Essays on Insurance Economics and the Regulation of Financial Markets

DISSERTATION
of the University of St.Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Management

submitted by

Przemysław Zbigniew Rymaszewski

from

Poland

Approved on the application of

Prof. Dr. Hato Schmeiser

and

Prof. Dr. Martin Eling

The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, May 13, 2011

The President:

Prof. Dr. Thomas Bieger

Acknowledgements

The work on my dissertation has been challenging, yet rewarding, and would not have been possible without the enduring support of a few people.

First, I would like to express my deepest gratitude to my supervisor, Prof. Dr. Hato Schmeiser, for his mentorship and invaluable support as well as a true open-door policy. Furthermore, I strongly appreciate the feedback from my co-supervisor, Prof. Dr. Martin Eling, who helped me to considerably improve the content of this doctoral thesis.

I am also grateful to all three of my co-authors for an excellent cooperation and many stimulating discussions.

I will always remember the time spent at the Institute of Insurance Economics of the University of St. Gallen, an institution offering excellent research and working conditions with strong ties to both academia and practice. I would like to thank its entire team. In particular, I am indebted to my colleagues and friends Alexander Braun, Roger Faust, Caroline Siegel, and Alexandra Zemp.

Finally, I could hardly underestimate the support from my parents, Maria and Julian, my brother, Sławomir, and my dear friends.

Zurich, July 2011

Przemysław Rymaszewski

Outline

Contents

List of Figures

How Risky Are Interest Rate Guarantees Embedded in Participating Life Insurance Contracts? The Case of Germany

List of Tables

Summary

This dissertation consists of five parts, each of which comprises an individual research paper. The first two parts analyze insurance guaranty funds from diverse perspectives. In the paper *Under What Conditions Is an Insurance Guaranty Fund Beneficial for Policyholders?*, we investigate the conditions, under which the introduction of a self-supporting insurance guaranty fund is advantageous for policyholders. It is shown that in an imperfect market setting and given homogeneous firms all policyholders can potentially benefit from the existence of an insurance guaranty fund to the same extent if they have the same underlying risk preferences and are charged identical premiums. However, in a more realistic heterogeneous setting, an insurance guaranty fund is in general no longer beneficial for all insureds in the same manner. Thus, its introduction is likely to cause systematic utility transfers between the policyholders of different insurance companies. Subsequently, we propose a framework for utility-based fund charges as a possible solution to this problem.

The paper *Insurance Guaranty Funds and Their Relation to Solvency Regulation* brings insurance guaranty funds into a different focus. Since similar institutions might in general interfere with the fulfillment of the goals followed by prudential regulation and supervision, we underline the need for an enhanced regulatory approach extended by issues connected with existing institutionalized run-off solutions. We argue that, if appropriately designed, insurance guaranty funds might improve the stability of the financial system, help to avoid market failures, support consumer protection as well as increase the overall degree of competition.

The third part of the dissertation, the research paper *A Traffic Light Approach to Solvency Measurement of Swiss Occupational Pension Funds*, deals with solvency measurement in the occupational pension sector. Based on the combination of a stochastic pension fund model and a traffic light signal approach, we propose a solvency test for occupational pension funds in Switzerland. Being designed as a regulatory standard model, the set-up is intentionally kept parsimonious and, assuming normally distributed asset returns, a closed-form solution can be derived. Despite its simplicity the framework comprises the essential risk sources needed in supervisory practice. Due to its ease of calibration, it is additionally well suited for the fragmented Swiss market, keeping costs of solvency testing at a minimum. To illustrate its application, the model is calibrated and implemented for a small sample of ten Swiss pension

funds. Moreover, a sensitivity analysis is conducted to identify important drivers of the shortfall probabilities for the traffic light conditions.

The fourth part of the dissertation consists of the research paper *Stock vs. Mutual Insurers: Who Does and Who Should Charge More?*, which is an empirical and theoretical examination of the relationship between the premiums of insurers in the legal form of stock and mutual companies. An analysis of panel data for the German motor liability insurance sector does not indicate that mutuals charge significantly higher premiums than stock insurers. Subsequently, a comprehensive model framework for the arbitrage-free pricing of insurance contracts is employed to compare stock and mutual insurance companies with regard to the three central magnitudes premium size, safety level, and equity capital. Although, from a normative perspective, there are certain circumstances in which the premiums of stock and mutual insurers should be equal, these situations would generally require the mutual to hold comparatively less capital. As this is inconsistent with our empirical results, it seems that the observed insurance prices are not arbitrage-free.

The risks of interest rate guarantees embedded in participating life insurance contracts are analyzed in the fifth research paper, titled *How Risky Are Interest Rate Guarantees Embedded in Participating Life Insurance Contracts? The Case of Germany.* As life insurance companies generally invest a significant part of their asset portfolio in bonds, we base our analysis on a term structure model. By means of a Monte Carlo simulation calibrated in line with empirical data for the German bond market, we are able to show that the interest rate guarantees offered in the German insurance market can be fulfilled to a very high probability using simple investment strategies based on investments in government bonds. Thus, we have reasons to believe that under certain conditions the risk resulting from interest rate guarantees is rather low.

Zusammenfassung

Die folgende Dissertation besteht aus fünf Teilen, die jeweils eine in sich geschlossene Forschungsarbeit enthalten. Die ersten zwei Kapitel setzen sich mit der Institution eines Insolvenzfonds in der Versicherungsbranche auseinander und stellen diese aus unterschiedlichen Perspektiven dar. In dem Arbeitspapier *Under What Conditions Is an Insurance Guaranty Fund Beneficial for Policyholders?* wird der Einfluss eines selbstragenden Insolvenzfonds in der Versicherungsindustrie auf die Vermögenssituation der Versicherungsnehmer auf unvollständigen Märkten analysiert. Innerhalb des untersuchten Rahmens stellen wir fest, dass die etwaigen Vorteile, die ein solcher Fonds mit sich bringen kann, generell nicht verursachungsgerecht auf die Versicherungsunternehmen und somit auf ihre Kunden alloziert werden können. Dies kann zu Nutzentransfers zwischen den teilnehmenden Unternehmen und deren Versicherungsnehmern führen. Als Lösung wird ein nutzenbasierter Ansatz zur Prämienbemessung vorgeschlagen.

Einen anderen Blickwinkel auf die unterschiedlichen Aspekte, die die Einführung eines Insolvenzfonds in der Versicherungsbranche mit sich bringen kann, liefert die Arbeit *Insurance Guaranty Funds and Their Relation to Solvency Regulation*. Da ähnliche Institutionen in der Regel das Erreichen der übergeordneten Aufsichtsziele entweder unterstützen oder verhindern können, wird in dieser Forschungsarbeit die Notwendigkeit eines integrierten Ansatzes im Bereich der Solvenzregulierung hervorgehoben. Dieser sollte gegenüber den heutigen Aufsichtssystemen um die Fragen der institutionalisierten Lösungen im Bereich der Konkurssicherung erweitert werden. Es wird argumentiert, dass entsprechend ausgestaltete Insolvenzfonds die Stabilität des Finanzsystems, die Vermeidung von Marktversagen, den Verbraucherschutz sowie den Wettbewerb in der Versicherungsindustrie fördern können.

Mit der Forschungsarbeit *A Traffic Light Approach to Solvency Measurement of Swiss Occupational Pension Funds* im dritten Teil der Dissertation liegt der Fokus auf der Solvenzmessung in der beruflichen Vorsorge. Es wird ein Solvenztest für schweizerische Vorsorgeeinrichtungen vorgeschlagen, welcher auf der Kombination eines stochastischen Pensionskassenmodells und eines Ampelsignalansatzes basiert. Im Sinne eines regulatorischen Standardmodells wird der Aufbau bewusst einfach gehalten. Zudem kann unter der Annahme normalverteiler Anlagerenditen eine geschlossene Lösung abgeleitet werden. Trotz der relativ geringen Komplexität deckt das System die wesentlichen in der aufsichtlichen Praxis erforderlichen Risikoquellen ab. Aufgrund der einfach durchzuführenden Ka-

librierung ist es zusätzlich gut zur Anwendung im fragmentierten Markt für schweizerische Vorsorgeeinrichtungen geeignet und hält die Kosten der Solvenzregulierung so gering wie möglich. Zur Veranschaulichung der Anwendung des Modells wird es mittels einer kleinen Stichprobe von zehn Vorsorgeeinrichtungen kalibriert und umgesetzt. Darüber hinaus wird eine Sensitivitätsanalyse durchgeführt, um wichtige Einflussfaktoren der Unterschreitungswahrscheinlichkeiten für die Ampelbedingungen zu identifizieren.

Der vierte Teil der Dissertation beinhaltet die Forschungsarbeit *Stock vs. Mutual Insurers: Who Does and Who Should Charge More?*, welche eine empirische und theoretische Analyse des Zusammenhangs zwischen den Prämien von Versicherern in der Rechtsform der Aktiengesellschaft und des Versicherungsvereins auf Gegenseitigkeit darstellt. Eine Auswertung von Paneldaten aus dem Bereich der deutschen Kfz-Haftpflichtversicherung liefert keinerlei Anhaltspunkte dafür, dass Versicherungsvereine signifikant höhere Prämien berechnen als Aktiengesellschaften. Im Anschluss wird ein umfassender Modellrahmen für die arbitragefreie Bewertung von Versicherungsverträgen eingesetzt, um Versicherer in Form von Aktiengesellschaften und Vereinen hinsichtlich der drei zentralen Grössen Prämienhöhe, Sicherheitsniveau und Eigenkapital zu vergleichen. Obwohl es aus normativer Sicht bestimmte Umstände gibt, in denen die Prämien einer Aktiengesellschaft und eines Versicherungsvereins auf Gegenseitigkeit gleich sein sollten, wäre es hierfür erforderlich, dass der Verein vergleichsweise weniger Kapital vorhält. Da dies mit den empirischen Ergebnissen nicht konsistent ist, scheinen die beobachteten Versicherungspreise nicht arbitragefrei zu sein.

Das fünfte und letzte Kapitel, *How Risky Are Interest Rate Guarantees Embedded in Participating Life Insurance Contracts? The Case of Germany*, setzt sich mit dem Risiko der Zinsgarantien auseinander, die den Versicherungsnehmern im Rahmen einer Kapitallebensversicherung üblicherweise gewährt werden. Da Lebensversicherungsunternehmen in der Regel einen erheblichen Teil ihres Asset-Portfolios in Anleihen investieren, beruht die Analyse auf der Modellierung eines Zinsstrukturmodells. Betrachtet wird ein Versicherungsunternehmen, das einfache Anlagestrategien in Staatsanleihen verfolgt. Mit Hilfe einer Monte-Carlo-Simulation, die mit empirischen Marktdaten für den deutschen Markt kalibriert wird, kann gezeigt werden, dass die Zinsgarantien auf dem deutschen Versicherungsmarkt mit einer sehr hohen Wahrscheinlichkeit erwirtschaftet werden können. Somit kann das Risiko, das aus den Zinsgarantien resultiert, in bestimmten Konstellationen als eher gering eingestuft werden.

Part I

Under What Conditions Is an Insurance Guaranty Fund Beneficial for Policyholders?

Abstract

In this paper, we derive conditions in an imperfect market setting, under which the introduction of a self-supporting insurance guaranty fund improves the position of the policyholders. In those cases where a guaranty fund is advantageous given homogeneous firms in the market, all policyholders benefit from it to the same extent, if they have the same underlying risk preferences and are charged identical premiums. In a more realistic heterogeneous setting, the introduction of an insurance guaranty fund is in general no longer beneficial for all policyholders in the same manner. Hence, systematic wealth transfers take place between the policyholders of different insurance companies. As a possible solution, and in order to counteract this effect, we introduce a framework for utility-based fund charges and discuss its implications for the insurance market.[1]

---

1 Introduction

The magnitude of losses throughout the current financial crisis has even jeopardized the existence of large financial institutions. Insolvency costs caused by the recent turbulence in the international financial markets not only affected equity and debt holders, but, through the necessity for major bail-outs, also affected taxpayers and the entire society. Regarding the insurance sector, these recent events revealed the need for a general reconsideration of regulation design in general and solvency measurement in particular. See, for example, the current development of the European Solvency II framework, e.g. CEIOPS (2009), for an overview. Since the aim of solvency regulation and supervision is to reduce the probability of insurer default to a predefined small, yet still positive level, further questions arise with regard to the case of an insurance company default and the coverage of associated insolvency costs. Making taxpayers pay for corporate insolvencies is hard to justify and may incentivize insurers to take more risks.

An insurance guaranty fund financed by all insurance companies in the market can be employed to force insurance companies to internalize the insolvency costs of the entire industry. Its introduction is only one of many possible approaches for the attempt to install a controlled run-off system within the insurance sector. Nevertheless, since insurance companies are not homogeneous and differ in risk brought into the insurance guaranty fund's pool, the calculation of risk-based premiums and the definition of possible pay-outs from the insurance guaranty fund become a very important task in this context. If these aspects are not considered—as is typically done in insurance practice[2]—adverse in-

---

[2]Most of the existing insurance guaranty fund schemes charge premiums that are not directly linked to insurer risks. An exception from this rule is the German life insurance guaranty fund scheme, where charges depend on company ranking according to their *financial capacity*, defined as an equity relative to solvency margin. However, under the currently valid regulatory framework, this approach cannot be treated as plenary risk-oriented. An overview on the funding of guaranty fund schemes within the European Union can be found in Oxera (2007, p. 34). See Feldhaus and Kazenski (1998, p. 44) for a detailed description of the still applicable U.S. based property-liability guaranty fund system. For U.S. regulation concerning life insurance sector, see Brewer-III et al. (1997, p. 305).

centives for insurers and extensive cross-subsidization between market participants can be expected.

In this paper, we examine the conditions under which the introduction of an insurance guaranty fund can be beneficial for policyholders. As a first step, we show that, if a contingent claim approach is applied in order to value the claims of the stakeholders of an insurance company, policyholders cannot benefit by the introduction of a fairly designed insurance guaranty fund. As a second step and in an imperfect market setting, we formally show under which conditions an insurance guaranty fund is advantageous for risk-averse policyholders. Possible diversification benefits through the introduction of an insurance guaranty fund are measured by an increase in the utility of the policyholders. The correlation between the payoff of the fund and the assets of the insurer, as well as the premium level in the fund, turn out to be important in order to draw benefits through the introduction of an insurance guaranty fund. If companies are homogeneous and diversification benefits arise through the insurance guaranty fund, the increase in utility is equally allocated to all participating policyholder collectives. However, we find that in the case of heterogeneous companies, an insurance guaranty fund is in general no longer beneficial—at least not to the same extent—for all policyholders of the different insurance companies on the market. As a possible solution to this problem, we introduce a concept of utility-based premium calculations within the fund and explain its effects formally and by means of a numerical example.

The remainder of this paper is organized as follows. In Section 2, a set of relevant issues related to the introduction of an insurance guaranty fund is presented. In addition, we give an overview of related literature. Section 3 concentrates on the introduction of an insurance guaranty fund in the case of an efficient and perfect market. If a perfect market is not given, which is considered in Section 4, we analyze the conditions under which an insurance guaranty fund can be beneficial. For the utility-based approach, we assume that the utility of a policyholder collective is described by the standard mean-variance utility function. In Section 5, we discuss different premium principles in the utility-based setting of Section 4. An exemplary payoff structure for the guaranty fund is given

and properties are derived in the case of homogeneous companies. In Section 6, numerical examples based on a Monte Carlo simulation are provided in order to illustrate the main findings. Finally, in Section 7 we set forth our conclusions and express an outlook.

2   Preliminary considerations and literature overview

Solvency rules should reduce the default probability of an insurer to a predefined level. However, due to general randomness as well as model risk, even a very solvent insurance company still remains exposed to bankruptcy. As a consequence of the systemic character of financial institutions, the aggregated costs of their insolvency can spill over to policyholders in general and to companies from the non-financial area in particular. Hence, costs related to the insolvency of financial institutions are one of many widely acknowledged reasons for their regulation. See, for example, the work by Mayes (2004, p. 516), with special focus on the banking sector. In contrast to other industries, the quality of most financial products depends instantaneously on the solvency level of the supplier.

In a competitive market with perfect information, policyholders would be able to entirely assess the risk profile of an insurer. In such a setting, one could believe that the choice of an insurer conditioned on the safety level is mainly of concern to the policyholder himself. However, possible claims of an ex-ante unknown third-party that cannot be covered because of the insurer insolvency, could contradict this reasoning, even if symmetric information between the different stakeholders is assumed. In some cases, this may even result in improper incentives for policyholders, while choosing an insurance company.

Since the potential direct and indirect costs for the economy induced by financial institutions insolvencies may be severe, governments have an incentive to regulate financial institutions. However, to avoid inappropriate incentives for equity and debt holders, the government should bindingly rule out, that taxpayers ultimately cover insolvencies. Pos-

sible inappropriate incentives may be compounded when shareholders interpret their position in the insurance company solely as a call option. Hence, in markets with asymmetric information, management of an insurance company can increase the market value of the company by raising the volatility of its assets. Whether they would be able to do it in a market with perfect information or not, depends on the overall insurance contract conditions. Since higher risk decreases the value of policyholder claims, in such a setting those insured would like to account for higher risk and demand lower insurance premiums from riskier insurance companies. If the contract conditions allow those insured to renegotiate the same, in the event that the risk of the insurance company increases, the insurer cannot increase its value by increasing its risk. If no contract renegotiation is possible, higher volatility will increase the market value of the insurer at the expense of the policyholders. However, if the nature of the insurance contract foresees a contract renewal, we can expect the policyholders to account for insurer behavior in the following period.[3] In general, similar incentives for enhanced risk-taking are one of many reasons for prudential regulation and supervision within insurance markets, essentially aiming at maintaining a minimum level of solvency.

One of the possibilities to reduce the risk appetite of an insurer in markets with imperfect information is to introduce an insurance guaranty fund, which could force the insurer to internalize the costs of its potential insolvency. Furthermore, an insurance guaranty fund could send signals to the insurer that bail-outs by the government are not intended in a situation of financial distress. This is due to the fact that a guaranty fund should lower insolvency costs for policyholders, thus making the run-off of an insurer more justifiable, from both the political and the social perspective. However, making insurance companies pay their entire insolvency expenses requires that an insurance guaranty fund be solely financed by the insurers and not, for example, (partly) by the taxpayer.

---

[3]Nevertheless, in such case the result is likely to depend on further aspects, for example, fixed versus indefinite number of periods, deeply analyzed within game theory (see, for example, Axelrod and Hamilton, 1981).

If the insurer is a mutual, policyholders also include the owners of the insurance company. This may lead to an alleviation of the classical risk incentive problem between equity and debt holders. Nevertheless, problems of third-party liabilities, imperfect information between policyholders and management, as well as the general economic costs of the insolvency of an insurer still remain.

Cummins (1988) argues that a well designed insurance guaranty fund should demand risk-based premium payments to avoid adverse incentives. Risk-adequate premiums may create a similar situation to the above described conditions of a market with perfect information and renegotiable insurance contracts, since insurers are not able to increase their market value only by enhancing risk. If the insurer is not charged according to its risk, the position is akin to the setting with asymmetric information. Insurance companies may still be able to increase their market value by raising the volatility of their assets. This problem, also denoted as the risk-subsidy effect, is analyzed by Lee et al. (1997), who provide sound empirical evidence for its significance within the U.S. property-liability insurance market. They do not find any significant influence of the so-called monitoring effect, which may occur if insurers are charged ex-post with risk-inadequate fees. In such a context, insurance companies should have a greater incentive to monitor their competitors. This relation is closely connected to the monitoring abilities of insurance companies and their policyholders. If the insurers are able to monitor other insurance companies more effectively than policyholders are capable of doing and are willing to, this effect can be expected to be crucial. Nevertheless, a system of ex-post charges cannot be organized in a risk-based way, due to the fact that the insolvent insurance company, which may have been the riskiest one, is typically not charged at all. This issue is extensively addressed by Han et al. (1997, p. 1119). Brewer-III et al. (1997) support the previous results for guaranty funds effectively funded by taxpayers, for the U.S. life insurance market. Downs and Sommer (1999) arrive at similar conclusions and extend their line of reasoning by insider ownership issues. Sommer (1996) provides empirical evidence for market discipline in the U.S. property-liability insurance market and puts it in the insurance guaranty fund context. He argues that fund charges based

on the amount of insurance premiums earned by an insurer may even strengthen the risk-subsidy incentives.

The calculation of risk-adequate premiums is one of the most important tasks in the context of an insurance guaranty fund. Cummins (1988) suggests a premium principle, based on option pricing theory. In his framework, the assets and liabilities of the company are modeled as diffusion processes. He interprets the value of the hedge provided by an insurance guaranty fund as the price of a European put option on the assets of the company with the value of liabilities as the strike.[4] In this manner, a closed-form solution can be derived for the fund charge because the value of an option with a volatile exercise price is equivalent to an option to exchange one asset for another (see Fischer, 1978 and Margrabe, 1978). Duan and Yu (2005) extend the one-period model from Cummins (1988) into a multi-period setting. They incorporate the interest rate risk and regulatory responses mandated by risk-based capital regulations.

However, in order to correctly apply the no-arbitrage option pricing framework in this specific context, one has to assume perfect markets, i.e., markets that, among others, provide adequate instruments for the replication of both the assets and liabilities of the company. Such an assumption implies that every market participant can diversify to the same extent. As we show in the following Section 3, in such a setting no advantages can be derived in principal through the existence of an insurance guaranty fund or an insurance company. More precisely, if additional transaction costs accrue, the implementation of an insurance guaranty fund would not be supported by the owners of the insurance company and its policyholders.

An insurance guaranty fund, the advantage of which can be identified as the ability of pooling diverse risks, is to some extent similar to an obligatory reinsurance. Borch (1962) analyzes an equilibrium price for transferring risks on the reinsurance market. Within the proposed setting, he shows that if every insurance company maximizes its utility, the market is unlikely to reach a Pareto optimal state. Borch (1962)

---

[4]See, for example, Merton (1977) for similar frameworks within the deposit insurance.

also claims that the existence of a price mechanism that will automatically lead to a Pareto optimal solution is improbable. It can only be reached if negotiations between the participating parties take place. However, additional assumptions about the negotiation patterns are needed. Mossin (1966) argues that this result is characteristic only for the—here analyzed—reinsurance market, where the price of a security depends mainly on the stochastic nature of the yield, not on the number of outstanding securities.[5]

3   Insurance guaranty funds in a contingent claims approach

Within the arbitrage-free setting of the contingent claims approach, the pooling of insurance claims in an insurance guaranty fund does not change the wealth situation of either policyholders or shareholders (see Doherty and Garven, 1986; Cummins, 1988). More precisely, if both stakeholder groups apply the same form of present value calculus and the stakes are priced in a fair way (the present value of future cash flow equals the initial contribution), there will be no advantage from an insurance guaranty fund. In what follows, we present this line of reasoning in more detail.

Consider a set $\mathcal{C} = \{1, \ldots, M\}$ of M mutual companies[6] denoted by $i = 1, \ldots, M$, that are active on the market, and define $W_i^{(0)}$ as the aggregated premium paid by the policyholders at time $t = 0$.

If there is no insurance guaranty fund, policyholders of the mutual i are entitled to two stochastic stakes at time $t = 1$, namely the position of those insured and the stake of the owner. The position of those insured, whose present value is denoted by $P_i^{(0)}$, grants the policyholders coverage of their aggregated stochastic claims $\tilde{S}_i^{(1)}$ each time the company remains solvent, i.e. the stochastic assets $\tilde{A}_i^{(1)}$ exceed the claims $\tilde{S}_i^{(1)}$. If

---

[5]Different aspects of reinsurance, other than the optimal risk sharing, for example, risk management know-how or monitoring (see, for example, Mayers and Smith, 1990; Plantin, 2006), seem not to be of high relevance in the insurance guaranty fund context.

[6]Similar arguments can be applied in case of a publicly traded company.

claims exceed the assets of the insurer at time $t = 1$, i.e. $\tilde{S}_i^{(1)} > \tilde{A}_i^{(1)}$, policyholders receive solely the market value of assets—in this case, the company is insolvent. The stake of the owner, whose present value is denoted by $E_i^{(0)}$, is determined residually by the difference between the value of the assets and the aggregated claims at $t = 1$. Formally, we can write the aggregated position of the policyholders in the mutual company i as

$$
\begin{aligned}
W_i^{(0)} &= P_i^{(0)} + E_i^{(0)} \\
&= \mathrm{PV}[\min(\tilde{A}_i^{(1)}, \tilde{S}_i^{(1)})] + \mathrm{PV}[\max(\tilde{A}_i^{(1)} - \tilde{S}_i^{(1)}, 0)] \\
&= \mathrm{PV}[\tilde{A}_i^{(1)}],
\end{aligned}
\tag{1}
$$

where PV denotes the present value. Notice that the present value of the stake of those insured $P_i^{(0)}$ can be rewritten as follows:

$$
\begin{aligned}
P_i^{(0)} &= \mathrm{PV}[\min(\tilde{A}_i^{(1)}, \tilde{S}_i^{(1)})] \\
&= \mathrm{PV}[\tilde{S}_i^{(1)}] - \mathrm{PV}[\max(\tilde{S}_i^{(1)} - \tilde{A}_i^{(1)}, 0)].
\end{aligned}
\tag{2}
$$

In this context, $\max(\tilde{S}_i^{(1)} - \tilde{A}_i^{(1)}, 0)$ stands for the insolvency put option, a measure of the safety level of the insurer (see, for example, Butsic, 1994).

If an insurance guaranty fund is introduced, the company i pays a fraction $\pi_i^{(0)}$ of the aggregated premium $W_i^{(0)}$ into the guaranty scheme, as an ex-ante charge. The insurance guaranty fund invests this premium on the capital market. This investment results in a stochastic cash flow $\tilde{\pi}_i^{(1)}$ at time $t = 1$. Hence, $\sum_{i \in \mathcal{C}} \tilde{\pi}_i^{(1)}$ constitutes the funds available within the insurance guaranty scheme at time $t = 1$. Since within our framework there is no external agent, for example the government, covering the claims of the policyholders in the event that those funds are insufficient, the insurance guaranty fund faces default risk. The systematic market risk, as well as the underwriting risk, are the only risk sources incorporated in this setting. Those are the relevant risks for the insurers, as well as the insurance guaranty fund. To simplify matters, we do not account for further idiosyncratic (unsystematic) risks, e.g., operational risks.

The present value of the company's assets $A_i^{(0),\star}$ in the case of an insurance guaranty fund being established is then given by

$$A_i^{(0),\star} = \text{PV}[\tilde{A}_i^{(1),\star}] = \text{PV}[\tilde{A}_i^{(1)}] - \text{PV}[\tilde{\pi}_i^{(1)}] = A_i^{(0)} - \pi_i^{(0)}, \qquad (3)$$

where $A_i^{(0)}$ denotes the assets of the insurer i at time $t = 0$ before the premium $\pi_i^{(0)}$ has been paid to the fund.

Notice that there is a general circularity problem connected with the calculation of $\pi_i^{(0)}$ in case it is derived in line with an insurer's overall risk. This is due to the fact that $\pi_i^{(0)}$ depends on the distribution $(\tilde{S}_i^{(1)} - \tilde{A}_i^{(1)})$ of the respective insurer. At the same time, $\pi_i^{(0)}$ influences the distribution of $(\tilde{S}_i^{(1)} - \tilde{A}_i^{(1),\star})$, which becomes the relevant distribution for premium calculation after the premium is charged. Hence, the risk-based premium can be derived solely as an approximation of the true risk-adjusted premium.

In the situation with an insurance guaranty fund and other things being equal, a positive fund charge lowers the value of the aforementioned policyholder and shareholder positions in the insurance company. This can be written as $P_i^{(0),\text{f}} < P_i^{(0)}$ and $E_i^{(0),\text{f}} < E_i^{(0)}$, respectively, where the superscript f denotes the values in the setup with the fund. In addition and because of $A_i^{(0),\star} = A_i^{(0)} - \pi_i^{(0)}$, the default probability of the insurance company ceteris paribus increases from $\text{Prob}(\tilde{A}_i^{(1)} < \tilde{S}_i^{(1)})$ to $\text{Prob}(\tilde{A}_i^{(1),\star} < \tilde{S}_i^{(1)})$. This is a direct consequence of charging insurers in advance. It is even more intense in case of a risk-based fund charge calculation, since companies with higher probability of default may be charged higher rates.

At the same time, policyholders of the mutual do have claims with respect to the insurance guaranty fund. If the mutual is insolvent, full coverage of the claims of the policyholders can be provided as long as the insurance guaranty fund is solvent. If too many insolvencies occur and not enough capital is available in the insurance guaranty fund, only a partial coverage is possible. Let us denote $\mathcal{P}_i^{(0)} = \text{PV}[\tilde{\mathcal{P}}_i^{(1)}]$ as the present value of the claims of the policyholders of company i against the guaranty fund. In addition, policyholders collectively enjoy an equity stake in the insurance guaranty fund. The present value of this position, from

the point of view of the policyholders of company i, can be written as $\mathcal{E}_i^{(0)} = \text{PV}[\tilde{\mathcal{E}}_i^{(1)}]$. The present value is positive whenever the probability that the fund will not be entirely exhausted at $t = 1$ is positive.

Hence, the present value of the claims of the policyholders of the mutual company i against the insurance guaranty fund can be denoted as

$$F_i^{(0)} = \text{PV}[\tilde{F}_i^{(1)}] = \text{PV}[\tilde{\mathcal{P}}_i^{(1)} + \tilde{\mathcal{E}}_i^{(1)}]. \tag{4}$$

Figure 1 summarizes, for times $t = 0$ and $t = 1$, from the point of view of the company i, the relevant cash flows and values of the different stakes in both setups without and with an insurance guaranty fund.

If the premium obtained from policyholders of the mutual company i is the same with and without an insurance guaranty fund, the following relation can be derived

$$
\begin{aligned}
W_i^{(0)} \;=\; W_i^{(0),\text{f}} \;&=\; P_i^{(0),\text{f}} + E_i^{(0),\text{f}} + \mathcal{P}_i^{(0)} + \mathcal{E}_i^{(0)} \\
&=\; \text{PV}[\min(\tilde{A}_i^{(1),\star}, \tilde{S}_i^{(1)})] \\
&\quad + \text{PV}[\max(\tilde{A}_i^{(1),\star} - \tilde{S}_i^{(1)}, 0)] + \text{PV}[\tilde{F}_i^{(1)}].
\end{aligned} \tag{5}
$$

The stake of the policyholder of company i can be broken down into the following elements:

$$
\begin{aligned}
P_i^{(0),\text{f}} \;&=\; \text{PV}[\min(\tilde{A}_i^{(1),\star}, \tilde{S}_i^{(1)})] \\
&=\; \text{PV}[\tilde{S}_i^{(1)}] - \text{PV}[\max(\tilde{S}_i^{(1)} - \tilde{A}_i^{(1),\star}, 0)].
\end{aligned} \tag{6}
$$

If $\text{PV}[\tilde{F}_i^{(1)}] = \pi_i^{(0)}$, the present value of the claims of the policyholders with respect to the fund equals the initial contribution. This condition is fulfilled in an arbitrage-free market. Otherwise systematic wealth transfers between different insurers would take place.

The safety level of an insurer changes ceteris paribus with the contribution $\pi_i^{(0)}$ in the fund. In addition, an agreed payoff scheme that defines the conditions under which payouts are made from the fund to the insurance companies will influence the distributions of $\tilde{\mathcal{P}}_i^{(1)}$ and $\tilde{\mathcal{E}}_i^{(1)}$. For instance, the probability that claims can be paid from the two sources (the insurer and the guaranty fund) will depend on the design of the
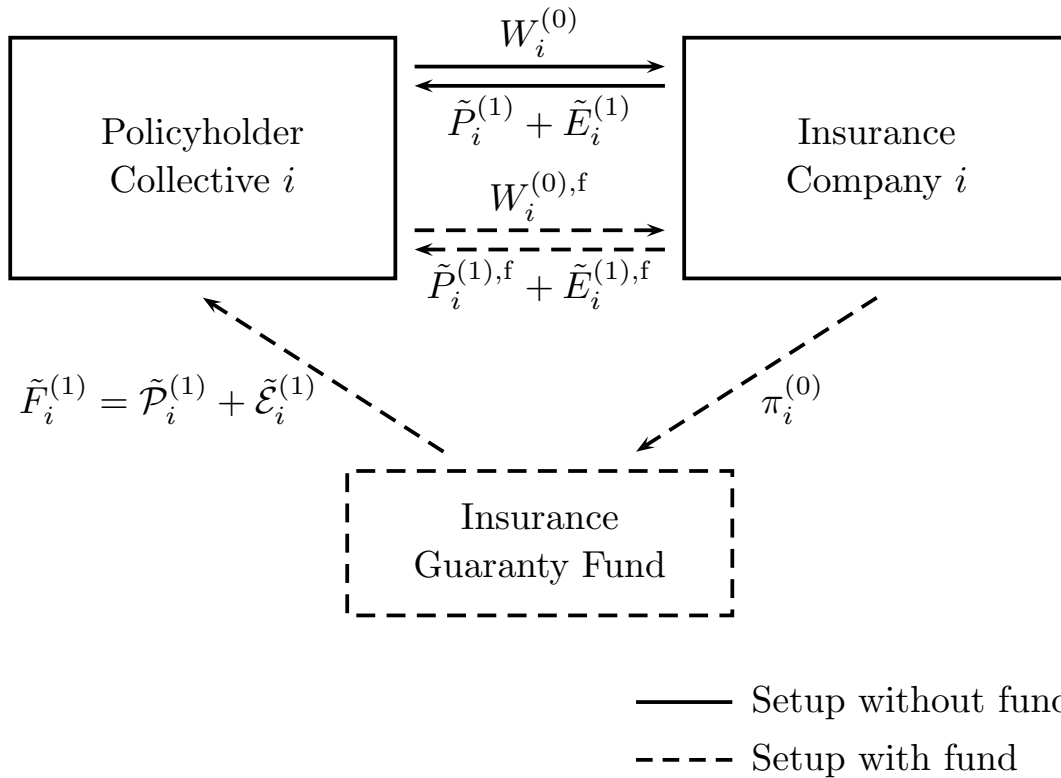
Figure 1: Illustration of the cash flows for company i and the stakeholder positions without and with an insurance guaranty fund at times $t = 0$ and $t = 1$. Notations: $W_i^{(0)}, W_i^{(0),\mathrm{f}}$ = aggregated premium paid by the policyholders without/with fund at $t = 0$; $\tilde{P}_i^{(1)}, \tilde{P}_i^{(1),\mathrm{f}}$ = position of the without/with fund at $t = 1$; $\tilde{E}_i^{(1)}, \tilde{E}_i^{(1),\mathrm{f}}$ = stake of the owners without/with fund at $t = 1$; $\pi_i^{(0)}$ = premium charged by the fund at $t = 0$; $\tilde{\mathcal{P}}_i^{(1)}$ = claims against the fund at $t = 1$; $\tilde{\mathcal{E}}_i^{(1)}$ = equity stake in the fund at $t = 1$.

payoff scheme and the premium payments in the fund. However, as long as the stakes are priced fairly, in this model setup the policyholders face neither an advantage nor disadvantage.

This situation may change, if the ability to diversify varies among distinct market participants. It can be the case in imperfect markets, where investors cannot replicate all possible future cash flows. In general, one may assume that insurance companies and insurance guaranty funds are able to diversify in a better way than policyholders. If this is the case, the potential pooling effect of an insurance guaranty fund may enhance

policyholders' wealth position and, hence, become one of the advantages in favor of its introduction. In such a setup, in order to value the diverse stakes, assumptions about the preferences of the investors are needed.

4 A utility-based approach

From now on, we refer to the insurance guaranty fund and the notations introduced in Section 3. Relevant notations of cash flows and stakes are illustrated in Figure 1. After introducing the wealth positions of the policyholder collectives, in the absence and after the introduction of the insurance guaranty fund, we set a minimal condition for possible payoff schemes. Finally, in a policyholder utility-based approach we discuss the resulting positions in different situations of risk-neutral and risk-averse policyholders.

4.1 Wealth position of the policyholder collectives

In a situation without an insurance guaranty fund, the wealth position $\tilde{W}_i^{(1)}$ of the policyholder group of the mutual company $i \in \mathcal{C}$ at time $t = 1$ is given by the sum of the position of the insured $\tilde{P}_i^{(1)}$ and the stake of the owner $\tilde{E}_i^{(1)}$ (compare with (1) in the contingent claims approach),

$$
\begin{aligned}
\tilde{W}_i^{(1)} &= \tilde{P}_i^{(1)} + \tilde{E}_i^{(1)} \\
&= \min(\tilde{A}_i^{(1)}, \tilde{S}_i^{(1)}) + \max(\tilde{A}_i^{(1)} - \tilde{S}_i^{(1)}, 0) \\
&= \tilde{A}_i^{(1)}, \qquad \forall i = 1, \dots, M.
\end{aligned}
\tag{7}
$$

Thus, the joint wealth position is equivalent to a long position in company assets. Thereby, we abstract from any other wealth positions and risk sources the policyholders of company i might face.

After the insurance guaranty fund is introduced and the fund premium $\pi_i^{(0)}$ is paid, at time 0, the company assets decrease. In this case, the available assets at time $t = 1$ are denoted by

$$
\tilde{A}_i^{(1),\star} = \tilde{A}_i^{(1)} - \tilde{\pi}_i^{(1)}, \qquad i = 1, \dots, M.
\tag{8}
$$

In order to simplify the analysis, we implicitly assume that the chosen asset allocation in the insurance guaranty fund for the respective premiums $\pi_i^{(0)}$, $i = 1, \ldots, M$, and in the insurance company i, is identical. With the introduction of the guaranty fund, claims against the fund, $\tilde{\mathcal{P}}_i^{(1)}$, as well as an equity stake in the fund, $\tilde{\mathcal{E}}_i^{(1)}$, arise and add up to the wealth position. Hence, the latter is equal to

$$
\begin{aligned}
\tilde{W}_i^{(1),\mathrm{f}} &= \tilde{P}_i^{(1),\mathrm{f}} + \tilde{E}_i^{(1),\mathrm{f}} + \tilde{\mathcal{P}}_i^{(1)} + \tilde{\mathcal{E}}_i^{(1)} \\
&= \min(\tilde{A}_i^{(1),\star}, \tilde{S}_i^{(1)}) + \max(\tilde{A}_i^{(1),\star} - \tilde{S}_i^{(1)}, 0) + \tilde{F}_i^{(1)} \\
&= \tilde{A}_i^{(1),\star} + \tilde{F}_i^{(1)}, \qquad \forall i = 1, \ldots, M,
\end{aligned}
\tag{9}
$$

where $\tilde{F}_i^{(1)} = \tilde{\mathcal{P}}_i^{(1)} + \tilde{\mathcal{E}}_i^{(1)}$ denotes the payoff of the insurance guaranty fund to the policyholder collective of company i, see Equation (4). A minimal requirement on the structure of the payoff is given below in Paragraph 4.2, an exemplary derivation is illustrated in Paragraph 5.3.

4.2 Payoff scheme of the guaranty fund

The cash flow $\tilde{F}_i^{(1)}$ of the insurance guaranty fund is dependent on the number of companies M, their asset and claim distributions, $\tilde{\mathbf{A}}^{(1)} = (\tilde{A}_i^{(1)})_{i \in \mathcal{C}}$ and $\tilde{\mathbf{S}}^{(1)} = (\tilde{S}_i^{(1)})_{i \in \mathcal{C}}$, the correlation structures between assets and claims, as well as the premiums charged by the fund, $\mathbf{\Pi}^{(0)} = (\pi_i^{(0)})_{i \in \mathcal{C}}$, and their stochastic distribution at time $t = 1$, $\tilde{\mathbf{\Pi}}^{(1)}$.

In order to ensure proper incentives, the minimal requirement on the obligatory insurance guaranty fund is that it has to be self-supporting, i.e., at time $t = 1$, the sum for all companies of the fund payoff equals the sum of the premiums collected and reinvested by the fund (according to the original asset allocation of the insurers):

$$
\sum_{i \in \mathcal{C}} \tilde{F}_i^{(1)} = \sum_{i \in \mathcal{C}} \tilde{\pi}_i^{(1)}.
\tag{10}
$$

This implies that, for example, there is no external agent (e.g., taxpayer) that would have to cover a part of the default risk through (contingent) payments to the fund.

Since the guaranty fund is required to be self-supporting, the derivation of an adequate structure for the payoff scheme is strongly determined. However, in general, various schemes can be derived, thereby implying different incentives for the market participants. In Paragraph 5.3, we derive and discuss one possible solution for the fund's payoff structure and prove that the proposed scheme guarantees that the fund is self-supporting.

4.3   Utility function of the policyholder collective

Let us assume that the utility of the policyholder collective for companies $i = 1, \ldots, M$, at time $t = 1$ is described by the standard mean-variance utility function of their respective stochastic wealth position. To simplify matters, we concentrate on the policyholder collective as a whole and do not model single policyholders.

In the setup without insurance guaranty fund, the wealth position of the policyholder collective of company i, $\tilde{W}_i^{(1)}$, is given by (7) and we introduce the corresponding utility

$$
\begin{aligned}
\phi_i^{(1)} &= \langle \tilde{W}_i^{(1)} \rangle - \frac{a_i}{2} \mathrm{var}(\tilde{W}_i^{(1)}) \\
&= \langle \tilde{A}_i^{(1)} \rangle - \frac{a_i}{2} \mathrm{var}(\tilde{A}_i^{(1)}),
\end{aligned}
\tag{11}
$$

where a_i defines the risk aversion parameter of the collective. Similarly, in the setting with the fund and based on the stochastic wealth position of the policyholder collective $\tilde{W}_i^{(1),\mathrm{f}}$, introduced in (9), we introduce the utility given by

$$
\begin{aligned}
\phi_i^{(1),\mathrm{f}} &= \langle \tilde{W}_i^{(1),\mathrm{f}} \rangle - \frac{a_i}{2} \mathrm{var}(\tilde{W}_i^{(1),\mathrm{f}}) \\
&= \langle \tilde{A}_i^{(1),\star} + \tilde{F}_i^{(1)} \rangle - \frac{a_i}{2} \mathrm{var}(\tilde{A}_i^{(1),\star} + \tilde{F}_i^{(1)}) \\
&= \langle \tilde{A}_i^{(1)} - \tilde{\pi}_i^{(1)} + \tilde{F}_i^{(1)} \rangle \\
&\quad - \frac{a_i}{2} \mathrm{var}(\tilde{A}_i^{(1)} - \tilde{\pi}_i^{(1)} + \tilde{F}_i^{(1)}).
\end{aligned}
\tag{12}
$$

The absolute change in policyholders' utility in company i, due to the introduction of the guaranty fund, is denoted by

$$
\begin{aligned}
\Delta_{\mathrm{a}}\phi_i^{(1)} &= \phi_i^{(1),\mathrm{f}} - \phi_i^{(1)} \\
&= \langle \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)} \rangle \\
&\quad - \frac{a_i}{2}\left[\mathrm{var}(\tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)}) + 2\mathrm{cov}(\tilde{A}_i^{(1)}, \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)})\right]. \quad (13)
\end{aligned}
$$

Risk-neutral investors

If the policyholder collective of company j is assimilated to a risk-neutral investor, who by definition does not adjust for risk while making its financial decisions, it would be indifferent between both setups, without or with the guaranty fund, if $\phi_j^{(1)} = \phi_j^{(1),\mathrm{f}}$. This condition implies that

$$
\Delta_{\mathrm{a}}\phi_j^{(1)} = \phi_j^{(1),\mathrm{f}} - \phi_j^{(1)} \overset{!}{=} 0. \qquad (14)
$$

Since in this case we have $a_j = 0$, we get from (13) with (14):

$$
\Delta_{\mathrm{a}}\phi_j^{(1)} = \langle \tilde{F}_j^{(1)} - \tilde{\pi}_j^{(1)} \rangle \overset{!}{=} 0 \quad \Leftrightarrow \quad \langle \tilde{F}_j^{(1)} \rangle = \langle \tilde{\pi}_j^{(1)} \rangle. \qquad (15)
$$

We conclude that the condition $\phi_j^{(1)} = \phi_j^{(1),\mathrm{f}}$ is fulfilled if and only if $\langle \tilde{F}_j^{(1)} \rangle = \langle \tilde{\pi}_j^{(1)} \rangle$. Since the guaranty fund is supposed to be self-supporting (10), which implies that $\sum_{i \in \mathcal{C}} \langle \tilde{F}_i^{(1)} \rangle = \sum_{i \in \mathcal{C}} \langle \tilde{\pi}_i^{(1)} \rangle$, Condition (15) is always accounted for on an aggregated level. Hence, if all investors are risk-neutral, there is no possibility of achieving a utility-based Pareto enhancement by introducing an insurance guaranty fund. In fact, in a self-supporting fund, violating Condition (15) for some company can improve the expected value of the mutual stake in one insurer only by (negatively) influencing at least some of the expected values of the payoff of the policyholders of the remaining companies. This means that, in such a case, some policyholders can benefit solely from the costs of other insured parties. This result is equivalent to a zero-sum game concept known from game theory (see, for example, Neumann and Morgenstern, 1953).

Risk-averse policyholders

In what follows, we assume risk-averse policyholders, i.e. $a_i > 0$, $\forall i \in \mathcal{C}$, and analyze potential benefits from the existence of an insurance guaranty fund by an analysis of the (absolute) change in utility $\Delta_{\mathrm{a}}\phi_i^{(1)}$.

