y) = \gamma y$ in the original specification of the model; thus, the contracting problem becomes a special case of section 2.3.2: the set of admissible contracts is restricted.[24] By proposition 2.3.3, the optimal contract is a sovereign debt contract; furthermore, a sovereign debt contract is admissible as it satisfies $O_1(y) = \gamma y$ for audited states. It follows that proposition 2.3.3 applies, so that

Proposition 2.5.1. *For $c_R = 0$ and $\gamma = 1$, the optimal contract is standard debt contract; for $c_R = 0$ and $\gamma < 1$ the optimal contract is a sovereign debt contract.*

Proof. Let $c_R = 0$ and $\gamma = 1$ and assume that an optimal contract exists. Then, by proposition 2.3.3, the optimal contract is a sovereign debt contract. As $\gamma = 1$, the sovereign debt contract coincides with the standard debt contract; If $c_R = 0$ and $\gamma < 1$, and there exists an optimal contract, then the optimal contract is a sovereign debt contract by proposition 2.3.3. □

Proposition 2.5.1 shows that the optimal contract is a sovereign debt contract. The conditions for the existence of an optimal contract are the same as in section 2.3.2, i.e. g cannot be too big, and γ cannot be too small. Proposition 2.5.1 also shows that renegotiable debt, as in Bolton and Jeanne (2009), is optimal if creditor coordination is costless ($c_R = 0$), and if creditor punishment is maximal ($\gamma = 1$).

If creditors cannot coordinate at all, then no renegotiation can take place and an audit leads to the same payoff as repudiation: creditors receive 0; the sovereign incurs a loss of γy, and loses her private benefit B. In an optimal contract, the payment to creditors in undisclosed states must compensate for the zero payment to creditors in all other states.[25] Furthermore, the contract must specify a constant contractual obligation across states. Any other contract leaves the sovereign with

[24] Stage 3 can be collapsed into Stage 2a of the original repayment game, by setting $O_1(y) = \gamma y$ for audited states (cf. section 3.2).

[25] Note that this implies that the conditions for existence of an optimal contract are more stringent than before.

an incentive to announce the state with the lowest contractual obligation and cannot satisfy truthful revelation. Let D denote the constant contractual obligation, or face value, of the optimal contract. Then expected payment is given by

$$EO_1 = D \int\limits_{\{y|D<\min\{\gamma y+B,y\}\}} f(y)dy$$

The sovereign only pays the debt at face value in states where her willingness-to-pay exceeds the debt's face value D. If creditor punishment is maximal, or $\gamma = 1$, then the willingness-to-pay of the sovereign equals her ability-to-pay. It follows that expected repayment is given by

$$EO_1 = D \int\limits_{D}^{\infty} f(y)dy$$

and the optimal contract corresponds to non-renegotiable debt as in Bolton and Jeanne (2009). If creditor punishment is less than maximal, or $\gamma < 1$, then the willingness-to-pay of the sovereign is smaller than her ability-to-pay and repayment is a political decision; expected repayment is

$$EO_1 = D \int\limits_{\frac{D-B}{\gamma}}^{\infty} f(y)dy$$

which corresponds to non-renegotiable debt à la Bolton and Jeanne (2009), but with a repudiation threshold that depends on creditor power, and on the private benefit of holding office. The following proposition summarises the discussion above.

Proposition 2.5.2. *For $c_R = \infty$ and $\gamma = 1$, the optimal contract corresponds to non-renegotiable debt; for $c_R = \infty$ and $\gamma < 1$ the optimal contract is a non-renegotiable debt contract with a repudiation threshold of $\frac{D-B}{\gamma}$.*

Proof. Omitted □

2.6 Conclusion

In this chapter, I analyse the problem of credit extension to a a sovereign borrower given that (i) there is no court, (ii) the sovereign knows better than creditors what her repayment capacity is, and (iii) disclosure of that information is politically costly. Recent events in Greece show the relevance of these issues: creditors did not have accurate information on the state of government finances and sovereign debt contracts proved difficult to enforce. In this setting, positive repayment is sustained by an economic and political penalty associated with repudiation; and by a political penalty associated with disclosure. These three penalties drive the optimal contract design.

I show that the sovereign borrower optimally issues a contract for which (i) the repayment profile is flat in high income states; (ii) there's a state contingent payment that depends on creditor power in low income states; and (iii) there is an audit (or disclosure) if the willingness-to-pay falls short of the *face value* of the contract.

The intuition for the optimal contract is that it saves on costly auditing, which is what the corporate finance literature has emphasised. The optimal contract itself, however, is different from what the corporate finance literature has found: it is still a debt contract, but the default decision and the repayment in case of default differ from standard debt due to the willingness-to-pay problem.

The optimal contract I derive explains some of the salient facts of sovereign borrowing. First, a sovereign's ability-to-pay is not the relevant constraint when it comes to repayment. It is the willingness-to-pay that determines repayment, and the willingness-to-pay depends jointly on the budget, creditor power, and the private benefit of holding office. Second, the payment to creditors depends on their power to punish repudiation. The most powerful creditor is the International Monetary Fund (IMF); historically, the IMF takes priority over all other creditors. The upshot is that, even if rates on IMF loans are lower than on other loans, IMF lending is not concessionary. Third, the sovereign chooses to issue plain bonds. A priori, this is puzzling, as Shleifer (2003) and others have argued. Why don't sovereign borrowers issue contracts that condition on future income? Such contracts are not optimal, because the auditing requirements would be prohibitively costly.

The current work can be extended and complemented in several directions. Empirically, the cross-sectional implications of the model should be taken to the data. In particular, do shifts in political power or ultimate holdings of sovereign debt lead to the secondary price responses that the model predicts? The theory can be extended to develop a fully dynamic model of sovereign debt; a model that endogenizes the cost of repudiation and allows the study of repayment and refinancing decisions in one framework.

References

ATKESON, A. (1991): "International Lending with Moral Hazard and Risk of Repudiation," *Econometrica*, 59(4), 1069–1089.

——— (2005): *And the Money Kept Rolling In (and Out)*. Public Affairs, New York.

BOLTON, P., AND O. JEANNE (2007): "Structuring and Restructuring Sovereign Debt: The Role of a Bankruptcy Regime," *Journal of Political Economy*, 115(6), 901–924.

——— (2009): "Structuring and Restructuring Sovereign Debt: The Role of a Seniority," *Review of Economic Studies*, 76, 879–902.

BORDER, K. C., AND J. SOBEL (1987): "Samurai Accountant: A Theory of Auditing and Plunder," *Review of Economic Studies*, 54, 525–540.

BORENSZTEIN, E., AND P. MAURO (2004): "The case for GDP-indexed bonds," *Economic Policy*, 19(38), 165–216.

BORENSZTEIN, E., AND U. PANIZZA (2009): "The Costs of Sovereign Default," IMF Staff Papers 56, International Monetary Fund.

——— (1989a): "A Constant Recontracting Model of Sovereign Debt," *Journal of Political Economy*, 97(1), 155–78.

——— (1989b): "Sovereign Debt: Is to Forgive to Forget?," *American Economic Review*, 79(1), 43–50.

COLE, H. L., AND P. J. KEHOE (1998): "Models of Sovereign Debt: Partial versus General Reputations.," *International Economic Review*, 39(1), 55 – 70.

DIAMOND, D. W. (1984): "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393–414.

DÍAZ-CASSOU, J., A. ERCE, AND J. J. VÁZQUEZ-ZAMORA (2008): "The Role of the IMF in Recent Sovereign Debt Restructurings: Implications for the Policy of Lending into Arrears.," Occasional Paper 805, Bank of Spain.

EATON, J., AND R. FERNANDEZ (1995): "Sovereign Debt," NBER Working Papers 5131, National Bureau of Economic Research.

EATON, J., AND M. GERSOVITZ (1981): "Debt with Potential Repudiation: Theoretical and Empirical Analysis," *Review of Economic Studies*, 48(2), 289–309.

FERNANDEZ, R., AND R. W. ROSENTHAL (1990): "Strategic Models of Sovereign-Debt Renegotiations," *Review of Economic Studies*, 57(3), 331–349.

GALE, D., AND M. HELLWIG (1985): "Incentive-Compatible Debt Contracts: The One-Period Problem," *Review of Economic Studies*, 52, 647–663.

——— (1989): "Repudiation and Renegotiation: The Case of Sovereign Debt," *International Economic Review*, 30(1), 3–25.

GROSSMAN, H. I., AND J. B. VAN HUYCK (1988): "Sovereign Debt as a Contingent Claim: Excusable Default, Repudiation, and Reputation," *American Economic Review*, 78(5), 1088–97.

GUEMBEL, A., AND O. SUSSMAN (2009): "Sovereign Debt without Default Penalties," *Review of Economic Studies*, 76(4), 1297–1320.

HART, O., AND J. MOORE (1998): "Default and Renegotiation: A Dynamic Model of Debt," *the Quarterly Journal of Economics*, 113, 1–41.

HELLWIG, C., AND G. LORENZONI (2009): "Bubbles and Self-Enforcing Debt," *Econometrica*, 77(4), 1137–1164.

HVIDE, H. K., AND T. E. LEITE (2010): "Optimal Debt Contracts under Costly Enforcement," *Economic Theory*, 44(1), 149–165.

INNES, R. D. (1990): "Limited Liability and Incentive Contracting with Ex-ante Action Choices," *Journal of Economic Theory*, 52, 45–67.

KLETZER, K. M., AND B. D. WRIGHT (2000): "Sovereign Debt as Intertemporal Barter," *American Economic Review*, 90(3), 621–639.

KRASA, S., AND A. P. VILLAMIL (1994): "Optimal Multilateral Contracts," *Economic Theory*, 4(2), 167–187.

——— (2000): "Optimal Contratcs when Enforcement is a Decision Variable," *Econometrica*, 68, 119–134.

MOOKHERJEE, D., AND I. PNG (1989): "Optimal Auditing, Insurance and Redistribution," *Quarterly Journal of Economics*, 104, 399–415.

REINHART, C. M., AND K. S. ROGOFF (2009): *This Time is Different: Eight Centuries of Financial Folly*. Princeton University Press, Princeton and Oxford.

SACHS, J., AND D. COHEN (1982): "LDC Borrowing with Default Risk," Working Paper 925, NBER.

SANDLERIS, G. (2008): "Sovereign Defaults: Information, Investment and Credit," *Journal of International Economics*, 76(2), 267–275.

SHLEIFER, A. (2003): "Will the Sovereign Debt Market Survive?," *American Economic Review Papers and Proceedings*, 93(2), 85–90.

STURZENEGGER, F., AND J. ZETTELMEYER (2006): *Debt Defaults and Lessons from a Decade of Crises*. MIT Press, Cambridge.

TOWNSEND, R. M. (1979): "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, 21, 265–293.

WORRALL, T. (1990): "Debt with potential repudiation," *European Economic Review*, 34(5), 1099–1109.

Chapter 3

Collective Pension Funds[1]

Credit Constraints and a Conflict over Risk and Contributions

Abstract. This chapter explores one rationale for pension funds. If individuals face credit constraints, i.e., if markets are incomplete, then a pension fund is able to improve on welfare by implementing participants' preferred investment strategy. A pension fund can do this if (i) participation is mandatory and (ii) it has access to a tax on human wealth. We show that implementation of the optimal allocation can be achieved through a defined-contribution (DC) pension scheme. We argue that, after a low stock return, such schemes can run into the same type of problems as underfunded defined-benefit (DB) schemes.

[1]This chapter is based on joint work with David Hollanders. For helpful comments, we thank Lans Bovenberg, Enrico Perotti, Bas Jacobs, and various seminar audiences.

3.1 Introduction

A pension is a payment stream that people receive upon retirement, i.e., when they leave the labor force. Rather than leaving it to individuals to save for their retirement, most advanced economies have pension systems in which individuals are required to participate. Common to such systems is that people contribute in their active working years, which entitles them to a pension benefit upon retirement.[2] But there is considerable variation between countries in how the pension system operates, how it is financed, and how pension benefits are determined.[3]

For example, some countries, like Germany, operate pay-as-you-go (PAYG) pension systems, where the active working population pays for the current retirees. Other countries, like the Netherlands, operate additional prefunded schemes, in which people *save* through pension contributions and receive a pension benefit that is set according to the pension contract. Again, pension contracts differ between countries. Roughly, one can distinguish between *defined-contribution* (DC) type contracts, where the pension benefit depends explicitly on investment returns (e.g. the famous 401(k) plans in the U.S.), and *defined-benefit* (DB) type contracts, where the pension is set according to a formula that may depend on average pay, years of employment, age at retirement, and other factors (e.g. the second pillar in the Netherlands).

Prefunded DB pension schemes—as found in the U.S., the Netherlands, and elsewhere—run into trouble when they are *underfunded*, i.e., when pension liabilities, which are fixed, exceed pension assets, which may fluctuate; the difference is called the *pension shortfall*. Pension shortfalls are an inherent risk—and recurring feature—of funded DB pension schemes.[4] When there is a pension shortfall, there are two ways to restore the pension system's solvency: pension entitlements can be cut; or, contributions can be raised.[5] While cutting entitlements hurts

[2]This seems an obvious requirement, but note that the first generation in a pay-as-you-go (PAYG) pension system receives a pension without having paid contributions.

[3]World Bank (1994) gives a useful categorization of pension systems into three pillars: a *state pension*, aimed at poverty reduction, and financed through taxes; an *occupational pension*, aimed to maintain the standard of living, and prefunded; and a *private pension*, allowing for individual supplements, also prefunded.

[4]For example, General Motor's defined-benefit pension plans reported a shortfall of $35 billion in 2011; this exceeded GM's market value.

[5]By cutting entitlements, pension funds decrease their liabilities; by raising contribu-

older generations the most—as they have accumulated most entitlements—raising contributions hurts younger generations more and shields retirees from losses. Clearly, ex-post, generations do not agree over who should pay for the shortfall; this must be agreed upon, ex ante, in the pension contract.[6] By default, and per definition, DB schemes put the risk of a pension shortfall on the working generation, implying that contributions must be raised in case of a shortfall. In practice, this may turn out to be infeasible for political or regulatory reasons: employers and workers may effectively resist an increase in contributions—which is a tax an labor—especially during a recession; if regulations require a quick return to solvency, as they often do, then pension funds may have no option than to cut entitlements.

To illustrate these issues, consider the Dutch case. After the 2008 credit crisis, more than half of the pension funds in the Netherlands were designated 'underfunded,' by the Dutch Central Bank. The decline in pension wealth led to controversy over who should pay to restore the solvency of the pension system—Dutch regulations required a return to solvency within 5 years. **?** calculates that the proposed policy mix of entitlement cuts and raised contributions hurts older generations the most.

In recognition of the intrinsic tensions in DB pension systems with mismatch risk, DB pension schemes are being replaced with DC pension schemes, cf. Goudzwaard et al. (2009). In this chapter, my aim is to explore one rationale for a DC pension scheme that has, so far, received little attention in the literature.[7] The rationale we explore is that pension funds exist to lift credit constraints and implement the optimal optimal life-cycle investment strategy of participating generations.

An important finding of the literature on modern life-cycle investment theory is that, over their life-cycle, individuals should hold a constant fraction of their total wealth in risky assets, with the remainder invested in a risk-free asset, cf. Bodie, Merton, and Samuelson (1992). The upshot is that individuals' optimal

tions—while keeping entitlements fixed—pension funds increase their assets.

[6]As **?** argues, the controversy over who should pay cannot be resolved *ex post*; it requires a model of optimal *ex ante* risk sharing.

[7]The notable exception is Bovenberg et al. (2007) ; Teulings and de Vries (2006) mention but do not pursue this rationale for pension funds.

investment strategy depends on their age: the young—who have human capital as well as financial capital—should invest their financial capital in a riskier manner than retirees—who have only financial capital. With their human capital, the young are naturally hedged against stock market risk.[8] Typically, these models require the young to take a leveraged position in the stock market, i.e., to borrow and invest the proceeds in the stock market.[9] If the young face credit constraints, this strategy is infeasible.

