Essays on Information Frictions in Economics

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Radoslaw Paluszynski

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Manuel Amador, Advisor
Timothy J. Kehoe, Co-Advisor

July, 2016

Acknowledgements

I am grateful to my advisors, Manuel Amador and Tim Kehoe, for their continuous support, encouragement and hard lessons throughout my doctoral studies. I am indebted to Motohiro Yogo for his excellent guidance and being the best supervisor I could possibly have asked for at the Minneapolis Fed. I would also like to thank Professors Rodney Smith and Terry Roe for serving on my Preliminary and Final committees, respectively.

This dissertation would never have been created without the unbounded support from the members of my family, in particular my parents, Wanda Paluszyńska and Witold Paluszyński, and my gramma Halina Paluszyńska. I also cannot express enough gratitude to Ruth and Frank Shaw for creating a true home for me here in Minnesota.

I have benefited immensely from the collaboration and friendship with my fellow grad students at the Minnesota Economics Department. I owe a special debt to my coauthor Pei Cheng Yu for always being there for me, and for teaching me a great deal of economics. For countless discussions and many shared adventures I also thank Joao Ayres, Zhifeng Cai, Richard Condor, Daniela Costa, Hyunju Lee, Boyoung Seo and George Stefanidis.

Doctoral studies would be a lot more challenging without the support of an amazing administrative staff. In particular, I am grateful to Kara Kersteter and Wendy Williamson at the Department of Economics for their continuous guidance through all the formalities, and to Caty Bach for her incredible effort and assistance during the job market year.

I would finally like to acknowledge the generous support of the Simler Fellowship and the Hutcheson-Lilly Dissertation Fellowship at the Economics Department. My work partially relied on the computational resources provided by the Minnesota Supercomputing Institute.

Dedication

To the loving memory of my grandfather, Jerzy Paluszyński (1921-2007).

Abstract

This dissertation consists of three chapters. The unifying topic of this collection is the impact of information frictions and lack of commitment on economic outcomes, in particular on prices.

The first two chapters are dedicated to the study of the European sovereign debt crisis of 2008-2014. This episode of global economic importance was marked with a surge in government bond yields on unprecedented scale among developed countries in modern history. The peak of the crisis occurred with a significant lag following the initial shocks to output, even though governments did not undertake the fiscal adjustments necessary to prevent a further increase in default risk. I show that these observations are at odds with the predictions of existing sovereign debt models and propose a new theory that features incomplete, symmetric (in Chapter 1) and asymmetric (in Chapter 2) information about the country's economic outlook. In a calibrated model, the delay arises as a result of the markets' learning process, while the government optimally postpones debt reduction in order not to send a negative signal about the underlying state of the economy.

In the third chapter, written jointly with Pei Cheng Yu, we study optimal pricing in markets characterized by long-term relationship and information asymmetries between the firm and its customers. As an example of such a market, we show that life insurance premiums have displayed a significant degree of rigidity over the past two decades. On average, prices took over 3 years to adjust and the magnitude of these one-time jumps exceeded 10%. This stands in sharp contrast with the dynamics of the corresponding marginal cost which exhibited considerable volatility since 1990 due to the movements in the interest and mortality rates. We build a model with consumer hold-up problem that captures these empirical findings. Price rigidity arises as an optimal response to the relationship-specific investment the consumers need to make before buying. The optimal contract takes the form of a simple cutoff rule: premiums are rigid for all cost realizations smaller than the threshold, and adjustments must be large and are only possible when cost realizations exceed it.

Contents

List of Tables

List of Figures

Chapter 1

Learning about Debt Crises

1.1 Introduction

The debt crisis in Europe of 2010-2012 has put into question a widespread belief about the developed countries' resilience to sovereign default. In retrospect it is important to understand the factors that led to very high debt accumulation before and during the crisis in the Eurozone periphery. Quantitative economic theories applied to that episode should also be able to replicate the path of interest rates on government bonds observed in years 2008-2014.

In particular, two facts about the timing of events during the recent debt crisis pose a challenge for existing models. First, the Eurozone governments did not reduce their foreign liabilities at the outset of the global recession in 2008. On the contrary, debt levels increased for some of the peripheral European countries in the period between the financial crisis and the actual debt crisis of 2010-2012. Such behavior contrasts with the predictions of most quantitative sovereign debt theories.[1] In those models, an endogenous sovereign default is driven by the occurrence of unexpected negative income shocks. Hence, following a sequence of low shocks the government has a strong motive to reduce external debt, both in order to stave off the possibility of default tomorrow, and to secure higher bond prices today.

[1] I am referring to the class of models based on the seminal framework of Eaton and Gersovitz (1981), e.g. Arellano (2008) or Chatterjee and Eyigungor (2012).

Second, markets did not initially express concern about sovereign defaults in Europe, in spite of the aforementioned lack of debt reductions. Government bond spreads[2] increased marginally during the first two years of the recession, while sovereign credit ratings stayed close to the risk-free level. A real stress did not appear in the European bond markets until much later, years 2011-2012, when the peripheral economies one by one experienced a surge in borrowing costs. This delay is puzzling given not only the expansionary fiscal policy of governments at the beginning of the crisis, but also the sheer size of the shocks to output.

To explain these facts, this paper introduces a new quantitative theory of sovereign debt based on a more general specification of the income process and on how markets learn about its realizations over time. I use a data set of historical GDP forecasts to show that market participants did not expect a major depression during the first two years of the crisis. Instead, the low output levels of 2008-2009 were perceived as merely a temporary downturn, another one of the several short-lived recessions in Europe's post-war history. Over time however, by monitoring the subsequent releases of national accounts data, investors realized that the current recession would be more severe than they had anticipated. Indeed, around 2011 we observe that GDP forecasts become much more pessimistic and, as it eventually turns out, more accurate. Once the lenders realized that a recovery was not to be expected any time soon they demanded a much higher compensation for default risk, leading to the sharp spikes in the interest rates observed around 2011-2012.

In order to support the research hypothesis laid out above I build a quantitative model of sovereign debt that features several novel elements. To begin, I introduce a regime-switching income process to better capture the distinct features of the European GDP data that contrast with those of the emerging markets studied by previous models. Then, I assume incomplete (and symmetric) information about the economy's current regime and allow the market participants to learn about it by Bayesian-updating. In a model calibrated to Portuguese data I show that, under this specification, the initial low income realizations do not cause a large increase in interest rates because markets perceive them as temporary

[2] The bond spread is defined as a difference between the interest rates paid by the given country's government bonds and a risk-free asset, in this case I use the German long-term government bonds. The spread is expressed in annual terms.

shocks, not a permanent regime switch. Moreover, the predicted reduction in foreign debt is attenuated as even the government expects the economy to rebound soon. Over time this belief drops however, and the market participants become convinced that the process has switched to a bad regime. As a result, we observe a delayed surge in interest rates combined with an ultimate reduction in government debt. Importantly, these predictions are consistent with the aforementioned evolution of market expectations over time. Early in the recession, while the belief of being in a good regime is still high, the model predicts excessively optimistic output forecasts. Over time though, as the belief drops, so does the expectation of recovery and the model-predicted forecasts become much more accurate, in line with empirical data.

While the benchmark model with incomplete information predicts a correct pattern of interest rates over time, it still induces the government to preventively cut down on debt at the outset of the recession. To account for this shortcoming, In Chapter 2 of this dissertation I further extend the model by assuming that government has private information about the current regime[3] and engages in a signaling game with international lenders. This assumption is motivated by the fact that lenders do not only learn independently from the incoming data, but they also extract information by observing governments' actions. When major news is revealed about a bad state of an economy previously perceived as risk-free (for example a drastic fiscal reform or an IMF bailout), its foreign lenders may decide to downgrade their belief about the country's fundamentals. Hence, an otherwise efficient action may be much less appealing in a world with asymmetric information and market learning because it conveys a negative signal about the underlying state. Considering this channel, the government's optimal policy in response to a sudden regime switch involves a pooling equilibrium in which a low-regime-economy government continues to issue high levels of debt in order to appear more resilient to the markets. At the same time, the lenders continue to learn about the regime slowly and eventually respond with abrupt spikes in the interest rate. It is worth noting that even though the two models (with symmetric and asymmetric information, respectively) deliver contrasting predictions for the government's behavior, they are based on the same path of the lenders' belief.

[3] This means that the government has a more precise expectation about the income realization next period, not that it observes any data privately.

More generally, in this paper I argue that current sovereign debt models are fundamentally misspecified when applied to the recent European crisis. The reason is that they typically approximate the economy's path of income with a simple autoregressive process, which does not allow for the variable expectations of recovery over time. Using empirical evidence I show that these expectations evolved during the European crisis episode and that it has important implications for the quantitative predictions of sovereign debt models. In particular, standard $AR(1)$ specifications of the income process based on pre- and post-2008 GDP data will overestimate the variance and underestimate the persistence parameter, relative how the lenders perceived them at the time. Consequently, the predicted bond spreads are based on mismatched (overly optimistic) market expectations of recovery.

1.1.1 Literature review

This paper is closely related to the quantitative sovereign debt literature, in particular one building on the seminal work of Eaton and Gersovitz (1981) and, more recently, Aguiar and Gopinath (2007) and Arellano (2008). The models presented in these papers set foundations for our understanding of the dynamics of sovereign default risk and the mechanics of equilibrium defaults observed in the world. However, because of their simplicity these models are not capable of jointly explaining the motivating observations in this paper.

A more recent wave of sovereign default publications, such as Chatterjee and Eyigungor (2012), Hatchondo and Martinez (2009) or Arellano and Ramanarayanan (2012), introduce long-duration bonds as means of getting the models closer to the data and in particular allowing them to better match the observed bond spread behavior. As I show in Section 1.3, a model with long-term bonds delivers a more realistic paths of interest rates and debt levels, although still quite far from replicating the data trends. Moreover, these models do not capture the time-varying expectations of recovery and consequently generate predictions based on mismatched beliefs of market participants.

Another branch of sovereign debt literature develops models of political economy to explain why under certain circumstances governments might find it optimal to not reduce debt in the face of a crisis. Conesa and Kehoe (2015) present a model with self-fulfilling debt crises in which upon a looming recession the government may optimally increase its foreign debt in order to gamble for a possible recovery in the future. However, this mechanism does not hinder the lenders' ability to evaluate the default risk early on and lower the bond price.

More generally, this paper relates to the literature on information frictions in macroeconomics. In particular, Boz, Daude and Durdu (2011) take the model of Aguiar and Gopinath (2007) and assume that market participants are unable to distinguish between the incoming permanent and transitory shocks. In a calibrated model, they show how a learning process can explain some of the observed differences between developed and emerging market economies. Importantly though, their paper approximates default risk with an exogenous process and thus is not suitable to explain the dynamic pattern of interest rates. Pouzo and Presno (2016) develop a model of sovereign debt in which international lenders are uncertain about the model specification and require an additional premium for potential default risk. This specification allows them to match the bond spread dynamics observed in the data closely, albeit at the expense of abandoning the well-established rational expectations framework.

Finally, Chapter 2 of this paper is also related to the sovereign debt literature with signaling. Cole, Dow and English (1995) develop a framework in which hidden government types determine the willingness to default. They show that in equilibrium repaying the debt serves as a tool for the good type to separate from the bad type and build reputation. Similarly, Sandleris (2008) shows that governments repay their debts to communicate information about the economy's fundamentals to the lenders. Alfaro and Kanczuk (2005) present a quantitative model with adverse selection to show that some countries may choose to delay default until the times are bad enough to make it look "excusable". D'Erasmo (2011) adds private information and government reputation to an otherwise standard quantitative sovereign debt model and shows that it is capable of producing much higher debt levels than many previous papers in the field.

The remainder of this chapter is structured as follows. Section 1.2 describes the puzzling observations about the timing of the European debt crisis. Section 1.3 applies the existing standard models of sovereign debt to simulate the debt crisis in Europe. Section 1.4 introduces the main model. Section 1.5 calibrates the model and applies it to simulate the European debt crisis.

1.2 Background

In this section, I document the two motivating observations about the European debt crisis. I explain how the pattern of bond spreads and external debt securities over time contrasts with the predictions of existing models (I formally show it in Section 1.3). Finally, I present new empirical evidence on the evolution of crisis using historical GDP forecasts and explain its potential to resolve the puzzle.

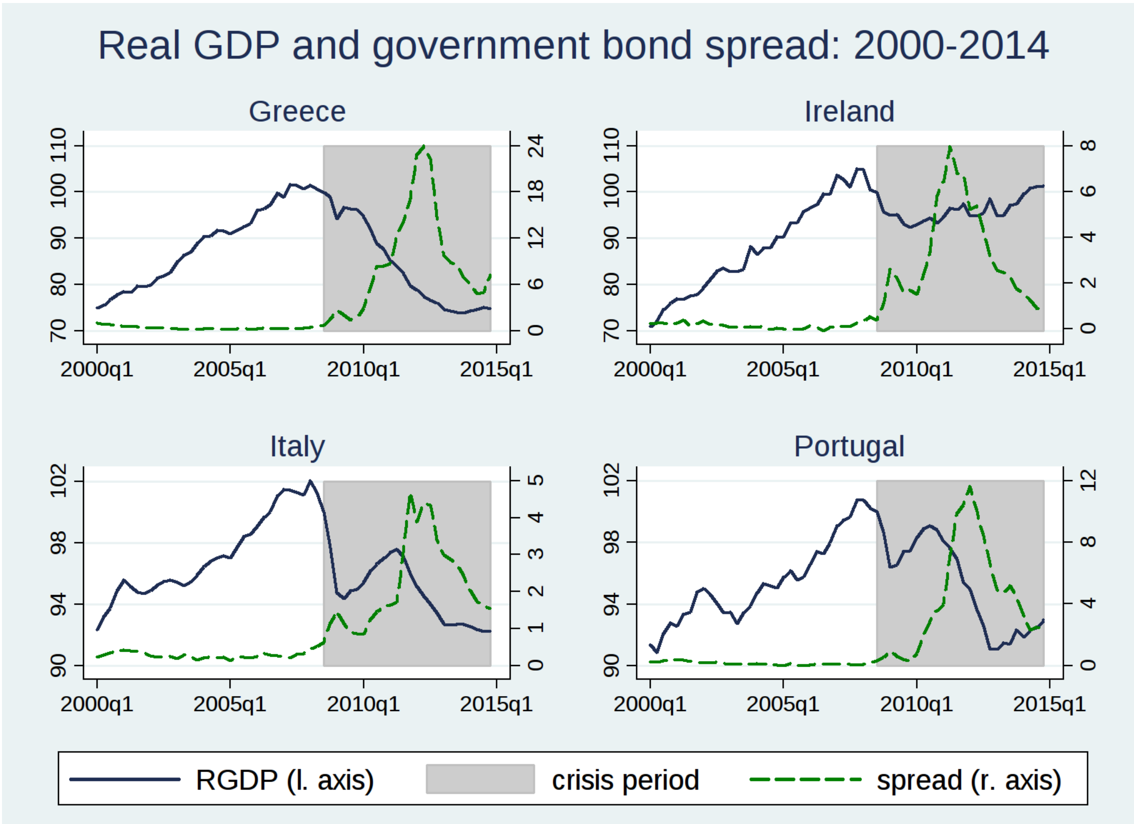
1.2.1 European debt crisis

Prior to 2008, the Eurozone's peripheral economies enjoyed a decade of relative prosperity and stable growth, fueled by the European integration, rising trade and the benefits of common currency. On the other hand, sovereign debt crises had predominantly been a problem of highly volatile emerging economies over the last few decades. As a result, the securities issued by governments of the developed European countries seemed risk-free to most financial market participants, in spite of the extraordinarily high debt stocks and low growth in these economies. The tendency to put excessive faith in Europe's ability to repay their debts continued even after the severe recession began in the second half of 2008. Figure 1.1 depicts the times series of real GDP for four[4] of the troubled European economies, along with the spreads on ten-year government bond yields.[5] As can be noticed, markets did not express much concern about the European governments' ability to repay for a long time following the initial slump in output. Instead, the bond spreads exhibited small "wiggles" when the financial crisis first hit, and only began a gradual increase afterwards, eventually leading to the dramatic spikes observed around 2011-2012.

The idea that financial markets have remained relatively calm for a long time following the outset of the crisis can also be grasped by examining the plot of output and sovereign bond ratings over time. Figure 1.2 presents plots of the countries' real GDP, together with

[4]In this figure, I focus on the cases of Greece, Ireland, Italy and Portugal. The fifth member of the colloquial GIIPS group, Spain, is left out for the purpose of clear exposition. Nevertheless, the empirical patterns discussed in this section can be observed for the Spanish case as well.

[5]The bond spread is defined as the difference between the annualized interest rate on the given ten-year government bonds and the interest rate on the German long-term bonds, assumed to be a risk-free asset.

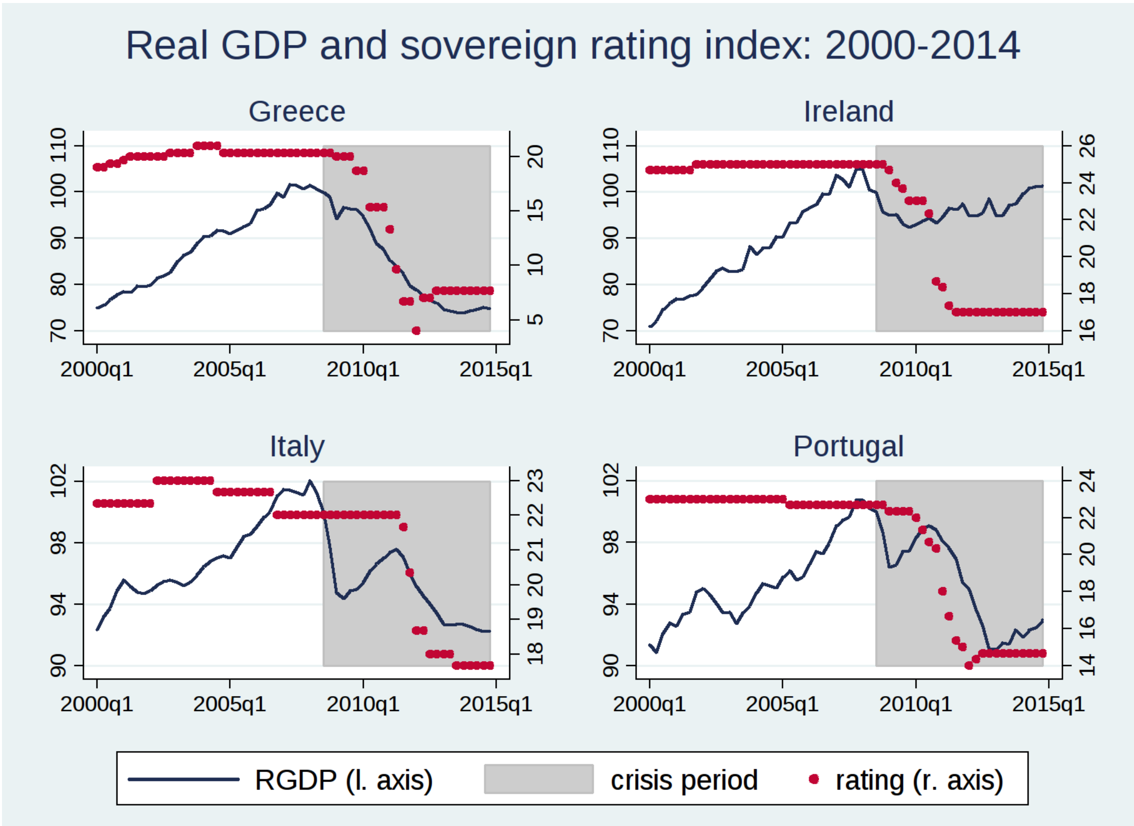


Note: The GDP series are in constant 2010 prices, and their values are normalized such that the third quarter of 2008 equals 100 (beginning of the financial crisis - shaded area). The bond spreads are expressed in percentage points and are acquired from the Bloomberg dataset.

Figure 1.1: GDP and bond spreads of the peripheral European economies: 2000-2014

the "sovereign bond rating index".[6] Even though the economies entered a severe recession as early as in the second quarter of 2008, markets continued to perceive their bonds as a relatively risk-free investment until about two years later. As a result, only around 2010-2011 do we observe a sequence of sovereign rating downgrades among peripheral European countries, indicating that market expectations about the governments' debt repayment probability had deteriorated significantly.

[6]The index is a simple weighted average of the three leading rating agencies (S&P, Moody's and Fitch), converted into a numerical scale from 0 to 25.



Note: The GDP series are in constant 2010 prices, and their values are normalized such that the third quarter of 2008 equals 100 (beginning of the financial crisis - shaded area). The bond rating index is constructed by converting the sovereign ratings of the three leading agencies - S&P, Moody's and Fitch - into a numerical scale from 0 to 25 and computing a simple average.

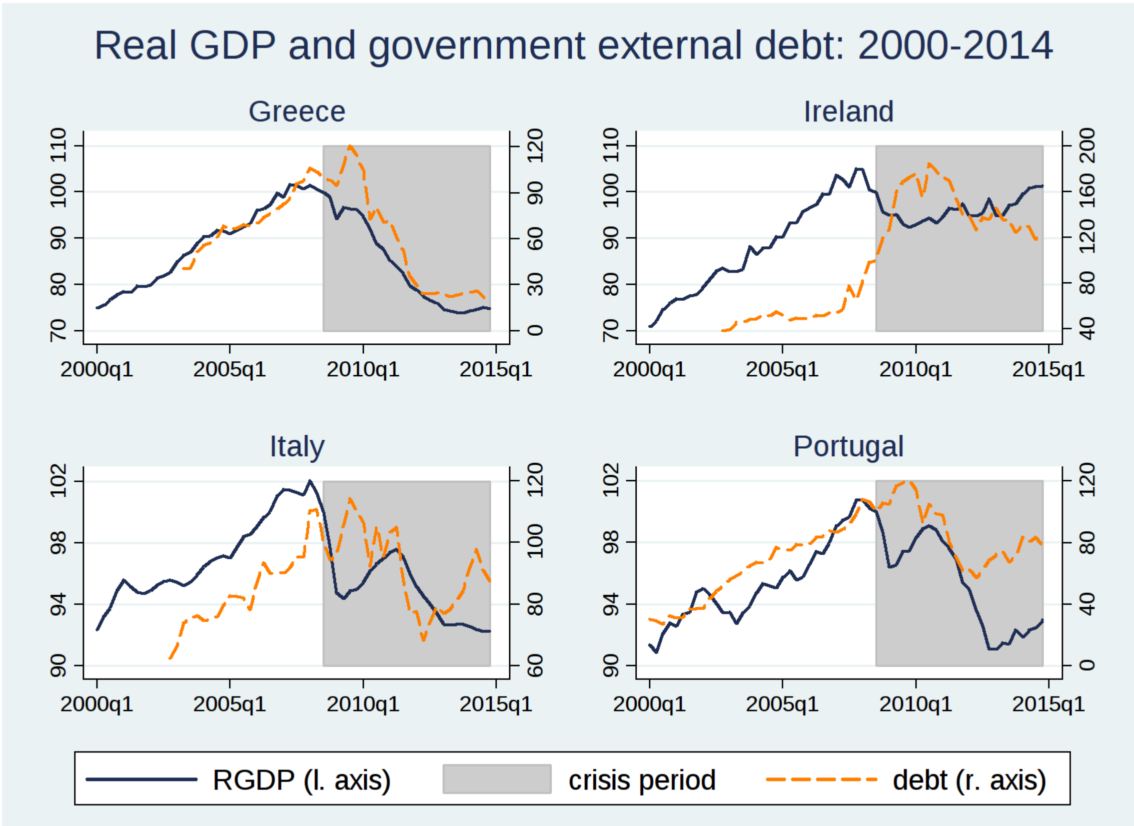
Figure 1.2: GDP and sovereign bond rating index of the peripheral European economies: 2000-2014

The evidence presented above suggests that financial market participants were not concerned about the risk of default during the first two years of recession. Considering the very high levels of government debt across Eurozone countries, this initially attenuated reaction of bond spreads seems to imply that markets viewed the shocks as more of a temporary downturn. Judging with hindsight, the actual debt crisis in Europe seems significantly delayed. Not surprisingly, as I formally show it in Section 1.3, existing sovereign debt models generally fail to predict such a pattern of bond spreads over time. This is because in a standard theory, a sovereign default is conditional on receiving unfavorable realizations of the income shock. It can easily be verified that already at the beginning of the Great Recession, the European economies incurred shocks to output comparable in their extremity to the one that triggered Argentina's default in the last quarter of 2001. Yet, despite generally much higher debt levels in Europe than in the emerging market economies, the interest rate only increased marginally on impact.

Another interesting aspect of the European crisis arises from examining the debt policy of national governments during the period of 2000-2014. Figure 1.3 contrasts the evolution of real GDP over time with external government debt of the four economies of interest. Two observations stand out in the graph. First, prior to 2008 all economies apart from Ireland exhibited gradually increasing paths of external debt. Such a steady rate of growth in debt continues despite the fact that the economies of Portugal or Italy had been slowing down since the early 2000s. Second, and more importantly for the present paper, when the low output shocks first hit in 2008, we see that governments respond by increasing their foreign debt on impact, and avoiding any major adjustments for the first two years of the crisis. The first sharp debt reductions occurred in years 2011-2012 and were mostly enforced by the international financial institutions (IMF and the European Commission), or an actual default in the case of Greece. As I also formally show in Section 1.3, this behavior contrasts the existing quantitative theories of debt. A government who faces an instant surge in the borrowing costs, as well as a looming possibility of default, has a strong motive to reduce the debt and to secure higher bond prices. I show that under a plausible calibration for Portuguese economy, a total predicted reduction in foreign debt service per quarter may reach 50 percent.

1.2.2 Market expectations during the recession

To shed more light on the source of the problem with current sovereign debt models, I look at important outside evidence on the expectations of the market participants. The beliefs about the distributions of future income shocks are a crucial element driving the interest rate in those models, and thus deserve particular attention.

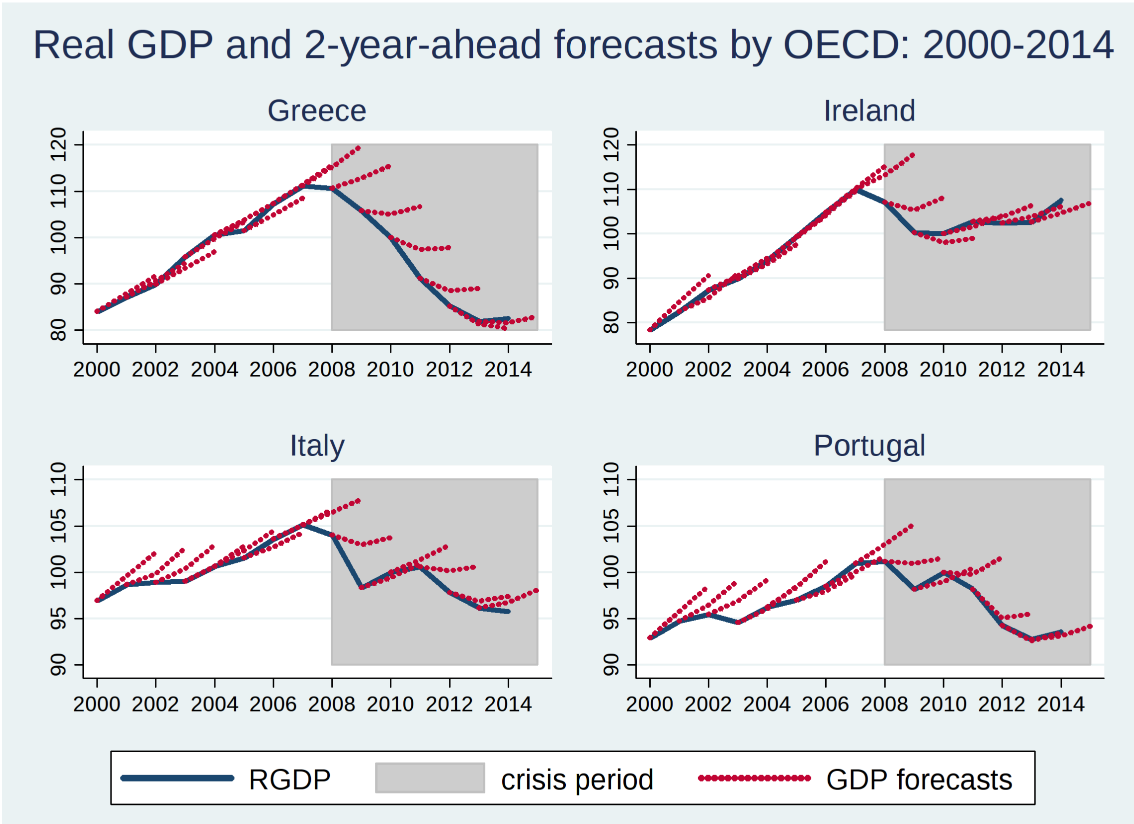


Note: The GDP series are in constant 2010 prices, and their values are normalized such that the third quarter of 2008 equals 100 (beginning of the financial crisis - shaded area). The debt series represent the external debt securities of the general government in constant 2010 prices, and its values are normalized such that the 2008:Q3 observation equals 100. The data are acquired from the World Bank's Quarterly External Debt Statistics and start at different points in time for different countries (hence the missing observations in the four figures above).

Figure 1.3: GDP and external debt of the peripheral European economies: 2000-2014

Figure 1.4 once again presents the plot of real GDP over time for the European countries, together with the GDP forecasts published every year by OECD.[7] As can be noticed for the period prior to 2008, while the European economies are still growing the forecasts tend

[7]Even though I use OECD data for calibration in this paper, I also verify that the same pattern of forecasts holds in other publicly available sources, most notably the IMF and the European Commission.



Note: The GDP series are annual and expressed in constant 2010 prices; their values are normalized such that the observation for 2010 equals 100. The red dotted lines represent one- and two-year ahead forecasts published by the OECD Economic Outlook (fall edition) and start in the year when each of them is made.

Figure 1.4: Forecast and actual GDP for the peripheral European economies: 2000-2014

to be precise (with a slight overshooting pattern for Italy and Portugal, whose output growth had been slowing down since early 2000s). When the financial crisis breaks out, the forecasts are still fairly optimistic, predicting a recovery in years 2008-2010. Over time however, as the GDP continues to plunge we also observe that the forecasts become flatter, indicating that the markets have realized the recovery of output cannot be expected in the short and medium term. From 2012 on, the forecasts essentially line up again with the subsequently realized data for all of the depicted economies. This is also the time when the

European bond markets undergo unprecedented turbulences, with surging interest rates and drastic reductions in debt levels, as it is documented in Figures 1.1 and 1.3. In the quantitative analysis part of this paper (Section 1.5.2) I present further evidence linking the markets' learning about future income with the outbreak of the debt crisis in 2011-2012.

In the next two sections I show that the process of learning about future income shocks evident in Figure 1.4 is an important missing element from the previous studies. What all those models have in common is a simplified treatment of the stochastic process for output, typically assumed to be an AR(1). Under such a specification, the economy always reverts to the long-run unconditional mean which is equal to the deterministic linear trend. As a result, during a recession, when the economy falls deep below the trend, agents in the model by construction predict a recovery of output in a uniform fashion over time. The speed of this expected recovery is determined by the persistence parameter of the process. Consequently, even if we calibrate a model so that it does a better job at predicting the pattern of interest rates during the crisis (e.g. a model with long-term bonds in which it takes a longer time for income realizations to reach the default set), it is still based on the mismatched beliefs of the market participants. In this sense, I argue that existing sovereign debt models lack the microfoundations that could help bring their predictions closer to the data, and make them consistent with the agents' actual expectations about future.

1.3 Existing models

In this section I consider the standard quantitative models of sovereign debt and I use them to simulate the European debt crisis in Portugal. The purpose of this exercise is to show the extent to which the main model developed in this paper can generate improved predictions relative to the existing theories.

1.3.1 Short-term debt model

As a first step, I consider the basic one-period debt model exactly as described in Arellano (2008), following her proposed calibration strategy. In what follows, I briefly recall key equations of the model written in a form with discrete income shocks.

Consider a small open economy with an endowment process of the form:

$$y_t = \rho y_{t-1} + \eta \varepsilon_t \tag{1.1}$$

where ε is an *i.i.d.* shock that follows standard normal distribution, while ρ and η are persistence and variance parameters, respectively. The general value of the government incorporates the choice of repayment or default:

$$v^0(b, y) = \max_{d \in \{0,1\}} \{(1-d)\, v^r(b, y) + d\, v^d(y)\} \tag{1.2}$$

The value associated with default is given by the following equation:

$$v^d(y) = u\big(\min(\hat{y}, y)\big) + \beta \sum_{y'} \pi(y', y)\big[\theta v^0(0, y') + (1-\theta) v^d(y')\big] \tag{1.3}$$

where $\pi(y', y)$ denotes the probability of moving from state y today to state y' tomorrow, and $\hat{y}$ is a parameter of the default penalty function. The value associated with repayment is defined as:

$$v^r(b, y) = \max_{b'} \left\{ u\big(c\big) + \beta \sum_{y'} \pi(y', y) v^0(b', y) \right\} \tag{1.4}$$

subject to

$$c = y - b + q(b', y) b' \tag{1.5}$$

International lenders are assumed to be perfectly competitive, risk-neutral and to have "deep pockets" in the sense that they can cover potentially large losses. Consequently, the government bond price is given by

$$q(b', y) = \frac{\sum_{y'} \pi(y', y)\big(1 - d(b', y')\big)}{1 + r^*} \tag{1.6}$$

I solve the model numerically using 300 grid points for assets and 41 grid points for the income space.[8] To calibrate the model, I follow the Arellano (2008) in assuming a CRRA utility function $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ with a standard parameter value of $\sigma = 2$. The risk-free interest rate is fixed at $r^* = 0.01$. The transition probabilities are derived using the AR(1)

[8]The state space used here is much more dense than in the original paper because, as Hatchondo, Martinez and Sapriza (2010) show, the results of sovereign debt models are typically sensitive to grid sizes.

discretization method of Rouwenhorst (1995) which is appropriate for highly persistent processes. The persistence and variance parameters of AR(1) are estimated from Portuguese real GDP data for 1974:Q2-2011:Q4, up until the height of the debt crisis, as in Arellano (2008)'s study for Argentina.[9] The resulting parameter values are 0.9816 and 0.0108, respectively. Notice that in particular the estimated persistence parameter is very high, making the process close to a random walk. This is a result of treating the Portuguese slump of 2008-2012 as an extremely unlucky sequence of bad shocks. By contrast, in the regime-switching model presented in Section 1.4, this episode arises from an underlying regime switch and thus the obtained persistence parameter is much lower.