Pooling of claims in an insurance guaranty fund is beneficial from the perspective of the policyholders of the mutual insurer i if

$$
\begin{aligned}
\Delta_{\mathrm{a}}\phi_i^{(1)} > 0 \quad\Leftrightarrow\quad & \langle \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)} \rangle - \frac{a_i}{2}\Big[\mathrm{var}(\tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)}) \\
& + 2\mathrm{cov}(\tilde{A}_i^{(1)}, \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)})\Big] > 0 \\
\Leftrightarrow\quad & \langle \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)} \rangle - \frac{a_i}{2}\Big[\mathrm{var}(\tilde{F}_i^{(1)}) + \mathrm{var}(\tilde{\pi}_i^{(1)}) \\
& - 2\mathrm{cov}(\tilde{F}_i^{(1)}, \tilde{\pi}_i^{(1)}) + 2\mathrm{cov}(\tilde{A}_i^{(1)}, \tilde{F}_i^{(1)}) \\
& - 2\mathrm{cov}(\tilde{A}_i^{(1)}, \tilde{\pi}_i^{(1)})\Big] > 0 \\
\Leftrightarrow\quad & \langle \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)} \rangle + a_i\Big[\mathrm{cov}(\tilde{A}_i^{(1)}, \tilde{\pi}_i^{(1)}) \\
& + \mathrm{cov}(\tilde{F}_i^{(1)}, \tilde{\pi}_i^{(1)}) - \mathrm{cov}(\tilde{F}_i^{(1)}, \tilde{A}_i^{(1)}) \\
& - \frac{1}{2}\mathrm{var}(\tilde{F}_i^{(1)}) - \frac{1}{2}\mathrm{var}(\tilde{\pi}_i^{(1)})\Big] > 0. \qquad (16)
\end{aligned}
$$

A close investigation of the inequality in (16) is helpful to analyze the main sources of diversification within the fund:

- The value of $\langle \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)} \rangle = \langle \tilde{F}_i^{(1)} \rangle - \langle \tilde{\pi}_i^{(1)} \rangle$ is dependent from the premiums paid by the companies $\mathbf{\Pi}^{(0)}$ and the payoff scheme $\tilde{F}_i^{(1)}$.[7]

- Since we assume that the insurance guaranty fund follows the same asset allocation for the premiums as the insurance companies, i.e. $\tilde{A}_i^{(1)}$ and $\tilde{\pi}_i^{(1)}$ have the same standardized stochastic distribution,

---

[7]For example, in the case of homogeneous companies discussed in Paragraph 5.4, see Property 2, equal premium charges for all companies imply that $\langle \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)} \rangle = 0$, $\forall i \in \mathcal{C}$ (see Statement (27)). In a general setting, a premium principle can define $\mathbf{\Pi}^{(0)}$ such that $\langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}_i^{(1)} \rangle$ (see the discussion in Section 5).

the correlation $\rho(\tilde{A}_i^{(1)}, \tilde{\pi}_i^{(1)}) = 1$, and $\text{cov}(\tilde{A}_i^{(1)}, \tilde{\pi}_i^{(1)})$ is always positive.

- Furthermore, since the correlation $\rho(\tilde{A}_i^{(1)}, \tilde{\pi}_i^{(1)}) = 1$, $\text{cov}(\tilde{F}_i^{(1)}, \tilde{\pi}_i^{(1)})$ is directly linked to $\text{cov}(\tilde{F}_i^{(1)}, \tilde{A}_i^{(1)})$. We have identity of the correlations $\rho(\tilde{F}_i^{(1)}, \tilde{\pi}_i^{(1)}) = \rho(\tilde{F}_i^{(1)}, \tilde{A}_i^{(1)})$, and hence the difference in $\text{cov}(\tilde{F}_i^{(1)}, \tilde{\pi}_i^{(1)})$ and $\text{cov}(\tilde{F}_i^{(1)}, \tilde{A}_i^{(1)})$ depends on the magnitudes of the assets $A_i^{(0)}$ and charged premium $\pi_i^{(0)} < A_i^{(0)}$ at time $t = 0$.

- The quantity $-\text{var}(\tilde{F}_i^{(1)}) - \text{var}(\tilde{\pi}_i^{(1)})$ is always negative.

A further interesting insight is that Condition (16) cannot be fulfilled whenever the insurance company invests only in risk-free assets. Since the insurance guaranty fund invests every $\pi_i^{(0)}$ following exactly the same investment strategy as the insurer i, in such case $\tilde{\pi}_i^{(1)}$ becomes deterministic, denoted by $\pi_i^{(1)}$. In this case all covariance terms, as well as $\text{var}(\tilde{\pi}_i^{(1)})$ in (16), are equal to zero. Due to the fact that the investment strategy does not affect the stochasticity of insurer claims, $\text{var}(\tilde{F}_i^{(1)})$ is in such case still positive. Hence, if $\langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}_i^{(1)} \rangle$, we get $\Delta_a \phi_i^{(1)} = -\frac{a_i}{2}\text{var}(\tilde{F}_i^{(1)}) \leq 0$. This implies that in order to achieve diversification, a positive asset return volatility is needed. Intuitively, if no insurance guaranty fund is established, an entirely risk-free investment strategy makes the wealth position of the policyholder collective deterministic. This is implied by the fact that the wealth position of a policyholder collective equals a long position in insurer assets (see Equation (7)). If an insurance guaranty fund is introduced, the wealth of a policyholder collective changes to a long position in reduced, but still deterministic, insurer assets plus the stochastic payoff from the insurance guaranty fund (see Equation (12)). Hence, if $\langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}_i^{(1)} \rangle = \pi_i^{(1)}$, the establishment of an insurance guaranty fund rises the volatility of the position of the policyholders without influencing its mean. Within the given setting, this is strictly a disadvantage for risk-averse policyholders with $a_i > 0$.

Since Equation (16) contains the (complex) relationship between $\tilde{F}_i^{(1)}$ and the asset $\tilde{\mathbf{A}}^{(1)}$, claim $\tilde{\mathbf{S}}^{(1)}$, and premium distributions $\tilde{\mathbf{\Pi}}^{(1)}$ of all companies, an explicit derivation of necessary conditions for a positive

diversification benefit is not practicable without loss of generality. However, from now on we will discuss the potential benefits from insurance guaranty funds given particular settings (e.g., homogeneous companies, identical premium charges). Moreover, we illustrate different setups with numerical simulations in Section 6.

Finally, let us point out that the changes in utility implied by the existing guaranty funds are not, in general, identical for all market players. Since most of the existing national guaranty funds charge premiums based on the premium income of the companies (e.g., USA, UK, France), or their net technical reserves (e.g., Germany), this neither guarantees that the fund is self-supporting, as we require in Equation (10), nor that different policyholder groups profit from an equal utility increase caused by the existence of the guaranty fund, as we discuss in Section 5.

5 Premium principles and payoff

The payoff distribution $\tilde{F}_i^{(1)}$ is strongly influenced by the premium principle used to derive $\pi_i^{(0)}$, $i \in \mathcal{C}$. Since we assume that the insurer and the insurance guaranty fund choose the same fixed risky asset allocation, an increase in fund premium payments $\pi_i^{(0)}$ will result in an increase in $\text{var}(\tilde{\pi}_i^{(1)})$. All elements in Equation (16) would be influenced by an alteration of $\pi_i^{(0)}$. Hence, for the policyholders of the insurer i with a specific risk aversion parameter $a_i > 0$, there may be a premium range where Inequality (16) is fulfilled. This premium range depends on the asset and claim distributions as well as their correlations between all insurers. In other words, we may calculate a premium $\pi_i^{(0)}$ that would set the utility of the policyholder to a given level. Since premium levels are directly linked to the safety level (ruin probability, expected shortfall) of the companies, charged premiums are bounded from above by in-force solvency regulations (e.g. Solvency II, Swiss Solvency Test). Furthermore, the effect of the premium level on the diversification benefit and the augmentation of the safety level if a guaranty fund is introduced, must be analyzed carefully.

5.1   Premiums in the general case

Whenever homogeneous companies are charged different premiums, or in the case of heterogeneous companies, the premium principle $\langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}_i^{(1)} \rangle$, sometimes referred to as the net risk premium, no longer holds in general for any $i \in \mathcal{C}$. Let us assume that we are able to find premiums $\pi_i^{(0)}$ such that the mentioned principle holds:

$$\text{Premiums } \mathbf{\Pi}^{(0)} = (\pi_i^{(0)})_{i \in \mathcal{C}} \text{ such that } \langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}_i^{(1)} \rangle, \forall i \in \mathcal{C}. \quad (17)$$

Even if all pooling participants are characterized by the same level of risk aversion $a_i > 0$, the effect of pooling is different for every participating insurer. This is due to diverging distributions of assets and claims as well as to the different correlation structure between those variables. In general, some policyholder collectives can be worse off after the introduction of an insurance guaranty fund. Even when all policyholders, of all companies, would benefit from an insurance guaranty fund (in this case, Inequality (16) would be fulfilled for all participants), the utility increase would in general differ among companies. This finding is intensified if, for example, different policyholder collectives valuate their own wealth positions according to different utility functions.

Using some kind of risk-based premium principle, based on the loss distribution $(\tilde{S}_i^{(1)} - \tilde{A}_i^{(1)})$ of the insurer i, would not change our general line of reasoning. In this case, the possibility that some companies will (and other will not) benefit from an insurance guaranty fund again cannot be excluded. We believe that, in an imperfect market, a non-arbitrary way of allocating the existing diversification benefits back to the pool participants, via a particular premium principle, is not possible, since we face a problem that is similar to the capital allocation dilemma, extensively discussed by the academic literature in the last few years (see, for example, Merton and Perold, 1993; Phillips et al., 1998; Myers and Read, 2001; Sherris, 2006; Gründl and Schmeiser, 2007; Ibragimov et al., 2010; Zanjani, 2010). It implies that there is no non-arbitrary way to allocate the benefits from diversifying the unsystematic risks within a pooled portfolio. Hence, the postulate of an insurance guaranty fund charging premiums that should be fair according to the risk born by

an insurer leads in general to some form of a utility transfer between the insurance companies pooled in the insurance guaranty fund. Such a transfer may under certain conditions be justified, if, for example, the proposed guaranty fund leads to a considerable reduction of agency problems, as described in Section 2.

We may further investigate the question of whether this result would change, if we applied a different utility function from the mean-variance utility function introduced in Section 4.3, which accounts solely for the first two central moments of the wealth distribution. In general, higher order central moments of the wealth distribution are not relevant for the maximization of the expected utility of the investor only if the latter either possesses a quadratic utility function or a normally distributed wealth position (see Chamberlain, 1983). Due to the general characteristics of $\tilde{F}_i^{(1)}$ we cannot expect policyholders' wealth position to be normally distributed. Hence, the assumption of mean-variance utility can be seen only as a simplified approach. Applying a different kind of a utility function would influence Condition (16) and, therefore, definitely have an effect on the question of whether or not, and to what extent, the insurance guaranty fund is beneficial for a specific policyholder collective. Nevertheless, the main point we want to show is the unavoidable wealth transfers that take place between the policyholder collectives. In the further analysis, we show that the main driver of possible cross-subsidization effects is the troublesome allocation of diversification benefits, which is not tied to a specific utility function.

5.2   Utility-based premiums in the general case

As stated above, the benefits of pooling claims within an insurance guaranty fund may differ widely among participants. Besides the used premium principle, the potential advantages or disadvantages from pooling are closely tied to the portfolio composition of the insurer, as well as the fund. One of the possible ways to derive the insurance guaranty fund premium is the premium calculation based on the individual utility of the participating policyholder collectives. More precisely, we could demand a premium calculation for all M companies leading to an equal

utility increase for all M participants.[8] Such a calculation is possible, if for each company $i \in \mathcal{C}$ there exists a premium $\pi_i^{(0)}$—accounting for the available amount of assets $A_i^{(0)}$ and solvency regulations in force—such that the preset (non-zero) utility increase can be reached.

The existing insurance guaranty funds, charging premiums calculated on their business volume or reserves, as mentioned above, may imply systematic disadvantages for some companies in the market. This makes an obligatory participation in a guaranty fund, beyond the consideration of risk-adequacy be, at least to some extent, advantageous for all policyholder collectives.

When the payoff structure for the guaranty fund is defined, as is, for example, given in Equation (22), the change in utility in $t = 1$, $\Delta_{\mathrm{a}}\phi_i^{(1)}$ (see (13)), from the setup without any to that with an insurance guaranty fund, can be calculated for all companies $i \in \mathcal{C}$ with given preferences a_i. The utility states $\phi_i^{(1)}$ and $\phi_i^{(1),\mathrm{f}}$, as well as the change in utility $\Delta_{\mathrm{a}}\phi_i^{(1)}$, are considered here as a function of the premiums $\pi_i^{(0)}$ charged for all companies: $\phi_i^{(1)} = \phi_i^{(1)}(\mathbf{\Pi}^{(0)})$, $\phi_i^{(1),\mathrm{f}} = \phi_i^{(1),\mathrm{f}}(\mathbf{\Pi}^{(0)})$, $\Delta_{\mathrm{a}}\phi_i^{(1)} = \Delta_{\mathrm{a}}\phi_i^{(1)}(\mathbf{\Pi}^{(0)})$.

We can evaluate the set $\mathcal{S}_K$ of possible premium combinations $\mathbf{\Pi}^{(0)}$ for given utility change parameter $K \in \mathbb{R}$, by the following procedure:

$$\text{Premiums } \mathbf{\Pi}^{(0)} = (\pi_i^{(0)})_{i \in \mathcal{C}} \text{ such that } \Delta_{\mathrm{a}}\phi_i^{(1)}(\mathbf{\Pi}^{(0)}) = K, \forall i \in \mathcal{C}. \quad (18)$$

Depending on the magnitude of the parameter K, this procedure yields a set of premium combinations $\mathbf{\Pi}^{(0)}$ such that the overall (absolute) change in utility $\Delta_{\mathrm{a}}\phi_i^{(1)}$ is equal for all participants (premium principle). An optimization calculus can define the premium combination such that the change in utility is maximized, i.e. the maximum of K such that there is a solution $\mathbf{\Pi}^{(0)} \in \mathcal{S}_K$.

If a solution exists, there may be more than one solution for the optimal premium combination, depending on the set of companies. Furthermore, premiums calculated in this way may in some particular cases lead to very high premiums and significantly increase the default prob-

---

[8]Our approach is to some extent similar to some liberal views on taxation already postulated in the nineteenth century (see Mill, 1848, p. 348).

abilities of some of the insurers. In this way, the insurance guaranty fund may fall foul of the effective supervision and prudential regulation (e.g., Solvency II, Swiss Solvency Test). To avoid such situations, the optimization should take into account the limitations of the individual company with respect to the particular asset and claim situations.

Moreover, a problem of different pool compositions may arise. It is unambiguously solved only if an obligation to enter the guaranty fund is established. Otherwise, some insurers may benefit by setting up a pool that would exclude some market participants. Another important issue may lie in the strategic behavior of an insurer, which, assuming that the firm is acquainted with all aspects of the premium calculation, could cause undesired actions on its part.

Furthermore, in practice we are very likely to face substantial problems with regard to the specification and calibration of the proposed model for premium calculation. Since the detailed determination of the adequate form and specific parameters of the utility functions of policyholder collectives may be very costly, hard to communicate, and in most cases not straightforward—due to its time and context dependence, see, for example, Kahneman and Tversky (1979); Farquhar (1984); Fennema and Van Assen (1998)—it is unlikely to be introduced in the supervisory practice.

Our solution is in general not Pareto efficient. However, as shown by Borch (1962), in this specific context a Pareto efficient solution can only be found if we assume a specific negotiation pattern of the participating companies (see Section 2).

The principle given by (18), yields premium combinations that equalize $\Delta_\mathrm{a}\phi_i^{(1)}$ for all participants. Other possible procedures with regard to the introduced utility include the equalization of the relative change in utility

$$\Delta_\mathrm{r}\phi_i^{(1)} = \frac{\Delta_\mathrm{a}\phi_i^{(1)}}{|\phi_i^{(1)}|}, \tag{19}$$

or, going further, the marginal change in utility $\Delta_\mathrm{a}\phi_i^{(1)}/\pi_i^{(0)}$, or the ratio $[\Delta_\mathrm{a}\phi_i^{(1)}/|\phi_i^{(1)}|]/\pi_i^{(0)}$. Different utility measures can replace $\Delta_\mathrm{a}\phi_i^{(1)}$ in (18). Moreover, we could consider utility functions based on the safety

level of the insurer and use them to define the premiums in the proposed framework. Related premium principles include, for example, the equalization of the default probabilities, or the expected shortfall values. For each principle, the obtained premium combinations must be analyzed carefully. A sensitivity analysis on the principles used, the differences in the charged premiums and the possible effects on the behavior of the companies is still due. In Section 6 we illustrate our proposal with some numerical examples.

5.3   Derivation of an exemplary payoff function

The exemplary payoff scheme for an insurance guaranty fund derived in this section is not only intended to be self-supporting, i.e. to fulfill Condition (10), but also to establish desirable incentives for the participating insurers. In general, a chance for premium refund to solvent companies, in the event that the insurance guaranty fund does not go bankrupt, can encourage the companies to not only limit their own risk, but also to monitor their rivals. Moreover, the issue of insurance companies that have gone bankrupt solely due to the existence of an insurance guaranty fund should be adequately addressed, within the payoff scheme, in order to reduce potential resistance against its introduction.

Let $\tilde{\delta}_i^{(1)}$ be the policyholder deficit in the setup with an insurance guaranty fund,

$$\tilde{\delta}_i^{(1)} = \tilde{S}_i^{(1)} - \tilde{A}_i^{(1),\star}, \qquad \forall i \in \mathcal{C}, \tag{20}$$

and $\tilde{\gamma}^{(1)}$ be the fund deficit,

$$\tilde{\gamma}^{(1)} = \sum_{i=1}^{M} \left(\tilde{S}_i^{(1)} - \tilde{A}_i^{(1),\star} \right)^+ - \sum_{i=1}^{M} \tilde{\pi}_i^{(1)}, \tag{21}$$

which inform about the funding situation and measure the safety level with regard to insolvency of the different companies and the insurance guaranty fund, at time $t = 1$.

Depending on their respective safety levels, companies are classified in different subsets $\mathcal{C}^{\Pi}$, $\mathcal{C}^0$, $\mathcal{C}^{\star} \subset \mathcal{C}$ with the values of their respective

policyholder deficits as criteria. The subsets of companies that are introduced below are illustrated in Figure 2:

- Let $\mathcal{C}^{\Pi} = \{i \in \mathcal{C} | \tilde{\delta}_i^{(1)} > \tilde{\pi}_i^{(1)}\}$, the subset of companies that become insolvent at time $t = 1$, even without paying the premium charged by the guaranty fund, i.e. $\tilde{S}_i^{(1)} > \tilde{A}_i^{(1)}$. Let $M^{\Pi} = |\mathcal{C}^{\Pi}|$, where $|\cdot|$ denotes the cardinality (here: number of companies in $\mathcal{C}^{\Pi}$).

- Let $\mathcal{C}^0 = \{i \in \mathcal{C} | 0 < \tilde{\delta}_i^{(1)} \leq \tilde{\pi}_i^{(1)}\}$, the subset of companies that, at $t = 1$, would remain solvent without the insurance guaranty fund, but become insolvent due to paying the fund premium. Let $M^0 = |\mathcal{C}^0|$.

- Let $\mathcal{C}^{\star} = \mathcal{C} \setminus (\mathcal{C}^{\Pi} \cup \mathcal{C}^0) = \{i \in \mathcal{C} | \tilde{\delta}_i^{(1)} \leq 0\}$, the subset of companies that remain solvent at $t = 1$ after paying the premium to the guaranty fund. Let $M^{\star} = |\mathcal{C}^{\star}|$.

We consider the payoff structure of the fund for all companies $i \in \mathcal{C}$ at time $t = 1$ defined by a function $\tilde{F}_i^{(1)} = \tilde{F}_i^{(1)}(\tilde{\mathbf{A}}^{(1)}, \tilde{\mathbf{S}}^{(1)}, \tilde{\mathbf{\Pi}}^{(1)})$, given by the following expression:

$$
\tilde{F}_i^{(1)} = \begin{cases}
\tilde{\delta}_i^{(1)} & \text{for } i \in \mathcal{C}^{\Pi} \text{ and if } \tilde{\gamma}^{(1)} \leq 0 \\[2mm]
\tilde{\kappa}_i^{(1)} & \text{for } i \in \mathcal{C}^0 \text{ and if } \tilde{\gamma}^{(1)} \leq 0 \\[2mm]
\dfrac{\tilde{\pi}_i^{(1)}}{\sum_{j \in \mathcal{C}^{\star}} \tilde{\pi}_j^{(1)}} \left[\sum_{j \in \mathcal{C}} \tilde{\pi}_j^{(1)} \right. & \\
\quad - \sum_{j \in \mathcal{C}^{\Pi}} \tilde{\delta}_j^{(1)} & \\
\quad \left. - \sum_{j \in \mathcal{C}^0} \tilde{\kappa}_j^{(1)} \right] & \text{for } i \in \mathcal{C}^{\star} \text{ and if } \tilde{\gamma}^{(1)} \leq 0 \\[2mm]
\dfrac{\tilde{\delta}_i^{(1)}}{\sum_{j \in (\mathcal{C}^{\Pi} \cup \mathcal{C}^0)} \tilde{\delta}_j^{(1)}} \sum_{j \in \mathcal{C}} \tilde{\pi}_j^{(1)} & \text{for } i \in \mathcal{C}^{\Pi} \cup \mathcal{C}^0 \\
& \text{and if } \tilde{\gamma}^{(1)} > 0 \\[2mm]
0 & \text{for } i \in \mathcal{C}^{\star} \text{ and if } \tilde{\gamma}^{(1)} > 0
\end{cases}
\tag{22}
$$

where, for $i \in \mathcal{C}^0$, whenever $\tilde{\gamma}^{(1)} \leq 0$, we define

$$
\tilde{\kappa}_i^{(1)} = \max \left(\tilde{\delta}_i^{(1)}, \frac{\tilde{\pi}_i^{(1)}}{\sum_{j \in \mathcal{C} \setminus \mathcal{C}^{\Pi}} \tilde{\pi}_j^{(1)}} \left[\sum_{j \in \mathcal{C}} \tilde{\pi}_j^{(1)} - \sum_{j \in \mathcal{C}^{\Pi}} \tilde{\delta}_j^{(1)} \right] \right).
\tag{23}
$$
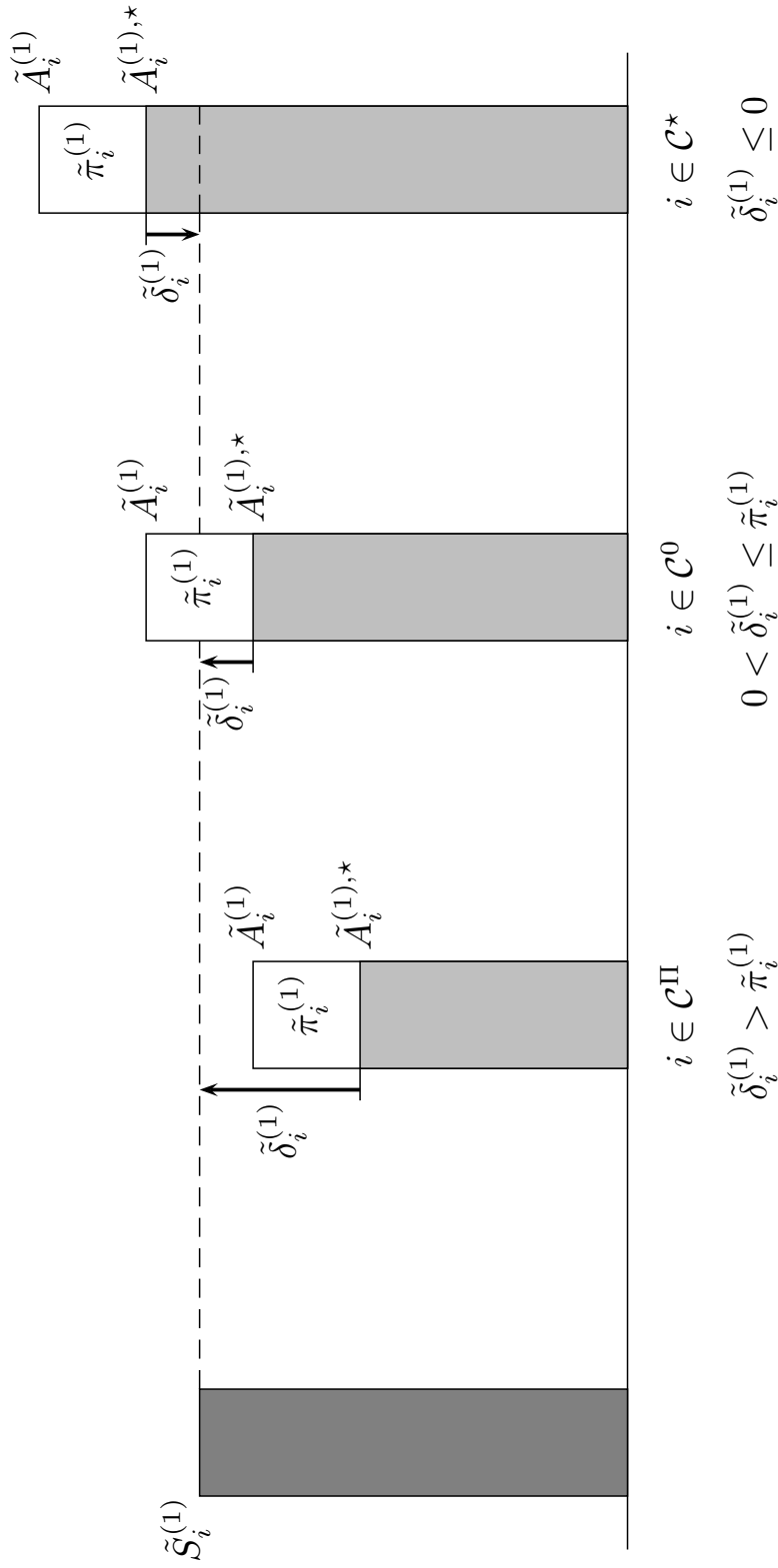
Figure 2: Illustration at time $t = 1$, for company i, of the three considered situations for $\tilde{\delta}_i^{(1)} = \tilde{S}_i^{(1)} - \tilde{A}_i^{(1),\star}$, where $\tilde{A}_i^{(1),\star} = \tilde{A}_i^{(1)} - \tilde{\pi}_i^{(1)}$, with their exemplary combinations of assets $\tilde{A}_i^{(1)}$ and charged and reinvested premium $\tilde{\pi}_i^{(1)}$ with respect to the claims $\tilde{S}_i^{(1)}$.

The first three lines in the definition of the payoff scheme of the fund, given by Equation (22), define the payoff in cases where the guaranty fund remains solvent, i.e. $\tilde{\gamma}^{(1)} \leq 0$, whereas the last two lines cover cases where the fund is strictly insolvent ($\tilde{\gamma}^{(1)} > 0$).

In the first case, where the fund remains solvent ($\tilde{\gamma}^{(1)} \leq 0$), we define first the payoff if company $i \in \mathcal{C}^{\Pi}$, i.e. faces a bankruptcy higher than the premium ($\tilde{\delta}_i^{(1)} > \pi_i^{(1)}$). In this situation, the payoff equals the realized insolvency put option with value $\tilde{\delta}_i^{(1)}$. In the second line, we consider the situation where $i \in \mathcal{C}^0$, i.e. where the company becomes insolvent only due to the premium payment. In this situation, the payoff equals at least the policyholder deficit $\tilde{\delta}_i^{(1)}$. However, if the fraction of the remaining assets in the fund allocatable to i (on a premium *pro-rata* basis) after covering the insolvencies of all companies $j \in \mathcal{C}^{\Pi}$ exceeds the insolvency $\tilde{\delta}_i^{(1)}$, the payoff is increased to equal this fraction. This ensures that "slightly" insolvent companies (after having paid a premium to the fund, see Figure 2) are not at a disadvantage. Finally, the third line of Equation (22) deals with the situation where the company is solvent at time $t = 1$: in this case, the residual assets after settling all deficits are distributed on a *pro-rata* basis of the respective premium payments ($\tilde{\pi}_i^{(1)}$ versus $\sum_{j \in \mathcal{C}^\star} \tilde{\pi}_j^{(1)}$). Notice that, given the third line, if no company becomes insolvent after the premium has been paid into the insurance guaranty fund, the fund reimburses the entire respective premium $\tilde{\pi}_i^{(1)}$ to each participating firm.

In the second case, where the guaranty fund goes bankrupt ($\tilde{\gamma}^{(1)} > 0$), a bankruptcy ratio is calculated for all companies ($\tilde{\delta}_i^{(1)}$ versus $\sum_{j \in (\mathcal{C}^\Pi \cup \mathcal{C}^0)} \tilde{\delta}_j^{(1)}$). If the company i is insolvent, only the fraction given by this ratio of the deficit $\tilde{\delta}_i^{(1)}$ is covered. If the company remains solvent, there is no payoff back to this company.

Proposition 1 *The payoff scheme of the fund, given by $\tilde{F}_i$ for companies $i \in \mathcal{C}$ defined in (22), yields a self-supporting guaranty fund, i.e., we have relation (10):*

$$\sum_{i \in \mathcal{C}} \tilde{F}_i^{(1)} = \sum_{i \in \mathcal{C}} \tilde{\pi}_i^{(1)}. \tag{10}$$

Proof Since $\mathcal{C} = \mathcal{C}^\Pi \cup \mathcal{C}^0 \cup \mathcal{C}^\star$ with the three subsets of $\mathcal{C}$ having empty intersection, we have

$$\sum_{i \in \mathcal{C}} \tilde{F}_i^{(1)} = \sum_{i \in \mathcal{C}^\Pi} \tilde{F}_i^{(1)} + \sum_{i \in \mathcal{C}^0} \tilde{F}_i^{(1)} + \sum_{i \in \mathcal{C}^\star} \tilde{F}_i^{(1)}. \tag{24}$$

In the case where $\tilde{\gamma}^{(1)} \leq 0$, we can write out (24) by introducing (22):

$$
\begin{aligned}
\sum_{i \in \mathcal{C}} \tilde{F}_i^{(1)} &= \sum_{i \in \mathcal{C}^\Pi} \tilde{\delta}_i^{(1)} + \sum_{i \in \mathcal{C}^0} \tilde{\kappa}_i^{(1)} \\
&\quad + \sum_{i \in \mathcal{C}^\star} \frac{\tilde{\pi}_i^{(1)}}{\sum_{j \in \mathcal{C}^\star} \tilde{\pi}_j^{(1)}} \left[\sum_{j \in \mathcal{C}} \tilde{\pi}_j^{(1)} - \sum_{j \in \mathcal{C}^\Pi} \tilde{\delta}_j^{(1)} - \sum_{j \in \mathcal{C}^0} \tilde{\kappa}_j^{(1)} \right] \\
&= \sum_{i \in \mathcal{C}^\Pi} \tilde{\delta}_i^{(1)} + \sum_{i \in \mathcal{C}^0} \tilde{\kappa}_i^{(1)} + \sum_{j \in \mathcal{C}} \tilde{\pi}_j^{(1)} - \sum_{j \in \mathcal{C}^\Pi} \tilde{\delta}_j^{(1)} - \sum_{j \in \mathcal{C}^0} \tilde{\kappa}_j^{(1)} \\
&= \sum_{i \in \mathcal{C}} \tilde{\pi}_i^{(1)}. \tag{25}
\end{aligned}
$$

Similarly, when $\tilde{\gamma}^{(1)} > 0$, relation (24) yields

$$\sum_{i \in \mathcal{C}} \tilde{F}_i^{(1)} = \sum_{i \in (\mathcal{C}^\Pi \cup \mathcal{C}^0)} \frac{\tilde{\delta}_i^{(1)}}{\sum_{j \in (\mathcal{C}^\Pi \cup \mathcal{C}^0)} \tilde{\delta}_j^{(1)}} \sum_{j \in \mathcal{C}} \tilde{\pi}_j^{(1)} + \sum_{i \in \mathcal{C}^\star} 0 = \sum_{i \in \mathcal{C}} \tilde{\pi}_i^{(1)}. \tag{26}$$

Combining both cases in (25) and (26), we complete the proof of Proposition 1.   ■

5.4 Particular case of homogeneous companies charged identical premiums

In the case where all M companies on the market are homogeneous in the sense of Property 2 and are charged an identical premium $\pi^{(0)}$—as they should be if we aim at charging risk-adjusted premiums—, we have $\langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}^{(1)} \rangle$, $\forall i \in \mathcal{C}$. If their policyholders have the same positive risk-aversion parameters $a_i = a > 0$, whenever this premium $\pi^{(0)}$ accompanies a fulfillment of Inequality (16), all groups of policyholders benefit to the same extent (compared to the situation without an insur-

ance guaranty fund). In Section 6, we illustrate this effect with the help of a numerical example.

Property 2 *In the case where all companies are homogeneous, i.e., when they have identical asset and claim distributions, as well as equal correlation structures between assets and claims, the aggregate-level relation (10) holds, on an individual company-level, with expected values, if all companies are charged the same premium:*

$$\pi_i^{(0)} = \pi^{(0)}, \forall i \in \mathcal{C} \quad \Rightarrow \quad \langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}_i^{(1)} \rangle = \langle \tilde{\pi}^{(1)} \rangle, \forall i \in \mathcal{C}. \tag{27}$$

Proof The homogeneity of the companies and the identity of the charged premiums imply that

$$\sum_{j \in \mathcal{C}} \langle \tilde{F}_j^{(1)} \rangle = M \langle \tilde{F}_i^{(1)} \rangle, \text{ i.e., } \langle \tilde{F}_i^{(1)} \rangle = \frac{1}{M} \sum_{j \in \mathcal{C}} \langle \tilde{F}_j^{(1)} \rangle, \forall i \in \mathcal{C}. \tag{28}$$

Similarly we have

$$\sum_{j \in \mathcal{C}} \langle \tilde{\pi}_j^{(1)} \rangle = M \langle \tilde{\pi}_i^{(1)} \rangle, \forall i \in \mathcal{C}. \tag{29}$$

Given Proposition 1 and Equation (10), we also have $\sum_{i \in \mathcal{C}} \langle \tilde{F}_i^{(1)} \rangle = \sum_{i \in \mathcal{C}} \langle \tilde{\pi}_i^{(1)} \rangle$. The result follows with (28) and (29). ∎

Property 3 *In the case of the exemplary payoff defined in (22) and where all companies are homogeneous, i.e., when they have identical asset and claim distributions, as well as equal correlation structures between assets and claims, we have*

$$\langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}_i^{(1)} \rangle, \forall i \in \mathcal{C} \quad \Rightarrow \quad \pi_i^{(0)} = \pi^{(0)}, \forall i \in \mathcal{C}. \tag{30}$$

Moreover, we have $\langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}_i^{(1)} \rangle = \langle \tilde{\pi}^{(1)} \rangle, \forall i \in \mathcal{C}$.

Proof We prove (30) by contradiction. We assume that $\langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}_i^{(1)} \rangle$, $\forall i \in \mathcal{C}$, and suppose that $\exists j \in \mathcal{C}$, such that $\pi_j^{(0)} \neq \pi^{(0)}$. Without any loss of generality, we can suppose that $\pi_j^{(0)} = \alpha \pi^{(0)}$, where $\alpha \geq 0$, $\alpha \neq 1$. We have $\langle \tilde{\pi}_j^{(1)} \rangle = \alpha \langle \tilde{\pi}^{(1)} \rangle$.

However, since (22), the payoff for a company i, $\tilde{F}_i^{(1)}$, is, in general, not proportional to the charged premium $\tilde{\pi}_i^{(1)}$. Hence, in the present case, we have $\langle \tilde{F}_j^{(1)} \rangle = \beta \langle \tilde{\pi}^{(1)} \rangle$, with $\beta \geq 0$, which, since in general $\beta \neq \alpha$, leads to a contradiction with the assumption $\langle \tilde{F}_j^{(1)} \rangle = \langle \tilde{\pi}_j^{(1)} \rangle = \alpha \langle \tilde{\pi}^{(1)} \rangle$. ∎

6 Numerical examples

In this section numerical results are reported and discussed. In the case of homogeneous companies, the sensitivities of the diversification benefit $\Delta_a \phi_i^{(1)}$ and of the default probabilities on different sets of companies are analyzed. In a setup with heterogeneous companies, we illustrate the premium levels resulting from the application of the utility-based calculation with several examples of sets of companies.

The following analysis implemented by means of a Monte Carlo simulation, with $N = 1\,000\,000$ iterations, allows us to explore how the diversification benefit and premium levels change with respect to varying input parameters. For the purpose of this simulation, we use the payoff scheme of the insurance guaranty fund derived in Section 5.3, and specified in Equation (22).

Asset as well as claim returns, denoted by $\tilde{r}_i^A$ and $\tilde{r}_i^S$, respectively, are modeled as normally distributed variables. Hence, with

$$\tilde{A}_i^{(1)} = A_i^{(0)} e^{\tilde{r}_i^A}, \tag{31}$$
$$\tilde{\pi}_i^{(1)} = \pi_i^{(0)} e^{\tilde{r}_i^A}, \tag{32}$$
$$\tilde{S}_i^{(1)} = S_i^{(0)} e^{\tilde{r}_i^S}, \tag{33}$$

the variables $\tilde{A}_i^{(1)}$, $\tilde{\pi}_i^{(1)}$, and $\tilde{S}_i^{(1)}$ at time $t = 1$ follow lognormal distributions.

For the numerical analysis, we always refer to a standard case for a company $i \in \mathcal{C}$ with the following parameterization: assets at $t = 0$, $A_i^{(0)} = 60$, charged premium $\pi_i^{(0)} = 5$, claims $S_i^{(0)} = 40$. The asset and claim return distributions are modeled with their expected value $\langle \tilde{r}_i^A \rangle = 0.15$, and standard deviation $\sigma(\tilde{r}_i^A) = 0.2$, respectively $\langle \tilde{r}_i^S \rangle = 0.1$,

$\sigma(\tilde{r}_i^S) = 0.15$. Furthermore, the correlation between asset and claim returns between different companies $i, j \in \mathcal{C}$ is set to $\rho(\tilde{r}_i^A, \tilde{r}_j^A) = 0.4$ and $\rho(\tilde{r}_i^S, \tilde{r}_j^S) = 0.3$, respectively. We assume the same risk aversion parameter $a_i = 2$ for all companies in their respective utility function, see Equations (11) and (12).

6.1 Diversification benefit in the case of homogeneous companies

We consider a set $\mathcal{C}$ of $M = 10$ homogeneous insurance companies. In this case, the fund charge calculation is proceeded so that $\langle \tilde{F}_i^{(1)} \rangle = \langle \tilde{\pi}_i^{(1)} \rangle = \langle \tilde{\pi}^{(1)} \rangle, \forall i \in \mathcal{C}$. This means, given Property 3, that all companies are charged the same premium, i.e., that we have $\pi_i^{(0)} = \pi^{(0)}, \forall i \in \mathcal{C}$.

In the following, we report the absolute change in utility $\Delta_a \phi_i^{(1)}$, introduced in Equation (13) and the influence of its composing elements $\langle \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)} \rangle$, $-\frac{a_i}{2} \text{var}(\tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)})$, and $-a_i \text{cov}(\tilde{A}_i^{(1)}, \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)})$. We analyze their sensitivity with respect to a variation of

- the correlation between asset returns for all insurers $\rho(\tilde{r}_i^A, \tilde{r}_j^A)$, see Figure 3,

- the correlation between claim returns $\rho(\tilde{r}_i^S, \tilde{r}_j^S)$, see Figure 4,

- the standard deviation of the asset returns $\sigma(\tilde{r}_i^A)$, see Figure 5,

- the premium $\pi^{(0)}$ charged by the fund, see Figure 6(a), and

- the number of companies M in the market, see Figure 6(b).

The findings of the sensitivity analyses are summarized as follows. Figure 3 shows that an increase of the correlations between the asset portfolios of the different insurers leads, other things being equal, to a decrease of the diversification benefit measured by $\Delta_a \phi_i^{(1)}$. Furthermore, the graph shows that the term $\text{cov}(\tilde{A}_i^{(1)}, \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)})$ plays a dominant role. At a correlation coefficient above $\rho(\tilde{r}_i^A, \tilde{r}_j^A) = 0.87$, the utility in the setup with the guaranty fund is below the utility in a setting without an insurance guaranty fund. This is due to the fact that, for an increasing correlation between insurer assets, less and less diversification can take place in the fund. This is an important issue, since we would typically expect high asset correlations within capital markets.
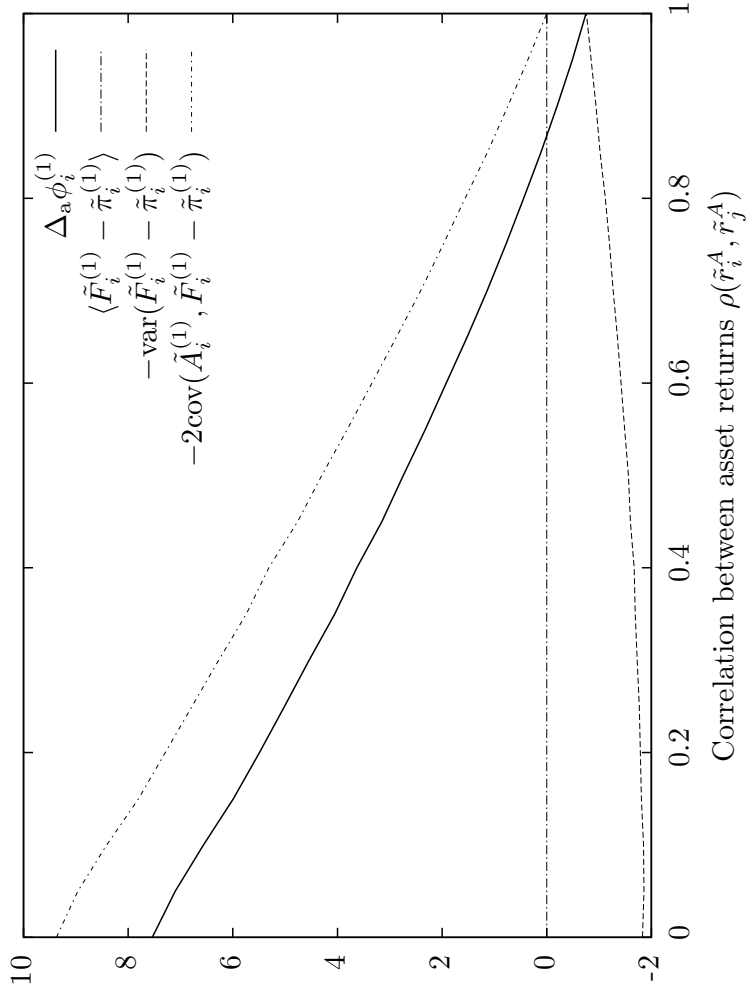
Figure 3: Sensitivity of the change in utility $\Delta_a \phi_i^{(1)}$ on the correlation between asset returns $\rho(\tilde{r}_i^A, \tilde{r}_j^A)$. Parameters: $M = 10$, $a_i = 2$, $\pi_i^{(0)} = 5$, $A_i^{(0)} = 60$, $\langle \tilde{r}_i^A \rangle = 0.15$, $\sigma(\tilde{r}_i^A) = 0.15$, $S_i^{(0)} = 40$, $\langle \tilde{r}_i^S \rangle = 0.1$, $\sigma(\tilde{r}_i^S) = 0.15$, $\rho(\tilde{r}_i^S, \tilde{r}_j^S) = 0.3$.

An increase of the correlation between the claims of the different insurance companies, implies a slight increase of the diversification benefit (arising from a slight decrease of the term $\text{var}(\tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)})$). However, its influence is less significant compared to the impact of variations in the correlations between insurer assets. This sensitivity analysis is shown in the graph of Figure 4.

In the analysis varying the standard deviation of asset returns $\sigma(\tilde{r}_i^A)$, we observe, see Figure 5, a slightly negative effect of pooling, $\Delta_\text{a}\phi_i^{(1)} < 0$, in a situation in which assets—in the insurance companies and, hence, in the guaranty fund—are invested risk-free, i.e. $\sigma(\tilde{r}_i^A) = 0$ (see also the discussion in Paragraph 4.3). Hence, within the model framework provided in this paper, a certain level of uncertainty in asset returns is needed for an insurance guaranty fund to become beneficial for policyholders.

Figures 6(a) and 6(b) illustrate the diversification effects with different premium charges $\pi_i^{(0)}$ and number M of market participants. In both cases, the initially quickly increasing diversification benefit eventually flattens out. Note that both composing elements of the change in utility, $\text{var}(\tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)})$ and $\text{cov}(\tilde{A}_i^{(1)}, \tilde{F}_i^{(1)} - \tilde{\pi}_i^{(1)})$, show similar trends, however, with opposite signs, for larger values of $\pi_i^{(0)}$, respectively M. In the case of increasing premium charges, this shows that no significant benefits can be obtained once the necessary funds for decreasing the default probabilities to a minimum level are collected. This feature can also be recognized in Figure 7(a), where the default probabilities are shown. Once the fund reaches a certain premium volume and the default probabilities of the companies get close to the default probability of the fund (given its asset base of the premiums collected), the cost of lowering the default probabilities, or increasing the diversification benefit, becomes relatively high. Both figures show that "naïve" diversification is limited and that there is a boundary where no significant benefits arise.