It is plausible that the young face credit constraints in private markets, as they do not have any collateral to offer to lenders.[10] If the young are credit constrained, then pension funds have a role to play: pension funds can implement the young's preferred investment strategy by extending them credit; thus, effectively alleviating the young's credit constraints.[11] There are two reasons why pension funds are better placed than private sector lenders to extend credit to the young: (i) participation is mandatory, reducing the adverse selection problem; and (ii) pension funds have access to a tax on human wealth, which allows them to enforce repayment. In effect, the pension fund helps to secure the human capital of participants as collateral, cf. Bovenberg et al. (2007).

Implementation of optimal investment strategies can be achieved by a DC pension scheme, where participants pay a fixed amount of contributions (e.g. yearly), and pension funds invest these contributions on behalf of participants. To implement the optimal strategy of its youngest participants, the pension fund needs to borrow on their behalf. A pension fund with only young participants would, thus, have to take a leveraged position in the stock market, i.e., take a short position in the risk-free asset. By contrast, pension funds with young and older participants would still take a net long position in the risk-free asset, as older generations

[8]This finding relies on the assumption that human capital is risk-free. If capital returns and wages are cointegrated, as **?** propose, then the young already hold a risky asset through their human capital.

[9]For example, Teulings and de Vries (2006) show that, at the beginning of one's career, a generation should borrow five times its yearly wage to invest in equity—similar results can be found in Bodie, Merton, and Samuelson (1992).

[10]The young wish to borrow against their human capital to invest in financial capital, but as **?** note: 'human capital alone does not collateralize major loans for reasons of moral hazard and adverse selection.'

[11]Bovenberg et al. (2007) provide a review of different rationales for pension funds; the review discusses alleviation of borrowing constraints as one possible rationale.

prefer strategies that are less risky.[12] Implementation of the optimal investment strategy can then be interpreted as the young borrowing from other participants at the risk-free rate, pledging their human capital as collateral.

To demonstrate how pension funds can improve welfare, we present a simple and stylized model with three overlapping generations of risk-averse individuals: the young, the middle-aged, and the old. The capital market is equally stylized, it consists of a risk-free bond and a risky stock; the equity premium is positive. To focus on the optimal investment decision, we fix the amount of savings in each period, i.e., we abstract from the labor supply and optimal savings decision, as in Gollier (2008). Savings then take the form of a per-period endowment, and the problem is simply to determine the optimal investment strategy of the young and middle-aged. The old consume their pension benefit, which is endogenous and depends on past stock returns. After deriving the complete markets benchmark, we introduce credit constraints, a form of market incompleteness. Market incompleteness, in turn, is the main rationale for pension funds.[13]

3.2 Model

The economy is populated by three overlapping generations, each of unit mass, called the *young* (y), the *middle-aged* (m), and the *old* (o). In their active working years, generations contribute to a pension fund: the young and middle-aged pay a fixed contribution, s, to a pension fund in each period. In retirement, generations consume a pension benefit: the old obtain a pension benefit, b, from the pension fund. The welfare of a generation is measured by its utility in retirement, and we assume that generations have a CRRA utility function,

$$u(b) = \frac{b^{1-\phi}}{1-\phi}$$

[12]One interpretation is the following: the pension fund operates an internal capital market where the young issue risk-free debt to the old and use the proceeds to invest in equity.

[13]Market incompleteness due to credit constraints is a key assumption of our analysis; without credit constraints, there is no rationale to have a pension fund in our model. Credit constraints are also assumed in related work by **?** and Gollier (2008).

where $\phi > 0$ is the coefficient of relative risk aversion.[14] After retirement, generations leave the model.

There is a simple capital market consisting of two financial assets: a risk-free asset, called *bond*, and a risky asset, called *stock*. The gross return of the bond is fixed over time and normalized to 1. The excess stock return is stochastic and serially uncorrelated over time. We denote the excess stock return in period t by $\tilde{r}_t$—the mnemonic is that random variables are denoted with a tilde, realized values without. For simplicity, we assume the excess return follows a Bernoulli distribution: with probability p, the excess return takes the value $r^h > 0$; with probability $1-p$, it takes the value $r^l < 0$. The equity premium is positive,

$$\mu := pr^h + (1-p)r^l > 0 \tag{3.2.1}$$

Within each period, the timing within each period is as follows: (i) individuals enter the period with their *financial reserve*, i.e., financial assets carried from the last period; (ii) an investment decision is made; (iii) the young and middle-aged pay their contributions, and the pension fund pays benefits to the old; (iv) returns materialize.

3.2.1 Complete Markets

As a benchmark, we examine each generation's optimal investment strategy if financial markets are complete in the sense that there are no credit constraints.[15] At the beginning of each period, the young and middle-aged decide how much to invest in the stock, and how much to invest in the risk-free bond. The old have no more decision to make, they simply consume their accumulated retirement wealth. Let α_t^i denote the monetary investment in the stock by generation $i = y, m$ in period t.

We define a generation's *retirement wealth* at the beginning of period t as

[14] As usual the utility function is given by $u(b) = \ln(b)$ for $\phi = 1$. Following Gollier (2008), we focus exclusively on individuals' optimal investment decision, and we abstract from their labor supply and savings decision. Hence contributions, s, are exogenously fixed, and individuals maximize their expected retirement benefit, b, which is endogenously determined.

[15] Financial markets are still incomplete in the sense that non-overlapping generations cannot share risks with each other, cf. **?**, **?**, and **?**.

the sum of (i) its *financial reserve*, i.e., the balance of its financial assets, which we denote by w_t^i; and (ii) its *human capital reserve*, i.e., the residual net present value of future contributions to the pension fund, which we denote by h_t^i for $i = y, m, o$. The old generation's retirement wealth consists only of their financial reserve, w^o; they have exhausted their human capital reserve. By contrast, the young generation's retirement wealth consists only of their human capital reserve, h^y; they have no financial reserve. The retirement wealth of the middle aged generation consists of both a financial reserve and a human capital reserve, $w^m + h^m$.

The dynamic investment problem can now be written as:

$$\max_{\alpha_t^y, \alpha_{t+1}^m} Eu(\tilde{b})$$

such that

$$\begin{aligned} w_t^y &= 0 \\ w_{t+1}^m &= \alpha_t^y (1+\tilde{r}_t) + (y - \alpha_t^y) \\ w_{t+2}^o &= \alpha_{t+1}^m (1+\tilde{r}_{t+1}) + (w_{t+1}^m + y - \alpha_{t+1}^m) \\ \tilde{b} &= w_{t+2}^o \end{aligned}$$

The solution to the investment problem is given by the following proposition.

Proposition 3.2.1. *The optimal dynamic investment strategy of each generation is to invest a constant fraction of retirement wealth in the risky asset, or* $\alpha_t^i = a^* \left(w_t^i + h_t^i \right)$ *for* $i = y, m$ *; where the optimal fraction is given by*

$$a^* = \frac{\left(\frac{-r^l (1-p)}{r^h p} \right)^{\frac{1}{\phi}} - 1}{r^l - r^h \left(\frac{-r^l (1-p)}{r^h p} \right)^{\frac{1}{\phi}}}$$

The corresponding expected lifetime utility for a young individual is

$$\begin{aligned} U^{cm} &= Eu(\tilde{b}) \\ &= \delta^2 u(h_1) \\ &= \delta^2 u(2s) \end{aligned}$$

with

$$\delta := E\left(1 + a^* \tilde{r}\right)^{1-\phi} \tag{3.2.2}$$

Proof. See the Appendix. □

The proposition gives the optimal investment decision for the young, the middle-aged, and the old; the proposition shows that, over their life-cycle, individuals optimally hold a constant fraction of their retirement wealth in the risky asset. This result is well-known from the literature on optimal life-cycle investment, cf. **?**, **?**, and Bodie, Merton, and Samuelson (1992). Simple comparative statics show that optimal investment in the risky asset, a^*, is increasing in the equity premium, μ, and decreasing in the coefficient of relative risk aversion, ϕ.

Note that δ can be interpreted as the gross per-period *utility return* of investing optimally in one period. A young individual's retirement wealth consists entirely of his human capital reserve, $h_1 = 2s$. In utility terms, each period of optimal investment yields a gross return of δ; hence $U^{cm} = \delta^2 u(2s)$. Using a first-order approximation we note that

$$\delta \approx 1 + (1-\phi) a^* \mu$$

In financial terms, the corresponding *certainty equivalent rate of return*, r^{ceq}, is then given by

$$u\left(h_1 \left(1 + r^{ceq}\right)^2\right) = \delta^2 u\left(h_1\right)$$

or

$$r^{ceq} = \delta^{\frac{1}{1-\phi}} - 1 \tag{3.2.3}$$

which is increasing in the equity premium, μ, and in the optimal risky asset exposure, a^*. The certainty equivalent excess return, r^{ceq}, is the riskless return that would leave the young as well off as optimally investing their retirement wealth.

To conclude the description of complete markets, note that implementation of the optimal investment strategy requires individuals to actively manage their portfolios: as individuals grow older, their financial portfolio optimally shows a decrease in risk. The young hold the riskiest portfolio; to obtain their optimal risk exposure, the young have to take a leveraged position in the risky asset, i.e., they have to borrow at the risk-free rate and invest the proceeds in the risky asset. In practice, this strategy is infeasible if the young are credit constrained.

3.2.2 Credit Constraints

In the following, we examine each generation's optimal investment policy if financial markets are incomplete in the sense that there are credit constraints. Specifically, we assume that individuals cannot use future contributions as collateral;[16] individuals then face a per-period budget constraint that equals their financial reserve, $\alpha_t^i \leq w_t^i$.

Compared to the complete market benchmark, these additional constraints lead to a welfare loss. The young, in particular, are affected: they cannot invest in the risky asset until the next period, as w^y. The middle-aged may also run into credit constraints, depending on the parameters of the model. We obtain a lower bound for the welfare loss by assuming that only the young are credit constrained,

Proposition 3.2.2. *With credit constraints, the maximum expected lifetime utility for a young individual is*

$$\begin{aligned} U^{cc} &= Eu(\tilde{b}) \\ &= Eu\left(2s\left(1+a^*\tilde{r}_{t+1}\right)\right) \\ &= \delta u(2s) \end{aligned}$$

and the minimum welfare loss due to credit constraints is

$$\left(\delta^2-\delta\right)u(2s)$$

[16]Without collateral, the market for stock-investment credit breaks down due to adverse selection and moral hazard problems, cf. Bovenberg et al. (2007) and Constantinides et al. (2002).

Proof. See the Appendix. □

The proposition is intuitive. The minimum welfare loss arises if individuals face credit constraints when they are young, but invest optimally when they are middle-aged. The welfare loss equals the opportunity cost of not being able to invest in the risky asset while young. In utility terms, this opportunity cost is the utility gross return δ; in financial terms, the opportunity cost is the certainty equivalent excess return, r^{ceq}, given by (3.2.3).

When markets are incomplete, there is a role for pension funds: they can extend credit to the young and implement the young's preferred investment strategy for them. This is the rationale for pension funds that we examine. There are two reasons why pension funds are better placed than the private sector to extend credit to the young: (i) compulsory participation alleviates adverse selection problems; and (ii) pension funds have access to a tax on human wealth.[17]

To implement the optimal allocation of participants, pension funds can operate *extended generational accounts*, a form of generational accounting where each generation pays contributions into their account; and gets the balance of its account upon retirement. The difference with strict generational accounts, as in Teulings and de Vries (2006), is that the pension fund borrows on behalf of the young. Note that, with extended generational accounts, the pension fund consists simply of the merged, accounts of the participating generations; its investment strategy is a weighted average of the strategies of the participating generations.

Consider the pension fund at time t. From proposition 3.2.1, the young want to invest an amount

$$\begin{aligned} \alpha_t^y &= a^* \left(w_t^y + h_t^y \right) \\ &= a^* 2s \end{aligned}$$

[17]Bovenberg at al. (2007) note: 'compulsory participation in collective pension schemes can alleviate adverse selection and moral hazard when young workers borrow against their human capital to invest in equity.'

in the risky asset. Likewise, the middle-aged want to invest an amount

$$\begin{aligned} \alpha_t^m &= a^* \left(w_t^m + h_t^m \right) \\ &= a^* 2y \left(1 + a^* r_{t-1} \right) \end{aligned}$$

which depends on the stock return in period $t-1$. It follows that the pension fund invests a total amount of $\alpha_t^y + \alpha_t^m = a^* 2y \left(2 + a^* r_{t-1} \right)$ in stock on behalf of the young and the middle-aged. Whether the pension funds' stock investment exceeds its financial reserve depends (i) on past stock returns, and (ii) the parameters of the problem. It is possible, in other words, that the pension fund as a whole has to take a leveraged position in the stock market. The investment strategy of the pension fund is a weighted average of the preferred strategies of the young and the old.

The scheme we described above is similar to *generational accounting*, where each generation pays contributions into their own account; the account's investments are separately administered; and generations get the balance of their account when they retire. A pension fund then simply consists of the merged, ring-fenced, savings accounts of its participants. But generational accounting in the strict sense, as proposed in Teulings and de Vries (2006), does not allow the young to borrow against their future contributions, i.e., they are still credit constrained.[18]

We have argued that pension funds, operating within a mandatory DC pension scheme, can improve on a laissez-faire market allocation by allowing the young to implement their optimal investment strategy, i.e., by alleviating their credit constraints. The lower bound welfare gain of introducing a pension fund is given by proposition 3.2.2. As long as the pension fund does not need to borrow in financial markets, the optimal allocation can be interpreted as the young borrowing from older participants at the risk-free rate, pledging their human capital as collateral. The pension fund acts to secure the young's human capital as collateral to facilitate trades in its internal capital market. We note that the ability of pension funds to increase welfare hinges crucially on the assumption that pension funds can collateralize the human capital of participants. This assumption may be problematic,

[18]Such a pension fund would not improve on welfare in our stylized model. It is motivated by hyperbolic discounting and sharing of longevity risks in Teulings and de Vries (2006).

as we discuss in the concluding section.

3.3 Conclusion

We summarize our argument as follows. If the young face credit constraints, i.e., if markets are incomplete, then a pension fund is able to improve on welfare by implementing the young's preferred investment strategy. A pension fund can do this because (i) participation is mandatory and (ii) because it has access to a tax on human wealth. In effect, the pension fund is able to collateralize the human capital of its participants; in particular, the pension fund is able to collateralize the human capital of the young. Introducing a pension fund, thus, increases welfare compared to a situation in which everyone invests for himself.

The welfare increase by the pension fund relies crucially on the credibility of young's future contribution policy. In practice, it may be difficult to raise contributions after a low stock market outcome—and in particular in a recession. Clearly, the young do not want to raise contributions ex post, and it may be politically unpalatable to do so in times of recession.[19] If ex-post, after a low stock market outcome, the pension fund cannot collect contributions, then the result is a distributional conflict, similar to those witnessed in underfunded DB pension schemes. It follows that the DC pension scheme we describe may run into the same problems as a DB pension scheme with mismatch risk, and that the ex-ante optimal risk level at the pension fund cannot be separated from the ex post contribution policy.

Indeed, one interpretation of the distributional conflict in the Netherlands is that the pension funds—which, de facto, run a combination of DB and DC pension contracts—took too much risk on behalf of the young, assuming that contributions could be raised in case of a pension shortfall. When this proved infeasible, the old realized that they had implicitly lent their capital to the young and wished to see a return. The young on the other hand, do not wish contributions to be raised. This *ex post* distributional conflict leads to an *ex ante* governance conflict at the

[19]The young workers face a commitment problem: ex ante they want to pledge human capital as collateral, but ex post they have no incentive to raise pension contributions.

pension fund.[20] The old wish to limit risk taking such that they are repaid in every contingency. Alternatively, the old may agree to letting the young take a high risk if pension contributions can indeed be raised ex post.

Several important caveats apply. First of all, we've abstracted from the labor-supply and from the savings-consumption decision of individuals to focus on the optimal investment decision. This is, of course, restrictive. For example, credit constrained individuals would save more to compensate for the lack in risk taking opportunities; thus the pension fund becomes less important.

In this chapter I focus exclusively on the preferred risk exposure of participants as a determinant of the investment strategy of a pension fund. Bikker, Broeders, Hollanders, and Ponds (2009) provide evidence that indeed age composition is an important determinant of a pension fund's asset allocation. But there may be others. For example, we have abstracted from agency problems that may be present at the pension fund. Reward systems may favor risky investment, as portfolio managers gain from profits but are sheltered from losses.