Parameter	**Value**	**Meaning**	**Source**
σ	2	Risk aversion	Literature
r^*	0.01	Risk-free rate	Literature
ρ	0.9816	Shock persistence	Estimation
η	0.0108	Shock volatility	Estimation
$\hat{y}$	0.958	Default penalty	Calibration
β	0.963	Discount factor	Calibration
θ	0.169	Re-entry probability	Calibration
Calibration targets		**Model**	**Data**
Debt service/GDP		8.46	8.47
St. dev. TB/GDP		1.36	1.61
Long-run def. prob.		1.53	1.50

Table 1.1: Calibrated parameters of the benchmark short-term debt model

The remaining three parameters of the model (β,$\hat{y}$,θ) are jointly calibrated as in Arellano (2008) by matching the following moments of the Portuguese economy: a long-run default probability of 1.5%,[10] an average debt-service-to-GDP ratio of 8.47% for 1995-2007, and

[9]The series are detrended allowing for a statistically significant breakpoint, see Section 1.5.2 for details.

[10]Obtaining a precise estimate for this moment is difficult due to insufficient recent historical experience. Standard & Poor's (2006) documents three sovereign default episodes for Portugal since 1800, while

the standard deviation of the trade balance of 1.61%. The model-generated moments are obtained by averaging across one thousand simulated paths of debt accumulation for the period of 1995:Q1-2012:Q1[11], feeding in the initial debt and output data for 1995:Q1. Table 1.1 summarizes the model calibration results. Notice that the moments are matched fairly well and the obtained parameters are close to the ones originally used by Arellano (2008).

I use the model to simulate the debt accumulation pattern and interest rate spread over the period 2000-2014 by feeding in the actual GDP realizations. Figure 1.5 presents the result of this experiment. As can be noticed, the debt level quickly increases in the early 2000s and starts dropping in late 2002, when the Portuguese economy first experienced a GDP slowdown. The bond spread exhibits several jumps in that period, corresponding to the major income shocks. Most importantly however, when the Great Recession begins in 2008 the bond spread shoots up leading to a default in the first quarter of 2009.

Such stark counterfactual predictions of the benchmark model are not particularly surprising. In this setup, the government is impatient and accumulates debt up to the steady state level very quickly (in the period 2000-2002) which it subsequently reduces upon receiving bad shocks. The predicted sovereign default at the beginning of the Great Recession results from the unprecedented scale of the income shocks Portugal experienced. To put it into perspective, in 2001:Q4 Argentina defaulted on its debt when it got hit by a shock 2.67 standard deviations below what it had expected (this is unlike 99.2% of all shocks drawn from the normal distribution). By comparison, the shock that hit Portugal in 2009:Q1 was 2.56 standard deviations below the mean, unlike 98.9% of normally distributed shocks. In a model entirely driven by income fluctuations it should not be surprising that a default is a predicted outcome.

Reinhard and Rogoff (2009) also add a debt crisis of 1828. While all of these episodes occurred in the 19th century and are thus not very informative, the resulting annual default probability of 1.5% appears to be a reasonable number, placing Portugal in terms of riskiness between Argentina (3%) and Germany (1%) or the UK (0.5%).

[11] While Portugal did not technically default on its debt during the recent crisis, I follow Arellano (2008) and consider a sample up to the peak of the bond spread which occurred in the first quarter of 2012.

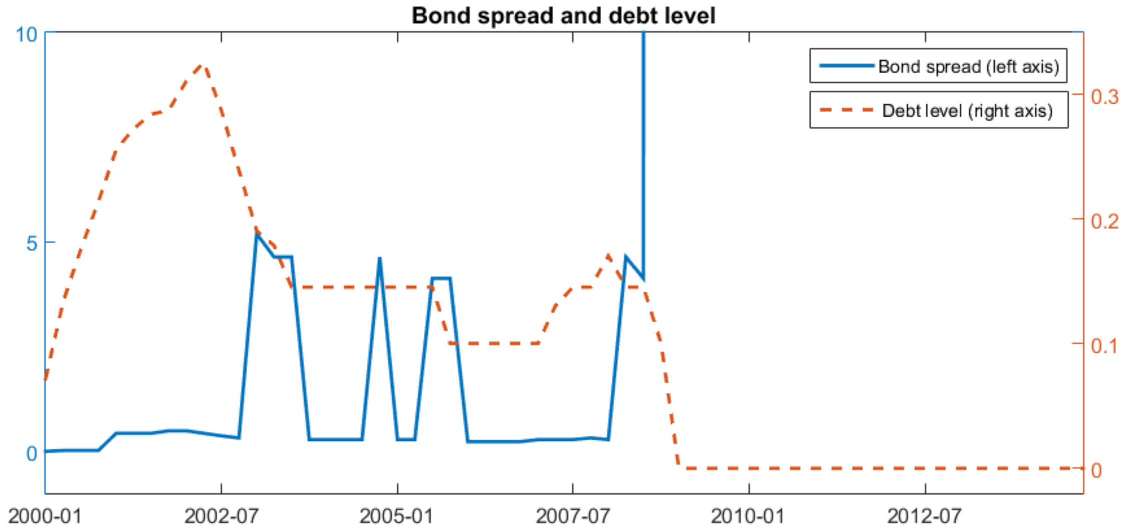


Figure 1.5: Simulated debt crisis in the benchmark short-term debt model

1.3.2 Long-term debt model

A problem with the model of Arellano (2008) is that it only allows the government to issue one-period non-contingent bonds, while in reality we observe sovereign debt with much longer maturities. As a result, such models require either an extremely low value of the discount factor (under the unrealistic assumption that the government rolls over the *entire* debt stock every quarter), or they must be calibrated to match the average debt service of the country only. Using the more conservative latter approach, it is shown in the previous section that such models fail to deliver the correct predictions for the European debt crisis.

In the present section, I introduce long-term sovereign debt in the spirit of Chatterjee and Eyigungor (2012) or Hatchondo and Martinez (2009). Specifically, I assume that each period a government bond matures with a probability given by parameter δ. With probability $1 - \delta$ the bond does not mature and then it pays a coupon κ. As a result of this assumption, in expectations the government debt matures at the rate of δ while the modeler only needs to keep track of a single state variable b. The government budget constraint analogous to (1.5) is therefore

$$c = y - \left[\delta + (1-\delta)\kappa\right] b + q(b', y)\left[b' - (1-\delta)b\right] \tag{1.7}$$

The government's value functions associated with repayment and default are analogous to equations (1.4) and (1.3), respectively. Following Chatterjee and Eyigungor (2012), I assume a non-linear specification of the output after default $y - \max\{0, d_0 y + d_1 y^2\}$, which facilitates bringing the model to the data. With long-term debt, the analog of the bond pricing equation (1.6) becomes

$$q(b', y) = \frac{\sum_{y'} \pi(y', y)\big(1 - d(b', y')\big)\left[\delta + (1-\delta)\big(\kappa + q(g(b', y'), y')\big)\right]}{1 + r^*} \tag{1.8}$$

where $g(\cdot)$ is the equilibrium debt policy function of the government. Formula (1.8) shows that the non-matured portion of government debt factors into the current bond price by allowing for the possibility of price change in the future, as a result of higher default risk.

I solve the model numerically using continuous debt choice methods described in Hatchondo, Martinez and Sapriza (2010). Similarly as in the previous section, I assume the risk aversion parameter of 2, the risk-free interest rate of 0.01, and the persistence and standard deviation parameters of the AR(1) process of 0.9816 and 0.0108, respectively. The maturing probability δ is fixed at 0.05, implying an average maturity of 20 quarters (or 5 years), while the coupon rate is set to 0.0125, implying an annual coupon of 5%. As in Chatterjee and Eyigungor (2012), I fix the probability of re-entry following a default at 0.0385, and I proceed to calibrate the remaining parameters of the model, (β, d_0, d_1) by jointly matching the following moments of Portugal's 1998:Q1-2014:Q4 data: average bond spread of 1.7%, standard deviation of the bond spread of 2.93, and average quarterly debt-to-GDP ratio of 1.77.[12] Table 1.2 summarizes the parameter values of the model. As can be noticed, the moment-matching exercise produces a reasonable match to the Portuguese experience.

Similarly as in the previous section, I use the calibrated model to simulate the actual debt crisis event by feeding in the initial debt level and the path of GDP realizations for

[12] Like in Chatterjee and Eyigungor (2012), I assume that only 70% of the sovereign debt is defaultable, due to the fact that we generally observe governments repaying a part of their debt before they re-enter credit markets.

Parameter	Value	Meaning	Source
σ	2	Risk aversion	Literature
r^*	0.01	Risk-free rate	Literature
θ	0.0385	Re-entry probability	Literature
κ	0.0125	Coupon rate	Literature
δ	0.05	Probability of maturing	Literature
ρ	0.9816	Shock persistence	Estimation
η	0.0108	Shock volatility	Estimation
d_0	-0.84	Default cost parameter	Calibration
d_1	0.96	Default cost parameter	Calibration
β	0.98	Discount factor	Calibration
Calibration targets		**Model**	**Data**
Debt/GDP		178	177
Average spread		1.91	1.70
St. dev. spread		2.52	2.93

Table 1.2: Calibrated parameters of the benchmark model with long-term debt

Portugal over the years 2000-2014. Figure 1.6 depicts the result of this exercise. As can be noticed, the long term debt provides much more realistic predictions for the path of debt accumulation and interest rate spreads prior to the Great Recession. As the debt reaches its height at the end of 2008 (equivalent to roughly 62% of annual debt-to-GDP), the country faces the large shocks to income and the interest rate spread shoots up to over 20%. In response, the government undertakes debt reduction, albeit in a much more smooth manner than in the case of short-term debt, evident in Figure 1.5. During the temporary recovery of 2009-2010 the interest rate spread falls back to a low level (still considerably higher than zero), but eventually in 2011Q3, upon receiving the most significant shock to GDP, the government chooses to default. While the latter is not an unreasonable prediction (recall that Portugal received a large-scale bailout from the IMF and the European Commission in May of 2011 which most likely prevented it from defaulting), the behavior of interest rates and the debt level at the outset of the Great Recession is clearly counterfactual.

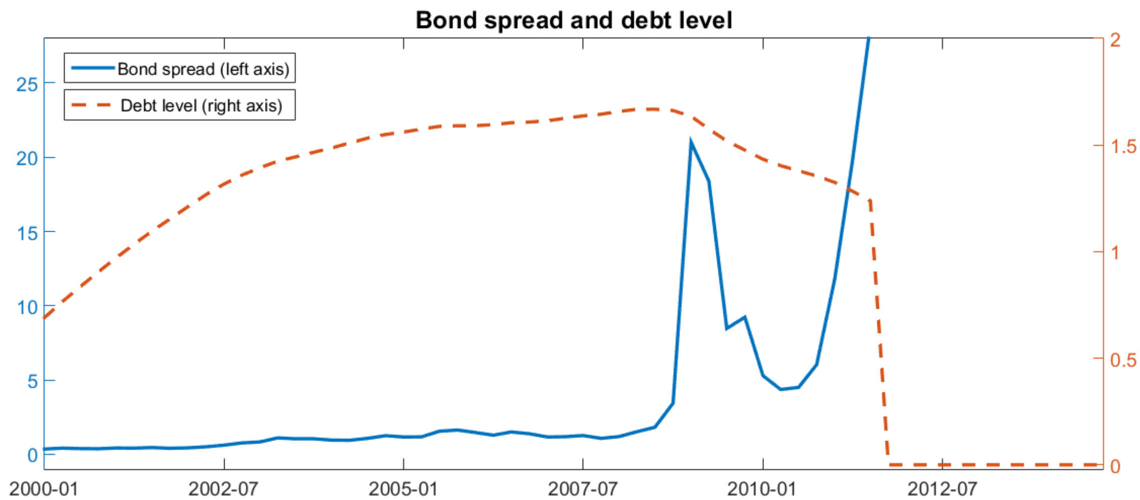
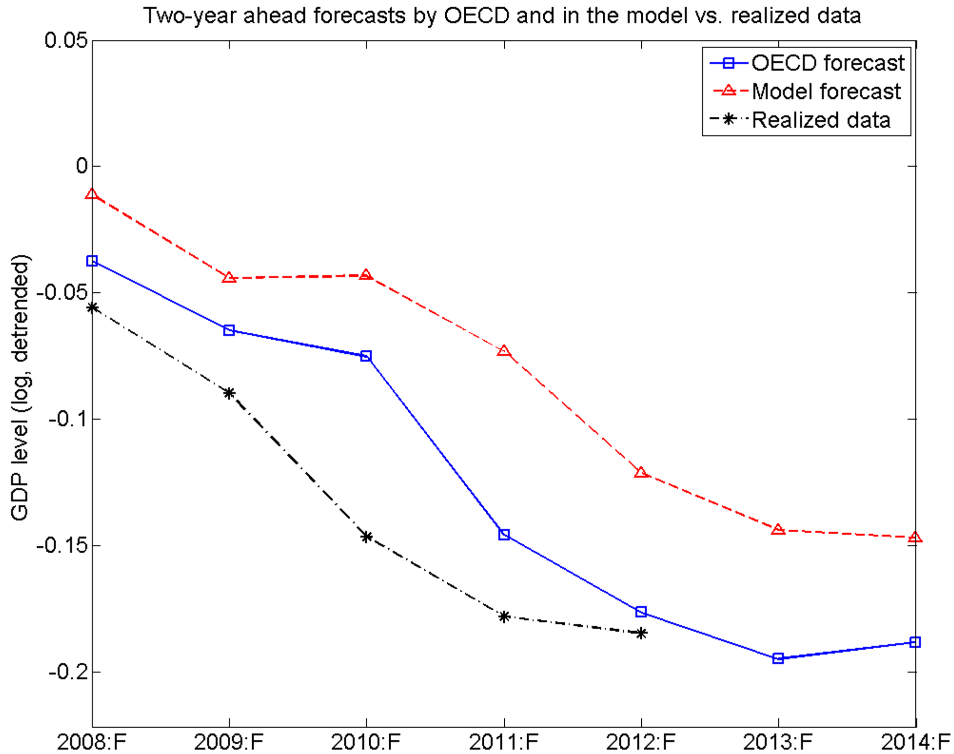


Figure 1.6: Simulated debt crisis in the benchmark long-term debt model

1.3.3 Misspecification of beliefs

On a more fundamental level, I argue that any model of sovereign debt based on a simple AR(1) process for income is fundamentally misspecified, due to the fact that it will mispredict the path of market expectations of recovery over time. Figure 1.7 is an analog of Figure 1.9 in Section 1.5 and presents the two-year-ahead forecasts generated by agents in the model (red dashed line) who base their judgment on the income process described in (1.1). These expectations are contrasted with the forecast data published by the OECD (see Section 1.5 for details), depicted by the blue solid line, and the factual data realized two years later (black dash-dotted line). As can be noticed, in each year during the crisis market participants tend to predict a recovery in a uniform fashion (due to the mean reversion of the process). In particular, in years 2011-2012 the model forecasts are still overly optimistic whereas in reality we observe a significant downward correction and improvement in accuracy. As I show in Section 1.5, the learning model proposed in this paper does a much better job at matching the beliefs of market participants during the debt crisis.



Note: Each point on the graph represents an annual detrended log-GDP level for two years ahead (for example, 2008:F corresponds to the GDP level in 2010). The solid blue line plots the actual OECD forecasts, while the dashed red line denotes the ones generated by the estimated model. Additionally, the dashed-dot black line shows the actual realized data that the corresponding forecasts refer to (only available until 2012:F, when the prediction for year 2014 was made).

Figure 1.7: Historical GDP forecasts: OECD- and model-generated predictions.

1.4 Model

In this section I present a model of sovereign debt that features an augmented specification of output process and partial (symmetric) information about its realizations. I then further augment it by adding asymmetric information in Chapter 2.

1.4.1 Economic environment

Consider a representative-agent small open economy with a benevolent sovereign government that borrows internationally from a large number of competitive lenders. Time is discrete and there is no production or labor. Instead, the economy faces a stochastic stream of endowment realizations. Markets are incomplete and the only asset available for trading is the one-period non-contingent bond.

Endowment process Suppose the country's endowment follows an autoregressive regime-switching process with two possible regimes, High and Low. Each of them is characterized by its own long-run mean. On the other hand, the persistence and variance parameters are assumed to be constant across regimes.[13] Specifically, the evolution of output, detrended with a deterministic long-run mean growth rate, is given by

$$y_t = \mu_j(1-\rho) + \rho y_{t-1} + \eta\varepsilon_t \tag{1.9}$$

where $\varepsilon_t \sim \mathcal{N}(0,1)$ is an *i.i.d.* random shock and $\rho, \eta, \{\mu_j\}_{j=L,H}$ are parameters of the two regimes. Regimes change according to a Markov process with the transition probability matrix given by

$$\Pi = \begin{bmatrix} \pi_L & 1-\pi_L \\ 1-\pi_H & \pi_H \end{bmatrix} \tag{1.10}$$

The specification of a bimodal stochastic process of endowment given in formula (1.9) is non-standard in the sovereign debt literature. It is motivated however by the growth pattern of the European economies in the recent decade, which was illustrated in Figures 1.1-1.3. This growth pattern differs considerably from the one of most emerging economies which exhibit frequent ups and downs around the trend.

[13] While this assumption may not necessarily be true, we do not have long enough times series to distinguish reliably between them. The assumption is important however for the single crossing property to hold in the model with asymmetric information (for details, see Chapter 2). Generally, it seems intuitive that the high-mean regime should be preferred to the low-mean one. Nevertheless, depending on the parameters, there may be states of nature in which this preference is reversed. For example, suppose that $\rho_H > \rho_L$ and $\eta_H < \eta_L$. If the income today is low enough then the high persistence and low volatility of the High regime implies that the economy will suffer for a number of periods, before slowly rebounding back to the high unconditional mean. In such case, being in the Low regime might be more desirable as it gives the government a higher chance of getting better income shock already next period. Consequently, the single crossing property is broken.

Preferences The representative household has preferences given by the expected utility of the form:

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{1.11}$$

where I assume the function $u(\cdot)$ is strictly increasing, concave and twice continuously differentiable. The discount factor is given by $\beta \in (0, 1)$.

Government In each period, the government chooses a consumption rule and the level of debt holdings in order to maximize the household's lifetime utility. The only asset available is the one-period zero-coupon bond. The government may save at an international risk free rate. If it decides to borrow, however, the government is not committed to repay the debt next period. Consequently, the bond is priced endogenously by risk-neutral lenders to account for the possibility of default. As it is commonly assumed in the sovereign debt literature, the government who refuses to honor its obligations faces an exogenous cost of default and is further excluded from borrowing in the financial markets, with a certain probability of being readmitted in every subsequent period.

Market clearing There is no storage technology and, under the aforementioned assumptions on the utility function, implies that the endowment is fully divided between current consumption and net borrowing. This market clearing condition is given by

$$c_t = y_t - b_t + q_t b_{t+1} \tag{1.12}$$

where q_t is the price of the bond b_{t+1} (to be repaid next period), and the negative value of b_t means that the government is a borrower.

Bond prices International lenders are perfectly competitive[14] and have "deep pockets" in the sense that potentially even large losses do not affect their decisions. In equilibrium the lenders make expected zero profit and as a result, the bond pricing formula compensates them only for the default risk implied in the government's decisions.

[14] This assumption is relaxed for one of the variants of the model featuring asymmetric information. See section 1.4.4.

1.4.2 Information structure

The two state variables mentioned so far, current bond holdings (b) and income (y), are standard in sovereign debt literature. In addition, the present model features another exogenous state, $z \in \{z_L, z_H\}$, which represents the regime (Low or High) in which the economy is currently operating. While all market participants are always aware of the present endowment realization, the model in this paper will feature incomplete information about the current regime. International lenders do not observe the regime, but instead form a belief p defined as the probability assigned to being in the High regime, formally $p = Prob(z = z_H)$. Intuitively, this variable can be thought of as market sentiment about the economy's expected income. As it is shown in the calibration section of this paper, the assumption that investors are never sure about the precise expectation of future output has strong empirical support. Following decades of stable (albeit low) growth with very moderate business cycle fluctuations, the European economies appeared resilient to major depression events. Hence, at the outset of the crisis even upon observing a sequence of very bad shocks, the markets are prone to believe in a prompt recovery. It takes some time until they realize the underlying regime has switched and are able to correctly assess the economy's expected future performance.

I consider two alternative assumptions about information available to the government. In the first variant, a Partial Symmetric Information model (PSI hereafter) the government has no knowledge of the current regime and shares with the markets a common belief p, defined as above. In this setup, Bayesian learning from the incoming output data is the only channel of updating the belief every period and it occurs symmetrically on the part of foreign lenders and the government. In the second variant, a Partial Asymmetric Information model[15] (PAI hereafter) the government knows the current regime exactly and the markets are aware of this informational advantage. In such case, the lenders get another channel of updating the belief by analyzing the observed debt allocation to infer the economy's regime by backward induction. This channel gives rise to another posterior belief $\widetilde{p}$ which is a function of the current state and the government's bond choice and feeds back into the price function. The government is aware of the signaling externality its

[15] A detailed analysis of this model is deferred until Chapter 2. Here, I limit the description to a brief overview for the purpose of comparison.

actions instill, and optimally chooses whether to engage in a separating equilibrium (i.e. one that reveals its identity) or pooling equilibrium (in which the lenders cannot infer its type and must rely on their own prior belief in evaluating a default probability).

As a direct benchmark for the main model, I also consider the typical Full Information (FI) variant, which is similar in nature to the model of Arellano (2008). The only difference boils down to a richer stochastic process for endowment specified in equation (1.9). As it has been mentioned before, the assumption of dichotomous output regimes appears particularly appropriate for the case of European economies, which tend to experience long streaks of persistent over- and under-performance relative to a long run trend. Moreover, I will argue that this augmented specification allows for variable expectations of future income which has been an important empirical factor during the European debt crisis.

1.4.3 Timeline

In every period, the timing of events is as follows:

1. The new regime $z \in \{z_L, z_H\}$ is drawn, with the probability distribution given by (1.10).

2. The new realization of endowment y is drawn, according to the newly updated regime z and conditional on its level from last period.

3. International lenders observe the endowment level y and mechanically form a new prior belief p about the output regime, conditional on the previous and current endowment realizations, as well as last period's posterior belief $\widetilde{p}$.[16]

4. Default and redemption decisions take place:

   - The government that has recently defaulted on its debt draws a random number to determine whether it can be readmitted to the financial markets.
   - The government that has recently been current on its debt decides whether to repay or default this period.

[16] In the present chapter, $\widetilde{p}$ is by construction equal to p.

5. Equilibrium allocations take place:

   - If the government defaults, it is excluded from financial markets this period and simply consumes its endowment, subject to a default penalty.
   - If the government repays, it chooses the new allocation of bonds b', while the lenders update their posterior belief $\widetilde{p}$ and post the bond price $q(b', y, \widetilde{p})$.

The treatment of the lenders' belief is a novel aspect of the present model and requires a further comment. In the general specification, in each period there are up to two stages of updating the lenders' belief. One of them results from mechanical application of Bayes' formula and always occurs at the beginning of the period, as soon as the new endowment realization has arrived. In that stage, the last period $\widetilde{p}$ is taken as input and the new state variable p is returned. The other stage is an outcome of the signaling interaction (described in detail in Chapter 2) between the lenders and the government, and is determined simultaneously with equilibrium allocations. Here, the state variable p is mapped into a new posterior belief $\widetilde{p}$. Consequently, both variables p and $\widetilde{p}$ constitute either a prior or a posterior belief, depending on the updating stage considered. For notational coherence, and given the chronological order of events within a period, throughout the paper I will refer to the state variable p as the prior, and the endogenous variable $\widetilde{p}$ as the posterior. Note that under the Partial Symmetric Information model the signaling interaction does not take place and naturally $\widetilde{p} = p$.

1.4.4 Recursive formulation

In the following section I formalize the economic environment by stating the problems faced by market participants in recursive form. To begin, define the vector of aggregate state variables that are common knowledge as $\mathbf{s} = (b, y, p)$. Notice that once we replace the state variable p with z, the model easily collapses to its full information variant which I consider as a benchmark in the quantitative analysis.

Government The government that is current on its debt obligations has the general value function given by

$$v^0(\mathbf{s}) = \max_{d \in \{0,1\}} \Big\{(1-d)v^r(\mathbf{s}) + dv^d(y, p)\Big\} \tag{1.13}$$

A sovereign who decides to default ($d = 1$) is excluded from international credit markets and has the probability θ of being readmitted every subsequent period. The associated default value is given by

$$v^d(y,p) = u\big(h(y)\big) + \beta \sum_{z \in \{z_L, z_H\}} \sum_{z' \in \{z_L, z_H\}} Prob(z)\, \pi(z'|z) \times \int f_{z'}(y', y) \left[\theta v^0(0, y', p') + (1-\theta) v^d(y', p')\right] dy' \tag{1.14}$$

subject to the law of motion for the lenders' beliefs

$$p'(y, y', p) = \frac{f_{z_H}(y', y) \times [p\, \pi(z_H|z_H) + (1-p)\, \pi(z_H|z_L)]}{\sum_{z'=z_L,z_H} f_{z'}(y', y) \times [p\, \pi(z'|z_H) + (1-p)\, \pi(z'|z_L)]} \tag{1.15}$$

and where the current regime probabilities are determined by the belief state variable p, i.e. $Prob(z_H) = p$ and $Prob(z_L) = 1 - p$. In equation (1.14), $h(\cdot)$ is a reduced-form representation of the output cost of defaulting[17]; $f_{z'}(y'|y)$ denotes the probability density of transitioning from state y to state y' given that tomorrow's regime is z'. The next period belief p', described in equation (1.15), depends on the current and future income realization, as well as the current period belief p. It is a simple application of Bayes' rule and takes into account a potential regime switch at the beginning of next period, in accordance to the transition matrix given by formula (1.10).

The value of the government associated with repayment of debt is given by

$$v^r(\mathbf{s}) = \max_{c,b'} \left\{ u(c) + \beta \sum_{z \in \{z_L, z_H\}} \sum_{z' \in \{z_L, z_H\}} Prob(z)\, \pi(z'|z) \int f_{z'}(y', y) v^0(\mathbf{s}')\, dy' \right\} \tag{1.16}$$

subject to

$$c = y - b + q(b', y, p) b' \tag{1.17}$$

$$b' \leq \bar{b}\, y + \xi b \tag{1.18}$$

$$p'(y, y', p) = \frac{f_{z_H}(y', y) \times [p\, \pi(z_H|z_H) + (1-p)\, \pi(z_H|z_L)]}{\sum_{z'=z_L,z_H} f_{z'}(y', y) \times [p\, \pi(z'|z_H) + (1-p)\, \pi(z'|z_L)]} \tag{1.19}$$

[17] Quantitative sovereign debt models typically assume an exogenous punishment in the case of default in order to facilitate calibration of the model to the data. For the specific functional form, see section 1.5.3.

where equation (1.17) is the budget constraint, equation (1.18) is the exogenous debt limit, and the law of motion for the lenders' beliefs is given by formula (1.19).[18]

The debt limit specified as inequality (1.18) requires a word of comment. It is not standard in the sovereign debt literature to subject governments to exogenous borrowing constraints. However, in contrast to the previous defaulters, European economies have been accumulating debt gradually over time. This observation is illustrated in Figure 1.8 which depicts the trade balance to GDP ratios for Portugal and Argentina, prior to its 2001 default. As can be noticed, while Argentina maintained a trade balance close to zero and with considerable volatility, Portugal kept a sizable negative trade balance for the entire period of 1995-2012. This poses a calibration issue, because the models of sovereign debt with no borrowing limits typically exhibit very rapid convergence of the debt level to its steady state due to government's impatience. As a result, some form of an exogenous debt ceiling is necessary in order to match the observed path of debt accumulation of the European economies prior to 2008. The general specification of inequality (1.18) follows a vast literature in international economics[19], the exact parameter values however will result as the outcome of endogenous calibration. Notice furthermore that such a form of debt constraint has a clear empirical counterpart. All members of the Eurozone have their fiscal policies bounded by the Maastricht treaty which imposes exogenous limits on debt-to-GDP ratio and the government deficit.[20]

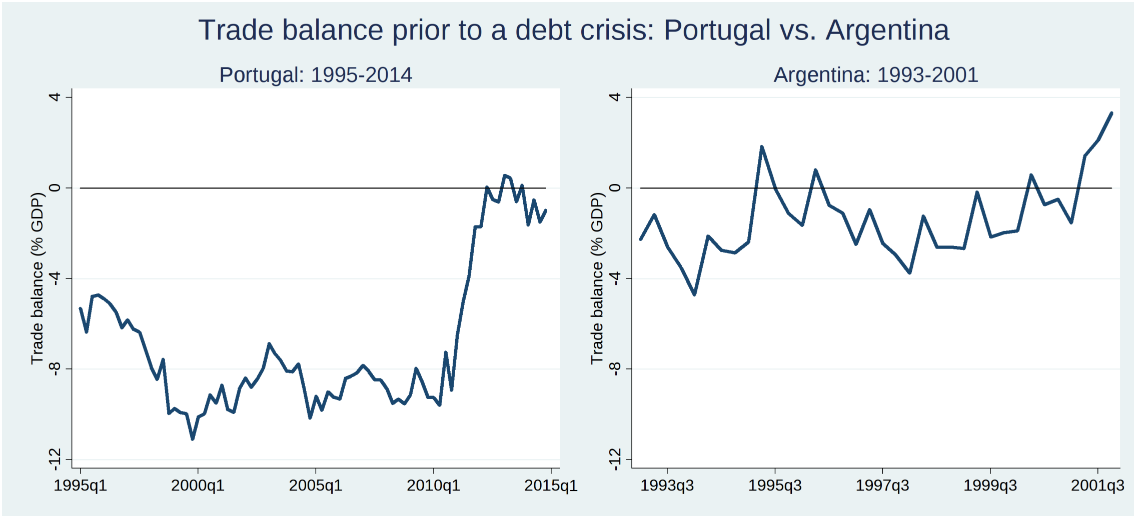
Having characterized the two value functions of the government, it is straightforward to derive the optimal default policy as a function of today's state variables

$$d(\mathbf{s}) = \begin{cases} 1, & \text{if } v^d(y,p) > v^r(\mathbf{s}) \\ 0, & \text{if } v^d(y,p) \leq v^r(\mathbf{s}) \end{cases} \tag{1.20}$$

[18]Note that equations (1.15) and (1.19) are identical as they represent the same Bayesian updating mechanism.

[19]For example, Guerrieri, Iacoviello and Minetti (2012) apply this form of debt constraint to their analysis of the European debt crisis.

[20]The Maastricht treaty stipulates punishment procedures against the violators of the fiscal policy rules. Naturally, before and especially during the crisis years several members exceeded the limit and did not face tangible consequences. To account for this fact, equation(1.18) assumes a flexible form of the debt limit and the exact parameter values are picked endogenously in calibration.



Source: OECD for Portugal, Ministry of Finance (MECON) for Argentina.

Figure 1.8: Debt accumulation patterns for different debt crises: Portugal vs. Argentina

Bond pricing Every period the lenders only observe (b, y) and share a market belief p. Although they do not see the current endowment regime z, they know its distribution and independently update their belief about it, as described by the law of motion in formulas (1.15) and (1.19). The denominator in those equations is always greater than zero and the resulting next period belief p' is strictly interior on the interval $(0, 1)$.

Having established a rule for updating the lenders' belief, we can turn to characterizing the bond price schedule as function of the selected debt level b', as well as the state variables y and p. As it is common in the quantitative models of sovereign debt, lenders are assumed to be competitive and risk-neutral, and the resulting equilibrium bond price is such that they make zero profit in expectation (corresponding to their imperfect information). The bond price function is

$$q\left(b', y, p\right) = \frac{1}{1+r^*}\left(\sum_{z'}\left[p\ \pi(z'|z_H) + (1-p)\ \pi(z'|z_L)\right]\int f_{z'}(y', y)\left[1 - d(\mathbf{s}')\right] dy'\right) \tag{1.21}$$

where $\mathbf{s}' = \left(b', y', p'(y, y', p)\right)$, $d(\cdot)$ is the optimal default decision derived in equation (1.20) and r^* is the risk-free rate of interest.

Concluding this section, Definition 1 introduces the standard concept of a Markov Perfect Bayesian Equilibrium. The posterior beliefs $p'(\cdot)$ must be specified for all agents in all states, and the agents' best responses must belong to the set of stationary Markov strategies.

Definition 1 *A Markov Perfect Bayesian Equilibrium for this economy consists of the government value functions $v^r(\mathbf{s})$, $v^d(y,p)$ and policy functions $c(\mathbf{s})$, $b'(\mathbf{s})$, $d(\mathbf{s})$; and the bond price schedule $q(b',y,p)$ such that:*

1. *Policy function d solves the government's default-repayment problem (1.13).*
2. *Policy functions $\{c, b'\}$ solve the government's consumption-saving problem defined in (1.16).*
3. *Bond price function q is such that the lenders make zero expected profit (subject to their imperfect beliefs).*

1.5 Quantitative analysis

In this section I calibrate the model to Portuguese data and discuss its basic mechanics. As an empirical test and the main result of this paper, I use the calibrated model to simulate an actual debt crisis episode and discuss how its predictions depend on the assumption of incomplete information, relative to a full information benchmark as well as the single-regime variants discussed in Section 1.3.

1.5.1 Data

I use the data for Portuguese economy as the case study for the theory developed in this paper.[21] Quarterly data for real GDP are taken from OECD and cover the period 1974:Q1-2014:Q4. Consumption and current account series, also from OECD, span the time

[21] The model could also be calibrated to other European economies discussed in the empirical part. The Portuguese episode is the most clear-cut case however, as it does not coincide with other major economic events such as the banking crisis in Ireland or the introduction of the OMT (Outright Monetary Transactions) program in the summer of 2012, at the height of the debt crises in Italy and Spain.

frame 1995:Q1-2014:Q2. Interest rates on government bonds are from Bloomberg (1995:Q1-2014:Q4), while the annual debt service data is acquired from The Economist Intelligence Unit (1995-2007). Historical GDP forecasts are taken from the OECD Economic Outlook (2008-2014).