Finally, in Figure 7, we analyze the risk by the policyholders that their liabilities may not be covered, namely, without guaranty fund given by $\mathcal{D}_i^{(1)} = \text{Prob}(\tilde{A}_i^{(1)} < \tilde{S}_i^{(1)})$, and with the payoff of the guaranty fund, $\mathcal{D}_i^{(1),\text{f}} = \text{Prob}(\tilde{A}_i^{(1)} - \tilde{\pi}_i^{(1)} + \tilde{F}_i^{(1)} < \tilde{S}_i^{(1)})$. We then compare these values to the default risk of the company after paying the premium to the fund, $\mathcal{D}_i^{(1),\star} = \text{Prob}(\tilde{A}_i^{(1),\star} < \tilde{S}_i^{(1)})$ and the default risk of the insurance guar-
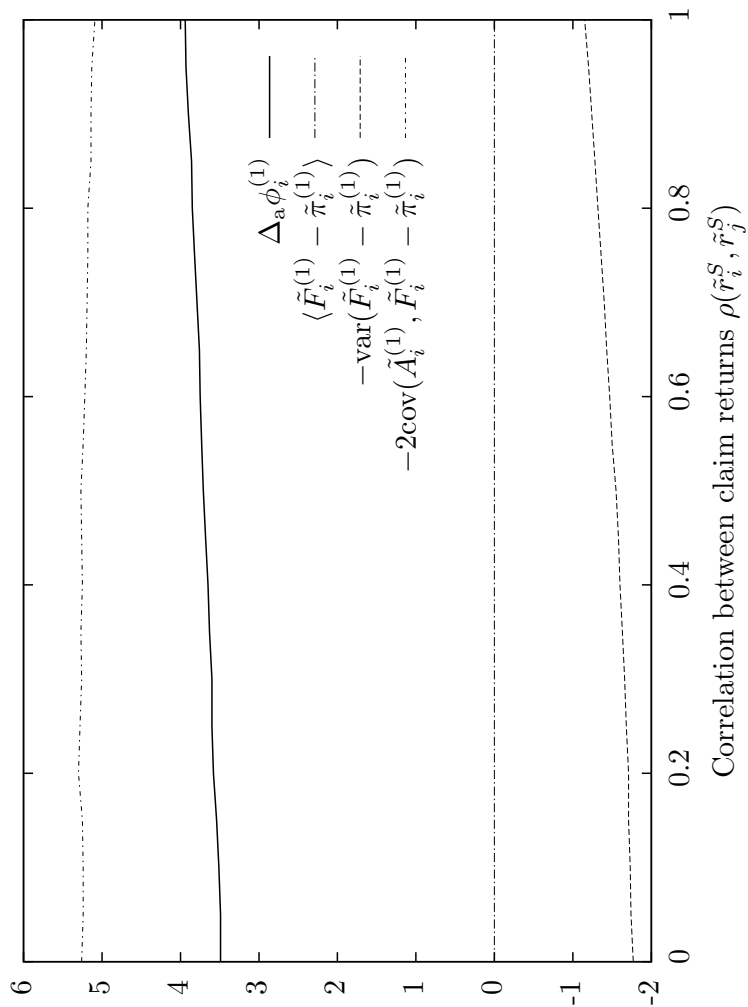
Figure 4: Sensitivity of the change in utility $\Delta_a \phi_i^{(1)}$ on the correlation between claim returns $\rho(\tilde{r}_i^S, \tilde{r}_j^S)$. Parameters: $M = 10$, $a_i = 2$, $\pi_i^{(0)} = 5$, $A_i^{(0)} = 60$, $\langle \tilde{r}_i^A \rangle = 0.15$, $\langle \tilde{r}_i^A \rangle = 0.2$, $\rho(\tilde{r}_i^A, \tilde{r}_j^A) = 0.4$ $S_i^{(0)} = 40$, $\langle \tilde{r}_i^S \rangle = 0.1$, $\sigma(\tilde{r}_i^A) = 0.15$, $\sigma(\tilde{r}_i^A) = 0.2$, $\rho(\tilde{r}_i^A, \tilde{r}_j^A) = 0.4$ $S_i^{(0)} = 40$, $\langle \tilde{r}_i^S \rangle = 0.1$, $\sigma(\tilde{r}_i^S) = 0.15$.
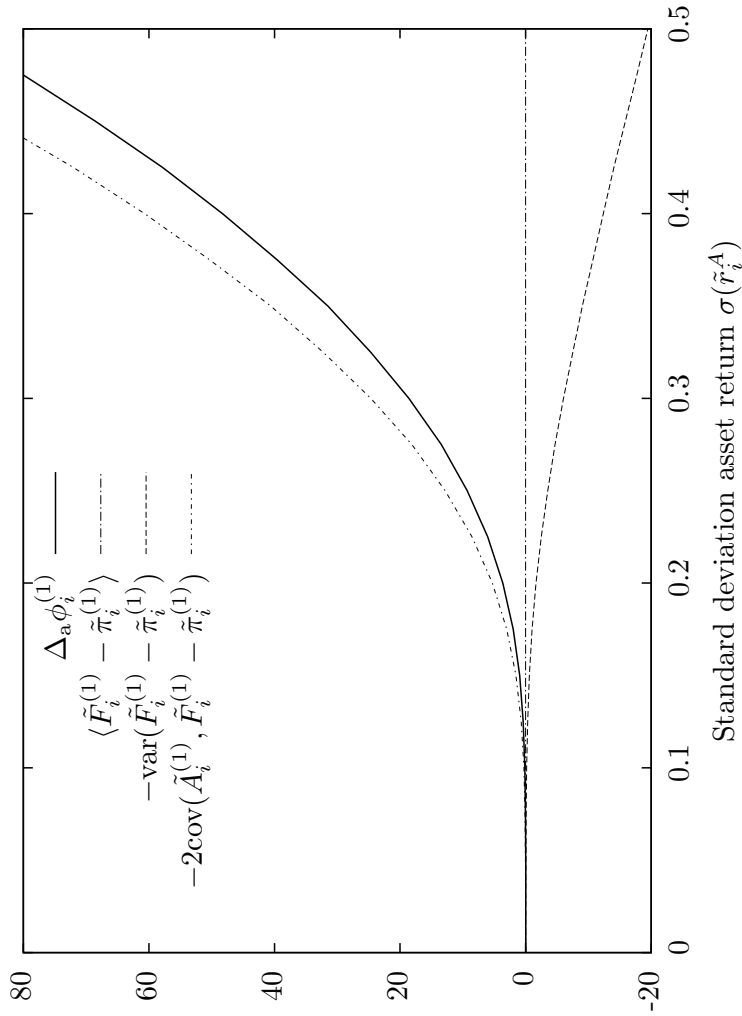
Figure 5: Sensitivity of the change in utility $\Delta_a \phi_i^{(1)}$ on the standard deviation of asset returns $\sigma(\tilde{r}_i^A)$. Parameters: $M = 10$, $a_i = 2$, $\pi_i^{(0)} = 5$, $A_i^{(0)} = 60$, $\langle \tilde{r}_i^A \rangle = 0.15$, $\rho(\tilde{r}_i^A, \tilde{r}_j^A) = 0.4$ $S_i^{(0)} = 40$, $\langle \tilde{r}_i^S \rangle = 0.1$, $\sigma(\tilde{r}_i^S) = 0.15$, $\rho(\tilde{r}_i^S, \tilde{r}_j^S) = 0.3$.
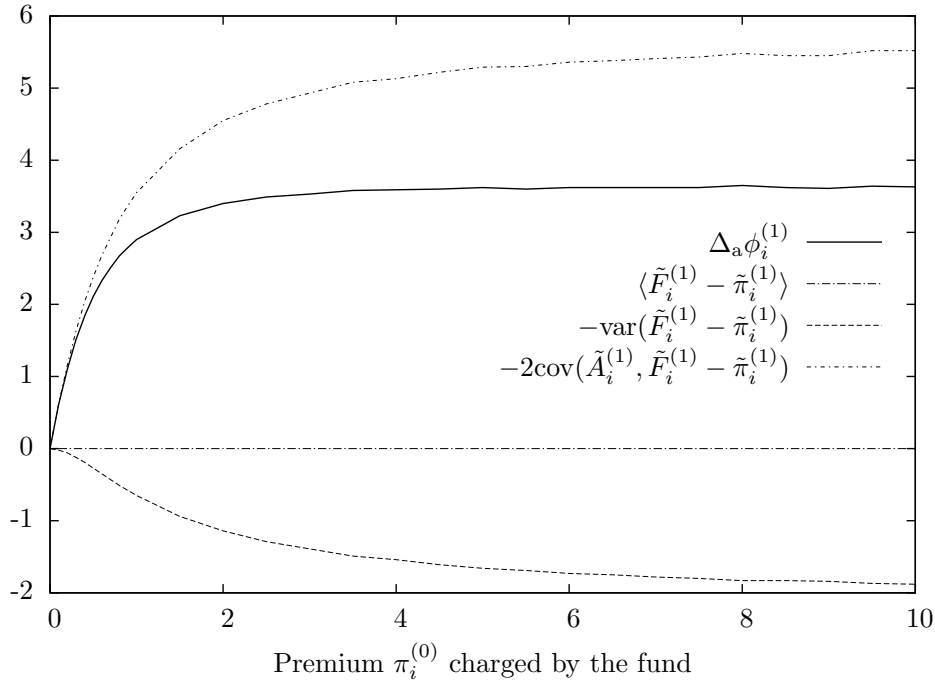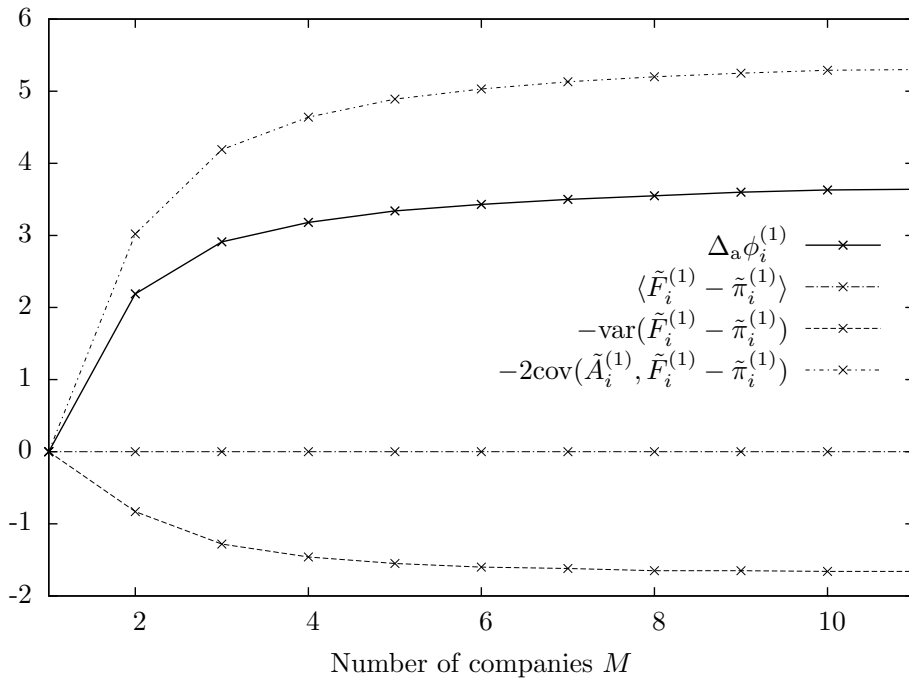
(a) Sensitivity on $\pi_i^{(0)}$

(b) Sensitivity on M

Figure 6: Sensitivity of the change in utility $\Delta_{\mathrm{a}}\phi_i^{(1)}$ (a) on the premium $\pi_i^{(0)}$ charged by the fund and, (b) the number of companies M. Parameters: $M = 10$, $a_i = 2$, $\pi_i^{(0)} = 5$, $A_i^{(0)} = 60$, $\langle \tilde{r}_i^A \rangle = 0.15$, $\sigma(\tilde{r}_i^A) = 0.2$, $\rho(\tilde{r}_i^A, \tilde{r}_j^A) = 0.4$, $S_i^{(0)} = 40$, $\langle \tilde{r}_i^S \rangle = 0.1$, $\sigma(\tilde{r}_i^S) = 0.15$, $\rho(\tilde{r}_i^S, \tilde{r}_j^S) = 0.3$.

(a) Sensitivity on $\pi_i^{(0)}$
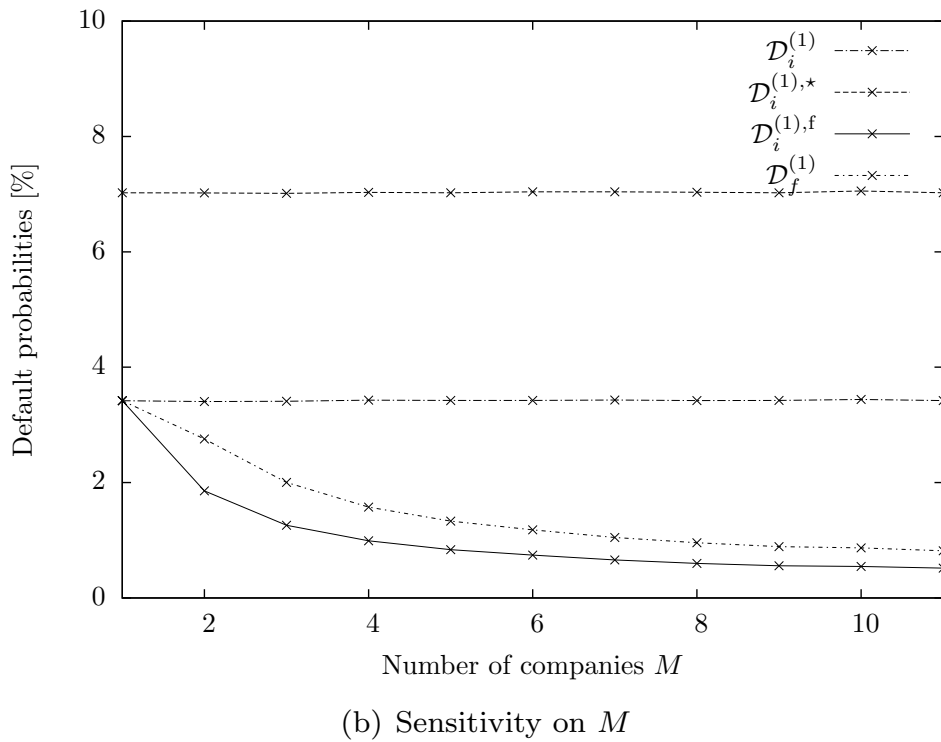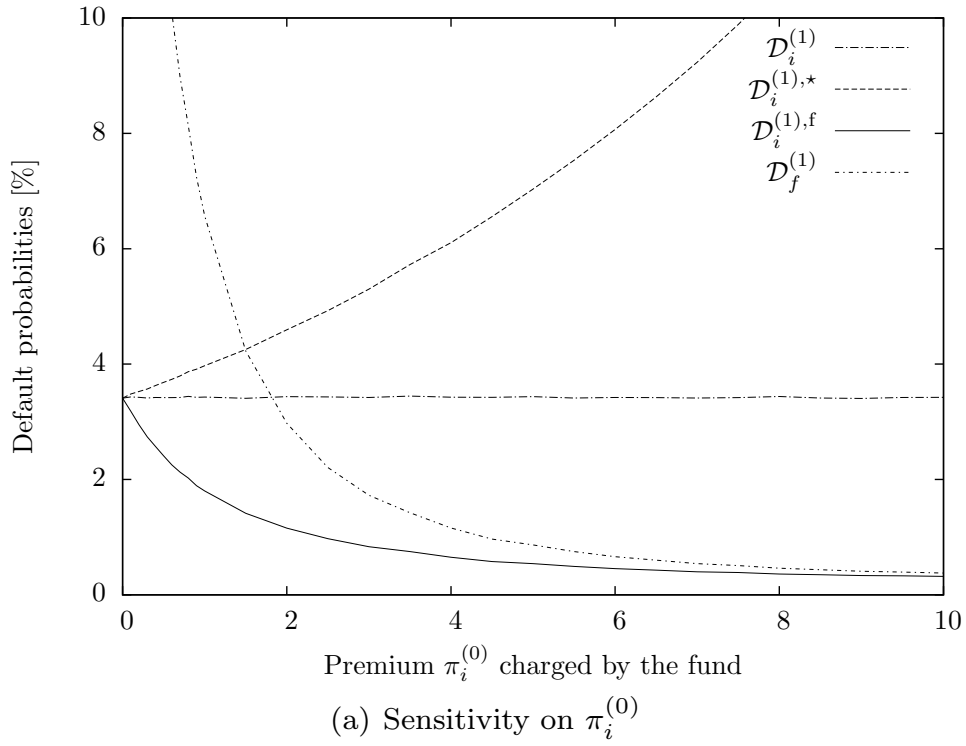
(b) Sensitivity on M

Figure 7: Sensitivity of the default probabilities on the premium $\pi_i^{(0)}$ charged by the fund and the number M of participating companies.

anty fund, $\mathcal{D}_{\mathrm{f}}^{(1)} = \mathrm{Prob}(\tilde{\gamma}^{(1)} > 0)$. As discussed earlier, the default risk for the liabilities of the policyholders is reduced, see the curve $\mathcal{D}_{i}^{(1),\mathrm{f}}$, to the default risk of the fund, $\mathcal{D}_{\mathrm{f}}^{(1)}$. After the fund premium is charged, we also notice an increase in the default risk $\mathcal{D}_{i}^{(1),\star}$ for the insurance company. In this manner, we can observe that particularly for insurers with intermediate solvency status, which are likely to go bankrupt only due to the premium payment ($i \in \mathcal{C}^0$ in Figure 2), the introduction of an insurance guaranty fund charging high premiums is similar to early closure rules within solvency frameworks. Since the increasing fund charges result in an increasing default risk of the insurance company, in practice, there should be a limit for the guaranty fund charges calibrated according to the solvency regulations in-force.

6.2   Calculation of utility-based premiums

From now on, we consider several sets of $M = 5$ insurance companies based on variations of the standard parameters introduced above and calculate their premiums. We apply the utility-based premium principle (18), use the relative utility increase $\Delta_{\mathrm{r}}\phi_{i}^{(1)}$ as a utility measure, and require $\Delta_{\mathrm{r}}\phi_{i}^{(1)} = K = 1.00\%$. The resulting premiums charged to different sets of companies, with different standard deviations of asset and claim returns, denoted by I and II, and of different sizes in terms of assets and claims, denoted by III through V, are given in Tables 1 and 2 and are compared to the situation with homogeneous companies. Note that the homogeneous case refers to the standard parameterization introduced earlier in this section.

The set I of heterogeneous companies in Table 1 considers companies with different asset return volatilities. We note that companies with higher volatilities are charged higher premiums compared to their peers. The set of companies II, considered in Table 1, refers to companies with different volatilities of claim returns: numerical optimization of the premium principle leads to higher premiums for companies with higher claim returns volatilities. Hence, within this specific example, the utility-based approach leads to appropriate incentives, since it accounts for the risk of the insurers.

companies	homogeneous			heterogeneous I			heterogeneous II		
i	$\sigma(\tilde{r}_i^A)$	$\sigma(\tilde{r}_i^S)$	$\pi_i^{(0)}$	$\sigma(\tilde{r}_i^A)$	$\sigma(\tilde{r}_i^S)$	$\pi_i^{(0)}$	$\sigma(\tilde{r}_i^A)$	$\sigma(\tilde{r}_i^S)$	$\pi_i^{(0)}$
1	0.20	0.15	0.49	0.20	0.15	0.45	0.20	0.15	0.51
2	—	—	—	0.16	—	0.22	—	0.11	0.16
3	—	—	—	0.18	—	0.31	—	0.13	0.30
4	—	—	—	0.22	—	0.64	—	0.17	0.69
5	—	—	—	0.24	—	0.82	—	0.19	0.82
$\Delta_{\mathrm{r}}\phi_i^{(1)}$	**1.00 %**			**1.00 %**			**1.00 %**		
$\sum_{i\in C}\pi_i^{(0)}$	2.45			2.44			2.48		
$\langle\boldsymbol{\Pi^{(0)}}\rangle$	0.49			0.49			0.50		
$\sigma(\boldsymbol{\Pi^{(0)}})$	0.00			0.24			0.27		
$\Delta_{\mathrm{r}}\phi_i^{(1)}/\langle\langle\boldsymbol{\Pi^{(0)}}\rangle\rangle$	2.04 %			2.04 %			2.00 %		

Table 1: Set of companies I and II with standard deviations on asset and claim returns. Premiums are calculated according to the principle $\Delta_{\mathrm{r}}\phi_i^{(1)}$ set at $K = 1.00\%$. Parameters: $M = 5$, $a_i = 2$, $A_i^{(0)} = 60$, $\langle\tilde{r}_i^A\rangle = 0.15$, $\rho(\tilde{r}_i^A, \tilde{r}_j^A) = 0.4$, $S_i^{(0)} = 40$, $\langle\tilde{r}_i^S\rangle = 0.1$, $\rho(\tilde{r}_i^S, \tilde{r}_j^S) = 0.3$.

The sets III to V of heterogeneous companies considered in Table 2 refer to companies of different sizes in assets and claims. In set III, where the initial asset base is different for the different companies (with equal claims base and identical asset and claim standard distributions), corresponding to a set of companies with range of claims ratios $S_i^{(0)}/A_i^{(0)}$ between 64.5% and 69.0% at $t = 0$, we note that the higher the assets, the lower the charged premium. In the case where all companies have the same initial asset base, but different magnitudes in claims (set IV, $S_i^{(0)}/A_i^{(0)}$ between 63.3% and 70.0%), we obtain that higher claims induce higher premiums. Finally, in set V, we consider companies of different sizes, where the initial asset and claim bases are scaled proportionally (the largest company $i = 5$ is four times larger than the smallest company $i = 2$), but asset and claim distributions are kept identical. Numerical simulation of this exemplary situation shows that larger companies are charged lower premiums.

Let us finally point out that, in the examples shown, all companies can expect an (equal) increase in utility due to the accordingly adapted premium charges, regardless of their asset or claim distributions. In comparing the reported results, we also note that the relative utility increase per premium is similar on an aggregate level (ratio $\sum_{i \in \mathcal{C}} \Delta_r \phi_i^{(1)} = M \Delta_r \phi_i^{(1)}$ over $\sum_{i \in \mathcal{C}} \pi_i^{(0)}$), whereas on an individual level the marginal relative utility increase $\Delta_r \phi_i^{(1)}/\pi_i^{(0)}$ is substantially varying.

7   Summary and outlook

The contingent claim approach is often suggested in the literature as an approach to derive risk-based premiums, which should be charged by an insurance guaranty fund for the protection it offers against insurer insolvencies. However, we make the point that within perfect markets, the introduction of such an insurance guaranty fund cannot improve the wealth position of the policyholders if all stakes are priced fairly. If we abandon the assumption of a perfect market, we also show that risk-neutral investors cannot benefit through the existence of a self-supporting insurance guaranty fund. Hence, if the roll-out of a self-supporting insurance

companies	heterogeneous III			heterogeneous IV			heterogeneous V		
i	$A_i^{(0)}$	$S_i^{(0)}$	$\pi_i^{(0)}$	$A_i^{(0)}$	$S_i^{(0)}$	$\pi_i^{(0)}$	$A_i^{(0)}$	$S_i^{(0)}$	$\pi_i^{(0)}$
1	60	40	0.50	60	40	0.50	60	40	0.83
2	58	—	0.84	—	38	0.06	30	20	0.95
3	59	—	0.65	—	39	0.27	45	30	0.94
4	61	—	0.33	—	41	0.74	90	60	0.50
5	62	—	0.18	—	42	0.98	120	80	0.12
$\Delta_{\mathrm{r}}\phi_i^{(1)}$	**1.00 %**			**1.00 %**			**1.00 %**		
$\sum_{i\in c}\pi_i^{(0)}$	2.50			2.55			3.34		
$\langle \mathbf{\Pi^{(0)}}\rangle$	0.50			0.51			0.67		
$\sigma(\mathbf{\Pi^{(0)}})$	0.26			0.36			0.36		
$\Delta_{\mathrm{r}}\phi_i^{(1)}/\langle\mathbf{\Pi^{(0)}}\rangle$	2.01 %			1.96 %			1.49 %		

Table 2: Set of companies III to V, with different sizes in terms of assets and claims. Premiums are calculated according to the principle $\Delta_{\mathrm{r}}\phi_i^{(1)}$ set at $K = 1.00\%$. Compare with the results from the homogeneous case in Table 1, $A_i^{(0)} = 60$, $S_i^{(0)} = 40$. Parameters: $M = 5$, $a_i = 2$, $\langle \tilde{r}_i^A \rangle = 0.15$, $\sigma(\tilde{r}_i^A) = 0.2$, $\rho(\tilde{r}_i^A, \tilde{r}_j^A) = 0.4$, $\langle \tilde{r}_i^S \rangle = 0.1$, $\sigma(\tilde{r}_i^S) = 0.15$, $\rho(\tilde{r}_i^S, \tilde{r}_j^S) = 0.3$.

guaranty fund implies transaction costs, its introduction is detrimental to those insured in both cases.

Matters may change in the more likely case of risk-averse investors and imperfect markets. The potential diversification benefit, which may be achieved by pooling claims in an insurance guaranty fund, may improve the wealth position of those insured. However, a diversification advantage (or disadvantage) measured through the increase in the utility of policyholders, is only equal for every single insurer if companies are homogeneous, have the same utility function and have an identical degree of risk aversion.

The problem of allocating possible diversification benefits attained in an insurance guaranty fund back to heterogeneous insurance companies, in an imperfect market setting with risk-averse policyholders, is similar to the capital allocation problem widely discussed in the academic literature over the last couple of years. It implies that there is no non-arbitrary way to allocate the benefits from diversifying the unsystematic risks within a pooled portfolio. Following this line of reasoning, no non-arbitrary allocation of the collective premium, within an insurance guaranty fund back to the different insurance companies, is possible if the conditions of a perfect market are not fulfilled. Different premium principles, based on the individual risk profile of an insurer and used to derive the fund charges, lead, in general, to a situation in which policyholders are treated unequally—in the sense of a utility increase—through the introduction of an insurance guaranty fund. Even if each policyholder enjoys a diversification benefit, some insured parties will benefit more than others. To counteract this effect, we introduce the concept of a utility-based premium calculation principle to derive charges for an insurance guaranty fund and discuss its implications within the insurance market. However, the difficulties in measuring the utility of policyholders essentially constrain its practical relevance.

On the one hand, our analysis reveals that, in general, wealth or utility transfers between policyholders of different insurers are unavoidable. In our setting, insurance guaranty funds may even be systematically unfavorable for all policyholders, if, for example, all insurers invest solely in risk-free assets, or in highly correlated assets. In our opinion, the

mentioned aspects can be seen as arguments against the introduction of an insurance guaranty fund. On the other hand, the analyzed literature often points out that insurance guaranty funds may have positive influence on the agency problems, within insurance markets. Examination of the interactions between both facets of the underlying problem should be of interest for further research. In addition, other solutions leading to a controlled run-off of an insurance company should be analyzed in more detail in the future.

References

Axelrod, R. and Hamilton, W. D. (1981). The Evolution of Cooperation. *Science*, 211(4489):1390–1396.

Borch, K. (1962). Equilibrium in a Reinsurance Market. *Econometrica*, 30(3):424–444.

Brewer-III, E., Mondschean, T. S., and Strahan, P. E. (1997). The Role of Monitoring in Reducing the Moral Hazard Problem Associated with Government Guarantees: Evidence from the Life Insurance Industry. *Journal of Risk and Insurance*, 64(2):301–322.

Butsic, R. P. (1994). Solvency Measurement for Property-Liability Risk-Based Capital Applications. *Journal of Risk and Insurance*, 61(4):656–690.

Chamberlain, G. (1983). A Characterization of the Distributions that Imply Mean-Variance Utility Functions. *Journal of Economic Theory*, 29(1):185–201.

Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS) (2009). CEOIPS Input to the European Commission work on Insurance Guarantee Schemes. *URL http://www.ceiops.eu*.

Cummins, J. D. (1988). Risk-Based Premiums for Insurance Guaranty Funds. *Journal of Finance*, 43(4):823–839.

Doherty, N. A. and Garven, J. R. (1986). Price Regulation in Property-Liability Insurance: A Contingent-Claims Approach. *Journal of Finance*, 41(5):1031–1050.

Downs, D. H. and Sommer, D. W. (1999). Monitoring, Ownership, and Risk-Taking: The Impact of Guaranty Funds. *Journal of Risk and Insurance*, 66(3):477–497.

Duan, J.-C. and Yu, M.-T. (2005). Fair Insurance Guaranty Premia in the Presence of Risk-Based Capital Regulations, Stochastic Interest Rate and Catastrophe Risk. *Journal of Banking and Finance*, 29(10):2435–2454.

Farquhar, P. H. (1984). Utility Assessment Methods. *Management Science*, 30(11):1283–1300.

Feldhaus, W. R. and Kazenski, P. M. (1998). Risk-Based Guaranty Fund Assessments. *Journal of Insurance Regulation*, 17(1):42–63.

Fennema, H. and Van Assen, M. (1998). Measuring the Utility of Losses by Means of the Tradeoff Method. *Journal of Risk and Uncertainty*, 17(3):277–296.

Fischer, S. (1978). Call Option Pricing When the Exercise Price is Uncertain, and the Valuation of Index Bonds. *Journal of Finance*, 33(1):169–176.

Gründl, H. and Schmeiser, H. (2007). Capital Allocation for Insurance Companies - What Good Is It? *Journal of Risk and Insurance*, 74(2):301–317.

Han, L.-M., Lai, G. C., and Witt, R. C. (1997). A Financial-Economic Evaluation of Insurance Guaranty Fund System: An Agency Cost Perspective. *Journal of Banking and Finance*, 21(8):1107–1129.

Ibragimov, R., Jaffee, D., and Walden, J. (2010). Pricing and Capital Allocation for Multiline Insurance Firms. *Journal of Risk and Insurance*, 77(3):551–578.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*, 47(2):263–291.

Lee, S.-J., Mayers, D., and Smith, C. W. (1997). Guaranty Funds and Risk-Taking Evidence from the Insurance Industry. *Journal of Financial Economics*, 44(1):3–24.

Margrabe, W. (1978). The Value of an Option to Exchange One Asset for Another. *Journal of Finance*, 33(1):177–186.

Mayers, D. and Smith, C. W. J. (1990). On the Corporate Demand for Insurance: Evidence from the Reinsurance Market. *Journal of Business*, 63(1):19–40.

Mayes, D. G. (2004). Who Pays for Bank Insolvency? *Journal of International Money and Finance*, 23(3):515–551.

Merton, R. C. (1977). An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantes. An Application of Modern Option Pricing Theory. *Journal of Banking and Finance*, 1(1):3–11.

Merton, R. C. and Perold, A. (1993). Theory of Risk Capital in Financial Firms. *Journal of Applied Corporate Finance*, 6(3):16–32.

Mill, J. S. (1848). *Principles of Political Economy*, volume 2. John W. Parker, West Strand, London.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4):768–783.

Myers, S. C. and Read, J. A. J. (2001). Capital Allocation for Insurance Companies. *Journal of Risk and Insurance*, 68(4):545–580.

Neumann von, J. and Morgenstern, O. (1953). *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 2nd edition.

Oxera (2007). Insurance Guarantee Schemes in the EU: Comparative Analysis of Existing Schemes, Analysis of Problems and Evaluation of Options. *URL http://ec.europa.eu.*

Phillips, R. D., Cummins, J. D., and Allen, F. (1998). Financial Pricing of Insurance in the Multiple-Line Insurance Company. *Journal of Risk and Insurance*, 65(4):597–636.

Plantin, G. (2006). Does Reinsurance Need Reinsurers? *Journal of Risk and Insurance*, 73(1):153–168.

Sherris, M. (2006). Solvency, Capital Allocation, and Fair Rate of Return in Insurance. *Journal of Risk and Insurance*, 73(1):71–96.

Sommer, D. W. (1996). The Impact of Firm Risk on Property-Liability Insurance Prices. *Journal of Risk and Insurance*, 63(4):597–636.

Zanjani, G. (2010). An Economic Approach to Capital Allocation. *Journal of Risk and Insurance*, 77(3):523–549.

Part II

Insurance Guaranty Funds and Their Relation to Solvency Regulation

Abstract

In this paper, we discuss the interdependencies between the present regulatory frameworks and different designs of insurance guaranty funds. We argue that these reciprocal effects constitute the need for an enhanced regulatory approach extended by issues connected with existing institutionalized run-off solutions. Particularly, appropriately designed guaranty funds have to be integrated into ongoing supervisory processes, supporting many of the broadly acknowledged goals of insurance prudential regulation and supervision. However, there is no easy answer regarding the calculation of risk-based premiums within insurance guaranty funds. Hence, the risk of cross-subsidization between different insurance companies remains for countries that have introduced insurance guaranty funds within their regulatory frameworks.[9]

---

1 Introduction

The last two decades have seen a substantial and almost revolutionary change in the regulation of financial markets. Two facets of this change—an increasingly global trend toward integrated prudential regulation and supervision and the pursuit of enhanced market discipline—are probably here to stay and are expected to play a large role in future regulatory frameworks covering the insurance, banking, and pension fund sectors. However, the recent financial crisis has revealed that there is still much work to be done in order to ensure the sustainable stability of international financial markets (see, e.g., CEIOPS, 2009b; the Geneva Association, 2010; IIF, 2010). The financial crisis has given rise to an increased interest in guaranty funds, which offer protection against possible insolvencies of financial institutions. Those institutions, currently intensively discussed (see, e.g., CEIOPS, 2009a; European Commission, 2010), are only one of many possible approaches for the organization of company's institutionalized run-off within the insurance industry.

Even though insurance guaranty funds are already employed as a means of policyholder protection in several countries (for an overview, see, e.g., Oxera, 2007), leading regulatory frameworks, both those being currently developed and those already in place, do not include guaranty funds as an essential part of the regulatory system, despite their potential to interact with regulatory schemes with the end result being either enhanced or reduced stability of the insurance sector.[10] In this paper, we discuss the interdependencies between different designs for guaranty funds and present regulatory frameworks based on market-consistent valuation of an insurer's balance sheet. We argue that if the legislator introduces an insurance guaranty fund, there will be a consequent need for specific regulation aimed at guaranty fund issues. For example, there is no easy answer as to how best to calculate risk-based premiums within insurance guaranty funds and hence the risk of cross-subsidization between different insurance companies remains for countries that have introduced insurance guaranty funds within their regulatory frameworks.

---

[10]For instance, the premium charged by the insurance guaranty fund influences the solvency level of the insurer.

The remainder of this paper is structured as follows. In Section 2, insurance guaranty funds are discussed in the general context of orderly run-off solutions for insurers. In Section 3, we focus on the main goals of insurance regulation and supervision and evaluate to what extent current solvency frameworks support those objectives. Furthermore, we analyze whether different designs for insurance guaranty funds could be used to facilitate achievement of regulatory goals. In Section 4, we analyze different aspects of a potentially holistic approach to solvency regulation, including issues related to the introduction of insurance guaranty funds. We conclude in Section 5.

2 Application for resolving failing insurers

In this section, we present the arguments often raised for and against predefined and continuously updated plans in the event of a company's default within the financial industry. We define a concept for an orderly run-off as a predefined and institutionalized plan for supervisor reactions to insurer insolvency that is different from liquidation procedures ordinarily employed in other, nonfinancial sectors. In addition, we describe different basic organizational forms of insurance guaranty funds and provide an overview of potential winding-up solutions within the insurance industry.

2.1 Run-off solutions in general

The recent financial crisis has shown that lack of an economically funded concept for an orderly run-off of systemically relevant financial institutions can result in extensive problems in times of serious financial distress (see, e.g., IIF, 2010). On the one hand, if there is no at least general plan for a controlled resolution of a financial institution, governments and regulators may be unprepared and thus less effective in abating the negative economic phenomena resulting from a company's default (see European Commission, 2010). On the other hand, the example of recent major bailouts within the banking industry shows that in case of substantial market distortions, it is often the taxpayer who ends up paying for

a financial institution's insolvency. This situation has the potential not only to provide inappropriate incentives for diverse market participants, but can also be viewed as inefficient from an economic point of view. Using public funds to support mismanaged financial intermediaries can lead to a serious misallocation of economic resources. Regulation that requires a company to have a continuously updated and clearly predefined framework for a controlled winding-up of its affairs in the event of insolvency may solve some of the underlying problems connected with default events of financial service providers. This is a particularly important issue in case of the so-called too-big-to-fail companies, where a collapse would be exceptionally drastic for the entire economic system.

Whether the same arguments are also relevant to the insurance sector depends strongly on whether and, if so, to what degree, the industry is host to systemic risk, a topic of intense current interest (see Section 3.1). If the insurance industry does have systemic characteristics, such would be an argument in favor of a controlled resolution approach for insurance companies. Depending on its specific structure and funding, a guaranty scheme based on transparent rules and accompanied by an efficient liquidation process of insurer assets might be a solution for the organization of an orderly, industry-financed run-off of an insurer facing a default. Such a scheme could be designed to provide specific incentives within insurance markets and, hence, among other results, influence financial market stability. However, a potential disadvantage is the possibility that institutionalized run-off solutions can cause serious wealth transfers between different policyholder groups (see, e.g., Rymaszewski et al., 2010).

In the following, we do not explicitly discuss solutions other than insurance guaranty funds for dealing with insurer defaults. However, insurance guaranty funds are not, of course, the only possible solution to this problem. In addition to diverse forms of state-funded methods, such as bailouts or nationalization, there are private-sector solutions, including obligatory reinsurance of policyholder claims (organized in the form of an insurance pool or managed by several reinsurance companies), as well as capital-market-oriented methods based on, e.g., collateralized debt obligations or credit default swaps.

2.2 Main economic characteristics of insurance guaranty funds

Insurance guaranty funds are either private or state-sector-based institutions the main goal of which is to protect policyholders from a (partial) default on their contingent claims against insurance companies. Their organization varies depending on country of origin.[11] In general, if an insurer goes bankrupt, an insurance guaranty fund secures the interests of the policyholders either by continuation or by an immediate termination of their contracts. In the former case, insurer claims are either transferred to an external agent or wound up by the fund itself. In the latter case, insureds receive cash compensation.

Depending on the jurisdiction, different funding methods are employed to cover the costs incurred by an insurance guaranty fund, i.e., compensation payments to policyholders, transfer payments covering the difference between the market value of an insurer's assets and liabilities, costs connected with winding-up insurer claims, and administrative costs. These costs are primarily financed by the insurance industry via charges incurred either ex-ante or post-assessment based on the realized costs. However, there are often governmental contributions toward these expenses as well, for example, a tax deduction for incurred fees. This is particularly interesting in light of the fact that guaranty funds are often designed as defined benefit schemes with regard to the amount of realized claims covered, frequently with caps. In many cases, insurance guaranty fund coverage is restricted to specific classes of insurance contracts, e.g., it might exclude commercial customers on the basis that they are expected to be more knowledgeable than retail clients. Often, legally independent funds are used separately to cover life and non-life lines of business.

---

[11]For an overview of existing insurance guaranty funds within diverse jurisdictions, see, e.g., Brewer-III et al. (1997); Feldhaus and Kazenski (1998); Oxera (2007); Bernier and Mahfoudhi (2010).

3 Congruence with goals of insurance regulation

Consumer protection is often stated to be the ultimate goal of insurance regulation and supervision.[12] However, since policyholders as well as some of the other stakeholders (e.g., society, employees) are particularly interested in the efficiency of the insurance market, issues such as stability, lack of disruption, and competitiveness within the insurance industry can also be seen as important facets of regulation (see Meier, 1991; European Commission, 1999). In the following, we investigate whether and, if so, how well current solvency frameworks support the essential goals of insurance regulation. In addition, we analyze how differently designed insurance guaranty funds can facilitate or disrupt achievement of these goals.

3.1 Stability of the financial system

Safeguarding the stability of the banking sector is one of the widely recognized reasons for regulating this sector. Regulation is deemed necessary because the highly interrelatedness of the banking system (e.g., intense reciprocal borrowing and lending) has the potential to render the entire system unstable. For example, liquidity problems induced by a bank run on one market player can rapidly result in difficulties for other institutions within the banking industry, which, in turn, can significantly jeopardize the economy and lead to momentous political and social problems.

On the one hand, Harrington (2009), CEA (2010), the Geneva Association (2010), and Radice (2010) argue that interconnection among insurance companies is much weaker than that between banks due to their different business model, and hence insurance companies can be a source of systemic risk only if they engage in extensive quasi-banking activity. On the other hand, a significant amount of research in this

---

[12]See, e.g., European Commission (1999) and FOPI (2004) for the motivation behind the development of the Solvency II framework and the Swiss Solvency Test, respectively. IAIS (2009, p. 7) mentions consumer protection as the *ultimate objective* of group-wide solvency assessment and supervision.

field supports the idea that an insurer's insolvency could cause significant problems for the entire financial sector (see, e.g., Fenn and Cole, 1994; Angbazo and Narayanan, 1996; Polonchek and Miller, 1999). Allen and Carletti (2006), as well as Allen and Douglas (2006), show that strong relationships between the banking and insurance sectors can lead to contagion between the two and increase the risk of financial crises, despite the fact that insurance companies face lapse risk (the insurance company equivalent of a bank run) to a rather small degree. Bernier and Mahfoudhi (2010) analyze the possible contagion effects within the insurance industry induced by post-assessed insurance guaranty funds.

In the following, we analyze market stabilizing incentives implied by current insurance prudential regulation and supervision, and compare their impact with that of an insurance guaranty scheme.

Restrictions on insurer's risk-taking

The aim of prudential insurance regulation and supervision is to lower insurer default probabilities. In this way, a restricted likelihood of insolvency for every insurance company operating in the market allows limiting the probability of market-wide financial distress. Risk-based capital requirements, such as those currently in place or still being developed (for an overview see, e.g., Eling and Holzmüller, 2008; IAIS, 2009; Klein and Wang, 2009), which link mandatory solvency capital to the amount of risk borne by an insurer might support this goal. On the other hand, risk-inadequate capital requirements may lead to improper incentives for insurers and possibly increase their appetite for risk-taking. This can happen if shareholders interpret their position in the insurance company solely as a call option on company assets with its liabilities as a strike price.[13] In such a setting and given asymmetric information, management of a publicly traded insurer can increase the company's market value by raising the volatility of its assets (see, e.g, Merton, 1977;

---

[13]Due to the fact that a bankrupt insurer remains always insolvent, theoretically, shareholders of an insurance company may see their stake in the firm as a barrier option (see, e.g., Black and Cox, 1976). Since, under certain conditions, the value of such an option might decrease in underlying's volatility, the insurer's shareholders might have less of an incentive toward risk-taking compared to the valuation based on a plain (European) call.

Cummins, 1988).[14] If the policyholders are not perfectly informed about the firm's actual solvency level, they might not be able to demand lower insurance premiums for the increased risk.[15] This despite the decreased value of their wealth position due to the higher value of the default put option, they are short in favor of the company's shareholders (see, e.g., Doherty and Garven, 1986; Butsic, 1994). Due to equity capital costs for the additional solvency capital, which is intended to reduce the company's risk, risk-adequate capital requirements directly penalize the insurer for taking additional risks.

Analogous considerations can be derived with respect to the risk adequacy of an insurance guaranty fund's charging and payoff system, which, if properly designed, can provide incentives for insurers similar to those provided by risk-oriented solvency capital requirements (see Cummins, 1988). On the one hand, a truly risk-oriented insurance guaranty fund requires charges closely linked to a company's risk profile. On the other hand, in order to ensure a genuine risk orientation, fund premiums should be paid solely by the participating companies and disbursed in advance.[16] Both systems aim at avoiding a situation where insurers do not entirely internalize their potential insolvency costs, either by transferring the costs to an external agent (e.g., the government and, by extension, to taxpayers) or by going bankrupt. In case the fund-charging system is not in this sense risk-adjusted, insurers may still increase their market value by taking additional risks. Lee et al. (1997) call this the risk-subsidy hypothesis. Given a poor congruence between a company's risk profile and insurance guaranty fund's charges, if the outcomes of the insurer's actions are favorable, the firm's shareholders receive the entire profit resulting from taking additional risks (or at least a part of it). However, in the event the company faces default, the firm's owners enjoy the limited

---

[14]Lee and Smith (1999) suggest that a similar effect occurs if an insurer intentionally underreserves its claims.

[15]A similar effect might occur in a market with perfect information if the policyholders cannot renegotiate their contract in the event the insurer changes its risk profile. However, this issue may be less important in a multi-period context, where the insureds can account for insurer behavior in the following periods (see, e.g., Axelrod and Hamilton, 1981).

[16]For an exemplary payoff scheme ensuring the self-supporting character of an insurance guaranty fund, see Rymaszewski et al. (2010).

liability of their stake and any deficit of assets over company liabilities is absorbed by the guaranty fund and partially by the policyholders.[17]

In general, a self-supporting insurance guaranty fund, as one possible solution for an orderly run-off of an insurance company, may have an additional and positive impact on the stability of financial markets, irrespective of the applied premium principle. Since the existence of an insurance guaranty fund effectively financed by the insurance industry should in most cases lower the negative outcomes for policyholders in the event of insurer default, the incentive for the government to bail out an insolvent insurer is likely to decrease. If the management of an insurance company is aware of this situation, it may change its risk-taking behavior accordingly.