Regarding the conceptual framework, our simple, stylized, model allows us to explore one rationale for pension funds. We show that a pension fund may be able to improve the welfare of credit constrained participants by letting them borrow within the pension fund. Our analysis could be extended to have an additional rationale for pension funds: to smooth risks between non-overlapping generations, as in Gollier (2008). In theory, this is an important benefit that pension funds can achieve. In practice, however, regulations—like the Dutch requirement to restore solvency within 5 years—inhibit such risk sharing. Furthermore, from a theory point of view, it is an open question how optimal risk sharing between non-overlapping generations is best achieved: through government policy (debt and tax) or through funded pension schemes. **?** present a stylized model in which optimal risk-sharing is achieved through both.

[20]This governance conflict resembles the conflict between debt- and equity holders in a firm that is close to financial distress Tirole and Dwatripont (1994). The young are, like equity holders, protected by limited liability, while the old have a claim on the pension fund that is debt-like.

References

BIKKER, J., D. BROEDERS, D. HOLLANDERS, AND E. PONDS (2009): "Pension funds' asset allocation and participant age: a test of the life-cycle model," Working papers 223, De Nederlandsche Bank.

BODIE, Z., R. MERTON, AND W. SAMUELSON (1992): "Labor supply flexibility and portfolio choice in a life cycle model," *Journal of Economic Dynamics and Control*, 16(3-4), 427–449.

GOLLIER, C. (2008): "Intergenerational risk-sharing and risk-taking of a pension fund," *Journal of Public Economics*, 92, 1463 – 1485.

TEULINGS, C., AND C. DE VRIES (2006): "Generational Accounting, Solidarity and Pension Losses," *de Economist*, 146, 63–83.

TIROLE, J., AND M. DWATRIPONT (1994): "A Theory of Debt and Equity," *Quarterly Journal of Economics*, 109(4), 1027–1054.

Chapter 4

Sand in the Wheels of Capitalism[1]

On the Political Economy of Capital Market Frictions

Abstract. This chapter develops a positive theory of capital market frictions, arising from a political conflict across different vintages of human capital. Older workers seek a political alliance to restrict the reallocation of capital between sectors, as this reduces their productivity and thus wages. Such an alliance is not feasible in a static framework, but may arise if capital market frictions are persistent over time. We show that a majority of voters chooses to restrict capital mobility if wealth is concentrated, and if technological obsolescence is high.

[1]This chapter is based on joint work with Enrico Perotti and Ernst-Ludwig von Thadden. For helpful comments we thank Philippe Aghion, Per Krusell, Enrique Schroth and seminar audiences at the ESWC 2010 Congress in Shanghai, and the EEA 2010 Congress in Glasgow.

4.1 Introduction

> *Political economists say that capital sets towards the most profitable trades, and that it rapidly leaves the less profitable non-paying trades. But in ordinary countries this is a slow process.*
>
> *(Bagehot, 1873)*

In a free market, capital moves naturally towards its most profitable use, leaving less productive activities. In reality, capital is reallocated fast in some countries, slow in others. Wurgler (2000) provides evidence that industries with better growth prospects invest more in countries that are more financially developed; these are also the countries in which declining sectors shrink faster. In a neoclassical economic framework, the financial sector should be functional to the needs of industry and trade, and these differences are attributed to institutional frictions in capital markets.

There are clear examples of institutional frictions in capital markets which are hard to reverse. Bankruptcy law, for instance, defines specific conditions to the assignment of assets from declining sectors. While in some countries bankruptcy law is designed to protect financial interests, in others–such as France and Italy–it explicitly instructs the liquidator to reassign capital in a manner which protects employment. As another example, state banking or specific financial regulators may be chosen on a mandate to protect traditional lending. For many years, banking in the U.S. was restricted to be local, assigning control over credit to established interests.

In this chapter, we adopt a political economy approach to explain the emergence and persistence of capital market frictions. In the tradition of classical political economy, we view the rules on capital reallocation as resulting from the political process that is shaped by economic interests. In practice, the allocation of capital across industries is heavily politicized, especially in more democratic countries. Political intervention may be *direct*, as when the government provides emergency loans or acquires companies outright; or *indirect*, as when the government adopts takeover regulations or bankruptcy laws that affect how much capital is reallocated.

Our starting insight is that labor is less mobile than capital. Here one should think of redeployable capital, such as land. While land can easily be redeployed, it is hard to retrain workers once they have acquired specific human capital. As human capital risk cannot be fully insured–for moral hazard reasons–workers are exposed to the risks that are specific to the sector they work in. The result of human capital specificity is a political conflict between citizens with different vintages of sunk human capital: agents with sector-specific human capital resist the reallocation of capital to newer sectors, as this leads to a reduction in their wages.

We show that in democracies a majority of the population wants to restrict capital mobility when the redistributive risk is large. This will be the case if wealth is concentrated, and if technological obsolescence is high. Young workers are the decisive, or pivotal, group in elections; they do not gain from capital market frictions immediately, but they would like to limit future capital reallocation, anticipating their old age, when they are less productive. A consumption smoothing motive then leads young workers' preferences to be partially aligned with the preferences of the old workers. Rapid technological change implies that the productivity gap between young and old workers is bigger, and therefore that the motive to impede capital reallocation is stronger.

An alliance against capital reallocation is never a political equilibrium when the capital market friction can be repealed at any time in the future. An alliance against capital reallocation can arise only if capital market frictions are persistent over time. So we posit that capital market frictions may occur if they can be introduced as institutional frictions–as opposed to reversible legislative choices.

4.1.1 Related Literature

Wurgler (2000) provides evidence that capital is reallocated more efficiently in countries with (i) more developed financial markets, (ii) a higher degree of minority investor protection, and (iii) a lesser extent of state ownership in the economy. Countries with deeper financial markets increase investment more in growing industries, and decrease investment more in declining industries. Our political explanation seeks to explain this pattern by endogenizing the resistance to capital reallocation.

This chapter is a contribution to the literature on the political determinants of financial market regulations and corporate governance, Pagano and Volpin (cf. 2005) or Perotti and von Thadden (2006). This literature emphasizes that the economic interests of capital investors can be subordinated to political considerations. In Pagano and Volpin (2005), labor forms an alliance with inside shareholders, in the contrast of a corporatist alliance with labor against financial investor's return. In Perotti and von Thadden (2006), a majority limits the ability of shareholders to allocate capital in order to limit risk for other stakeholders. This result arises only in more unequal societies, as in this chapter. Other related papers include Krusell and Ríos-Rull (1996) and Saint-Paul (2002), who study the political support for technological innovation, and labor market flexibility. The capital market plays no role in these papers, while it is central in this chapter. We offer an alternative channel to advance stakeholder interests: capital market frictions.

Hassler, Mora, Storesletten, and Zilibotti (2003) study the political support for a distortionary welfare state. The welfare state distorts private incentives to invest in education, which in turn gives rise to a constituency that supports the welfare state. Hassler et al. (2003) provide an example of how repeated majority voting in an OLG model can generate persistence in support of an inefficient welfare policy, we provide another. A key difference in our model is that all constituencies vote, whereas the young are disenfranchised in Hassler et al. (2003) Another difference is that our current framework does not allow for dynamic feedback of political choices through incentives as in Hassler et al. (2003); the composition of constituencies is fixed in this chapter.

Our approach is close to Azariadis and Galasso (2002), who study the political support for intergenerational transfers from young to old generations (such as pay-as-you go pension systems). They show that the young generation, who form a majority, may choose to set positive transfers if they can expect to receive a transfer when old. Our approach differs in two important aspects. First, we study political support for distortionary policies, whereas intergenerational transfers are efficient in Azariadis and Galasso (2002). Second, the young are a majority in Azariadis and Galasso (2002), while the outcome of our voting game is more complex: based on age and wealth differences, we identify four distinct voter classes, none of which forms a majority. Young workers are decisive because their preferences are less extreme than the preferences of other voter classes, not because of their number.

4.2 Model

We use a repeated two-period overlapping-generations model with an infinite horizon. Production requires capital and labor, and takes place in two sectors: a sector of young firms that employ the young generation; and a sector of old firms that employ the old generation. Labor is sector-specific while a fixed supply of capital can be used by all firms. We ignore capital growth in order to focus on the question how capital is allocated among different sectors. Time is denoted by subscripts $t = 0, 1, 2, ...,$; the different sectors are denoted by superscripts $j = Y, O$.

4.2.1 Production

At each time, there is a unit mass of identical firms in the young sector and a unit mass of identical firms in the old sector. All firms exist for two period; they use a vintage technology to produce a common consumption good that cannot be stored or saved. As young firms use the latest technology, they are more productive than old firms.

Production in each sector is given by a sector-specific productivity factor θ^j and a general production function F; production in the j-sector is

$$\theta^j F(K^j, L^j), \;\; j = Y, O$$

where K^j and L^j denote the amounts of capital and labor used in sector j, $\theta^O < \theta^Y$,and the price of output is normalized to 1. The production function F satisfies the common conditions, i.e., (i) production is increasing in both factors, at a decreasing rate; (ii) capital and labor are complementary factors of production; and (iii) the Inada conditions are satisfied.[2] Firms maximize profits facing competitive factor and output markets. Firms hire workers in competitive segmented labor

[2]Formally, F satisfies (i) $F_K, F_L > 0$ and $F_{KK}, F_{LL} < 0$, (ii) $F_{LK} = F_{KL} > 0$, and (iii) $\lim_{K \to 0} F_K = \lim_{L \to 0} F_L = \infty$, and $\lim_{K \to \infty} F_K = \lim_{L \to \infty} F_L = 0$.

markets and sell their output in a competitive output market. Reallocated capital is subject to a politically determined capital market friction.

Due to human capital specificity, the labor market is segmented: old firms hire old workers and pay wages w_t^O; young firms hire young workers and pay wages w_t^Y. The capital market has two important features: the cost of last period's retained capital is r_t, while firms pay an additional cost, c_t, if they wish to employ additional capital this period. The cost c_t represents a pure deadweight loss; it drives a wedge between the interest rate that capitalists receive and the cost of capital that firms pay. We refer to c_t as the capital market friction or, simply the reallocation cost.

As there are two costs of capital in the economy–one for retained capital and one for newly obtained capital–a firm's capital cost depends on the capital stock at the start of each period. We denote the initial capital stock of j-firms in period t by $\hat{K}_t^j$. Profits by firms of age j in period t are then given by

$$\theta^j F(K_t^j, L_t^j) - r_t K_t^j - c_t \max(0, K_t^j - \hat{K}_t^j) - w_t^j L_t^j$$

This is a standard expression for firm profits, except for the third term. Firms pay a marginal cost of capital $r_t + c_t$, if they want to attract capital beyond the initial stock of $\hat{K_t}^j$.

Young firms don't have retained capital ($\hat{K}_t^Y = 0$) and must attract all capital at a unit cost of $r_t + c_t$. When young firms turn into old firms–and the young generation turns into the old generation–they retain last period's capital ($\hat{K}_t^O = K_{t-1}^Y$). As old firms are less productive than young firms, there is an economic rationale for capital reallocation from old to young firms. When old firms go extinct, the capital they previously employ comes available to use elsewhere.

4.2.2 Agents

At each time, there are two generations of agents, the young and the old, each of unit mass. Young agents work in young firms, old agents work in old firms. All agents inelastically supply labor normalized at $\bar{L} = 1$ per period.

The fixed capital stock, $\bar{K} > 0$, is owned the *capitalists*, a fraction η of the old generation. Capitalists receive all firm profits and interest payments; they are identical and diversified.[3] It follows that capitalists receive

$$w_t^O + \frac{r_t + \Pi_t}{\eta}$$

where Π_t denotes aggregate firm profits. When capitalists die, a subset of their children inherits the capital stock; a fraction η of the young workers turns into old capitalists.

The population falls into four groups with identical lifetime income: young workers (YW), old workers (OW), young workers that will be capitalists (YC), and old capitalists (OC). The fraction of capitalists $\eta \in (0,1)$ is a measure of inequality among the old: higher η means more capitalists and less wealth per capitalist.

As income cannot be saved or stored, agents consume all income in each period. The lifetime utility of the young generation at time t is given by

$$U_t^{YW} := u(w_t^Y) + \delta u(w_{t+1}^O) \tag{4.2.1}$$

for young workers, and by

$$U_t^{YC} := u\left(w_t^Y\right) + \delta u\left(w_{t+1}^O + s_{t+1}\right) \tag{4.2.2}$$

for young capitalists, where $\delta \varepsilon (0,1]$ is the time discount factor; u is a standard felicity function with $u' > 0$ and $u'' < 0$; and $s_t := \frac{1}{\eta}\left(r_t \bar{K} + \Pi_t\right)$. Remaining lifetime utility of the old generation at time t is then

$$U_t^{OW} = u(w_t^O) \tag{4.2.3}$$

for the old workers, and

$$U_t^{OC} = u(w_t^O + s_t) \tag{4.2.4}$$

for the old capitalists. Note that agents do not optimize over economic choices, as they don't save and supply labor inelastically. Instead, agents optimize over

[3] All the debt and equity in the economy is owned by the capitalists, who all hold the same portfolio of assets.

political choices, by choosing a capital market friction in each period.

4.2.3 Interaction

Firms and agents interact in competitive factor and product markets. The product market is competitive and the price of the unique consumption good is normalized to 1. Each segment of the labor market is competitive and wages, w_t^j, adjust until the markets for young and old workers clear. There is one market for capital; it is competitive in the sense that the interest rate, r_t adjusts until the market clears, but transactions on this market are subject to the reallocation cost, c_t.

The reallocation cost is set by a vote: preceding market interaction, agents vote over c_t in each period. We return to the voting process in section 4, when we discuss political equilibrium. Timing in each period is as follows,

1. an *initial allocation* of capital is inherited from the previous period;

2. agents *vote* over the capital market friction c_t;

3. *economic activity* results in a new allocation of capital; and

4. agents get their *payoff*, i.e., their wage and capital income.

The political conflict has two dimensions: there is a class conflict between capitalists and workers; and there is a generational conflict between young and old. Old workers workers stand to lose most from free capital mobility: their wage drops, as capital is reallocated from old firms to young firms. Old capitalists too see their labor income drop, but their capital income increases. Preferences of the young generation depend on the nature of the capital market frictions, in particular, whether they are linked over time. Young workers may vote in favor of a positive reallocation cost, if they expect it to prevail until they are old. We analyze policy preferences and the resulting political equilibria in section 4.4. First we characterize the set of economic equilibria for a given sequence of capital market frictions.

4.3 Economic Equilibrium

4.3.1 Existence and Characterization

For a given sequence of capital market frictions,

Definition. An *economic equilibrium* is given by a sequence of factor prices and capital allocations $E = \{r_t, w_t^Y, w_t^O, K_t^Y, K_t^O\}_{t=0}^{\infty}$ such that in every period (i) firms maximize profits, and (ii) markets clear.

We prove the existence and uniqueness of an economic equilibrium for any sequence of capital market frictions $\{c_t\}_{t=0}^{\infty}$ with $0 \leq c_t < \infty$. In each period, firms take all prices as given and maximize the period profits.[4] Young firms start without capital, and pay the reallocation cost on each unit of capital they employ. They solve

$$\max_{K_t^Y, L_t^Y} \theta^Y F(K_t^Y, L_t^Y) - (r_t + c_t)K_t^Y - w_t^Y L_t^Y$$

which leads to standard first-order conditions

$$\theta^Y F_K(K_t^Y, L_t^Y) = r_t + c_t \tag{4.3.1}$$

$$\theta^Y F_L(K_t^Y, L_t^Y) = w_t^Y \tag{4.3.2}$$

and corresponding capital and labor demand, $K_t^Y(r_t, w_t^Y)$ and $L_t^Y(r_t, w_t^Y)$. Old firms retain the capital they employed last period; they solve

$$\max_{K_t^O, L_t^O} \theta^O F(K_t^O, L_t^O) - r_t K_t^O - c_t \max(0, K_t^O - \hat{K}_t^O) - w_t^O L_t^O$$

which leads to standard first-order condition

$$\theta^O F_L(K_t^O, L_t^O) = w_t^O \tag{4.3.3}$$

and corresponding labor demand $L_t^O(w_t^O, r_t)$. Note that the profit function of old firms is not differentiable at $\hat{K}_t^O$ and capital demand depends on whether old firms

[4]Firms maximize profits that accrue to their owners, i.e. the capitalists. As all capitalists are old, firms maximize current profits only.

adjust their capital. If old firms acquire extra capital in equilibrium, then K_t^O must satisfy

$$\theta^O F_K(K_t^O, L_t^O) = r_t + c_t \tag{4.3.4}$$

which is consistent if old firms indeed scale up, i.e., if

$$\theta^O F_K(K_t^O, L_t^O) < \theta^O F_K(\hat{K}_t^O, L_t^O)$$

or

$$r_t < \theta^O F_K(\hat{K}_t^O, L_t^O) - c_t$$

Old firms increase their capital if the interest rate is sufficiently small; likewise, old firms decrease their capital if the interest rate is sufficiently big, or

$$r_t > \theta^O F_K(\hat{K}_t^O, L_t^O).$$

For intermediate values of the interest rate, old firms keep using the capital they from last period, $\hat{K}_t^O = K_{t-1}^Y$.