1.5.2 Calibration of the output process

I start by bringing to the data the bimodal endowment process introduced in (1.9). Calibrating this equation is by far the most significant challenge in the present paper, as it determines markets' speed of learning about potential regime switch. Recall that in this paper I interpret the high regime as "normal" times, while the low regime should be thought of as a major depression. However, because we do not have enough historical data to account for such episodes it is not a valid approach to estimate a Markov-switching AR(1) process based on the available national accounts data. The result of such estimation will not be capable of matching the size and frequency of large depressions, such as the one we have observed for European economies since 2008. Instead, the obtained low regime will represent one of the regular recessions Europe has experienced in the recent decades. I formally show this in Appendix A by estimating equation (1.9) with a variant of the expectation maximization algorithm of Hamilton (1989).

Instead, in this paper I proceed to calibrating equation (1.9) in three steps. First, I fix the probabilities of switching between regimes based on the historical experience. By all accounts, the recession in southern European countries has been the worst since the Great Depression.[22] This gives us roughly 70 years or 280 quarters of high regime duration.[23] Conversely, the Great Depression lasted for about 10 years, or 40 quarters, which I use to calibrate the expected low regime duration. The resulting probabilities of staying in the high and low regimes are therefore 0.9964 and 0.975, respectively.[24]

[22] For Portugal, the case study of this paper, this historical reference appears fully justified. It comes close to satisfying the defining criteria of a great depression established by Kehoe and Prescott (2002). Also Reis (2013) compares the Portuguese episode to the US Great Depression and Japan's lost decade.

[23] I ignore World War II in my calculations as the model is not designed to account for such events.

[24] The exact specification of these probabilities is not crucial for the results because, given their predetermined values, I use another data source to discipline the size of the depression. What matters is to capture

As a second step, I use historical GDP data to calibrate the persistence and variance parameters of the AR(1) process. The unconditional mean in the high regime μ_H is normalized to zero. Then, for every given choice of the low regime mean μ_L (think of it as a generic step in the estimation algorithm), I use Maximum Likelihood to estimate ρ and η in (1.9) on the GDP data from 1974:Q2 to 2014:Q4, assuming that the regime switches from high to low in 2008:Q3.[25] The data is in constant prices and the linear trend is removed.[26] The resulting estimates for persistence and variance of the output process (for the optimal choice of μ_L, to be discussed below) are 0.953 and 0.011, respectively.

Finally, it remains to select the unconditional mean of the low regime. This number must represent the markets' belief about the depth of the ensuing depression in years 2008-2014. To capture this information, I use the historical GDP forecast data from OECD's Economic Outlook. The data are released twice annually (spring and fall) in the form of projected growth rates for the next two years. I take the two-year-ahead forecast[27] from every fall edition of the Economic Outlook between 2008-2014, convert it into a prediction about GDP level, take logarithm and remove the previously estimated trend (as to make the forecasts directly comparable to the model's endowment).

Next, I generate corresponding forecasts using the PSI variant of my model[28] for every given set of parameters of the output equation. I feed in the sequence of actual GDP observations for Portugal between 2000 and 2014 and create a sequence of corresponding

the right order of magnitude - the low regime ought to be rare and long enough so that it clearly distinguishes from a regular economic downturn. On the other hand, it should be feasible within a generation's lifetime.

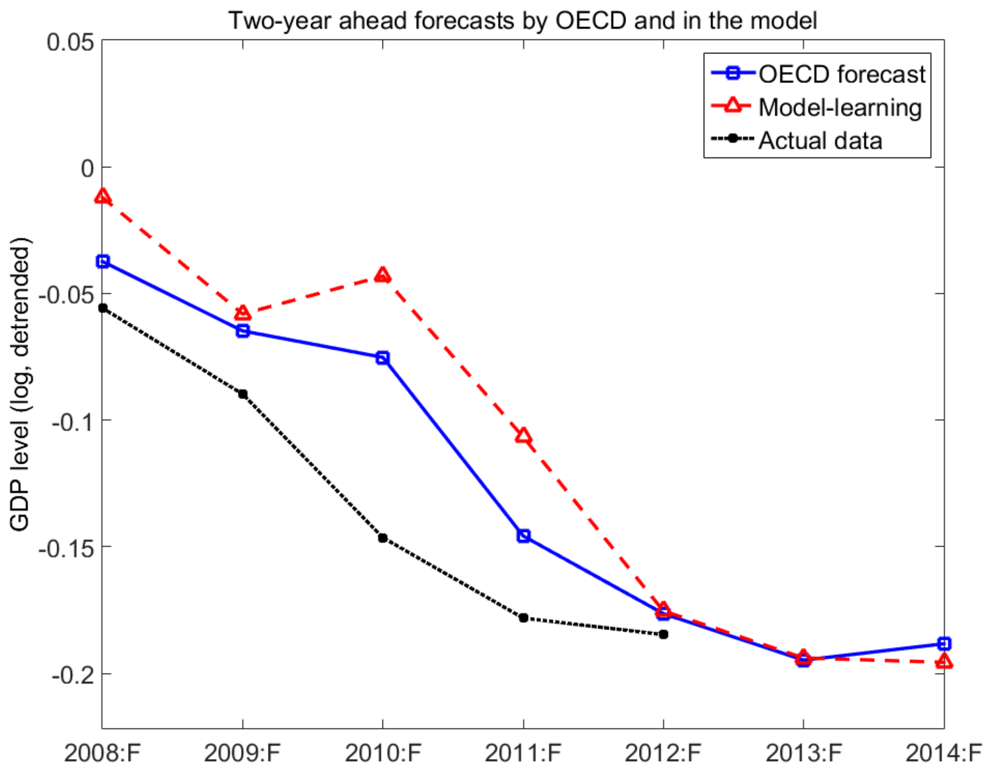
[25]While the precise beginning of the crisis is not indisputable, it is commonly associated with the bankruptcy of Lehman Brothers in September 2008. Around that time we can also observe the beginning of a slump in GDP series for most European countries.

[26]Because of the relatively long sample, I follow Bai and Perron (1998) to identify a statistically significant structural break in the growth rate of Portuguese economy in 1999:Q4. The estimated quarterly growth rate is 0.8% prior to 1999:Q4 and drops to 0.39% ever since.

[27]This is the maximum horizon of forecasts published by OECD and the European Commission. Even though the IMF publishes its forecasts for up to five years ahead, I use the shorter horizon to ensure consistency of the data across the three sources.

[28]As it will be shown in section 1.5.5, the models with symmetric and asymmetric information feature almost the same path of beliefs over time, so this assumption does not affect the results of estimation.

beliefs that result from Bayesian updating. Then, for every third quarter between 2008 and 2014 (the fall edition of OECD's Economic Outlook is usually published in November) I generate a sequence of model-implied quarterly forecasts given the current belief and GDP level, and convert them into annual levels. I finally choose the mean μ_L so to minimize the sum of squared differences between the forecasts from the data and those implied by the model.



Note: Each point on the graph represents an annual detrended log-GDP level for two years ahead (for example, 2008:F corresponds to the GDP level in 2010). The solid blue line plots the actual OECD forecasts, while the dashed red line denotes the ones generated by the calibrated model. Additionally, the dashed-dot black line shows the actual realized data that the corresponding forecasts refer to (only available until 2012:F, when the prediction for year 2014 was made).

Figure 1.9: Historical GDP forecasts: OECD- and model-generated predictions.

Figure 1.9 depicts the result of matching model-generated forecasts to the historical ones published by OECD. As can be noticed, the match is relatively close for the first two years of the crisis, and almost perfect for the most recent three years. For the middle two years of the crisis, 2010 and 2011, the two forecasts diverge in that the model agents are too optimistic about the recovery. This distinction is reasonable however given a very simple two-regime structure of the model. It can be implied from the forecasts data that in 2010 and 2011 the OECD was predicting the economy to stay in a persistent recession of moderate magnitude. The model is not capable of replicating this because the only two regimes available are "normal times" or "major depression". This divergence is acceptable however, because what really matters for calibration of the model is to capture the magnitude of depression anticipated by the markets. This goal is achieved by selecting $\mu_L = -0.253$.

The use of historical GDP forecast data is relevant for the story presented in this paper because it conveys well the process of markets' learning about regime switch. Figure 1.9 also shows that especially in the middle part of the crisis (years 2009-2011) the markets' expectations about future growth turned out to be way too optimistic compared to the actually realized data. This discrepancy disappears with the 2012 forecasts which matches the factual 2014 level of GDP almost perfectly. Those mistaken forecasts at the height of the crisis have recently led the OECD to publish a working paper to evaluate the source of errors. In a "Post Mortem", Pain *et al.* (2014) write:

> *GDP growth was overestimated on average across 2007-12, reflecting not only errors at the height of the financial crisis but also errors in the subsequent recovery. (...)*
>
> *The repeated assumption that the euro crisis would dissipate over time, and that sovereign bond yield differentials would narrow, has been a more important source of error. (...)*
>
> *The OECD was not alone in finding this period particularly challenging. The profile and magnitude of the errors in the GDP growth projections of other international organisations and consensus forecasts are strikingly similar.*

In their *ex post* reflection, the OECD points to the repeated expectation of a swift recovery as the main source of forecast errors. This can precisely be interpreted as learning about the regime switch in my model. The uninformed agents must first observe a number of bad output realizations to start believing that the economy has switched to an extremely rare disaster regime. In order to not appear as the only culprit, the OECD also emphasizes that the excessively optimistic forecasts have been common among other influential forecasters associated with international organizations. I verify this claim using the publicly available data of historical forecasts from the European Commission and the International Monetary Fund. In both datasets we can indeed observe a similar pattern of forecast accuracy over time as presented in Figure 1.9. For this reason, I take the OECD data as proxy of the market-wide expectation about future growth in my model. Table 1.3 summarizes the results of calibrating the stochastic endowment process specified in equation (1.9).

				Transition Prob.	
Regime	Mean μ	Persistence ρ	St. dev. η	Low	High
Low	-0.253	0.953	0.011	0.975	0.025
High	0.00	0.953	0.011	0.004	0.996

Table 1.3: Parameters of the regime-switching endowment process

1.5.3 Functional forms and calibration

To select the remaining structural elements of the model I mostly follow the general trends in the literature. A representative household's utility is a CRRA function of the form $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, with the risk aversion parameter set at the standard level of 2. The risk-free interest rate is set equal to 1% (quarterly value) and the probability of re-entry after default is fixed at 10% following Aguiar and Gopinath (2006). The output cost of default is specified as $h(y) = \max\left\{0, -d_0 y + d_1 y^2\right\}$ following Chatterjee and Eyigungor (2012). The two parameters of the default penalty function, d_0 and d_1, together with the discount factor β are selected using the Simulated Method of Moments. The economy's income path in years 1995-2014 is simulated 1000 times, starting from the GDP and debt levels observed in 1995:Q1 and under a restriction that the regime switches from High to Low in 2008:Q3. The targeted moments of the Portuguese economy are: average debt service-to-GDP ratio in 1995-2007[29] of 8.47%; the correlation of log GDP and log consumption in 1995-2014 of 0.953; and standard deviation of the bond spread in 1995-2014 of 2.77%. Finally, the two parameters of the EU-imposed debt limit are chosen to match the coefficients from the regression of formula (1.18), assuming that the constraint binds, on the annual debt service-to-GDP series from 1995-2007. Table 1.4 presents a summary of the parameter values used to compute the model's equilibrium under the two variants of the information structure considered.

Symbol	**Meaning**	**Model FI**	**Model PSI**	**Source**
σ	Risk aversion	2	2	Literature
r^*	Risk-free rate	0.01	0.01	Literature
θ	Re-entry probability	0.10	0.10	Literature
d_0	Default cost par.	-1.058	-0.808	Calibration
d_1	Default cost par.	1.138	0.910	Calibration
β	Discount factor	0.921	0.945	Calibration
$\bar{b}$	Debt constraint par.	0.005	0.005	Estimation
ξ	Debt constraint par.	1.014	1.014	Estimation
Calibration targets		**Model FI**	**Model PSI**	**Data**
corr(cons,GDP)		0.94	0.96	0.95
st. dev. (spread)		2.66	1.69	2.80
mean debt serv./GDP ('95-'07)		9.21	8.66	8.47

Note: The consumption data is detrended using a common GDP trend.

Table 1.4: Calibration of structural parameters of the model

[29]This time frame is chosen so that the government's debt policy following the 2008 recession is an endogenous outcome of the model, not a target in calibration.

1.5.4 Characterization of the equilibrium

In the following section I first characterize some of the key properties of the equilibrium, and then show how the model's simulated behavior compares with actual data. The model is solved numerically by value function iteration on a computer cluster with 32 nodes. The detailed algorithm for computing the equilibrium is described in Appendix B.

It is instructive to begin with examining the policy functions of the PSI model. Figure 1.10 presents the default and debt policy functions for different levels of prior belief. In line with basic intuition, the left hand side panel confirms that higher belief about being in the good regime induces the government to default in smaller number of states on the income-debt grid. This relationship is strictly monotonic in the level of prior belief. A natural corollary to this result is that international lenders will offer higher prices to the government they assign a larger probability of being in the high regime. The right hand side panel of Figure 1.10 shows that higher prior belief leads the government to borrow more. In this model, agents are impatient and would rather consume today than tomorrow. When making their debt decisions though, they need to weigh their impatience against the expected income level in the future. A higher chance of being in economic depression next period implies that the government must restrict its consumption today and reduce foreign debt in order to decrease the probability of defaulting tomorrow and to secure a high bond price today. Consequently, higher market belief has a strictly monotonic, increasing effect on the optimal debt level. On the other hand, notice that this relationship is not necessarily linear and depends on the current income level.

Figure 1.11 plots the prices of government bonds as functions of the next period debt allocation, for three different fixed levels of the lenders' belief. It can easily be noticed that information about the regime is important in determining the default risk and leads to large differences in the offered bond prices. The highest dash-dot black line represents the prices in a state where markets know the economy is in the high regime. By contrast, the lowest solid blue line represents prices of the government whose type is strongly believed to be low. Because the risk of default is much higher during a major depression, this government is not surprisingly offered low prices for its debt. Finally, the middle dashed red line denotes bond prices in a middle-belief state, where the lenders unsure about the

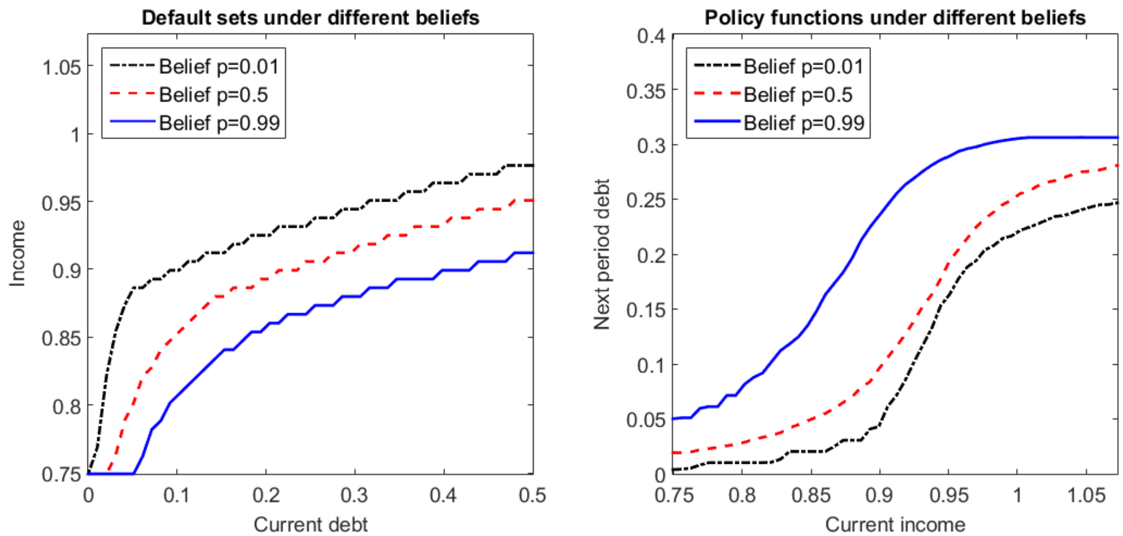


Figure 1.10: Default sets and bond price policy functions in the symmetric information world

type of government issuing bonds. In that case, they rely on their own assessment of default risk and offer prices lying somewhere in between the two extreme cases of full knowledge of the regime.

1.5.5 Simulating the European debt crisis

In this section, I use the two calibrated model to analyze the timing and pattern of events during the debt crisis in Portugal. I start with two benchmark cases, a full-information version of the current model, as well as the Arellano (2008) model with an added debt constraint. I then present a main result with the PSI variant and discuss the main difference.

Full Information case

I start by feeding in the actual detrended GDP observations for Portugal into the benchmark Full Information version of the model. The upper panel of Figure 1.12 presents the simulated evolution of the debt level and the bond spread in that world. The economy

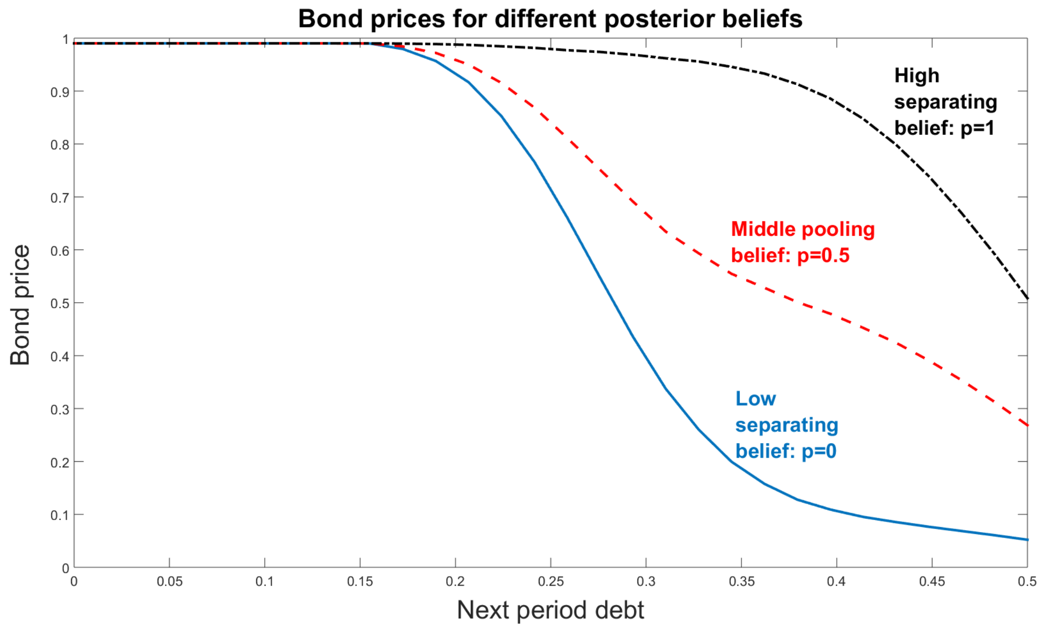


Figure 1.11: Bond price as function of next period debt for different beliefs of the lenders

is started with year-2000 level of debt service observed in the data and gradually accumulates it throughout the 2000s, up to the level of around 17%. At the same time, the bond spread is very close to zero in all periods prior to 2008, implying a zero probability of default during that time. When the crisis first hits in the third quarter of 2008, the bond spread jumps up, while the government drastically reduces its debt. As can be noticed then, the Full Information version of the model delivers predictions opposite to what we observe in the data for the European debt crisis. The key assumption of the analysis is that the regime switches from High to Low in the third quarter of 2008, and both the government and the lenders are aware of this. This information structure is confirmed in the lower panel which shows that the market belief is equal to 1 before and 0 after that period. Overall, the result displayed in Figure 1.12 should be taken with a grain of salt. It is based on an unrealistic assumption that all agents in the economy know exactly about the adverse regime switch. A much more plausible benchmark is the actual single-regime model of Arellano (2008) (augmented with an exogenous debt limit (1.18) and recalibrated to facilitate a direct comparison). It is studied in the following subsection.

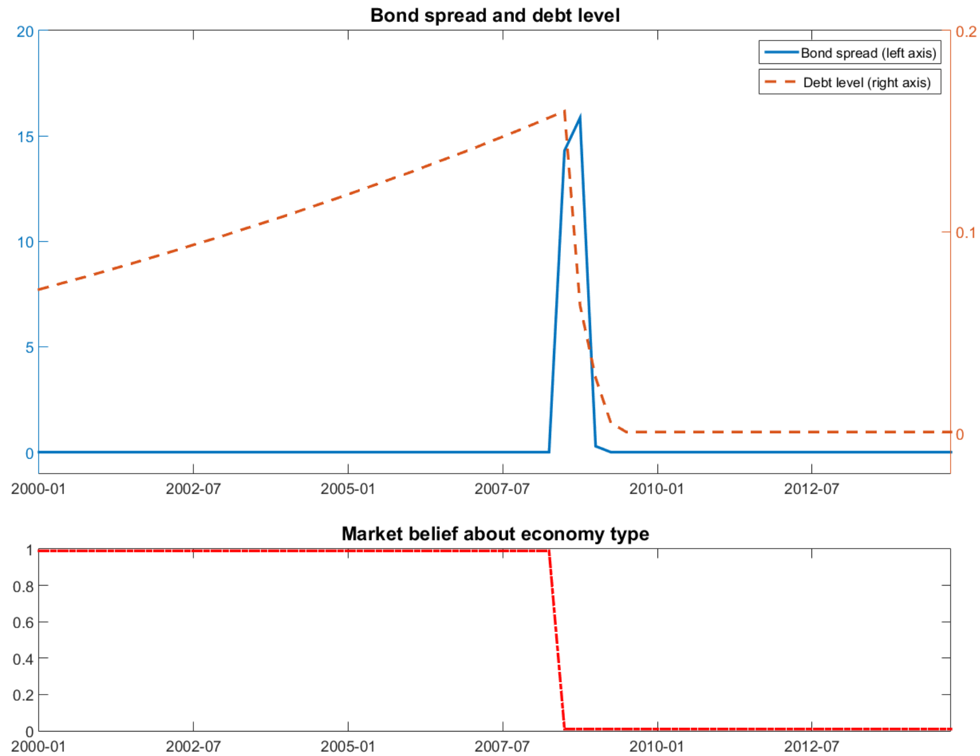


Figure 1.12: Simulated debt crisis in the Full Information model

One-regime case with exogenous debt constraint

To make a better benchmark case, I now use the model as in Arellano (2008), where the government's problem is subject to an exogenous debt limit of the form $b' \leq \bar{b}y + \xi b$. In contrast to the analysis in Section 1.3.1, I now use the same calibration algorithm as for the main model, albeit keeping the AR(1) parameters from Table 1.1 unchanged. In particular, I fix the probability of re-entering the credit markets at 10% and I use the augmented specification of the default penalty, $h(y) = \max\left\{0, -d_0 y + d_1 y^2\right\}$ of Chatterjee and Eyigungor (2012). The parameters of this function, d_0 and d_1 are calibrated jointly with the discount factor to match the three moments: debt-service-to-GDP ratio of 8.47% for 1995-2007, standard deviation of the bond spread for 1995-2014 fo 2.796 and the correlation fo GDP and consumption for this period of 0.95. Parameters of the debt constraint, $\bar{b}$ and ξ, are chosen to match the debt accumulation path of Portugal prior to the crisis (as described in more details in the main text). Table 1.5 summarizes the parameter values.

Parameter	Value	Meaning	Source
σ	2	Risk aversion	Literature
r^*	0.01	Risk-free rate	Literature
θ	0.10	Re-entry probability	Literature
d_0	-1.041	Default cost par.	Calibration
d_1	1.119	Default cost par.	Calibration
β	0.967	Discount factor	Calibration
$\bar{b}$	0.005	Debt constraint par.	Estimation
ξ	1.014	Debt constraint par.	Estimation
ρ	0.9816	Shock persistence	Estimation
η	0.0108	Shock volatility	Estimation
Calibration targets		**Model**	**Data**
corr(cons,GDP)		0.98	0.95
st. dev. (spread)		2.20	2.79
mean debt serv./GDP ('95-'07)		8.81	8.47

Table 1.5: Calibrated parameters in the one-regime model with debt constraint

I then use the model to simulate a debt crisis episode in Portugal, exactly in the same way as in Section 1.5, and I plot it in Figure 1.13. As can be noticed, the economy accumulates debt gradually in years 2000-2008. When the recession first hits in 2008:Q4 and 2009:Q1, the bond spread jumps sharply on impact, forcing the government to reduce its short-term debt by a half. After a mild recovery of 2009-2010, the interest rate surges once again at the end of 2011 (by roughly the same magnitude as at the beginning) and the government eventually reduces its debt almost to zero, reflecting the fact from the data that the Portuguese government has been holding negligible amounts of short-term debt since the height of its crisis in 2012. In conclusion, the model of Arellano (2008) is not well suited to analyze the European debt crisis episode, even if we augment it by an exogenous debt constraint to discipline the accumulation of debt prior to 2008.

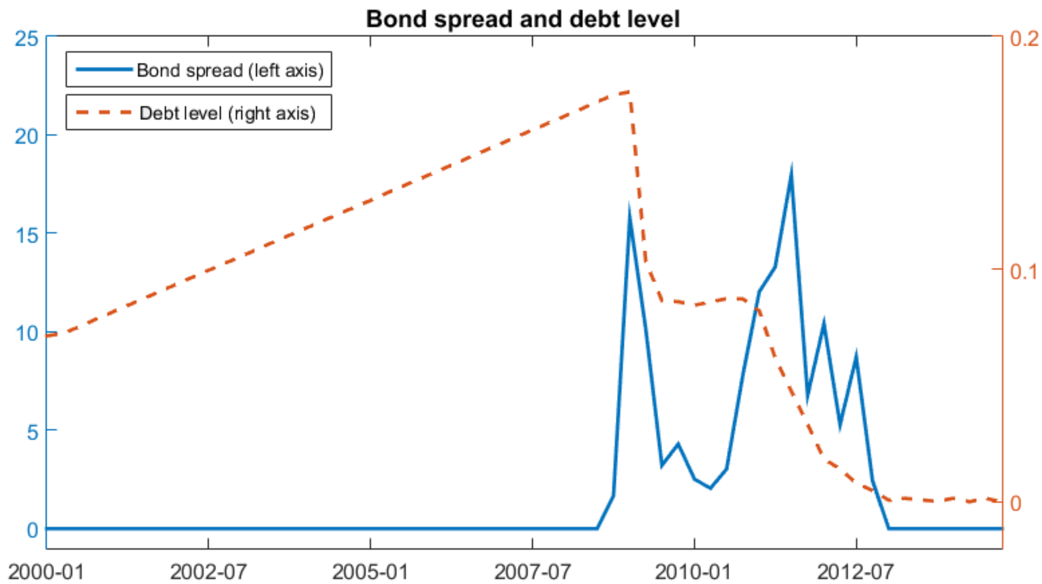


Figure 1.13: Simulated debt crisis in the one-regime model with exogenous debt limit

Partial Symmetric Information case

Now I conduct the simulation under the assumption that the market and the government share a common belief about the output regime in which the economy is operating. The upper panel of Figure 1.14 presents the evolution of the debt level and the bond spread in the model with incomplete symmetric information. The economy is once again started with its initial debt service level and gradually accumulates it throughout the 2000s, up to the level of around 17%. The bond spread is once again equal to zero in all periods prior to 2008. However, in this world when the crisis begins and the regime switches to Low, the market belief only drops partially, which can be viewed in the lower panel of Figure 1.14. As a result, the bond spread only increases slightly, while the government reduces its debt by a smaller amount as compared with the FI case, and a similar amount to the case of one-regime AR(1) process with exogenous debt constraint. Subsequently however, when the bad income shocks hit the economy again at the beginning of 2011, the markets and the government finally start to realize that this prolonged downturn must in fact be driven by the underlying regime switch, and reduce their belief correspondingly. As a result, the bond spread goes on a sharp increase, much higher in magnitude than initially in 2009:Q1.

This exercise shows therefore that market learning is a mechanism capable of delivering the delay in the outset of a debt crisis, as it happened for the peripheral European economies during 2008-2014.

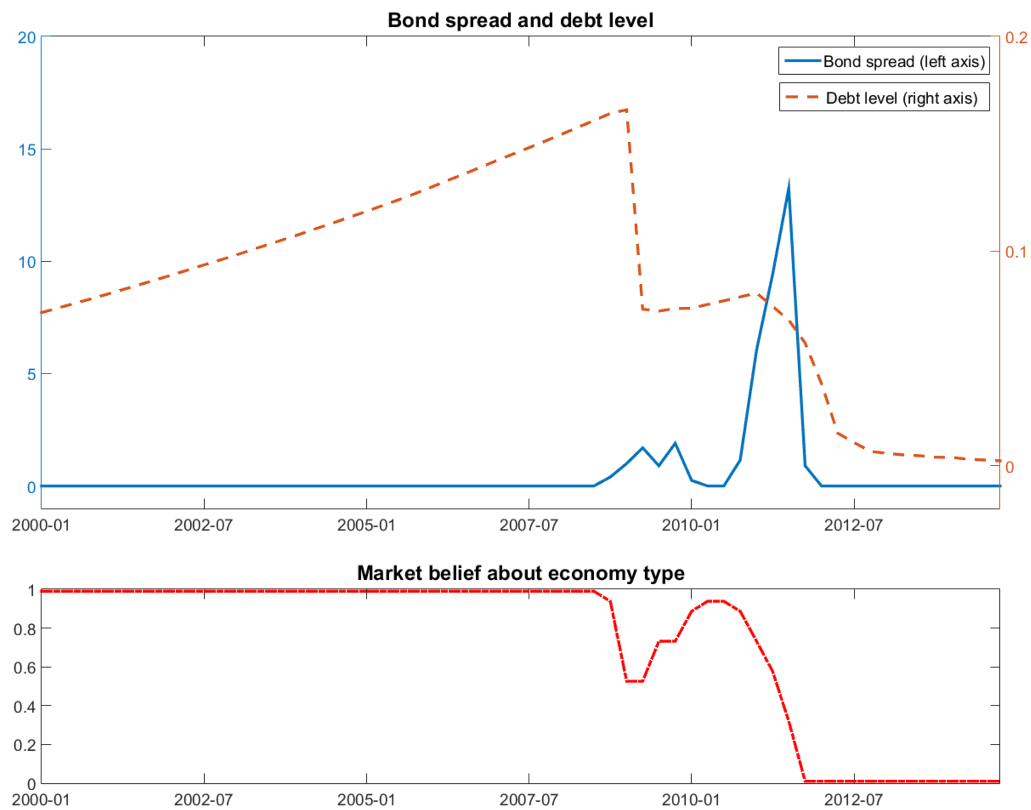


Figure 1.14: Simulated debt crisis in the Partial Symmetric Information model

Chapter 2

Debt Crises and Asymmetric Information

2.1 Introduction

In Chapter 1 I discussed the model of incomplete information, in which all the market participants (government and the foreign lenders) learn symmetrically about the underlying regime shift during the Great Recession. While the main result presented in Figure 1.14 delivers a correct pattern for the dynamics of interest rate spreads, it has the obvious shortcoming in predicting a counterfactual sharp debt reduction upon receiving the negative shocks in 2008-2009. A natural extension of that model is to consider an information structure that is asymmetric in that the government knows the current regime exactly, while the foreigners do not. Instead, they observe the country's debt levels and try to extract the hidden information from them. As I will show in the present chapter, such a model is capable of generating much closer predictions for the dynamics of debt level during the crisis.

The assumption of the government's informational advantage over the markets is empirically justified in two ways. First, governments often have access to confidential (or at least very inaccessible) data about the performance of key industries in the economy. This has in particular been a factor during Ireland's recent debt crisis. While the markets learned very late about the bad condition of the country's banking sector, the government arguably had access to this information beforehand and could have anticipated the ensuing crisis. Second, and more importantly, over the last 15 years the bonds of European governments have been increasingly accumulated by non-resident holders, often located in distant financial centers around the world. As Andritzky (2012) shows, the fraction of non-resident bondholders reached a maximum in 2007 for the four European economies considered and varied between 50% for Greece, up to over 90% for Ireland. It is reasonable to assume that international investors who hold the bonds of many different governments in their portfolios, have less precise expectation about the evolution of output than the governments themselves. In addition, over the last two decades we have observed a steady increase in the fraction of Exchange Traded Funds among financial market participants around the world. These so-called index-trackers are passive investors who do not conduct fundamental analysis of assets in their portfolio and instead rely on the index-wide composition. However, upon receiving particularly bad news about the expected default probability on some of the bonds in their portfolio, ETFs may decide to sell them and thus contribute to the ignition of the crisis.

While a government clearly has some degree of informational advantage over the markets, the assumption in this paper that it can observe the current endowment regime exactly is naturally an extreme one. However, as it is argued in section 1.4.4, this assumption is necessary to develop a tractable method of computing a signaling equilibrium. In an alternative setup, the government would only receive a noisy private signal about the underlying regime switch. The problem is that in such a framework the state variable z would be reinterpreted as the government's private belief with continuous support in $(0, 1)$. As a result, the model would feature a continuum of government types unobserved by the markets. While my current solution method described in the computational appendix does not extend to a continuum of types, it is an important area for future research.

The remainder of this chapter is structured as follows. Section 2.2 introduces the model with asymmetric information and discusses the solution method. Section 2.3 presents the quantitative analysis. Section 2.4 discusses the role of bailouts and main policy implications coming from the model. Section 2.5 concludes.

2.2 Model with asymmetric information

In this section, I briefly present the model with asymmetric information and highlight the main differences relative to the formulation shown in Chapter 1. I also discuss the equilibrium concept and solution method adopted for the purpose of numerical analysis in the following section.

2.2.1 Recursive formulation

In the following section briefly go over the modified equations of the model and emphasize the differences with corresponding specifications from Chapter 1. As before, the vector of aggregate state variables that are common knowledge is $\mathbf{s} = (b, y, p)$. Variable z is now private information of the government and will therefore be separated from the other ones by semicolon. Notice that this general specification can be collapsed easily to either of the former two variants of the model, FI or PSI, by removing one of the state variables (p or z, respectively).