Influence on monitoring

Under certain conditions, the management of a company may increase its market value by taking additional risks only if the market is not free of frictions, in particular, if there is information asymmetry between market participants. Hence, reduction of friction with regard to the information known to different market players may be used to reduce the incentives for extensive risk-taking. This is why enhanced reporting requirements are an integral part of many modern solvency frameworks, which is this way aim at supporting market discipline (e.g., Solvency II and Swiss Solvency Test).

The impact of an insurance guaranty fund on the scale of information asymmetries in an insurance market depends, among others, on its charging and payoff structure. If policyholders will bear only a part of insolvency costs—and they are aware of this fact[18]—they might have weaker incentives to monitor their own insurers compared to the setting without an insurance guaranty fund. Hence, even within markets

---

[17]Hence, the impact of the risk-subsidy effect is likely to depend on the legal structure of the insurer and can be expected to be less significant for mutual insurance companies, where the equity and policyholder stakes are inseparable. In their empirical study, Lee et al. (1997) do not find any significant influence of the insurance guaranty fund introduction on the risk-taking behavior of mutuals.

[18]The European Commission (2010) claims that policyholders have a limited understanding of information about insurance guaranty funds.

characterized by perfect information and with self-supporting, but risk-inadequately designed insurance guaranty funds in place, single policy-holders are likely to have less incentive to avoid very risky protection providers. This is due to the fact that stakeholders (in particular, policyholders and/or shareholders) of less risky insurance companies pay, at least in part, for their hazardous choice. Such a decrease of monitoring incentives is even stronger if there is an external agent covering a part of the policyholder deficit in case it occurs. However, if the policyholders still bear at least a part of the insolvency costs, incentives to monitor their insurance company may weaken but never entirely disappear (see Cummins and Sommer, 1996).

Further monitoring effects may arise if the fund charges are not in-curred upfront, but post-assessed based on realized policyholder deficits. In such a setting, stakeholders of all insurance companies might be very interested in monitoring their rivals, since bankruptcy costs will be borne by the surviving members of the industry. Lee et al. (1997) call this the monitoring hypothesis. Nevertheless, the impact of this effect mainly depends on the level of market discipline within the market. Introduction of a post-assessed insurance guaranty fund may have a positive effect on the stability of the financial system due to additional monitoring incentives only if markets are relatively incapable of disciplining the insurers.[19]

Current situation

To the best of our knowledge, all extant insurance guaranty fund schemes charge premiums that are not (directly) linked to insurer risk.[20] None of

---

[19]For empirical evidence of market discipline with regard to the price of insurance see, e.g., Sommer (1996), Cummins and Danzon (1997), and Phillips et al. (1998). Fenn and Cole (1994) find evidence of market discipline with regard to the price of insurance company stock.

[20]See Oxera (2007, p. 34), for an extensive overview of the insurance guaranty fund schemes within the European Union. A discussion of U.S. regulation of the life insurance sector can be found in Brewer-III et al. (1997, p. 305). See Lee and Smith (1999) for a description of the U.S.-based property-liability guaranty fund system. Within the German life insurance guaranty fund scheme, charges depend on a company's ranking based on its financial capacity, defined as equity relative to solvency margin. Oxera (2007) classifies this approach as risk-based. However, in our opinion, a similar method of premium calculation cannot be viewed as plenary risk-oriented under the currently valid Solvency I regulatory framework.

the current schemes fulfill the previously defined postulates for a strict risk orientation. Even in cases when the premium charges are incurred upfront (as currently done in, e.g., France, Germany, and Denmark; see Oxera, 2007, p. 39), there are no examples of insurance guaranty funds charging risk-adequate premiums. In most cases, the premiums charged by a guaranty fund are calculated in relation to the premiums earned by the insurer or the amount of its technical reserves. The former approach may have a particularly adverse influence on the stability of insurance markets characterized by a high level of market discipline. Since in such case the less reliable insurers earn lower premiums than the solvent firms, the fund charge calculation is an exact opposite of the risk-adjusted approach (see Sommer, 1996). However, current discussion on the design of insurance guaranty schemes within the European Union allows us to surmise that the future harmonized approach within the European market is likely to be based on risk-adjusted, mainly ex-ante charged contributions (see European Commission, 2010).

Several empirical studies conducted in the United States confirm that insurance guaranty funds characterized by risk-inadequate charging and payoff schemes may increase the instability of the financial system. Lee et al. (1997) find a significant impact of the risk-subsidy effect within the property-liability insurance market. They do not find any significant influence of the monitoring effect. Similar results for guaranty funds effectively funded by taxpayers within the U.S. life insurance market are presented by Brewer-III et al. (1997). Downs and Sommer (1999), as well as Lee and Smith (1999), arrive at similar conclusions. In an earlier study, Munch and Smallwood (1980) cannot confirm a significant causal dependence between the introduction of an insurance guaranty fund and the likelihood of company default.

3.2 Avoidance of market failure

In an extreme case, information asymmetries between both parties to an insurance contract may lead to a situation when nearly all suppliers of

insurance have a strong incentive to exit the market.[21] Since the economic and social costs of a market failure concerning insurance supply could be substantial, it is important that prudential regulation and supervision appropriately addresses this issue. Due to the fact that current solvency frameworks formally follow the goal of constraining insurer default probabilities, insurance companies are to some extent restricted to offering a minimum quality level of products. Hence, the probability of a market failure induced by the problem of adverse selection decreases if policyholders believe in the effectiveness of governmental regulation. The threat of market failure resulting from adverse selection is also likely to lessen due to the enhanced reporting requirements associated with regulatory schemes.

The probability of a market failure induced by adverse selection may also decrease if the introduced insurance guaranty fund provides incentives for a restricted range of product quality on the insurance market and the insureds perceive the efficiency of this constraint. Additional incentives for risk reduction can be achieved by introducing insurance guaranty funds that charge risk-adjusted premiums. Depending on the premium principle used to derive the fund charge, it is possible to reduce insurer willingness to take additional risk. An additional market disciplining mechanism mitigating the impact of adverse selection can be based on publication of fund charges paid by the insurers.[22]

3.3   Consumer protection

The general idea that customers of financial service providers need some kind of protection is mainly motivated by the existence of information asymmetries between both parties signing an insurance contract. Never-

---

[21]Akerlof (1970) shows, in an example of a market for second-hand cars, that, given asymmetric information, it may be disadvantageous for a sound supplier to remain active in the market if customers do not perceive the above-average quality of its products. For issues connected with information asymmetries within insurance markets, see, e.g., Rothschild and Stiglitz (1976). An empirical investigation into asymmetric information within insurance markets was conducted by Chiappori and Salanié (2000).

[22]In fact, the problem of adverse selection can be alleviated even more if insurers were explicitly allowed to use information about premium payments as a signaling instrument (see, e.g., Spence, 1973).

theless, despite the fact that insurers are also affected by missing information (e.g., about the true behavior of a policyholder), the main focus of insurance safety nets is on policyholders, who are assumed to be the less knowledgeable partner to the contract. However, even if the insurance market was free of information asymmetries between insurers and their clients, possible claims of ex-ante unknown third parties could still occur and this is an issue that may need special consideration in terms of regulation.

In this context, the role of solvency frameworks in defending the interests of insurance company customers is based on reduction of information asymmetries between policyholders and insurers as well as limitation of risks borne by the insurers. The latter effect might result from the fact that within markets with asymmetric information, an increase of insurer risk does not necessarily lead to a decrease of the insurance premium, even despite the increased value of the insurer's default put option, which the policyholders are short in, but the value of which they cannot perfectly observe. Hence, restrictions on insurer risk positions as well as increased market discipline could allow relatively risk-averse insureds to avoid choosing an insurance company with a risk profile that is not congruent with their preferences.

Protection of policyholders as well as third-party claimants against the negative outcomes of an insurance company's insolvency is broadly acknowledged as the main goal of an insurance guaranty fund (see, e.g., Krogh, 1972; Oxera, 2007). The fund is intended to guard the insureds against risk resulting from pure randomness, which is an integral part of the insurance business. As a so-called last resort measure, an insurance guaranty fund at least partially covers the costs resulting from default events whose probability of occurrence can be reduced by prudential regulation and supervision only to a predefined small, yet still positive, level. Hence, policyholder collectives of insurers participating in an appropriately designed insurance guaranty fund theoretically should enjoy only unsystematic wealth transfers in the event their insurance company suffers a positive deficit of assets over claims.

In contrast, systematic wealth transfers between the policyholders of different insurance companies can occur if the fund payoff structure de-

pends on the joint default distribution of all insurers participating in the insurance guaranty scheme. In this situation, there is no way to guarantee that systematic cross-subsidization between the insureds of different insurance companies will not take place. Rymaszewski et al. (2010) show that in the real-world of incomplete markets and heterogeneous insurance companies, systematic wealth transfers between diverse insurers cannot be avoided. Hence, some of the policyholders might be disadvantaged by the introduction of an insurance guaranty fund if it does not significantly lower the extent of the agency problems described in Section 3.1. Even if pooling of risks within an insurance guaranty fund leads to positive diversification effects, allocating them back to the participating insurers is in general not possible in a nonarbitrary way. Therefore, we cannot rule out the possibility that some policyholder groups will be better off than the others under such a scheme.

Therefore, whether insurance guaranty funds in individual cases support the goal of consumer protection depends on the regulator's exact definition of this objective. If the major aim of an insurance guaranty fund is understood solely as securing the policyholders in extreme bad states of nature, every payoff structure financially supporting the insureds given insurer default and characterized by a positive (or neutral) influence on agency issues will meet the goal of consumer protection, despite the potential for cross-subsidization problems. However, in our opinion, analyzing the introduction of an insurance guaranty fund without taking into account wealth transfer issues that can impact the entire financial position of policyholders is insufficient, since it fails to consider a possibly substantial part of hedging costs borne by (some of) those insured.

3.4   Competition

Lack of competition in the insurance sector may have an undesirable effect not only on the wealth position of policyholders, but also on the economy as a whole. Joskow (1973) argues that deficiency of competition within insurance markets may lead to inefficient capital allocation and restricted insurance supply. Furthermore, a decreased level of competi-

tion is likely to lower insurers' incentives to monitor their rivals, which may result in a restricted level of market discipline.

On the one hand, it can be argued that insurance regulation decreases the level of competition, since mandatory safety levels limit the products and services insurance companies can offer. The effect may be similar if solvency requirements include binding market access restrictions, for instance, with regard to minimum capital requirements.[23]

On the other hand, there are several reasons to believe that the introduction of solvency frameworks, which are based on recent developments in the field of risk management and regulation, might have a positive influence on competition in insurance markets. Enhanced reporting standards may have a desirable impact on the disciplining capabilities of financial markets and hence an auxiliary influence on the level of competition. Furthermore, converging regulatory rules between different jurisdictions—for example, within the future Solvency II framework—might lower market access barriers as well as provide a level playing field between competing insurers (see European Commission, 2010). Hence, harmonization of regulatory systems is likely to strengthen competition.[24]

In general, insurance guaranty funds, in their role as institutions securing the orderly run-off of insurance companies, are expected to ease market exit of inefficient and uncompetitive insurers. This is chiefly due to the fact that introduction of an insurance guaranty fund reduces the social and political costs of insurer insolvency. However, if the premiums of an insurance guaranty fund are charged ex-ante, new market entry may be restricted, depending on the premium principle applied to derive these premiums, for instance, if there are floors for the fund charges implied by insurance guaranty fund's fixed costs.

Ligon and Thistle (2007) discuss the potential influence of insurance guaranty funds effectively underwritten by the government on the or-

---

[23]This matter is at least partially allowed for within the Solvency II framework by the principle of proportionality, which is aimed at ensuring that the regulatory framework reflects the nature, scale, and complexity of the insurance business—in particular, that is appropriately addresses the issue of small and less risky insurance companies (see European Parliament, 2009, p. 9).

[24]Note that if international financial markets are strongly interconnected the enhanced competition may have a negative impact on their stability.

ganizational structure within the insurance industry. They prove that under certain conditions it might be rational for an insurance provider to offer different insurance products within separate monolines rather than conduct itself as a multiline company. However, the influence of this effect on the level of competition is difficult to assess without a loss of generality.

The issues discussed concerning an insurance guaranty fund's risk orientation show that the interactions between the goals of solvency frameworks and the shape of insurance guaranty funds are not limited to risk-based solvency capital and the fund's premium calculation or to the resulting incentives for preventing excessive risk-taking. Moreover, the resulting monitoring incentives are to some extent directly linked to efforts aimed at enhancing market discipline with the goal of increasing market stability. In our opinion, both aspects justify integrating insurance guaranty funds into solvency frameworks.

4   Holistic view of solvency issues

4.1   Possible interactions with the supervisory process

In general, current solvency frameworks aim at providing a holistic approach to solvency measurement. Among others, they are expected to account for the intra-group effects and interactions between different business lines (see, e.g., FOPI, 2004, 2006; European Parliament, 2009). Nevertheless, they incorporate strategies for an orderly run-off (in the event of insurance company bankruptcy) to only a restricted degree. They do not explicitly aim at addressing the applied institutionalized run-off solutions, as insurance guaranty funds. This even despite the complementarity of both concepts.

For instance, both Solvency II and the Swiss Solvency Test account for a potential company run-off only in regard to the market-consistent estimation of insurer assets and liabilities. Such an approach, allowing for a market value margin over the best estimate of insurer liabilities, is intended to ensure that a company's claims are covered with an appro-

priately high probability, either by the company itself or by a third party willing to take over the insurance company's portfolio under the current market conditions. However, policyholders are still potentially exposed to a total or partial default of their contingent claims. Using such a market-oriented insurance regulation as the only measure of consumer protection will make insureds' financial position default-free only if there is no possibility of a so-called jump-to-default and given continuous supervision of insurer activity.

An extended view of an insurance company's solvency can fill the gap not covered by the current solvency frameworks, which aim at substantially restricting but not entirely neutralizing insurer default probabilities. We argue that, depending on their design, insurance guaranty funds may either support or disrupt achievement of the goals of prudential regulation and supervision. The effectiveness of this strategy—indicated by several empirical studies[25]—depends not only on the specific guaranty fund payoff structure but also on the level of market discipline within the insurance sector. From the perspective of a supervisory authority, integration of current insurance guaranty funds into the regulatory process could allow a more comprehensive view of the incentives for insurers and policyholders that are inherent in each element of the insurance safety net. In the case of inconsistencies, an extended approach could facilitate their detection and provide additional instruments to address them.

In particular, an analysis of the interrelations between guaranty funds' charging and payoff systems and risk-based capital requirements may provide essential insights into the influence of insurance guaranty funds on insurer default probabilities. This specific dependence is particularly important since the introduction of an insurance guaranty fund in general leads to an increase in the likelihood of insurer insolvency. This phenomenon is due to the fact that insurance companies are charged premiums that are used to cover deficits faced by policyholders in case of a default.

Moreover, in our opinion and for similar reasons, in case the local regulatory system does not include an institution comparable to an insurance guaranty fund, but there is a different institutionalized strategy

---

[25]See Section 3.1.

for coping with insurer run-off, the regulator should always be aware of and systematically analyze the potential interactions between the solvency framework and the existing resolution approach.

Explicit consideration of insurance guaranty funds and/or other run-off solutions is particularly important when different jurisdictions are attempting to develop a harmonized system of insurance regulation, e.g., Solvency II within the European Union. Different resolution approaches within each member state may impair the desired compatibility of a comprehensive regulatory system, even despite the congruency of the solvency capital requirements.[26]

4.2   Interactions with regard to premium calculation

In general, introduction of an insurance guaranty fund on incomplete financial markets (almost) certainly causes wealth transfers between the participating collectives of policyholders (see Rymaszewski et al., 2010). This effect is due to the general impossibility of allocating diversification benefits resulting from the pooling process to the participating insurance companies. In this context, we face a problem similar to the capital allocation issue intensively discussed in the literature the last couple of years (see, e.g., Phillips et al., 1998; Myers and Read, 2001; Sherris, 2006; Gründl and Schmeiser, 2007; Ibragimov et al., 2010; Zanjani, 2010) that cannot be solved by a premium principle based, e.g., on risk of a single insurer, i.e., oriented on an insurer's loss distribution (e.g., value at risk). The degree of cross-subsidization depends on the scheme employed to calculate fund charges.[27]   Nevertheless, if the establishment of an insurance guaranty fund is likely to lead to a decrease of agency costs, its introduction may still be beneficial for policyholders. This is because

---

[26]Harmonization of insurance guaranty schemes within the European Union is postulated by de Larosière et al. (2009) and the European Commission (2010), as well as recommended within the Solvency II Directive (see European Parliament, 2009).

[27]A nonarbitrary solution for premium calculation is possible only under the assumption of perfect financial markets (see Cummins, 1988; Duan and Yu, 2005). However, Rymaszewski et al. (2010) show that under the contingent claim approach, a self-supporting insurance guaranty fund does not result in any benefit to policyholders.

agency problems can be at least partially restricted by a risk-based fund premium calculation (see Section 3.1).

Derivation of the guaranty fund's premium based on the standard or internal models used for calculation of the risk-based capital requirements is one possible solution for adjusting the charges paid by insurers to the risk they are bearing. One possible advantage of a this approach is the potential economies of scale, which would lower the direct costs of establishing the insurance guaranty fund. However, a simple adoption of the models used to calculate solvency capital requirements might in some cases be problematic because the solvency frameworks aim at reducing the default probability of a single insurer and hence account mainly for insurer-specific risk.

In the context of an insurance guaranty fund, the important advantage of which is the pooling of default risks, an insurer's risk contribution to the underlying portfolio might be the relevant factor for premium calculation. This will be the case if the payments an insurance guaranty fund will make to an insurance company depend not only on its own realized deficit, but also on other insurers' shortfalls, which is always the case when an insurance guaranty fund is organized as a self-supporting institution, implying that available funds are restricted. Thus, if the solvency frameworks do not appropriately account for the interdependencies between assets and claims of all insurers participating in an insurance guaranty fund—particularly if those interrelations are significant—simply adopting the models employed within the solvency frameworks may insufficiently account for pooling issues.

On the other hand, allowing for diverse methods of premium calculation might result in diverse views of a company's risk profile and, hence, reduce the model risk within the regulatory process.[28] If the model applied for the purpose of fund premium calculation explicitly allows for interdependencies between assets and claims of the participating insurers and adjusts for an insurer's risk contribution to the insurance guaranty fund's portfolio, the insights from fund charge estimates can

---

[28]See Rymaszewski et al. (2010) for an alternative, utility-based method for guaranty fund premium calculation. For different risk-oriented solutions within the banking sector see, e.g., European Commission (2008).

be of use within the entire supervisory process.[29]   An enhanced analysis of the interrelations between the portfolios of different insurers may have a positive effect on financial market stability. Since the current regulatory frameworks explicitly aim at constituting an integral part of insurers' risk-management systems, implementation of complementary risk-assessment models may result in further advantages from an insurer's perspective.

5   Conclusion

The recent financial crisis has shown that default events of systematically relevant financial institutions can pose a very real threat to the economy. Insurance guaranty funds, both as currently in existence and as possible developments, are one possible and widely discussed solution to resolve failing service providers within the insurance industry. Since the introduction of an insurance guaranty fund is connected with diverse incentives for market participants that interact with the goals of prudential regulation and supervision, we find it appropriate and important to analyze the effects of their establishment and existence on the functionality of the current solvency frameworks. In this way, a permanent analysis of issues related to currently in-force insurance guaranty funds into systems of solvency supervision can be applied to support the congruence of actions within both schemes. In this paper, we discuss how insurance guaranty funds can either support or disrupt achievement of the goals of insurance regulation. In case a different institutionalized resolution approach is employed within the insurance industry, the effects of its implementation should also be taken into account. Such an integrated approach will facilitate a holistic treatment of issues relevant to insurance regulation, which is the major objective of modern solvency frameworks.

---

[29]As a practicable approach, under the assumption of well-diversified asset portfolios correlations between insurers' assets can be derived from adequate market data. However, such a method can hardly be applied for insurers' claims.

References

Akerlof, G. A. (1970). The Market for Lemons: Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, 84(3):488–500.

Allen, F. and Carletti, E. (2006). Credit Risk Transfer and Contagion. *Journal of Monetary Economics*, 53(1):89–111.

Allen, F. and Douglas, G. (2006). *Systemic Risk and Regulation, in: The Risk of Financial Institutions by M. Carey and R. M. Stulz*. Chicago.

Angbazo, L. A. and Narayanan, R. (1996). Catastrophic Shocks in the Property-Liability Insurance Industry: Evidence on Regulatory and Contagion Effects. *Journal of Risk and Insurance*, 63(4):619–637.

Axelrod, R. and Hamilton, W. D. (1981). The Evolution of Cooperation. *Science*, 211(4489):1390–1396.

Bernier, G. and Mahfoudhi, R. M. (2010). On the Economics of Postassessments in Insurance Guaranty Funds: A Stakeholders' Perspective. *Journal of Risk and Insurance*, (forthcoming).

Black, F. and Cox, J. C. (1976). Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. *Journal of Finance*, 31(2):351–367.

Brewer-III, E., Mondschean, T. S., and Strahan, P. E. (1997). The Role of Monitoring in Reducing the Moral Hazard Problem Associated with Government Guarantees: Evidence from the Life Insurance Industry. *Journal of Risk and Insurance*, 64(2):301–322.

Butsic, R. P. (1994). Solvency Measurement for Property-Liability Risk-Based Capital Applications. *Journal of Risk and Insurance*, 61(4):656–690.

Chiappori, P.-A. and Salanié, B. (2000). Testing for Asymmetric Information in Insurance Markets. *Journal of Political Economy*, 108(1):56–78.

Comité Européen des Assurances (CEA) (2010). Insurance: A Unique Sector. Why Insurers Differ from Banks. *URL http://www.cea.eu.*

Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS) (2009a). *CEOIPS Input to the European Commission work on Insurance Guarantee Schemes. URL http://www.ceiops.eu.*

Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS) (2009b). *Lessons Learned from the Crisis: Solvency II and Beyond. URL http://www.ceiops.eu.*

Cummins, J. D. (1988). Risk-Based Premiums for Insurance Guaranty Funds. *Journal of Finance*, 43(4):823–839.

Cummins, J. D. and Danzon, P. M. (1997). Price, Financial Quality, and Capital Flows in Insurance Markets. *Journal of Financial Intermediation*, 6(1):3–38.

Cummins, J. D. and Sommer, D. W. (1996). Capital and Risk in Property-Liability Insurance Markets. *Journal of Banking and Finance*, 20(6):1069–1092.

de Larosière, J., Balcerowicz, L., Issing, O., Masera, R., Mc Carthy, C., Nyberg, L., Pérez, J., and Ruding, O. (2009). Report of the High-Level Group on Financial Supervision in the EU. *URL http://ec.europa.eu.*

Doherty, N. A. and Garven, J. R. (1986). Price Regulation in Property-Liability Insurance: A Contingent-Claims Approach. *Journal of Finance*, 41(5):1031–1050.

Downs, D. H. and Sommer, D. W. (1999). Monitoring, Ownership, and Risk-Taking: The Impact of Guaranty Funds. *Journal of Risk and Insurance*, 66(3):477–497.

Duan, J.-C. and Yu, M.-T. (2005). Fair Insurance Guaranty Premia in the Presence of Risk-Based Capital Regulations, Stochastic Interest Rate and Catastrophe Risk. *Journal of Banking and Finance*, 29(10):2435–2454.

Eling, M. and Holzmüller, I. (2008). An Overview and Comparison of Risk-Based Capital Standards. *Journal of Insurance Regulation*, 26(4):31–60.

European Commission (1999). The Review of Overall Financial Position of an Insurance Undertaking. *URL http://ec.europa.eu.*

European Commission (2008). Risk-Based Contributions in EU Deposit Guarantee Schemes: Current Practices. *URL http://ec.europa.eu.*

European Commission (2010). White Paper on Insurance Guaranty Schemes. *URL http://ec.europa.eu.*

European Parliament (2009). Directive of the European Parliament and of the Council on the Taking-up and Pursuit of the Business of Insurance and Reinsurance (Solvency II). *URL http://register.consilium.europa.eu.*

Federal Office of Private Insurance (FOPI) (2004). White Paper of the Swiss Solvency Test. *URL http://www.finma.ch.*

Federal Office of Private Insurance (FOPI) (2006). Technical Document on the Swiss Solvency Test. *URL http://www.finma.ch.*

Feldhaus, W. R. and Kazenski, P. M. (1998). Risk-Based Guaranty Fund Assessments. *Journal of Insurance Regulation*, 17(1):42–63.

Fenn, G. W. and Cole, R. A. (1994). Announcements of Asset-Quality Problems and Contagion Effects in the Life Insurance Industry. *Journal of Financial Economics*, 35:181–198.

Geneva Association (2010). Systemic Risk in Insurance An Analysis of Insurance and Financial Stability. *URL http://www.genevaassociation.org.*

Gründl, H. and Schmeiser, H. (2007). Capital Allocation for Insurance Companies - What Good Is It? *Journal of Risk and Insurance*, 74(2):301–317.

Harrington, S. E. (2009). The Financial Crisis, Systemic Risk, and the Future of Insurance Regulation. *URL http://www.naic.org.*

Ibragimov, R., Jaffee, D., and Walden, J. (2010). Pricing and Capital Allocation for Multiline Insurance Firms. *Journal of Risk and Insurance*, 77(3):551–578.

Institute of International Finance (IIF) (2010). Global Approach to Resolving Failing Financial Firms: An Industry Perspective. *URL http://www.iif.com.*

International Association of Insurance Supervisors (IAIS) (2009). Issues Paper on Group-Wide Solvency Assessment and Supervision. *URL http://www.iaisweb.org.*

Joskow, P. L. (1973). Cartels, Competition and Regulation in the Property-Liability Insurance Industry. *Bell Journal of Economics and Management Science*, 4(2):375–427.

Klein, R. W. and Wang, S. (2009). Catastrophe Risk Financing in the United States and the European Union: A Comparative Analysis of Alternative Regulatory Approaches. *Journal of Risk and Insurance*, 76(3):607–637.

Krogh, H. C. (1972). Insurer Post-Insolvency Guaranty Funds. *Journal of Risk and Insurance*, 39(3):431–450.

Lee, S.-J., Mayers, D., and Smith, C. W. (1997). Guaranty Funds and Risk-Taking Evidence from the Insurance Industry. *Journal of Financial Economics*, 44(1):3–24.

Lee, S.-J. and Smith, M. L. (1999). Property-Casualty Insurance Guaranty Funds and Insurer Vulnerability to Misfortune. *Journal of Banking and Finance*, 23(9):1437–1456.

Ligon, J. A. and Thistle, P. D. (2007). The Organizational Structure of Insurance Companies: The Role of Heterogeneous Risks and Guaranty Funds. *Journal of Risk and Insurance*, 74(4):851–862.

Meier, K. J. (1991). The Politics of Insurance Regulation. *Journal of Risk and Insurance*, 58(4):700–713.

Merton, R. C. (1977). An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees. An Application of Modern Option Pricing Theory. *Journal of Banking and Finance*, 1(1):3–11.

Munch, P. and Smallwood, D. E. (1980). Solvency Regulation in the Property-Liability Insurance Industry: Empirical Evidence. *Bell Journal of Economics*, 11(1):261–279.

Myers, S. C. and Read, J. A. J. (2001). Capital Allocation for Insurance Companies. *Journal of Risk and Insurance*, 68(4):545–580.

Oxera (2007). Insurance Guarantee Schemes in the EU: Comparative Analysis of Existing Schemes, Analysis of Problems and Evaluation of Options. *URL http://ec.europa.eu.*

Phillips, R. D., Cummins, J. D., and Allen, F. (1998). Financial Pricing of Insurance in the Multiple-Line Insurance Company. *Journal of Risk and Insurance*, 65(4):597–636.

Polonchek, J. and Miller, R. K. (1999). Contagion Effects in the Insurance Industry. *Journal of Risk and Insurance*, 66(3):459–475.

Radice, M. P. (2010). Systemische Risiken im Versicherungssektor? Betrachtungen mit Fokus auf die schweizerische Assekuranz. *FINMA (Swiss Financial Market Supervisory Authority) Working Paper*.

Rothschild, M. and Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *Quarterly Journal of Economics*, 90(4):629–649.

Rymaszewski, P., Schmeiser, H., and Wagner, J. (2010). Under What Conditions Is an Insurance Guaranty Fund Beneficial for Policyholders? *Journal of Risk and Insurance*, (forthcoming).

Sherris, M. (2006). Solvency, Capital Allocation, and Fair Rate of Return in Insurance. *Journal of Risk and Insurance*, 73(1):71–96.

Sommer, D. W. (1996). The Impact of Firm Risk on Property-Liability Insurance Prices. *Journal of Risk and Insurance*, 63(4):597–636.

Spence, M. (1973). Job Market Signaling. *Quarterly Journal of Economics*, 87(3):355–374.

Zanjani, G. (2010). An Economic Approach to Capital Allocation. *Journal of Risk and Insurance*, 77(3):523–549.

Part III

A Traffic Light Approach to Solvency Measurement of Swiss Occupational Pension Funds

Abstract

In this paper, we combine a stochastic pension fund model with a traffic light approach to solvency measurement of occupational pension funds in Switzerland. Assuming normally distributed asset returns, a closed-form solution can be derived. Despite its simplicity, we believe the model comprises the essential risk sources needed in supervisory practice. Due to its ease of calibration, it is well suited for a regulatory application in the fragmented Swiss market, keeping costs of solvency testing at a minimum. We calibrate and implement the model for a small sample of ten Swiss pension funds in order to illustrate its application and the derivation of traffic light signals. In addition, a sensitivity analysis is conducted to identify important drivers of the shortfall probabilities for the traffic light conditions. Although our analysis concentrates solely on Switzerland, the approach could also be applied to similar pension systems.[30]

1 Introduction

The recent crisis in the global financial markets hit not only banks and insurers but also the pension fund industry. The resulting underfunding of a large number of pension schemes triggered a discussion about the rearrangement of prudential regulation and supervision for occupational pension funds in Switzerland. The obligatory character of occupational pension plans for the majority of Swiss employees, the large volume invested through them (according to the Swiss Federal Statistical Office, in 2008 the aggregated book value of assets was approximately equal to the Swiss GDP), as well as significant social costs linked to potential insolvencies demonstrate that this debate is not exclusively political. Instead, a solvency test for pension funds is of considerable relevance to employees, employers, and pensioners. Supervision and regulation of pensions in Switzerland is currently conducted at the cantonal level.[31] The main task of these regulators is to ensure that the pension funds comply with the legal requirements. Besides, they receive the annual reports and the report of an independent occupational pension expert, whose duty is the valuation of a fund's technical liabilities. The expert also examines whether or not a fund is able to cover its liabilities. Comprehensive solvency regulation, however, is not present for occupational pension funds in Switzerland, although banks and insurance companies have to adhere to Basel II and the Swiss Solvency Test (SST), respectively.[32] This paper is an attempt to address this issue. We suggest an efficient solvency test for occupational pension funds, providing condensed information for the stakeholder groups instead of prescribing regulatory capital. For this purpose, we adopt a model for pension funds under stochastic rates of return and combine it with a traffic light approach, allowing an efficient comparison of the risks inherent in different funds as well as a comprehensible communication of results of the solvency test. This signal based approach can be used not only to support the supervisory process, but also to facilitate an increased level of market discipline. However, more transparency within the pension fund market can intensify the latter

---

[31]See also Gugler (2005).
[32]See, e.g., Eling et al. (2008).

only if insureds are both capable of interpreting the received signal and of taking actions as a consequence of the information they receive.

The literature with regard to stochastic pension fund modeling has been strongly influenced by the work of O'Brien (1986, 1987) and Dufresne (1988, 1989, 1990). While the former proposes a continuous-time approach, the early model of Dufresne operates in a discrete-time environment. This original discrete-time model has subsequently been applied and extended in several papers. Haberman (1992) introduces time delays with regard to additional contributions for unfunded liabilities and, in a consecutive paper, Haberman (1993a) examines the effects of changes in the valuation frequency for the pension fund's assets and liabilities. Furthermore, Zimbidis and Haberman (1993) use the model with time delays to derive expectations and variances for fund and contribution level distributions. In two additional publications, Haberman (1993b, 1994) drops the assumption of independent and identically distributed (iid) asset returns in favor of a first-order autoregressive process and utilizes the model to compare different pension funding methods. In contrast to the discrete-time focus of the majority of papers, Haberman and Sung (1994) present and employ a continuous-time model to simultaneously minimize an objective function for contribution rate and solvency risk. Haberman (1997) reverts to a discrete-time version with iid asset returns and analyses funding approaches to control contribution rate risk of defined benefit pension funds. Cairns (1995) extends previous work by turning to the fund's asset allocation strategy as a means of controlling funding level variability. In a later paper, Cairns (1996) presents a pension fund model in continuous-time with continuous adjustments to the asset allocation and contribution rate. A similar model but with stochastic benefit outgo is discussed in Cairns (2000), while Cairns and Parker (1997) apply a discrete-time approach and compare the effect of a change from iid to autoregressive returns on the variability of funding level and contribution rates. Finally, Bédard and Dufresne (2001) show that the dependence of successive rates of return can have a considerable effect on the model results in a multi-period setting.

The model we present is based on the discrete-time framework which has been frequently employed in the literature in order to analyze issues,

such as contribution rate risk or behavior of the funding level over time. However, it has not been previously considered in the context of solvency measurement. We adapt the model as to capture the particularities associated with the occupational pension fund system in Switzerland and demonstrate that its simplicity and ease of calibration are advantages for an application as a regulatory standard model in this fragmented market. The model enables us to estimate shortfall probabilities which are then funneled into a traffic light approach in order to send a signal to stakeholder groups, which carries condensed information about a fund's financial strength and is straightforward to interpret, even for less sophisticated claim holders. Although the scope of our analysis is limited to Switzerland, both the model itself and the insights from its application can be transferred to similar pension systems.

The remainder of this paper is organized as follows. Section 2 sets the stage with a brief introduction to the particularities of Switzerland's occupational pension fund system. The stochastic pension fund model which forms the basis for the proposed solvency test is presented in Section 3, while Section 4 explains the traffic light approach to solvency measurement. Section 5 comprises an exemplary calibration of the model and illustrates its application by computing shortfall probabilities and deriving the traffic lights for a small sample of ten Swiss pension funds. A sensitivity analysis is then conducted in Section 6 in order to identify important drivers of the shortfall probabilities for the traffic light conditions. Section 7 focuses on the supervisory review process. Some additional considerations concerning a potential implementation in Switzerland are provided in Section 8. Finally, in Section 9, we conclude.

2   The particularities of the Swiss occupational pension fund system

The Swiss pension system comprises three pillars.[33] The first pillar is earnings-related and embedded in the public social security scheme; the second pillar relates to the mandatory occupational pension fund sys-

---

[33]See also Brombacher Steiner (1999).

tem;[34] the third pillar consists of additional benefits that need to be accumulated individually.[35] Our paper focuses on the second pillar which is governed by the Swiss occupational pension law (abbreviated in German: BVG and BVV2). At the heart of the second pillar, which, apart from retirement pensions, also provides widow(er) and invalidity pensions, are the occupational pension funds (in German: *Vorsorgeeinrichtungen*).

The vast majority of occupational pension funds in Switzerland takes the legal form of private trusts, where the employees have a right of parity participation in the administrative council (Art. 55 BVG).[36] Apart from single-employer pension funds, which are run exclusively for the employees of one company, the specific structure of the Swiss economy with many small and medium-sized businesses necessitates so-called multi-employer pension funds (in German: *Sammeleinrichtungen*).[37] This relieves small businesses from the burden of setting up their own pension fund, because they can join a multi-employer fund which bundles the occupational pension schemes of several independent firms.[38] A change of pension fund can only be completed by the employer with the agreement of the majority of employees. The second pillar is covered by a guarantee fund (in German: *Sicherheitsfonds BVG*), with the main purpose of subsidizing schemes with an adverse age structure and guaranteeing the obligatory payments of defaulted funds.

Compulsory pension contributions are based on the so-called coordinated salary[39] (in German: *koordinierter Lohn*) of the employee and the employer has to bear at least half of each installment (Art. 8 and Art. 66 BVG).[40] These regular payments are credited to a pension account (in German: *Altersguthaben*) and at least compounded with an obligatory minimum rate of return (currently 2 percent). Once the insured reaches

---

[34]Participation in the occupational pension system is mandatory for all employees of age 18 or older who earn a minimum annual salary of 20'520 CHF (Art. 7 BVG).

[35]See, e.g., OECD (2009).

[36]Pension funds of the federation, cantons, and municipalities are institutions under public law.

[37]See Swiss Federal Statistical Office (2009).

[38]Employers are obliged to either establish a firm-specific pension fund or to join multi-employer fund with the consent of their employees (Art. 11 BVG).

[39]Currently the coordinated salary is the part of an employee's annual income between 23'940 and 82'080 CHF.

[40]Voluntary payments in excess of the compulsory contributions are possible.

the retirement age of 65 for men or 64 for women (Art. 13 BVG), the obligatory pension annuity is calculated by multiplying the annuity conversion rate, which is currently 6.8 percent, with the final balance of the pension account (Art. 14 BVG).[41] The Swiss Federal Council (in German: *Der Schweizerische Bundesrat*) determines both the minimum interest rate and the conversion rate at two- and ten-year intervals, respectively.[42] In general, Swiss occupational pension funds can be set up either as defined contribution or as defined benefit plans.

One important aspect of the occupational pension fund system in Switzerland is that funds are legally allowed to temporarily operate with a deficit of assets relative to liabilities (Art. 65c BVG). Such an underfunding of liabilities is indicated by the coverage ratio, i.e., the proportion of the market value of assets over technical liabilities, falling below 100 percent (Art. 44 BVV2). However, the tolerance of a temporary underfunding is strictly linked to the condition that a pension fund continues its ongoing obligatory pension payments and takes action to restore full coverage within an adequate time horizon. In addition, the pension fund has to promptly inform the regulator, the employer, the employees, and the pensioners about the magnitude and causes of the asset shortage as well as countermeasures that have been initiated. The pension fund has to eliminate the deficit itself as the guarantee fund can merely intervene in case of insolvency (Art. 65d BVG). For this purpose, the fund can raise additional contributions from the employer and the employees to rectify the deficit. If and only if all other actions prove insufficient, the fund is allowed to go below the obligatory minimum interest rate by up to 0.5 percent for no longer than 5 years.

---

[41] The pension funds can decide to provide annuities over and above the obligatory level.

[42] When determining the minimum interest rate, the Swiss Federal Council takes into account the recent development of the returns of marketable investments, with a particular focus on government bonds, corporate bonds, equities, and real estate (Art. 15 BVG). Mortality improvements are accounted for through an adjustment of the conversion rate.

3   The model framework

We suggest building a solvency framework for occupational pension funds around underfunding probabilities, at the center of which we need a stochastic pension fund model. While advanced internal models could be allowed for the supervision of pension funds with sophisticated risk management know-how and processes, the requirements of a regulatory standard model suggest an approach that concentrates on the most essential risk drivers. The complexity of such a standard model should be kept within adequate limits so that the introduction of the solvency regulation does not cause an unjustifiably large increase in personnel and infrastructure cost, especially for smaller occupational pension funds. Apart from that, a properly developed simple model is capable of capturing the main determinants of pension fund activity.[43] Moreover, the feasibility of the whole concept depends on sufficient data being available for calibration. This is more likely to be the case for an approach which entirely relies on observable variables such as accounting figures. With these considerations in mind, we decide in favor of a discrete-time model that ensures universal applicability, cost-efficient implementation, and straightforward calibration.[44] The model we present is based on the work of Cairns and Parker (1997). In the following, we adapt it to the specific characteristics of occupational pension funds in Switzerland and combine it with a traffic light approach for the assessment of shortfall probabilities in order to construct a pragmatic solvency test.

Consider a one-period evaluation horizon and continuous compounding. If the occupational pension fund is assumed to have a stationary membership and all cash flows are exchanged at the beginning of the period, the asset process of the pension fund can be described as follows:

$$\tilde{A}_1 = \exp(\tilde{r}_1)\left(A_0 + C_0 - B_0\right), \tag{34}$$

where

- $\tilde{A}_1$: stochastic market value of the assets in $t = 1$,

---

[43]See, e.g., Cairns and Parker (1997).

[44]Equivalent formulations in continuous time can be found in the literature. See, e.g., Cairns (1996).

- $\tilde{r}_1$: stochastic return on the assets between $t = 0$ and $t = 1$,

- A_0: assets in $t = 0$,

- C_0: contributions for the period between $t = 0$ and $t = 1$,

- B_0: benefit payments for the period between $t = 0$ and $t = 1$.

The aggregated asset return consists of normally distributed returns for each asset class in the fund's portfolio:

$$\tilde{r}_1 = \sum_{i=1}^{n} w_i \tilde{r}_i, \tag{35}$$

with $\tilde{r}_i \sim N(\mu_i, \sigma_i), \forall i \in \{1, \ldots, n\}$, where

- w_i: portfolio weight for asset class i,

- $\tilde{r}_i$: return of asset class i between $t = 0$ and $t = 1$,

- n: number of asset classes in the portfolio.

Note that for some asset classes, the assumption of normally distributed returns is merely an approximation (see, e.g., Officer, 1972). However, it will enable us to derive a closed-form solution, which we consider a very valuable aspect of a standard solvency model.

Since occupational pension funds commonly have a large pool of employees and pensioners, their liabilities are fairly well diversified and consequently relatively stable. Hence, the crucial source of risk is constituted by a pension fund's asset allocation and a deterministic approach for the liabilities is justifiable. In general, the value of the liabilities in $t = 0$ is the present value of the stochastic future cash flows from the fund to those insured. These cash outflows are estimated actuarially, taking into account the age structure and mortality profile of the fund as well as the targeted rate of return, which needs to be equal to or greater than the obligatory minimum. Although an actuarial technical interest rate is commonly used in this context, it is more adequate to apply the current interest rate term structure. Therefore, we incorporate the market value of the liabilities into our model and define the corresponding yield as the valuation rate of interest i_v. Issues resulting from

a potential misestimation of the pension liabilities will be addressed in Section 6. If the liabilities are assumed to be continuously compounded at i_v, we have the following relationship:

$$L_1 = \exp(i_\mathrm{v})\left(L_0 + RC_0 - B_0\right), \tag{36}$$

where

- L_1: market value of the liabilities in $t = 1$,

- i_v: interest rate for the valuation of the liabilities,

- L_0: market value of the liabilities in $t = 0$,

- RC_0: regular contributions for the period between $t = 0$ and $t = 1$.

The assumptions of normally distributed asset returns and deterministic liabilities could be relaxed by resorting to numerical solutions, e.g., via a Monte-Carlo simulation framework. In that case, many different distributional assumptions and dependency structures could be incorporated. Similarly, a numerical solution would allow the introduction of a longer time horizon and intermediate time steps or a continuous-time framework.[45]

The contributions between $t = 0$ and $t = 1$, C_0, consist of two distinct elements:

$$C_0 = RC_0 + AC_0, \tag{37}$$

with

$$AC_0 = \alpha \max\left[L_0 - A_0, 0\right],\ {}^{[46]} \tag{38}$$

where

- AC_0: additional contributions between $t = 0$ and $t = 1$ for the recovery of a deficit in $t = 0$,

- α: fraction of the deficit in $t = 0$, which will be covered between $t = 0$ and $t = 1$.

---

[45]See Bühlmann (1996) for the calculation of ruin probabilities in a similar context, applying a multi-dimensional geometric Brownian motion for the asset dynamics.

[46]Note that a negative value of $L_0 - A_0$ implies a positive fluctuation reserve or a positive amount of uncommitted funds.

At the beginning of each period due additional contributions are determined based on the current deficit of assets relative to liabilities. Hence, AC_0 also accounts for additional contributions remaining from prior deficits. Consider, e.g., a deficit in $t = -1$. The resulting additional contribution AC_{-1} will increase the value of the assets in $t = 0$, A_0, which then forms the basis for the calculation of AC_0. Therefore, if AC_{-1} together with the development of the assets and liabilities between $t = -1$ and $t = 0$ was sufficient to eliminate the deficit, there will be no need for further additional contributions and AC_0 will be zero.

Additional contributions are subject to two restrictions. First of all,

$$\alpha \geq \alpha^{\min} = \frac{1}{\theta}, \tag{39}$$

which implies

$$AC_0^{\min} = \alpha^{\min} \max\left[L_0 - A_0, 0\right], \tag{40}$$

where

- θ: maximum number of years for the elimination of the deficit (set by the regulator),

- $\alpha^{\min}$: minimum fraction of the deficit in $t = 0$, which needs to be covered between $t = 0$ and $t = 1$.

The restriction in Inequality (39) implies that deficits have to be eliminated within an adequate time horizon (see Section 2), which will be set by the regulator through the choice of θ.[47] As a consequence, additional contributions in the period under consideration must not fall below a certain minimum, $AC_0^{\min}$, as defined in (40), since otherwise the elimination of the deficit would take too long. Intuitively, the fewer years available for the fund to restore its coverage ratio at least to unity, the higher the scheduled additional contributions for each year have to be.

Furthermore,

$$A_0 \geq A_0^{\min} = \beta L_0, \tag{41}$$

---

[47] In Switzerland this time period is not legally defined. In current practice, however, a five-year span seems to have emerged as convention.

which implies

$$AC_0^{\max} = \max\left[L_0 - \beta L_0, 0\right] = \max\left[(1 - \beta) L_0, 0\right], \qquad (42)$$

with

- β: lowest acceptable coverage ratio $A_0^{\min}/L_0$ (set by the regulator).