In each period, young and old agents inelastically supply labor, normalized to 1, in their respective labor markets. As both segments of the labor market must clear in equilibrium, we obtain the equilibrium wage rate as a function of capital from (4.3.2) and (4.3.3):

$$\begin{aligned} w_t^Y &= \theta^Y F_L(K_t^Y, 1) && (4.3.5) \\ w_t^O &= \theta^O F_L(K_t^O, 1). && (4.3.6) \end{aligned}$$

Sector wages increase, or decrease, along with an increase, or decrease, of sector capital.

Capital demand of young firms follows from (4.3.1) and, letting $g(K) := F_K(K, 1)$, we can write

$$K_t^Y(r_t) = g^{-1}\left(\frac{r_t + c_t}{\theta^Y}\right) \tag{4.3.7}$$

Our earlier discussion shows that capital demand of old firms is

$$K_t^O(r_t) = \begin{cases} g^{-1}\left(\frac{r_t+c_t}{\theta^O}\right) & \text{if} \quad r_t < \underline{r}_t \\ g^{-1}\left(\frac{r_t}{\theta^O}\right) & \text{if} \quad r_t > \bar{r}_t \\ \hat{K}_t^O & \text{if} \quad \underline{r}_t \leq r_t \leq \bar{r}_t \end{cases} \tag{4.3.8}$$

with $\underline{r}_t \equiv \theta^O g(\hat{K}_t^O) - c_t$, and $\bar{r}_t \equiv \theta^O g(\hat{K}_t^O)$. Total capital demand then is

$$\varphi_{c_t}(r_t) := K_t^Y(r_t) + K_t^O(r_t) \tag{4.3.9}$$

and the capital market clears if $\varphi_{c_t}(r_t) = \bar{K}$. The following lemma shows that a market clearing interest rate exists and is unique in each period.

Lemma 4.3.1. *For $0 \leq c_t < \infty$, there exists a unique market clearing interest rate $r_t^* = r_t^*(c_t)$.*

Proof. For any $0 \leq c_t < \infty$, φ_{c_t} is continuous, strictly decreasing, and piecewise differentiable in r_t (with kinks at $\underline{r}_t$ and $\bar{r}_t$). Furthermore, by the Inada conditions we have

$$\lim_{r \to \infty} \varphi_{c_t}(r_t) = 0 \quad \text{and} \quad \lim_{r_t \to -c_t} \varphi_{c_t}(r_t) = \infty$$

Hence by the continuity of φ_{c_t}, there is $r_t^* > -c_t$ such that

$$\varphi_{c_t}(r_t^*) = \bar{K} \tag{4.3.10}$$

By the strict monotonicity of φ_{c_t}, r_t^* is unique. □

From the market clearing interest rate, r_t^*, equilibrium capital demands and sector wages readily follow.[5] Hence, we have proven the following proposition:

Proposition 4.3.1. *For any sequence of capital market frictions $\{c_t\}_{t=0}^{\infty}$ with $0 \leq c_t < \infty$, there exists a unique economic equilibrium E.*

The previous argument, and in particular capital demand of old firms, given by (4.3.8), shows that the economic equilibrium is characterized by the initial

[5] If the market clearing interest rate is negative, we set it to 0, which implies that some capital remains unused. We show in section 4.4.1 that no voter class wishes to set c_t so high as to induce a negative interest

capital of the old firms, $\hat{K}_t^O$, and the prevailing capital market friction in the capital market, c_t. We characterize firm behavior in equilibrium in the following.

Depending on the equilibrium interest rate, old firms adjust their capital, down- or upward, or they keep the capital of last period. We investigate these cases in turn. Old firms scale up capital if and only if the interest rate is sufficiently small, or $r_t^* < \underline{r}_t$, cf. (4.3.8). Market clearing then reads

$$g^{-1}\left(\frac{r_t^* + c_t}{\theta^Y}\right) + g^{-1}\left(\frac{r_t^* + c_t}{\theta^O}\right) = \bar{K} \tag{4.3.11}$$

which implicitly gives the equilibrium interest rate, r_t^*; old firms' capital is given by

$$\underline{K}_t^O := g^{-1}\left(\frac{r_t^* + c_t}{\theta^O}\right) \tag{4.3.12}$$

It follows that old firms scale up if and only $\hat{K}_t^O < \underline{K}_t^O$. Note that, through (4.3.11), old firms' equilibrium capital, $\underline{K}_t^O$, does not depend on c_t.

Similarly, old firms scale down capital if and only if the interest rate is sufficiently big, or $r_t^* > \bar{r}_t$. Market clearing then reads

$$g^{-1}\left(\frac{r_t^* + c_t}{\theta^Y}\right) + g^{-1}\left(\frac{r_t^*}{\theta^O}\right) = \bar{K} \tag{4.3.13}$$

which gives the interest rate. Old firms' capital is given by

$$\bar{K}_t^O := g^{-1}\left(\frac{r_t^*}{\theta^O}\right) \tag{4.3.14}$$

so that old firms scale down if and only if $\hat{K}_t^O < \bar{K}_t^O$. Note that equilibrium capital of old firms, $\bar{K}_t^O = \bar{K}^O(c_t)$, is a function of the capital market friction, c_t.

Finally, old firms keep their initial capital if and only if

$$\underline{K}_t^O \leq \hat{K}_t^O \leq \bar{K}_t^O(c_t)$$

in which case the equilibrium interest rate is given by

$$r_t^* = \theta^Y g(\bar{K} - \hat{K}_t^O) - c_t \tag{4.3.15}$$

This concludes the description of old firms in equilibrium.

Equilibrium behavior of young firms is easily described: as they have no initial capital, they must adjust capital upward. It follows that equilibrium capital of young firms is given by

$$g^{-1}\left(\frac{r_t^* + c_t}{\theta^Y}\right)$$

where r_t^* depends on the equilibrium behavior of old firms, i.e., on $\hat{K}_t^O$ and c_t.

As an illustration, consider the economic equilibrium if there is no capital market friction, i.e., if $c_t = 0$. Then there is only one cost of capital in the economy and capital market clearing reads

$$g^{-1}\left(\frac{r_t^*}{\theta^Y}\right) + g^{-1}\left(\frac{r_t^*}{\theta^O}\right) = \bar{K}$$

As old firms are less productive than young firms, old firms employ less capital than young firms in equilibrium if there is no capital market friction. If the capital market friction is positive ($c_t > 0$), then the decision of old firms to adjust capital can go either way–as we have shown–and there is no guarantee that old firms employ less capital than young firms. But in a dynamic economic equilibrium, firms in the old sector do not scale up, except possibly in the first or second period..

Lemma 4.3.2. *In economic equilibrium, old firms do not scale up if $t \geq 2$.*

Proof. Consider an arbitrary economic equilibrium and fix $t \geq 2$. We prove that $\hat{K}_t^O \geq \underline{K}_t^O$, so that old firms do not scale up in period t. For our argument we consider the economic equilibrium at time $t-2$ and assume that $c_{t-2} = 0$. Then old firm equilibrium capital is given by

$$K_{t-2}^{O*} = \underline{K}^O$$

and is minimized across all state-parameter pairs $(\hat{K}_{t-2}^O, c_{t-2})$. By market clearing, it follows that young firm capital $\left(K_{t-2}^{Y*}\right)$ is maximized across state-parameter pairs, and we denote it by $\bar{K}^Y$. Note that $\bar{K}^Y$ exceeds $\underline{K}^O$, as young firms are more productive than old firms. Moving forward one period, it follows that $\hat{K}_{t-1}^O > \underline{K}^O$, which means that old firms do not scale up in period $t-1$. Hence equilibrium old firm capital satisfies $K_{t-1}^{O*} \leq \bar{K}^Y$, and market clearing implies that equilibrium young firm capital satisfies $K_{t-1}^{Y*} \geq \underline{K}^O$. We have shown that the minimum value K_{t-1}^{Y*} can take in equilibrium exceeds $\underline{K}^O$. □

The intuition for Lemma 4.3.2 is that, independent of the sequence of capital market frictions, young firms choose to employ more capital than what they will need when they are old firms.

4.3.2 Steady States

For arbitrary sequences of frictions $\{c_t\}_{t=0}^{\infty}$, capital market activity has little structure. As we are mostly interested in stable political outcomes that yield constant sequences of c_t, we look for steady state equilibria.

Definition. A steady state is an economic equilibrium such that $(K_t^Y, K_t^O) = (K_{t+1}^Y, K_{t+1}^O)$ for all t.

In a steady state, the amount of capital that is reallocated from old firms to young firms is constant over time; it is positive by lemma 4.3.2. The next proposition gives all steady states that exist for a constant sequence of capital market frictions.

Proposition 4.3.2. *Given a constant sequence of capital market frictions* $\{c\}_{t=0}^{\infty}$, *there exists a bound* $\bar{c} > 0$ *such that*

1. for $c \in [0, \bar{c}]$, the unique steady state equilibrium is given by

$$K^{Y*} = g^{-1}\left(\frac{r^* + c}{\theta^Y}\right), \quad K^{O*} = g^{-1}\left(\frac{r^*}{\theta^O}\right) \tag{4.3.16}$$

   with r^* given by the market clearing condition,

$$g^{-1}\left(\frac{r^* + c}{\theta^Y}\right) + g^{-1}\left(\frac{r^*}{\theta^O}\right) = \overline{K} \tag{4.3.17}$$

   (a) for $c \geq \bar{c}$, the unique steady state equilibrium is given by

$$K^{Y*} = K^{O*} = \frac{1}{2}\bar{K} \tag{4.3.18}$$

   with $r^* = \theta^Y g(\frac{1}{2}\bar{K}) - c$.

Proof. For 1: Let time t capital allocations be given by (4.3.16). We check that the same allocation of capital results in period $t+1$, for which it suffices to check that old firms scale down in period $t+1$. Old firms scale down if and only if

$$\hat{K}^O_{t+1} = K^{Y*}_t > \bar{K}^O_{t+1} \tag{4.3.19}$$

where $\bar{K}^O_{t+1} = \bar{K}^O(c)$ is the equilibrium cut-off level defined by (4.3.14). Note that condition (4.3.19) holds for $c = 0$, since without frictions the more productive young firms attracts more capital than the old firms. Taking the total derivative of (4.3.17) with respect to c shows that $r^*_t + c$ is strictly increasing in c. It follows that K^{Y*}_t is strictly decreasing in c. Furthermore, $\bar{K}^O_{t+1}$ is strictly increasing in c as noted after (4.3.14). Hence, the proposed equilibrium is a steady state for $c \in [0, \bar{c}]$, with $\bar{c}$ defined by

$$K^{Y*}_t(\bar{c}) = \bar{K}^O(\bar{c}) \tag{4.3.20}$$

Now consider the steady state for $c = \bar{c}$. By the definition of $\bar{c}$, we have $\hat{K}^O = \bar{K}^O$ which implies that old firms do not adjust capital, and so $K^O = \hat{K}^O$. Since we also have $\hat{K}^O = K^Y$, it follows that $K^O = K^Y = \frac{1}{2}\bar{K}$ in this boundary case steady state. With (4.3.16), we then obtain r^* and $\bar{c}$:

$$r^* = \theta^O g\left(\frac{1}{2}\bar{K}\right)$$

and

$$\bar{c} = (\theta^Y - \theta^O) g\left(\frac{1}{2}\bar{K}\right) \tag{4.3.21}$$

For 2: let $c > \bar{c}$ and the time t allocations be given by (4.3.18). This allocation induces the same capital allocation in period $t+1$, if

$$\underline{K}^O \leq \frac{1}{2}\bar{K} < \bar{K}^O_{t+1}$$

The first inequality holds trivially by (4.3.11) and (4.3.12), the second holds by the definition of $\bar{c}$ since $\bar{K}^O_{t+1} = \bar{K}^O(c)$. Finally, note that every economic equilibrium is unique by lemma 4.3.1. □

Proposition 4.3.2 gives all steady state equilibria that exist for constant sequences of capital market frictions $\{c\}_{t=0}^{\infty}$. We denote these steady states by E_c.[6] As a benchmark, consider the steady state equilibrium that results if the capital market is frictionless, E_0. If the sequence of capital market frictions is given by $c_t = 0$ for all t, then all capital has the same rental price. It follows from proposition 4.3.2 that young firms' capital is maximized in this steady state–and old firms' capital minimized. As total capital is constant and young firms are more productive than old firms, E_0 is the steady state in which maximum output is achieved. Hence $c = 0$ would be chosen by a social planner if lump sum transfers are available.

In any period, a positive capital market friction reduces output, an economic inefficiency. The economic inefficiency, however, does not constitute a Pareto inefficiency, as old workers stand to benefit from it. We rule out vote buying as a means to restore the economically efficient outcome, i.e., we rule out that the capitalists compensate the workers to vote for a free (or frictionless) capital market. As Acemoglu (2003) has argued, there is an essential hold-up problem that prevents such trades from taking place: if the workers vote for $c_t = 0$, the capitalists have no incentive to compensate them ex post; likewise, the workers have no incentive to vote for $c_t = 0$, if they receive the compensation upfront.

4.4 Political Equilibrium

We endogenize the sequence of capital market frictions by treating the frictions as politically determined. To include politics, we extend the economic model by a simple majority vote in each period. In order to obtain closed form solutions, we assume that production is Cobb-Douglas, with decreasing returns to scale; production is given by

$$F(K,L) = K^{\alpha}L^{\beta} \tag{4.4.1}$$

with $0 < \alpha + \beta < 1$. We derive steady state capital allocations and factor prices in appendix A.3.1.

[6] From the Proof of Prop 4.3.2, we see that a steady state with $K^Y = K^O = \frac{1}{2}\bar{K}$ also exists for non constant sequences $\{c_t\}_{t=0}^{\infty}$ as long as $\bar{c} < c_t < \infty$. We do not consider these in the following.

4.4.1 Policy Preferences

Voter preferences follow from the lifetime utility functions (4.2.1) - (4.2.4) and equilibrium factor prices. Remember that for a given value of the reallocation cost, c_t, the economic equilibrium at time t is fully characterized by the initial capital of old firms, $\hat{K}_t^O$.[7] The next lemma establishes two useful properties of the equilibrium interest rate

Lemma 4.4.1. *In each period, the equilibrium interest rate r_t^* is strictly decreasing in the reallocation cost c_t, and $r_t^* + c_t$, is nondecreasing in c_t.*

Proof. (i) Let $0 < c_t < \infty$. By lemma 4.3.1, there is a unique $r_t^*(c_t)$ such that

$$\varphi_{c_t}(r_t^*(c_t)) = \bar{K}$$

Now consider a $c_t' > c_t$. An inspection of (4.3.9) shows that $\varphi_{c_t'} < \varphi_{c_t}$ uniformly, so $\varphi_{c_t'}(r_t^*(c_t)) < \bar{K}$. Again by lemma 4.3.1, there is a unique $r_t^*(c_t')$ such that

$$\varphi_{c_t'}(r_t^*(c_t')) = \bar{K}$$

Because $\varphi_{c_t'}$ is strictly monotonically decreasing, it follows that $r_t^*(c_t') < r_t^*(c_t)$. (ii) Suppose that $r_t^*(c_t) + c_t$ is decreasing in c_t. Then result (i) implies that $K_t^{Y^*}$ and $K_t^{O^*}$ must increase in c_t which contradicts market clearing. □

The following lemma summarizes the comparative statics of factor prices, sector capital and voter class income with respect to c_t. In order to have the necessary generality to analyze out-of-steady-state deviations, the lemma is formulated for arbitrary economic equilibria.