Government The government that is current on its debt obligations has the general value function given by

$$v^0(\mathbf{s}; z) = \max_{d \in \{0,1\}} \left\{ (1-d) v^r(\mathbf{s}; z) + d v^d(y, p; z) \right\} \tag{2.1}$$

A sovereign who decides to default ($d = 1$) is excluded from international credit markets and has the probability θ of being readmitted every subsequent period. The associated default value is given by

$$\begin{aligned} v^d(y, p; z) = & u\left(h(y)\right) + \beta \sum_{z' \in \{z_L, z_H\}} \pi(z'|z) \\ & \int f_{z'}(y', y) \left[\theta v^0(0, y', p'; z') + (1-\theta) v^d(y', p'; z') \right] \, dy' \end{aligned} \tag{2.2}$$

subject to the laws of motion for the lenders' beliefs

$$p'(y, y', \widetilde{p}) = \frac{f_{z_H}(y', y) \times [\widetilde{p}\, \pi(z_H|z_H) + (1-\widetilde{p})\, \pi(z_H|z_L)]}{\sum_{z'=z_L, z_H} f_{z'}(y', y) \times [\widetilde{p}\, \pi(z'|z_H) + (1-\widetilde{p})\, \pi(z'|z_L)]} \tag{2.3}$$

$$\widetilde{p} = \psi^d(\mathbf{s}) \tag{2.4}$$

In equation (2.2), ψ^d is a function that depends on today's state and describes the posterior belief of the lenders in response to an observed default. The next period belief p', described in equation (2.3), depends on the current and future income realization, as well as the current period posterior belief $\widetilde{p}$. It is analogous to the Bayes' formula presented in equation (1.15), but it takes into account the possibility of a change in belief as a result of the signaling interaction between the government and the lenders.

The value of the government associated with repayment of debt is given by

$$v^r(\mathbf{s};z) = \max_{c,b'} \left\{ u(c) + \beta \sum_{z' \in \{z_L, z_H\}} \pi(z'|z) \int f_{z'}(y', y) v^0(\mathbf{s}'; z')\, dy' \right\} \tag{2.5}$$

subject to

$$c = y - b + q(b', y, \widetilde{p})b' \tag{2.6}$$

$$b' \leq \bar{b}\, y + \xi b \tag{2.7}$$

$$\widetilde{p} = \psi^r(b'; \mathbf{s}) \tag{2.8}$$

$$p'(y, y', \widetilde{p}) = \frac{f_{z_H}(y', y) \times [\widetilde{p}\, \pi(z_H|z_H) + (1-\widetilde{p})\, \pi(z_H|z_L)]}{\sum_{z'=z_L,z_H} f_{z'}(y', y) \times [\widetilde{p}\, \pi(z'|z_H) + (1-\widetilde{p})\, \pi(z'|z_L)]} \tag{2.9}$$

where the laws of motion for the lenders' beliefs are given by formulas (2.8)-(2.9).[1] Because of the signaling interaction, the posterior belief $\widetilde{p}$ depends not only on the current state, but also on the next period bond allocation selected by the government. As a result, in this model the bond price q varies not only with the level of future borrowing and the current income (as it typically does in the models of sovereign debt), but also with the lenders' posterior belief. This is due to the fact that markets attempt to read the signal embedded in the bond allocation and map it to one of the government types. Note that, by contrast, in the PSI version of the model both posterior functions ψ^d and ψ^r are always equal to the prior p.

With the updated value functions of the government in place, the default decision is analogous to the one derived in equation (1.20), except for the fact that it now also depends on the state variable z.

[1]Note that equations (2.3) and (2.9) are identical as they represent the same Bayesian updating mechanism.

International Lenders Under asymmetric information, the government's debt choice conveys a signal about the economy's type and international lenders therefore gain another channel of updating their belief. In equilibrium, the posterior belief function is given by applying Bayes' rule

$$\psi^r(b'; \mathbf{s}) = \frac{Prob(b'|z = z_H) \times p}{Prob(b'|z = z_H) \times p + Prob(b'|z = z_L) \times (1 - p)} \tag{2.10}$$

Notice that formula (2.10) may return one of the three possible values for posterior belief, $\{0, 1, p\}$. In the case of 0 or 1, all uncertainty that period is resolved and we have a separating equilibrium - the lenders are able to associate the observed bond allocation with a government type exactly. However, the allocation might also be pooling in the sense that both types of government settle on the same debt amount and the lenders learn nothing from observing that choice. In such case, the posterior belief ψ^r is simply equal to the prior. Notice again that under the assumption of incomplete but symmetric information, the government itself is unaware of the current regime and equation (2.10) trivially boils down to $\psi^r(b'; \mathbf{s}) = p$.

The bond price schedule is now a function of any fixed posterior belief $\widetilde{p}$, in addition to next period debt b' and current income y. The lenders are risk-neutral and the resulting equilibrium bond price reflects their perceived default probability of the government. The bond price function is

$$q\left(b', y, \widetilde{p}\right) = \frac{1}{1 + r^*} \left(\sum_{z'} \left[\widetilde{p}\, \pi(z'|z_H) + (1 - \widetilde{p})\ \pi(z'|z_L) \right] \int f_{z'}(y', y) \left[1 - d(\mathbf{s}'; z')\right] dy' \right) \tag{2.11}$$

where $\mathbf{s}' = \left(b', y', p'(y, y', \widetilde{p})\right)$.

A novel aspect of the model with asymmetric information that is worth emphasizing is the dependence of bond prices on the level of borrowing in both the next period (as in the standard debt models) as well as the current one. This occurs through the posterior $\widetilde{p}$ which in turn is a function of the current debt level, as specified in equation (2.8). Consequently, even if we hold the typical arguments b' and y fixed, the bond price will be different for various levels of the current debt b, depending on whether the government decides to separate or pool.

In solving the model with asymmetric information, the problem of adverse selection arises due to the fact that the government knows its own type, and the lenders do not. In this setup, a contract is a four-tuple $\{b'_H, b'_L, \psi^r(b'_H;\mathbf{s}), \psi^r(b'_L;\mathbf{s})\}$ and consists of the bond allocations offered by the lenders to each government type, along with the resulting posterior beliefs. For notational convenience, define the lifetime utility of type-i government, $i \in \{H, L\}$, of choosing a debt amount b' under a fixed posterior belief $\widetilde{p}$ as

$$W_i(\mathbf{s}; b', \widetilde{p}) = u\big(y - b + q(b', y, \widetilde{p})b'\big) + \beta \sum_{z' \in \{z_L, z_H\}} \pi(z'|z_i) \int f_{z'}(y', y) v^0(\mathbf{s}'; z')\, dy' \quad (2.12)$$

Because the government can choose which offers to accept, we can restrict our attention to the contracts that are incentive compatible in the sense that

$$W_L\big(\mathbf{s}; b'_L, \psi^r(b'_L;\mathbf{s})\big) \geq W_L\big(\mathbf{s}; b'_H, \psi^r(b'_H;\mathbf{s})\big) \quad (2.13)$$

$$W_H\big(\mathbf{s}; b'_H, \psi^r(b'_H;\mathbf{s})\big) \geq W_H\big(\mathbf{s}; b'_L, \psi^r(b'_L;\mathbf{s})\big) \quad (2.14)$$

The incentive compatibility constraints (2.13)-(2.14) guarantee that, when solving its problem, the government of type i will not choose an offer designed for the other type. Notice that the High type's constraint will never bind in this setup because it cannot have incentive to pretend it is a Low type (due to a strictly monotonic relationship between a type and the probability of default, implied by the assumed income process in (1.9)). By contrast, the Low type's constraint (2.13) may or may not bind for the optimal contract. If it does not, then we have a separating allocation in which the two government types select different debt amounts and fully reveal their identities to the lenders. If the constraint binds however, we have a pooling allocation in which $b'_H = b'_L$ and $\psi^r(b'_H;\mathbf{s}) = \psi^r(b'_L;\mathbf{s}) = p$, and as a result the lenders are unable to extract any information from the observed choice of the government.

To solve for the optimal contract, we need to maximize the government's lifetime utility subject to incentive constraints (2.13)-(2.14) and the lenders' zero profit conditions. However, according to the classic argument in the microeconomic literature by Rothschild and Stiglitz (1976), a Nash equilibrium may not exist in a model of insurance with adverse selection.[2] Intuitively, if the measure of bad types in the economy is small (i.e. prior

[2]The original discussion on the existence of equilibria in models with adverse selection referred to the

belief p is large), a profitable pooling deviation may exist in which the high type agrees to subsidize the low type in exchange for a more desirable debt level than in the fully separating allocation that does not violate the IC constraint (2.13). However, pooling cannot be sustained in an equilibrium of the Rothschild-Stiglitz model as the lenders will have incentive for cream-skimming of the high type, deviating from the optimal pooling allocation. This means that they will try to attract the high type government to increase borrowing by offering a much better price schedule. They will do so taking the behavior of their competitors as given, i.e. under the assumption that the low type will still be offered the current pooling contract. Naturally, this deviation is not sustainable in equilibrium as it will render the pooling contract unprofitable (only the low type will remain to select it) and lead to its subsequent withdrawal. Then, once the low type in turn increases the debt level to mimic the high type, the new contract will itself become unprofitable. As a result, we can construct an infinite sequence of such deviations and no Nash equilibrium exists for this economy.

The literature has come up with several ways to overcome the negative result of Rothschild and Stiglitz (1976). One of the most appealed to solutions is to move away from the perfectly competitive structure of the market and to relax the assumption that any lender takes the actions of his competitors as given. In this spirit, Wilson (1977) proposed the notion of *anticipatory equilibrium* which exists always, even when the Rothschild-Stiglitz equilibrium does not. In this concept, the lenders make a decision to introduce a new contract under the conjecture that other lenders will immediately withdraw their current contracts that have become unprofitable as a result of offering the new one. In doing so, the lenders in the Wilson world are more sophisticated than the ones in the Rothschild-Stiglitz world who naïvely assume that other lenders will not respond to their attempted cream-skimming deviation. Consequently, the anticipatory equilibrium features a pooling allocation that maximizes the high type's lifetime utility whenever the Rothschild-Stiglitz equilibrium fails to exist. By contrast, if the Rothschild-Stiglitz equilibrium does exist then it features the same optimal allocation as the anticipatory equilibrium.[3]

insurance markets. Because the model developed in this paper is effectively an insurance model of debt, I illustrate the problem using the sovereign lending framework.

[3]For an intuitive proof, see Seog (2010, ch. 7).

In order to assure existence of the equilibrium, in this paper I follow Wilson (1977)[4] and the subsequent generalizations of this approach of Miyazaki (1977) and Spence (1978) by restricting the set of admissible deviations that the lenders may attempt. Intuitively, a government is only offered a high-type-revealing contract if at the selected allocation b' the low type has no incentive to deviate from its own low-type-revealing contract. If a high type government chooses b' such that the low type would find it optimal to mimic its behavior and get $\psi^r(b';\mathbf{s}) = 1$ as a result, the lenders will respond with a pooling contract, i.e. $\psi^r(b';\mathbf{s}) = p$. More specifically, the Miyazaki-Wilson-Spence (MWS)[5] contract can be found as a solution to the following maximization problem

$$\max_{b'_H, b'_L} W_H\big(\mathbf{s}; b'_H, \psi^r(b'_H;\mathbf{s})\big) \tag{2.15}$$

subject to

$$\begin{aligned} &\text{Formulas} \quad (2.10), (1.21), (2.13) \\ &W_L\big(\mathbf{s}; b'_L, \psi^r(b'_L;\mathbf{s})\big) \geq \max_{b'} W_L\big(\mathbf{s}; b', \widetilde{p} = 0\big) \end{aligned} \tag{2.16}$$

The first three constraints have been defined earlier in this section. Equation (2.10) specifies that the posterior belief function based on the observed government's choice is derived with Bayes' rule. Equation (1.21) defines the bond price as a function of the posterior belief using the lenders' zero profit condition (corresponding to their possibly imperfect knowledge about the government's type). Inequality (2.13), is the low type's incentive compatibility constraint. Finally, inequality (2.16) states that the low type must be at least as well-off as with the optimal separating Rothschild-Stiglitz allocation. This condition is a form of participation constraint and is necessary to ensure that the low type would not prefer to deviate from the optimal pooling allocation. When (2.16) holds with equality, the Miyazaki-Wilson-Spence contract corresponds to the Rothschild-Stiglitz contract. When

[4] Similar approach has been adopted by previous related studies, e.g. D'Erasmo (2011). In addition, a growing branch of recent literature modifies the underlying game structure such that a Wilson equilibrium arises endogenously (this can be achieved by adding an additional stage to the extensive form game in which initial contracts can be withdrawn upon observing the contracts offered by competitors); see e.g. Mimra and Wambach (2011) or Netzer and Scheuer (2014).

[5] For a review of this concept, see Seog (2010, ch. 7) or Mimra and Wambach (2014).

it is slack, the resulting MWS allocation exhibits pooling[6] such that the low type is fully insured and subsidized by the high type, while the latter is only partially insured relative to the full-information contract.

The last issue in discussing the lenders' strategy involves specification of their posterior belief in response to an observed default, $\psi^d(\mathbf{s})$. Because default sets are strictly monotonic with respect to the belief (due to the fact that the two regimes only differ in terms of the unconditional mean), when the high type defaults, so does the low type and thus $\psi^d(\mathbf{s}) = p$, i.e. the lenders do not update their prior. However, when the low type defaults, information may be revealed or not, depending on what the high type does in equilibrium. For computational simplicity, I assume that the low type always expects its identity to be revealed following a default. While this assumption may be not true in some particularly bad states, it cannot quantitatively alter the default incentives of the low type. Given the commonly assumed low values of parameter θ (probability of being readmitted to financial markets after default - typically around 0.1), it does not matter much how we model the lenders' reaction to an observed default. This is because during the expected 10 periods of staying in exclusion from financial markets, the lenders' independent belief will have diverged away from the one updated on impact. Notice also that without this assumption, the low type's default value would depend on the current level of debt, greatly increasing the computational burden.

Definition 2 modifies the Markov Perfect Bayesian Equilibrium introduced in Chapter 1. In this equilibrium the posterior beliefs of agents must be specified at all states and for all strategies of other players (including those involving off-equilibrium actions). The agents' best responses must belong to the set of stationary Markov strategies.

Definition 2 *A Markov Perfect Bayesian Equilibrium for this economy consists of the government value functions* $v^r(\mathbf{s}; z)$, $v^d(y, p; z)$ *and policy functions* $c(\mathbf{s}; z)$, $b'(\mathbf{s}; z)$, $d(\mathbf{s}; z)$;

[6]Note that in Miyazaki's original article pooling is not an equilibrium outcome due to the possibility of cross-subsidization among the coexisting types. In my model however, this is not feasible as there simultaneously exists only one government. Consequently, a contract in this framework is what Miyazaki (1977) refers to as a "singleton wage structure" in footnote 21 and pooling is an admissible outcome.

the posterior belief functions under repayment and default $\psi^r(b';\mathbf{s})$ *and* $\psi^d(\mathbf{s})$*; and the bond price schedule* $q(b', y, \psi^r(b';\mathbf{s}))$ *such that:*

1. *Policy function* d *solves the government's default-repayment problem (1.13).*
2. *Policy functions* $\{c, b'\}$ *solve the government's consumption-saving problem defined in (2.15).*
3. *Bond price function* q *is such that the lenders make zero expected profit (subject to their imperfect beliefs).*
4. *Posterior belief functions* ψ^r*,* ψ^d *are updated according to Bayes' rule whenever possible.*

The last part of the above definition requires a word of comment. When the government plays an off-equilibrium action by choosing a debt amount different from the equilibrium policy function, Definition 1 does not specify the resulting posterior belief. However, the solution concept I follow in this paper (Miyazaki-Wilson-Spence) imposes certain restrictions on such beliefs in order to guarantee existence of an equilibrium. In particular, the high-type-revealing belief $\psi^r = 1$ is only allowed for high enough debt amounts (such that the IC constraint (2.13) is non-binding) leading to a least-cost separating equilibrium, as in Cho and Kreps (1987). In addition, the pooling belief $\psi^r = p$ can arise only for debt levels at least as large as the high type's optimal pooling allocation. If any lower debt level is attempted, the lenders are assumed to believe it could only come from the low type.

2.3 Quantitative results

In this section I recalibrate the model to Portuguese data and discuss how the predictions for bond spreads and debt level change under asymmetric information, relative to the variants of the model analyzed in Chapter 1. The numerical algorithm used to compute the equilibrium of this model is described in Appendix B.

I hold the parameters of the output process from Table 1.3 unchanged and follow the strategy for structural calibration described in Section 1.5.3 to bring the model with asymmetric information to the data. Table 2.1 summarizes the obtained parameter values, together with matching of the moments. It can be observed that while the exact parameters are different for the PAI variant, the quality of the model's match to the data is comparable to the FI and PSI variants discussed in Chapter 1.

Symbol	**Meaning**	**Model PAI**	**Source**
σ	Risk aversion	2	Literature
r^*	Risk-free rate	0.01	Literature
θ	Re-entry probability	0.10	Literature
d_0	Default cost par.	-0.763	Calibration
d_1	Default cost par.	0.887	Calibration
β	Discount factor	0.952	Calibration
$\bar{b}$	Debt constraint par.	0.005	Estimation
ξ	Debt constraint par.	1.014	Estimation
Calibration targets		**Model PAI**	**Data**
corr(cons,GDP)		0.96	0.95
st. dev. (spread)		1.73	2.80
mean debt serv./GDP ('95-'07)		8.90	8.47

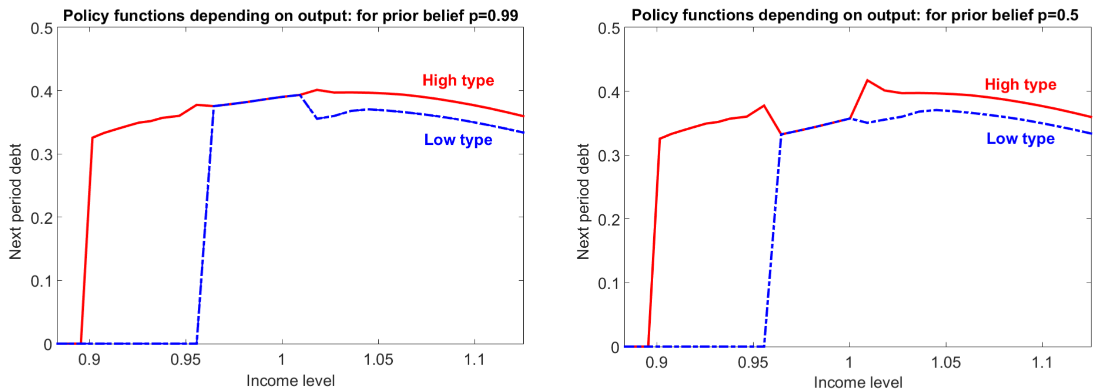
Note: The consumption data is detrended using a common GDP trend.

Table 2.1: Calibration of structural parameters of the PAI model

Following the approach from Chapter 1, I next turn to characterizing the policy function of the present model. The relationship between beliefs and the direction of government's decisions described in Section 1.5 highlights the nature of the single crossing property in my model with asymmetric information. In the previous theories of debt with private information and signaling it is typical that a bad-type borrower takes on more debt, and the good-type borrower can separate itself by saving more. This is because in those models borrower types are characterized by preferences, the good type usually being a more prudent or patient one. This is not the case in the present model. All governments have the same preference parameters and are only differentiated by the regime in which their economy is operating. As a result, a high-regime-economy government expects larger income in the future and can afford more consumption and debt today.

Going back to Figure 1.11 also gives an intuitive interpretation of why governments are reluctant to adjust their debts accordingly at the outset of a major depression. Suppose a high-regime government is running a large debt level of around 0.45 in a separating equilibrium. When the regime switches to low, the government would ideally like to reduce the debt by moving upwards on the black curve (say, to 0.35) to account for the expected low income realizations in the upcoming future. If it does so however, it will instead get revealed as the low type and see the price on its bonds fall all the way to the bottom blue line. In this situation the government has two choices: either to maintain a large debt stock by mimicking the high type, or alternatively, accept the revelation of type and conduct a much more sizable debt reduction (say, to 0.2) to secure a reasonable price for the bonds it sells. One interpretation for the European debt crisis is therefore that governments were willing and able to adjust their debt policies, but not to the extent necessary to guarantee stable interest rates after the markets fully learn about the economy's new regime.

Figures 2.1(a) and 2.1(b) show the mechanics of the separating and pooling equilibria in the model. The two plots contrast debt policies of the two types for different prior beliefs, as functions of current income level. As can be noticed, for the lowest output realizations today, both government types default and are excluded from further borrowing, which is why their debt policy equals to zero. As the level of income rises, the High type finds it optimal to repay the debt and beyond that threshold enjoys an interval of mostly undistorted debt levels. At some point in the income range however, the low-type government also decides to repay the debt and engages in a pooling equilibrium, by mimicking the High type's optimal policy. As a result, the two lines merge on the graphs. Notice that at the endpoints of the pooling interval the high type makes an "escape", by increasing the borrowing level to prove itself to the markets and prevent the low type from mimicking its behavior. Eventually, for endowment realizations large enough, the two types decide to separate and settle their debt policy on their individually optimal levels corresponding to a full information world.



(a) High belief: $p = 0.99$ (b) Medium belief: $p = 0.5$

Figure 2.1: Policy functions for different levels of prior belief under asymmetric information

As can also be noticed from analyzing Figures 2.1(a) and 2.1(b), the prior market belief matters for the level at which the policy function in a pooling equilibrium is established. For a very high prior, the low type government is the one to take the entire burden of mimicking. This is because the low type has potentially a lot to gain from confusing the markets and is willing to distort its full-information benchmark debt level a lot. At the same time the high type's loss in that case is almost negligible (posterior belief drops from 1 to 0.99). Conversely, for a much lower prior belief, the high type government optimally reduces the debt because it cannot secure high enough prices for its bonds any more due to the lenders' confusion. Meanwhile, the low-type government has much less to gain from mimicking. Thus, the optimal pooling debt level is set much closer to the low type's optimal amount in a full information world.

Figures 2.1(a) and 2.1(b) provide a key interpretation of the debt crisis in Europe through the lenses of the asymmetric information model. For high enough income levels the equilibrium induces separation of types and all information is revealed to the markets. As income drops and the regime switches from high to low the government finds it optimal to maintain a high level of debt while the market belief is still high in a hope that the regime will switch back in the future and no drastic fiscal adjustment will be necessary. As the income stays low for a longer time, the lenders begin to update their prior belief downwards forcing the government to reduce its debt upon lower bond prices. Eventually, the government either defaults or settles on its benchmark full-information level of debt as the markets become fully convinced that the economy is in a major depression.

In the third step, I incorporate information asymmetry and conduct a simulation exercise of the debt crisis, analogous to the ones presented in Section 1.5 of Chapter 1. Figure 2.2 presents the results of simulating the debt crisis using the model in which the government knows exactly the current regime exactly. Similarly as in the previous two cases, the government accumulates debt slowly during the period of 2000-2008, while the bond spread is equal to zero. When the crisis hits in 2008:Q4, the newly-switched low-type government decides to engage in a pooling equilibrium and mimic the high type's optimal policy. In turn, the lenders must rely on their own prior belief to evaluate the default probability and hence we obtain a small initial increase in the bond spread. The separation of types occurs much later and causes the bond spread to increase sharply, while the debt is reduced to a low-type optimal level under a separating equilibrium.

The difference between the cases of symmetric and asymmetric information is subtle at first glance, but important. Under symmetric information, government debt is initially reduced by much larger extent because of two factors. First, higher interest rates make bond issuance less attractive. Second, the increased possibility that the regime might be low encourages the government to preventively reduce debt, in line with the policy functions presented in the right hand side panel of Figure 1.10. In the asymmetric information case however, the latter effect is not present because the high type government whose debt policy is being mimicked since 2008:Q3 is not concerned about the possible regime switch. Consequently, the observed debt reduction is only driven by the higher interest rate and it is much smaller in magnitude.

It is worth emphasizing that the result with asymmetric information is delivered under almost the same path of market belief over time as in the PSI model (recall that in the pooling equilibrium lenders evaluate bond prices based on their own prior belief, just like under symmetric information). The difference only arises at the end of the pooling episode where the revelation of types occurs one quarter earlier with asymmetric information and is more abrupt (market belief drops to zero). That is to say, the mimicking behavior of

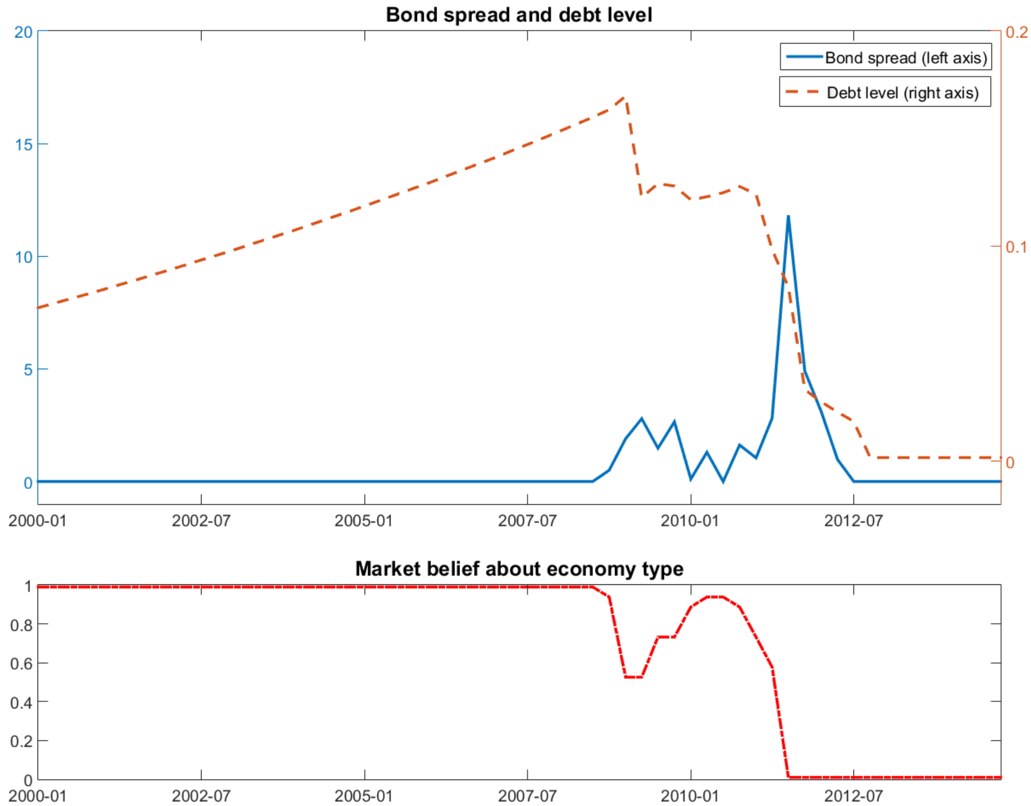


Figure 2.2: Simulated debt crisis in the Partial Asymmetric Information model

the low type government does not affect the lenders' belief in any way. Instead, it prevents the full separation of types and has a significant impact only on the observed holdings of government debt.

Even though the asymmetric information model predicts some reduction in debt, it should be viewed as a conservative result that greatly depends on the assumed calibration of the output process. Figure 2.2 shows that while the model predicts a drop in belief to almost 0.5 at the outset of the crisis, this was not the case in Figure 1.2. In reality, the sovereign rating of most European governments suffered only minor downgrades in years 2008-2010, reflecting a large confidence in debt repayment prospects on the part of international lenders. Hence, if we allowed a higher variance in the calibration presented in Table 1.3, the resulting market learning would be slower, the initial belief drop milder, and the high type government would have even less incentive to reduce debt in the pooling equilibrium.

2.4 Discussion

In this section, I discuss the role of bailouts during the debt crisis in Europe and show how they fit with the theory presented in this paper. I also highlight the main policy implications.

2.4.1 Bailouts and the European debt crisis

One of the crucial factors during the debt crisis in Europe has been the presence of emergency loans from the international institutions such as the International Monetary Fund, the European Central Bank and the European Commission. The scale of these bailouts was historically unprecedented. Table 2.2 lists emergency lending programs announced for the Eurozone members in years 2010-2012 and compares them to the largest official loans prior to 2008. As can be noticed, the total disbursements promised as part of the emergency packages ranged from over 40 percent of GDP for Ireland and Portugal, up to almost 100 percent of GDP for Greece, corresponding to $40-70$ percent of the total debt of these economies. By contrast, in one of the largest such interventions prior to 2008, Mexico was bailed out by the US government in 1995 with the total amount of roughly 15 percent of its GDP at the time.

It is tempting to interpret these bailouts as an alternative explanation for the small increase in the interest rates in Europe between 2008 and 2010. For example, Aguiar and Gopinath (2006) propose an extension to their sovereign debt model which incorporates third-party bailouts. They argue that bailouts essentially provide an upper bound on the creditors' losses in the case of default and result in an interest schedule that is lower and less sensitive to additional borrowing. They also show that including a bailout in the simulations of Argentina's business cycle results in over 60 percent lower standard deviation of the interest

Country	Time	Total loans	Gov. debt	Loans/Debt
Bailouts for the Eurozone members:				
Greece	May.2010	99.65	148.3	0.67
Ireland	Nov.2010	42.36	91.2	0.46
Portugal	May.2011	43.82	107.0	0.41
Spain	Jul.2012	4.02	86.0	0.05
Largest bailouts prior to 2008:				
Mexico	Jan.1995	14.54	47.98	0.30
Turkey	Dec.2000	7.13	43.81	0.16
Brazil	Nov.1998	4.73	27.91	0.17
Argentina	Oct.2001	2.98	37.22	0.08

Sources: Total amounts of emergency packages for the Eurozone countries are taken from the IMF and the European Commission data. The largest bailouts prior to 2008 are based on the IMF accounts and additional sources as reported in the official announcements (the amounts actually disbursed were in fact much lower). The GDP and external debt stock data are taken from the World Development Indicators.

Table 2.2: Emergency loans and total government debt before and after 2008 (% of GDP)

rate than in the benchmark case with no bailouts. Their model provides a theory for why Portugal, a member of the Eurozone with implicit guarantees from the European institutions, had experienced a significantly lower volatility of the bond spread prior to 2008. By itself however, it does not explain why the surge in the interest rate occurred with such a delay. Suppose that, following Aguiar and Gopinath (2006), the equilibrium bond price is

$$q_t = \frac{1}{1+r^*}\left\{ \min\left[1, a^*\right] + E_t\big(1 - d_{t+1}\big) \max\left[1 - a^*, 0\right] \right\} \tag{2.17}$$

where a^* denotes a fraction of the economy's debt guaranteed by the external institutions. Then, given that Portugal's GDP and foreign debt were roughly at the same level at the beginning of 2009 as in mid-2011 (which can be verified by looking at Figures 1.1 and

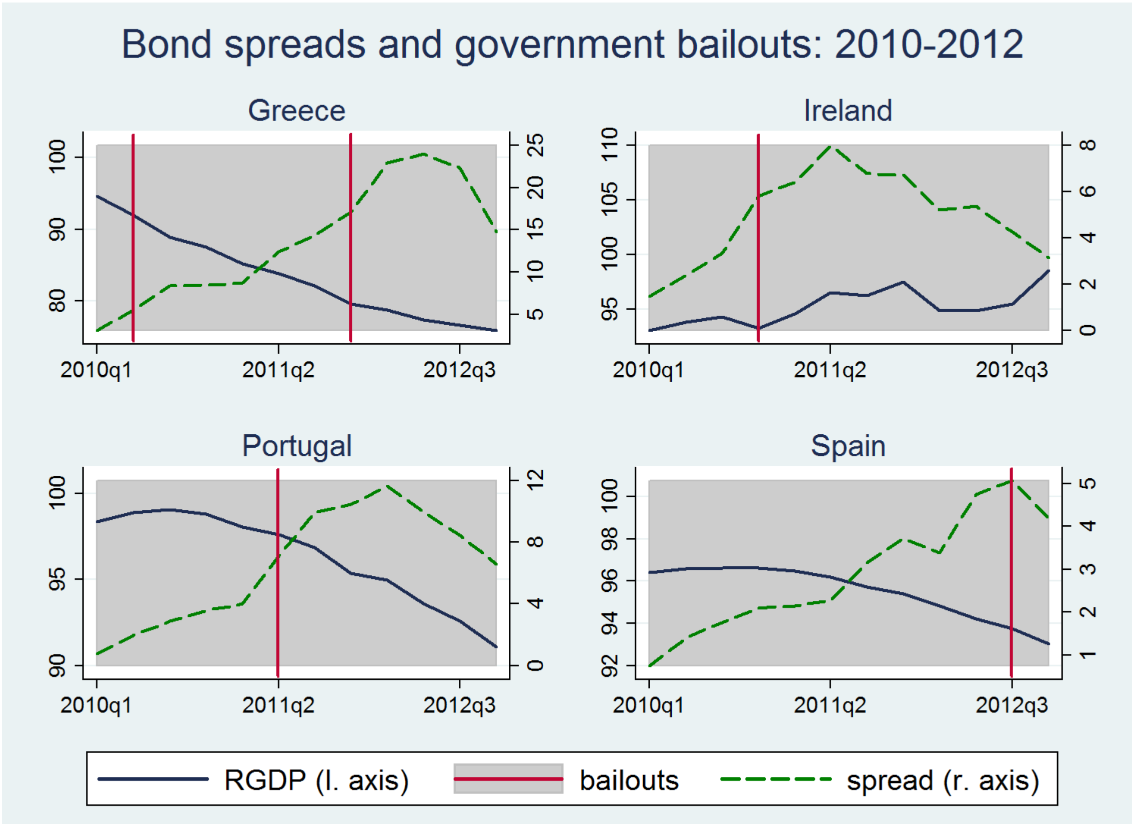
1.3), it seems puzzling why the interest rate increases so much more in the latter case. It is plausible to assume though that the exact value of a^* is not known to the lenders *ex ante* and the expectation of it changes over time. In other words, foreigners may not know to what extent the European Union will be willing to cover the debts of its troubled members, and their doubts increase over time resulting in an upward pressure on the interest rate. However, as we look more closely at the timing of the bailout announcements for particular countries, this hypothesis also seems questionable. Figure 2.3 presents the plot of real GDP and bond spreads for the four economies listed in Table 2.2,[7] zoomed in for the period 2010-2012 and with the announcements of emergency loans marked with vertical lines. Two observations from Figure 2.3 are particularly relevant. First, notice that the bailout announcements are well spread out in time, i.e. Greece receives one first in May 2010,[8] followed by Ireland in November 2010 and Portugal in May 2011. It is therefore hard to argue that the surge in the yield on Portuguese bonds was mostly driven by a growing uncertainty around the value of a^* given that the European Commission had proven in several cases by that time that it was willing to provide large-scale loans to its members of roughly similar size and importance as Portugal, in order to prevent them from defaulting.[9]

The second observation evident in Figure 2.3 is that interest rates *increase* on impact of the announcement of an emergency loan and only reach a peak up to three quarters following it (the only exception is Spain, whose bailout coincided with the introduction of OMT, a general debt-repurchasing program of the European Central Bank, and in itself was significantly smaller in size). Even if we accept the argument that the increase in bond spreads *prior to* the announcement of the Portuguese bailout was driven by the uncertainty if the EU would be willing to grant it, once it has been observed it should stabilize the interest rate on impact (with a possible further increase later, contingent on the country

[7] Relative to Figure 1.1, Italy is replaced with Spain because it did not receive an official bailout.