Excessive additional contributions are disputable, since they transfer the investment risk from the pensioners to the employees and employers. Accordingly, Inequality (41) accounts for the fact that deficits can only be healed by means of additional contributions up to a certain amount. For instance, consider a case in which the value of assets falls to zero. Clearly, a restructuring of the pension fund is not feasible in this case. Hence, in order to protect those insured from having to pay an unacceptably large amount of additional contributions into a pension fund in major distress, we define a lower limit for the assets $A_0^{\min}$ (a fixed percentage of the pension fund's liabilities), which puts a cap on additional contributions per period. This amount, termed $AC_0^{\max}$, is defined in Equation (42) and based on β, i.e., the lowest coverage ratio acceptable by the regulator. Usually, $0 \leq \beta \leq 1$, and the lower β, the higher the maximum amount of additional contributions that can be charged by the pension fund in any given period.[48] If the assets fall below the threshold $A_0^{\min}$, the fund will ceteris paribus be unable to rectify the deficit within a single period. In addition, if $AC_0^{\min}$ exceeds $AC_0^{\max}$, which is theoretically possible, particularly for high values of α (low values of θ) and β, the pension fund faces an existential funding problem, since it would be required to collect a larger amount of additional contributions than it is actually allowed to.

Under the above assumptions, the assets at the end of the evaluation period are log-normally distributed with:

$$\mathrm{E}\left[\tilde{A}_1\right] = \mathrm{E}\left[\exp\left(\tilde{r}_1\right)\left(A_0 + C_0 - B_0\right)\right], \qquad (43)$$

---

[48]Note that theoretically β could also exceed one. In such a case, additional contributions would be ruled out by our model framework. To see this, refer to Equation (42).

which is equivalent to

$$E\left[\tilde{A}_1\right] = \exp\left(E\left[\tilde{r}_1\right] + \frac{\text{var}\left[\tilde{r}_1\right]}{2}\right)(A_0 + C_0 - B_0), \qquad (44)$$

and

$$\text{var}\left[\tilde{A}_1\right] = \text{var}\left[\exp\left(\tilde{r}_1\right)(A_0 + C_0 - B_0)\right], \qquad (45)$$

$$\text{var}\left[\tilde{A}_1\right] = (A_0 + C_0 - B_0)^2 \exp\left(2E\left[\tilde{r}_1\right] + \text{var}\left[\tilde{r}_1\right]\right)$$
$$\times \left\{\exp\left(\text{var}\left[\tilde{r}_1\right]\right) - 1\right\}. \qquad (46)$$

Hence, in order to calculate the first two central moments, which entirely determine the asset distribution in $t = 1$ under the assumption of normally distributed returns, estimates for $E[\tilde{r}_1]$ and $\text{var}[\tilde{r}_1]$ are required. Using Equation (35), mean and variance for the returns of the aggregated asset portfolio can be calculated in the following manner:

$$E\left[\tilde{r}_1\right] = E\left[\sum_{i=1}^{n} w_i \tilde{r}_i\right] = \sum_{i=1}^{n} w_i E\left[\tilde{r}_i\right] \qquad (47)$$

and

$$\text{var}\left[\tilde{r}_1\right] = \sigma_{\tilde{r}_1}^2 = \text{var}\left[\sum_{i=1}^{n} w_i \tilde{r}_i\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \rho_{\tilde{r}_i,\tilde{r}_j} \sigma_{\tilde{r}_i} \sigma_{\tilde{r}_j}, \qquad (48)$$

where $\rho_{\tilde{r}_i,\tilde{r}_j}$ denotes the correlation coefficient between the returns of asset class i and j.

4 The traffic light approach

There are several ways to implement a solvency framework. The regulator could, for example, prescribe that each pension fund needs to set aside regulatory capital based on the outcome of a solvency test. Such an approach is common in the banking and insurance industries. In the case of occupational pension funds, however, which do not posses equity capital, this is rather problematic as the funds would need to build up

reserves from contributions. On the other hand, pension funds have the possibility to demand additional contributions from employers and employees, which is similar to authorized equity capital of corporations that can be drawn in predefined cases. The risk of not being able to raise this capital when needed is negligible, since it resembles a tax levied by the government. Thus, we believe that the prescription of regulatory capital is not the most suitable approach for pension funds. Instead, our proposal is oriented towards early alert. For solvency measurement purposes, we combine the previously introduced pension fund model with a concept akin to a value-at-risk framework and funnel the results into a so-called traffic light approach.

As discussed in the previous section, the model delivers a deterministic value for the liabilities at the end of the analyzed period. Using this value as a threshold in conjunction with the asset distribution, we can derive shortfall probabilities for the pension fund under consideration. These probabilities could be compared to reference probabilities ψ, e.g., default rates from rating agency data, in order to generate a signal for the regulator and the insured. Various categorizations for such a signal are conceivable. As a straightforward solution, we suggest the following:

- *green:*
$$\Pr\left(\tilde{A}_1 \leq L_1\right) \leq \psi, \tag{49}$$

- *yellow:*
$$\Pr\left(\tilde{A}_1 + AC_1^{\max} \leq L_1\right) \leq \psi, \tag{50}$$

- *red:*
$$\Pr\left(\tilde{A}_1 + AC_1^{\max} \leq L_1\right) > \psi, \tag{51}$$

where $AC_1^{\max}$ denotes the maximum amount of additional contributions which can be charged by the pension fund in $t = 1$. $AC_1^{\max}$ is deterministic, since it is based on the value of the liabilities in $t = 1$.[49] If the probability of underfunded liabilities in $t = 1$ is smaller than the preset reference probability ψ, the pension fund is assigned a green light.

---

[49] Alternatively, different reference probabilities could be chosen for all three conditions.

In addition, if the assets and the maximum additional contributions in $t = 1$ are only insufficient to cover the liabilities with a probability lower than ψ, the light is yellow. In this case the fund is able to suppress the probability of underfunded liabilities in $t = 1$ below the reference probability through its option of charging additional contributions. Finally, the red light comes up if the probability that the assets plus $AC_1^{\max}$ fall short of the liabilities exceeds ψ.

5 Implementation and calibration

5.1 Input data

A major advantage of the model is its low implementation cost due to the use of readily available data. In this section, we illustrate that even for smaller occupational pension funds with less sophisticated risk management techniques in place, it should be straightforward to calibrate and implement the model. For the purpose of calibration, we rely on accounting figures from the funds' annual reports. In practice, pension funds and regulators would be able to use superior data from their management accounting and financial planning units or databases. As such internal data is not available to us, we deem annual reports to be the most adequate and reliable source. Note that this approach is subject to certain limitations. As defined in Section 3, a solvency test for pension funds should theoretically be based on market values of assets and liabilities. This is in line with the latest developments in risk management practice as well as supervisory frameworks for the insurance sector (Solvency II and the SST). Yet, figures derived from annual reports are, in general, not consistent with market values. In particular, reported pension liabilities are commonly valued using a technical interest rate, i.e., an actuarial rate instead of the prevailing term structure. Consequently, we substitute i_v in Equation (36) with the technical interest rate i_tec applied by each fund. The appropriate level of the technical interest rate is currently controversially discussed in Switzerland. More specifically, some pension funds seem to be reluctant to reduce it as to reflect the low interest rate environment which resulted from the financial crisis

2007/2008, implying an even greater discrepancy between market and book values of the liabilities. Nonetheless, we believe that the following numerical illustration of the proposed solvency framework offers useful insights.

Tables 3 to 6 show the parameter values we collected for ten occupational pension funds in Switzerland.[50] It is important to note that coverage ratios, assets, liabilities, technical interest rates, and portfolio weights for 2007 and 2008 have been extracted from annual reports of the same year. In contrast to that, 2008 and 2009 figures have been used for contributions and benefits of 2007 and 2008, respectively, assuming that the funds can reliably forecast their values at the beginning of the period.[51] Since the market values of the funds' assets could not be directly obtained from their annual reports, they have been estimated by multiplying the reported coverage ratios (A_0/L_0) with the book values of the liabilities. Furthermore, we decided to conduct the analysis based on seven broad asset classes. Tables 5 and 6 contain the portfolio weights each fund assigns to the these asset classes.[52] Note that the asset allocation of some pension funds is fairly concentrated. The implications of this issue together with the effect of insufficient diversification within the subportfolio for each asset class will be addressed in Section 6.2. While the market environment in 2007 was still relatively stable, the 2008 figures reflect the major turbulences caused by the global financial crisis. Thus, this dataset enables us to apply the solvency test in two different economic settings. In addition, we included single as well as multi-employer funds to further increase the informative value of our calculations.

It could be discussed whether the parameters for the asset class return distributions should be preset by the regulator, thereby reducing discretionary competencies to a minimum. However, taking into account the ease of estimation and regulatory verification of these parameter values,

---

[50]The funds were made anonymous.

[51]Our model treats these magnitudes as deterministic (see Section 3), thus assuming that they can be perfectly forecasted. If this is unlikely to be true, the model could be revised by incorporating benefits and contributions as stochastic variables.

[52]If deemed necessary, the solvency test could be based on a more detailed categorization of the asset side.

mil. CHF	Fund 1	Fund 2	Fund 3	Fund 4	Fund 5	Fund 6	Fund 7	Fund 8	Fund 9	Fund 10
A_0/L_0	111%	103%	104%	130%	115%	104%	110%	116%	107%	102%
A_0	11'591.90	3'136.98	247.65	14'585.39	16'996.47	1'173.10	1'498.20	6'582.35	34'703.20	13'589.43
L_0	10'415.01	3'048.57	237.67	11'176.54	14'792.40	1'130.16	1'360.16	5'688.67	32'524.09	13'309.92
C_0	1'175.49	260.93	50.42	631.57	750.00	179.18	61.21	1'200.24	1'360.62	891.41
RC_0	919.13	249.60	50.42	631.57	750.00	79.45	61.21	1'200.24	1'183.44	891.41
AC_0	256.36	12.33	0.00	0.00	0.00	99.72	0.00	0.00	177.18	0.00
B_0	787.37	201.88	15.58	837.47	797.20	87.57	90.87	711.13	2'359.56	952.76
i_{tec}	3.75%	4.00%	3.00%	4.00%	4.00%	3.00%	3.50%	3.50%	4.00%	4.00%

Table 3: Input parameters for the sample funds in 2007

mil. CHF	Fund 1	Fund 2	Fund 3	Fund 4	Fund 5	Fund 6	Fund 7	Fund 8	Fund 9	Fund 10
A_0/L_0	100%	88%	85%	105%	97%	88%	86%	98%	96%	88%
A_0	10'934.19	2'866.88	230.28	11'712.17	14'822.70	1'109.74	1'174.94	6'207.89	30'290.15	11'704.35
L_0	10'923.27	3'257.81	270.60	11'186.41	15'265.40	1'261.07	1'373.40	6'317.17	31'611.52	13'285.30
C_0	951.45	275.08	47.18	615.18	720.80	168.87	53.28	1'000.23	1'357.44	983.90
RC_0	937.62	261.17	47.18	615.18	720.80	106.50	48.56	1'000.23	1'357.44	983.90
AC_0	13.83	13.91	0.00	0.00	0.00	62.37	4.72	0.00	0.00	0.00
B_0	816.38	213.94	13.39	759.69	787.90	95.05	108.06	782.97	2'113.04	882.45
i_{tec}	3.75%	4.00%	3.50%	4.00%	4.00%	3.00%	3.50%	3.50%	3.50%	3.50%

Table 4: Input parameters for the sample funds in 2008

%	Fund 1	Fund 2	Fund 3	Fund 4	Fund 5	Fund 6	Fund 7	Fund 8	Fund 9	Fund 10
Bonds (intl.)	9.7	14.4	33.5	13.2	18.3	23.0	8.9	12.3	23.5	15.3
Bonds (CH)	24.1	23.9	8.1	24.1	15.6	32.0	8.9	34.6	46.3	30.5
Stocks (intl.)	15.8	12.7	21.9	30.6	17.4	16.6	34.3	20.1	13.2	9.2
Stocks (CH)	4.4	19.0	6.2	7.8	11.2	13.6	5.5	15.9	7.3	18.3
Real Estate	11.4	15.4	6.5	9.0	23.8	2.5	25.0	12.7	5.4	10.3
Alternatives	9.7	4.0	10.3	12.5	4.3	6.9	7.7	2.2	0.0	8.1
Cash	25.0	10.7	13.4	2.9	9.4	5.4	9.6	2.2	4.3	8.4

Table 5: Asset allocations of the sample funds in 2007

%	Fund 1	Fund 2	Fund 3	Fund 4	Fund 5	Fund 6	Fund 7	Fund 8	Fund 9	Fund 10
Bonds (intl.)	9.2	15.8	15.0	22.4	5.2	22.7	6.4	13.5	19.4	18.5
Bonds (CH)	33.4	28.0	29.8	31.8	31.7	29.5	10.0	37.3	53.1	37.0
Stocks (intl.)	8.5	11.9	18.4	15.9	11.0	14.4	25.1	13.2	12.6	6.7
Stocks (CH)	2.4	13.4	5.1	3.7	9.2	12.5	5.1	11.7	6.9	13.5
Real Estate	10.8	16.6	7.8	9.1	28.1	7.4	30.5	14.8	6.1	10.5
Alternatives	17.1	2.4	7.7	10.2	4.5	9.2	7.5	4.7	0.0	8.3
Cash	18.6	11.9	16.4	6.9	10.2	4.4	15.4	4.7	2.0	5.6

Table 6: Asset allocations of the sample funds in 2008

we suggest they should be determined by the pension funds themselves. Therefore, means, volatilities, and pair-wise correlations for the return distributions of the seven asset classes have been estimated from capital market time series data. To this end, we have chosen broad indices as representatives for each asset class.[53] The S&P U.S. Treasury Bond Index and the SBI Swiss Government Bond Index have been selected as proxies for the international and Swiss government bond markets, respectively. International equities are represented by the MSCI World, while the Swiss Market Index (SMI) is employed for the Swiss equity market. Real estate returns are provided through the Rued Blass Swiss REIT Index and the HFRI Fund Weighted Composite Index serves as a broad measure for the alternative investments universe. Finally, the Swiss 3M Money Market Index is used as an indicator for the development of cash holdings. Distribution moments as well as a correlation matrix based on monthly returns for these indices from January 1, 1997 to December 31, 2007 are exhibited in Tables 7 and 8. Based on the the simplifying assumption that the pension funds can perfectly hedge exchange rate fluctuations at a negligible cost, we have not converted the time series of the three U.S. Dollar denominated indices into Swiss Francs. Since hedging foreign currency investments against exchange rate risk is very common[54] in the asset management industry and the trading costs for the necessary foreign exchange (FX) instruments such as futures and options are relatively small,[55] we believe this to be an acceptable approach for our purpose. The effect of imperfect FX hedging will be considered in Section 6.6.

---

[53]Wherever available, total return indices have been used to account for coupons and dividends.

[54]As an example, consider an investment in a foreign currency denominated government bond. If left completely unhedged, this would be an outright speculation on the exchange rate, as the returns in the investor's home currency will be dominated by exchange rate movements, implying that the asset does not exhibit the typical characteristics of a government bond.

[55]Flat fees for FX futures trades at the Chicago Mercantile Exchange (CME), the largest regulated FX marketplace worldwide, can be as low as 0.11 USD, depending on membership and volume. For more information see `http://www.cmegroup.com`.

	S&P US Treasury Bond Index	SBI Swiss Gov. Bond Index	MSCI World	SMI Index	Rued Blass Swiss REIT Index	HFRI Fund Weighted Composite	Swiss 3M Money Mkt. Index
Currency denomination	USD	CHF	USD	CHF	CHF	USD	CHF
Mean return (%)	5.83	3.63	7.02	8.50	4.98	10.32	1.60
Volatility (%)	3.71	3.51	14.04	17.22	7.67	7.19	0.29

Table 7: Annualized means and standard deviations for the seven asset classes

	S&P US Treasury Bond Index	SBI Swiss Gov. Bond Index	MSCI World	SMI Index	Rued Blass Swiss REIT Index	HFRI Fund Weighted Composite	Swiss 3M Money Mkt. Index
S&P US Treasury Bond Index	1.00						
SBI Swiss Gov. Bond Index	0.57	1.00					
MSCI World	-0.27	-0.22	1.00				
SMI Index	-0.30	-0.17	0.74	1.00			
Rued Blass Swiss REIT Index	0.00	0.21	0.27	0.26	1.00		
HFRI Fund Weighted Composite	-0.19	-0.16	0.77	0.50	0.22	1.00	
Swiss 3M Money Mkt. Index	0.21	0.12	-0.21	-0.14	-0.07	-0.14	1.00

Table 8: Correlation matrix for the seven asset classes

5.2   Results

In order to be able to interpret the shortfall probabilities with the traffic light approach presented in Section 4, we need to determine reference probabilities. A straightforward approach is to refer to historic default rate data as commonly collected and maintained by the large rating agencies. Consequently, we suggest constructing probability intervals for rating categories based on default rate experience. The regulator could then set a minimum target rating for occupational pension funds, which is linked to the threshold probability.

In the following we use global corporate cumulative default rates from 1981 to 2008 provided by Standard & Poor's and establish intervals for the one-year default probabilities as shown in Table 9.[56] As a reasonable minimum target rating for pension funds, we propose the lowest investment grade rating category: BBB. As discussed in Section 2, participation in the occupational pension fund system in Switzerland is not voluntary. In addition, the volume invested through contributions of employers and employees is significant. As a result, occupational pension funds bear much responsibility for an individual's retirement provisions. Against this background, their financial strength should be demanded to be investment grade. Otherwise the uncertainty for those insured would be considerable, while they are not free to entrust their money with other financial institutions of their choice. Moreover, from the perspective of regulators and financial market participants, it would be very difficult to argue why pension funds should be allowed a notably lower credit quality than other financial institutions such as banks or insurance companies.

Having determined the reference probability ψ to be 0.99 percent (lower bound of BBB), we can now run the model calculations and interpret the results. For each fund, the probabilities for the traffic light conditions in 2007 and 2008 as well as the associated test outcomes (pass/fail) are presented in Tables 10 and 11, respectively. The calculations for the yellow condition have been conducted based on a β of 0.95 and 0.90, i.e.,

---

[56]See Standard & Poor's (2009). While the use of specific default rates for the investment industry in general or the pension fund market segment in particular would be preferable, we need to rely on the rather high-level data available to us. Nevertheless, in case of an introduction in practice, it would be advisable for the regulator to cooperate with rating agencies in order to access a more precise database.

%	AAA	AA	A	BBB	BB	B
lower bound	0.00	0.03	0.08	0.24	0.99	4.51
upper bound	0.03	0.08	0.24	0.99	4.51	25.67

Table 9: One-year default probabilities for different rating classes, see Standard & Poor's (2009)

we limited the maximum additional contributions per period to 5 and 10 percent of the liabilities.[57] First, we observe that four out of ten funds fail the green condition in 2007, although none has underfunded liabilities at the outset (the lowest coverage ratio among the sample funds in 2007 was 102 percent, see Table 3). When inspecting the coverage ratio these four funds actually reported in 2008 (see Table 4), we find that all of them in fact suffer from underfunded liabilities, ranging from 2 to an alarming 15 percent. A fund with a 15 percent deficit of assets relative to liabilities is in a serious state, since, even for the lower β of 0.90, it cannot be restructured through additional contributions in a single period. Taking into account that the current convention in Switzerland is a maximum of 5 years to eliminate the deficit, the fund needs additional contributions of at least 3 percent per year. These are already close to the 5 percent upper limit which we applied in the analyses for the yellow condition, underscoring the severity of this situation. Hence, a failure of the green condition is only acceptable in exceptional cases and should instantly trigger heightened attention from the supervisor as well as those insured. In this context it should also be emphasized, that a need for refinements to the traffic light approach is not automatically constituted by the fact that Fund 3 ends up with a deficit in excess of $AC_1^{\max}$ in 2008, although it was assigned a yellow light in the previous period. The proposed solvency test is exclusively based on probabilities. Therefore, by all means, a fund can pass one or both traffic light conditions and still end up with substantial unfunded liability at the end of the period. Being assigned a green or yellow light only means that the probability for

---

[57]Recall the definition of β from Equation (42). As discussed, it is ultimately up to the regulator to set a value for β.

%	Fund 1	Fund 2	Fund 3	Fund 4	Fund 5	Fund 6	Fund 7	Fund 8	Fund 9	Fund 10
green condition	0.01	18.36	7.29	0.00	0.10	0.07	3.28	0.24	0.70	21.42
	pass	**fail**	**fail**	pass	pass	pass	**fail**	pass	pass	**fail**
yellow condition	0.00	3.02	0.47	0.00	0.00	0.00	0.45	0.01	0.00	3.06
$\beta = 0.95$	pass	**fail**	pass	pass	pass	pass	pass	pass	pass	**fail**
yellow condition	0.00	0.18	0.01	0.00	0.00	0.00	0.03	0.00	0.00	0.13
$\beta = 0.90$	pass	pass	pass	pass	pass	pass	pass	pass	pass	pass

Table 10: Probabilities and test outcomes for the traffic light conditions in 2007

%	Fund 1	Fund 2	Fund 3	Fund 4	Fund 5	Fund 6	Fund 7	Fund 8	Fund 9	Fund 10
green condition	29.46 **fail**	99.44 **fail**	99.96 **fail**	4.36 **fail**	66.93 **fail**	83.99 **fail**	99.32 **fail**	48.31 **fail**	81.83 **fail**	99.71 **fail**
yellow condition $\beta = 0.95$	1.23 **fail**	91.07 **fail**	97.74 **fail**	0.09 pass	23.06 **fail**	43.52 **fail**	94.09 **fail**	10.94 **fail**	26.29 **fail**	92.20 **fail**
yellow condition $\beta = 0.90$	0.00 pass	53.69 **fail**	72.41 **fail**	0.00 pass	2.41 **fail**	8.34 **fail**	72.87 **fail**	0.65 pass	1.19 **fail**	50.28 **fail**

Table 11: Probabilities and test outcomes for the traffic light conditions in 2008

the respective event is sufficiently low. However, if similar discrepancies are detected for a large number of funds in the context of a comprehensive quantitative impact study prior to the introduction of the solvency test, its overall configuration and calibration should be reconsidered.

The second point we learn from Table 10 is that a β of 0.95 is more than enough to compress the probabilities for the yellow condition to very low levels for almost all funds. Evidently, this effect is even stronger for $\beta = 0.90$. While the probabilities are virtually zero for the financially sounder pension funds, even those which did not conform to the green condition seem to be able to comply with the yellow condition without difficulties. This illustrates an important point, which had already been mentioned in Section 4: the option to demand additional contributions implies that a large part of the pension funds' investment risk is ultimately borne by employees and employers. Consequently, in case of an actual introduction of the proposed approach in practice, the supervisory authority should carefully determine the upper limit on additional contributions.

A further insight we gain from Table 3 is that merely comparing the coverage ratios, as currently done in supervisory practice in Switzerland, is generally insufficient to capture the risk profile of pension funds. To see this, compare Fund 3 and Fund 6. Both are characterized by an equal coverage ratio of 104 percent in 2007 (see Table 5). However, only Fund 6 passes the green condition of our proposed solvency test (see Table 10). The reason is simple: a comparison of the coverage ratio does not take into account differences in asset allocation and the option to charge additional contributions.

Finally, when examining the results of the solvency test for 2008 in Table 11, we notice that the financial crisis has strongly influenced the condition of the pension funds in our sample. Since all funds except Fund 1 and Fund 4 already exhibit underfundings at the beginning of the period, their probabilities for both the green and yellow condition have increased considerably. As a result, not a single pension fund is able to pass the green condition and only one (Fund 4) passes the yellow condition based on a β of 0.95. Although the analyzed sample is rather small, this illustrates that the financial crisis has left the Swiss pension fund sector in a dramatic situation.

6   Sensitivity analysis

In this section, we explore the main drivers of the shortfall probabilities for the traffic light approach. These are relevant for the regulator in various ways, including the political discussion about the state of the Swiss occupational pension fund sector, the supervisory determination of model variables, and the decision about measures in case a pension fund fails the green or yellow condition of the solvency test. We base the analysis on a standard (representative) pension fund, the input data for which can be found in Table 12.[58] This data is mainly based on 2007 average figures from the Swisscanto (2008) pension fund survey, comprising 265 occupational pension funds in Switzerland, and has been complemented and cross-checked with annual report data from our sample.[59] The standard pension fund under consideration is financially sound at the beginning of the period with a coverage ratio of 110 percent and a fairly balanced asset allocation.

Input parameters		Asset Allocation	
A_0/L_0	110%	Bonds (intl.)	13%
A_0	11'000	Bonds (CH)	27%
L_0	10'000	Stocks (intl.)	18%
C_0	1'000	Stocks (CH)	10%
RC_0	1'000	Real Estate	15%
AC_0	-	Alternatives	7%
B_0	750	Cash	10%
i_{tec}	4%		

Table 12: Parameters for a representative pension fund

---

[58] Unless noted otherwise, β has been set to 0.95.

[59] The Swisscanto series of surveys is published on an annual basis and contains representative data with regard to the structure, performance, capitalization and portfolio allocation of the participating funds.

6.1 Portfolio weight allocated to equities

The first sensitivity we examine is related to the proportion of equities in the pension fund's portfolio. In this context we proceed as follows: from the original asset allocation in Table 12, we calculate the weight of each asset class with regard to the remaining part of the portfolio if stocks (international and Swiss) are excluded. As an example, consider the category alternative investments: aside from stocks, the remaining asset classes together make up 72 percent of the portfolio, 7 percent of which are alternative investments. Consequently, alternative investments are assigned a "residual" weight of $7/72 = 9.7$ percent for the analysis. In the same fashion, we get 18.1 percent for international bonds, 37.5 percent for Swiss bonds, 20.8 percent for real estate, 9.7 percent for alternatives, and 13.9 percent for cash. We then successively calculate the shortfall probabilities associated with the traffic light conditions for an increasing portfolio weight of stocks, beginning with zero and ending with the legal limit of 50 percent. In every case, the percentage is equally shared between Swiss and international equities. For each allocation, the remainder of the portfolio is distributed among the other asset classes according to the previously calculated residual weights. Figure 8 shows the results.

As one would expect, the shortfall probabilities generally increase in the portfolio weight assigned to equities. The probabilities associated with the pension fund's original portfolio composition as shown in Table 12 are represented through a point and a triangle on the curves at the 0.28 position,[60] while the threshold probability of 0.99 percent has been indicated by the dotted horizontal line. In its current state this average pension fund evidently passes the green condition with ease. On the one hand, we observe that the increase of the probability curve for the green condition is quite strong, revealing a critical portfolio weight for equities of 0.34, i.e., well below the legal limit of 0.5. On the other hand, Figure 8 reveals that the fund would in no case fail the yellow condition, even for the highest possible allocation to stocks. Hence, as already suspected in

---

[60]Note that these points are in fact slightly off the curve since at the position 0.28 the curve has been calculated with 0.14 allocated to international and 0.14 allocated to Swiss stocks, while the original asset allocation shows 0.18 and 0.1, respectively.
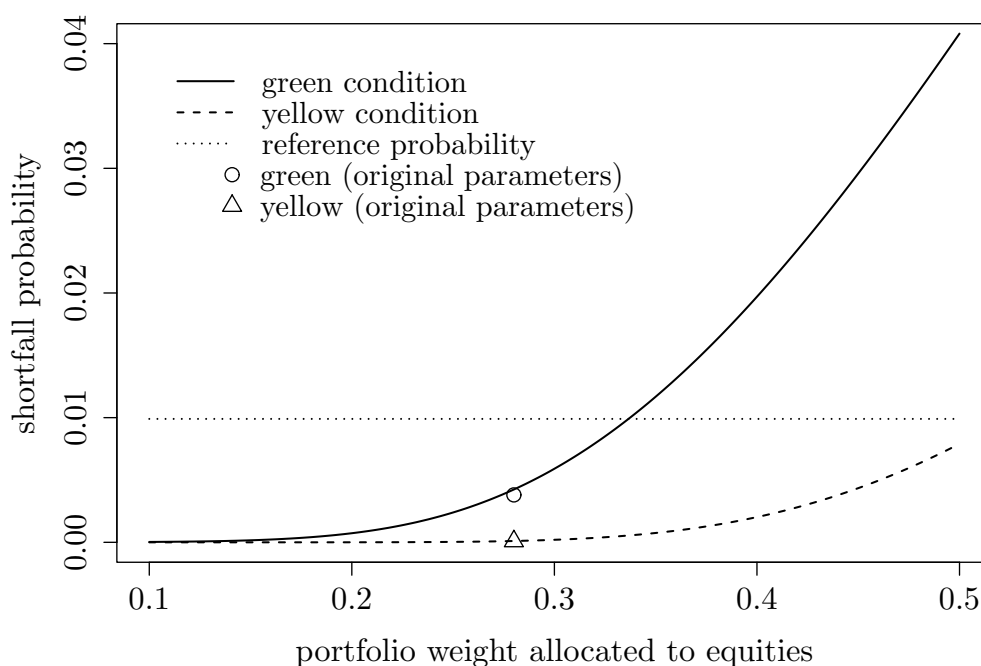
Figure 8: Sensitivity analysis: asset allocation

the previous section, allowing additional contributions up to 5 percent of the liabilities within a single year bears a considerable potential to suppress the shortfall probabilities. These results have an important implication for the regulator. One of the prevalent regulatory actions in case of a failure of the green traffic light condition should be an in-depth analysis of the portfolio composition of the respective pension fund with a focus on the more volatile asset classes such as equities. This could be followed by a dialog between the fund and the regulator to agree on an optimization of the portfolio to lower the probability of failing the green condition while still retaining reasonable return potential.

6.2   Concentration in asset class subportfolios

In Section 5.1 we calibrated the model based on indices (well-diversified portfolios), representing various asset classes. This approach implicitly assumes that pension funds adequately diversify their investments within the subportfolio of each asset class. In practice, a basic degree of diversi-

fication should be ensured, since pension funds have to obey mandatory limits on their asset positions. The equity portfolio, e.g., cannot account for more than 50 percent of a fund's total assets. In addition, within this equity portfolio, the maximum investment per individual stock (domestic or international) is currently limited to 5 percent of the total assets. Yet, with its calibration relying on indices, the solvency test might not be well suited for an application to pension funds which hold insufficiently diversified subportfolios. Thus, we briefly illustrate the impact of concentration issues within subportfolios, using domestic equity holdings as an example. For this purpose, we form an equally weighted portfolio (naïve diversification), consisting of an increasing number of stocks which are drawn from the constituents of the SMI Index. First, the portfolio only contains a single stock. Additional stocks are then successively added in random order until the portfolio contains a total of eight stocks.[61] For each step, we recalculate mean and volatility of the domestic equity portfolio (Table 7, column 4) as well as the correlations with the other asset classes (Table 8, line 4 and column 4) and recalibrated the model accordingly. The resulting shortfall probabilities for the solvency test are illustrated in Figure 9, together with the original case based on the complete SMI (20 stocks). As expected, the shortfall probabilities tend to decrease with a rising number of equally weighted stock holdings in the portfolio, i.e., with a decreasing asset concentration.[62] More specifically, for the relatively small number of six equities in the portfolio, the shortfall probabilities are already fairly close to those of the original case with the SMI as domestic equity portfolio. Therefore, only very high degrees of asset concentration in the subportfolios should result in a notable distortion of the proposed solvency test. However, if a pension fund naïvely diversifies its holdings over at least half a dozen stocks, the use of indices for calibration purposes seems to be a valid approach. To further underscore this, note that the standard pension fund underlying

---

[61]The final portfolio consists of the following equities, mentioned in the sequence in which they have been added: Credit Suisse Group, Adecco SA, Roche Holding AG, Holcim Ltd., SGS SA, Nestle SA, Swatch Group AG, and Swiss Re.

[62]Note that while we observe a general decrease in the shortfall probability for a growing number of stocks in the portfolio, it can sometimes slightly increase when a new stock is added, depending on the order of inclusion. This effect occurs due to changes on the overall correlation structure in the portfolio.
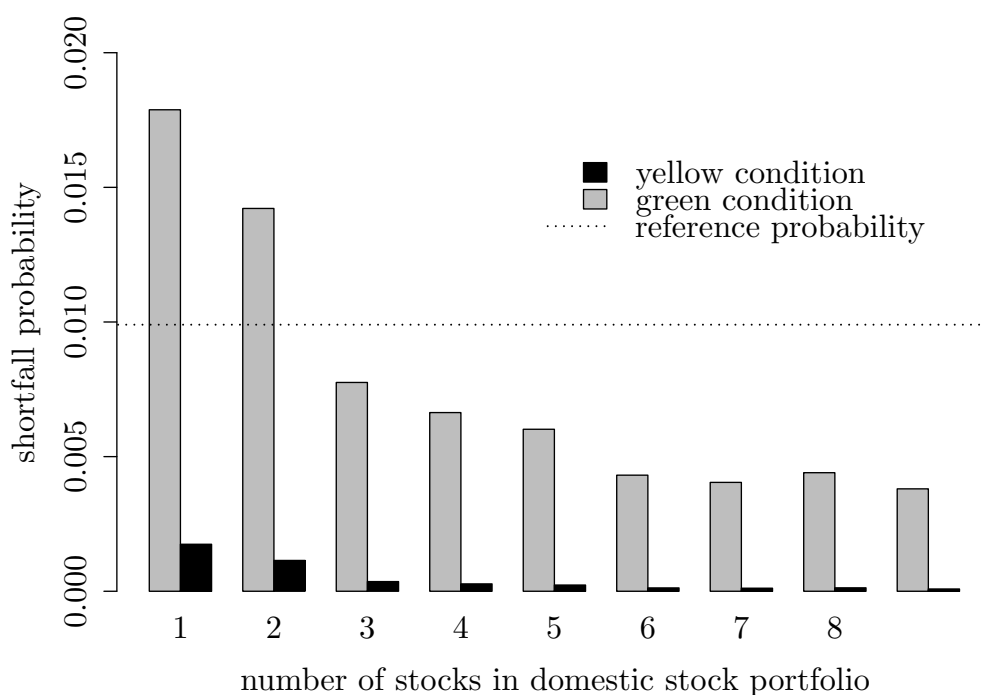
Figure 9: Sensitivity analysis: asset concentration in the domestic stock portfolio

the sensitivity analyses in this section has a total of 18 (international) + 10 (domestic) = 28 percent invested in equities (see Table 12). Taking into account the legal limit of 5 percent per individual stock, this implies that the fund needs to have at least $28/5 = 5.6 \approx 6$ different stocks in its portfolio. Similarly, consider a hypothetical pension fund which invested the legal maximum of 50 percent of its portfolio in equities. As a result, its equity holdings would need to consist of a minimum of $50/5 = 10$ different stocks. Nevertheless, if for some reason a subportfolio is extremely concentrated, the model should be recalibrated accordingly.

6.3 Misestimation of liabilities

Another interesting question centers around the valuation of liabilities. As explained in Sections 4 and 5, the model at the heart of our approach to measuring pension fund solvency treats the liabilities as deterministic and relies on input figures which are reported by the pension funds them-

selves. A current discussion in the Swiss pension fund system revolves around the technical interest rate, which serves as a discount rate for the stochastic future cash outflows in the context of an actuarial valuation of the liabilities. It has repeatedly been stated that many funds hesitated to lower their technical interest rate in lockstep with the development of the term structure, thereby understating the present value of their technical liabilities.[63] In addition, despite various hedging techniques broadly applied in practice (see, e.g., Mao et al., 2008; Yang and Huang, 2009), there is a remaining uncertainty about future mortality improvements and their modeling. Obviously, a potential misestimation of the liabilities will have consequences for the results of the proposed solvency test. Thus, Figure 10 displays the sensitivity of the shortfall probabilities with regard to the estimation error of the technical liabilities. We observe a pattern similar to the effect of a change in the portfolio weight of stocks examined in Section 6.1. Again, the graph comprises a point and a triangle, representing the shortfall probabilities for the original value of the liabilities. In the area to the left of these points, where the liabilities are found to have been overestimated its liabilities, the actual shortfall probabilities rapidly decline towards zero. The opposite is true for an underestimation, however. If the liabilities were a mere 1.6 percent higher than originally estimated, the fund would already breach the threshold for condition green. Beyond that estimation error, the increase of the probabilities becomes even steeper. Again, the whole curve for the yellow condition lies below the reference probability. A practical insight associated with these results is that the supervisory review should include an in-depth analysis of the methodology, assumptions, and database which the pension funds employ to estimate their liabilities. In case the supervisor has reasons to doubt the precision of the estimates, the outcome of the solvency test would have to be adjusted.

6.4   Coverage ratio

Next, we consider the probabilities' sensitivity to the coverage ratio of the pension fund at the beginning of the period. Figure 11 shows the results

---

[63]See, e.g., Swisscanto (2008).
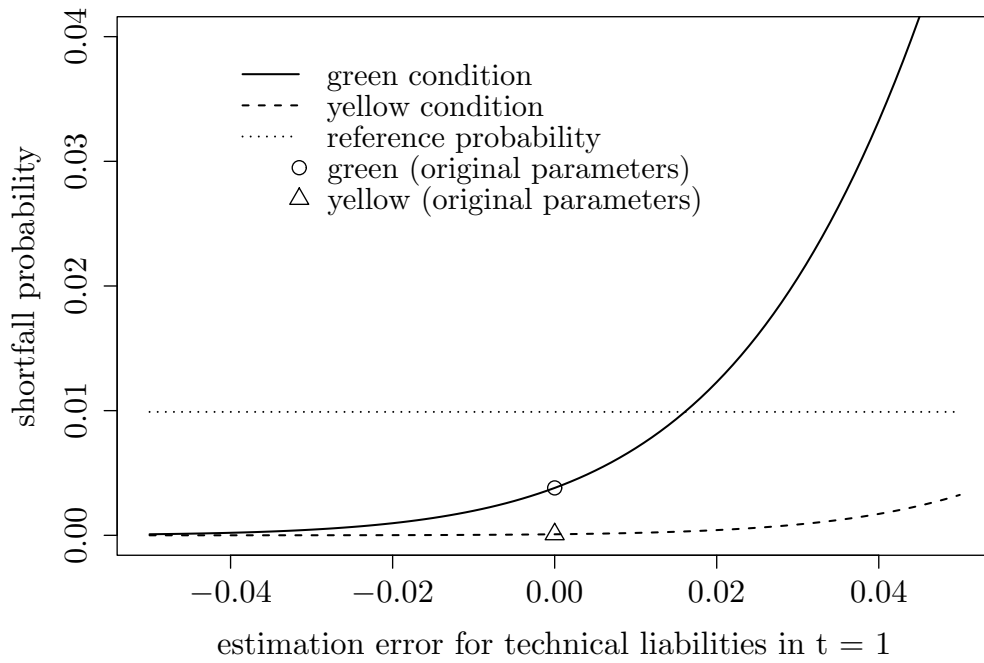
Figure 10: Sensitivity analysis: liabilities
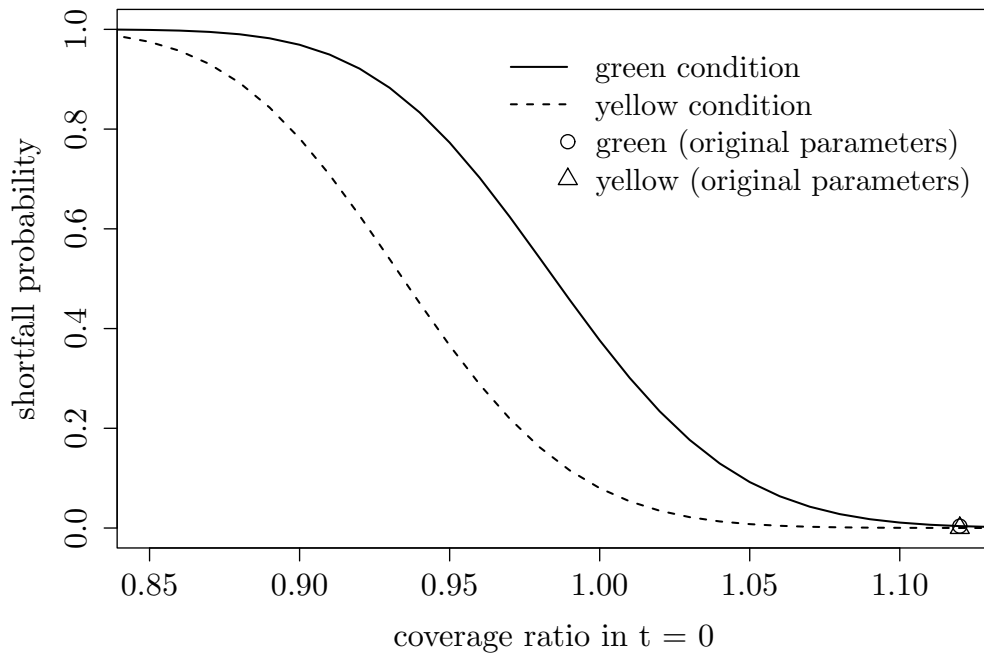
Figure 11: Sensitivity analysis: coverage ratio

for coverage ratios varying from 1.1 down to 0.85. Again, the values of 0.3803 percent for the green condition and 0.0085 percent for the yellow condition associated with the original coverage ratio of 1.1 (see Table 12) are represented by a point and a triangle on the curves.[64] The results for coverage ratios over and above 1.1 are not particularly interesting as the shortfall probabilities quickly become very small. Similarly, we observe that for coverage ratios of below 0.9, the probabilities are very close to 1 and thus far beyond any reasonable threshold. Some more attention should be devoted to the region between 0.9 and 1.1. Just below 1.1, both curves initially exhibit a slightly negative slope, which then sharply increases in magnitude below 1.05 for the curve representing the green and below 1.0 for the curve representing the yellow condition. This is an important result: pension funds with a coverage ratio of below 1.05 are relatively likely to end up with underfunded liabilities at the end of the period. In addition, if their liabilities are just about covered at the beginning of the period, even the probability for failing the yellow condition grows to levels where it begins to be perceivable. This has important implications for the Swiss occupational pension fund system. In particular, the common practice of letting pension funds continue their business with dramatically underfunded liabilities without a specifically tight supervisory review and careful amendments to their overall strategy has to be considered inadequate from the point of view of modern risk management and solvency regulation principles.

6.5   Lowest acceptable coverage ratio

Furthermore, we examine the sensitivity of the shortfall probability for the yellow condition with regard to β, i.e., the lowest coverage ratio acceptable by the supervisor. The results are depicted in Figure 12. Consistent with our previous analyses, the curve begins at a β of 0.95 which is associated with a near zero probability (0.0085 percent) of an underfunding after additional contributions (marked by a triangle). For an increasing β, however, we observe a non-linear rise in the probability. A β of 0.97, for example, is already associated with a probability of

---

[64]Due to their small difference relative to the scale chosen for the overall graph, these points appear as one.
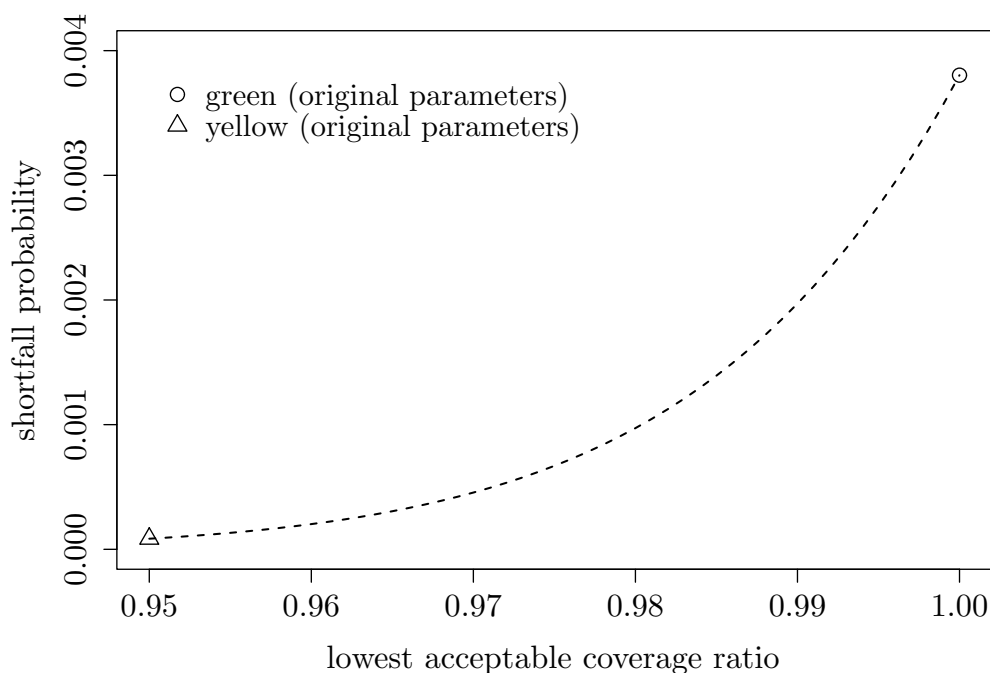
Figure 12: Sensitivity analysis: lowest coverage ratio acceptable by the supervisor

0.046 percent, which is more than five times the above value. When β approaches 1, i.e., additional contributions are ruled out, the probability reaches the value of 0.3803 percent associated with the green condition (marked by a point). This suggests that the impact of each percentage of additional contributions allowed to fix deficits is relatively strong.