Lemma 4.4.2. *Let $\bar{c}_t$ be defined by*

$$\hat{K}_t^O = \bar{K}_t{}^O(\bar{c}_t). \tag{4.4.2}$$

Then (i) K_t^Y and w_t^Y are strictly decreasing in $c_t \in [0, \bar{c}_t]$ and constant for $c_t > \bar{c}_t$; (ii) K_t^O and w_t^O are strictly increasing in $c_t \in [0, \bar{c}_t]$ and constant for $c_t > \bar{c}_t$; (iii) $w_t^O + s_t$ is decreasing in c_t.

[7]From lemma 4.3.2 we know that $\hat{K}_t^O \geq \underline{K}^O$, because old firms do not scale up in equilibrium.

Proof. See appendix A.3.2 □

The function $\bar{K}_t{}^O$ that yields the equilibrium cut-off value $\bar{K}_t{}^O(\bar{c}_t)$ has been defined after (4.3.14). Note that $\bar{c}_t$ is the boundary value of the capital market friction at which old firms keep their initial capital $\hat{K}_t^O$: for lower values of the friction, old firms scale down.[8] The proof of the lemma is straightforward but long and we provide it in the appendix.

Lemma 4.4.2 shows that total income of old capitalists is decreasing in c_t, even as wages may be increasing; the positive wage effect is dominated by the negative capital income effect.[9] Note that once c_t is fixed for a given period, the equilibrium interest rate r_t^* and all other time t equilibrium values readily follow. Hence we may write lifetime utility of agents as a function of the capital market frictions that prevail today and tomorrow:

$$U_t^i = U_t^i(c_t, c_{t+1}) \quad \text{for} \quad i\varepsilon\{OW, OC, YW, YC\}$$

We note that income of the old generation does not, in fact, depend on c_{t+1}. The following definition of single peakedness goes back to Black (1948).

Definition. Policy preferences of voter class $i\varepsilon\{OW, OC, YW, YC\}$ are single-peaked if the following statement is true:

for any

$$c^i \varepsilon Arg \max_{c_t} U^i(c_t, c_{t+1})$$

if $c'' \leq c' \leq c^i$, or if $c^i \leq c' \leq c''$, then $U^i(c'', c_{t+1}) \leq U^i(c', c_{t+1})$.

As is well-known, with single-peaked preferences we can apply a median voter theorem: for existence of an equilibrium of the simple majority vote, it suffices that all voters have single-peaked preferences. Preferences of the old generation can easily be characterized.

[8]Since we consider deviations out-of-steady-state here, we may restrict attention to values $\hat{K}_t^O \geq \frac{1}{2}\bar{K}$ in the following (cf. proposition 4.3.16). It follows that $\bar{c}_t \geq \bar{c}$, where $\bar{c}$, given by (4.3.20), is the boundary value of the steady state friction for which no reallocation of capital takes place.

[9]This holds for any value of η.

Lemma 4.4.3. *In each period, preferences of the old generation are single-peaked, and their preferred policies are*

$$Arg \max_{c_t} U_t^{OW}(c_t, c_{t+1}) = [\bar{c}_t, \infty)$$

for the old workers, and

$$arg \max_{c_t} U_t^{OC}(c_t, c_{t+1}) = 0$$

for the old capitalists.

Proof. The preferences over c_t follow from lemma 4.4.2. Because the utility of the *OC* and *OW* class is monotonic in c_t, it follows that they are also single peakedness. □

Turning to the preferences of the young, we note that lifetime utility of the young generation depends on current policy c_t as well as future policy c_{t+1}. Keeping tomorrow's friction fixed, we obtain the young's preferences over c_t.

Lemma 4.4.4. *Preferences of the young generation over c_t are single-peaked, and preferred policies are*

$$arg \max_{c_t} U_t^{YW}(c_t, c_{t+1}) = 0$$

and

$$arg \max_{c_t} U_t^{YC}(c_t, c_{t+1}) = 0$$

Proof. Let c_{t+1} be fixed. The YW class and the YC class have the same income w_t^Y in period t. By lemma 4.4.2, w_t^Y is strictly decreasing in c_t. Hence, the utility of the YW and YC are monotonically decreasing in c_t which is sufficient for single peakedness.. □

A majority consisting of old capitalists and all the young favor a capital market without frictions. The intuition is that wages in the Y-sector and capitalist income are maximized in a frictionless capital market. Importantly, we have derived preferences keeping future policy fixed. If the future friction c_{t+1} depends on the currently prevailing friction c_t, then preferences of the young change while preferences of the old are still given by lemma 4.4.3.

Lemmas 4.4.3 and 4.4.4 show that young capitalists achieve maximum lifetime utility if $c_t = 0$ and $c_{t+1} = 0$. Young workers on the other hand achieve maximum utility if $c_t = 0$ and $c_{t+1} \geq \bar{c}_{t+1}$. Young workers prefer a frictionless capital market when young and wish to see the maximum friction when old.

We turn to the question in which steady state the utility U_t^i of the different voter classes $i\varepsilon\{OW, OC, YW, YC\}$ is maximized. Note that in steady state the reallocation cost is time independent, as are the utilities, so that we may write $U^i = U^i(c)$. It follows from the above that U^{YC} and U^{OC} are maximized in the steady state without friction, E_0. Old worker utility U^{OW} is maximized in steady states with no capital reallocation, i.e. E_c with $c \geq \bar{c}$. Young workers are the only group who may have more interesting preferences as we show in the following.

Consider young worker utility $U^{YW} = u(w^Y) + \delta u(w^O)$. Steady state wages follow from (4.3.5) and (4.3.6) and the steady state capital allocations we derived in section A.3.1. We have

$$w^Y = \theta^Y \beta \left(\frac{\alpha \theta^Y}{r(c) + c} \right)^{\frac{\alpha}{1-\alpha}} \;\; ; \;\; w^O = \theta^Y \beta \left(\frac{\alpha \theta^O}{r(c)} \right)^{\frac{\alpha}{1-\alpha}}$$

for $c \in [0, \bar{c}]$, and

$$w^Y = \theta^Y \beta \left(\frac{1}{2} \bar{K} \right)^{\frac{\alpha}{1-\alpha}} \quad ; \quad w^O = \theta^Y \beta \left(\frac{1}{2} \bar{K} \right)^{\frac{\alpha}{1-\alpha}}$$

for $c > \bar{c}$. The following lemma gives the preferred steady state policy of the young worker.

Lemma 4.4.5. *Steady state preferences of the young worker are single-peaked. Their preferred policies are*

- $c^{YW} = 0$ if $\frac{u'(w^Y)}{\delta u'(w^O)} > \frac{r(c)}{r(c)+c}$ on $(0, \bar{c})$,
  - $c^{YW} = \bar{c}$ if $\frac{u'(w^Y)}{\delta u'(w^O)} < \frac{r(c)}{r(c)+c}$ on $(0, \bar{c})$,
  - otherwise c^{YW} is given by

$$\frac{u'(w^Y)}{\delta u'(w^O)} = \frac{r(c^{YW})}{r(c^{YW}) + c^{YW}} \tag{4.4.3}$$

Proof. Consider $U^{YW}(c)$ on $(0,\bar{c})$. Taking the derivative of U^{YW} with respect to c, we get

$$\frac{dU^{YW}}{dc} = u'(w^Y)\frac{dw^Y}{dc} + \delta u'\left(w^O\right)\frac{dw^O}{dc} \tag{4.4.4}$$

Substituting $\frac{dw^Y}{dc}$ and $\frac{dw^O}{dc}$ in (4.4.4) yields

$$\frac{dU^{YW}}{dc} = \frac{-\beta\alpha}{1-\alpha}\left[u'(w^Y)\theta^Y\frac{\left(K^Y\right)^{\alpha}}{r+c}\left(\frac{dr}{dc}+1\right) + \delta u'\left(w^O\right)\theta^O\frac{\left(K^O\right)^{\alpha}}{r}\frac{dr}{dc}\right]$$

where K^Y and K^O are the steady state capital allocations given by (A.3.2) and r follows from $K^Y + K^O = \bar{K}$. Implicit differentiation of the capital market clearing condition with respect to c shows that

$$\frac{dr}{dc} + 1 = -\left(\frac{\theta^O}{\theta^Y}\right)^{\frac{1}{1-\alpha}}\left(\frac{r+c}{r}\right)^{\frac{1}{1-\alpha}}\frac{dr}{dc}$$

which we use to obtain

$$\frac{dU^{YW}}{dc} = \frac{-\beta}{1-\alpha}\frac{dr}{dc}\frac{1}{r}K^O[\delta u'\left(w^O\right)r - u'\left(w^Y\right)(r+c)]$$

Since r and K^O are nonnegative and $\frac{dr}{dc} < 0$, the first part of this expression is positive for all c. Hence $\frac{dU^{YW}}{dc} \geq 0$ if and only if

$$\delta u'\left(w^O\right)r - u'\left(w^Y\right)(r+c) \geq 0 \tag{4.4.5}$$

Note that $u'\left(w^Y\right)(r+c)$ strictly increases in c while $\delta u'\left(w^O\right)r$ strictly decreases in c. It follows that if for some c condition (4.4.5) is satisfied with equality then it is the unique utility maximizing steady state policy. Otherwise, either $\delta u'\left(w^O\right)r - u'\left(w^Y\right)(r+c) > 0$ for all $c\varepsilon(0,\bar{c})$ so that $Arg\max_c U^{YW}(c) = [\bar{c},\infty)$; or $\delta u'\left(w^O\right)r - u'\left(w^Y\right)(r+c) < 0$ for all $c\varepsilon(0,\bar{c})$ so that

$$Arg\max_c U^{YW}(c) = \{0\}$$

Finally note that preferences are single peaked. □

Lemma 4.4.5 shows that young workers may prefer a steady state with a positive capital market friction. A positive c allows the young worker to smooth consumption over his lifetime; the capital market friction works as a savings technology. Equation (4.4.3) can be interpreted as an optimal savings condition. If utility is linear, for example, there is no consumption smoothing motive and (4.4.3) cannot be satisfied. Condition (4.4.3) shows that c^{YW} depends on the functional form u, the discount rate δ and technological factors θ^Y and θ^O. Implicit derivation of c^{YW} gives the following comparative statics.

Lemma 4.4.6. *c^{YW} is increasing in δ; decreasing in θ^O; and increasing in θ^Y.*

Lemma 4.4.6 is intuitive: the consumption smoothing motive is increased if future consumption is valued more (increase in δ) and if the wage gap between young and old age is bigger (decrease in θ^O; increase in θ^Y). This concludes our description of voter preferences.

4.4.2 Majority Voting

In this section we analyze equilibria under *pure majority rule*. A pure majority rule is defined by three characteristics: (i) democracy is direct so that voters directly choose their preferred capital market friction; (ii) voters vote sincerely; and (iii) there is an open agenda so that all alternatives are considered in the vote. Every period t, all agents cast a vote over the capital market friction c_t.

Formally, an action at time t for a member of voter class i is a capital market friction $a_t^i \in [0,\infty)$, we also refer to actions as votes. At time t, the publicly known history of the game is $h_t = (c_0, .., c_{t-1}) \varepsilon H_t$ with $H_t = [0,\infty)^t$. A strategy for voter class i at time t is given by a mapping $v_t^i : H_t \rightarrow [0,\infty)$.

Definition. A *political economic equilibrium* is a sequence of factor prices and allocations $E = \{r_t, w_t^Y, w_t^O, K_t^Y, K_t^O\}_{t=0}^{\infty}$, supplemented with a voting strategy profile $v = \{v_t^{YW}, v_t^{YC}, v_t^{OW}, v_t^{OC}\}_{t=0}^{\infty}$, such that (i) v is an equilibrium of the voting game; and (ii) E is an economic equilibrium given the sequence of reallocation costs $\{c_t\}_{t=0}^{\infty}$ in the outcome of v.

We focus on steady state equilibria that can be politically supported, i.e. steady state equilibria E_c for which there is a voting strategy v such that $\{E_c, v\}$ is a political economic equilibrium.

Open-Loop Equilibrium

As a benchmark we consider the political equilibrium that results if voters play open-loop strategies, i.e. strategies that do not depend on history.

Proposition 4.4.1. *If voters play open-loop strategies, then the unique political economic equilibrium is given by E_0 and the voting strategy profile $v = \{v_t^{YW} = 0, v_t^{YC} = 0, v_t^{OW} = \bar{c}_t, v_t^{OC} = 0\}_{t=0}^{\infty}$.*

Proof. Strategy profile v follows from sincere voting and voter preferences derived in lemma 4.4.3 and lemma 4.4.4. We see that a majority of agents votes for $c_t = 0$ in every period. The corresponding steady state is E_0. □

Capital market frictions cannot arise in the political equilibrium under open-loop strategies. Capitalists achieve maximum lifetime utility in this equilibrium. The young worker class, while decisive, does not achieve her maximum lifetime income (which is achieved for $c = c^{YW}$, cf. lemma 4.4.5). Open-loop strategies leave no scope for cooperation between young workers in different periods. This result is reminiscent of results in Sjoblom (1985) and Azariadis and Galasso (2002), who show that social security cannot be supported if voters play open-loop strategies.

Subgame Perfect Equilibrium

Once we allow for richer voting strategies, cooperation between subsequent young generations can be achieved. In particular, any steady state in which the young workers get a higher utility than their open loop utility can be supported as a political equilibrium.

Proposition 4.4.2. *There exists $c^* \in [0, \bar{c}]$ such that for every $c \in [0, c^*]$, the steady state equilibrium E_c can be politically supported.*

Proof. Consider the set

$$D := \left\{c\varepsilon[0,\infty) | U^{YW}(c,c) \geq U^{YW}(0,0)\right\}$$

D is nonempty, as $0 \in D$; D is closed, as U^{YW} is continuous; and D is convex, as U^{YW} is single-peaked. It follows that D is a closed interval which contains c^{YW}, the preferred lifetime policy of the young workers (cf. lemma 4.4.5). Now, let $c^* := \min\{\bar{c}, \sup D\}$ and choose an arbitrary $c \in [0, c^*]$. Then E_c can be politically supported. Consider the voting strategy profile $v^* = \{v_t^{YW}, v_t^{YC}, v_t^{OW}, v_t^{OC}\}_{t=0}^{\infty}$ such that

$$\begin{aligned}
v_t^{YW} &= \begin{cases} c & \text{if} \quad c_{t-s} = c \quad \text{for} \;\; s = 1, \ldots t \\ 0 & \qquad\qquad \text{otherwise} \end{cases} \\
v_t^{YC} &= v_t^{OC} = 0 \\
v_t^{OW} &= \bar{c}_t
\end{aligned}$$

With this strategy profile the voting game equilibrium is c in every period provided the play started with $c_0 = c$. The best deviation of the YW class at time t is to set $c_t = 0$ given that $v_{t+1}^{YW} = 0$. This deviation is not profitable since $U_t^{YW}(0,0) \leq U_t^{YW}(c,c)$. It remains to check that $v_{t+1}^{YW} = 0$ is incentive compatible. It is because $U_{t+1}^{YW}(0,0) > U_{t+1}^{YW}(c_{t+1},0)$ for any $c_{t+1} > 0$. We have shown that v^* is a subgame perfect strategy profile of the repeated voting game. We conclude that $\{E_c, v^*\}$ is a political economic equilibrium. □

The political equilibria with subgame-perfect voting strategies can be interpreted as arising from *social contract* which allows for cooperation between current and future young. The familiar logic for cooperation is that voters cooperate as long as they expect future generations to honor the social contract if they do so themselves.

Since there are multiple equilibria of the repeated voting game, there is no guarantee that cooperation will be achieved. Following the literature on dynamic political economy models, we can reduce the multiplicity of equilibria in two ways. First, in the spirit of Azariadis and Galasso (2002), we may impose more structure on the political model. Second, we may restrict the solution concept as in Hassler et al. (2003) and related work that studies policies that can be sustained without commitment (e.g. Klein, Krusell, and Ríos-Rull, 2008). We pursue both routes in the following.