[8] In fact, if we consider the non-Euro area members of the European Union, also Hungary (November 2008), Latvia (December 2008) and Romania (May 2009) obtained sizable bailout packages and managed to avoid a sovereign default.

[9] This argument is more compelling for Italy or Spain, for whom the peak of the debt crisis occurred after observing the eventual Greek default in March of 2012 and whose economies are significantly larger than in any of the previous cases.



Note: The GDP series are in constant 2010 prices, and their values are normalized such that the third quarter of 2008 equals 100. The bond spreads are expressed in percentage points. Bailouts are marked for the quarter when the official announcement of an emergency package was first made.

Figure 2.3: Bond spreads and announcement of bailouts during the crisis: 2010-2012

not receiving the scheduled disbursements or not recovering from the recession). Instead, the observed sharp increase in the bond spread is consistent with the theory of asymmetric information presented in this paper. A government whose actions convey a signal about the state of the economy's underlying fundamentals is reluctant to request official assistance early during the crisis (in addition to not reducing the debt) as that would reveal its true type.[10] Later, when the lenders' belief independently drops to a level low enough, the government has little to gain from further pooling and decides to separate by requesting a bailout and reducing the debt. When the lenders update their information fully, the observed outcome is a sharp increase in the interest rate resulting from a higher expected probability of default. Showing this result explicitly however involves computing a signaling equilibrium under two choice variables (bonds and emergency loans) and is a challenging extension for future research.

2.4.2 Policy implications

This paper shows that in a world where fiscal policy decisions convey a signal about the underlying state of economic fundamentals, the government faces a choice to either conduct a comprehensive fiscal reform and thus to secure high bond prices, or to keep running large debt levels and pretend that the economic fundamentals are sound. Any action in between may actually have an adverse effect on the interest rates by signaling low expectations of future income to the markets. In the presence of political frictions,[11] a drastic debt reduction is often not feasible and the government may decide to gamble by postponing fiscal adjustments, as well as the requests for emergency loans. An important policy implication coming out of this theory is that, for international bailouts to be effective, they generally need to be more elastic and appealing to governments as a preemptive measure, rather than the ultimate lifeline.[12] If it is more common to observe countries use IMF credit lines flexibly to prevent potential debt crises then the signaling effect of such actions would also be greatly attenuated.

More generally, the trade-off between a radical fiscal reform and lack thereof discussed above depends crucially on the parameters of the model, in particular on the government's impatience. The relatively high values of the discount factor used in sovereign debt models

[10] This idea, albeit novel to formal economic research, has long been discussed by financial markets professionals. For example, commenting on the IMF bailout for Brazil in September of 2002, the Bloomberg Businessweek magazine wrote: *The irony of IMF credit lines is that while investors are glad the money is there if needed, they don't want to see governments actually dip into it. It's seen as sign of desperation.*

[11] For a review of sovereign debt models with political economy, see Hatchondo and Martinez (2010).

[12] Boz (2011) documents that in the data prior to 2007, sovereign borrowing from International Financial Institutions tends to be intermittent and countercyclical.

can be interpreted as an effect of political frictions (Amador (2003)) and give rise to a gambling motive in government's decisions. To prevent such behaviors, it seems desirable to consider more fiscal policy rules that would automatically stabilize the debt at *ex ante* optimal levels and thus discharge the information load of the government's discretionary choices.

2.5 Conclusion

In their seminal contribution, Lucas and Sargent (1979) make the following remark about general equilibrium macroeconomic models:

> *It has been only a matter of analytical convenience and not of necessity that equilibrium models have used the assumption of stochastically stationary shocks and the assumption that agents have already learned the probability distributions they face. Both of these assumptions can be abandoned, albeit at a cost in terms of the simplicity of the model.*

This paper shows that learning about the probability distributions of future income shocks was an important factor during the European debt crisis. It had an impact not only on the observed asset prices, but also the real variables such as government debt and consumption. I show that an otherwise standard quantitative model of sovereign debt can be augmented to incorporate this learning process and match the markets' gradually evolving beliefs over time. As a result, we obtain a delayed pattern of bond spread increases during the Great Recession in Europe. In addition, I argue that if the government is endowed with more precise expectations about the economy's future income, it has strong incentives to keep its foreign debt high in order not to send a negative signal about the bad state of economic fundamentals. That is, it tends to delay the necessary fiscal reforms and also delay requests for external emergency loans (such as the IMF bailout), because of the temptation to pretend that it expects a prompt recovery.

The two-channel learning mechanism developed in this paper is more general and has potential applications in other fields of economics. In particular, researchers working on the problems of political economy often analyze situations in which a government deliberately postpones the necessary reforms. At the same time it maintains a high level of popular support and is only ousted from power when the crisis breaks out. The model developed in this paper offers a simple interpretation for such situations. The government naturally knows more about the fundamentals of the economy than members of the society do, who only learn about it slowly by observing prices and wages. The government fails to introduce reforms because doing so would signal a bad state to the voters and result in an immediate loss of popular support. This pooling equilibrium cannot last forever though. At some point a separation occurs in which either the economy improves and the authorities stay in power, or the society fully learns about the bad state and replaces the authorities. The exact specification of such a political economy model is an interesting avenue for future research.

Chapter 3

Pay What Your Dad Paid: Commitment and Price Rigidity in the Market for Life Insurance

3.1 Introduction

Traditional economic theories predict that prices in a competitive economic environment evolve accordingly to changes in the underlying marginal cost of production. However, empirical literature provides evidence that this process is often sluggish. In response, a vast sticky price literature has emerged attempting to provide understanding for the observed movements in prices. These models are often based on simple mechanical frictions that lack theoretical underpinnings. In this paper, motivated by empirical findings from the life insurance market, we propose a novel theory of price rigidity in which the company optimally commits to a constant pricing schedule for a certain range of marginal cost variations.

Life insurance data is particularly suitable to test pricing theories. The contracts are simple, the data on historical premiums is readily available and the marginal cost for that industry can be estimated easily. Our attention focuses on the renewable level-term form of insurance. These contracts require a down payment of premium at the moment of signing and stay in force for a pre-defined period, typically between one and twenty years. After the term expires, customers face a premium schedule that increases with age and are allowed to renew the policy without undergoing a medical reclassification. Thus, as it is pointed out by Hendel and Lizzeri (2003), level-term contracts are characterized by one-sided commitment of the company, but not the consumer. Table 3.1 presents the structure of a typical level-term insurance, commonly referred to as the Annual Renewable Term (ART), for the first 10 policy years. It is important to notice that the insurer only commits to an upper bound on the future premiums ("Guaranteed Maximum" column), which vastly exceed the amounts that can be expected in a market equilibrium. At the same time, however, the contract stipulates a projected path of premiums based on the rates currently offered to older individuals in the same category ("Non-Guaranteed Current" column). This schedule is not binding though, and the company may change it at any point in the future. The only question is: will it?

Table 3.1: Structure of an Annual Renewable Term (ART) contract

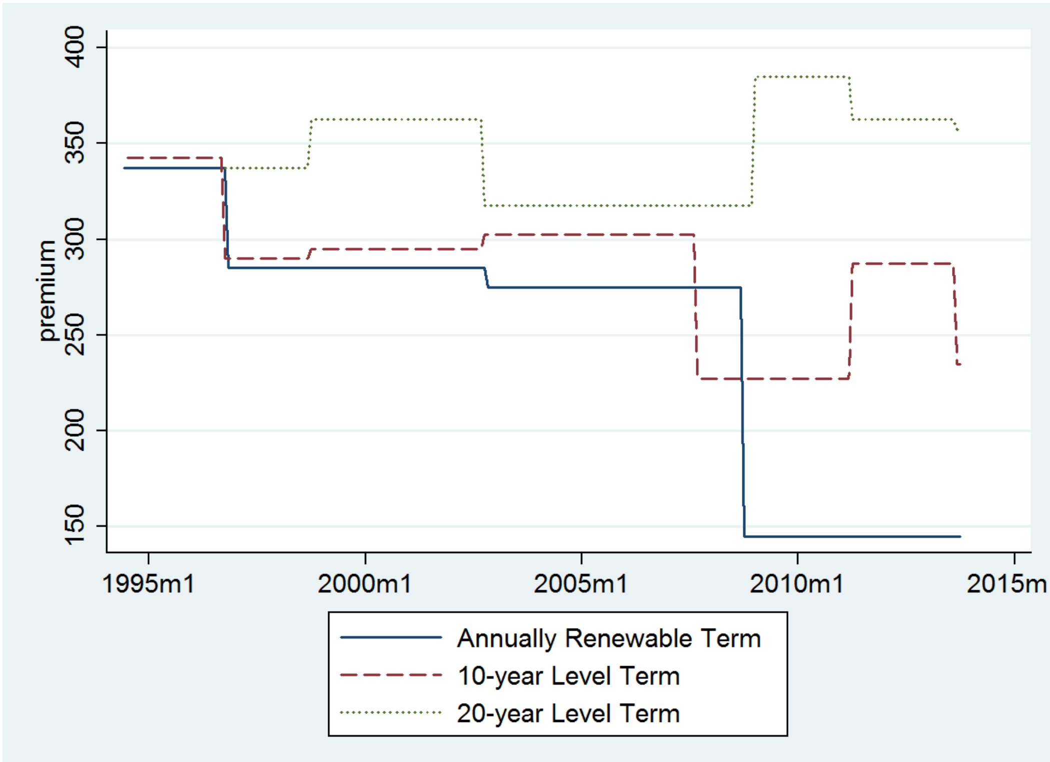
Age	**Face Value**	**Guaranteed Maximum Contract Premium**	**Non-Guaranteed Current Contract Premium**
30	250,000	265.00	265.00*
31	250,000	517.50	267.50*
32	250,000	517.50	267.50*
33	250,000	530.00	270.00*
34	250,000	557.50	272.50*
35	250,000	587.50	280.00*
36	250,000	627.50	292.50*
37	250,000	672.50	307.50*
38	250,000	722.50	325.00*
39	250,000	780.00	350.00*

Sample contract offered by United Heritage Life Insurance Company.

Source: Compulife Software, October 2013.

In this paper we document that life insurance companies in the US have displayed a striking commitment to their own non-binding promises since 1990. In a data set we construct from an insurance quotation software, the premiums are extremely rigid over time, with the overall probability of a monthly premium change amounting to just 2.56%. This implies an average premium duration of roughly 39 months, placing life insurance on the far-right tail of the price change frequency distribution documented by Bils and Klenow (2004). Figure 3.1 presents an example of premium evolution for different types of policies. These prices are characterized by long periods of rigidity and infrequent, but sharp, adjustments. Remarkably, in the data set we observe a few policies that have kept a constant premium for at least 20 years, rendering it possible, as the title of this paper goes, that a son could pay the same amount for his life insurance as the father used to! We furthermore use the data on historical interest and mortality rates to estimate the underlying marginal cost for the life insurance industry. We find that this cost is non-stationary and highly volatile over time. These empirical patterns from the life insurance market therefore present a puzzle in the light of elementary economic reasoning and create a need for a new theory.

We propose a new theoretical framework that could explain the phenomenon exhibited in the data. The model addresses the issue of premium rigidity by introducing a transaction cost invested by new consumers before signing the contract. At the same time, existing policyholders have already made an investment in their relationship with the insurance company from previous premium payments. Once the investment is sunk, the consumers are locked in and the insurance company has incentive to increase premiums from their originally promised value by exactly the amount invested. This increase would not decrease the demand for life insurance once the investments have been made. This is because the transaction cost creates a kink in the demand for insurance, introducing a time inconsistency problem for the insurance company. The insurance company also faces stochastic cost shocks that are not observable to the consumers. Therefore, consumers are unsure whether premium changes are due to being held-up or the cost shock. As a result, the insurance companies design premium profiles or schedules that leave no doubt to the consumers that deviations in the premium from its promised value are due to changes in the underlying cost. In other words, the premium schedule has to be incentive compatible for the insurance company.



Sample premiums offered by the National Life Insurance Company of Vermont.

Source: Compulife Software, 1990-2013.

Figure 3.1: Premiums over time for different level-term policies

The main predictions of the model are that optimal premiums display rigidity, and that premium hikes need to be sufficiently large to be incentive compatible. To see why premiums are rigid, notice that if we neglect incentive compatibility, the optimal change in premiums under a mild cost shock is small. However, small changes in premiums do not affect the demand for life insurance once the relationship-specific investments are sunk, so the only incentive compatible premium profile involves rigidity over mild cost shocks. The insurance company is able to regain flexibility in premiums once the cost shocks are sufficiently large, since the corresponding optimal change in premiums is large and would also adversely affect the demand. Therefore, the premium hike needs to be of significant magnitude so that the negative effect on demand is severe and it can only be attributed to changes in cost and not because of any ulterior motives. We show that the optimal premium profile has a simple cutoff rule, where premiums are rigid for cost realizations below a threshold, but flexible above it. Our model explains why the level-term insurance sold by insurance companies have a non-guaranteed premium schedule that afford them the room to be flexible, but at the same time the finalized premiums rarely deviate from it.

The remainder of the paper is structured as follows. Section 2 presents a comprehensive literature review. Section 3 discusses the construction of our data set and summarizes the main findings about price rigidity in the life insurance market. Section 4 introduces the theoretical model to study the phenomena documented in Section 3. Section 5 presents the solution to the model and discusses its main predictions. In Section 6 we address the alternative hypotheses that could also explain our empirical phenomenon and present some evidence to rule them out. Section 7 concludes our findings and discusses the broader implications of the theory developed in the paper. The Appendix contains the proofs to theorems and lemmas presented in the main text, as well as a description of the method to estimate the marginal cost of life insurance.

3.2 Literature Review

Our paper builds upon several strands of economic literature, which we discuss briefly in the following section. Recently there has been much interest in the life insurance market, starting with Hendel and Lizzeri (2003). They use the data on life insurance to test predictions of the Harris-Holmstrom type of model with symmetric learning about the evolution of the insured's health and lack of commitment for the buyer. Because of learning on the company's side, short-term policies induce a risk of reclassification - should the consumer's health deteriorate, the new contract will involve much higher premiums. On the other hand, long-term policies are infeasible due to one-sided commitment on the policyholder's side. The solution, as predicted by the model and confirmed by the data, is front-loading of premiums. Hendel and Lizzeri show that virtually all life insurance policies available on the market exhibit some degree of front-loading which consequently affects lapsation (i.e. rate of voluntary termination of coverage) in a negative way. Daily, Hendel and Lizzeri (2008) and Fang and Kung (2010) have pursued a similar line of research by considering the effects of a life settlement market on the optimal life insurance contract. Alternatively, Gottlieb and Smetters (2013) develop a model with naive policyholders who underestimate the extent of their income shocks to explain the front-loading of premiums. When policyholders are hit with an unexpected income shock early in their life-cycle and respond by lapsing, the insurance company can make a profit if the premium structure is front-loaded.

The aforementioned papers developed theoretical models for the premium structure faced by a *fixed individual* over time. In contrast, our work tries to explain the premium evolution of a life insurance contract for a *fixed age group* over time. To the best of our knowledge, very few papers have explored this dimension of life insurance contracts, and none have documented the phenomenons highlighted in this paper.

More recently, Koijen and Yogo (2015) investigate the unusual pricing behavior of insurance companies following the financial crisis of 2008. They find that life insurers sold long-term policies at greatly reduced premiums relative to their actuarially fair value, resulting in negative average markups across different companies. The authors propose a theory according to which this phenomenon can be explained as a consequence of financial frictions and statutory reserve regulation. Essentially, life insurers were able to improve their required capital holdings by selling discounted long-term policies. In a subsequent contribution, Koijen and Yogo (2016) develop a similar framework to assess the effect of the so-called shadow insurance on the market outcome observed in the US. nsurance companies have recently been ceding large amounts of liabilities to their affiliated reinsurers, situated in both on- and offshore locations characterized by less restrictive capital regulation. By doing so, the operating insurance companies are able to avoid the cost of keeping the required level of risk-based capital. As a result, the supply of policies in the market for life insurance is significantly expanded, at the cost of much higher impairment probabilities than the ones considered by rating agencies. The findings from these two papers help us understand some of the price drops observed in our sample in the late 2000s, as documented in Section 3.

The theory presented in this paper relates to Klemperer (1987)'s original idea of consumer switching costs as a factor differentiating the ex ante identical products. The existence of a transaction cost (which we also assume throughout this paper) leads to consumer lock-in and allows the company to extract future monopoly rents. Because of that, firms may often be interested in pricing their products for new customers below marginal cost in order to attract as many of them as possible. In our model, the transaction cost incurred before the contract is signed can also act as a switching cost in future periods. However, price rigidity is not consistent with the standard switching cost model.

Our theoretical model is also related to several papers in the mechanism design literature. Amador, Werning and Angeletos (2006) consider the optimal consumption and savings decision for a two-period problem with time inconsistency and taste shocks. Their analysis differs from ours in two dimensions. In their framework, it is not optimal for the agent to choose consumption and savings flexibly because of a systematic indulgence in present consumption which they regret ex-post. In our model, the insurance company suffers from time inconsistency because of a hold-up problem which is inherent in the market for life insurance.

Secondly, their agents do not incur a cost for indulgence, so it is optimal to set an upper bound for present consumption. However, in our model, we show that for firms it is possible to act optimally when the shock to cost is high, since the loss of consumers from a higher associated price is costly and would help establish credibility. In other words, for high cost shocks, the optimal level of profit can be attained since the prices can be credibly adjusted in a way that the insurance company fully internalizes the consequences of a price hike. This logic is similar to Koszegi (2005). Also related is Athey, Atkeson and Kehoe (2005), who investigate the optimal discretion that should be allowed for a monetary authority. Similar to Amador, Werning and Angeletos (2006), and in contrast to our analysis, it is optimal to set an upper bound on the inflation rate for the monetary authority because they lack credibility.

3.3 Life Insurance Prices

3.3.1 Data Construction

We construct a sample of life insurance premiums from Compulife Software, a commercial quotation system used by insurance agents. The updated programs are released monthly, spanning the period from May 1990 until October 2013. For each of the 282 months collected, we recover the premiums for 1-, 5-, 10- and 20-year renewable term policies offered by different companies[1]. Even though Compulife is not a complete data set, it covers most of the major life insurers with an A.M. Best rating of at least A-. As the default customer profile we use a 30-year-old male, in the "regular" health category, purchasing a policy at a face value of $250,000 in California. The choice of this particular state is by Compulife's recommendation, due to a relatively large population and wide representation of insurance companies. The obtained sample consists of 55,829 observations[2] on the premium levels for 578 different policies offered by 234 insurance companies. Naturally, over the course of 23 years these firms tend to disappear or merge, as well as discontinue their old products and launch new ones. For this reason, even though we keep track of such transformations whenever possible, each product is observed on average for just over 96 months (with a median of 84).

3.3.2 Historical Premiums

Table 3.2 provides a statistical description of price rigidity in our data set. Among 578 distinct insurance products that appear for at least 12 continuous months in the sample, only 369 change their premium amount at all. In total there are just 1432 price changes, consisting of 580 hikes and 852 drops. It is also important to notice that these premium changes, whenever they occur, tend to be of large magnitude, on average amounting to over 10%. The probability of a price change in any month is 2.56%, resulting in an average

[1] Insurance firms often offer several policies of the same type in parallel. In such case, we keep the lowest price assuming it would be the consumer's optimal choice.

[2] Because of frequent incompleteness of Compulife databases (especially in the 1990s), we impute the prices whenever a discontinuity appears for up to most 12 months. As a result, we have a total of 562 imputations which constitute roughly 1% of the final sample size. We also drop all the products that are observed for less than 12 continuous months.

premium duration of roughly 39 months. This figure includes a vast number of companies that do not adjust prices even once. This may be a deliberate business strategy, but it may also result from other, not market-oriented factors[3]. Hence, we also calculate the statistics for the subsample of insurance policies that undergo at least one price change. Among those products the probability of a monthly price adjustment increases slightly, but still remains low at 3.5%, resulting in average duration of almost 28 months. It should furthermore be noticed though that products whose price changes in the data set, also tend to stay around for a longer time (114 months as opposed to an unconditional average of 96 months). This observation suggests that certain insurance policies tend to be discontinued (and supposedly replaced by new ones) rather than deviate from the previously promised premium schedule.

In order to visualize these findings, Figures 3.2(a) and 3.2(b) take a closer look at the distribution of premium durations and adjustment magnitudes. The first chart depicts a standard view of a distribution of durations with significant positive skewness and a long right tail reaching up to 20 years! Each bin on the histogram represents 6 months, which means that a majority of premiums last between 6 and 30 months, but only a few adjust sooner than that. The second chart presents the distribution of relative sizes of price adjustments. As it is clear from the summary statistics in Table 3.2, premium drops occur more often and are of slightly larger magnitude. The adjustments reach as much as 50% in both directions.

In a final piece of data analysis, we explore the distribution in insurance premiums in our sample by examining the relative price dispersion. Figure 3.3 sketches a histogram of all prices, where the average premium in every month and of every term length is normalized to 100. The most striking feature of the graph is the long right tail which implies that some life insurance policies are offered at a price 2.5 times as high as the average in that category, at the given point in time. More generally, even though life insurance may seem to be a rather homogeneous financial product, we observe a significant dispersion across different policies. This may be attributed to varying terms and conditions of different policies (we aggregate

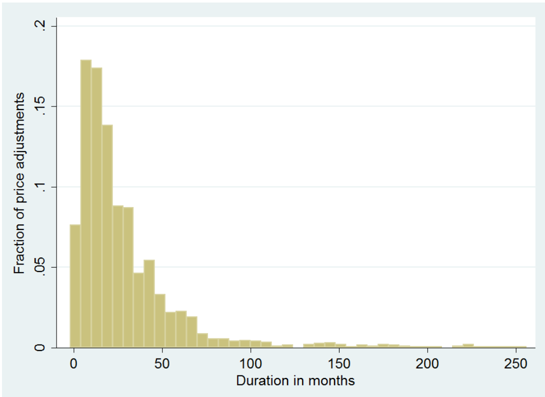
[3] For instance, an insurance company that has no interest in selling certain types of policies may still offer them as a reference for tax authority.

Table 3.2: Price rigidity in the sample

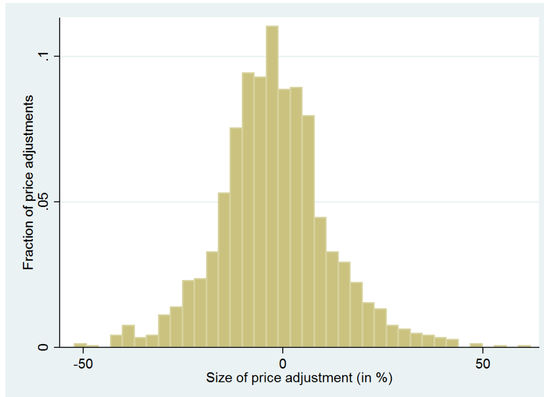
Total number of observations	55,829
Total number of insurance products observed	578
Total number of products that change price	369
Total number of price changes	1432
Total number of price hikes	580
Total number of price drops	852
Average magnitude of price change (in %)	10.74
Average magnitude of price hike (in %)	10.58
Average magnitude of price drop (in %)	10.85
Whole sample:	
probability of a monthly price change (in %)	2.56
median probability of a price change (in %)	1.73
average number of observations per product	96.59
Excluding the companies that never adjust:	
probability of a monthly price change (in %)	3.41
median probability of a price change (in %)	3.23
average number of observations per product	113.82

all products under the category "renewable level-term"), as well as imperfectly competitive economic environment in which life insurance companies operate. These imperfections may include search frictions (Hortaçsu and Syverson (2004)), information frictions or product differentiation (e.g. with respect to company reputation or brand loyalty).

Thus far we have illustrated the rigidity of life insurance premiums for a fixed profile of consumers. This means that every month a new 30-year-old non-smoking male, in "regular" health category, can purchase an ART policy for the same price. It is important to understand however, that the premium level this customer expects to pay at renewal is equal to the amount currently offered to the corresponding 31-year-olds (the "Non-Guaranteed Current Contract Premium" column of Table 3.1). Because life insurers only



(a) Histogram of premium durations



(b) Histogram of adjustment sizes

Figure 3.2: Distribution of premium durations and adjustment sizes

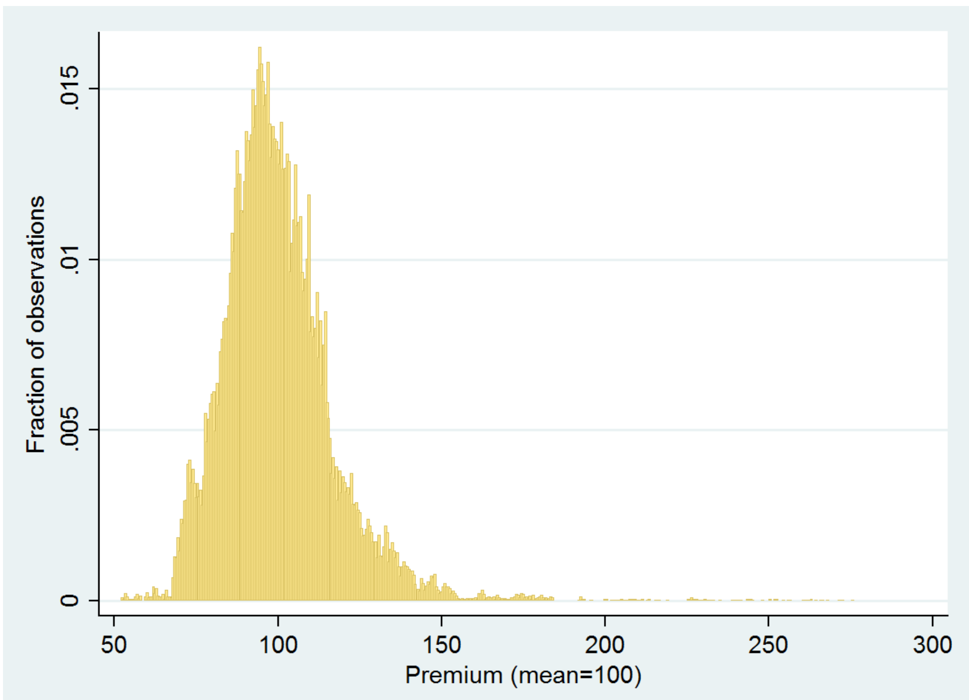


Figure 3.3: Distribution of insurance premiums, relative to the cross-sectional average

adjust the entire premium schedule, rather than individual rates for selected ages, in most cases these non-guaranteed promises are being kept by the companies. More generally, price rigidity in the market for life insurance can be thought of in two dimensions - one, where a new customer of the same age is offered the same price every period; and two, where an existing policyholder renews the policy for the same amount as, at the moment of buying, his older counterparts did.

3.3.3 Marginal Cost Estimation

In this section we contrast the historical premiums discussed so far with the marginal cost faced by life insurance companies over the time. We will approximate the life insurance company's marginal cost by estimating the net premium of a level-term policy (also known as the actuarially fair value). A precise description of the method based is provided in the Appendix. Intuitively, a net premium can be thought of as outcome of a zero-profit condition faced by the insurance company. Figure 3.4 presents a stylized illustration of an insurer's cash flow structure. A level-term insurance policy is effective from the moment the first premium is paid, period t, and stays in force for as long as the customer keeps renewing it. Beyond the predefined term, premiums are increasing with age and the benefit is paid out by the company at the moment of death of the insured, denoted $t+n$. In order to break-even, the company must acquire a portfolio of risk-free assets to replicate the present expected value of its future balances. P_t is therefore such that the expected present value of cash flows between the company and the policyholder are equalized, i.e. $\sum_{s=0}^{N} \frac{E_t[P_{t+s}]}{R_t^s} = \sum_{s=0}^{N} \frac{E_t[B_{t+s}]}{R_t^s}$.

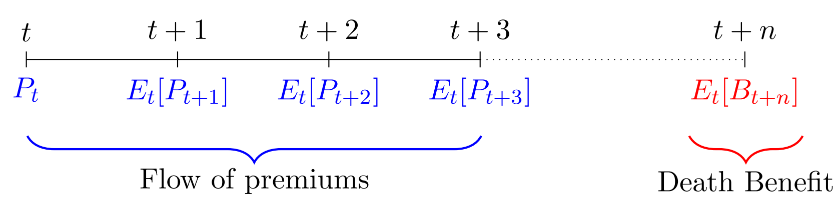


Figure 3.4: Stylized structure of a life insurer's expected cash flows

Figure 3.5 plots the evolution of net premium for an ART policy, from May 1990 until October 2013, together with the HP trend. It ranges from as low as \$118 (in February

2011) up to $173 (in December 2008), with a standard deviation of 10.5. The net premium exhibits considerable fluctuations over time that depend, by and large, on movements in the interest and mortality rates. In particular, a downward trend can be noticed in the early 2000s (due to a relatively large drop in mortality rates), as well as sharp hikes after November 2008 caused by the interest rate shocks. The recent financial crisis episode contributes to this volatility significantly. While the standard deviation of the HP-filtered net premium is 5.33 for the entire sample, it drops to 3.19 if we disregard all observations after October 2008.

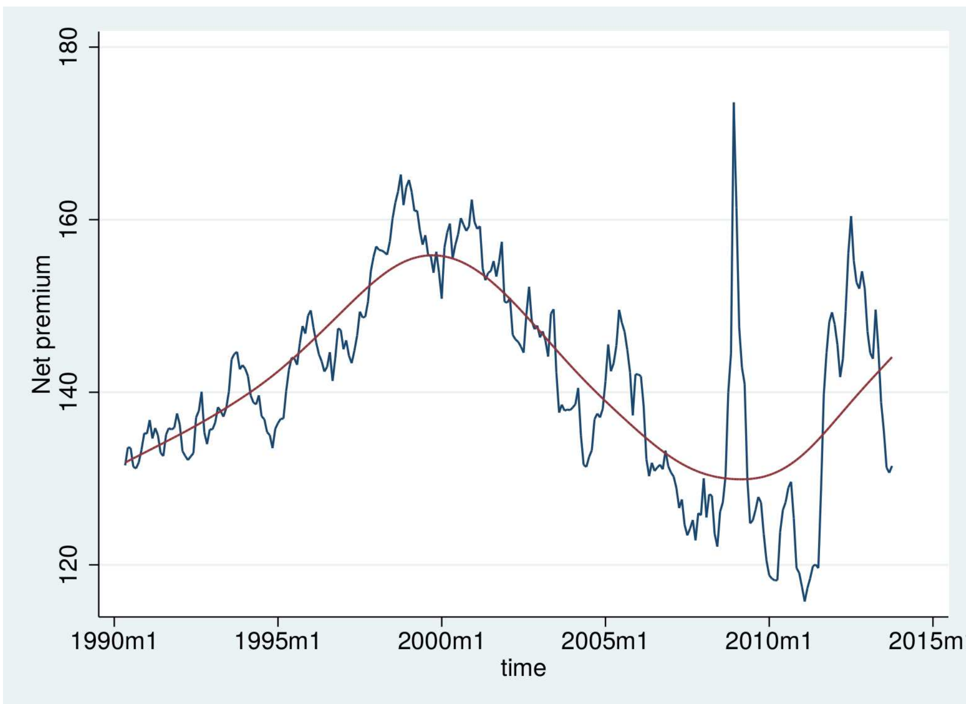


Figure 3.5: Net premium for an Annual Renewable Term policy over time

3.4 The Model

In this section, we present a dynamic life insurance pricing model. We start by briefly presenting the main idea of the model along with its basic intuition. The two key ingredients of the model are the presence of a one-sided hold-up problem of the consumers, and the private stochastic cost faced by the insurance company.

The hold-up problem comes from the consumers' need to make an investment (time forgone while searching, attending medical exams and filling out questionnaires, as well as the risk of having claims denied during the contestability period) before formally purchasing life insurance. We will call this investment the transaction cost. This creates an incentive for the life insurance company to increase its premiums after the transaction cost is sunk. Another source of the hold-up problem arises from a risk of health reclassification. An existing policyholder may face the risk of being reclassified in the future by other insurance companies after switching. With adverse selection, the insurance company would want to increase the premiums since the existing pool of policyholders would likely face higher premiums if they choose to switch due to health deterioration. As a result, the life insurance company is confronted with a time inconsistency problem, and it can respond by setting a legally binding guaranteed premium schedule to demonstrate commitment.

However, the life insurance company would also like to retain a certain degree of flexibility to respond to the private stochastic cost it faces. Due to this problem, it may not be optimal for the life insurance company to commit to a guaranteed premium schedule. Instead, it may wish to retain some degree of flexibility in the premium schedule in response to potentially large movements in the cost.

These two features of the model highlight the main trade-off. The life insurance company would need to announce an incentive compatible premium schedule such that consumers are convinced that upward premium movements are due to cost shocks and not opportunistic behavior. The solution for the life insurance company is to commit to a rigid premium for small cost shocks, and to change its premiums only when the cost shocks are large enough. The life insurance company can afford to change its premiums only when the increase in premiums decrease the demand for insurance to an extent that it couldn't possibly be profitable to do so for any other reason.

3.4.1 The Setup

Health, Demand and Transaction Cost

Consider a two period model where a continuum of consumers choose whether to purchase insurance in the first period and whether to renew in the second period. In each period, the consumers face a mortality risk of $m_t \in (0,1)$, where $t \in \{1,2\}$, which is common knowledge.