Since both the current and lowest acceptable coverage ratio have a strong influence on the shortfall probability for the yellow condition, we finally want to consider their joint impact in order to assess which combinations have counterbalancing or strengthening effects (see Figure 13). A very important observation is, that for β below approximately 0.96, the yellow condition becomes rapidly less binding, even if we assume that the fund already begins the period with a relatively weak coverage ratio of around 1. For a β of 0.90, the shortfall probability for condition yellow is virtually negligible until the coverage reaches 0.95 where it begins to rise sharply. In contrast, if β is set to 1 (no additional contributions allowed), even a relatively small underfunding leads to large shortfall
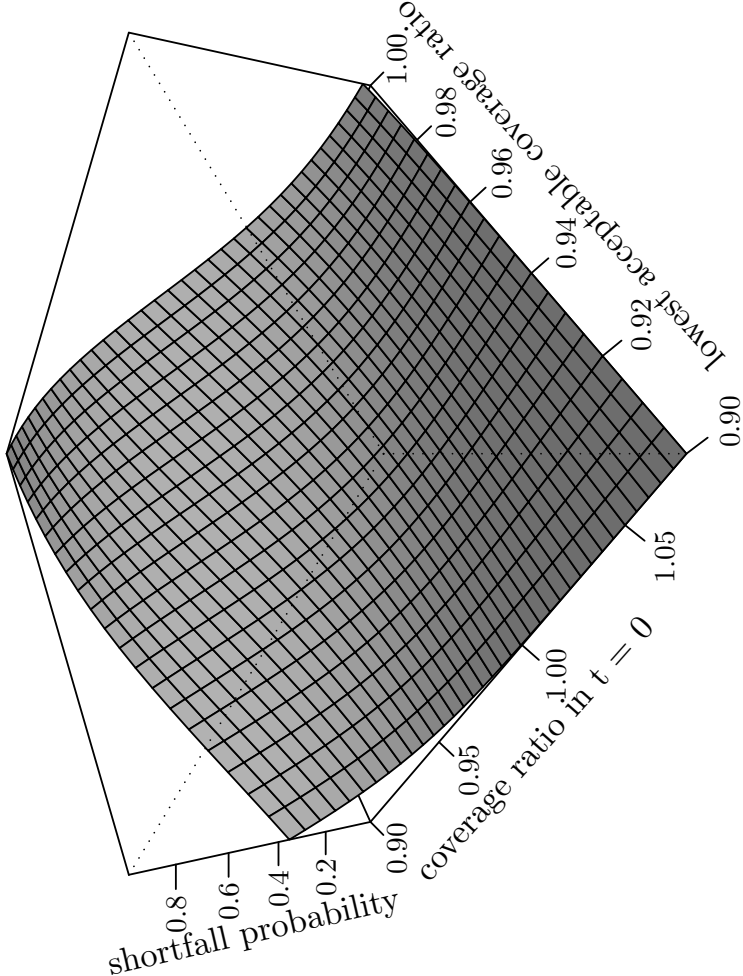
Figure 13: Sensitivity analysis: actual and lowest acceptable coverage ratio (yellow condition)

probabilities, which reach 100 percent around the coverage ratio of 0.90. Thus, as previously suspected, the regulator's choice of β has a crucial impact on the bindingness of the yellow traffic light condition. Simply allowing pension funds with a low coverage ratio to draw on large amount of additional contributions per period provides them with a convenient means to continue business without significant revisions to their asset or risk management practices. This somewhat contradicts the purpose of a solvency test and essentially means that premium payers subsidize pensioners, an effect which is generally not intended within the second pillar of the Swiss pension system.

6.6 Exchange rate risk

In Section 5.1 we mentioned that the U.S. Dollar denominated indices have not been converted to Swiss Francs for the model calibration. For this to be an adequate approach, pension funds would need to hedge out major exchange rate fluctuations in their asset portfolios at an immaterial cost. In this section, we relax the assumption of a perfectly FX hedged portfolio and analyze the impact of exchange rate risk on the shortfall probabilities. To this end, we convert the time series of the three U.S. Dollar denominated indices[65] to Swiss Francs and compute the associated returns. For each index, we then calculate weighted averages of the returns of the original time series (U.S. Dollars) and the returns of the converted time series (Swiss Francs), applying weights of 100, 75, 50, 25, and 0 percent. These weights are meant to reflect the percentage of the foreign currency denominated portfolio which has been hedged against exchange rate risk.[66] Accordingly, a 100 percent weight on the returns of the U.S. Dollar index time series reflects a situation where the whole portfolio is immune to exchange rate fluctuations, whereas a 100 percent weight on the returns of the Swiss Franc converted time series

---

[65] These are the S&P U.S. Treasury Bond Index, the MSCI World, and the HFRI Fund Weighted Index. See Table 7.

[66] While this is a rather general analysis, abstracting from a detailed characterization of the associated transactions with regard to strategy, timing, instruments, volumes, strike prices, etc., we believe it to be satisfactory in this context. A more elaborate treatment of FX hedging issues in the asset management industry is beyond the scope of this paper.
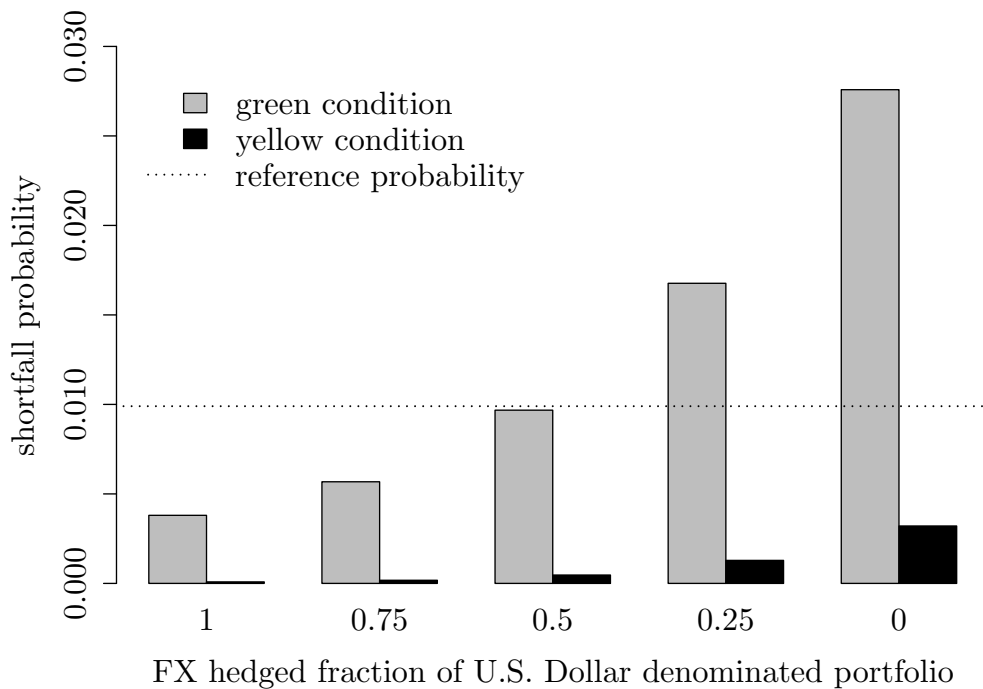
Figure 14: Sensitivity analysis: FX hedging of the U.S. Dollar denominated portfolio

implies no FX hedging activity at all. For all combinations in between, the exchange rate risk is assumed to be partially hedged. Subsequently, means, volatilities, correlation matrix, and the resulting shortfall probabilities are recalculated for each case (see Figure 7 for the results). For both the green and yellow condition the shortfall probabilities expectedly rise with the exchange rate exposure. If foreign investments in the pension fund's portfolio remain entirely unhedged, the probability for the green condition turns out to be more than seven times higher than in the case of a perfect FX hedge. However, we also see that the pension fund would still be assigned a green light if it protects only half of its foreign asset holdings against exchange rate risk. In addition, the yellow condition is passed in every case, even without any FX hedging activities. From these insights we conclude that a model calibration based on foreign currency denominated indices should be valid, as long as pension funds hedge a large part of their foreign asset portfolio against exchange rate fluctuations. If this is not the case, a model recalibration should be requested and monitored by the supervisory authority.

7   Supervisory review and actions

In analogy to Basel II and the SST, the approach we introduced and illustrated throughout the previous sections should be embedded into a comprehensive supervisory review process. As part of this process, occupational pension funds should be obliged to report and comment on certain key figures resulting from the application of the supervisory model in regular intervals. This quantitative solvency report could be accompanied by a qualitative judgment of risks which are not explicitly covered by the model framework, such as credit and operational risk.

In order to react properly to the risk and solvency situation of pension funds, the regulator should possess a variety of competencies. According to the degree of compliance with the traffic light conditions, a certain catalog of measures could be decided. For pension funds which are assigned a green light, the regulator could stick to periodic reviews focused on the adequacy of the regulatory standard model. As illustrated in Section 6.2, a recalibration could become necessary in certain cases.

If a pension fund hands in a regulatory report with a yellow light, it should be subjected to closer scrutiny. This could, for instance, comprise a comprehensive check-up of the fund's assets, liabilities, liquidity, and cash flow profile with a particular focus on valuation methodologies and assumptions. In addition, such funds could be put on a regulatory watch list, resulting in a shortened reporting interval. The requirement to design a concept for financial restructuring is also a potential measure to be imposed on funds in the yellow category. Such a concept would need to cover the asset and liability side, demonstrating how a solid solvency situation can be restored through a combination of portfolio adjustments as well as capital replenishment by means of additional contributions. In any case, the regulator would have to ensure that the lower and upper limit for additional contributions is obeyed.

If a fund is assigned a red light, more drastic consequences would be necessary. These could comprise constraints to the management's ability to choose its asset allocation with the aim of preventing the fund from incurring additional investment risks. Otherwise the problem of "gambling for resurrection" could arise, meaning that the fund management tries

to rescue the institution through large bets. Furthermore, the regulator should be authorized to issue directives to the management of red light funds. Moreover, the inclusion of additional contributors should be suspended until the fund has been restored to an acceptable level of solvency. This protects prospective fund members from the excessive subsidization of current pensions through their contributions. Finally, the regulator should have the ability to replace the board and fund management of highly distressed pension funds with a special administrator.

Beyond that, rules with regard to the publication and dispersion of these easily interpretable solvency signals could increase transparency, and, given the receivers can appropriately react to the information, promote market discipline. Hence, apart from the supervisory authority, receivers of the signal should be employers, employees, transaction partners, and the general public. Due to a reduction of information asymmetries, pension funds with an abnormally high shortfall probability would thus have to face public scrutiny and reactions of their business partners.

8 Some notes on a potential introduction in Switzerland

An important organizational requirement for pension funds which would arise from a concrete introduction of the solvency test is the recruitment of personnel with an adequate background for the application and maintenance of stochastic pension fund models. Further requirements relate to the necessary infrastructure for running the model, including databases and software. In order to align the fund manager's interest with that of the insured, the former should benefit from the prevention of yellow signals. This could, for example, be achieved by linking his variable compensation to a combination of fund performance and traffic light signals.

As explained in Section 2, Swiss occupational pension funds take the legal form of private trusts, which have very limited possibilities of self-supervision. A corporation, in contrast, has bodies such as the board and annual meeting, which serve supervisory purposes. Conse-

quently, the introduction of a regulatory framework for occupational pension funds could be complemented with a fundamental reformation of the legal forms they can adopt. The recommended traffic light approach would then receive additional disciplinary weight through board and shareholders of the corporation as receivers of the signal.

The degree of market discipline emanating from the traffic light approach strongly depends on its familiarity to stakeholder groups and the expected consequences of bad signals such as the potential threat of many insured wanting to change their pension fund. However, employees are currently not free in their choice, which greatly reduces this sort of pressure. Therefore it needs to be discussed whether the introduction of the solvency regulation should be accompanied by a liberalization of the market itself, enabling a free choice of the pension fund. The downside would be, that the situation of an already distressed fund could further deteriorate in case a large number of insureds wants to redeem their holdings. Nevertheless, we believe that more flexibility in this regard is warranted and would be an important step towards an efficient regulation of Swiss occupational pension funds.

Finally, the regulator could conceive of establishing higher barriers to entry for pension schemes. These could, for example, be fit and proper conditions for the individuals managing the pension fund, as common for employees in other branches of the financial services industry such as banking. Participation in the Swiss pension fund market is currently not tied to specific criteria. Setting prerequisites would likely lead to a consolidation, reducing the current number of funds from approximately 2'500 in 2008[67] to a lower number which can be supervised more efficiently.

9 Conclusion

We adopt a stochastic pension fund model and combine it with a traffic light approach for solvency measurement purposes. The calibration and implementation of the model with a small sample of ten pension funds illustrates its application for the computation of probabilities and

---

[67]See Swiss Federal Statistical Office under `http://www.bfs.admin.ch`.

derivation of traffic light signals. The model adequately captures the particularities associated with the occupational pension fund system in Switzerland. Due to its efficiency and ease of calibration it is well suited as a regulatory standard model in this very fragmented market, keeping costs of the solvency test at a minimum, even for small pension funds with less sophisticated risk management know-how and infrastructure. In addition, the sensitivity analysis identifies important drivers of the shortfall probabilities and can thus assist the regulator with regard to specific decisions associated with the configuration of the framework.

However, some questions remain in respect to model design and calibration. First, we did not explicitly account for credit risk in the fund's asset portfolio. Therefore, the supervisory authority should exercise additional care with regard to solvency test results for pension funds with a relatively high proportion of default-able instruments, such as corporate bonds, in their portfolio. Second, an incorporation of stochastic liabilities and different statistical distributions for the modeled asset classes could be discussed, although a departure from the associated assumptions would necessitate a switch from the closed-form to a numerical solution. Third, portfolio diversification and foreign currency exposure have to be borne in mind as critical factors with regard to the proposed calibration procedure. Finally, a practical implementation would need to be preceded by a comprehensive quantitative impact study for the majority of Swiss pension funds. Overall, we consider this straightforward framework to be an adequate first step towards a state-of-the-art solvency regulation of occupational pension funds in Switzerland.

References

Bédard, D. and Dufresne, D. (2001). Pension Funding with Moving Average Rates of Return. *Scandinavian Actuarial Journal*, 1:1–17.

Brombacher Steiner, M. V. (1999). The Swiss Pension Scheme: Long-term Security Through Reform. *Geneva Papers on Risk and Insurance - Issues and Practice*, 24(4):488–494.

Bühlmann, H. (1996). *Mathematical Methods in Risk Theory*. Berlin et al., 2 edition.

BVG (2009). Bundesgesetz über die berufliche Alters-, Hinterlassenen- und Invalidenvorsorge (June 1). *URL http://www.admin.ch*.

BVV2 (2009). Verordnung über die berufliche Alters-, Hinterlassenen- und Invalidenvorsorge (June 1). *URL http://www.admin.ch*.

Cairns, A. J. G. (1995). Pension Funding in a Stochastic Environment: The Role of Objectives in Selecting an Asset Allocation Strategy. *Heriot-Watt University Working Paper*.

Cairns, A. J. G. (1996). Continuous-Time Pension Fund Modeling. *Heriot-Watt University Working Paper*.

Cairns, A. J. G. (2000). Some Notes on the Dynamics and Optimal Control of Stochastic Pension Fund Models in Continuous Time. *Astin Bulletin*, 30(1):19–55.

Cairns, A. J. G. and Parker, G. (1997). Stochastic Pension Fund Modelling. *Insurance: Mathematics and Economics*, 21(1):43–79.

Dufresne, D. (1988). Moments of Pension Contributions and Fund Levels When Rates of Return are Random. *Journal of the Institute of Actuaries*, 115:535–544.

Dufresne, D. (1989). Stability of Pension Systems When Rates of Return are Random. *Insurance: Mathematics and Economics*, 8(1):71–76.

Dufresne, D. (1990). The Distribution of a Perpetuity, with Applications to Risk Theory and Pension Funding. *Scandinavian Actuarial Journal*, 9:39–79.

Eling, M., Gatzert, N., and Schmeiser, H. (2008). The Swiss Solvency Test and its Market Implications. *Geneva Papers on Risk and Insurance - Issues and Practice*, 33(3):418–439.

Gugler, P. (2005). The Integrated Supervision of Financial Markets: The Case of Switzerland. *Geneva Papers on Risk and Insurance - Issues and Practice*, 30(1):128–143.

Haberman, S. (1992). Pension Funding with Time Delays: A Stochastic Approach. *Insurance: Mathematics and Economics*, 11(3):179–189.

Haberman, S. (1993a). Pension Funding: The Effect of Changing the Frequency of Valuations. *Insurance: Mathematics and Economics*, 13(3):263–270.

Haberman, S. (1993b). Pension Funding with Time Delays and Autoregressive Rates of Investment Return. *Insurance: Mathematics and Economics*, 13(1):45–56.

Haberman, S. (1994). Autoregressive Rates of Return and the Variability of Pension Contributions and Fund Levels for a Defined Benefit Pension Scheme. *Insurance: Mathematics and Economics*, 14(3):219–240.

Haberman, S. (1997). Stochastic Investment Returns and Contribution Rate Risk in a Defined Benefit Pension Scheme. *Insurance: Mathematics and Economics*, 19(2):127–139.

Haberman, S. and Sung, J.-H. (1994). Dynamic Approaches to Pension Funding. *Insurance: Mathematics and Economics*, 15(2-3):151–162.

Mao, H., Ostaszewski, K. M., and Wang, Y. (2008). Risk Analysis of Mortality Improvement: The Case of Chinese Annuity Markets. *Geneva Papers on Risk and Insurance - Issues and Practice*, 33(2):234–249.

O'Brien, T. (1986). A Stochastic-Dynamic Approach to Pension Funding. *Insurance: Mathematics and Economics*, 5(2):141–146.

O'Brien, T. (1987). A Two-Parameter Family of Pension Contribution Functions and Stochastic Optimization. *Insurance: Mathematics and Economics*, 6(2):129–134.

OECD (2009). Pensions at a Glance: Retirement Income Systems in OECD Countries. *URL http://www.oecd.org*.

Officer, R. R. (1972). The Distribution of Stock Returns. *Journal of the American Statistical Association*, 67(340):807–812.

Standard & Poor's (2009). 2008 Annual Global Corporate Default Study and Rating Transitions. *URL http://www.standardandpoors.com.*

Swiss Federal Statistical Office (2009). Pensionskassenstatistik: Struktur der beruflichen Vorsorge in der Schweiz. *URL http://www.bfs.admin.ch.*

Swisscanto (2008). Schweizer Pensionskassen. *URL http://www.swisscanto.ch.*

Yang, S. S. and Huang, H.-C. (2009). The Impact of Longevity Risk on the Optimal Contribution Rate and Asset Allocation for Defined Contribution Pension Plans. *Geneva Papers on Risk and Insurance - Issues and Practice*, 34(4):660–681.

Zimbidis, A. and Haberman, S. (1993). Delay, Feedback and Variability of Pension Contributions and Fund Levels. *Insurance: Mathematics and Economics*, 13(3):271–285.

Part IV

Stock vs. Mutual Insurers: Who Does and Who Should Charge More?

Abstract

In this paper, we empirically and theoretically analyze the relationship between the insurance premium of stock and mutual companies. Evaluating panel data for the German motor liability insurance sector, we do not find evidence that mutuals charge significantly higher premiums than stock insurers. If at all, it seems that stock insurer policies are more expensive. Subsequently, we employ a comprehensive model framework for the arbitrage-free pricing of stock and mutual insurance contracts. Under the chosen set-up, the formulae for the premium and the present value of the equity of a stock insurer are nested in our more general model. Based on a numerical implementation of the framework, we then compare stock and mutual insurance companies with regard to the three central magnitudes premium size, safety level, and equity capital. Although we identify certain circumstances under which the mutual's premium should be equal to or smaller than the stock insurer's, these situations would generally require the mutual to hold less capital than the stock insurer. This being inconsistent with our empirical results, it appears that policies offered by stock insurers are overpriced relative to policies of mutuals. While our analysis focuses on the insurance context, the insights can be transferred to other industries where mutual companies are an established legal form.[68]

---

[68]This paper has been written jointly with Alexander Braun and Hato Schmeiser.

1   Introduction

Private insurance firms in many insurance markets can be organized either as mutual or stock insurance companies. Similar to policyholders of a stock insurance company, those of a mutual insurer are obliged to pay the insurance premium which, in turn, entitles them to an indemnity payment contingent on the occurrence of a loss. Apart from that, however, several important differences between these two legal forms exist (see, e.g., Smith and Stutzer, 1990). First of all, in contrast to stock insurers, mutuals are in fact owned by their policyholders. By paying the respective premium, the buyer of a mutual policy becomes a so-called member, which is economically equivalent to simultaneously acquiring a policyholder and an equityholder stake in the firm.[69] As a result, those insured by a mutual are usually granted direct or indirect participation in the administrative bodies and should thus be able to exert influence on business decisions. To establish a similar position, policyholders of stock insurance companies would need to acquire additional ownership rights by purchasing the company's common stock. Unlike the shareholders of a stock insurer, however, members of a mutual cannot simply sell their equity stake. This is due to the fact that, in practice, it is not explicitly differentiated from the policyholder stake and a secondary market does not exist. Hence, the only way to fully realize the value of the equity are liquidation or demutualization of the company, which would need to be enacted collectively by a majority of the members.[70] A further difference to stock insurers is, that mutual members can expect occasional premium refunds if the company is profitable. These payouts are economically akin to the dividends a stock insurer distributes to its shareholders. Finally, stock insurance companies cannot draw on their policyholders to recover financial deficits, whereas the membership in a mutual insurer might be associated with the obligation to make additional premium payments contingent on the firm being in financial distress. These additional premiums are virtually authorized capital,

---

[69]Rasmusen (1988) describes rights and obligations resulting from a membership in savings and loan associations, credit unions, and mutual savings banks.

[70]In the course of a demutualization, the insurer changes its legal form and is transformed into a stock company.

i.e., equity which has not been paid in yet (see, e.g., Mayers and Smith, 1988). Since the legal form determines these rights and obligations associated with the purchase of an insurance contract, it should ceteris paribus result in different arbitrage-free prices for policies, covering identical claims.

While there is a large body of literature, dealing with various aspects of mutual and stock companies, to the best of our knowledge, there has not yet been a rigorous empirical and theoretical analysis of the relationship between the premium of stock and mutual insurers. Therefore, in this paper, we want to shed some light on this research question by evaluating panel data for the German motor liability insurance sector. In addition, we contribute to the literature by employing a contingent claims model framework to consistently price stock and mutual insurance contracts. For this purpose, we split the arbitrage-free mutual insurance premium into an ownership and policyholder stake, both of which are further decomposed into distinct option-theoretic building blocks. The model explicitly takes into account the restricted ability of members to realize the value of their equity stake as well as the mutuals' right to charge additional premiums in times of financial distress, which will be termed recovery option in the course of this paper. Under the chosen set-up, the formulae for the premium and the present value of the equity of a stock insurer are nested in our more general model. Moreover, we derive conditions, under which the premiums of a stock and a mutual insurance company should theoretically be equal. Finally, combining our empirical and theoretical results, we are able to derive relevant economic implications. While we apply our model within the insurance context, its insights can be transferred to other industries where mutual companies are an established legal form such as credit unions and pension funds.

The remainder of this paper is organized as follows. Section 2 contains a comprehensive overview of previous literature on issues surrounding stock and mutual insurance companies. In Section 3, we apply panel data methodology to provide some empirical evidence with regard to the relationship between the premiums of stock and mutual insurers. Aiming to explain these empirical results by means of normative theory, in Section 4 we develop our contingent claims model framework, beginning

with the simple and well-established case of the stock insurance company. Subsequently, we consider a mutual insurer with recovery option and fully realizable equity, before formally describing the general case with partial participation in future equity payoffs. Section 5 comprises a comprehensive numerical analysis which forms the basis for our normative findings. In Section 6, we integrate our empirical and theoretical results and discuss relevant economic implications. Finally, in Section 7, we conclude.

2   Literature overview

The literature comparing stock and mutual insurance companies has predominantly dealt with agency issues associated with the legal form. Coase (1960) argues that the ownership structure of a company, which is determined by property rights constituting the discretionary power of control, is relevant only in the presence of transaction costs. This is due to the fact that conflicts of interest between different stakeholders may arise and entail costs, which depend on the extent of discretion as well as established control mechanisms. Ownership structure is identified as one possible means of control. In this spirit, Mayers and Smith (1981) develop a positive theory on insurance contracting, extending the fundamental work of Jensen and Meckling (1976) on agency theory. They analyze incentives resulting from the different ownership arrangements of stock and mutual insurers and discuss two kinds of potential conflicts between parties brought together in an insurance firm. On the one hand, asymmetric information and the call option-like payoff profile associated with the shareholder position in a stock insurer imply that the equity value increases with the risk inherent in the company.[71] At the same time, however, riskier assets are detrimental to the position of the policyholders, giving rise to the so-called owner-policyholder conflict. Against this background, the company's owners will seek to establish efficient sanction mechanisms, ensuring that the management acts in their inter-

---

[71]The notion that the equity stake in a company can be interpreted as a call option on its assets, struck at the face value of the liabilities, was introduced by Merton (1974).

est. Consequently, agency costs occur and impair economic efficiency compared to a setting without transaction costs. In contrast to that, since owners and policyholders within a mutual insurance company coincide, agency costs can be reduced.[72] On the other hand, stock insurers provide more efficient sanction mechanisms to tackle the so-called owner-management conflict, which results from diverging incentives between shareholders and company executives. In addition to being held responsible by the organizational bodies of the insurance company, which are controlled by its owners, poorly performing executives of a stock insurer must fear market discipline such as, for instance, hostile takeovers.[73] The reason is that, in contrast to a mutual insurer, the equity of a stock insurer is freely tradable and not linked to a particular insurance policy. Hence, agency costs resulting from the so-called owner-manager conflict can be expected to be higher for mutuals.[74] Assuming that a large number of decision makers (owners) cannot coordinate as efficiently as single entity or individual, the costs of control can be expected to rise with the granularity of the equity stake. While the majority of shares of publicly listed corporations are frequently owned by large blockholders, only a marginal fraction of the ownership rights is allocated to each member of a mutual firm. Thus, internal sanction mechanisms are likely to be more effective for stock than for mutual insurers. Accordingly, from the policyholder perspective, the optimal choice of legal form should depend on the trade-off between agency costs arising from the owner-policyholder and the owner-manager conflict. Therefore, Mayers and Smith (1981, 1988, 1994) argue that stock firms should be more prevalent in activities that involve significant managerial discretion, since, in this context, potential owner-manager conflicts are most severe (also see Pottier and Sommer, 1997). In contrast to that, mutuals should theoretically prevail in the long-term lines of business that are usually encumbered with a

---

[72]Also see Garven (1987). Similar incentives can be achieved by including participation rights in the stock insurance contracts (see, e.g., Garven and Pottier, 1995).

[73]The new owner normally exchanges the board of the company (see, e.g., Mayers and Smith, 1988).

[74]Fama and Jensen (1983a,b) argue that a further mechanism to control management is the fact that assets of all mutual financial institutions need to be redeemed on demand of their members. However, we assent to the arguments raised by Smith and Stutzer (1990), who suggest that this is not the case within the insurance context.

more significant owner-policyholder conflict potential, such as the life insurance sector (see Hansmann, 1985; Mayers and Smith, 1988).

A number of empirical articles support the previously explained agency-theoretic considerations. Lamm-Tennant and Starks (1993) provide evidence for the owner-policyholder conflict by showing that stock insurers are generally riskier than mutual insurance companies. This is coherent with the results of Lee et al. (1997), who analyze both legal forms in the context of insurance guaranty funds. Furthermore, the greater potential for the owner-manager conflict in mutuals is illustrated by Greene and Johnson (1980), who conduct a survey in which they analyze policyholder awareness of the rights resulting from the ownership stake in a mutual insurance company. Compared to the holders of publicly traded stock, members of the analyzed mutual companies were less aware of their voting rights and appeared to exercise less control. Similarly, Wells et al. (1995) find that, in contrast to managers of stock insurers, those of mutuals have a higher free cash flow at their disposal, implying a greater opportunity to waste cash on unprofitable investments. Further evidence for the owner-manager conflict in the context of mutual and stock insurers is provided by Mayers and Smith (2005), who document that mutual company charters are more likely to contain provisions which limit the range of operating policies of the firm. Zou et al. (2009) observe that, probably owing to their inferior management-control mechanisms, mutuals tend to pay significantly lower dividends than stock insurers. Finally, analyzing data from the property-liability insurance sector, He and Sommer (2010) find that, compared to stock insurers, the board of mutuals generally comprises a larger fraction of outside directors. They argue that additional monitoring through outside directors is necessary since ownership and control in mutuals are separated to a greater extent, thus increasing agency costs arising from the owner-manager conflict.[75]

Another major strand of literature deals with changes in the legal form of an insurer. Fletcher (1966) as well as Mayers and Smith

---

[75]The owner-manager conflict in the context of mutual and stock banks has, e.g., been considered by Gropper and Hudson (2003) who provide evidence for considerable expense-preference behavior in mutual savings and loans associations based on a U.S. wide sample.

(1986) focus on mutualization issues. However, much more research has been conducted on the demutualization process. A survey by Fitzgerald (1973) identifies economic pressure as the main reason for the conversion of small property-liability insurers into stock companies, while Viswanathan and Cummins (2003) view access to capital as a major driver for demutualization. Furthermore, Carson et al. (1998) find the level of free cash flow to be significantly related to the probability of mutual firms transforming into stock companies. Zanjani (2007) analyzes macroeconomic and regulatory conditions under which mutual insurance companies have been formed in order to explain the observed evolution of the whole U.S. life insurance industry from the mutual towards the stock insurer form. He concludes that tight state regulation did not coincide with a demise of the mutual form. Instead, a general rise in founding capital requirements seems to have harmed mutuals due to their very limited access to external funding. Moreover, Erhemjamts and Leverty (2010) argue that the incentive to demutualize differs by the type of conversion: full demutualization versus mutual holding company. Finally, in their empirical study of U.S. life insurers, McNamara and Rhee (1992) find that increased efficiency seems to be an important reason for demutualization.

The question of efficiency differences between stock and mutual firms has been further examined by several other authors. Spiller (1972) finds evidence that ownership structure is a determinant of performance. While Jeng et al. (2007) present mixed results with regard to efficiency improvements implied by changes of the legal form, Cummins et al. (1999) find mutuals to be less cost-efficient.[76] Furthermore, in their study based on Spanish insurance market data, Cummins et al. (2004) identify differences in efficiency between stocks and mutuals only for small mutual insurance companies. Harrington and Niehaus (2002) focus on dissimilarities concerning capital structure, which may result from the costs of raising new capital and Viswanathan (2006) finds initial public

---

[76]Iannotta et al. (2007) conducted a similar study for the banking industry. Controlling for company characteristics as well as geography and time, they find that mutual banks are less profitable than stock banks. Moreover, they provide evidence for a higher loan quality among mutuals compared to stock and public sector banks.

offerings of mutuals to be significantly underpriced. The latter result is confirmed by Lai et al. (2008).

Besides agency-theoretic considerations, (de)mutualization, and efficiency implied by the legal form, various other topics related to stock and mutual insurers have been explored in the literature. Differences in the contractual structure of policies offered by mutual and stock insurers are examined by Smith and Stutzer (1990, 1995). They argue that information asymmetries rather than agency problems are the major determinant for the types of contracts offered by mutuals. The parallel existence of different legal forms of insurance companies is justified, amongst others, by self-selection of those insured. In addition, Cass et al. (1996) consider how a Pareto optimal risk allocation can be achieved through mutual insurance in the presence of individual risk. Ligon and Thistle (2005) point out that issues arising from asymmetric information can restrict the size of mutual institutions. Using an equilibrium model in which mutuals can exclusively offer fully participating policies, Friesen (2007) shows that stock companies can only provide partially participating insurance when their shareholders require premiums that ensure a fair return on equity. Finally, Laux and Muermann (2010) demonstrate that, by linking policies to the provision of capital, mutuals can resolve free-rider and commitment issues faced by stock insurers.

3   Empirical analysis

In this section, we want to empirically investigate whether the legal form of an insurance company is a determinant of the premium it charges. To ensure comparability, the insurance product under consideration needs to be as homogeneous as possible. Therefore, our sample is based on annual accounting figures for the German motor vehicle liability insurance sector.[77] The data has been obtained from Hoppenstedt, a major provider of company information for a wide variety of industries of the German economy. To ensure consistency, we have carried out cross-checks with the annual reports of the respective insurers. The sample

---

[77]Specialty insurers have been excluded.

consists of 99 stock and 14 mutual insurers for which repeated observations over a differing number of time periods between 2000 and 2006 are available. Hence, we are working with unbalanced panel data, covering 532 and 87 firm years for stock and mutual insurance companies, respectively. Table 18 contains some descriptive statistics on the panel dataset. We measure the price of insurance by means of the average annual gross premium ($AvPrem$), which is obtained by dividing the total annual premium volume in the motor liability business line of each firm by the respective number of contracts.[78] Within the analysis, we control for various additional factors which are likely to influence the insurance price. The average annual loss ($AvLoss$), defined as the amount of losses in the motor insurance line divided by the number of contracts, is used as a proxy for underwriting risk. In a similar manner, the average annual costs of the motor liability business line ($AvCosts$) are employed to account for differences in the efficiency of the companies. Furthermore, we include the equity ratio (EqR), i.e., the book value of equity divided by the book value of the assets, as well as the log total premium volume in a given year (LTP) to control for capital structure and size effects.

The Lagrange multiplier (pooling) test, conducted in line with Gouriéroux et al. (1982), suggests significant cross-sectional and time effects in our data.[79] In this case, the pooled ordinary least squares (OLS) estimator is known to be inefficient: it does not fully exploit the information inherent in panel datasets (see, e.g., Petersen, 2008). Instead, more sophisticated models are needed to make the most effective use of our data. Based on the Hausman test (see Hausman, 1978) with a χ^2 test statistic of 483.70 and four degrees of freedom, we reject the random effects (RE) model. A likely reason for this outcome are significant

---

[78]An alternative measure for the insurance price is the economic premium ratio (EPR) which has been suggested by Winter (1994) and is frequently used in the literature (see, e.g., Gron, 1994; Cummins and Danzon, 1997; Phillips et al., 2006). For a given business line of an insurer, the EPR is the ratio of premium revenues net of expenses and policyholder dividends relative to the estimated present value of losses (see Phillips et al., 2006). Since, in the subsequent chapters, we are interested in the mutual premium which includes an equityholder and a policyholder stake, policyholder dividends cannot be excluded. In addition, our data does not cover line-specific estimates for the present value of losses. Hence, we control for underwriting risk by incorporating average annual losses into our regression equations.

[79]We compute a χ^2 test statistic of $2,134.13$, with two degrees of freedom.

Panel A: *Pooled stocks and mutuals*

	abbreviation	mean	std. dev.	minimum	maximum	skewness	kurtosis
Average annual premium	AvPrem	256.6744	57.8718	110.8053	577.1687	0.5238	5.5277
Average annual loss	AvLoss	225.6424	63.3354	54.6619	514.9426	0.7426	4.5793
Average annual costs	AvCosts	40.4934	23.1798	5.8454	287.5176	3.1525	26.7360
Equity ratio	EqR	0.2244	0.1127	0.0387	0.7432	1.3999	5.5725
Log-Total premiums	LTP	19.0946	1.4775	13.4902	22.9681	-0.2079	2.8635

Panel B: *Mutuals*

	abbreviation	mean	std. dev.	minimum	maximum	skewness	kurtosis
Average annual premium	AvPrem	206.7204	40.2694	149.1075	327.2924	1.1263	3.9781
Average annual loss	AvLoss	180.2888	45.4460	110.6300	371.1519	2.4661	10.3262
Average annual costs	AvCosts	27.1455	13.3401	5.8454	53.9230	-0.0729	1.9150
Equity ratio	EqR	0.2798	0.1253	0.1138	0.5804	0.8573	2.7524
Log-Total premiums	LTP	19.3349	1.2750	16.2262	21.1092	-0.7238	3.3359

Panel C: *Stocks*

	abbreviation	mean	std. dev.	minimum	maximum	skewness	kurtosis
Average annual premium	AvPrem	264.8436	56.2097	110.8053	577.1687	0.4961	6.3869
Average annual loss	AvLoss	233.0593	62.7851	54.6619	514.9426	0.6260	4.8106
Average annual costs	AvCosts	42.6762	23.7181	9.3966	287.5176	3.2660	26.9900
Equity ratio	EqR	0.2154	0.1079	0.0387	0.7432	1.5321	6.5753
Log-Total premiums	LTP	19.0553	1.5054	13.4902	22.9681	-0.1359	2.8267

Descriptive statistics for the variables which enter the empirical analysis. In Panel A, all available data has been pooled, whereas Panel B and C refer to the separate subsamples of mutual and stock insurers. The underlying currency is Euro.

Table 13: Descriptive statistics of the data

correlations between unit-specific components and regressors, implying an inconsistent RE (and pooled OLS) estimator. While a fixed effects (FE) model with unit-specific intercept terms could handle this sort of correlation, the so-called FE within estimator is based on a transformation of the regression equation into deviations from individual means and is thus incapable of capturing the impact of time-invariant variables (see, e.g., Wooldridge, 2010). This a serious issue since our analysis is focused on the legal form, which, if at all, changes very rarely.

Therefore, we decide to apply the Hausman-Taylor estimator, an instrumental variables approach combining characteristics of FE and RE models (see Greene, 2007; Verbeek, 2008). It is capable of handling correlations between independent variables and unobserved unit-specific effects and enables us to estimate coefficients for time-invariant regressors. Consider the following linear regression equation:

$$
\begin{aligned}
AvPrem_{it} \;=\; & \mu + \beta_1 AvLoss_{it} + \beta_2 AvCosts_{it} + \beta_3 EqR_{it} \\
& + \beta_4 LTP_{it} + \beta_5 Stock_i + u_i + \epsilon_{it}.
\end{aligned}
\tag{52}
$$

where μ is the intercept and $Stock_i$ is a time-invariant dummy variable representing the legal form of insurer i, which is set to one for stock and zero for mutual companies. The u_i are $N-1$ (here: 112) unit-specific fixed effects and ϵ_{it} denotes the independent and identically distributed error term. In order to estimate this model, Hausman and Taylor (1981) propose the following instruments: exogenous regressors, i.e., those explanatory variables that are uncorrelated with the unit-specific effects, are their own instruments. In addition, endogenous time-varying and time-invariant regressors are instrumented by their own individual means (over time) and those of the exogenous time-varying regressors, respectively.[80] Hence, the analysis requires at least as many exogenous time-varying as there are endogenous time-invariant regressors, i.e., one in our case. Based on a correlation test between the above explanatory variables and their unit-specific components from a fixed effects model, we identify EqR as exogenous.

---

[80] A more detailed treatment of the Hausman-Taylor estimator is beyond the scope of this paper. The reader is referred to advanced panel data texts such as Hsiao (2002), Baltagi (2005), and Wooldridge (2010).

An alternative three-stage procedure for estimation of time-invariant variables in panel data models named fixed effects vector decomposition (FEVD) has been proposed by Plumper and Troeger (2007). Originated in the empirical political science literature, FEVD quickly gained popularity among researchers in various fields. Although the authors provided Monte Carlo simulation results to underline the apparent favorable characteristics of their estimator, it has recently been severely criticized. In particular, FEVD standard errors have been shown to be systematically too small and the estimator is inconsistent if time-invariant variables are correlated with unit-specific effects (see Breusch et al., 2010 and Greene, 2010). Despite these major shortcomings, we decide to additionally apply this method for comparison purposes.

Table 14 contains the estimation results for the Hausman-Taylor approach, the FEVD procedure, as well as a simple FE model.[81] Apart from EqR, all time-varying regressors seem to be key determinants of the insurance premium, since they are associated with statistically significant coefficients for each of the three estimators. For the time-invariant variable $Stock$, in contrast, we get diverging results. While the Hausman-Taylor estimator does not indicate a significant difference in the average premium of stock and mutual insurers, the FEVD coefficient suggests that mutuals tend to charge significantly less. Taking into account the above-mentioned limitations of FEVD, we are evidently more confident in the Hausman-Taylor estimate. For our purpose, however, it is sufficient to conclude that observed premiums are either approximately equal, or stock insurer policies tend to be more expensive. To put it differently, we do not find evidence that mutuals charge higher premiums than stock insurers. Throughout the remainder of this paper we want to adopt a normative stance and explore whether this empirical phenomenon is consistent with fair insurance prices as suggested by contingent claims theory.

---

[81]The heteroskedasticity and autocorrelation consistent (HAC) covariance matrices of Andrews (1991) as well as Driscoll and Kraay (1998) have been applied.

	Hausman-Taylor	FEVD Procedure	Fixed Effects Model
(Intercept)	-213.4151***	-237.3012***	—
	(-2.6692)	(-12.1466)	
AvLoss	0.3420***	0.3469***	0.3420***
	(15.4295)	(9.9042)	(10.9533)
AvCosts	0.6053***	0.5994***	0.6053***
	(7.3825)	(6.1891)	(3.9955)
EqR	20.0231	15.7489*	20.0231
	(1.0095)	(1.9075)	(0.5184)
LTP	19.2463***	18.7959***	19.2463***
	(7.0319)	(17.3699)	(7.3742)
Stock	-3.9429	33.7803***	—
	(-0.0470)	(14.7292)	

Coefficients and t-statistics (in parentheses) for Hausman-Taylor estimator, the FEVD procedure, and the standard FE model. The average annual premium ($AvPrem$) is regressed on the following set of explanatory variables: average annual losses ($AvLoss$), average annual costs ($AvCosts$), equity ratio (EqR), and logged total premium (LTP). Hausman-Taylor and FEVD additionally include the time-invariant variable legal form ($Stock$). ***, **, and * denote statistical significance on the 1, 5, and 10 percent confidence level.

Table 14: Estimation results

4   Model framework

In this section we present a general contingent claims model framework for insurance companies based on the seminal work of Merton (1974) as well as Doherty and Garven (1986). Assume that the firm runs for a single period and all stakes are paid in full at the outset. The economy is characterized by perfect capital markets, i.e., there are no bid-ask spreads, transaction costs, short-selling constraints, taxes or other market frictions. We begin with the relatively simple case of the stock insurance company (Section 4.1), which is then incrementally generalized to include the specifics of mutual insurers. In Section 4.2, we introduce the recovery option, i.e., the right to demand additional payments in times of financial distress. Similarly, in Section 4.3, we further extend

our model by allowing for incomplete participation of members in the mutual's equity payoffs.

4.1 Stock insurer claims structure

Equity stake

An insurance firm in the legal form of a corporation (stock insurer) is bankrupt, if the market value of the assets A_1 available at the end of the period is insufficient to cover its claims costs (losses) L_1, i.e., $A_1 < L_1$. Due to the limited liability of the owners, the equity in $t = 1$ is worth zero in this case. Therefore, the payoff profile of the equity stake equals that of a European call option on the company's assets, struck at the value of the claims. Hence, the present value of the equity of a publicly traded stock insurer EC_0, which is a function a parameter set $\mathcal{P}$, can be expressed as follows

$$
\begin{aligned}
EC_0^{\mathrm{S}} &= e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left[\max\left(A_1 - L_1; 0\right)\right] \\
&= e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left(A_1 - L_1\right) + DPO_0^{\mathrm{S}},
\end{aligned}
\tag{53}
$$

where $\mathrm{E}_0^{\mathbb{Q}}$ denotes the conditional expectation in $t = 0$ under the risk-neutral measure $\mathbb{Q}$, r is the riskless interest rate, and $\mathcal{P}$ contains the relevant parameters for any specific option pricing framework.[82] The call option payoff is equivalent to a long position in the assets and a short position in the claims costs $(A_1 - L_1)$ plus the value of the so-called default put option of the stock insurer (DPO^{S}). To see this refer to Figure 15. The default put option is a proxy for the expected bankruptcy cost and therefore a measure for the safety level of the firm from the policyholder perspective (see Doherty and Garven, 1986). Its present value $DPO_0^{\mathrm{S}} = DPO_0^{\mathrm{S}}(\mathcal{P})$ is equal to

$$
DPO_0^{\mathrm{S}} = e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left[\max\left(L_1 - A_1; 0\right)\right].
\tag{54}
$$

---

[82]Under the Black and Scholes (1973) model, e.g., the parameter set $\mathcal{P}$ would contain the initial value of the assets, the level of claims costs (i.e., the option's strike price), the asset volatility, the risk-free interest rate, as well as the time to maturity.
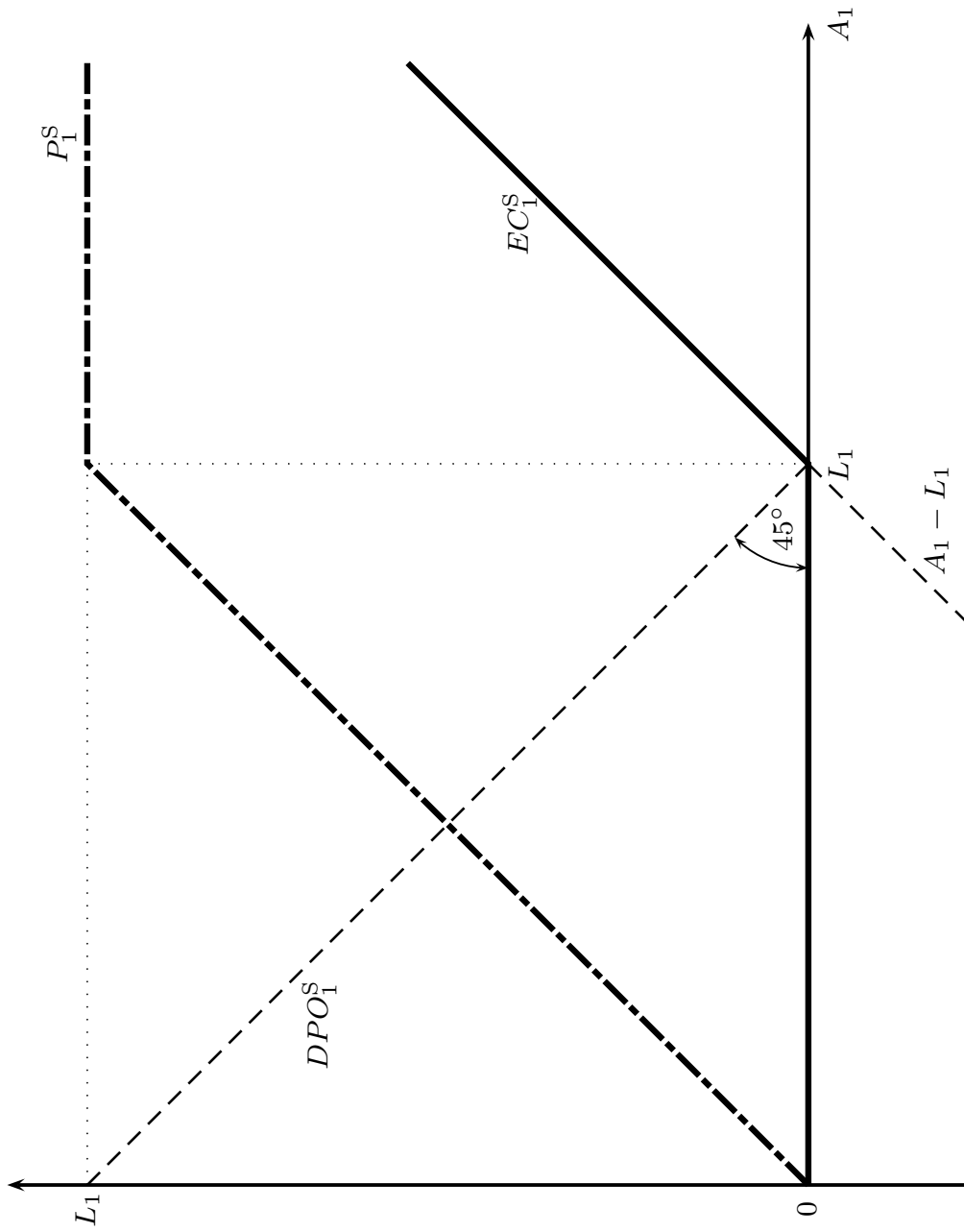
Figure 15: Payoff to the equityholders EC_1^S (solid line) and policyholders P_1^S (dotdashed line) of a stock insurance company in $t = 1$. The dashed lines illustrate the elements of the replicating portfolio ($A_1 - L_1$ and DPO_1^S).