Policy Persistence

We restrict the political model and assume that voters cannot overturn policy in every period. Instead, voters at time t set a persistent policy that lasts throughout their lifetime (i.e. $c_t = c_{t+1}$). Note that this means that the next generation of voters is disenfranchised.[10]

As agents vote sincerely, actions follow from voter preferences derived in section 4.4.1. Recall that utility of old capitalists is maximized for $c_t = 0$. Utility of young capitalists is maximized for $c_t = 0$ and $c_{t+1} = 0$. Hence all capitalists vote for a zero friction, or

$$a_t^{OC} = a_t^{YC} = 0$$

If capitalists form a majority (i.e. $\eta > \frac{1}{2}$), the policy outcome is $c_t = c_{t+1} = 0$. Consequently E_0 is the unique steady state that can be politically supported if capitalists are a majority. If capitalists are a minority (i.e. $\eta < \frac{1}{2}$)–the more realistic case that we focus on in the following–then worker preferences are decisive for the political equilibrium. We have seen before that utility of the old worker class is maximized if all capital $\hat{K}_t^O$ is retained in the old firms where they work. It follows that

$$a_t^{OW} = \bar{c}_t$$

As for the young worker class, they face a tradeoff: a higher capital market friction leads to a decrease in their wage when young $\left(w_t^Y\right)$ and an increase in their wage when old $\left(w_{t+1}^O\right)$. If young voters can set the policy for two periods, then they will vote *as if* choosing among steady state utility levels.[11]

Lemma 4.4.7. *Consider a vote at time t under policy persistence. Then young workers choose $a_t^{YW} = c^{YW}$, where c^{YW} is given by lemma 4.4.5.*

[10] However, we show that disenfranchised median voters are better off. The intuition is that the YW class achieves higher utility if the next generation can be bound to its choice.

[11] It is not a priori clear whether the economy reaches a new steady state right after the vote at time t. We show in the appendix that it does.

Proof. See appendix A.3.2 □

The young worker is the median voter, with preferences in between those of all the capitalist (who are a minority) and the old workers. We have the following proposition,

Proposition 4.4.3. *The unique political equilibrium if voters choose persistent policies is given by* $E_{c^{YW}}$ *and voting strategy profile* $v = \{v_t^{YW} = c^{YW}, v_t^{YC} = 0, v_t^{OW} = \bar{c}_t, v_t^{OC} = 0\}_{t=0}^{\infty}$.

Proof. Strategy profile v follows from sincere voting and lemmas 4.4.3, 4.4.4 and 4.4.7. All preferences are single peaked so that the median voter is decisive. It follows that $c_t = c^{YW}$ in every period and the corresponding economic equilibrium is $E_{c^{YW}}$. □

The unique political equilibrium that arises if voters choose persistent policies features the capital market friction c^{YW}, which may be strictly positive (cf. lemma 4.4.5). The young worker class, which is pivotal in the vote, achieves maximum lifetime utility in this political equilibrium.

Markovian Policies

Markovian equilibria are subgame-perfect equilibria in which the the policy variable is a time-invariant function of the state variable. In the present context, the natural state variable is $k_t := \hat{K}_t^O = K_{t-1}^Y$. We restrict the choice of the decisive voter class to a Markovian policy function,

$$c_t = \mu(k_t) \tag{4.4.6}$$

and assume that $\mu(.)$ is time invariant and differentiable.

The transition function T gives next period's state variable as a function of this period's state variable and current policy. The transition function follows from firm behavior derived in section 4.3.1, evaluated for a Cobb-Douglas production function, so

$$T : [0, \bar{K}] \times [0, \infty) \to [0, \bar{K}]$$

is given by

$$T(k_t, c_t) = \begin{cases} \dfrac{(\theta^Y)^{\frac{1}{1-\alpha}}}{(\theta^Y)^{\frac{1}{1-\alpha}} + (\theta^O)^{\frac{1}{1-\alpha}}} \bar{K} & \text{if} \quad k_t < \underline{K}^O \\ \bar{K} - k_t & \text{if} \quad \underline{K}^O \leq k_t < \bar{K}^O(c_t) \\ \left(\frac{\alpha\theta^Y}{r_t^*(c_t)+c_t}\right)^{\frac{1}{1-\alpha}} \text{ with } r_t^*(c_t) \text{ given by} & \\ \left(\frac{\alpha\theta^Y}{r^*(c_t)+c_t}\right)^{\frac{1}{1-\alpha}} + \left(\frac{\alpha\theta^O}{r^*(c_t)}\right)^{\frac{1}{1-\alpha}} = \bar{K} & \text{if} \quad k_t \geq \bar{K}^O(c_t) \end{cases}$$

Our transition function takes a simple form as it is stationary and does not depend on future policy c_{t+1}. Young workers now solve

$$\begin{aligned} \mu(k_t) &= \arg\max_{c_t} U_t^{YW}(c_t, c_{t+1}; k_t) \\ \text{subject to } c_{t+1} &= \mu(k_{t+1}), \quad c_t \in [0, \infty] \\ k_{t+1} &= T(k_t, c_t) \end{aligned}$$

In words, the function μ must yield a c_t such that utility of the YW class is maximized, taking into account any effect the choice of c_t has on future policy through the state variable k_{t+1}.

To develop some intuition for the solution of this problem, consider trivial Markovian policy functions of the form

$$\mu(k_t) = C$$

Then the optimization problem reduces simply to

$$\max_{c_t} U_t^{YW}(c_t, C)$$

which, by lemma 4.4.4, is uniquely solved for $c_t = 0$. It follows that

$$\mu(k_t) = 0$$

is the solution to the young worker's problem. Consequently, $c_t = 0$ is the outcome of the voting game for any t. Thus we have shown that E_0, the steady state without frictions, is politically supported by the trivial Markovian voting strategy $\mu(k_t) = 0$. A priori, there could be other steady states that are politically supported. But we are able to rule this out in the following. Note that in steady state we have

$$K^Y(c) \geq \bar{K}^O(c)$$

which implies that the transition function T reduces to

$$h(c) = T(k,c) = \left(\frac{\alpha\theta^Y}{r+c}\right)^{\frac{1}{1-\alpha}}$$

a strictly decreasing function of c on $[0,\bar{c}]$. Since we must also have

$$c = \mu(h(c))$$

we see that $\mu(k) = h^{-1}(k)$. But this cannot be a solution to the young worker's optimization problem since

$$\arg\max_{c_t} U_t^{YW}(c_t, c_t)$$

is independent of k, cf. lemma 4.4.5. We conclude that there are no steady states, besides E_0, that can be politically supported using Markovian policy functions. The impossibility to politically support steady states (other than E_0) through Markovian voting strategies results from the fact that the pivotal young workers' optimal choice is independent of the state variable. Hence, the form of c_t imposed by (4.4.6) can only lead to trivial solutions.

4.4.3 Illustration: CRRA Utility

As an illustration, we solve for the political equilibrium under persistent policy voting with CRRA utility. Let the felicity function u be given by

$$u_g(w) := \begin{cases} \frac{w^{1-g}}{1-g} & \text{for} \quad g > 0, g \neq 1 \\ \ln w & \text{for} \quad g = 1 \end{cases}$$

We derive the policy outcome if voters choose persistent policies (cf. also proposition 4.4.3)

Lemma 4.4.8. *If voters choose persistent policies, the unique outcome c^{YW} of the voting game in every period is given by*

$$c^{YW} = \begin{cases} 0 & \text{if} \quad \delta A^{\frac{-g}{1+g(\gamma-1)}} \leq 1 \\ \left(\delta A^{\frac{-g}{1+g(\gamma-1)}} - 1\right) r & \text{if} \quad 1 < \delta A^{\frac{-g}{1+g(\gamma-1)}} < \frac{\theta^Y}{\theta^O} \\ \bar{c} & \text{if} \quad \delta A^{\frac{-g}{1+g(\gamma-1)}} \geq \frac{\theta^Y}{\theta^O} \end{cases}$$

with $\gamma := \frac{1}{1-\alpha}$ *and* $A := \left(\frac{\theta^O}{\theta^Y}\right)^{\gamma}$.

Proof. Taking the derivative of U^{YW} with respect to c we have shown (cf. lemma 4.4.5) that $\frac{dU^{YW}}{dc} \geq 0$ if and only if

$$u'\left(w^Y\right)(r+c) - \delta u'\left(w^O\right) r \leq 0 \tag{4.4.7}$$

With CRRA utility this rewrites as

$$\beta^{-g}\alpha^{-g(\gamma-1)}\left(\theta^Y\right)^{-g\gamma}\left[(r+c)^{1+g(\gamma-1)} - \delta A^{-g} r^{1+g(\gamma-1)}\right] \leq 0$$

so that lifetime utility of the young is increasing in the reallocation cost iff

$$\left(\frac{r+c}{r}\right) \leq \delta A^{\frac{-g}{1+g(\gamma-1)}} \tag{4.4.8}$$

The utility maximizing policy now follows from this condition. First note that $(\frac{r+c}{r})$ is increasing in c; and that we have $(\frac{r+\bar{c}}{r}) = \frac{\theta^Y}{\theta^O}$, so that

$$1 \leq \frac{r+c}{r} \leq \frac{\theta^Y}{\theta^O}$$

for $c\varepsilon[0,\bar{c}]$. We see that if

$$\delta A^{\frac{-g}{1+g(\gamma-1)}} \leq 1$$

then U^{YW} is decreasing in c so that $c^{YW} = 0$. Likewise if

$$\delta A^{\frac{-g}{1+g(\gamma-1)}} \geq \frac{\theta^Y}{\theta^O}$$

then U^{YW} is increasing in c so that $c^{YW} = \bar{c}$ Finally if

$$1 < \delta A^{\frac{-g}{1+g(\gamma-1)}} < \frac{\theta^Y}{\theta^O}$$

then $\frac{dU^{YW}}{dc}$ switches sign on $[0, \bar{c}]$ and c^{YW} is given by

$$c^{YW} = \left(\delta A^{\frac{-g}{1+g(\gamma-1)}} - 1\right) r(c^{YW})$$

□

4.5 Conclusion

In this chapter we model political support for distortionary capital market frictions. We show that workers in democracies may successfully oppose the reallocation of capital to newer sectors, as this affects their labor rents.

Besides class difference, we identify age difference as a source of political conflict. The political conflict exists as long as capital and labor are complementary factors of production, as long as human capital is less mobile than physical capital, and as long as human capital risk cannot be fully insured.

We identify young workers as the decisive voter class–under the plausible assumption that capitalists are a minority. Young workers are decisive because their preferences are less polarized than preferences of other groups in society. Young workers are hurt by capital market frictions in the short term, but may still favor them to smooth consumption over their lifetime. Young workers prefer a higher friction (i) if technology grows at a faster pace,(ii) if they place more weight on the future, and (iii) if they are more risk averse; the result holds as long as young workers expect the capital market friction to persist.

A special case is when capitalists form a majority; in this case, our model predicts that no capital market frictions arise. Broad capital market participation is found in some democracies, in particular those with funded pension schemes. It would be interesting to see if capital reallocation is less restricted in democracies with fully funded pension systems.

The voting process is critical for the outcome of the voting game: the political equilibrium depends on the ability of a current majority to establish a persistent policy. When policies may be overturned in each period, the model features multiple equilibria. By contrast, the equilibrium prediction is the unique outcome favored by the young worker if the outcome of the vote is irreversible.

Our main contribution is to explain why unrestricted capital mobility may be opposed in democracies as a result of the wealth and age distribution in a country's population. We identify young workers as the decisive class in society. What makes them pivotal is not their number–they are a minority just as every other voter class–but the fact that their preferences are the least extreme.

A crucial assumption in our model is that agents cannot save, so a capital market friction is necessary for self insurance. If we reinterpret the drop in wages as a decline in employment, then our findings hold even if households can store their income.[12] The capital market friction then becomes an unemployment insurance that young workers take out in order to have a quiet life.

References

ACEMOGLU, D. (2003): "Why not a political Coase theorem? Social conflict, commitment, and politics," *Journal of Comparative Economics*, 31, 620–652.

AZARIADIS, C., AND V. GALASSO (2002): "Fiscal Constitutions," *Journal of Economic Theory*, 103, 255–281.

BAGEHOT, W. (1873): *Lombard Street: A Description of the Money Market*. King & Co, London.

BLACK, D. (1948): "On the Rationale of Group Decision-making," *Journal of Political Economy*, 56(1), 23–34.

[12] A simple way to introduce unemployment in our model is through sticky wages.

HASSLER, J., J. V. R. MORA, K. STORESLETTEN, AND F. ZILIBOTTI (2003): "The Survival of the Welfare State," *The American Economic Review*, 93(1), 87–112.

KLEIN, P., P. KRUSELL, AND J.-V. RÍOS-RULL (2008): "Time-Consistent Public Policy," *Review of Economic Studies*, 75(3), 789–808.

KRUSELL, P., AND J.-V. RÍOS-RULL (1996): "Vested Interests in a Positive Theory of Stagnation and Growth," *Review of Economic Studies*, 63, 301–329.

PAGANO, M., AND P. F. VOLPIN (2005): "The Political Economy of Corporate Governance," *American Economic Review*, 95(4), 1005–1030.

PEROTTI, E. C., AND E.-L. VON THADDEN (2006): "The Political Economy of Corporate Control and Labor Rents," *Journal of Political Economy*, 114(1), 145–174.

SAINT-PAUL, G. (2002): "The Political Economy of Employment protection," *Journal of Political Economy*, 110(3), 672–704.

SJOBLOM, K. (1985): "Voting for social security," *Public Choice*, 45, 225–240.

TABELLINI, G. (1991): "The Politics of Intergenerational Redistribution," *The Journal of Political Economy*, 99(2), 335–357.

WURGLER (2000): "Financial Markets and the Allocation of capital," *Journal of Financial Economics*, 58, 187–214.

Appendix A

Appendix: Proofs

This appendix contains the proofs of Chapter 2 (*Incentive-Compatible Sovereign Debt*), Chapter 3 (*Collective Pension Funds*), and Chapter 4 (*Sand in the Wheels of Capitalism*).

A.1 Proofs of Chapter 2

Proof of Proposition 2.3.3:

Proof. Let (O_1, I_d) be an optimal contract and let D be the constant value of O_1 when $I_d(y) = 0$. Consider a new contract $(\tilde{O}_1, \tilde{I}_d)$ given by

$$\tilde{I}_d(y) = \begin{cases} 0 & \text{if} \quad \tilde{D} \leq \min\{\gamma y + B, y\} \\ 1 & \text{if} \quad \tilde{D} > \min\{\gamma y + B, y\} \end{cases}$$

and

$$\tilde{O}_1(y) = \begin{cases} \tilde{D} & \text{if} \quad \tilde{I}_d(y) = 0 \\ \gamma y & \text{if} \quad \tilde{I}_d(y) = 1 \end{cases}$$

and suppose first that $\tilde{D} = D$. If $\tilde{I}_d(y) = I_d(y)$, then the construction of $\tilde{O}_1$ implies that $\tilde{O}_1(y) \geq O_1(y)$. If $\tilde{I}_d(y) < I_d(y)$, i.e. if $\tilde{I}_d(y) = 0$ and $I_d(y) = 1$, then it follows from proposition 2.3.2 that

$$O_1(y) \leq D \leq \tilde{O}_1(y)$$

Furthermore, we can rule out $\tilde{I_d}(y) > I_d(y)$. To see this, suppose that y is such that $\tilde{I_d}(y) = 1$ and $I_d(y) = 0$. Then we know that $O_1(y) = D$, but this cannot be, as we also know that $\gamma y + B < D$ from $\tilde{I_d}(y) = 1$, and an optimal contract must be repudiation proof. This proves that $\tilde{O}_1(y) \geq O_1(y)$ if $\tilde{D} = D$.

Now, one can choose $\tilde{D} \leq D$ such that the investor participation constraint is still satisfied. By construction, the resulting contract $(\tilde{O}_1, \tilde{I_d})$ satisfies truthful revelation and is repudiation-proof–like any sovereign debt contract. As $\tilde{I_d}(y) \leq I_d(y)$, it must be optimal.