Consumers are heterogeneous in their private valuations of the life insurance contract $r = (r_1, r_2)$. Variable r_t can be thought of as linear utility from owning a policy in period t. The private valuations are distributed according to a continuous and differentiable joint distribution function $h(\cdot,\cdot)$, with bounded support $[\rho, R]^2$ and $\rho > 0$. We assume that R is sufficiently large such that there still is positive demand even when a firm facing the largest possible cost realization charges the optimal monopoly premium. We will denote the cumulative marginal distribution function as $H_t(\cdot)$ for period t private valuations. We will also assume that the hazard rate is increasing for H_t, for $t \in \{1,2\}$, which is a common assumption to ensure a downward sloping demand curve. We will assume that only the distribution of consumers' valuations is common knowledge, so that life insurance companies are unable to write individual specific contracts. We normalize the value of not owning life insurance to zero.

We focus on consumers who seek to obtain life insurance coverage for multiple periods. This is obtained by assuming that the expected value of renewing a policy in the second period is greater than the value of dropping out, normalized to zero, or switching. The reason behind excluding the consumers who only demand life insurance for one period is that they are not held-up by the company, and they would most likely purchase a term policy without the option to renew. These so-called "non-renewable" contracts are usually cheaper than renewable level-term insurance and thus are more likely to attract such customers.

Before becoming a policyholder, a consumer needs to invest a transaction cost of $\mu > 0$. As mentioned before, there are various costs captured within the transaction cost parameter

μ. For simplicity, we assume that μ does not vary with age or time. The policyholder also needs to invest the transaction cost if he decides to switch companies in the second period, but would not need to if he renews. We will elaborate on the outside options when we discuss the commitment problem of the consumers.

The model implicitly assumes the existence of other firms providing life insurance coverage. However, we do not explicitly model the strategic interaction among firms. This modeling choice enables us to focus on the relationship between the insurance company and its customers and on the contracts they sign.

Cost Shocks and Premium Schedules

For each period, the life insurance company faces a stochastic unit cost c_t, which is randomly drawn from a continuous and differentiable cumulative distribution function $G_t(\cdot)$ with bounded support $[\underline{c}, \overline{c}]$, for $t \in \{1, 2\}$. Notice that the unit cost does not vary with the size of the insured pool. We will assume that $\overline{c}$ is sufficiently large, so that large negative cost shocks have a strictly positive probability of occurring. We also assume that the cost realizations are independent.

The insurance company chooses premium schedules $P_1(c_1)$ and $P_2(c_2; c_1)$ as a function of the possible cost realizations. The second period premium $P_2(c_2; c_1)$ depends on c_1 because the first period premium affects the pool of existing policyholders before the second period. In essence, the premium schedule is a mapping $P_1 : [\underline{c}, \overline{c}] \to \mathrm{R}$ and $P_2 : [\underline{c}, \overline{c}]^2 \to \mathrm{R}$. A life insurance contract is defined as the premium schedules and the face value of the contract $\{P_1(c_1), P_2(c_2; c_1), F\}$. For simplicity, we take the face value as given, so the equilibrium contract that we solve for is $\{P_1(c_1), P_2(c_2; c_1)\}$, with a predefined F. For illustrative purposes, we also assume that c_1 is known, which simplifies the contract to $\{P_1, P_2(c_2), F\}$.

One-sided Commitment and Selection

A policyholder can choose to switch to another company in the second period, if there are other companies that offer better expected premiums. We will denote the expected premium of the outside option to be $\underline{P}_2$. We assume that the policyholder cannot revert back to its original contract if he is reclassified to a different risk group after switching.

An existing policyholder may experience health deterioration and hence face higher premiums in the second period. We assume that the expected premium of the outside option is correlated with the second period valuation r_2. More specifically, the premium of the outside option is denoted as $\underline{P}_2(r_2)$, a stochastic variable for any given second period valuation r_2. Furthermore, we assume that the biased selection only applies to consumers who were not covered in the first period.

We also assume that a consumer cannot find a better insurance contract than the one offered unless he experiences some changes in health status after the first period. This can only happen if the consumer is a policyholder in the first period. In essence, a consumer who chooses to forgo purchasing insurance in the first period, or a newcomer who chooses to forgo purchasing in the second period, will not find a better insurance deal in the market, and hence receive no coverage for that period.

Timing

Figure 3.6 provides a timeline with a detailed account of the sequence of events in the model. The insurance company's decisions are shown at the top of the timeline, while the consumer's decisions are shown at the bottom. Shocks are marked in italics. The dashed line represents the underlying events that take place implicitly before and following the story described in our model. The solid part of the timeline depicts the sequence of events that we focus on in this framework.

In the first period, the insurance company announces the schedule for period t premiums that depends on the realization of c_t. Consumers proceed to make their investment decision. After the cost shock c_1, the existing consumers decide whether to buy the policy as insurance against the impending death risk at the end of the first period. The model will only describe the purchasing decision of the consumer and take the first period cost shock c_1 as given.

At the beginning of the second period, consumers who did not purchase in the first period can choose to invest μ in order to buy in the second period. After the investment μ is

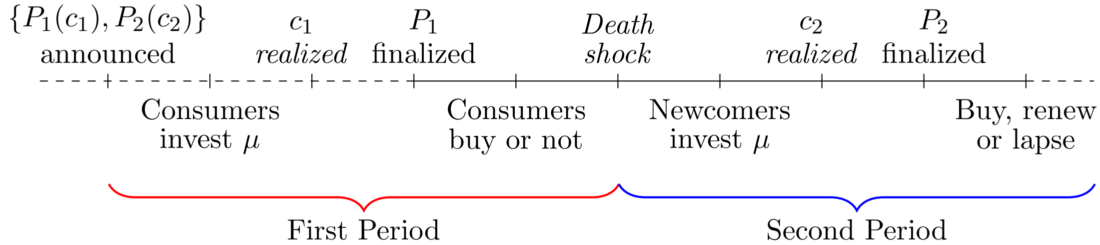


Figure 3.6: Timing of events

sunk, the cost shock is realized and the insurance company finalizes their premium P_2. Having observed P_2, the newcomers can choose to actually sign the contract or not, while the existing policyholders decide whether to renew, switch or lapse.

3.4.2 Characterizing the Demand

Let $P_1(c_1)$ and $P_2(c_2; c_1)$ be the premium schedules announced by the life insurance company, where c_1 is given. For simplicity, we will refer to the premium schedules as $\{P_1, P_2(c_2)\}$. We use backward induction to characterize the demand for life insurance. Consumers, irrespective of whether they are newcomers or existing policyholders, would purchase life insurance in the second period if the realized second period premium is less than their second period valuation: $r_2 \geq P_2(c_2)$. Therefore, newcomers would choose to invest if the following inequality holds

$$V_2^N(r_2) \equiv \Pr\left(P_2(c_2) \leq r_2\right) \mathrm{E}\left[r_2 - P_2(c_2) \mid P_2(c_2) \leq r_2\right] - \mu \geq 0. \tag{3.1}$$

In other words, the newcomers choose to invest in the second period if the expected payoff from owning life insurance less the transaction cost, denoted $V_2^N(r_2)$, is greater than the outside value of 0.

We focus on the premiums that are efficient. In essence, the expected payoff for the consumers monotonically increases with r_2. Afterwards, we will show that the optimal premium schedules are indeed efficient in Proposition 1. We call a second period valuation for which inequality (3.1) binds the second period threshold valuation of the newcomers, and denote it as $\bar{r}_2^N$ which is unique. This implies that consumers with $r_2 \geq \bar{r}_2^N$ would invest μ in the second period, if they didn't purchase in the first period.

A consumer would also evaluate the merits of renewing against the benefit of switching prior to purchasing in the first period. The expected value of switching is

$$V_2^S(r_2) \equiv \Pr\left(\underline{P}_2(r_2) \leq r_2\right) \mathrm{E}\left[r_2 - \underline{P}_2(r_2) \mid \underline{P}_2(r_2) \leq r_2\right] - \mu.$$

We assume that for any second period valuation r_2, $V_2^S(r_2)$ is strictly positive. Since our model focuses on consumers with coverage needs that extend to multiple periods, a consumer who purchases life insurance in the first period would have to satisfy the following inequality

$$V_2^R(r_2) \equiv \Pr\left(P_2(c_2) \leq r_2\right) \mathrm{E}\left[r_2 - P_2(c_2) \mid P_2(c_2) \leq r_2\right] \geq V_2^S(r_2). \tag{3.2}$$

In other words, for an existing policyholder with valuation r_2 the second period expected payoff from renewing is at least as high as the expected payoff from switching. Notice that an existing policyholder who renews does not need to invest μ. As a result, combining (3.1) and (3.2), we have the following relationship between the payoffs for a newcomer and for an existing policyholder

$$V_2^R(r_2) = V_2^N(r_2) + \mu.$$

We call a second period valuation for which inequality (3.2) binds the second period threshold valuation for policyholders, and denote it as $\bar{r}_2^E$. This implies that consumers who bought coverage in the first period have at least a second period valuation of $\bar{r}_2^E$. If not, then they wouldn't have invested in the first period. Assuming that $V_2^R(r_2)$ and $V_2^S(r_2)$ only cross each other once, $\bar{r}_2^E$ is unique.

In the first period, following the investment of μ, a consumer with valuation (r_1, r_2) such that $r_2 \geq \bar{r}_2^E$ will actually purchase life insurance if the following inequality is satisfied

$$(r_1 - P_1) + (1 - m_1)\, V_2^R(r_2) \geq (1 - m_1) \max\left\{V_2^N(r_2), 0\right\}. \tag{3.3}$$

First note that $V_2^R(r_2)$ is always greater or equal to zero. The inequality states that consumers would purchase life insurance in the first period if the value of purchasing coverage in the first period with the expected benefit of renewing in the second period is greater than the expected payoff from delaying the purchasing to next period. Notice the consumers who purchase in the first period do not need to invest the transaction cost μ again in the second period, which factors into the consumer's decision to purchase. Also, since we assumed that a consumer without coverage in the first period would not experience a change in health status, the expected payoff from delaying the purchase to the second period does not include $V_2^S(r_2)$.

Next, we examine the purchasing decision of consumers with different second period valuations. For a consumer with $r_2 \geq \bar{r}_2^N$, inequality (3.3) can be expressed as $r_1 \geq P_1 - (1 - m_1)\,\mu$. On the other hand, for consumers with $r_2 < \bar{r}_2^N$, inequality (3.3) can be written as $r_1 \geq P_1 - (1 - m_1)\,V_2^R(r_2)$, which is dependent on the second period valuation r_2. For $r_2 \geq \bar{r}_2^N$, let $\delta = P_1 - (1 - m_1)\,\mu$ which is independent of r_2. For $r_2 < \bar{r}_2^N$, let $\chi(r_2) = P_1 - (1 - m_1)\,V_2^R(r_2)$ which is monotonically decreasing in r_2 under our efficiency assumption. Since if $r_2 < \bar{r}_2^N$ we have $V_2^N(r_2) < \mu$, then a consumer with lower second period valuation requires a higher first period valuation for him to purchase first period coverage.

We assume $\mu > V_2^S(\bar{r}_2^E)$. Under this assumption, it immediately follows that $\bar{r}_2^N > \bar{r}_2^E$ and there exists a benefit for early coverage, since the cost for postponing is high. As a result, there exists a positive mass of policyholders who will not buy insurance in the second period unless they can renew the policy purchased in the first period. In other words, some consumers with $r_2 < \bar{r}_2^N$ and sufficiently high r_1 buy in the first period to avoid the higher cost incurred by postponing.[4]

Figure 3.7 presents a stylized graphical illustration of the consumers' investment decisions depending on their private valuation (r_1, r_2). In the graph, the red area represents the mass of consumers who pay the transaction cost in period one. The blue area represents the newcomers who make the investment of μ in the second period. It should be noticed that the consumers' eventual purchase decision depends on the announced prices in both

[4] Another case is $\mu \leq V_2^S(\bar{r}_2^E)$. For this case, $\bar{r}_2^N \leq \bar{r}_2^E$ and many consumers with low first period valuation would prefer to defer purchasing coverage till the second period. There is no benefit in purchasing in the first period unless the consumer's first period and second period valuations are both sufficiently high. It should be noted that the main results on the optimal premium schedule would still hold for this case.

periods. Also, $\chi(\cdot)$ is represented on the graph as a straight line for simplicity, in fact it may be a nonlinear function.

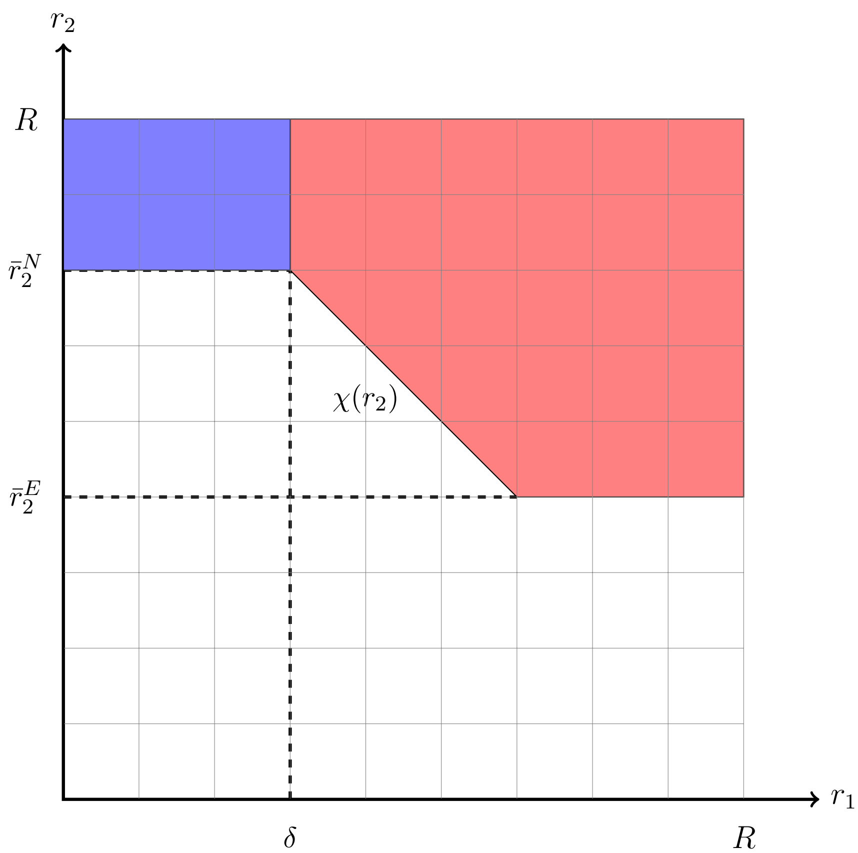


Figure 3.7: Distribution of consumers and their investment decisions

We derive the first period demand function by integrating the probability distribution function of the consumers

$$D_1(P_1, P_2(c_2)) = \int_{\bar{r}_2^N}^{R} \int_{\delta}^{R} h(r_1, r_2)\, dr_1 dr_2 + \int_{\bar{r}_2^E}^{\bar{r}_2^N} \int_{\chi(r_2)}^{R} h(r_1, r_2)\, dr_1 dr_2. \tag{3.4}$$

The first part of the demand represents the consumers with $r_2 \geq \bar{r}_2^N$, while the second part stands for the consumers with $r_2 < \bar{r}_2^N$. The demand is weakly decreasing in P_1.

To characterize the demand in the second period, we first focus on the demand of the newcomers, which can be expressed as

$$D_2^{buy}(P_1, P_2(c_2)) = (1 - m_1) \int_{\max\left\{\bar{r}_2^N, P_2(c_2)\right\}}^{R} \int_0^{\delta} h(r_1, r_2) dr_1 dr_2. \quad (3.5)$$

Similarly, the demand function is weakly decreasing in the second period premium $P_2(c_2)$. In particular, the newcomers' demand stays constant when the premiums are below $\bar{r}_2^N$. In other words, the company would not attract new consumers by lowering the second period premiums below $\bar{r}_2^N$. However, they may still choose to do so in order to attract more policyholders to renew their contract.

The demand of the existing policyholders for renewing is

$$\begin{aligned} D_2^{renew}(P_1, P_2(c_2)) =& (1 - m_1) \int_{\max\left\{\bar{r}_2^N, P_2(c_2)\right\}}^{R} \int_{\delta}^{R} h(r_1, r_2) dr_1 dr_2 \\ &+ (1 - m_1) \int_{\max\left\{\bar{r}_2^E, min\{P_2(c_2), \bar{r}_2^N\}\right\}}^{\bar{r}_2^N} \int_{\chi(r_2)}^{R} h(r_1, r_2) dr_1 dr_2. \end{aligned} \quad (3.6)$$

The first part of the demand represents the policyholders with $r_2 \geq \bar{r}_2^N$, while the second part of the demand stands for the policyholders with $r_2 < \bar{r}_2^N$. The min function in the lower integral limit makes sure the demand is non-negative. The demand is weakly decreasing in $P_2(c_2)$. In particular, the policyholders' demand is constant when the premiums are below $\bar{r}_2^E$.

We define the second period demand as the aggregate demand of the newcomers and the existing policyholders

$$D_2(P_1, P_2(c_2)) = D_2^{buy}(P_1, P_2(c_2)) + D_2^{renew}(P_1, P_2(c_2)).$$

As intimated above, the company cannot increase demand by lowering the premiums below $\bar{r}_2^E$. As a result, the demand becomes perfectly inelastic for any second period premium smaller than the threshold.

3.4.3 Incentive Compatibility

The insurance company faces a different demand once the consumers' investment μ is sunk. Before the investment of μ, the demand in the second period is downward sloping for all prices above the investment cost μ and the insurance company would like to charge a monopoly price. We refer to this pre-investment demand as the *ex-ante* demand and the demand following the investment of μ as the *ex-post* demand. (The kink in the *ex-ante* demand comes from the fact that even if the insurance company sets a premium of zero, only the consumers with valuation of at least μ would consider buying.) The *ex-ante* and *ex-post* demand are different which creates a time inconsistency problem for the insurance company. In particular, after the investment of μ, the insurance company loses the incentive to announce low premiums since it cannot attract any new consumers. Furthermore, the existing policyholders partially revealed their second period valuation by investing and purchasing in the first period. By taking advantage of this opportunity, the insurance company has incentive to increase second period premiums up to $\bar{r}_2^E$, because it knows the demand is inelastic for any prices below it.

The disparity between the *ex-ante* demand and the *ex-post* demand is the reason why the insurance company needs to commit to keeping its promises. However, private cost shocks create an incentive to tailor the premium according to the shocks. This tension generates a trade-off between commitment and flexibility. To resolve this tension, in our model the life insurance company disciplines its pricing behavior by setting incentive compatible premium schedules. In essence, the company's choice of the finalized premium amount is restricted to the promised schedule that corresponds to the true cost realization.

To begin, we divide the possible cost realizations into three regions. We can define the following cost regions for the second period cost shocks

$$\begin{aligned}
\mathcal{C}_2^h &= \left\{ c_2 \mid P_2(c_2) \geq \bar{r}_2^N \right\}, \\
\mathcal{C}_2^m &= \left\{ c_2 \mid \bar{r}_2^E < P_2(c_2) < \bar{r}_2^N \right\}, \\
\mathcal{C}_2^l &= \left\{ c_2 \mid P_2(c_2) \leq \bar{r}_2^E \right\},
\end{aligned}$$

and let $\underline{\mathcal{C}}_2 = \mathcal{C}_2^l$ and $\overline{\mathcal{C}}_2 = \mathcal{C}_2^m \cup \mathcal{C}_2^h$.

We will proceed by formulating the incentive compatibility constraints for the second period. Using the newly defined cost regions we can express the incentive compatibility constraints as

$$[P_2(c_2) - c_2]\, D_2(P_1, P_2(c_2)) \geq [P_2(\tilde{c}_2) - c_2]\, D_2(P_1, P_2(\tilde{c}_2)), \tag{3.7}$$

$\forall c_2 \in \mathcal{C}_2^i, \tilde{c}_2 \in \mathcal{C}_2^j$ and $i, j \in \{h, m, l\}$. There are a total of nine second period incentive compatibility constraints. Three of those are to deter deviations within the same cost regions (when $i = j$). Three are to deter downward deviations, and the rest are to deter upward deviations. The incentive compatibility constraints for the second period are written out in full in the appendix.

We will now briefly describe how the incentive compatibility works in the context of our model. The constraints serve to reduce the possible set of premiums the life insurance company can choose from in each period. This set is announced at the beginning of the first period, and it is implicitly assumed that in both periods, any finalized premium that is not within the set can be subjected to legal measures. The only requirement so far is that the set be incentive compatible.

3.4.4 Equilibrium Definition and the Optimization Problem

We model the premium schedule as a mapping from the cost shock realizations announced by the insurance company to the dollar premium amount. We consider a sequential equilibrium of this model. The *sequential* equilibrium has the insurance company choosing the premium schedule in each period while taking into account the future premiums chosen by the future insurance company (a future-self). In the language of mechanism design, a sequential equilibrium has the insurance company choosing a *static* premium that maximizes profit at every history given its future-self will do the same. In essence, a sequential equilibrium results from the solution of two separate static maximization problems in our model.

The sequential equilibrium is a contract $\{P_1, P_2(c_2)\}_{c_2 \in [\underline{c}, \overline{c}]}$ that solves the second period static optimization problem

$$\Pi_2(P_1) \equiv \max \int_{\underline{c}}^{\overline{c}} (P_2(c_2) - c_2) D_2(P_1, P_2(c_2)) dG_2(c_2), \tag{3.8}$$

subject to (3.7) and

$$P_2(c_2) \geq 0, \forall c_2 \in [\underline{c}, \overline{c}], \tag{3.9}$$

and the first period static optimization problem

$$\max(P_1 - c_1)D_1(P_1, P_2(c_2)) + \Pi_2(P_1), \tag{3.10}$$

$$P_1 \geq 0. \tag{3.11}$$

Inequalities (3.9) and (3.11) ensure that the premium is non-negative. In equilibrium, (3.9) and (3.11) will not bind. However, with $\mu > 0$, it is conceivable that the life insurance company may want to charge a negative premium in the first period to increase the demand for life insurance as described in Klemperer (1987).

We solve for the sequential equilibrium premium schedules by the backward induction method.

3.5 Characterization of the Optimal Premium Schedule

In this section, we characterize the optimal premium schedule. In section 5.1, we obtain some preliminary properties of the life insurance premium. We show that it is rigid for certain realizations of the cost shock, and the optimal premium schedule contains a jump in the premium levels. In section 5.2, we provide a cookbook method for computing the optimal premium schedule using standard mechanism design tools.

3.5.1 Qualitative Features of the Optimal Premium Schedule

To characterize the equilibrium premium schedule, we begin by examining the incentive compatibility constraints. We will first show that the only incentive compatible premiums for cost realizations within $\underline{\mathcal{C}}_2$ are constant.

Lemma 3 *For $c_2 \in \underline{\mathcal{C}}_2$, premiums do not vary with cost.*

Proof: See Appendix.

Lemma 3 tells us that the premium schedules are not sensitive to cost shocks within the set $\underline{\mathcal{C}}_2$. This result comes from the incentive compatibility constraints (3.7) when $i = j = l$. This result gives us a glimpse of the rigidity result. However, it could still be the case that the set $\underline{\mathcal{C}}_2$ is of measure zero, and incentive compatible premium schedules can depend on price for almost all cost shocks. The first part of the following proposition rules this out, and provides a full description of the set of incentive compatible premiums.

Proposition 1 *The set of incentive compatible premiums for the second period has the following properties*

i. $\underline{\mathcal{C}}_2$ *has strictly positive measure.*

ii. For $c_2 \in \overline{\mathcal{C}}_2$, *the company charges monopoly premiums.*

iii. There exists c_2^h *and* c_2^m *with* $c_2^h > c_2^m$ *such that for all* $c_2 \in [c_2^m, c_2^h)$ *we have* $P_2(c_2) = \bar{r}_2^N$.

iv. The optimal premium schedule for the second period $P_2(c_2)$ *is weakly increasing in* c_2.

v. For all $c_2 \in \underline{\mathcal{C}}_2$ *and for all* $c_2' \in \overline{\mathcal{C}}_2$, *we have* $c_2 < c_2'$.

Proof: See Appendix.

Part (i) combined with Lemma 3 delivers the rigidity result. Therefore, it is incentive compatible for premiums to be unresponsive to certain cost shocks. We will denote $\bar{P}_2 = P_2(c_2)$ and the demand $\bar{D}_2 = D_2(P_1, P_2(c_2))$ for all $c_2 \in \underline{\mathcal{C}}_2$.

Part (ii) shows that the insurance company would choose flexible premiums in the cost region $\overline{\mathcal{C}}_2$. In particular, the insurance company would charge a monopoly premium in the second period for cost realizations in $\overline{\mathcal{C}}_2$. Since the company is setting a monopoly premium for the second period within the cost region $\overline{\mathcal{C}}_2$, then it is not profitable to deviate from a cost realization within $\overline{\mathcal{C}}_2$ and report a different cost that is also in the same set. (More specifically, we can refer to the incentive compatibility constraints in the appendix and see that (C.2), (C.3), (C.5) and (C.7) will always hold for the second period.)

Part (ii) also shows that the analysis can be simplified by consolidating $\mathcal{C}_2^m$ and $\mathcal{C}_2^h$. Let $P_2^*(c_2)$ denote the monopoly premium for a given cost shock c_2. This allows us to rewrite the downward deviating incentive compatibility constraints ((C.6) and (C.8)) into the following single incentive compatibility constraint

$$[P_2^*(c_2) - c_2]\, D_2(P_1, P_2^*(c_2)) \geq \left(\bar{P}_2 - c_2\right) \bar{D}_2, \forall c_2 \in \overline{\mathcal{C}}_2, \forall \tilde{c}_2 \in \underline{\mathcal{C}}_2. \tag{3.12}$$

Similarly, we can also rewrite the upward deviating incentive compatibility constraints ((C.9) and (C.10)) into the following incentive compatibility constraint

$$\left(\bar{P}_2 - c_2\right) \bar{D}_2 \geq [P_2^*(\tilde{c}_2) - c_2]\, D_2(P_1, P_2^*(\tilde{c}_2)), \forall c_2 \in \underline{\mathcal{C}}_2, \forall \tilde{c}_2 \in \overline{\mathcal{C}}_2. \tag{3.13}$$

Part (iii) says that the transaction cost creates a kink in the demand for life insurance in the second period. This is due to the fact that in the second period, the firm needs to consider the willingness to purchase for both the newcomers and the existing policyholders above the premium $\bar{r}_2^N$, but only the existing policyholders for any premium below $\bar{r}_2^N$. This generates a region of strictly positive measure of cost realizations for which the premium does not change according to cost.

Part (iv) says that the shape of the optimal premium schedule is monotonically increasing, but only weakly since the premium is rigid for a strictly positive section of cost realizations.

Part (v) combined with part (iv) says that there exits a cutoff c_2^T such that $\underline{\mathcal{C}}_2 = \left[\underline{c}, c_2^T\right]$ and $\overline{\mathcal{C}}_2 = (c_2^T, \overline{c}]$. In particular, for all low cost realizations where $c_2 \leq c_2^T$, the optimal premium does not change according to cost. However, for high cost realizations where $c_2 > c_2^T$, the optimal premium is weakly increasing in cost.

Proposition 1 provides us with a description of the optimal premium schedule. To summarize, we have so far showed that the incentive compatibility constraint is binding for realizations and deviations within $\underline{\mathcal{C}}_2$, and non-binding for all others except for (3.12) and (3.13). We can further simplify the set of incentive compatible contracts with the following lemma.

Lemma 4 *The incentive compatibility constraints (3.12) and (3.13) only bind when the cost is c_2^T.*

Proof: See Appendix.

Lemma 4 allows us to reduce the incentive compatibility constraints (3.12) and (3.13) to a single binding constraint

$$\left[\bar{P}_2 - c_2^T\right] \bar{D}_2 = \left[P_2^*(c_2^T) - c_2^T\right] D_2\left(P_1, P_2^*(c_2^T)\right). \tag{3.14}$$

Another important observation derived from Lemma 4 is that there is an upward discrete jump in the optimal premium schedule for the second period: $P_2^*(c_2^T) > \bar{P}_2$. This result is summarized in the following proposition.

Proposition 2 *The optimal premium schedule has an upward discrete jump at c_2^T such that $P_2^*(c_2^T) > \bar{r}_2^E > \bar{P}_2$.*

Proof: See Appendix.

The insurance company can increase its premiums within a certain range ex-post without changing the demand of the consumers. This is similar to the kinked demand literature in industrial organization. The logic presented above shows us that the life insurance company can convince the consumers that it is increasing the premiums by announcing a premium schedule that is flexible only when the cost shock is sufficiently large. When the cost shock is large, the life insurance company will charge a high price, and Proposition 2 tells us that this price must be high enough so that it would decrease the demand of the consumers. In other words, flexibility is attainable for cost shocks beyond c_2^T because the premium above the threshold induces lapsation since $P_2^*(c_2^T) > \bar{r}_2^E$. When the decline in demand due to high lapsation rates is significant, the life insurance company will not be able to profit from the hold-up problem. Therefore, the life insurance company can credibly respond to cost shocks only when it causes a simultaneous decrease in the demand for insurance. Figure 3.8 presents a stylized illustration of the model's main insights using the second period premium as an example.

Figure 3.8 also demonstrates how the premium rigidity in cost region $\underline{\mathcal{C}}_2$ is not driven by the same mechanics as the premium rigidity generated by the kinked demand in cost region $[c_2^m, c_2^h)$. The premium rigidity for costs below c_2^T are borne out of the insurance company's

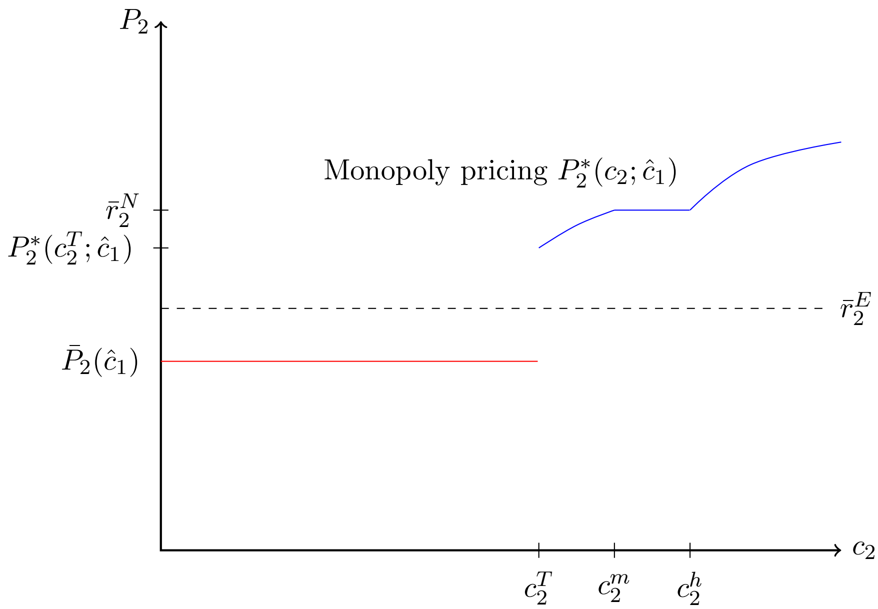


Figure 3.8: Second period incentive compatible premium profile for a given $\hat{c}_1$

need for commitment to avoid holding-up the consumers when the actual cost is realized. This type of rigidity, as we have shown in Proposition 2, can generate a discontinuity in the optimal premium schedule. The premium rigidity for costs within $[c_2^m, c_2^h)$ follows exclusively from the non-differentiability of the demand function at premium $\bar{r}_2^N$. As we can see, the premium schedule is continuous at cost realizations c_2^m and c_2^h. Our paper focuses on the former mechanism for generating premium rigidity.

3.5.2 Computing the Optimal Premium Schedule

In this section, we adopt the standard mechanism design tools to compute the optimal premium schedule. The methods being used here can be found in Krishna Krishna (2010), which were developed in Myerson Myerson (1981). We will take a step back and rewrite the incentive compatibility constraints. First, we introduce some helpful notation. We can rewrite the incentive compatibility constraints as

$$\pi(c_2) \geq \pi(\hat{c}_2, c_2), \forall c_2, \hat{c}_2 \in [\underline{c}, \overline{c}],$$

where $\pi(\hat{c}_2, c_2)$ is the profit of the insurance company when it announces a cost of $\hat{c}_2$ while the true cost is c_2, and $\pi(c_2) = \pi(c_2, c_2)$ is when the company reports truthfully.

Notice that the usual monotonicity of $D_2(P_1, P_2(c_2))$ with respect to second period cost c_2 necessary for incentive compatibility is trivially satisfied. Since the demand function is non-increasing in second period premiums and by part (iv) of Proposition 1, the second period demand is non-increasing in second period costs.

By Proposition 1, the expected profit conditional on the cost realization being greater than the threshold c_2^T is

$$\mathrm{E}\left[\pi(c_2)\middle| \, c_2 \geq c_2^T\right] = \int_{c_2^T}^{\overline{c}} \pi^*(c_2)dc_2,$$

where $\pi^*(c_2)$ denotes the profit under monopoly pricing. It is also straight forward to show that

$$\mathrm{E}\left[\pi(c_2)\middle| \, c_2 \leq c_2^T\right] = G_2(c_2^T)\pi^*(c_2^T) + \bar{D}_2 \int_{\underline{c}}^{c_2^T} G_2(c_2)dc_2$$

We can now formulate the optimization problem of the life insurance company. By (3.14), the insurance company chooses the threshold c_2^T and the rigid premium $\bar{P}_2$ that maximizes

$$G_2(c_2^T)\pi^*(c_2^T) + \bar{D}_2 \int_{\underline{c}}^{c_2^T} G_2(c_2)dc_2 + \int_{c_2^T}^{\overline{c}} \pi^*(c_2)dc_2$$

subject to (3.14).

Notice that the optimization problem of the insurance company highlights the trade-off between commitment to a single rigid premium and retaining a flexible premium schedule. First, we discuss why the insurance company would like to choose a single rigid premium. If the premium schedule is a singleton, both the last term in the objective function and (3.14) disappear. In other words, the insurance company doesn't need to pay the incentive costs for having the option to change its premiums, as highlighted by Proposition 2. Finally, to see why the insurance company might also want to retain some flexibility, notice that as $c_2^T \to \underline{c}$, the first term in the insurance company's expected profit increases. Also, since $\pi(c_2)$ is a decreasing convex function and by the envelope theorem, $\pi'(c_2) = -D_2(P_1, P_2(c_2))$ for all $c_2 \in [\underline{c}, \overline{c}]$ where $\pi(c_2)$ is differentiable, we must have decreasing demand with respect to c_2. Therefore, the life insurance company would like to retain some flexibility to increase the first three terms of the optimization problem if the increase in $\bar{D}_2$ is faster than the decrease in $G_2(c_2^T)$ as c_2^T decreases. However, if the gains from having flexibility is outweighed by the incentive costs, then it would be optimal for the company would choose a singleton for the premium schedule. We can easily rule out this corner solution by assuming that $\bar{c}$ is sufficiently large and relatively high cost shocks happen with a large probability.