Policyholder stake

If the stock insurer is solvent at time $t = 1$, the insurance company fully indemnifies policyholders for their incurred losses. In case of bankruptcy, however, policyholders only receive the part of their claims which is covered by the remaining market value of the assets in $t = 1$. Based on this payoff profile, the present value of the policyholder stake and thus the fair premium π_0^{S} of a stock insurer, $P_0^{\mathrm{S}} = P_0^{\mathrm{S}}(\mathcal{P})$, is:

$$P_0^{\mathrm{S}} = \pi_0^{\mathrm{S}} = e^{-r} \mathrm{E}_0^{\mathbb{Q}} (L_1) - DPO_0^{\mathrm{S}}. \tag{55}$$

The first term represents the present value of expected future claims costs and corresponds to a default-free insurance premium. The second term is the value of default put option. This relation implies that stock insurers with a higher (lower) default risk should charge lower (higher) premiums π_0^{S}. In the absence of arbitrage, the contribution of the equityholders and policyholders in $t = 0$ will be equal to EC_0^{S} and $P_0^{\mathrm{S}} = \pi_0^{\mathrm{S}}$, respectively, implying that the purchase of each stake is associated with a net present value of zero. The insurance company then invests the sum $A_0 = EC_0^{\mathrm{S}} + \pi_0^{\mathrm{S}}$ in the capital markets.

4.2 Mutual insurer claims structure: full participation in equity payoff

Equity stake

As discussed in Section 1, one important aspect in which mutuals may differ from stock insurance companies is their potential right to demand additional premiums in times of financial distress. Provided a mutual insurer exhibits such a recovery option and its members fully participate in the payoff profile of the equity, the present value of the mutual's equity stake, EC_0^{Mf}, can be expressed as

$$EC_0^{\mathrm{Mf}} = e^{-r} \mathrm{E}_0^{\mathbb{Q}} (A_1 - L_1) + RO_0 + DPO_0^{\mathrm{M}}, \tag{56}$$

where RO_0 equals the present value of the recovery option and DPO_0^{M} denotes the present value of the default put option of the mutual insurer. Comparing Equations (53) and (56), we notice that these two option

components replace DPO_0^{S}. Due to the recovery option, the default put option of the mutual insurer ceteris paribus differs from its stock insurer counterpart (see Figure 16 for a graphical illustration). In particular, the mutual insurer remains solvent as long as the recovery option has not been fully exhausted. Accordingly, the assets in $t = 1$ have to fall under a lower default threshold $X = L_1 - C^{\mathrm{max}}$ than for the stock insurer before bankruptcy is declared and the remaining assets are distributed among those members with valid claims. C^{max} denotes the upper limit on additional payments which can be charged through the recovery option.[83] Formally, the present value of the mutual insurer's default put option, $DPO_0^{\mathrm{M}} = DPO_0^{\mathrm{M}}(\mathcal{P}, C^{\mathrm{max}})$, is defined as

$$DPO_0^{\mathrm{M}} = PO_0^{\mathrm{X}} + BPO_0 \tag{57}$$

where
$$PO_0^{\mathrm{X}} = e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left(PO_1^{\mathrm{X}}\right) = e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left[\max\left(X - A_1; 0\right)\right], \tag{58}$$

and
$$BPO_0 = e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left(C^{\mathrm{max}}\mathbf{1}_{A_1 < X}\right). \tag{59}$$

$\mathbf{1}$ is the indicator function, which equals one if $A_1 < X$ and zero otherwise. PO_0^{X} is a simple European put option with strike price X and BPO_0 is a cash-or-nothing binary put option which reflects the fact that, in the instance in which the mutual insurer becomes insolvent, the assets will have already dropped below the claims by an amount of C^{max}. In other words, in case of a mutual insurer bankruptcy, losses on the policy-holder stake will be at least C^{max}. By comparing the respective payoff profiles in Figure 16, we notice that generally $PO_0^{\mathrm{X}} \leq DPO_0^{\mathrm{M}} \leq DPO_0^{\mathrm{S}}$. In addition, the smaller C^{max}, the more valuable DPO_0^{M} and, in the special case of $C^{\mathrm{max}} = 0$ (i.e., $X = L_1$), we get $PO_0^{\mathrm{X}} = DPO_0^{\mathrm{M}} = DPO_0^{\mathrm{S}}$ and $BPO_0 = 0$.

Figure 17, depicts the payoff profile for two different specifications of the recovery option. We define the standard (basic) recovery option as

---

[83]C^{max} is usually defined in a company's charter. In our model, it can be easily adjusted to account for members' potential default risk or reluctance to pay additional premiums.
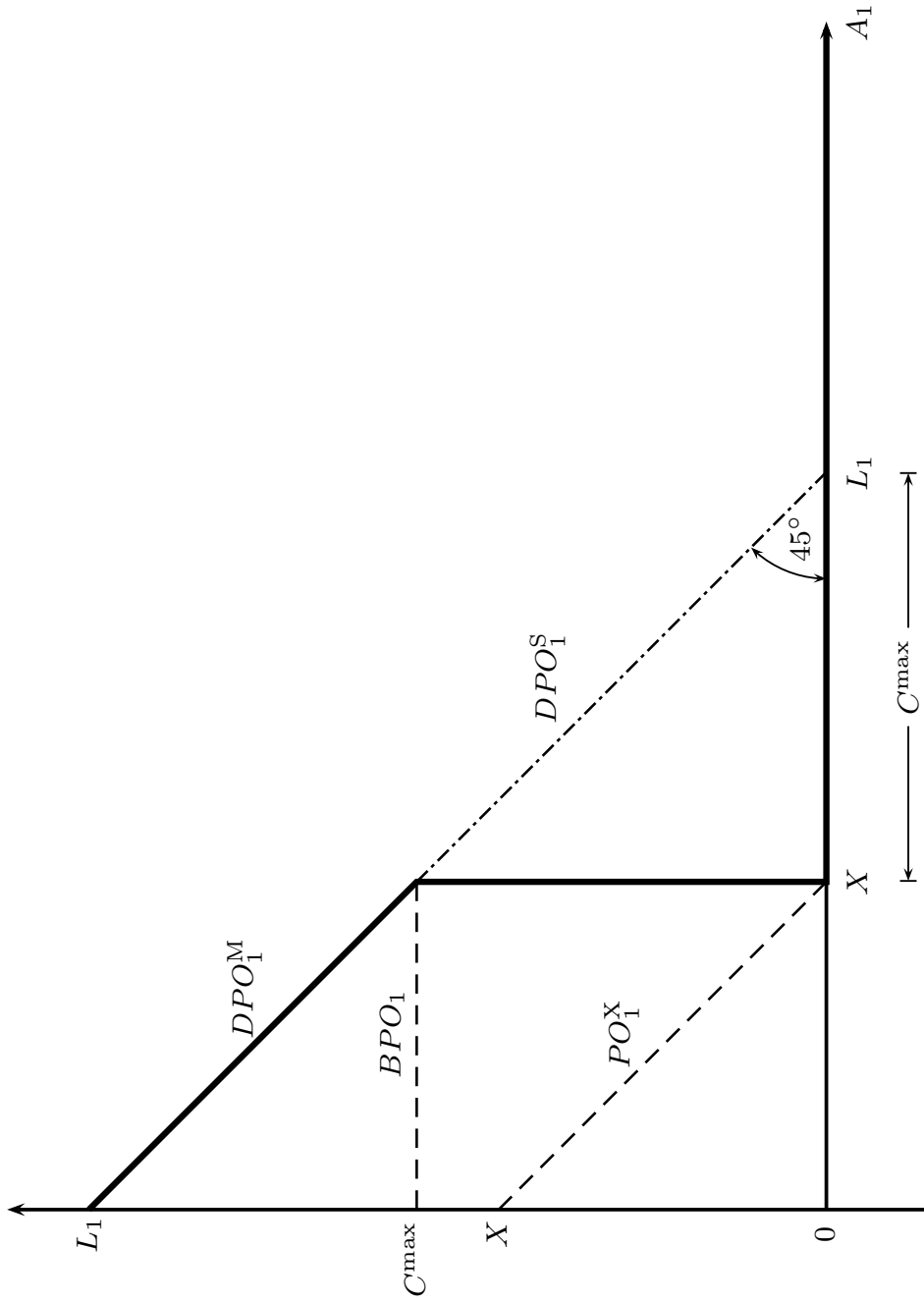
Figure 16: Mutual insurer default put option payoff in $t = 1$ (DPO_1^M, solid line). The dashed lines indicate the elements of the replicating portfolio (BPO_1 and PO_1^X). For comparison purposes, the dotdashed line illustrates the default put option of a stock insurer with an identical claim structure (DPO_1^S).
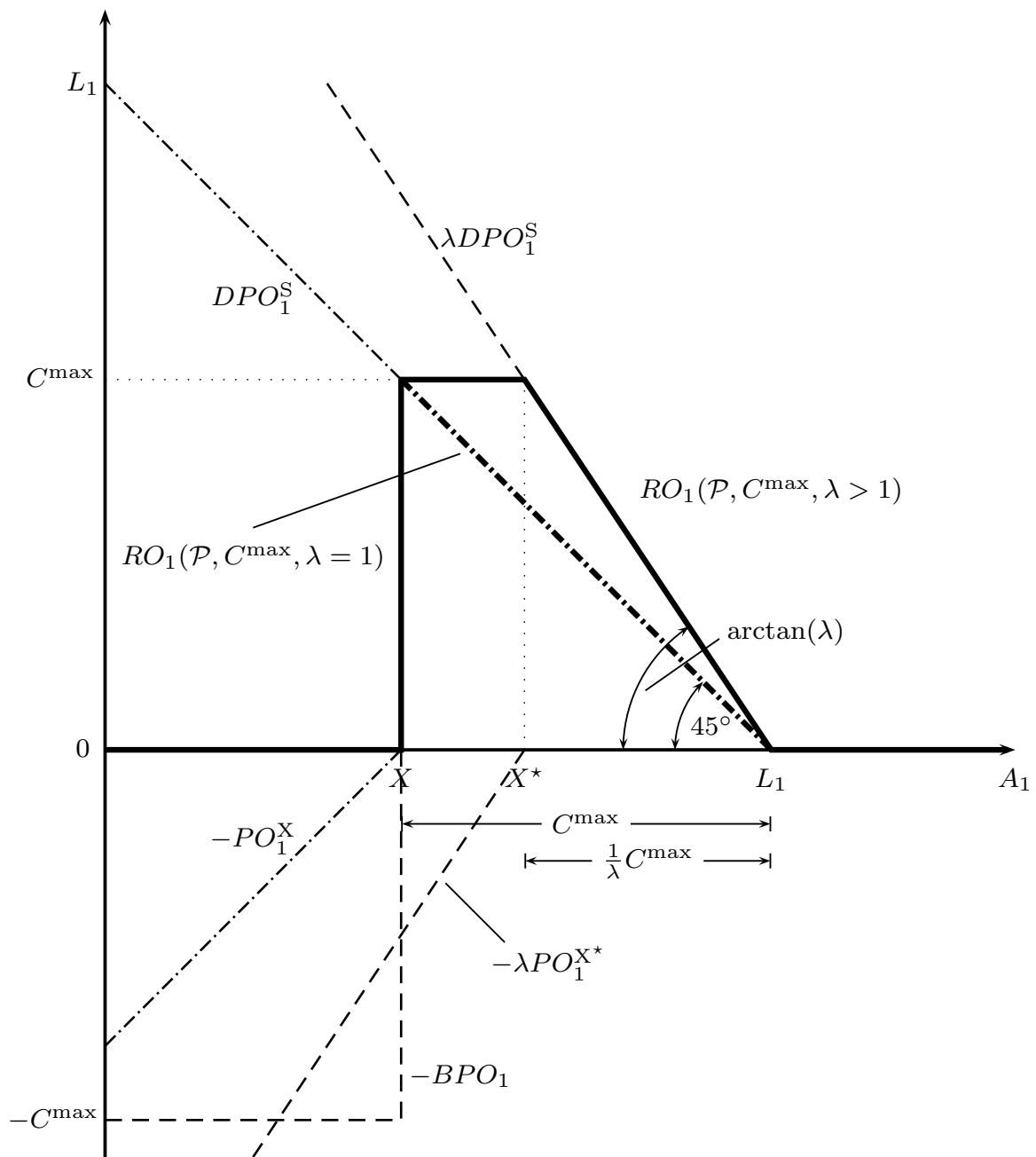
Figure 17: Mutual insurer recovery option payoff in $t = 1$: $RO_1(\mathcal{P}, C^{\max}, \lambda = 1)$ (bold dotdashed line) and $RO_1(\mathcal{P}, C^{\max}, \lambda > 1)$ (bold solid line). The thin dotdashed and dashed lines illustrate the respective replicating portfolios.

one which allows to raise no more than the exact amount of the missing capital. Its present value, $RO_0 = RO_0(\mathcal{P}, C^{\mathrm{max}})$, can be expressed as

$$
\begin{aligned}
RO_0 &= DPO_0^{\mathrm{S}} - DPO_0^{\mathrm{M}} \\
&= DPO_0^{\mathrm{S}} - PO_0^{\mathrm{X}} - BPO_0.
\end{aligned}
\tag{60}
$$

and thus equals a long position in DPO_0^{S} and a short position in DPO_0^{M}. To put it differently, instead of the stock insurer's default put option, the owners of a mutual insurer hold a combination of the recovery option and the default put option of the mutual, implying that the value DPO_0^{S} is perfectly decomposed into RO_0 and DPO_0^{M}, i.e., $DPO_0^{S} = RO_0 + DPO_0^{M}$. Consequently, the equity of the stock insurer and the equity of the mutual do not differ in value. However, ceteris paribus mutual members enjoy a higher safety level of their policies since the probability that their insurance claims in $t = 1$ are paid in full is greater than for the stock firm. Intuitively, the recovery option works as follows: when $X \leq A_1 \leq L_1$, i.e., if the assets in $t = 1$ fall below the claims by an amount less than C^{max} such that the recovery option is sufficient to rectify the deficit, $L_1 - A_1$ is demanded from policyholders. This is exactly enough additional capital to eliminate the shortage. Note that the lower C^{max}, the less valuable RO_0 and for $C^{\mathrm{max}} = 0$, RO_0 is worthless. In contrast to that, $C^{\mathrm{max}} = L_1$ is associated with the maximum value of the recovery option, while the default put option of the mutual insurer has no value in this case. Therefore, C^{max} determines how the value of the stock insurer default put option is split into DPO_0^{M} and RO_0.

Theoretically, a distressed mutual insurer might be allowed to collect more than just the missing capital from its members, implying that the firm can build up a reserve. By adjusting Equation (60) we can extend our model framework to account for this special case, which will be called excess of loss recovery option. The following is a more general expression for the present value of the recovery option, $RO_0 = RO_0(\mathcal{P}, C^{\mathrm{max}}, \lambda)$,

$$
RO_0 = \lambda DPO_0^{\mathrm{S}} - \lambda PO_0^{\mathrm{X}^{\star}} - BPO_0,
\tag{61}
$$

where

$$PO_0^{X^\star} = e^{-r} E_0^{\mathbb{Q}} \left(PO_1^{X^\star} \right) = e^{-r} E_0^{\mathbb{Q}} \left[\max \left(X^\star - A_1; 0 \right) \right], \qquad (62)$$

with $X^\star = L_1 - \frac{1}{\lambda} C^{\max}$ and $\lambda \in [1; \infty)$. Consequently, in the general case, the recovery option is a position of λ units of DPO^S long, λ units of $PO^{X^\star}$ short, and BPO short. The parameter λ constitutes a straightforward charging rule and denotes the multiple of additional payments over the deficit. For $\lambda = 2$, e.g., the mutual is able to charge policyholders twice the deficit and build up a reserve from the surplus. The impact of λ on the recovery option payoff profile is illustrated in Figure 17. Intuitively, the higher this multiple, the lower the distance between L_1 and $X^\star$, the steeper the slope of RO_1 in this interval, and the smaller the amount by which the assets have to fall below the claims so that the mutual will simply collect $C^{\max}$.[84] Analogously to PO^X, $PO^{X^\star}$ is a European put option with strike $X^\star$, which depends on λ. Clearly, if $\lambda = 1$ we have $PO_0^{X^\star} = PO_0^X$. In Figure 17, we see that $RO_1(\mathcal{P}, C^{\max}, \lambda > 1) > RO_1(\mathcal{P}, C^{\max}, \lambda = 1)$, which implies $RO_0(\mathcal{P}, C^{\max}, \lambda > 1) > RO_0(\mathcal{P}, C^{\max}, \lambda = 1)$.

To sum up, if the recovery option is designed to simply eliminate a given deficit ($\lambda = 1$), the payoff profile of the overall equity stake of a mutual insurer will ceteris paribus be the same as for the stock insurer. However, if the recovery option is specified so that the mutual can charge multiples of a given deficit ($\lambda > 1$), its equity stake will be relatively more valuable. This can be easily seen by comparing the payoff profile for the equityholders of a mutual insurer with excess of loss recovery option (Figure 18) to that plain call option shape we saw for the stock insurer in Figure 15. In any case, the value of the equity stake EC_0^{Mf} depends not only on the parameter set $\mathcal{P}$ but also on the specific characteristics of the recovery option as represented by $C^{\max}$ and λ.

---

[84] For $\lambda \to \infty$, the slightest deficit will induce the mutual insurer to charge additional payments of $C^{\max}$.
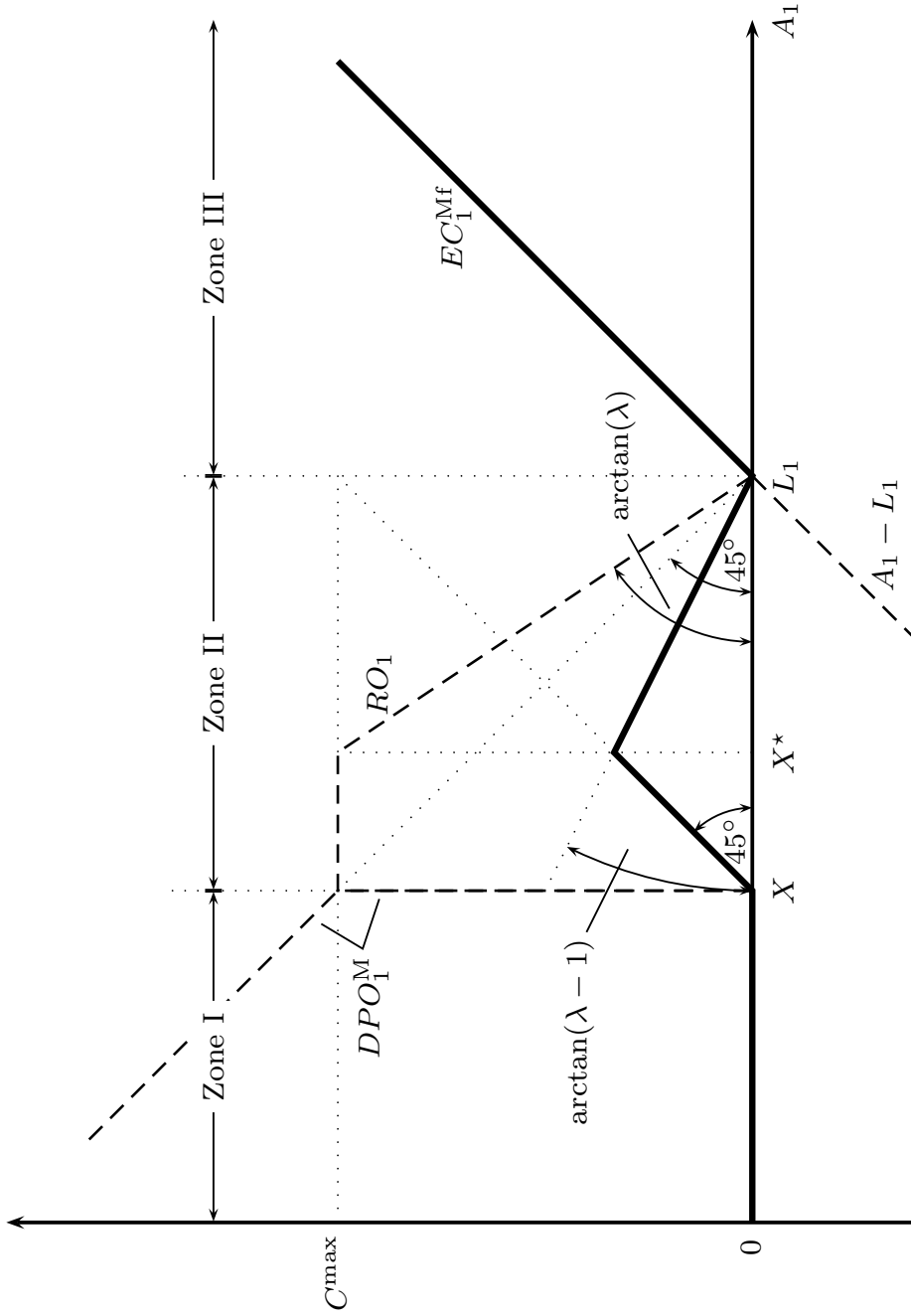
Figure 18: Mutual insurer equity payoff (full participation) in $t = 1$ (EC_1^{Mf}) for $\lambda > 1$ (solid line). The dashed lines indicate the elements of the replicating portfolio.

Policyholder stake

Consistent with its equity, we define the present value of the policyholder stake of a mutual insurer as

$$P_0^{\mathrm{M}} = e^{-r}\mathrm{E}_0^{\mathbb{Q}}(L_1) - RO_0 - DPO_0^{\mathrm{M}}. \tag{63}$$

Again, $e^{-r}\mathrm{E}_0^{\mathbb{Q}}(L_1)$ is the fair insurance premium without default risk and instead of DPO_0^{S} we have a short position in the combination of the recovery option and the default put option of the mutual. If, at the end of the period, the assets have fallen below the claims costs L_1 but not the mutual's default threshold X, i.e., $X < A_1 \leq L_1$, the policyholder stake of the mutual insurance company is associated with an equal or a higher financial loss than that of a stock insurer. This is due to the fact that the mutual charges $\lambda(L_1 - A_1)$ through the recovery option, while the insolvency of the stock insurer results in a policyholder deficit of $L_1 - A_1$. Therefore, generally $DPO_0^S \leq RO_0 + DPO_0^M$ and $P_0^M \leq P_0^S$. More specifically, the policyholder stake of a mutual insurance company is less valuable than that of an otherwise identical stock insurer when it contains a recovery option with $\lambda > 1$ (see Figure 19). Similarly, we know from Equation (56) that the present value of the equity stake increases for more expensive recovery options. Hence, an excess of loss recovery option essentially redistributes value from the policyholder to the equity stake. If $\lambda = 1$, in contrast, we have $DPO_0^S = RO_0 + DPO_0^M$ and consequently $P_0^S = P_0^M$ (refer to Equation (60)).

Arbitrage-free premium

Since, through the purchase of a policy in a mutual insurance company, one acquires the equity and the policyholder stake at the same time, the fair premium of a mutual insurer must comprise the arbitrage-free price
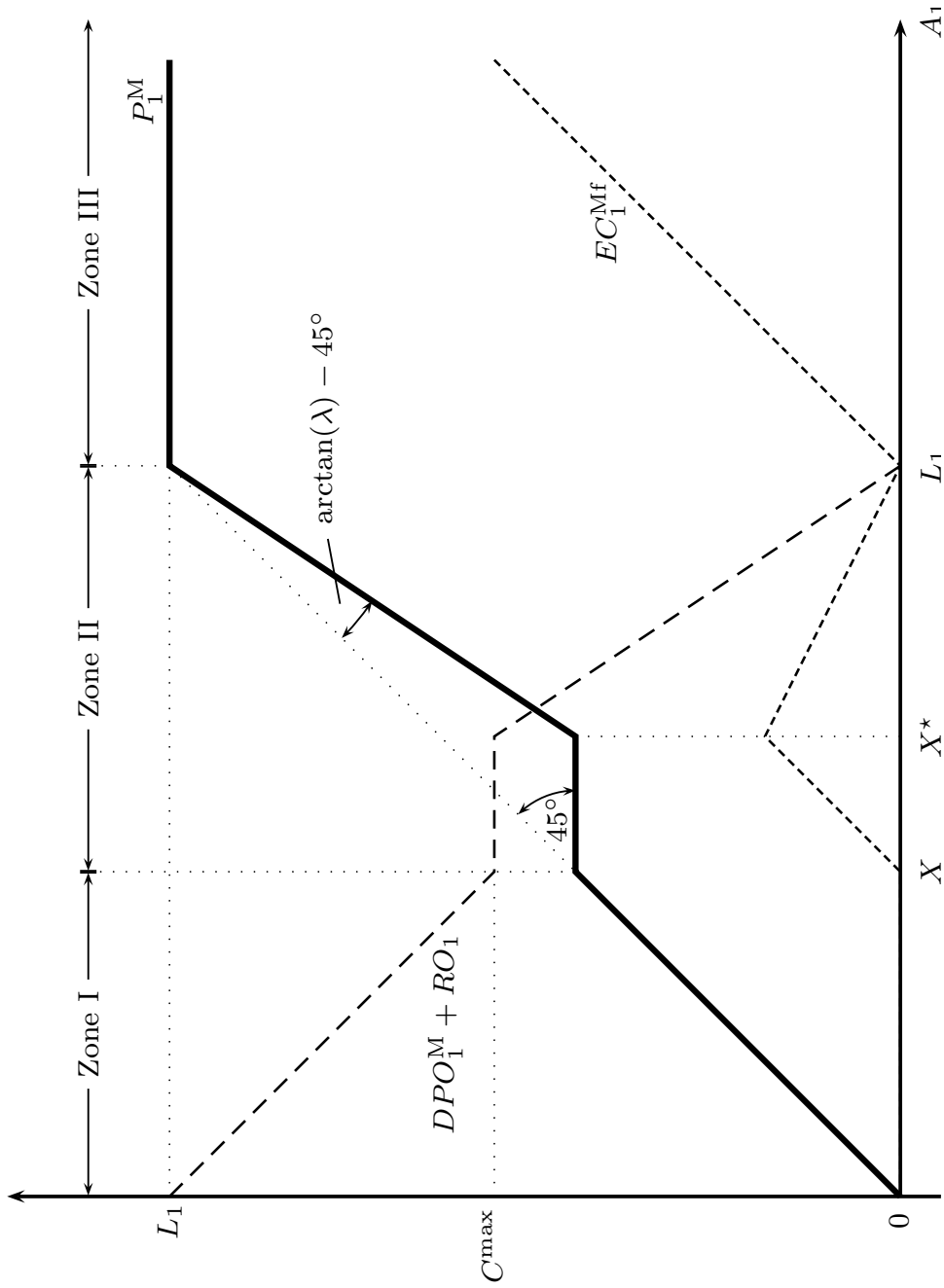
Figure 19: Mutual insurer policyholder stake payoff in $t = 1$ (P_1^{M}) for $\lambda > 1$ (solid line). The payoff profile of the default put option (DPO_1^{M}), the recovery option (RO_1), and the equity stake given full participation (EC_1^{Mf}) are drawn as dashed lines for comparison purposes.

of both components. Accordingly, we define $\Pi_0^{\mathrm{M}} = \Pi_0^{\mathrm{M}}(\mathcal{P}, C^{\mathrm{max}}, \lambda)$ as follows

$$
\begin{aligned}
\Pi_0^{\mathrm{M}} &= P_0^{\mathrm{M}} + EC_0^{\mathrm{Mf}} \\
&= \underbrace{e^{-r}\mathrm{E}_0^{\mathbb{Q}}(L_1) - RO_0 - DPO_0^{\mathrm{M}}}_{\text{policyholder stake}} \\
&\quad + \underbrace{e^{-r}\mathrm{E}_0^{\mathbb{Q}}(A_1 - L_1) + RO_0 + DPO_0^{\mathrm{M}}}_{\text{equityholder stake}} \\
&= e^{-r}\mathrm{E}_0^{\mathbb{Q}}(A_1).
\end{aligned}
\tag{64}
$$

Thus, if members fully participate in the equity payoffs, purchasing a policy from a mutual insurer is equivalent to acquiring a position the company's assets. Policyholders of an otherwise identical stock insurer additionally would have to buy the common stock of the company in order to establish the same payoff profile.

4.3   Mutual insurer claims structure: partial participation in equity payoff

Equity stake

There is generally no secondary market for ownership stakes in mutual insurance companies. As a consequence, payoffs from the equity stake of a mutual insurer and thus its present value crucially depend on the premium refund policy of the management and the ability of the members to prompt an initial public offering (IPO) or break-up of the company. Let α be the payout (premium refund) ratio and p_{L} the probability of demutualization or liquidation of the company.[85] The impact of these parameters on the payoff profile of the equity stake depends on the zones indicated in Figures 18, 19, and 20 which are determined by the realiza-

---

[85]Under agency-theoretic considerations the firm's management generally has a preference to retain as much capital in the company as possible. This aspect of the so-called owner-manager conflict lowers the premium refund ratio α. Furthermore, in contrast to a corporation, there are no blockholders in a mutual insurer. Therefore, p_{L} will depend on the members' ability to coordinate an agreement on the demutualization or liquidation of the firm.

tions of assets and claims in $t = 1$. If $A_1 < X$, i.e., the assets have
fallen below the default threshold (Zone I), the mutual is insolvent, bro-
ken up and the remaining assets are distributed to its members. Thus,
the equity stake is worthless and neither α nor p_L are relevant in Zone
I. Furthermore, if $X < A_1 < L_1$ (Zone II) the mutual insurer exercises
the recovery option to charge additional payments (via the policyholder
stake). It is safe to assume that a mutual in financial distress will re-
frain from premium refunds, implying that members can only fully re-
alize the equity payoff via an IPO or the liquidation of the company.
Hence, in Zone II only p_L has an influence on the present value of the
equity. Finally, in Zone III, where the company is solvent and does
not need to exercise the recovery option, members receive the whole eq-
uity value with probability p_L, or a premium refund of $\alpha(A_1 - L_1)$ with
probability $(1 - p_\mathrm{L})$. We summarize these two cases in the parameter
$\gamma = p_\mathrm{L} + (1 - p_\mathrm{L})\alpha$, which can be interpreted as the expected value of
the equity stake in Zone III, normalized to unity. Since $\alpha \in [0; 1]$ and
$p_\mathrm{L} \in [0; 1]$, we get $\gamma \in [0; 1]$. Under this set-up, the present value of a
mutual's equity stake in the general case (recovery option and partially
realizable equity), $EC_0^\mathrm{M} = EC_0^\mathrm{M}(\mathcal{P}, C^\mathrm{max}, \lambda, p_L, \alpha)$, can be described as
follows:

$$
\begin{aligned}
EC_0^\mathrm{M} \;=\; & \underbrace{0}_{\text{Zone I}} + \underbrace{p_\mathrm{L}\left[(\lambda - 1)DPO_0^\mathrm{S} - \lambda PO_0^{X^\star} + PO_0^X\right]}_{\text{Zone II}} \\
& + \underbrace{\gamma e^{-r}\mathrm{E}_0^\mathbb{Q}\left[\max\left(A_1 - L_1; 0\right)\right]}_{\text{Zone III}} \\
\;=\; & \gamma\left[e^{-r}\mathrm{E}_0^\mathbb{Q}\left(A_1 - L_1\right) + DPO_0^\mathrm{S}\right] \\
& + p_\mathrm{L}\left(\lambda DPO_0^\mathrm{S} - \lambda PO_0^{X^\star} + PO_0^X - DPO_0^\mathrm{S}\right) \\
\;=\; & \gamma e^{-r}\mathrm{E}_0^\mathbb{Q}\left(A_1 - L_1\right) + \gamma DPO_0^\mathrm{S} \\
& + p_\mathrm{L}\left(RO_0 + DPO_0^\mathrm{M} - DPO_0^\mathrm{S}\right) \\
\;=\; & \gamma e^{-r}\mathrm{E}_0^\mathbb{Q}\left(A_1 - L_1\right) - \left(p_\mathrm{L} - \gamma\right)DPO_0^\mathrm{S} \\
& + p_\mathrm{L}\left(RO_0 + DPO_0^\mathrm{M}\right).
\end{aligned}
\tag{65}
$$

For a graphical verification of this expression refer to Figure 20. As mentioned previously, the equity payoff in the interval $[0, X]$ (Zone I) is zero. Furthermore, the payoff profile between X and L_1, i.e., in Zone II, is characterized by an asymmetric butterfly spread, consisting of $(\lambda - 1)$ units of DPO_0^S long, λ units of $PO_0^{X^\star}$ short and one unit of PO_0^X long.[86] Finally, the equity payoff in Zone III is equal to a long stake in γ units of a simple call option on the assets with strike price L_1. Recalling Equation (53), we realize that this call option is exactly the one describing the equity value of a stock insurance company. As also illustrated in Figure 20, the consideration of the parameters p_L and γ, which were introduced above, results in a flattening of the payoff of the mutual insurer's equity stake in Zones II and III. In the absence of arbitrage, members of a mutual insurance company anticipate that they can only partially access future cash flows arising from the equity stake, implying a reduction of its present value. The difference between EC_0^{Mf} and EC_0^M – represented by the shaded area in Figure 20 – is the discount in the present value of the equity stake resulting from the incomplete participation of the current members in its future payoff. In our contingent claims framework, this "non-realizable" equity, $EC_0^{Mn} = EC_0^{Mn}(\mathcal{P}, C^{\max}, \lambda, p_L, \alpha)$, has a price in $t = 0$ equal to

$$
\begin{aligned}
EC_0^{Mn} \;=\;& EC_0^{Mf} - EC_0^M \\[2mm]
=\;& \underbrace{e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left(A_1 - L_1\right) + RO_0 + DPO_0^M}_{\text{equity given full participation}} \\[2mm]
& \left.\begin{array}{l}
- p_L\left(RO_0 + DPO_0^M\right) \\[1mm]
- \gamma e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left(A_1 - L_1\right) \\[1mm]
+ \left(p_L - \gamma\right) DPO_0^S
\end{array}\right\} \quad \text{realizable equity} \\[2mm]
=\;& \left(1 - \gamma\right) e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left(A_1 - L_1\right) + \left(p_L - \gamma\right) DPO_0^S \\[1mm]
& + \left(1 - p_L\right)\left(RO_0 + DPO_0^M\right).
\end{aligned}
\tag{66}
$$

In order to comprehensively understand the effect of p_L, α, and, in turn, γ, on the value of the equity stake of the mutual insurance company, we consider two special cases. First of all, the expression for the present

---

[86]Note that for $\lambda = 1$, we have a standard put option butterfly spread.

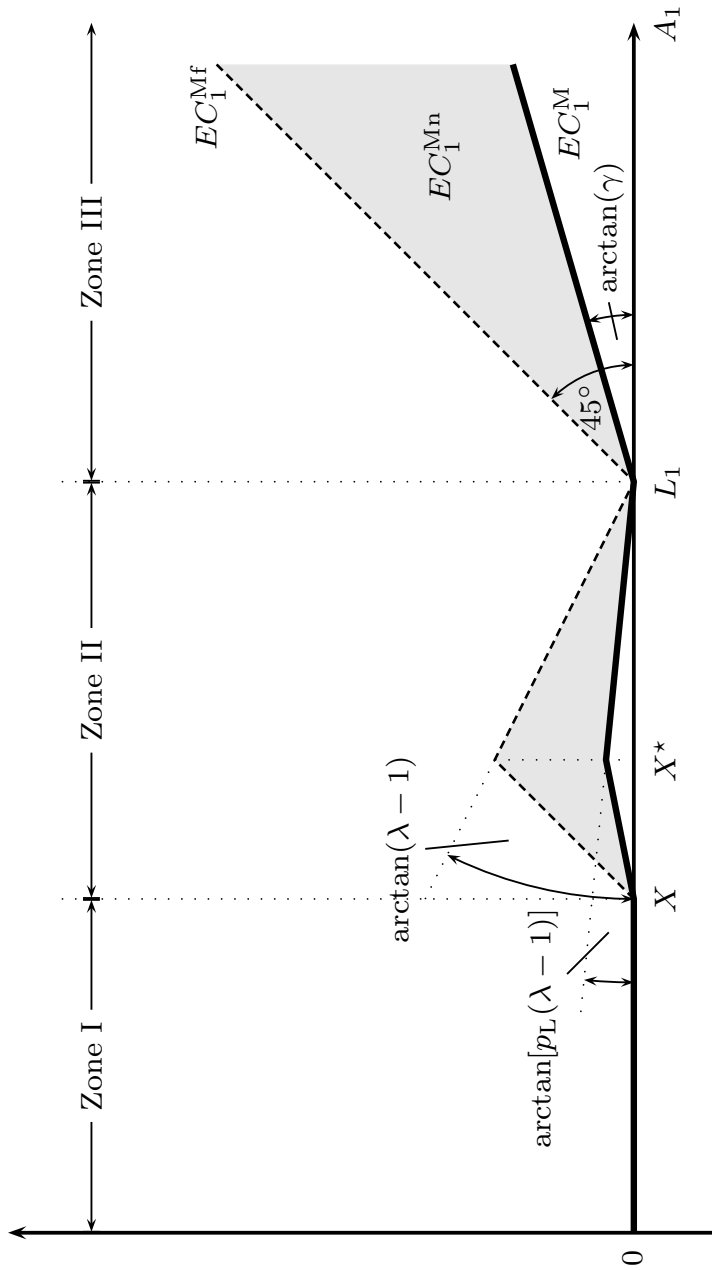Figure 20: Mutual insurer (expected) equity payoff in $t = 1$ in case of partial equity participation (EC_1^M) for $\lambda > 1$, p_L, and $\alpha < 1$ (solid line). The dashed line illustrates the equity stake given full participation (EC_1^{Mf}) for comparison purposes.

value of the equity stake under full participation, i.e., Equation (56), is nested in the more general Equation (65). The exact payoff profile we saw in Figure 18 (EC_1^{Mf}) in the previous section can only be realized in the special case of $p_L = 1$, i.e., full participation in the equity payoff stream in all three zones. To see this note that $p_L = 1$ directly results in $\gamma = 1$ such that Equation (65) becomes Equation (56) and Equation (66) collapses to zero: there is no non-realizable equity component. Apart from that, $p_L < 1$ and $\alpha = 1$ also results in $\gamma = 1$: the mutual distributes the whole equity to its members when it is solvent, but whenever the recovery option is exercised, there are no premium refunds and participation in the equity payoff is contingent on the probability of liquidation p_L. Consequently, the first term in Equation (66) disappears and the remainder reduces to $(1 - p_{\mathrm{L}}) \left(RO_0 + DPO_0^{\mathrm{M}} - DPO^{\mathrm{S}} \right)$. This means that only the excess value of the recovery and default put option of a mutual over the default put option of a stock insurer constitutes non-realizable equity.[87] While, in this case, the payoff profile in Zone III is the same as for full equity participation, we get a flatter curve in Zone II. Overall, in the arbitrage-free setting, realizable equity EC_0^{M} and non-realizable equity EC_0^{Mn} will always sum up to EC_0^{Mf}, while p_L and α govern the size of these components relative to each other.

Policyholder stake

While participation in the future cash flows of the equity stake might be limited, there are no such restrictions associated with the policyholder stake. In other words, the present value of the policyholder stake remains the same as in Section 4.2, comprising the default-free premium $e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left(L_1\right)$ as well as a short position in the recovery and the default put option of the mutual. Thus, we have the same expression as in Equation (63):

$$P_0^{\mathrm{M}} = e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left(L_1\right) - RO_0 - DPO_0^{\mathrm{M}}. \tag{67}$$

---

[87] In case there is no such excess value, i.e., for $\lambda = 1$, the non-realizable equity is zero and the equity stake of the mutual equals that of the stock insurer.

Arbitrage-free premium

As explained at the end of Section 4.2, the arbitrage-free premium of a mutual insurer must comprise the present values of both the equity and the policyholder stake. In the general case of partial participation in the equity payoff, however, the price of the equity stake splits into a realizable and a non-realizable component. Yet, the overall level of the mutual insurance company premium remains unchanged and equals the expected discounted value of the firm's assets in $t = 1$. Consequently, in the general case, we replace Equation (64) with an alternative expression for $\Pi_0^{\mathrm{M}} = \Pi_0^{\mathrm{M}}(\mathcal{P}, C^{\max}, \lambda)$:

$$
\begin{aligned}
\Pi_0^{\mathrm{M}} \;=\;& P_0^{\mathrm{M}} + EC_0^{\mathrm{M}} + EC_0^{\mathrm{Mn}} \\
\;=\;& \underbrace{e^{-r}\mathrm{E}_0^{\mathbb{Q}}(L_1) - RO_0 - DPO_0^{\mathrm{M}}}_{\text{policyholder stake}} \\
& \left.\begin{aligned}
&+ \gamma e^{-r}\mathrm{E}_0^{\mathbb{Q}}(A_1 - L_1) \\
&- (p_{\mathrm{L}} - \gamma)\, DPO_0^{\mathrm{S}} \\
&+ p_{\mathrm{L}}\left(RO_0 + DPO_0^{\mathrm{M}}\right)
\end{aligned}\right\} \quad \text{realizable equity stake} \\
& \left.\begin{aligned}
&+ (1 - \gamma)\, e^{-r}\mathrm{E}_0^{\mathbb{Q}}(A_1 - L_1) \\
&+ (p_{\mathrm{L}} - \gamma)\, DPO_0^{\mathrm{S}} \\
&+ (1 - p_{\mathrm{L}})\left(RO_0 + DPO_0^{\mathrm{M}}\right)
\end{aligned}\right\} \quad \text{non-realizable equity} \\
\;=\;& e^{-r}\mathrm{E}_0^{\mathbb{Q}}(A_1). && (68)
\end{aligned}
$$

4.4   Claims structure relationships

Below we briefly illustrate the theoretical impact of recovery option and limited participation in equity payoffs on the premium of a mutual insurer relative to a comparable stock insurer. Imagine two insurance firms with the exact same underlying assets and claims: one is founded as a corporation and the other one adopts the legal form of a mutual. Figure 21 depicts the general relationship between the claim structures of these two companies in four distinct cases, characterized by different configurations of recovery option and equity participation. As defined in Equation (55), the marginal premium charged by a stock insurer equals
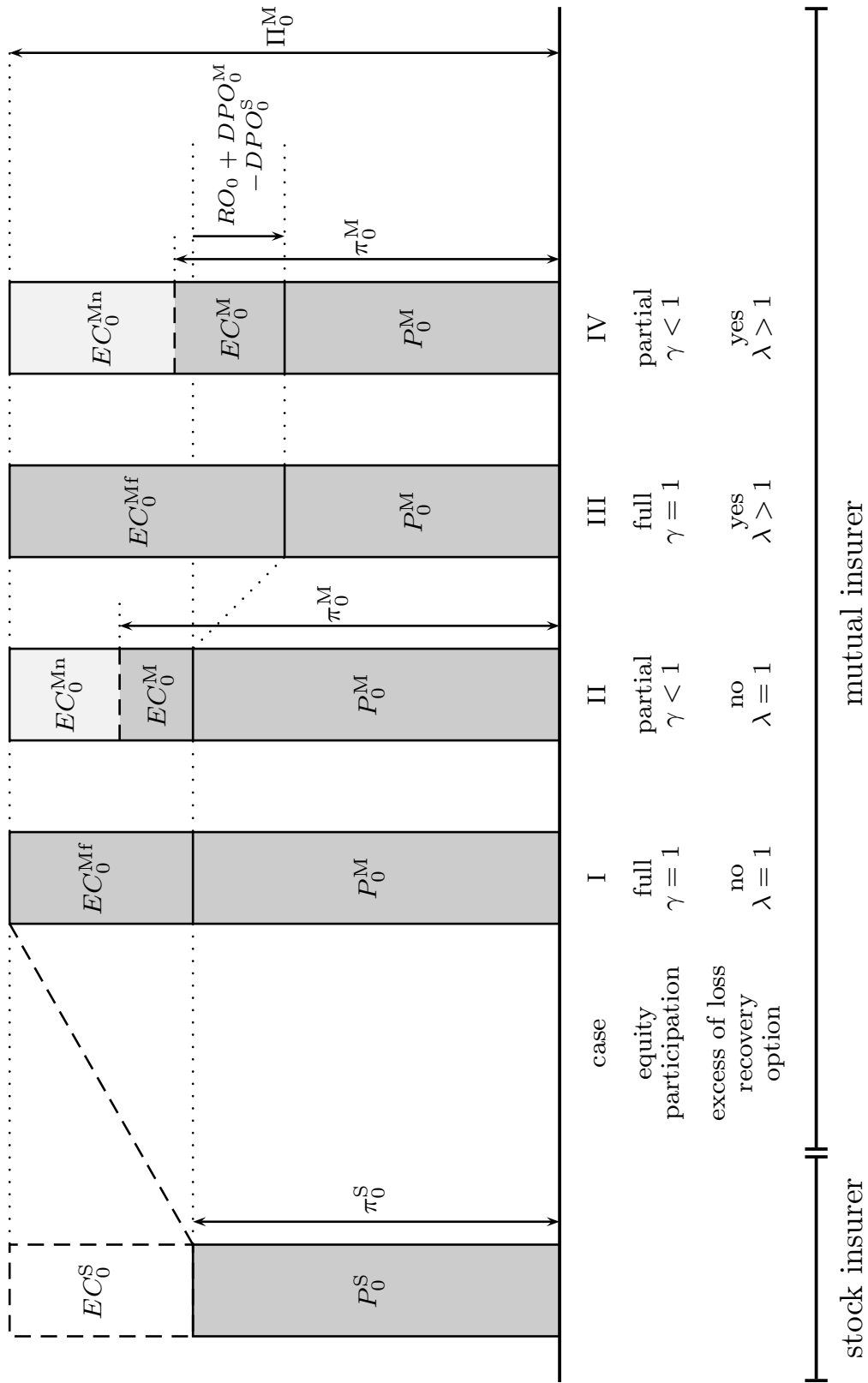
Figure 21: Comparison of premiums

the value of its policyholder stake. The premium of the mutual insurer corresponding to this reference case, however, depends on the appointed setting.