Since both (O_1, I_d) and $(\tilde{O}_1, \tilde{I_d})$ are optimal contracts, we have

$$E(I_d - \tilde{I_d})B = 0$$

Consider the state observation function $I_d(y)$. For all states $y \in \left[0, \frac{D-B}{\gamma}\right)$ we must have $I_d(y) = 1$, since $I_d(y) = 0$ would mean that $O_1(y) = D$ which contradicts repudiation-proofness. There may be more states for which $I_d(y) = 1$, as we only know that $I_d(y) \geq \tilde{I_d}(y)$. Let T_2 denote the set of those states, so $T_2 = \left\{ y \geq \frac{D-B}{\gamma} \middle| I_d(y) = 1 \right\}$. We see that

$$EI_d = \int_0^{\frac{D-B}{\gamma}} 1 f(y) dy + \int_{T_2} 1 f(y) dy$$

furthermore we have

$$E\tilde{I_d} = \int_0^{\frac{\tilde{D}-B}{\gamma}} 1 f(y) dy$$

Now since $\tilde{D} \leq D$ and $B > 0$, it follows that (i) $D = \tilde{D}$ and (ii) T_2 has probability mass zero; hence we see that $I_d = \tilde{I_d}$ almost surely. It follows that, as $EO_1 = E\tilde{O}_1$, we must also have that $O_1 = \tilde{O}_1$ almost surely, and I conclude that the optimal contract is a sovereign debt contract. $\square$

A.2 Proofs of Chapter 3

A.2.1 Proof of Proposition 3.2.1

The dynamic investment problem to solve is[1]

$$\max_{\alpha_t,\alpha_{t+1}} Eu(\tilde{b})$$

such that

$$\begin{aligned} w_t &= 0 \\ w_{t+1} &= \alpha_t(1+\tilde{r}_t)+(y-\alpha_t) \\ w_{t+2} &= \alpha_{t+1}(1+\tilde{r}_{t+1})+(w_{t+1}+y-\alpha_{t+1}) \\ \tilde{b} &= w_{t+2} \end{aligned}$$

First, we rewrite this problem in recursive form

$$v_t(w_t) = \max_{\alpha_t} Ev_{t+1}\left((w_t+y)+\alpha\tilde{r}\right) \tag{A.2.1}$$

where v_t, the remaining-value function, is a function of the financial reserve, w_t. We know that $v_{t+2}(w) = u(w) = \frac{w^{1-\phi}}{1-\phi}$, as individuals consume their financial reserve in retirement. Note that $v_t(0)$ is the expected utility in retirement of a young individual at time t who invests optimally throughout his life. Optimal investment, α_t, is a function of the single state variable, w_t.

We consider the trial solution function $v_{t+1}(w_{t+1}) = \gamma_{t+1}\frac{(w_{t+1}+h_{t+1})^{1-\phi}}{1-\phi}$, where $\gamma_{t+1} > 0$ is a scalar and $h_{t+1}$is the human capital reserve of an individual. Our trial solution implies that

$$\begin{aligned} v_{t+1}\left((w_t+y)+\alpha_t\tilde{r}\right) &= \gamma_{t+1}\frac{\left((w_t+y)+\alpha_t\tilde{r}+h_{t+1}\right)^{1-\phi}}{1-\phi} \\ &= \gamma_{t+1}\frac{\left(w_t+h_t+\alpha_t\tilde{r}\right)^{1-\phi}}{1-\phi} \end{aligned}$$

[1]We drop the superscripts to save on notation.

so that the first-order condition reads as

$$\gamma_{t+1} E\tilde{r}\left(w_t + h_t + \alpha_t \tilde{r}\right)^{-\phi} = 0$$

Solving for optimal investment in the risky asset yields

$$\alpha_t\left(w_t\right) = a^*\left(w_t + h_t\right)$$

where

$$a^* := \frac{\left(\frac{-r^l(1-p)}{r^h p}\right)^{\frac{1}{\phi}} - 1}{r^l - r^h\left(\frac{-r^l(1-p)}{r^h p}\right)^{\frac{1}{\phi}}} \tag{A.2.2}$$

Finally, it follows from $(A.2.1)$ that

$$\begin{aligned} v_t\left(w_t\right) &= \gamma_{t+1} E\frac{\left(w_t + h_t + a^*\left(w_t + h_t\right)\tilde{r}\right)^{1-\phi}}{1-\phi} \\ &= \delta\gamma_{t+1}\frac{\left(w_t + h_t\right)^{1-\phi}}{1-\phi} \end{aligned}$$

where

$$\delta := E\left(1 + a^*\tilde{r}\right)^{1-\phi}$$

so that our trial solution is correct with $\gamma_t = \delta\gamma_{t+1}$. Note that

$$\delta \approx 1 + (1-\phi)a^*\mu > 1$$

where we've used a first-order approximation. We conclude that, conditional on an optimal investment strategy, expected lifetime utility of the young at time t is

$$\begin{aligned} v_t(0) &= \delta^2 u(h_t) \\ &= \delta^2 u(2y) \end{aligned}$$

□

A.2.2 Proof of Proposition 3.2.2

Incomplete markets give rise to additional constraints, $\alpha_t = 0$ and $\alpha_{t+1} \leq w_{t+1}$. To obtain a lower bound for the welfare loss due to credit constraints, we assume that the middle-aged do not face constraints, only the young do. Then the new investment problem is

$$\max_{\alpha_{t+1}} Eu\left(\tilde{b}\right)$$

such that

$$\begin{aligned} w_t &= 0 \\ w_{t+1} &= y \\ w_{t+2} &= \alpha_{t+1}\left(1+\tilde{r}_{t+1}\right)+\left(w_{t+1}+y-\alpha_{t+1}\right) \\ \tilde{b} &= w_{t+2} \end{aligned}$$

It is easy to see that the middle aged will invest $a^*(2y)$ so that

$$w_{t+2} = 2y\left(1+a^*\tilde{r}_{t+1}\right)$$

and

$$\begin{aligned} Eu\left(\tilde{b}\right) &= Eu\left(2y\left(1+a^*\tilde{r}_{t+1}\right)\right) \\ &= E\left(1+a^*\tilde{r}_{t+1}\right)^{1-\phi} u(2y) \\ &= \delta \quad u(2y) \end{aligned}$$

With proposition 4.3.2 we see that the lower bound for the welfare loss due to credit constraints is

$$\left(\delta^2-\delta\right)u(2y)$$

□

A.3 Proofs of Chapter 4

A.3.1 Cobb-Douglas Production

We derive the steady state equilibria for a Cobb-Douglas production economy, where F is given by

$$F(K,L) = K^{\alpha} L^{\beta} \tag{A.3.1}$$

with $0 < \alpha + \beta < 1$. Steady-state capital allocations for $c \in [0, \bar{c}]$ are

$$K^{Y*} = \left(\frac{\alpha \theta^{Y}}{r^{*}(c) + c} \right)^{\frac{1}{1-\alpha}}; \quad K^{O*} = \left(\frac{\alpha \theta^{O}}{r^{*}(c)} \right)^{\frac{1}{1-\alpha}} \tag{A.3.2}$$

with $r^{*}(c)$ given by capital market clearing condition

$$\left(\frac{\alpha \theta^{Y}}{r^{*}(c) + c} \right)^{\frac{1}{1-\alpha}} + \left(\frac{\alpha \theta^{O}}{r^{*}(c)} \right)^{\frac{1}{1-\alpha}} = \bar{K}$$

The boundary value $\bar{c}$ follows from (4.3.21) and is given by

$$\bar{c} = \alpha \frac{\theta^{Y} - \theta^{O}}{\left(\frac{1}{2}\bar{K}\right)^{1-\alpha}}$$

For $c > \bar{c}$ we have steady state capital allocations

$$K^{Y*} = \left(\frac{\alpha \theta^{Y}}{r^{*}(c) + c} \right)^{\frac{1}{1-\alpha}};, \quad K^{O*} = \frac{1}{2}\bar{K}$$

where capital market clearing condition

$$\left(\frac{\alpha \theta^{Y}}{r^{*}(c) + c} \right)^{\frac{1}{1-\alpha}} = \frac{1}{2}\bar{K}$$

allows us to obtain the equilibrium interest rate explicitly,

$$r^{*}(c) = \frac{\alpha \theta^{Y}}{\left(\frac{1}{2}\bar{K}\right)^{1-\alpha}} - c$$

Next we determine what steady state equilibria can be politically supported.

A.3.2 Proof of lemma 4.4.2

Consider the equilibrium interest rate in period t. For $c_t < \bar{c}_t$, we have $\hat{K}_t^O \geq \bar{K}_t^O$ so that old dirms scale down and the interest rate is given by

$$\left(\frac{\alpha\theta^Y}{r_t + c_t}\right)^{\frac{1}{1-\alpha}} + \left(\frac{\alpha\theta^O}{r_t}\right)^{\frac{1}{1-\alpha}} = \bar{K}$$

Implicit differentation yields

$$-\frac{1}{1-\alpha}\frac{K_t^Y}{r_t + c_t}\left(\frac{dr_t}{dc_t} + 1\right) - \frac{1}{1-\alpha}\frac{K_t^O}{r_t}\frac{dr_t}{dc_t} = 0 \tag{A.3.3}$$

which rewrites as

$$\frac{dr_t}{dc_t} = -\frac{K_t^Y}{K_t^Y + K_t^O(\frac{r_t + c_t}{r_t})}$$

so that $\frac{dr_t}{dc_t}\varepsilon(-1,0)$. For $c_t \geq \bar{c}_t$, we have $K_t^O = \hat{K}_t^O$ and the interest rate is given by

$$\left(\frac{\alpha\theta^Y}{r_t + c_t}\right)^{\frac{1}{1-\alpha}} = \bar{K} - \hat{K}_t^O$$

$$r_t = \frac{\alpha\theta^Y}{\left(\bar{K} - \hat{K}_t^O\right)^{1-\alpha}} - c_t$$

and we see that $\frac{dr_t}{dc_t} = -1$. Now for (i), capital in the Y-sector is given by

$$K_t^Y = \left(\frac{\alpha\theta^Y}{r_t + c_t}\right)^{\frac{1}{1-\alpha}}$$

Taking the derivative with respect to c_t gives

$$\frac{dK_t^Y}{dc_t} = \frac{-1}{1-\alpha}\frac{K_t^Y}{r_t + c_t}\left(\frac{dr_t}{dc_t} + 1\right)$$

so that $\frac{dK_t^Y}{dc_t} < 0$ for $c_t \in [0, \bar{c}_t)$ and $\frac{dK_t^Y}{dc_t} = 0$ for $c \geq \bar{c}_t$. Wages in the Y-sector are

$$w_t^Y = \theta^Y \beta \left(K_t^Y\right)^\alpha$$

Taking the derivative with respect to c_t yields

$$\frac{dw_t^Y}{dc_t} = -\frac{\beta}{1-\alpha} K_t^Y \left(\frac{dr_t}{dc_t} + 1\right) \tag{A.3.4}$$

so that $\frac{dw_t^Y}{dc_t} < 0$ for $c_t \in [0, \bar{c}_t]$; $\frac{dw_t^Y}{dc_t} = 0$ for $c_t > \bar{c}_t$. For (ii), capital in the O-sector is given by

$$K_t^O = \left(\frac{\alpha\theta^O}{r_t}\right)^{\frac{1}{1-\alpha}}$$

for $c_t < \bar{c}_t$; and

$$K_t^O = \hat{K}^O$$

for $c_t \geq \bar{c}_t$. Hence $\frac{dK_t^O}{dc_t} > 0$ for $c_t \in [0, \bar{c}_t)$ and $\frac{dK_t^O}{dc_t} = 0$ for $c_t \geq \bar{c}_t$. Wages in the O-sector are

$$w_t^O = \theta^O \beta \left(K_t^O\right)^\alpha$$

Taking the derivative with respect to c_t gives

$$\frac{dw_t^O}{dc_t} = -\frac{\beta}{1-\alpha} K_t^O \frac{dr_t}{dc_t} \tag{A.3.5}$$

so that $\frac{dw^O}{dc_t} > 0$ for $c_t \varepsilon [0, \bar{c}_t)$ and $\frac{dw_t^O}{dc_t} = 0$ for $c_t \geq \bar{c}_t$. For (iii), first consider profits. Let π_t^Y and π_t^O denote profits in the Y- and O-sector respectively. Then

$$\pi_t^Y = \theta^Y \left(K_t^Y\right)^\alpha - (r_t + c_t) K_t^Y - w_t^Y$$

and

$$\pi_t^O = (\theta^O)(K_t^O)^\alpha - r_t K_t^O - w_t^O$$

We take the derivative of π_t^Y with respect to c_t and obtain

$$\frac{d\pi_t^Y}{dc_t} = \frac{\alpha+\beta-1}{1-\alpha} \left(\frac{dr_t}{dc_t} + 1\right) K_t^Y$$

so that $\frac{d\pi_t^Y}{dc_t} < 0$ for $c_t \in [0, \bar{c}_t)$ and $\frac{d\pi_t^Y}{dc_t} = 0$ for $c \geq \bar{c}_t$. Similarly, for O-sector profits, we get

$$\frac{d\pi_t^O}{dc_t} = \frac{\alpha + \beta - 1}{1 - \alpha} \frac{dr_t}{dc_t} K_t^O$$

so that $\frac{d\pi_t^O}{dc_t} > 0$ for $c_t \in [0, \bar{c}_t)$ and $\frac{d\pi_t^O}{dc_t} > 0$ for $c_t \geq \bar{c}_t$. Turning to total profits, $\Pi_t = \pi_t^Y + \pi_t^O$, we have

$$\frac{d\Pi_t}{dc_t} = \frac{\alpha + \beta - 1}{1 - \alpha} \left[\left(\frac{dr_t}{dc_t} + 1 \right) K_t^Y + \frac{dr_t}{dc_t} K_t^O \right] \tag{A.3.6}$$

and we see that Π_t is decreasing in c_t iff

$$\frac{dr_t}{dc_t} \bar{K} + K_t^Y \geq 0$$

Recall that for $c_t \in [0, \bar{c}_t)$ we have

$$\frac{dr_t}{dc_t} = -\frac{K_t^Y}{K_t^Y + K_t^O (\frac{r_t + c_t}{r_t})}$$

so that $-\frac{dr_t}{dc_t} \bar{K} \leq K_t^Y$ and total profits are nonincreasing in c_t. For $c_t \geq \bar{c}_t$, we have $\frac{dr_t}{dc_t} = -1$ so that total profits are increasing in c_t. This result is due to the fact that r_t declines in c_t while the allocation of capital does not change in this range of capital market frictions. Hence the cost of capital goes down for O-firms, stays the same for Y-firms, as production in both sectors remains the same. While profits may increase in c_t, capital income cannot. Recall that capital income is given by

$$s_t = \frac{r_t \bar{K} + \Pi_t}{\eta}$$

For $c_t \in [0, \bar{c}_t)$, we have $\frac{d\Pi_t}{dc_t} \leq 0$ so that $\frac{ds_t}{dc_t} \leq 0$. For $c_t \geq \bar{c}_t$ we have

$$\frac{ds_t}{dc_t} = \frac{1}{\eta} \left(-\bar{K} + \frac{1 - \alpha - \beta}{1 - \alpha} K_t^O \right)$$

so that $\frac{ds_t}{dc_t} < 0$. Now, old capitalists income is given by

$$w_t^O + s_t$$

Let $c_t \in [0, \bar{c}_t)$, then we have

$$\begin{aligned}\frac{d(w_t^O + s_t)}{dc_t} &= -\frac{\beta}{1-\alpha}K_t^O\frac{dr_t}{dc_t} + \frac{1}{\eta}\left(\frac{\beta}{1-\alpha}\frac{dr_t}{dc_t}\bar{K} + \frac{\alpha+\beta-1}{1-\alpha}K_t^Y\right)\\ &< -\frac{\beta}{1-\alpha}K_t^O\frac{dr_t}{dc_t} + \frac{\beta}{1-\alpha}\frac{dr_t}{dc_t}\bar{K}\\ &< 0\end{aligned}$$

Next, for $c_t \geq \bar{c}_t$, $\frac{dw_t^O}{dc_t} = 0$ and $\frac{ds_t}{dc_t} < 0$ which shows part (iii).