It is worth repeating that the optimal solution for the cutoff c_2^T is not $\underline{c}$. To see why this is the case, notice that

$$\bar{D}_2 = (1-m_1)\left[\int_{\bar{r}_2^N}^{R}\int_0^R h(r_1,r_2)dr_1dr_2 + \int_{\bar{r}_2^E}^{\bar{r}_2^N}\int_{\chi(r_2)}^{R} h(r_1,r_2)dr_1dr_2\right],$$

where by the definition of $\bar{r}_2^E$ and Proposition 1,

$$\bar{r}_2^E = \frac{V_2^S(\bar{r}_2^E)}{G(c_2^T)} + \bar{P}_2.$$

When $c_2^T \to \underline{c}$, we have $\bar{r}_2^E \to \infty$, since $V_2^S(r_2) > 0$ for all values of r_2. As a result, if c_2^T is too small, (3.14) would be violated and $\bar{D}_2$ would be negative.

Finally, to solve for the optimal cutoff c_2^T and the rigid premium $\bar{P}_2$, we derive the first order conditions and use (3.14) to express $\bar{P}_2$ as a function of c_2^T. The result is summarized by the next proposition.

Proposition 3 *The cutoff c_2^T and the rigid premium $\bar{P}_2$ in the optimal second period premium schedule are given by (3.14) and the following equation:*

$$\frac{G_2(c_2^T)}{g_2(c_2^T)}\left[1+\left(1-G_2(c_2^T)\right)\right]D_2\left(P_1,P_2^*(c_2^T)\right) - \int_{c_2^T}^{\bar{c}}\left[1-G_2(c_2)\right]D_2\left(P_1,P_2^*(c_2)\right)dc_2 \quad (3.15)$$
$$= \Phi(c_2^T)\bar{D}_2 + \frac{G_2(c_2^T)}{g_2(c_2^T)}\left[1-G_2(c_2^T)\right]c_2^T\frac{\partial\bar{D}_2}{\partial c_2^T},$$

where $\Phi(c_2^T) = \left[1-2G_2(c_2^T)\right]c_2^T + \frac{G_2(c_2^T)}{g_2(c_2^T)} + \int_{\underline{c}}^{c_2^T} G_2(c_2)dc_2$.

Proof: See Appendix.

Proposition 3 gives the two equations that pin down the optimal premium schedule $\left(c_2^T, \bar{P}_2\right)$ for the second period. For any given optimal $c_2^T \in (\underline{c}, \overline{c})$, (3.14) gives us the the optimal second period rigid premium level $\bar{P}_2$. Notice that (3.15) gives us the optimal second period cutoff c_2^T. First off, the left hand side of (3.15) is negative if $c_2^T = \underline{c}$ while the right hand side is strictly positive. Therefore, the optimal second period premium schedule is never fully flexible, confirming part (i) of Proposition 1.

Finally, once we have the optimal second period premium schedule, we are now ready to characterize the optimal first period premium using backward induction. This is presented in our next proposition.

Proposition 4 *Given the optimal second period premium schedule $P_2(c_2)$, the optimal first period premium is characterized by the following equation*

$$D_1(P_1, P_2(c_2)) + (P_1 - c_1)\frac{\partial D_1}{\partial P_1} + \frac{\partial \Pi_2}{\partial P_1} = 0. \tag{3.16}$$

Furthermore, let P_1^{pNR} be the optimal premium for the non-renewable life insurance contract without transaction cost μ and P_1^{pR} be the optimal premium for the renewable life insurance contract without transaction cost. We have $P_1^{pNR} = P_1^{pR} > P_1$.

Proof: See Appendix.

The second part of Proposition 4 states that the life insurance company will charge a lower first period premium than compared to an environment without transaction costs. To see why this is the case, we can imagine a static market. The demand for the contract would be $D(P) = 1 - H(P + \mu)$ and the optimal premium would be described by $P = \frac{1-H(P+\mu)}{h(P+\mu)} + c$. Since the hazard rate for the distribution of private values is assumed to be increasing, then $P(\mu > 0) < P(\mu = 0)$. The same logic applies to Proposition 4. This shows that the transaction cost μ generates a first period "teaser" rate to attract consumers. This is similar to results exhibited in switching cost models like Klemperer (1987).

3.6 Alternative Explanations

In this section, we attempt to briefly describe some alternative theories that could potentially explain our empirical findings and then attempt to rule them out.

3.6.1 Menu Cost

We begin by investigating the relationship between premium duration and the relative size of adjustment. A plausible hypothesis addressing the observed phenomena is that life insurance companies face a heterogeneous menu cost to changing their premiums, in response to an industry-wide cost shock. In such a world, many companies would postpone any changes until the deviation of the listed price from marginal cost is large enough, and then adjust correspondingly. In that case, the life insurance market should exhibit a positive relationship between the duration of premiums and the size of changes. This relationship is depicted in Figure 3.9, where each point represents a single incidence of a price adjustment. Consequently, the magnitude of change is plotted against the duration of premium. As can be noticed, all points are scattered without a clear pattern and the correlation between the two variables (captured by the regression line) is about 0.13. We conclude that the menu cost is not a relevant theory to explain price rigidities observed in the life insurance market.

3.6.2 Adverse Selection and Avoiding the 'Death Spiral'

A potential explanation that could account for the rigid premiums and the infrequent premium changes observed in the data would be the presence of adverse selection. One way for adverse selection to cause such a phenomenon in the premiums observed would be for the following to concurrently happen. If policyholders possess private information on their health status, high risk individuals may increase their coverage needs relative to low risk individuals. This creates a need for the insurance company to increase their premiums to cover the higher expenses. However, if the price elasticity of demand for life insurance is higher for low risk individuals than high risk individuals, then a small increase in prices may trigger a 'death spiral' that could severely deteriorate the quality of the pool of policyholders. As a result, as long as the increase in coverage demanded by the high risk individuals do not burden the operation of the insurance company, it is reluctant to change its premiums.

This theory is appealing since insurance markets are believed to be plagued by adverse selection. However, Cawley and Philipson (1999) found no evidence of adverse selection

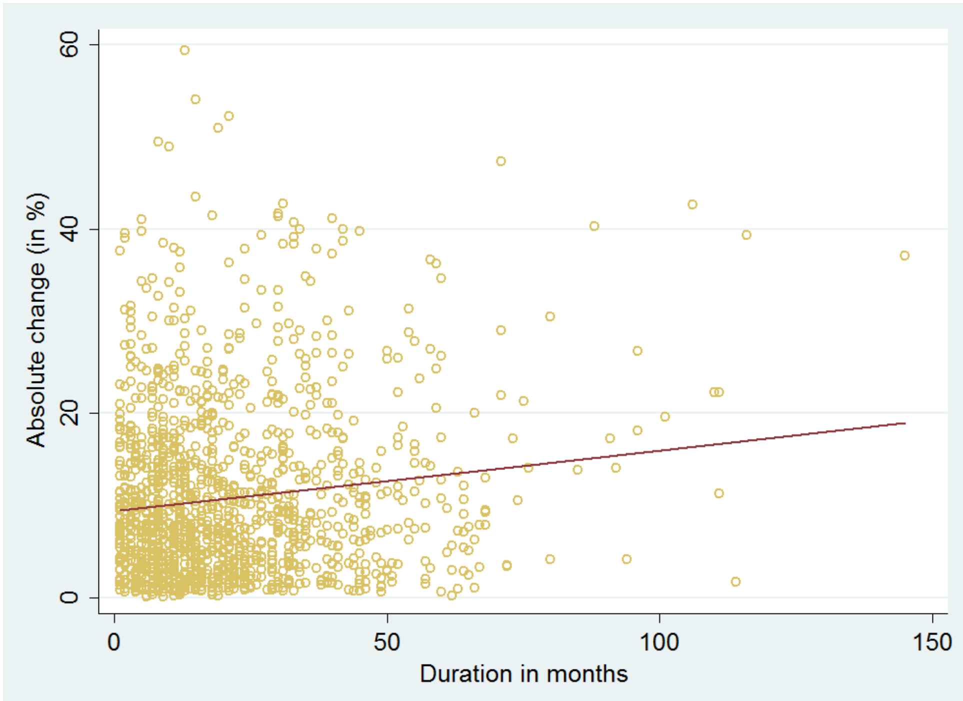


Figure 3.9: Premium duration and adjustment size in the life insurance market

in term life insurance. Without the presence of adverse selection, it is difficult to explain the pricing pattern exhibited by the insurance companies as an attempt to avoid the death spiral.

Even if we assumed the presence of adverse selection in the term life insurance market, Pauly *et al.* (2003) showed that the elasticity of demand for ART with respect to changes in risk is relatively small. In other words, it is unlikely a high risk individuals would be signing a substantially larger amount of insurance coverage compared to low risk individuals. Furthermore, they showed that the price elasticity of demand for ART is also sufficiently small and would require a severe adverse selection problems (significant portion of the policyholders to acquire a substantial difference in information about health status) to trigger a death spiral.

3.7 Conclusion

We show that the market for life insurance has exhibited a remarkable degree of price rigidity over the past two decades. Firms that changed premiums in the analyzed sample did so on average every 39 months, preferring one-time jumps of large magnitude to more frequent and gradual price adjustments. We build a theoretical model to explain this phenomenon, based on the fundamental assumption that consumers are locked-in due to a relationship-specific investment. In line with what we find in the data, the model predicts that premiums will remain constant for a wide range of cost shock realizations, while any potential changes take the form of discrete jumps.

The theory in this paper can be applied more generally to environments plagued by the hold-up problem. This may include worker compensations, trading between upstream and downstream firms, international trade and other settings where the hold-up problem has been documented. The key insight of the model is that there is a cost in indulging in one's temptation, and it is optimal to give in to this temptation only when the cost is sufficiently large. The theory could also be used in models with time inconsistencies. For example, it could potentially shed light on the optimal degree of discretion delegated to a monetary authority with private information and temptation to stimulate the economy through surprise inflation.

References

Aguiar, Mark and Gita Gopinath (2006): "Defaultable debt, interest rates and the current account", *Journal of International Economics* 69, 6483.

Aguiar, Mark and Gita Gopinath (2007): "Emerging Market Business Cycles: The Cycle Is the Trend," *Journal of Political Economy*, 115, 69-102.

Alfaro, Laura and Fabio Kanczuk (2005): "Sovereign debt as a contingent claim: a quantitative approach", *Journal of International Economics* 65, 297-314.

Amador, Manuel (2003): "A Political Economy Model of Sovereign Debt Repayment", *unpublished manuscript*.

Amador, Manuel, Ivan Werning and George-Marios Angeletos (2006): "Commitment vs. Flexibility", *Econometrica*, 74, 365-396.

Andritzky, Jochen (2012): "Government Bonds and Their Investors: What Are the Facts and Do They Matter?", *IMF Working Paper*, 12/158.

Arellano, Cristina (2008): "Default Risk and Income Fluctuations in Emerging Economies", *American Economic Review*, 98, 690-712.

Arellano, Cristina and Ananth Ramanarayanan (2012): "Default and Maturity Structure in Sovereign Bonds", *Journal of Political Economy*, 120, 187-232.

Athey, Susan, Andrew Atkeson and Patrick Kehoe (2005): "The Optimal Degree of Monetary Policy Discrestion", *Econometrica*, 73, 1431-1476.

Bai, Jushan and Pierre Perron (1998): “Estimating and Testing Linear Models with Multiple Structural Changes,” *Econometrica*, 66, 47-78.

Bils, Mark and Peter J. Klenow (2004): “Some Evidence on the Importance of Sticky Prices”, *Journal of Political Economy*, 112, 947-985.

Boz, Emine (2011): “Sovereign default, private sector creditors, and the IFIs”, *Journal of International Economics*, 83, 70-82.

Boz, Emine, Christian Daude and C. Bora Durdu (2011): “Emerging markets business cycles: Learning about the trend”, *Journal of Monetary Economics*, 58, 616-631

Cawley, John and Tomas Philipson (1999): “An Empirical Examination of Information Barriers to Trade in Insurance”, *American Economic Review*, 89, 827-846.

Chatterjee, Satyajit and Burcu Eyigungor (2012): “Maturity, Indebtedness, and Default Risk”, *American Economic Review*, 102, 2674-2699.

Cho, In-Koo and David M. Kreps (1987): “Signaling Games and Stable Equilibria”, *The Quarterly Journal of Economics*, 102, 179-221.

Cole, Harold L., James Dow, and William B. English (1995): “Default, Settlement, and Signalling: Lending Resumption in a Reputational Model of Sovereign Debt”, *International Economic Review*, 36, 365-385.

Conesa, Juan Carlos and Timothy J. Kehoe (2015): “Gambling for Redemption and Self-Fulfilling Debt Crises”, *NBER Working Papers*, 21026.

Daily, Glenn, Igal Hendel and Alessandro Lizzeri (2008): “Does the Secondary Life Insurance Market Threaten Dynamic Insurance?”, *American Economic Review Papers and Proceedings*, 98, 151-156.

D’Erasmo, Pablo (2011): “Government Reputation and Debt Repayment in Emerging Economies”, *unpublished manuscript.*

Eaton, Jonathan and Mark Gersovitz (1981): “Debt with potential repudiation”, *Review of Economic Studies* 48, 289-309.

Fang, Hanming and Edward Kung (2010): "How Does Life Settlement Affect the Primary Life Insurance Market?". *NBER Working Paper* No. 15761.

Gottlieb, Daniel and Kent Smetters (2013): "Lapse-Based Insurance", *Working Paper*.

Guerrieri, Luca, Matteo Iacoviello and Raoul Minetti (2012): "Banks, Sovereign Debt and the International Transmission of Business Cycles", *NBER International Seminar on Macroeconomics*.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright (2007): "The U.S. Treasury Yield Curve: 1961 to the Present", *Journal of Monetary Economics*, 54, 2291-2304.

Hamilton, James D. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica* 57, 357-384.

Hatchondo, Juan Carlos and Leonardo Martinez (2009): "Long-duration bonds and sovereign defaults", *Journal of International Economics*, 79, 1717-125.

Hatchondo, Juan Carlos and Leonardo Martinez (2010): "The politics of sovereign defaults", *Economic Quarterly*, Federal Reserve Bank of Richmond, 96, 291-317.

Hatchondo, Juan Carlos, Leonardo Martinez and Horacio Sapriza (2010): "Quantitative properties of sovereign default models: Solution methods matter", *Review of Economic Dynamics*, 13, 919-933.

Hendel, Igal and Alessandro Lizzeri (2003): "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance". *Quarterly Journal of Economics*, 118, 299-328.

Hortaçsu, Ali and Chad Syverson (2004): "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds". *Quarterly Journal of Economics*, 119, 403-456.

Huntington, Henry S. (1958): "Derivation of Premium Rates for Renewable Term Insurance", *Transactions of Society of Actuaries*, 10, 329-351.

Kehoe, Timothy J. and Edward C. Prescott (2002): "Great Depressions of the Twentieth Century", edited volume of *Review of Economic Dynamics*, 5, 1-18.

Kim, Chang-Jin (1994): "Dynamic Linear Models with Markov-switching", *Journal of Econometrics* 60, 1-22.

Klemperer, Paul (1987): "Markets with Consumer Switching Costs", *Quarterly Journal of Economics*, 102, 375-394.

Koijen, Ralph S. J. and Motohiro Yogo (2015): "The Cost of Financial Frictions for Life Insurers", *American Economic Review*, 105, 445-475.

Koijen, Ralph S. J. and Motohiro Yogo (2016): "Shadow Insurance". *Econometrica*, 84, 1265-1287.

Koszegi, Botond (2005): "On the feasibility of Market Solutions to Self-Control Problems", *Swedish Economic Policy Review*, 12, 71-94.

Krishna, Vijay (2010): "Auction Theory", San Diego, Academic Press.

Krusell, Per, and Anthony Smith (2003): "Consumption-savings decisions with quasi-geometric discounting", *Econometrica*, 71, 365-375.

LIMRA (1996): "Long-Term Ordinary Lapse Survey - U.S.", Connecticut.

Lucas, Robert E. and Thomas J. Sargent (1979): "After Keynesian Macroeconomics", *Federal Reserve Bank of Minneapolis Quarterly Review*, 3, 1-16.

Mimra, Wanda and Achim Wambach (2014): "New Developments in the Theory of Adverse Selection in Competitive Insurance", *The Geneva Risk and Insurance Review*, 39, 136-152.

Mimra, Wanda and Achim Wambach (2011): "A Game-Theoretic Foundation for the Wilson Equilibrium in Competitive Insurance Markets with Adverse Selection", *CESifo working paper*, 3412.

Miyazaki, Hajime (1977): "The Rat Race and Internal Labor Markets", *The Bell Journal of Economics*, 8, 394-418.

Myerson, Roger (1981): "Optimal Auction Design", *Mathematics of Operations Research*, 6, 58-73.

Netzer, Nick and Florian Scheuer (2014): "A Game Theoretic Foundation of Competitive Equilibria with Adverse Selection", *International Economic Review*, 55, 399-422.

Pain, Nigel, Christine Lewis, Thai-Thanh Dang, Yosuke Jin, and Pete Richardson (2014): "OECD Forecasts During and After the Financial Crisis: A Post Mortem", *OECD Economics Department Working Papers*, 1107.

Pauly, Mark and Kate Withers and Krupa Subramanian-Viswanathan and Jean Lemaire and John Hershey and Katrina Armstrong and David Asch (2003): "Price Elasticity of Demand for Term Life Insurance and Adverse Selection," *NBER Working Paper* No. 9925.

Pouzo, Demian and Ignacio Presno 2016: "Sovereign Default Risk and Uncertainty Premia", *American Economic Journal: Macroeconomics*, forthcoming.

Reinhart, Carmen M. and Kenneth S. Rogoff (2009): "This Time is Different: Eight Centuries of Financial Folly", *Princeton University Press.*

Reis, Ricardo (2013): "The Portuguese Slump and Crash and the Euro Crisis", *Brookings Papers on Economic Activity*, 46, 143-193.

Rothschild, Michael and Joseph Stiglitz (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", *The Quarterly Journal of Economics*, 90, 629-649.

Rouwenhorst, K. Geert (1995): "Asset pricing implications of equilibrium business cycle models", *In: Cooley, T.F. (Ed.), "Frontiers of Business Cycle Research"*, Princeton University Press, 294-330.

Sandleris, Guido (2008): "Sovereign defaults: Information, investment and credit", *Journal of International Economics*, 76, 267-275.

Seog, S. Hun (2010): “The Economics of Risk and Insurance”, *Wiley-Blackwell*.

Spence, Michael (1978): “Product differentiation and performance in insurance markets”, *Journal of Public Economics*, 10, 427-447.

Standard & Poor’s (2006): “Default Study: Sovereign Defaults At 26-Year Low, To Show Little Change In 2007”, *Ratings Direct report*.

Wilson, Charles (1977): “A model of insurance markets with incomplete information”, *Journal of Economic Theory*, 16, 167-207.

Appendix A

Appendix to Chapter 1

In this section I show the results of estimating the regime-switching income process directly from the GDP data, as opposed using the information from historical forecasts. Consider again the output equation of the form

$$y_t = \mu_j(1-\rho_j) + \rho_j y_{t-1} + \eta_j \varepsilon_t \tag{A.1}$$

where $\varepsilon_t \sim \mathcal{N}(0,1)$ is an $i.i.d.$ random shock and $\{\rho_j, \eta_j, \mu_j\}_{j=L,H}$ are parameters of the two regimes. Notice that for the purpose of generality I also allow the variance and persistence parameters to be different across the regimes. Regimes change according to a Markov process with the transition probability matrix given by

$$\Pi = \begin{bmatrix} \pi_L & 1-\pi_L \\ 1-\pi_H & \pi_H \end{bmatrix} \tag{A.2}$$

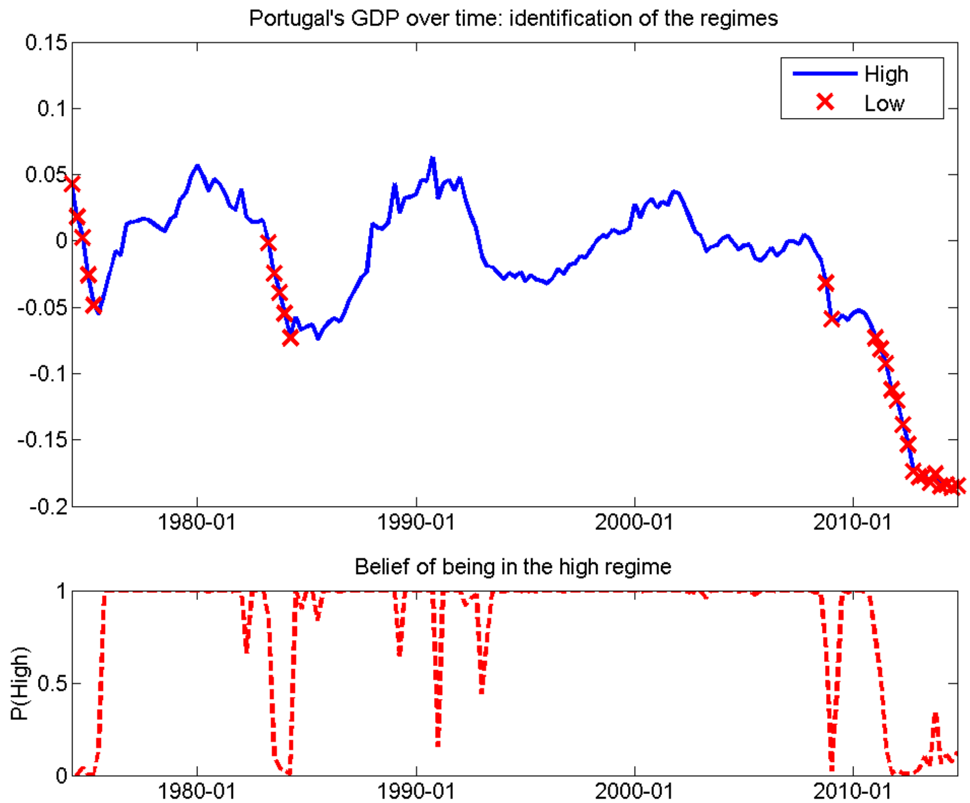
I estimate this process using 1974:Q2-2014:Q4 data for Portugal's GDP, detrended in the same way as it is described in the main text. I use a variant of the Expectation Maximization algorithm proposed by Hamilton (1989) which iteratively maximizes the expected likelihood of observing the sample $Y_T = \{y_1, y_2, ..., y_T\}$. The procedure consists of two generic steps, which are repeated subsequently until convergence. In the first step, given some starting vector of parameters $\theta^n = \{\rho_j^n, \eta_j^n, \mu_j^n\, \pi_{jj}^n\}_{j=L,H}$ the algorithm calculates the smoothed probabilities of being in regime j, conditional on the entire data sample, i.e. $P(z_t = j|Y_T; \theta^n)$ (for details, see Kim (1994)). In the second step, the expected log-likelihood function $\sum_{j=1}^{2}\sum_{t=1}^{T} P(z_t = j|Y_T)\log f(y_t, \theta^n)$ is maximized to obtain an updated vector of parameters θ^{n+1}. The algorithm is terminated when the difference between the parameter vectors obtained in two consecutive iterations is small enough.

Table A.1 summarizes the estimation results. It is worth noticing that the obtained unconditional means in the two regimes are very similar to the numbers obtained in the main text of the paper. The persistence and variance parameters are slightly lower than the ones in the paper. The reason for this is that the transition probabilities between the regimes are much larger now. As a result, a part of the volatility and persistence of the output process is captured by relatively more frequent regimes changes. The switching probabilities obtained here imply the expected duration of the high and low regimes to be 35 and 7 quarters, respectively.

				Transition Prob.	
Regime	Mean μ	Persistence ρ	St. dev. η	Low	High
Low	-0.2244	0.9091	0.0061	0.8576	0.1424
High	0.0171	0.9471	0.0090	0.0285	0.9715

Table A.1: Estimated parameters of the regime-switching endowment process

The upper panel of Figure A.1 plots the detrended Portuguese GDP over time with the periods of identified low regime marked with a red ex. The regimes are determined using "smoothed probabilities" which take into account information from the whole sample (i.e. observations from the past and the future). The lower panel depicts the induced belief of being in the high regime over time (derived with simple Bayesian updating). As can be noticed, the low regime switches on during almost any GDP slump since 1974. Historically these recessions were followed by a prompt recovery, hence the low expected duration of the low regime. It is also important to note that Bayesian learning under this specification is extremely fast - whenever a regime switches, uninformed agents are able to infer it within 1-2 periods. Due to such a fast learning process, the parameters listed in Table 1.3 are not compatible with the theory presented in my paper.

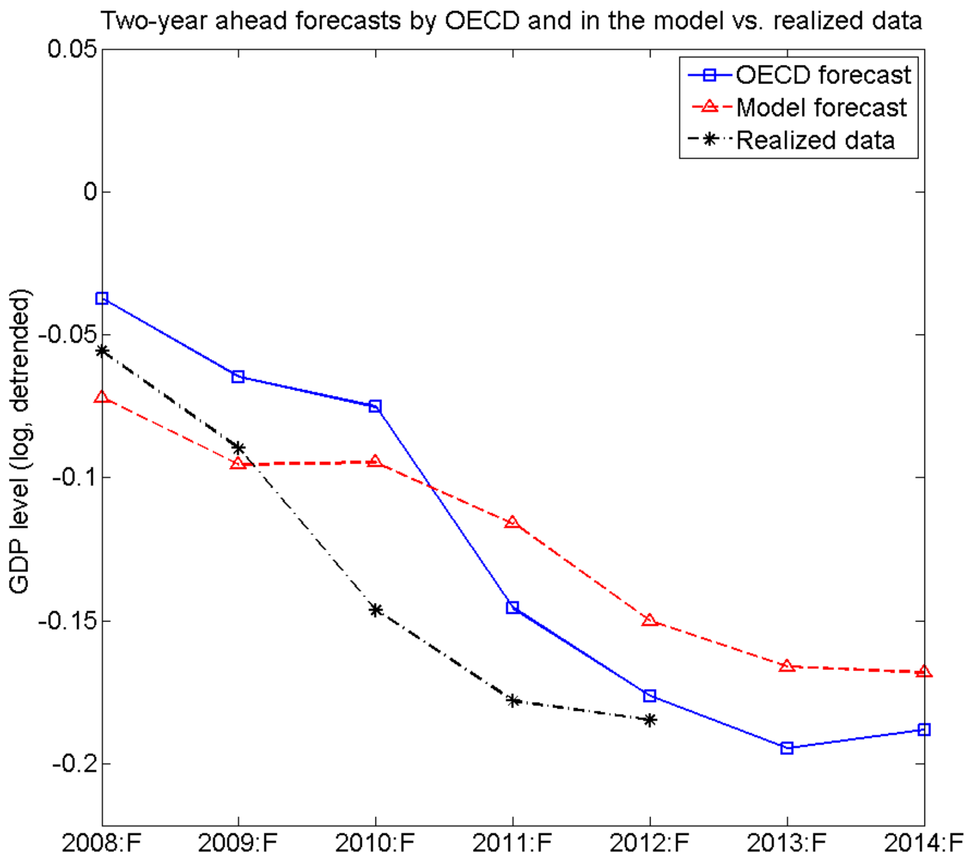


Note: The upper panel plots the GDP data (in logs, detrended) and marks the periods identified as belonging to the low regime with red x's. For any period, the regime is assumed to be low if the corresponding smoothed probability (as described in the text above) falls below 0.5. The lower panel plots the evolution of belief over time, which is updated using Bayes' formula.

Figure A.1: Identification of the regime switches over time for Portugal's GDP

To determine whether these estimates of the income process can capture the actual beliefs of market participants, I once again generate two-year-ahead forecasts using the parameters listed in Table 1.3 and a belief path presented in the lower panel of Figure A.1. Figure A.2 shows the matching of the actual forecasts published by OECD (blue line) and the model generated ones (red dashed line). Both lines are then contrasted with actual data realizations two years lated (black dash-dotted line). As can be noticed, under this set of estimates the model performs really poorly at matching the OECD's expectations of recovery. In particular, early in the recession it predicts much more pessimistic forecasts than in reality (even more than the actual data realizations), while towards the end of it forecast become overly optimistic. The reason is that under these estimates the regimes are not particularly persistent and the prediction for two years ahead is always close to the unconditional mean across the regimes. Moreover, some forecasts are particularly off due to very fast learning. By the fall of 2008 the belief has already dropped quite significantly, leading to a much more pessimistic forecast for 2010 than even the realized data two years later. By contrast, in the most recent part of the crisis (2013-2014) the belief begins to increase again, due to the slight improvement in the arriving GDP data, and consequently, the generated predictions are overly optimistic.

From the evidence presented above, I conclude that applying standard econometric techniques to plain GDP data is not a valid way to obtain meaningful estimates of the regime-switching income process. Such parameter values lead to very fast learning of the market participants and, most importantly, they completely mispredict the evolution of the expectations of recovery. The reason why standard econometric tools are not relevant here is simply the lack of adequate data. Prior to 2008 we do not have any systematic data covering major depression episodes (such as the Great Depression) and as a result, the induced low regime picks up certain downturns that would otherwise be considered standard business cycle fluctuations.



Note: Each point on the graph represents an annual detrended log-GDP level for two years ahead (for example, $2008 : F$ *corresponds to the GDP level in 2010). The solid blue line plots the actual OECD forecasts, while the dashed red line denotes the ones generated by the estimated model. Additionally, the dashed-dot black line shows the actual realized data that the corresponding forecasts refer to (only available until* $2012 : F$*, when the prediction for year 2014 was made).*

Figure A.2: Historical GDP forecasts: OECD- and model-generated predictions.

Appendix B

Appendix to Chapter 2

In this section I outline the algorithm used to compute an equilibrium of the model under different assumptions on the information structure. The model is solved numerically using parallel programming on a computer cluster with 256 processors.

The algorithm is based on the methods proposed in Hatchondo, Martinez and Sapriza (2010) and allows for a continuous choice of next period bonds. The model is solved using value function iteration on equally-spaced grids over the state variables b, y, and p. Continuation values are computed by approximating the corresponding integrals using Gauss-Legendre quadrature and cubic spline interpolation to evaluate off-grid income realizations in the next period. The expected value function is approximated piecewise, taking into account the default threshold and the cutoff income realization above which the default cost is positive. The main computational burden in this model comes from the fact that any given income realization results in Bayesian updating of the belief and actually requires a two-dimensional interpolation for $\big(y', p'(y, y', \widetilde{p})\big)$.

The general procedure to compute an equilibrium proceeds as follows:

1. Start with the initial guesses for v_0^d and v_0^r equal to the continuation values in the last period of the finite-horizon version of the model.[1]

[1] The actual infinite-horizon equilibrium in the model is approximated as the limit of the finite-horizon version. The number of periods is large enough so that the difference between the value and policy functions

2. For every pair of state variables (b, y) and each economy regime z solve the government's problem as specified in equation (2.1) under full separation of types, i.e. if $z = z_L$ $(z = z_H)$ then $\widetilde{p} = 0$ $(\widetilde{p} = 1)$, respectively. Obtain the corresponding optimal policy functions denoted as $b^{*'}_{L,0}$ and $b^{*'}_{H,1}$.

   - If the assumed information structure implies full information, use $b^{*'}_{L,0}$ and $b^{*'}_{H,1}$ as final policy functions for the iteration, update the corresponding repayment value $v^r_1(\mathbf{s}; z)$ and proceed to step 6.

3. For every state $\mathbf{s} = (b, y, p)$ solve the government's problem under incomplete information, i.e. $\widetilde{p} = p$. Obtain the optimal policy functions denoted as $b^{*'}_{L,p}$ and $b^{*'}_{H,p}$.

   - If the assumed information structure implies partial *symmetric* information, the policy functions are computed such that even the government has no knowledge of the current regime and evaluates continuation values in expectation over current z. As a result, use $b^{*'}_{L,p} = b^{*'}_{H,p}$ as the final policy function for this iteration, update the corresponding repayment value $v^r_1(\mathbf{s}; z)$ and proceed to step 6.
   - If the assumed information structure implies partial *asymmetric* information, check whether $b^{*'}_{H,p}$ satisfies the participation constraint (2.16). If so, then $b'_{pool} \equiv b^{*'}_{H,p}$ is the optimal pooling allocation. If not, find a debt level $\hat{\hat{b}}'$ at which the participation constraint binds for the low type, i.e.

$$W_L(\mathbf{s}; \hat{\hat{b}}', \widetilde{p} = p) = W_L(\mathbf{s}; b^{*'}_{L,0}, \widetilde{p} = 0) \tag{B.1}$$

   where W_L is the lifetime utility defined in equation (2.12). Then, $b'_{pool} \equiv \hat{\hat{b}}'$ is the optimal pooling allocation.

4. For every state $\mathbf{s} = (b, y, p)$ calculate a debt allocation $\hat{b}'$ at which the low type's incentive compatibility constraint (2.13) binds. Formally, $\hat{b}'$ is such that:

$$W_L(\mathbf{s}; \hat{b}', \widetilde{p} = 1) = W_L(\mathbf{s}; b^{*'}_{L,0}, \widetilde{p} = 0) \tag{B.2}$$

---

at two subsequent iteration are negligible. This strategy is used to avoid the common problem of multiplicity of Markov Perfect equilibria for infinite-horizon economies discussed in Krusell and Smith (2003).

Equation (B.2) states that the lifetime utility of choosing $\hat{b}'$ and pretending to be a high type is equal to the lifetime utility of selecting a benchmark separating contract of the low type. Define $b'_{sep} \equiv \max\{\hat{b}', b^{*'}_{H,1}\}$ as the high type's optimal separating allocation.

5. Solve the high type's problem by choosing whether to separate or pool, i.e.

$$v_1^r(b, y, p; z_H) = \max\left\{W_H(\mathbf{s}; b'_{sep}, \widetilde{p} = 1), W_H(\mathbf{s}; b'_{pool}, \widetilde{p} = p)\right\} \tag{B.3}$$

and use the corresponding debt choice as the high type's final policy function for the iteration. Record the low type's policy depending on the high type's preceding selection.

   - In the case of separating equilibrium, the low type's policy function is $b^{*'}_{L,0}$.
   - In the case of pooling equilibrium, the low type's policy function is b'_{pool}.

6. Compare the two repayment values. If $||v_1^r - v_0^r|| < \varepsilon$ for some desired tolerance level ε then iterations are over and the value function has converged. Otherwise, set $v_0^r = v_1^r$, update the value of default by iterating on equation (2.2) and go back to step 2.

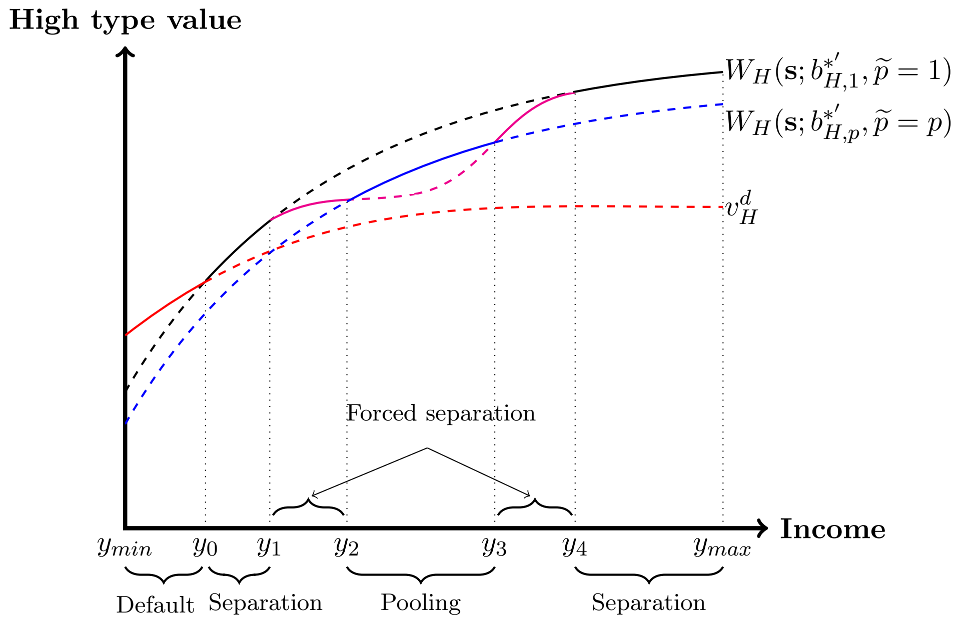
The above algorithm naturally results in non-standard value functions that may feature discontinuities and jumps related to the separating and pooling equilibria arising for different states. Figure B.1 presents a stylized illustration of value functions of the two types that depend on current income level (underpinning the actual policies depicted in Figure 2.1). The two panels present the high and low type's general value (marked with a solid line) as a combination of the lifetime utility under different equilibrium outcomes. The black lines denote each type's lifetime utility from a separating contract calculated in step 2 of the algorithm, while the blue lines represent the utility from a pooling contract in which the low adopts high type's debt amount $b^{*'}_{H,p}$ obtained in step 3.[2] As can be noticed on the figures, for very low income levels lying in the interval (y_{min}, y_0), both government types find it optimal to default. For income falling between (y_0, y_1), the high type repays and

[2]Assume for simplicity that the belief p is high enough so that the participation constraint (2.16) is slack.

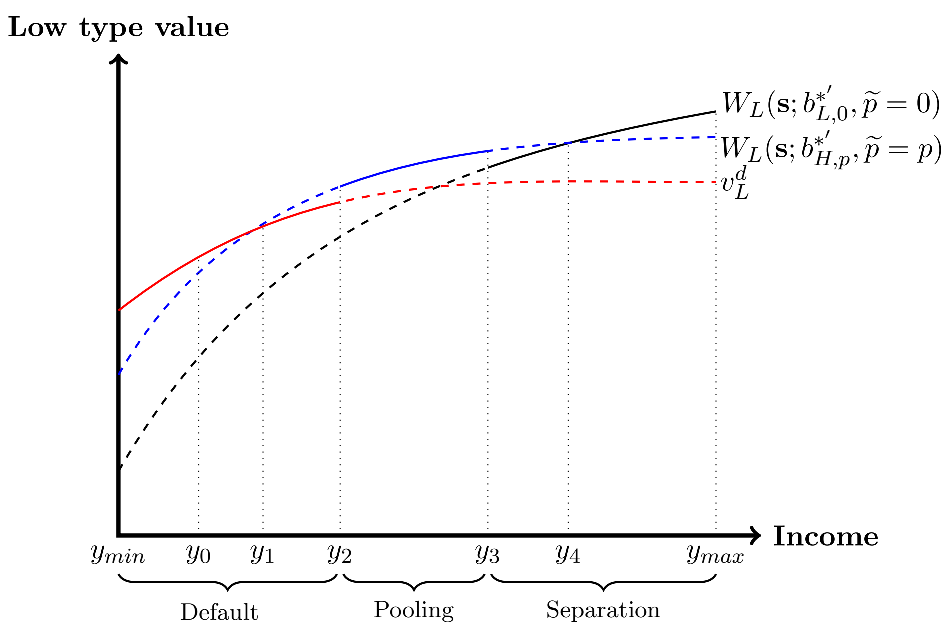
selects its benchmark separating policy $b^{*'}_{H,1}$ while the low type prefers to default even if mimicking was a feasible choice. Next, for current shock in the interval (y_1, y_2), the low type would like to repay and mimic the high type. This is not possible however, as the high type prevents pooling by increasing next period debt up to $\hat{b}'$, the borderline mimicking amount computed in step 4. Merely repaying the debt is no longer enough to separate, and consequently the high type's value over this interval (marked with the solid pink line) is lower than what it would be in the full information world (dashed black line). Because the low type's IC constraint (2.13) binds here, I call this equilibrium outcome a "forced separation".

For the income shock lying between (y_2, y_3) we have a pooling equilibrium. In this interval the low type is most determined to mimic and therefore the high type's value of separation falls below the lifetime utility of a pooling contract. As a result, both types settle on the optimal policy equal to $b^{*'}_{H,p}$. As can be noticed from Figure B.1(b) the low type's value exhibits an upward jump. That is because the decision whether to separate or pool is fully in control of the high type and remains exogenous (but very beneficial) to the low type.

As current income further rises above y_3, the interest rates get closer to a risk-free rate and the low type's incentive to mimic gets weaker. As can be noticed in Figure B.1(a), the value of separation exhibits an inflection point beyond which it rises faster than the high type's utility from pooling. Consequently, in the interval (y_3, y_4) we have another "forced" separating equilibrium in which the high type increases the debt up to $\hat{b}'$ in order to prove its identity to the lenders. Finally, for output levels above y_4 we obtain a regular separating equilibrium where the IC constraint (2.13) is slack and the low type has very little to gain from distorting its own benchmark policy $b^{*'}_{L,0}$ derived in step 2 of the algorithm.



(a) **High type**



(b) **Low type**

Figure B.1: Value functions in the asymmetric information model

Appendix C

Appendix to Chapter 3

Estimating the marginal cost of a life insurance company

In what follows, let $m_{t,n,\bar{n}}$ denote the period t mortality rate of age n individuals who bought life insurance at age $\bar{n}$[1], and let N be the maximum attainable age according to the corresponding mortality tables. Let $R_t(i)$ be the (annualized) interest rate on zero-coupon risk-free securities with maturity i at time t. Further, denote by $\ell_{x,\bar{n},i}$ the lapsation rate after the ith year of an x-year level-term insurance policy, where $x \in \{1, 5, 10, 20\}$, purchased at age $\bar{n}$ (i.e. the probability that a policyholder fails to pay in time the outstanding premium). Finally, let $\gamma_{t,n,\bar{n}}$ stand for the growth rate in premium between age n and $n+1$ for a schedule posted at time t[2]. We define the net premium for an x level-term policy acquired at age n per dollar of death benefit as

$$V_t^x(n) \equiv \left(\sum_{i=1}^{N-n} \frac{\prod_{j=0}^{i-2}(1 - m_{t,n+j,n})(1 - \ell_{x,n,j+1}) m_{t,n+i-1,n}}{R_t^i(i)} \right) \left(1 + \sum_{i=1}^{N-n-1} \frac{\prod_{j=0}^{i-1}(1 - m_{t,n+j,n})(1 - \ell_{x,n,j+1}) \gamma_{t,n+j,n}}{R_t^i(i)} \right)^{-1} \tag{C.1}$$

[1] It is important to keep track of different cohorts of the insured due to adverse selection, i.e. individuals who have already held a policy tend to have significantly higher mortality rates than the same-age newcomers.

[2] Notice that, by construction, $\gamma_{t,n,\bar{n}} = 1$ for all n such that $mod(n - \bar{n}, x) \neq 0$, i.e. depending on the policy term, premiums are allowed to increase only every x years.

The net premium of a renewable term life insurance is difficult to calculate and to the best of our knowledge there is no agreed-upon way to do so. The first problem is that the premium schedule is increasing in age and different companies apply different growth schemes. The second problem relates to lapsations which are not modeled or predicted easily. In the pricing of other life insurance products (e.g. universal or whole life insurance), characterized by significant front-loading of premium schedules, lapsation rates are often disregarded since a policyholder's incentive to lapse decreases over time. With short-term renewable policies though, lapsation is an important factor because most customers acquire them for a limited number of years only. For this reason, in equation (C.1) we propose a modified version of the standard formula (see e.g. Koijen and Yogo (2015)), in that we account for lapsation as well as the premium schedule that increases in age.

In order to obtain actuarially fair growth rates of the premium, we implement the following algorithm. First, compute formula (C.1) assuming the policy is a whole-life insurance (premiums are constant forever, i.e. $\gamma_{t,n+j,n} = 1,\ \forall j > 0$). Having obtained an increasing sequence of actuarial values at renewal dates across different age profiles $V_t^x(n)$, $V_t^x(n+1)$, ..., $V_t^x(N-1)$, we next calculate the resulting actual growth rates, i.e. $\gamma_{t,n+j,n} = V_t^x(n+j+1)/V_t^x(n+j),\ \forall j > 0$. Finally, we plug the growth rates obtained in this way in formula (C.1) to compute the net premium across time. It is important to notice that a life insurance company can choose among many different premium growth patterns which are all actuarially fair (i.e. they all equate the present expected value of cash flows between the company and the consumer). The advantage of our method is that it pins down the growth rates that result purely from increased mortality rates due to aging.

In our calculation of the net premium we use the mortality tables issued by the American Society of Actuaries. We apply the 1980 Commissioners Standard Ordinary (CSO) table for all years prior to January 2001, the 2001 Valuation Basic Table (VBT) prior to January 2008 and the 2008 VBT for the time period following January 2008. We use geometric averaging on the monthly basis to smooth the transition between any two vintages of the mortality tables. It is important to emphasize that these tables are created based on the actual mortality rates among the insured rather than the general population. For this reason, they account for a potential adverse selection in the market for life insurance.[3] As for the lapsation data, we use the rates published for yearly renewable term products in LIMRA (1996). As the risk-free interest rate we use the U.S. Treasury zero-coupon yield curve[4].

Incentive Compatibility Constraints: From (3.7), we can write out all nine incentive compatibility constraints for the second period for any $\hat{c}_1$:

$$[P_2(c_2) - c_2]\, D_2(P_1, P_2(c_2)) \geq [P_2(\tilde{c}_2) - c_2]\, D_2(P_1, P_2(\tilde{c}_2)), \forall c_2, \tilde{c}_2 \in \mathcal{C}_2^h \tag{C.2}$$

$$[P_2(c_2) - c_2]\, D_2(P_1, P_2(c_2)) \geq [P_2(\tilde{c}_2) - c_2]\, D_2(P_1, P_2(\tilde{c}_2)), \forall c_2, \tilde{c}_2 \in \mathcal{C}_2^m \tag{C.3}$$

$$[P_2(c_2) - c_2]\, D_2(P_1, P_2(c_2)) \geq [P_2(\tilde{c}_2) - c_2]\, D_2(P_1, P_2(\tilde{c}_2)), \forall c_2, \tilde{c}_2 \in \mathcal{C}_2^l \tag{C.4}$$

$$[P_2(c_2) - c_2]\, D_2(P_1, P_2(c_2)) \geq [P_2(\tilde{c}_2) - c_2]\, D_2(P_1, P_2(\tilde{c}_2)), \forall c_2 \in \mathcal{C}_2^h, \forall \tilde{c}_2 \in \mathcal{C}_2^m \tag{C.5}$$

$$[P_2(c_2) - c_2]\, D_2(P_1, P_2(c_2)) \geq [P_2(\tilde{c}_2) - c_2]\, D_2(P_1, P_2(\tilde{c}_2)), \forall c_2 \in \mathcal{C}_2^h, \forall \tilde{c}_2 \in \mathcal{C}_2^l \tag{C.6}$$

$$[P_2(c_2) - c_2]\, D_2(P_1, P_2(c_2)) \geq [P_2(\tilde{c}_2) - c_2]\, D_2(P_1, P_2(\tilde{c}_2)), \forall c_2 \in \mathcal{C}_2^m, \forall \tilde{c}_2 \in \mathcal{C}_2^h \tag{C.7}$$

$$[P_2(c_2) - c_2]\, D_2(P_1, P_2(c_2)) \geq [P_2(\tilde{c}_2) - c_2]\, D_2(P_1, P_2(\tilde{c}_2)), \forall c_2 \in \mathcal{C}_2^m, \forall \tilde{c}_2 \in \mathcal{C}_2^l \tag{C.8}$$

$$[P_2(c_2) - c_2]\, D_2(P_1, P_2(c_2)) \geq [P_2(\tilde{c}_2) - c_2]\, D_2(P_1, P_2(\tilde{c}_2)), \forall c_2 \in \mathcal{C}_2^l, \forall \tilde{c}_2 \in \mathcal{C}_2^h \tag{C.9}$$

$$[P_2(c_2) - c_2]\, D_2(P_1, P_2(c_2)) \geq [P_2(\tilde{c}_2) - c_2]\, D_2(P_1, P_2(\tilde{c}_2)), \forall c_2 \in \mathcal{C}_2^l, \forall \tilde{c}_2 \in \mathcal{C}_2^m \tag{C.10}$$

Proof of Lemma 3: Since we defined $\underline{\mathcal{C}}_2 = \mathcal{C}_2^l = \left\{c_2 \mid P_2(c_2) \leq \bar{r}_2^E\right\}$, by (3.6) and (3.5) we have the following demand for $c_2 \in \underline{\mathcal{C}}_2$,

$$D_2(P_1, P_2(c_2)) = (1 - m_1)\left[\int_{\bar{r}_2^N}^{R}\int_0^R h(r_1, r_2)dr_1 dr_2 + \int_{\bar{r}_2^E}^{\bar{r}_2^N}\int_{\chi(r_2)}^{R} h(r_1, r_2)dr_1 dr_2\right].$$

The insurance company takes the threshold valuations as given, so for cost realizations in $\underline{\mathcal{C}}_2$, the demand is independent of the variations in second period premiums.

Suppose $c_2 \in \underline{\mathcal{C}}_2$ is the true cost realization and the insurance company contemplates reporting a cost of $\tilde{c}_2 \in \underline{\mathcal{C}}_2$, then by (C.4) the incentive compatible premium schedule requires $P_2(c_2) \geq P_2(\tilde{c}_2)$. Now suppose $\tilde{c}_2$ is the true cost realization and the insurance company

[3] Cawley and Philipson (1999) found no strong evidence of adverse selection in the term life insurance.
[4] Taken from Gürkaynak, Sack and Wright (2007) and averaged for each month.

contemplates reporting a cost of c_2, then the incentive compatible premium schedule requires $P_2(c_2) \leq P_2(\tilde{c}_2)$. Therefore, we have $P_2(c_2) = P_2(\tilde{c}_2)$ for all cost realizations and deviations within the set $\underline{\mathcal{C}}_2$. ■

Proof of Proposition 1: For part (i), we first assume that $\underline{\mathcal{C}}_2$ is of measure zero. This yields $\Pr\left[P_2(c_2) \leq \bar{r}_2^E\right] = 0$, then by (3.2) we have that $\bar{r}_2^E$ is not the second period threshold valuation of the existing policyholders, because the outside option is strictly positive. By efficiency, this implies that there exists a new threshold valuation $\bar{r}_2^{E'} > \bar{r}_2^E$ such that (3.2) binds and $\Pr\left[P_2(c_2;\hat{c}_1) \leq \bar{r}_2^{E'}\right] > 0$. We can redefine $\underline{\mathcal{C}}_2$ in terms of the new threshold $\bar{r}_2^{E'}$ and it is not measure zero, which is a contradiction.

To see why part (ii) is true, we start by analyzing (C.2) and notice that in this cost region

$$D_2(P_1, P_2(c_2)) = (1 - m_1)\left[1 - H_2\left(P_2(c_2)\right)\right]. \tag{C.11}$$

We will assume that $P_2(c_2)$ is incentive compatible and differentiable on $\mathcal{C}_2^h$, then for a given $c_2 \in \mathcal{C}_2^h$ the following must be true

$$c_2 = \arg\max_{\hat{c}_2 \in \mathcal{C}_2^h} \left(P_2(\hat{c}_2) - c_2\right) D_2(P_1, P_2(\hat{c}_2)). \tag{C.12}$$

Therefore, (C.12) implies that the life insurance company will set a monopoly premium in the cost region $\mathcal{C}_2^h$, which is represented in the following fixed point problem:

$$P_2^*(c_2) = \frac{1 - H_2(P_2^*(c_2))}{h_2(P_2^*(c_2))} + c_2, \forall c_2 \in \mathcal{C}_2^h. \tag{C.13}$$

Notice that the incentive compatible second period premium schedule for cost realizations within the set $\mathcal{C}_2^h$ and the set itself do not depend the announced first period cost realization.

Similarly, we can examine (C.3) and notice that in this cost region the demand is

$$D_2(P_1, P_2(c_2)) = (1 - m_1)\left\{\left[1 - H_2\left(\bar{r}_2^N\right)\right] + \int_{P_2(c_2)}^{\bar{r}_2^N} \int_{\chi(r_2)}^{R} h(r_1, r_2) dr_1 dr_2\right\}. \tag{C.14}$$

Following a similar argument, the life insurance company will set a monopoly premium $P_2^*(c_2)$ in the cost region $\mathcal{C}_2^m$ according to the following equation

$$\left\{\left[1 - H_2\left(\bar{r}_2^N\right)\right] + \int_{P_2^*(c_2)}^{\bar{r}_2^N} \int_{\chi(r_2)}^{R} h(r_1, r_2) dr_1 dr_2\right\} = \left[P_2^*(c_2) - c_2\right] \left\{\int_{\chi(P_2^*(c_2))}^{R} h(r_1, P_2^*(c_2)) dr_1\right\}. \tag{C.15}$$

The incentive compatible second period premium for cost realizations within the set $\mathcal{C}_2^m$ and the set itself depends on the announced first period cost realization.

Notice that (C.5) and (C.7) hold trivially since the premiums are chosen to maximize each independent segment of the demand which also maximizes the whole demand for $c_2 \in \overline{\mathcal{C}}_2$.

Part (iii) follows from the usual kinked demand result. We can equate the solution from (C.13) to equal $\bar{r}_2^N$ to find c_2^h, and apply the same process to (C.15) to find c_2^m. We now show that $c_2^h > c_2^m$.

From (C.13) we have the following relationship between $\bar{r}_2^N$ and c_2^h

$$\bar{r}_2^N = \frac{1 - H_2(\bar{r}_2^N)}{h_2(\bar{r}_2^N)} + c_2^h. \tag{C.16}$$

Similarly, since $\chi(\bar{r}_2^N) = \delta$, then from (C.15) we have the following relationship between $\bar{r}_2^N$ and c_2^m

$$1 - H_2(\bar{r}_2^N) = (\bar{r}_2^N - c_2^m) \int_{\delta}^{R} h(r_1, \bar{r}_2^N) dr_1. \tag{C.17}$$

We substitute (C.16) into (C.17) and by the definition of marginal probability distribution for $h_2(\cdot)$ we can derive the following

$$\left(\bar{r}_2^N - c_2^h\right) \int_0^R h(r_1, \bar{r}_2^N) dr_1 = (\bar{r}_2^N - c_2^m) \int_{\delta}^{R} h(r_1, \bar{r}_2^N) dr_1.$$

Since $\delta \geq 0$, we have $c_2^h \geq c_2^m$. Next, we show that for any c_2 in $[c_2^m, c_2^h)$ the optimal premium is $\bar{r}_2^N$.

To show rigidity within $[c_2^h, c_2^m)$, suppose there exists a $c_2 \in (c_2^m, c_2^h)$ such that the optimal premium $P_2(c_2)$ is strictly greater than $\bar{r}_2^N$. Since the optimal premium has to be incentive compatible, we have the following

$$\begin{aligned}(\bar{r}_2^N - c_2^h)D_2(P_1, \bar{r}_2^N) &\geq \left[P_2(c_2) - c_2^h\right] D_2(P_1, P_2(c_2)) \\ &= \left[P_2(c_2) - c_2\right] D_2(P_1, P_2(c_2)) - (c_2^h - c_2)D_2(P_1, P_2(c_2)) \\ &\geq (\bar{r}_2^N - c_2)D_2(P_1, \bar{r}_2^N) - (c_2^h - c_2)D_2(P_1, P_2(c_2)) \\ &> (\bar{r}_2^N - c_2)D_2(P_1, \bar{r}_2^N) - (c_2^h - c_2)D_2(P_1, \bar{r}_2^N) \\ &= (\bar{r}_2^N - c_2^h)D_2(P_1, \bar{r}_2^N).\end{aligned}$$

The first and third inequality follows from the incentive compatibility constraints. The fourth inequality follows from a weakly decreasing demand and our assumption that $P_2(c_2) > \bar{r}_2^N$ and $c_2 < c_2^h$. We have a contradiction.

We now assume there exists a $c_2 \in (c_2^m, c_2^h)$ such that the optimal premium $P_2(c_2)$ is strictly smaller than $\bar{r}_2^N$. Similarly, with incentive compatibility, we can show

$$\begin{aligned}(\bar{r}_2^N - c_2^m)D_2(P_1, \bar{r}_2^N) &\geq \left[P_2(c_2) - c_2^m\right] D_2(P_1, P_2(c_2)) \\ &= \left[P_2(c_2) - c_2\right] D_2(P_1, P_2(c_2)) + (c_2 - c_2^m)D_2(P_1, P_2(c_2)) \\ &\geq (\bar{r}_2^N - c_2)D_2(P_1, \bar{r}_2^N) + (c_2 - c_2^m)D_2(P_1, P_2(c_2)) \\ &> (\bar{r}_2^N - c_2)D_2(P_1, \bar{r}_2^N) + (c_2 - c_2^m)D_2(P_1, \bar{r}_2^N) \\ &= (\bar{r}_2^N - c_2^m)D_2(P_1, \bar{r}_2^N).\end{aligned}$$

The first and third inequality follows from the incentive compatibility constraints. The fourth inequality follows from a weakly decreasing demand and our assumption that $P_2(c_2) < \bar{r}_2^N$ and $c_2 > c_2^m$. We have a contradiction.

To prove part (iv), suppose $\exists c_2$ and $\bar{\epsilon} > 0$ such that $\forall \epsilon \in (0, \bar{\epsilon})$, we have $P_2(c_2+\epsilon) < P_2(c_2)$. Notice that with part (i) and (ii), we can simply analyze the two cases: (i) $c_2 \in \underline{\mathcal{C}}_2$ and $c_2 + \epsilon \in \overline{\mathcal{C}}_2$, and (ii) $c_2 \in \overline{\mathcal{C}}_2$ and $c_2 + \epsilon \in \underline{\mathcal{C}}_2$. We first examine case (i). Since (3.12) and (3.13) must hold, we have the following

$$\begin{aligned}\left[P_2(c_2+\epsilon) - (c_2+\epsilon)\right] D_2\left(P_1, P_2^*(c_2)\right) &\geq \left[\bar{P}_2 - (c_2+\epsilon)\right] \bar{D}_2 \\ &> \left[\bar{P}_2 - c_2\right] \bar{D}_2 \\ &\geq \left[P_2(c_2+\epsilon) - c_2\right] D_2\left(P_1, P_2^*(c_2)\right),\end{aligned}$$

which is a contradiction since $\epsilon > 0$. Note the first and third inequality come from (3.12) and (3.13) respectively.

Next, we examine case (ii). Since (3.13) must hold, we have the following

$$\begin{aligned}[\bar{P}_2 - (c_2+\epsilon)]\bar{D}_2 &\geq [P_2^*(c_2) - (c_2+\epsilon)]\, D_2(P_1, P_2^*(c_2)) \\ &= [P_2^*(c_2) - c_2]\, D_2(P_1, P_2^*(c_2)) - \epsilon D_2(P_1, P_2^*(c_2)) \\ &\geq [\bar{P}_2 - c_2]\bar{D}_2 - \epsilon D_2(P_1, P_2^*(c_2)).\end{aligned}$$

The second inequality follows from (3.12). The analysis implies that $D_2(P_1, P_2^*(c_2)) \geq \bar{D}_2$. Since the demand function is weakly decreasing, it must be the case that $P_2^*(c_2) \leq \bar{r}_2^E$ and $c_2 \in \underline{\mathcal{C}}_2$. Therefore, (C.4) is violated if $P_2^*(c_2) > \bar{P}_2$ and $c_2 \in \underline{\mathcal{C}}_2$, which gives us the contradiction.

Part (v) follows immediately by examining (3.12) for the second period premium. For the second period premium, we know that $P_2^*(c_2) > \bar{P}_2$ for any $c_2 \in \overline{\mathcal{C}}_2$. Furthermore, by part (iii), we know that the optimal premium schedule is weakly increasing in c_2. Therefore, the result follows. ■

Proof of Lemma 4: We will first claim that the following must be true:

$$\left(\bar{P}_2 - c_2^T\right)\bar{D}_2 = \left[P_2^*(c_2^T) - c_2^T\right] D_2\left(P_1, P_2^*(c_2^T)\right).$$

The expression above has to hold, since if not, then there exists $\epsilon > 0$ such that we have a $c_2 \in (c_2^T - \epsilon, c_2^T + \epsilon)$ where one of the incentive compatibility constraints would not hold.

We first show that (3.13) holds trivially. For any $c_2 \in \underline{\mathcal{C}}_2$ and $\tilde{c}_2 \in \overline{\mathcal{C}}_2$, we have the following

$$\begin{aligned}[P_2^*(\tilde{c}_2) - c_2]\, D_2\left(P_1, P_2^*(\tilde{c}_2)\right) &= \left[P_2^*(\tilde{c}_2) - c_2^T\right] D_2\left(P_1, P_2^*(\tilde{c}_2)\right) + \left(c_2^T - c_2\right) D_2\left(P_1, P_2^*(\tilde{c}_2)\right) \\ &\leq \left[P_2^*(c_2^T) - c_2^T\right] D_2\left(P_1, P_2^*(c_2^T)\right) + \left(c_2^T - c_2\right) D_2\left(P_1, P_2^*(\tilde{c}_2)\right) \\ &= \left(\bar{P}_2 - c_2^T\right)\bar{D}_2 + \left(c_2^T - c_2\right) D_2\left(P_1, P_2^*(\tilde{c}_2)\right) \\ &\leq (\bar{P}_2 - c_2^T)\bar{D}_2 + \left(c_2^T - c_2\right)\bar{D}_2 \\ &= (\bar{P}_2 - c_2)\bar{D}_2.\end{aligned}$$

The first inequality follows from the insurance company's optimization problem. The second equality follows from (3.14). The second inequality follows from the monotonically decreasing demand function and parts (iv) and (v) of Proposition 1 and the fact that $c_2 \leq c_2^T$.

Next, we will show that (3.12) holds. $\forall c_2 \in \overline{\mathcal{C}}_2$,

$$\begin{aligned}
[P_2^*(c_2) - c_2]\, D_2\left(P_1, P_2^*(c_2)\right) &\geq \left[P_2^*(c_2^T) - c_2\right] D_2\left(P_1, P_2^*(c_2^T)\right) \\
&= \left[P_2^*(c_2^T) - c_2^T\right] D_2\left(P_1, P_2^*(c_2^T)\right) - (c_2 - c_2^T) D_2\left(P_1, P_2^*(c_2^T)\right) \\
&= \left(\bar{P}_2 - c_2^T\right) \bar{D}_2 - (c_2 - c_2^T) D_2\left(P_1, P_2^*(c_2^T)\right) \\
&= \left(\bar{P}_2 - c_2\right) \bar{D}_2 + (c_2 - c_2^T)\left[\bar{D}_2 - D_2\left(P_1, P_2^*(c_2^T)\right)\right] \\
&\geq \left(\bar{P}_2 - c_2\right) \bar{D}_2.
\end{aligned}$$

The first inequality follows from the insurance company's optimization problem. The second equality follows from (3.14). The second inequality follows from the the definition of $\overline{\mathcal{C}}_2$ and from the monotonically decreasing demand function and the fact that $c_2 \geq c_2^T$.
■

Proof of Proposition 2: We first show that there is a discrete jump at c_2^T where $\bar{P}_2 < P_2^*(c_2^T)$. Note that Proposition 1 has ruled out the case where $\bar{P}_2 > P_2^*(c_2^T)$. We proceed by ruling out the case for $\bar{P}_2 = P_2^*(c_2^T)$.

Suppose that $\bar{P}_2 = P_2^*(c_2^T)$, then by Lemma 4, we have that (3.14) holds, which implies the following

$$\bar{D}_2 = D_2\left(P_1, P_2^*(c_2^T)\right).$$

By (3.5) and (3.6), we have $\bar{r}_2^E = P_2^*(c_2^T)$. Therefore, $\mathrm{E}\left[\bar{r}_2^E - P_2(c_2) \mid P_2(c_2) \leq \bar{r}_2^E\right] = 0$, which violates the definition of $\bar{r}_2^E$. We have a contradiction and this proves that there is a discontinuity in the premium schedule at c_2^T.

To show that $\bar{r}_2^E > \bar{P}_2$, notice that by the definition of $\bar{r}_2^E$ shown in (3.2) can be expressed as

$$\left(\bar{r}_2^E - \bar{P}_2\right) G(c_2^T) = V_2^S(\bar{r}_2^E).$$

Since the expected value of switching is assumed to be strictly positive, $\bar{r}_2^E > \bar{P}_2$ follows.

Next, to show that $P_2^*(c_2^T) > \bar{r}_2^E$, first observe that by Lemma 4 and the fact that $\bar{P}_2 < P_2^*(c_2^T)$, it must be the case that $\bar{D}_2 > D_2\left(P_1, P_2^*(c_2^T)\right)$. By the definition of second period demand, we have

$$\frac{\bar{D}_2}{1-m_1} = \left[\int_{\bar{r}_2^N}^R \int_0^R h(r_1,r_2)dr_1dr_2 + \int_{\bar{r}_2^E}^{\bar{r}_2^N} \int_{\chi(r_2)}^R h(r_1,r_2)dr_1dr_2\right],$$

and

$$\frac{D_2(P_1,P_2^*(c_2^T))}{1-m_1} = \left[\int_{\max\left\{\bar{r}_2^N, P_2^*(c_2^T)\right\}}^R \int_0^R h(r_1,r_2)dr_1dr_2 + \int_{\max\left\{\bar{r}_2^E, min\{P_2^*(c_2^T), \bar{r}_2^N\}\right\}}^{\bar{r}_2^N} \int_{\chi(r_2)}^R h(r_1,r_2)dr_1dr_2\right].$$

Suppose $\bar{r}_2^E \geq P_2^*(c_2^T)$, then $\bar{D}_2 = D_2\left(P_1, P_2^*(c_2^T)\right)$, which is a contradiction. As a result, at the cutoff, $P_2^*(c_2^T) > \bar{r}_2^E > \bar{P}_2$ must hold. ■

Proof of Proposition 3: Using 3.14, we can express the incentive compatible $\bar{P}_2$ as a function of c_2^T. Hence, we can also express $\bar{D}_2$ as a function of the second period cutoff c_2^T. To complete the proof, we need to take the first order condition with respect to c_2^T. ■

Proof of Proposition 4: For a given optimal second period premium, the optimal first period premium is characterized by the first order condition (4). Also, since the second period premium is optimal, $\frac{\partial \Pi_2}{\partial P_1}$ can be found using the envelope theorem.

To show the last part of the proposition, notice that P_1^{pNR} is characterized by (3.16) without the second period premium, and P_1^{pR} is characterized by (3.16) with a first period demand function without transaction cost μ. When $\mu = 0$, no initial investment is required and there is cost to switching. In other words, there are no dynamic considerations, so consumers buy if and only if $r_t \geq P_t$ for each period. Therefore, in the absence of health shocks, a renewable contract is the same as a non-renewable contract, so they are priced the same: $P_1^{pNR} = P_1^{pR}$.

Finally, when $\mu = 0$, we have $\frac{\partial \Pi_2}{\partial P_1} = 0$ due to the fact that maximizing period by period is equivalent to maximizing present discounted value. However, if $\mu > 0$, by the definition of δ and equations (3.5) and (3.6), the envelope theorem yields us $\frac{\partial \Pi_2}{\partial P_1} < 0$.

More specifically,

$$\frac{\partial \Pi_2}{\partial P_1} = -(1-m_1)\int_{\max\left\{\bar{r}_2^E, \min\left\{P_2(c_2), \bar{r}_2^N\right\}\right\}} h(\chi(r_2), r_2)dr_2.$$

Let D_1^p denote the first period demand for life insurance in an environment without transaction cost, then

$$D_1^p(P_1) = \int_{P_1}^{R} \int_0^R h(r_1, r_2) dr_1 dr_2,$$

then we have

$$\frac{\partial D_1^p}{\partial P_1} = -\int_0^R h(P_1, r_2) dr_2.$$

Also, by (3.4), we have

$$\frac{\partial D_1}{\partial P_1} = -\int_{\bar{r}_2^N}^{R} h(\delta, r_2) dr_2 - \int_{\bar{r}_2^E}^{\bar{r}_2^N} h(\chi(r_2), r_2) dr_2.$$

Thus, for (3.16) to hold, it must be the case that $P_1 < P_1^{pNR} = P_1^{pR}$. ■