A.3.3 Proof of Lemma 4.4.7

To prove lemma 4.4.7,we first prove two auxiliary lemmas that give the economic equilibrium in periods t and $t+1$. Since we consider out-of steady-state dynamics we assume that the equilibrium at time $t-1$ is given by steady state allocations for some arbitrary $c \in [0, \bar{c}]$. The first lemma describes the economic equilibrium after a downward change in policy:

Let the economic equilibrium at time $t-1$ be given by steady state values for some $c \leq \bar{c}$. Consider a downward change in policy $c_t = c_{t+1} \leq c$, then

$$K_t^Y = \left(\frac{\alpha\theta^Y}{r_t^* + c_t}\right)^{\frac{1}{1-\alpha}}$$

and

$$K_{t+1}^O = \left(\frac{\alpha\theta^O}{r_t^*}\right)^{\frac{1}{1-\alpha}}$$

with r_t^* given by $K_t^Y + K_t^O = \bar{K}$, furthermore $K_{t+1}^Y = K_t^Y$ and $K_{t+1}^O = K_t^O$.

Since $c_t \leq c$, we have $\bar{K}^O(c) \geq \bar{K}^O(c_t)$, where $\bar{K}^O$ is the equilibrium cut-off value function defined after (4.3.14). It follows that

$$\hat{K}_t^O = K_{t-1}^Y \geq \bar{K}^O(c_t)$$

so that K_t^Y and K_t^O are as posed. Note that because $c_t \leq c$, we have $K_t^O < K_{t-1}^O$ by lemma 4.4.1 and hence $K_t^Y > K_{t-1}^Y$ by market clearing . We must verify that the same allocation obtains in period $t+1$. Moving forward one period we have

$$\hat{K}_{t+1}^O = K_t^Y > \bar{K}^O(c_{t+1})$$

so that again old firms scale down, the same interest rate obtains (i.e. $r_t^* = r_{t+1}^*$), and allocations are as posed.

Lemma A.3.3 shows that a downward change in policy results in steady state values that correspond to a lower reallocation cost. Informally, we can say that changing policy downward moves the economy to a new steady state corresponding to the new value of the friction c_t. The same need not be true for an upward policy change as the next lemma shows.

Let the economic equilibrium at time $t-1$ be given by steady state values for some $c \leq \bar{c}$. Consider an upward change in policy $c_t = c_{t+1} > c$, then

(i) if $c_t \leq \bar{c}$ we have

$$K_t^Y = \left(\frac{\alpha\theta^Y}{r_t^* + c_t}\right)^{\frac{1}{1-\alpha}} \quad ; \quad K_t^O = \left(\frac{\alpha\theta^O}{r_t^*}\right)^{\frac{1}{1-\alpha}}$$

with r_t^* given by $K_t^Y + K_t^O = \bar{K}$, furthermore $K_{t+1}^Y = K_t^Y$ and $K_{t+1}^O = K_t^O$;

(ii) if $\bar{c} < c_t \leq \bar{c}_t$ we have

$$K_t^Y = \left(\frac{\alpha\theta^Y}{r_t^* + c_t}\right)^{\frac{1}{1-\alpha}} \quad ; \quad K_t^O = \left(\frac{\alpha\theta^O}{r_t^*}\right)^{\frac{1}{1-\alpha}}$$

with r_t^* given by $K_t^Y + K_t^O = \bar{K}$, furthermore $K_{t+1}^Y = K_t^O$ and $K_{t+1}^O = K_t^Y$; and

(iii) if $c_t > \bar{c}_t$ we have

$$K_t^Y = \bar{K} - \hat{K}_t^O \quad ; \quad K_t^O = \hat{K}_t^O$$

furthermore $K_{t+1}^O = K_t^Y$ and $K_{t+1}^Y = K_t^O$.

For (i): suppose $c < c_t \leq \bar{c}$. Then also $c_t \leq \bar{c}_t$, where $\bar{c}_t$ is given by (4.4.2). Hence we have

$$K_{t-1}^Y = \hat{K}_t^O \geq \bar{K}^O(c_t)$$

by the monotonicity of $\bar{K}^O$. It follows that old firms scale down and K_t^Y and K_t^O are as posed. Consider period $t+1$. Since $c_t = c_{t+1} \leq \bar{c}$, we have

$$K_t^Y(c_t) = \hat{K}_{t+1}^O \geq \bar{K}^O(c_{t+1})$$

by the definition of $\bar{c}$, given by (4.3.21), and the monotonicity in of K^Y and $\bar{K}^O$ in c_t. Hence K_{t+1}^Y and K_{t+1}^O are as posed.

For (ii): suppose $\bar{c} < c_t \leq \bar{c}_t$. Then we have

$$\hat{K}_t^O \geq \bar{K}^O(c_t)$$

so that old firms scale down and allocations are as posed. Moving forward one period it follows from $\bar{c} < c_t = c_{t+1}$ that

$$K_t^Y(c_t) < \bar{K}^O(c_{t+1})$$

Hence old firms do not adjust capital and $K_{t+1}^O = K_t^Y$. By market clearing then $K_{t+1}^Y = K_t^O$.

For (iii): suppose $\bar{c}_t < c_t$, then $\hat{K}_t^O < \bar{K}^O(c_t)$ so that old firms do not adjsut capital. It follows that $K_t^O = \hat{K}_t^O$ and, by market clearing, $K_t^Y = \bar{K} - \hat{K}^O$. In period $t+1$, since $\bar{c} < c_t = c_{t+1}$ we have $\hat{K}_{t+1}^O < \bar{K}^O(c_{t+1})$ and so $K_{t+1}^O = K_t^Y$ and $K_{t+1}^Y = K_t^O$.

Lemma A.3.3 shows that a small upward change in policy ($c_t \leq \bar{c}$) results in steady state allocations that correspond to a higher friction. Now, consider lifetime utility U_t^{YW} of the young worker at time t. Lemma A.3.3 also implies that young workers will not vote for a higher friction than $\bar{c}$. To see this note that K_t^Y and K_{t+1}^O are strictly decreasing in c_t for $\bar{c} < c_t \leq \bar{c}_t$; they are constant in c_t for $c_t > \bar{c}_t$. Hence young workers strictly prefer $c_t = \bar{c}$ over any $c_t > \bar{c}$.

With the auxiliary lemmas, we can now proof lemma 4.4.7.The choice of a persistent friction, $c_t = c_{t+1} \in [0, \infty]$, uniquely determines the equilibrium interest rates, r_t^* and r_{t+1}^*, and hence the economic equilibrium at time t and $t+1$ (cf lemma 4.3.1). Let c_t^{YW} denote the preferred policy of the YW class. We have shown that $c_t^{YW} \varepsilon [0, \bar{c}]$ and that, if this policy is set, the economy attains steady state values corresponding to the friction c_t^{YW}. It follows that $c_t^{YW} = c^{YW}$, where

c^{YW} is given by lemma 4.4.5. By sincere voting we have $a_t^{YW} = c^{YW}$, which concludes the proof.

Summaries

English Summary

This dissertation investigates capital market frictions across three themes, presented in three stand-alone chapters (Chapters 2-4). Chapter 2 deals with sovereign debt. Recent experience in the EU shows that it can be complex to enforce the repayment promises of states. Furthermore, governments are better informed about their repayment capacity than creditors are. In this dissertation, I argue that enforcement and information problems explain why states issue debt contracts that frequently lead to debt crises. Such contracts are optimal because they save on costly audits by creditors.

Chapter 3 deals with collective pension funds. It is often argued that pension funds can enhance the welfare of their participants. In this dissertation, I highlight one rationale for pension funds based on credit constraints. I then argue that pension funds' actual ability to increase welfare may be limited due to an agency problem.

Chapter 4 deals with political intervention in capital markets. Financial liberalization and expanded access to capital are historically seen as signs of greater freedom. Yet many democratic states choose to restrain the resource allocation called for by free capital markets. In this dissertation, I argue that democracies may choose to introduce restraints on free capital markets—thereby favouring income stability over economic growth—depending on demographical context, the distribution of wealth, and the rate of technological progress.

Nederlandse Samenvatting

Dit proefschrift onderzoekt kapitaalmarktfricties in drie op zichzelf staande hoofdstukken met verschillende thema's (hoofdstukken 2-4). Hoofdstuk 2 gaat in op overheidsfinanciering en overheidsschuld. Recente ervaring in de EU laat zien dat betalingsbeloftes van staten dikwijls niet af te dwingen zijn. Overheden zijn daarnaast beter geïnformeerd over hun terugbetalingscapaciteit dan hun schuldeisers. In dit proefschrift betoog ik dat handhavings- en informatieproblemen verklaren waarom staten zich financieren met contracten die regelmatig leiden tot heronderhandelingen en schuldencrises. Dergelijke contracten zijn toch te zien als optimaal, omdat ze besparen op dure audits door de schuldeisers.

Hoofdstuk 3 gaat over collectieve pensioenfondsen. Er wordt vaak beweerd dat collectieve pensioenfondsen het welzijn van hun deelnemers kunnen verhogen. In dit proefschrift richt ik me op één mogelijkheid daarvoor: collectieve pensioenfondsen, zoals die onder andere in Nederland te vinden zijn, kunnen het welzijn van deelnemers verhogen door de kredietbeperkingen van jonge deelnemers op te heffen. Aansluitend betoog ik dat een agency probleem de daadwerkelijke welzijnsverhogende bijdrage van pensioenfondsen in de praktijk kan beperken.

Hoofdstuk 4 gaat over politieke interventie in kapitaalmarkten. Financiële liberalisering en uitbreiding van de toegang tot kapitaal zijn historisch gezien tekenen van een grotere vrijheid. Toch zijn er veel democratische staten die ervoor kiezen om de vrije kapitaalmarkt aan banden te leggen. In dit proefschrift betoog ik dat democratieën kunnen kiezen om beperkingen in te voeren op de vrije kapitaalmarkt–waarmee ze feitelijk inkomensnivellering verkiezen boven economische groei–afhankelijk van de demografische context, kapitaalmarktparticipatie, en de snelheid van technologische vooruitgang.

Deutsche Zusammenfassung

Die vorliegende Arbeit untersucht Kapitalmarkt Friktionen in drei, in sich geschlossene, Kapitel mit verschiedenen Themen (Kapitel 2-4). Kapitel 2 beschäftigt sich mit der Staatsverschuldung. Jüngste Erfahrungen in der EU zeigen, dass die Rückzahlungsversprechen von Staaten oft nicht durchsetzbar sind. Zudem sind

Regierungen besser über ihre Rückzahlungsfähigkeit informiert als ihre Gläubiger. In Kapitel 2 dieser Arbeit argumentiere ich dass die beschriebene Informations- und Durchsetzungsprobleme erklären können warum Staaten sich über Verträge finanzieren die regelmäßig neu verhandelt werden müssen und zu Schuldenkrisen führen. Solche Verträge können dennoch optimal sein, weil sie auf kostspielige Prüfungen durch die Gläubiger sparen.

Kapitel 3 beschäftigt sich mit Rentenkassen. Es wird oft argumentiert, dass Rentenkassen mit verpflichtende Teilnahme–so wie es zum Beispiel in den Niederlanden üblich ist–die Wohlfahrt ihrer Teilnehmer steigern können. In dieser Dissertation, untersuche ich eine mögliche Begründung für diese Behauptung; eine Begründung die sich basiert auf Kreditbeschränkungen von jüngeren Teilnehmerkohorten. Anschließend argumentiere ich dass ein Agentur Problem den Wohlfahrt förderlichen Beitrag der Rentenkassen in der Praxis begrenzt.

Kapitel 4 beschäftigt sich mit politischem Eingreifen auf den Kapitalmärkten. Liberalisierung der Finanzmärkte und einen erweiterten Zugang zu den Kapitalmärkten werden historisch als Anzeichen für eine größere Freiheit gesehen. Dennoch entscheiden viele demokratische Staaten sich um den freien Kapitalverkehr Einhalt zu gebieten. In dieser Dissertation argumentiere ich dass Demokratien sich für Beschränkungen des freien Kapitalverkehrs entscheiden–und so Einkommensstabilität über wirtschaftlichen Wachstum begünstigen–in Abhängigkeit von den demographischen Kontext, von der Verteilung des Reichtums und von der Rate des technischen Fortschritts.

Acknowledgements

In the course of writing this dissertation, I have benefited from the insights, guidance, support, and criticism of many. I wish to express my gratitude here, for lack of a sounder currency to repay them.

Enrico Perotti and Ernst-Ludwig von Thadden supervised my efforts, and I am most indebted to them. I thank Enrico and Elu for their guidance and for allowing me to write this dissertation. The *capital wedge* paper, which we co-authored, gave me a unique opportunity for some on-the-job training. Enrico's creativity and out-of-the-box thinking particularly inspired me. Elu's commitment to research, relentless questioning, and refusal to accept easy answers, taught me to always dig deeper. His enthusiasm, gentle manner, and sheer intelligence equally stood out and I can only hope that some of it has rubbed off on me over the years.

I want to thank the other committee members, Arnoud Boot, Pieter Gautier, Bas Jacobs, and Enrique Schroth, for their valuable feedback and for taking the time to serve as my readers.

I am grateful to Casper de Vries who supervised my master thesis and awakened my enthusiasm for research; and to Roald Ramer who supervised my bachelor thesis in Mathematics and encouraged me to apply to the Tinbergen Institute. I thank Bruno Biais and Martin Hellwig, two inspiring examples whom I was lucky enough to meet and discuss my work with. I also thank Robin de Vilder, my erstwhile math teacher, who has been an instrumental guide throughout my studies.

At the Tinbergen Institute, where I started to read economics, it would have been impossible to survive the notorious M.Phil. program without the help and companionship of Iwan Bos, David Hollanders, Jacob de Hoop, Freek van der Linden, Sandra Phlippen, Bernd Schwaab, Nick Vikander, and Razvan Vlahu.

In Amsterdam, at the Finance Group, I was fortunate to share space and thoughts with, among others, Raj Iyer, Christel Karsten, Jeroen Ligterink, Florian Peters, Enrique Schroth, and Zacharias Zautner. I thank Raj for an exciting summer of research in India; I thank Enrique for teaching me the theatrics of the profession.

After Amsterdam, Mannheim became my second academic home. At the economics department I benefited from interaction—both academic and social—with, among others, Agustin Arias, Petyo Bonev, Daniel Harenberg, Petra Loerke, Dries de Smet, Philipp Zahn, and in particular Andras Niedermayer—who housed me on his couch when I first arrived.

Both in Amsterdam and in Mannheim, I have found that the secretarial offices are the heart of the department. I wish to thank Yolanda Carrasco, and Jolinda Gompel from the University of Amsterdam; and Nuray Mamaç, Anja Mayer, and Sabine Münch from the University of Mannheim for all the visible and invisible things they have done over the years.

My *paranimfen*, Marnix Jansen and Thierry Baudet, a doctor and a lawyer, deserve special thanks for standing by my side during the defense of my dissertation, for their help on the final draft, and for setting an exceptional standard.

I thank my friends who never take me too seriously and who are there when it counts. I cannot mention them all here, but I must thank Klaas Beks, Vica Bogaerts, Elias Fels, Marie-Claire van Hessen, Martijn Hillenius, David Janssen, Wouter Ritmeester, Menno Schröder, and Michel Topper for their continuing friendship and support.

To balance academic life, I run—as do many PhD students. Running is my cure-all therapy and I thank my running mates David Moszkowicz, and Wietze Smid, for the hundreds of kilometers we ran together. Other people who have contributed significantly to my well being include Wessel Kirsten, who kept me running; and Theo Verbeek, who provided invaluable counsel in all matters.

I am deeply grateful to my entire family: to the Bersems, to the Van Overmeires, to Oma and Piet, to the De Vilders, and especially to my parents, Mirjam and Erich, for their unconditional love and support throughout my life and studies. Thank you! I would also like to remember here my brother and buddy Christian, whose passing last year was a severe blow to us all; his presence remains deeply missed.

Finally, I thank Ziva, my love, for giving me space; space to write and space to be.

Kurzlebenslauf

- **2007 - 2012**: Ph.D. student at the Center for Doctoral Studies in Economics, University of Mannheim; and the Finance Group, University of Amsterdam (joint doctorate).
- **2004 - 2007**: M.Sc. in Economics at the Tinbergen Institute, University of Amsterdam.
- **2000 - 2004**: B.Sc. in Mathematics at the University of Amsterdam.