In Case I, the mutual insurance company is either not allowed to charge additional premiums at all (i.e., $C^{\max} = 0$, which results in $DPO_0^{\mathrm{S}} = DPO_0^{\mathrm{M}}$)[88] or the amount of additional premiums is restricted to the actual deficit $L_1 - A_1$ (i.e., $\lambda = 1$, which results in $DPO_0^{\mathrm{S}} = DPO_0^{\mathrm{M}} + RO_0$)[89]. In addition, the equity stake of the mutual is fully realizable ($p_{\mathrm{L}} = 1$ and, hence, $\gamma = 1$). Comparing Equation (55) and (63), we see that under these circumstances $P_0^{\mathrm{S}} = P_0^{\mathrm{M}}$: there is no difference between the value of the policyholder stakes of a stock and a mutual insurer. Moreover, comparing Equation (53) and (56), we notice that both equity stakes have the same value, i.e., $EC_0^{\mathrm{S}} = EC_0^{\mathrm{Mf}}$. Due to the fact that the equity of the mutual can be entirely realized and there are no additional contributions in excess of a loss, the rights of mutual members are economically identical to those of the combined policyholder and ownership stake of the stock insurer. In other words, since policyholders and owners coincide, the position in a mutual could be replicated by simply purchasing both an insurance contract and an appropriate amount of shares of the stock insurer. Hence, the aggregate premium Π_0^{M} charged by a mutual should equal the premium of a stock insurer, π_0^{S}, plus the value of its equity EC_0^{S}.

In Case II, the mutual insurer's company charter excludes additional premiums in excess of a loss ($\lambda = 1$). However, its equity stake cannot be fully realized ($\gamma < 1$). Since, in an arbitrage-free market, rational individuals anticipate this, the ownership stake of the mutual insurer is separated into a realizable and a non-realizable component and prospective mutual members are generally not willing to provide the latter. Consequently, the mutual premium is now π_0^{M}, i.e., Π_0^{M} net of the non-realizable equity EC_0^{Mn}. The full premium Π_0^{M} can only be demanded if members are being compensated for EC_0^{Mn}, e.g., through a binding right to payments from future policyholders upon the beginning

---

[88] Refer to Equations (57) to (60).
[89] Refer to Equation (60).

of their membership in the mutual.[90] It is important to note that, in any case, the non-realizable equity needs to be paid in for the company to be founded at all. This is due to the fact that less initial equity than EC_0^{Mf} is associated with a lower expected payoff in $t = 1$. In anticipation of this consequence, individuals will further reduce their willingness to pay, eventually reaching an equilibrium where the value of both stakes is zero. In this situation, the mutual insurance company cannot find customers if it charges a positive arbitrage-free premium, since every insurance policy would be associated with a negative net present value. Thus, if the members do not provide EC_0^{Mn}, an external third party such as a founding capital provider, whose capital repayment is contractually guaranteed, would need to step in instead.

Case III represents the claims structure of the mutual if its equity is fully realizable ($p_{\text{L}} = 1$ and, hence, $\gamma = 1$) and its recovery option allows to charge additional premiums over and above the actual loss ($\lambda > 1$). Due to the excess of loss recovery option, the value of the equity position increases and the value of the policyholder position decreases compared to the stock insurer (and Case I) by an amount equal to the difference between $RO_0 + DPO_0^{\text{M}}$ and DPO_0^{S}. The higher λ, the bigger the shift between both stakes. Since the recovery option solely redistributes value between the stakes, the overall amount of assets within the company is unchanged. Therefore, the overall mutual premium remains equal to Π_0^{M}.

Finally, the combined effect of partially realizable equity ($\gamma < 1$) and excess of loss recovery option ($\lambda > 1$) is illustrated in Case IV. Again, the equity stake splits into EC_0^{Mn} plus EC_0^{M} and π_0^{M} denotes the full mutual premium less the present value of the non-realizable equity. In contrast to Case II, however, both equity components are slightly more expensive, since value is shifted from the policyholder to the equity stake via the excess of loss recovery option. As before, a non-zero arbitrage-free solution can only be achieved if the non-realizable equity is paid in

---

[90]In a multiperiod framework such compensation payments could be conducted at the end of each period. For instance, the current members (from $t = 0$ to $t = 1$) would need to receive the right to be paid an amount of EC_1^{Mn} in $t = 1$ by the members of the following period ($t = 1$ to $t = 2$). This right is worth $EC_0^{\text{Mn}} = e^{-r}\text{E}_0^{\mathbb{Q}}\left(EC_1^{\text{Mn}}\right)$ today.

as well. Hence, more expensive recovery options (for a higher λ), are ceteris paribus associated with a more valuable (non-realizable) equity and a lower π_0^M.

5   Numerical analysis

In this section, we concretely describe assets and claims costs as continuous-time stochastic processes and present closed-form solutions for the various option prices on which our model framework is based. Subsequently, we provide a brief numerical example to further illustrate the model mechanics as well as the effect of recovery options and equity participation on the premium of a mutual insurer. In addition, based on the numerical implementation of our model framework, we derive normative insights with regard to feasible combinations of premium, safety level, and capital structure of stock and mutual insurance companies.

5.1   Option pricing formulae

Suppose that assets are traded continuously in time and that the term structure of interest rates is flat and deterministic. The insurance companies' assets are assumed to be stochastic and their dynamics are modeled by the following Geometric Brownian Motion under the risk-neutral measure $\mathbb{Q}$:

$$\frac{\mathrm{d}A_t}{A_t} = r\mathrm{d}t + \sigma_\mathrm{A}\mathrm{d}W_{\mathrm{A}t}^{\mathbb{Q}}, \tag{69}$$

where the drift is given by the risk-free interest rate r, σ_A denotes the volatility of the assets, and $\mathrm{d}W_{At}$ is a standard Wiener process under $\mathbb{Q}$.[91] The insurer's claims are assumed to be deterministic: $L_0 = e^{-r}L_1$.[92]

---

[91] In this set-up, asset returns are normally distributed. While, in most cases, this is merely an approximation of the empirically observed distributions (see, e.g., Officer, 1972; Akgiray and Booth, 1988; Lau et al., 1990), it simplifies matters by allowing us to apply closed-form solutions. Since insurance companies tend to hold a considerable fraction of bonds in their investment portfolios, an alternative set-up could include term structure models (see, e.g., Vasicek, 1977; Cox et al., 1981).

[92] This decision is made for reasons of computational simplicity. Since the model framework in Section 4 has been deliberately kept on a general level, different assumptions for the asset and claims dynamics as well as associated option-pricing frameworks can be applied without loss of generality.

Under these assumptions, closed-form solutions for the present values of the various European options described in Section 4 are available (see Black and Scholes, 1973). In line with the one-period model from Section 4, the present value of the stock insurer default put option struck at L_1 (DPO_0^S) can be computed as follows:

$$
\begin{aligned}
DPO_0^S &= e^{-r}E_0^{\mathbb{Q}}\left(DPO_1^S\right) = e^{-r}E_0^{\mathbb{Q}}\left[\max\left(L_1 - A_1; 0\right)\right] \\
&= e^{-r}L_1\Phi(-d_1) - A_0\Phi(-d_2),
\end{aligned}
\tag{70}
$$

where $\Phi(x)$ is the cumulative distribution function of the standard normal distribution and

$$
d_1 = \frac{\ln(A_0/L_1) + r - \sigma_A^2/2}{\sigma_A},
$$

$$
d_2 = \frac{\ln(A_0/L_1) + r + \sigma_A^2/2}{\sigma_A}.
$$

In addition, the present value of the put option PO_0^X in Equation (58), which is one of the two building blocks of the default put option of a mutual insurer (DPO^M), can be calculated using the following formula:

$$
\begin{aligned}
PO_0^X &= e^{-r}E_0^{\mathbb{Q}}\left(PO_1^X\right) = e^{-r}E_0^{\mathbb{Q}}\left[\max\left(X - A_1; 0\right)\right] \\
&= e^{-r}X\Phi(x_1) - A_0\Phi(x_2) \\
&= e^{-r}(L_1 - C^{\max})\Phi(-x_1) - A_0\Phi(-x_2),
\end{aligned}
\tag{71}
$$

where

$$
x_1 = \frac{\ln\left[A_0/(L_1 - C^{\max})\right] + r - \sigma_A^2/2}{\sigma_A},
$$

$$
x_2 = \frac{\ln\left[A_0/(L_1 - C^{\max})\right] + r + \sigma_A^2/2}{\sigma_A}.
$$

The second building block of the DPO^M is a cash-or-nothing binary put option which pays $C^{\max}$ if $A_1 < X$ and zero otherwise. Rubinstein

and Reiner (1991) show that the price of this binary put option is equal
to:

$$BPO_0 = e^{-r}C^{\max}\Phi(-x_1). \tag{72}$$

Using Equations (71) and (72), the formula for the present value of
the default put option of the mutual insurer (DPO_0^{M}) can be derived:

$$
\begin{aligned}
DPO_0^{\mathrm{M}} &= PO_0^{\mathrm{X}} + BPO_0 \\
&= e^{-r}(L_1 - C^{\max})\Phi(-x_1) - A_0\Phi(-x_2) + e^{-r}C^{\max}\Phi(-x_1) \\
&= e^{-r}L_1\Phi(-x_1) - A_0\Phi(-x_2). \tag{73}
\end{aligned}
$$

This formula somehow resembles Equation (70), which describes the
price of the default put option of a stock insurer. Yet, the probabili-
ties with which the parameters $e^{-r}L_1$ and A_0 are weighted differ. To
grasp the intuition behind this, recall from Section 4.2 that the assets
A_1 have to fall below the threshold X before the default put option of
the mutual insurer is in the money. Contingent on $A_1 < X$, however,
the payoff profiles of DPO^{M} and DPO^{S} are congruent (refer back to
Figure 16): in the area $A_1 < X$, both options pay $L_1 - A_1$. As a result,
the formula for DPO_0^{M} includes $e^{-r}L_1$ and A_0, but weighted with the
probabilities $\Phi(-x_1)$ and $\Phi(-x_2)$ instead of $\Phi(-d_1)$ and $\Phi(-d_2)$.

Finally, to calculate the value of the recovery option in the general
case (i.e., for $\lambda > 1$), we additionally need the closed-form solution for
the put option $PO_0^{\mathrm{X}^\star}$. Following the same rationale as above, we get

$$
\begin{aligned}
PO_0^{\mathrm{X}^\star} &= e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left(PO_1^{\mathrm{X}^\star}\right) = e^{-r}\mathrm{E}_0^{\mathbb{Q}}\left[\max\left(X^\star - A_1; 0\right)\right] \\
&= e^{-r}X^\star\Phi(z_1) - A_0\Phi(z_2) \\
&= e^{-r}(L_1 - \frac{1}{\lambda}C^{\max})\Phi(-z_1) - A_0\Phi(-z_2), \tag{74}
\end{aligned}
$$

with

$$z_1 = \frac{\ln\left[A_0/(L_1 - \frac{1}{\lambda}C^{\max})\right] + r - \sigma_{\mathrm{A}}^2/2}{\sigma_{\mathrm{A}}},$$

$$z_2 = \frac{\ln\left[A_0/(L_1 - \frac{1}{\lambda}C^{\max})\right] + r + \sigma_{\mathrm{A}}^2/2}{\sigma_{\mathrm{A}}}.$$

Combining Equations (70), (72), and (74), the value of the recovery option can be expressed as:[93]

$$
\begin{aligned}
RO_0 &= \lambda DPO_0^{\mathrm{S}} - \lambda PO_0^{\mathrm{X}^\star} - BPO_0, \\
&= \lambda\left[e^{-r}L_1\Phi(-d_1) - A_0\Phi(-d_2)\right] \\
&\quad - \lambda[e^{-r}(L_1 - \frac{1}{\lambda}C^{\max})\Phi(-z_1) - A_0\Phi(-z_2)] \\
&\quad - e^{-r}C^{\max}\Phi(-x_1) \\
&= \lambda e^{-r}L_1\Phi(-d_1) - \lambda A_0\Phi(-d_2) - \lambda e^{-r}L_1\Phi(-z_1) \\
&\quad + \lambda A_0\Phi(-z_2) + e^{-r}C^{\max}\Phi(-z_1) - e^{-r}C^{\max}\Phi(-x_1) \\
&= \lambda\left\{e^{-r}L_1\left[\Phi(-d_1) - \Phi(-z_1)\right] - A_0\left[\Phi(-d_2) - \Phi(-z_2)\right]\right\} \\
&\quad + e^{-r}C^{\max}\left[\Phi(-z_1) - \Phi(-x_1)\right]. \quad\quad (75)
\end{aligned}
$$

5.2 The impact of recovery option and participation in equity payoff

Having determined asset and claims dynamics as well as the associated option pricing formulae, the equity and policyholder stake of mutual and stock insurance companies can now be valued. Table 15 contains the basic input parameters used in our numerical examples and the results for the stock insurer.

The first three columns of Table 16 illustrate the impact of the recovery option in a mutual insurance company with full participation in the equity payoff stream ($p_L = 1$ and $\alpha = 1$).[94] For $\lambda = 1$, i.e., no excess of loss recovery option, the value of the default put option of the stock insurer DPO_0^{S} (0.2481) perfectly splits into RO_0 (0.2463) and DPO_0^{M} (0.0018). In addition, equity EC_0^{Mf} (30.2481) and policyholder stake P_0^{M} (69.7519) of the mutual are worth the same as those of the stock insurer shown in Table 15: although the two companies differ in terms of legal form, they are economically identical in this case. Since, through a membership in the mutual, one acquires both stakes, the mutual premium ($\Pi_0^{\mathrm{M}} = 30.2481 + 69.7519 = 100$) equals the present value of the

---

[93]Note that $[\Phi(-d_1) - \Phi(-z_1)]$ is $\Pr(X^\star < A_1 < L_1)$ and $[\Phi(-z_1) - \Phi(-x_1)]$ is $\Pr(X < A_1 < X^\star)$.

[94]These numerical results correspond to Case I and III in Section 4.4.

A_0	100	initial value of the assets
$L_0 = e^{-r}L_1$	70	initial value of the liabilities
σ_A	0.20	volatility of the asset returns
r	0.03	risk free rate
DPO_0^{S}	0.2481	present value of the default put option of the stock insurer
EC_0^{S}	30.2481	present value of the equity of the stock insurer
$P_0^{\mathrm{S}} = \pi_0^{\mathrm{S}}$	69.7519	present value of the policyholder claims in the stock insurer

Table 15: Input parameters and resulting values for DPO_0^{S}, EC_0^{S}, and P_0^{S}

assets. For an increasing λ, however, we observe a non-linear growth in RO_0, resulting in a value of 0.2708 in case 110 percent of a deficit can be demanded from mutual members, i.e., $\lambda = 1.1$. In this case, the sum $RO_0 + DPO_0^{\mathrm{M}}$ (0.2726) is almost ten percent higher than DPO_0^{S} (0.2481). Furthermore, in Table 16 we see that the excess of loss recovery option redistributes value from the policyholder to the equityholder stake since P_0^{M} falls and EC_0^{Mf}, EC_0^{Mn}, as well as EC_0^{M} rise in λ. Analogously to the default put option of the stock insurer, DPO_0^{M} measures the safety level of a mutual insurance company's policyholder stake. As DPO_0^{M} remains the same (0.0018) for all values of λ and is always lower than DPO_0^{S} (0.2481), the mutual insurer with recovery option has a higher safety level than the otherwise identical stock insurer.

The three columns in the center of Table 16 show the case where mutual members partially participate in the equity payoff ($p_L = 0.1$ and $\alpha = 0.1$).[95] Again, for $\lambda = 1$, we have $RO_0 + DPO_0^{\mathrm{M}} = DPO_0^{\mathrm{S}} = 0.2481$. This time, however, the total equity value EC_0^{Mf} (30.2481) splits into a realizable component EC_0^{M} (5.7471) and a non-realizable component EC_0^{Mn} (24.5010). The former is considerably lower than the latter, since the figures are based on a fairly low premium refund rate and probability of liquidation. As in the previous case, a rise in λ implies a more

---

[95]See Case II and IV in Section 4.4.

λ	$p_L = 1, \alpha = 1$			$p_L = 0.1, \alpha = 0.1$			$p_L = 0, \alpha = 0$		
	1.00	1.05	1.10	1.00	1.05	1.10	1.00	1.05	1.10
DPO_0^M	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018	0.0018
RO_0	0.2463	0.2586	0.2708	0.2463	0.2586	0.2708	0.2463	0.2586	0.2708
$DPO_0^M + RO_0$	0.2481	0.2604	0.2726	0.2481	0.2604	0.2726	0.2481	0.2604	0.2726
EC_0^M	30.2481	30.2604	30.2726	5.7471	5.7484	5.7496	0.0000	0.0000	0.0000
EC_0^{Mn}	0.0000	0.0000	0.0000	24.5010	24.5120	24.5230	30.2481	30.2604	30.2726
EC_0^{Mf}	30.2481	30.2604	30.2726	30.2481	30.2604	30.2726	30.2481	30.2604	30.2726
P_0^M	69.7519	69.7396	69.7274	69.7519	69.7396	69.7274	69.7519	69.7396	69.7274
Π_0^M	100.0000	100.0000	100.0000	100.0000	100.0000	100.0000	100.0000	100.0000	100.0000

Table 16: Impact of the excess of loss recovery option and equity participation ($C^{max} = 25$)

expensive recovery option. The associated value redistribution reduces P_0^M and increases both components of the equity stake. Consistent with our choice of p_L and α, however, EC_0^{Mn} absorbs a relatively larger share. Although the ten percent likelihood of liquidation assumed for this numerical example probably has to be considered relatively high from a real world perspective, the realizable equity stake has already become quite small. Consequently, even lower values for p_L, which are perfectly conceivable, would result in a situation where virtually the whole equity is attributed to the non-realizable component. As explained in Section 4.4, the capital in such a case would need to be provided by a third party, since, under the arbitrage-free framework applied, mutual members would not be prepared to incur a negative net present value investment. Again, $DPO_0^M = 0.0018 < DPO_0^S = 0.2481$ for all λ. Therefore, as in the previous example, the mutual insurer's policyholder stake exhibits a higher safety level than that of the stock insurer. The last three columns of Table 16 contain the numerical results when the equity stake is not realizable at all ($p_L = 0$ and $\alpha = 0$). Obviously, in this case, the whole equity value is attributed to the non-realizable component.

Apart from λ, the maximum amount of additional premiums $C^{\max}$ is a key determinant of the recovery option value and has a direct impact on the safety level of the firm. Table 17 illustrates that a recovery option does not exist if $C^{\max} = 0$.[96] Instead, the default put option of the mutual insurance company is exactly the same as for a stock insurer ($DPO_0^M = DPO_0^S = 0.2481$). The higher $C^{\max}$, i.e., the less binding the upper limit on additional payments, the more valuable becomes the recovery option. In addition, an increase in $C^{\max}$ simultaneously results in a decline of DPO_0^M, implying an improving safety level. For $C^{\max} = 40$, we get $RO_0 = 0.2726$ and the mutual's default put option is (almost) worthless because the value of the assets in $t = 1$ would have to drop by more than 40 below the value of the claims for it to be in the money. As in Table 16, the decrease in P_0^M due to the incremental growth in RO_0 is counterbalanced by an increased value of the equity stake (realizable and non-realizable component).

---

[96]See Figure 17 in Section 4.2.

C^{max}	0	10	20	30	40
DPO_0^{M}	0.2481	0.1016	0.0103	0.0002	0.0000
RO_0	0.0000	0.1604	0.2612	0.2727	0.2729
$DPO_0^{\mathrm{M}} + RO_0$	0.2481	0.2620	0.2715	0.2729	0.2729
EC_0^{M}	5.7471	5.7485	5.7495	5.7496	5.7496
EC_0^{Mn}	24.5010	24.5135	24.5221	24.5233	24.5233
EC_0^{Mf}	30.2481	30.2620	30.2716	30.2729	30.2729
P_0^{M}	69.7519	69.7380	69.7284	69.7271	69.7271
Π_0^{M}	100.0000	100.0000	100.0000	100.0000	100.0000

Table 17: Impact of the maximal amount of additional contributions (C^{max}) given partial participation of members in the equity payoffs of the mutual firm: $p_L = 0.1$, $\alpha = 0.1$, $\lambda = 1.1$

5.3 Stock vs. mutual insurers: premium, safety level, and equity capital

In the following, we compare a stock and a mutual insurance company with identical underlying assets and claims with regard to the three central magnitudes premium size, safety level, and equity capital, considering cases with and without recovery option as well as full and partial participation in equity payoffs. Again, the calculations have been based on the parameter values in Table 15. While other configurations would change the magnitude of the observed effects, their direction remains the same.

We begin with the case where the mutual insurer does not have a recovery option ($C^{\mathrm{max}} = 0$) and its equity stake can be fully realized by the members. In Figure 22, the arbitrage-free mutual and stock insurer premiums have been plotted against the value of the respective equity stakes. Under the arbitrage-free framework used, both curves must start at zero. Let us first look at the solid curve, which represents equity-premium-combinations for the stock insurer and equity-policyholder-stake-combinations for the mutual insurer. As the amount of initial equity capital is raised, the stock insurer premium converges
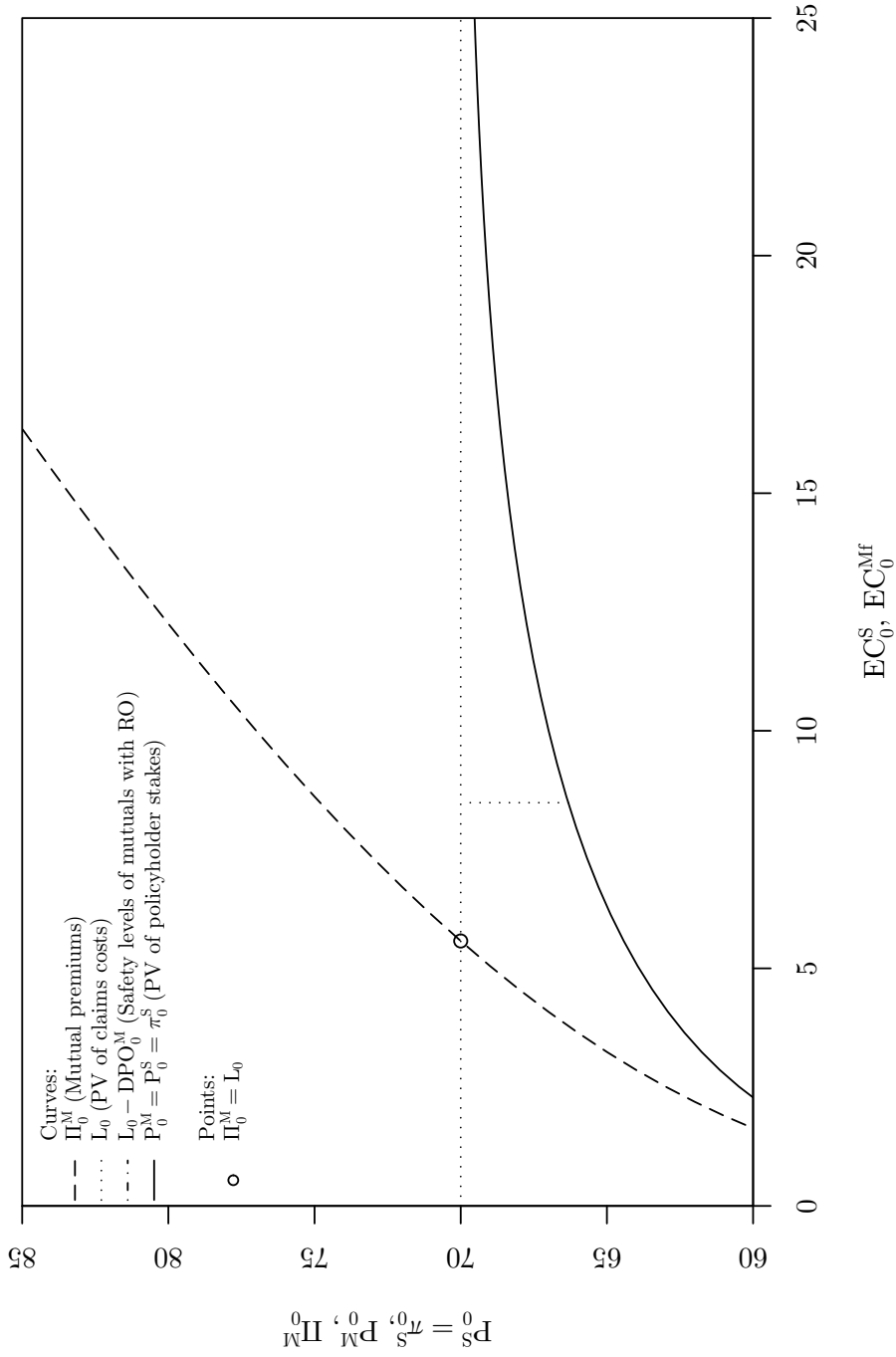
Figure 22: Premium comparison for full participation in equity payoffs and no recovery option

towards the present value of the claims costs L_0 (represented by a dotted horizontal line).[97] For any amount of equity capital, the distance between L_0 and the solid curve equals the present value of the stock insurer's default put option (DPO_0^S), which, in this case, is identical to that of the mutual insurer (DPO_0^M) since it is assumed that the latter does not have a recovery option. The vertical dotted line is meant to serve as a concrete example. As a consequence, if they are identically capitalized, mutual and stock insurer offer contracts with the same safety level. In addition, more equity capital is associated with a decline in DPO_0^S $(= DPO_0^M)$ due to the fact that a larger equity buffer reduces the likelihood of the assets dropping below the claims costs at the end of the period. The dashed curve represents premiums of the mutual insurer. Since members have to purchase both stakes, it lies strictly above the solid curve. Thus, in the absence of a recovery option, if both companies hold the same amount of equity capital and members of the mutual insurer can fully participate in its equity stake, then they should be charged higher premiums than the policyholders of the stock insurer. Another relevant observation is related to the point where the Π_0^M-curve intersects the L_0-line (marked by a small circle). If the mutual insurer holds more initial equity capital than associated with this point, its premium must be strictly higher than that of the stock insurer, no matter how well capitalized the latter is. This is due to the fact that the π_0^S-curve converges to but never exceeds L_0.

Next, we introduce a basic recovery option $(C^{\max} > 0, \lambda = 1)$, while still allowing for full participation in the equity payoffs of the mutual insurer. As discussed in Section 4.2, the recovery option enables mutual insurers to stay solvent and satisfy all claims, even if their equity capital is fully exhausted. More specifically, a mutual insurer is bankrupt only if the deficit of assets relative to liabilities exceeds the limit on additional premiums $(C^{\max})$, which implies $DPO_0^S = RO_0 + DPO_0^M$ or $DPO_0^M < DPO_0^S$. This is illustrated in Figure 23, where we now have an additional dotdashed curve, reflecting the safety levels of the mutual

---

[97]In Section 4.1 we explained that the fair stock insurance premium equals the present value of the policyholder stake, i.e., $\pi^S = P_0^S$. Besides, L_0 is the default-free premium.

insurer. While DPO_0^S is still represented by the distance from L_0 to the solid curve, the distance between L_0 and the dotdashed curve equals DPO_0^M. Since the dotdashed lies strictly above the solid curve, the mutual insurer with recovery option exhibits a strictly better safety level than the identically capitalized stock insurer. In other words, the mutual insurer with recovery option needs less equity capital to achieve the same safety level as the stock insurer.[98] Furthermore, analogously to Figure 22, the mutual must charge a higher premium than the stock insurer if it holds more equity capital than associated with the intersection of the Π_0^M-curve and the L_0-line. Consequently, safety level and premium of a well-capitalized insurance company should be higher if it adopts the legal form of a mutual. In contrast to the results in Figure 22, however, we now find capitalizations for which the premium of the mutual can be equal to or lower than that of the stock insurer with an identical safety level. To see this, we focus on the intersection between the dashed (Π_0^M) and dotdashed curve ($L_0 - DPO_0^M$), which has been highlighted by a black dot. If the mutual insurance company holds precisely this much equity capital, it exhibits the same safety level and charges the same premium as the stock insurer with the amount of equity capital which corresponds to the black triangle.[99] Right of the black dot, the mutual charges more and left of the black dot it charges less than the stock insurer with the same safety level.

In Figure 24, we account for limited participation in the equity payoff stream of the mutual insurance company by splitting its capital into the realizable and the non-realizable component.[100] However, as explained in Section 4.4, both components need to be paid in for the company to be able to begin business. Therefore, the x-axis is still based on the full value of the mutual's equity EC_0^{Mf} and the dotdashed curve, representing safety levels of the mutual insurer, is unaffected by this change. In contrast to that, however, the non-realizable equity is excluded from the

---

[98]Intuitively, the recovery option can be interpreted as an equity substitute. Hence, in most jurisdictions mutual insurers can—to some extent—account for their recovery option when calculating solvency capital charges.

[99]To find the latter, follow an imaginary horizontal line from the black dot to the right until it reaches the π_0^S-curve.

[100]The non-realizable equity is calculated based on $p_L = 0.1$ and $\alpha = 0.1$.
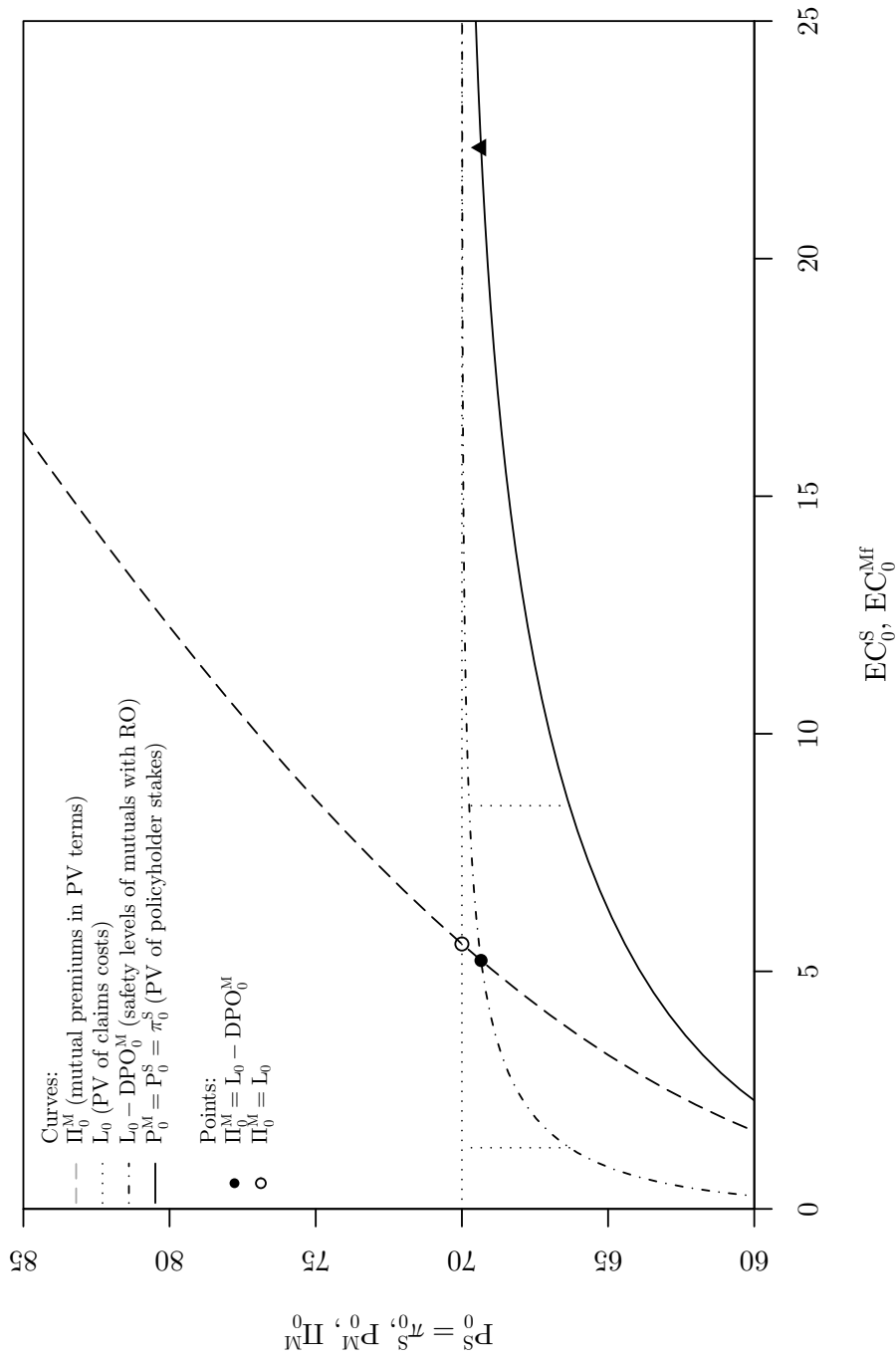
Figure 23: Premium comparison for full participation in equity payoffs and recovery option
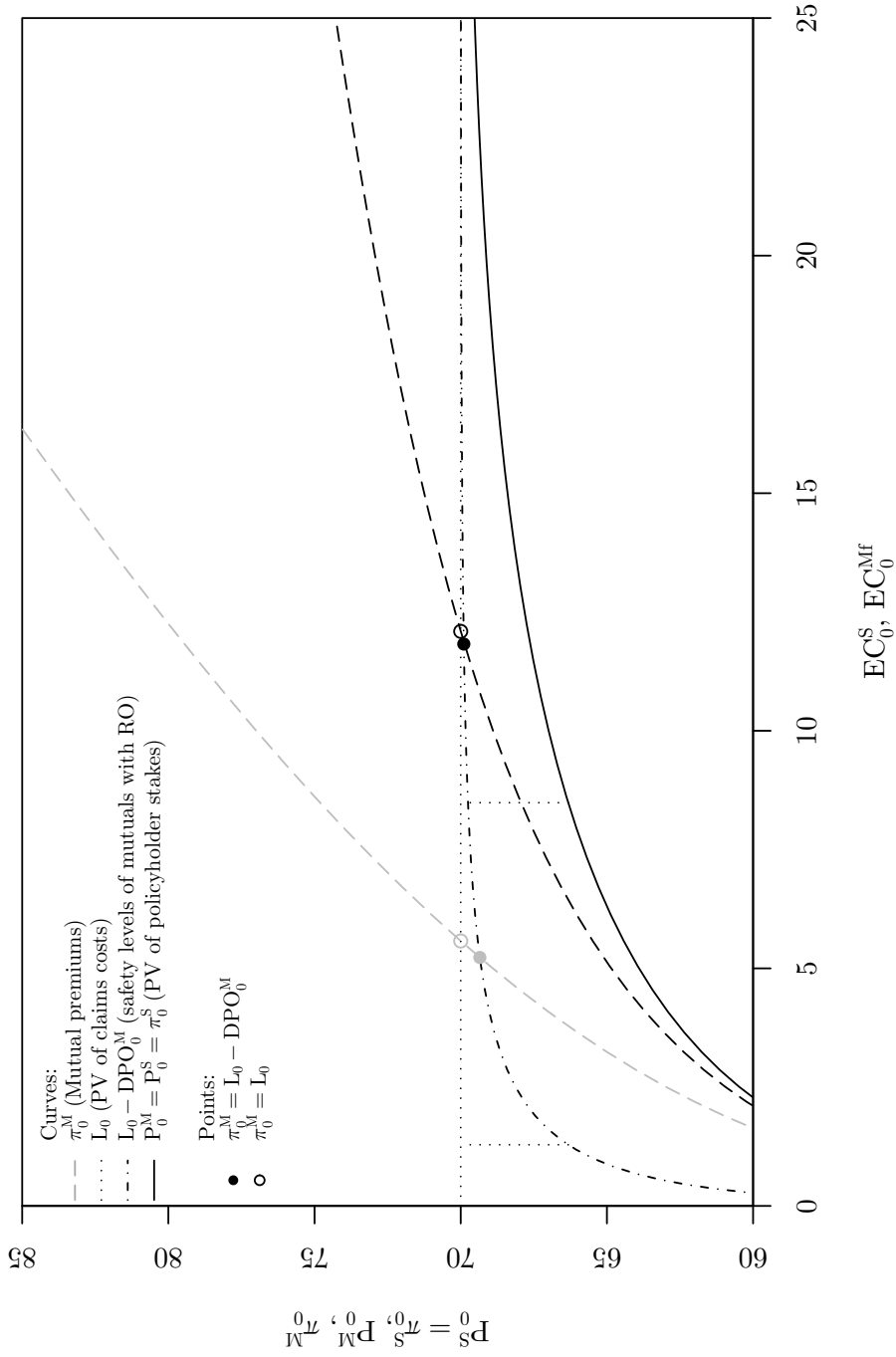
Figure 24: Premium comparison for partial participation in equity payoffs and recovery option

mutual premium, meaning that π_0^M instead of Π_0^M is shown on the y-axis. The intuition behind this proceeding is that members are either compensated by an amount equal to the present value of the non-realizable equity, or the latter is provided by a third party, e.g., a founding capital provider. Since, for each amount of initial equity capital, the mutual premium is now lower than in the case of full equity participation (Figures 22 and 23), the π_0^M-curve has a smaller slope than the Π_0^M-curve (plotted in light grey). Hence, for a decreasing probability of liquidation p_L, the premiums of the mutual insurance company converge to those of the identically capitalized stock insurer as the non-realizable equity is not borne by the members. Besides, the dashed curve (Π_0^M) now intersects the dotted line (L_0) further to the right such that we have a broader range of capitalizations, which allow the mutual to match the premium of the stock insurer. Similarly, the intersection between the dashed (π_0^M) and the dotdashed curve ($L_0 - DPO_0^M$) has been shifted to the right, implying a larger set of capital structures of the stock insurer for which the mutual is able to provide less expensive policies with the same safety level.[101]

6   Economic implications

Due to competition in insurance markets one might expect the premiums of stock and mutual companies not to differ significantly (see, e.g., Mayers and Smith, 1988). This view is partially supported by the empirical evidence we presented in Section 3. Despite the different results for two common estimators we were able to conclude that, in any case, mutuals do not charge higher premiums than stock firms. If at all, it seems that stock insurer policies are more expensive. We can now combine these empirical findings with the normative results from the previous section to derive economic implications with regard to the relationship of stock and mutual insurer premiums. To begin with, we sum up un-

---

[101] The black triangle, representing that particular capitalization of the stock insurer for which the mutual can chose to match both its safety level and its premium, is now outside the scale of Figure 24.

der which specific circumstances the contingent claims model framework
supports the equality of premiums.

First of all, in the absence of a recovery option and if members can
fully participate in a mutual insurer's equity payoffs, its premium can
only be similar to that of a stock insurer when its capitalization and
safety level are very low (Figure 22). However, such a scenario is un-
likely to occur in practice since, in most jurisdictions, solvency regulation
frameworks ensure a minimum safety level for insurance companies.[102]
If the equity of a mutual insurer without recovery option is only par-
tially realizable, i.e., its premium curve in Figure 22 becomes flatter,
there are more likely to be capitalizations which allow the mutual to
charge the same or a lower premium than the stock insurer, while still
conforming to the applicable solvency standards. By holding less equity
capital than the stock insurer, however, the mutual would also maintain
a comparatively lower safety level.

Secondly, even in the presence of a recovery option, the mutual com-
pany with fully realizable equity is only able to match or undercut the
prices of the stock insurer when featuring less initial equity capital. Yet,
despite the generally smaller equity buffer, the mutual's safety level
could be lower than, similar to, or even higher than that of the stock in-
surer with the same premium, depending on its specific capitalization.[103]
Again, the practical relevance of this scenario depends on the lower limit
for the safety level as established by the applicable solvency regulation.
However, due to the recovery option, it is less likely that all capital struc-
tures which enable mutuals to charge less than stock insurers are ruled
out.

Finally, reconsider the situation where the equity of mutuals with
recovery option is only partially realizable (Figure 24). As before, of-
fering policies for the same or a lower premium than the stock insurer
requires that the mutual commands less equity. However, in this case
it will be more likely that the mutual also complies with the respective

---

[102]Within our model framework, a minimum safety level is equivalent to an upper
limit on the present value of the default put option. Hence, it could be reflected in
the Figures 22 to 24 by means of a vertical line, the area to the left of which would
not be admissible under the prevailing solvency standards.

[103]Consider the area around the black dot in Figure 23.

solvency standards since for any given capital structure, a larger fraction of non-realizable equity is associated with a lower mutual premium.[104] As mentioned in Section 1, there are generally no liquid secondary markets for ownership stakes of mutuals. Consequently, the non-realizable equity should be rather large, leading us to believe that this might the most relevant case from a practical perspective.

To sum up, while our arbitrage-free model does not generally exclude the possibility of the mutual premium being lower than the stock insurer premium, in any case, such an outcome would require the mutual to hold less equity capital than the stock insurer. Within the empirical analysis, however, we explicitly controlled for capital structure effects as well as other premium determinants such as underwriting risk and administration costs. Thus, it appears that the empirically observed prices are not arbitrage-free in the sense of the applied contingent claims approach. In other words, from a normative perspective, policies offered by stock insurers seem to be overpriced relative to policies of mutuals. Since this situation is not a theoretical equilibrium, it can only prevail due to further factors which are exogenous to our model. One such aspect might be that we consider stakes in present value terms while observed mutual premiums are quoted as up-front cash flows, i.e., net of the recovery option value which can be viewed as an ex-post premium component. However, due to its rather low value compared to the overall mutual premium (see numerical analysis in Section 5), it is safe to assume that the recovery option has a minor impact on the results. Moreover, the deviation from the theoretical premium relationship could be caused by superior marketing and sales efforts of stock companies. Although this might be a valid reason for the persistence of economic rents, its impact is difficult to assess in the absence of specific empirical work on the subject.

Another point to be taken into account is that asymmetric information can be an important issue in insurance markets. As explained in Section 4.4, the notion of perfectly informed individuals which underlies our contingent claims context implies that a mutual insurance company

---

[104]Recall that, relative to Figure 23, the black dot and triangle in Figure 24 are shifted to the right.

would not be able to attract customers if its premium includes the present value of non-realizable equity. Yet, in a situation where prospective mutual members are unaware of economic differences associated with the legal form of insurance companies, they are unable to correctly assess the value of a policy. Therefore, the deviation from our arbitrage-free results might occur because mutual members do not have enough information or are not financially literate enough to determine the fair price of both stakes included in the mutual premium. Asymmetric information could lead to a scenario in which individuals actually pay for all or part of the non-realizable equity without being compensated in some form. Evidently, this would imply a transfer of wealth to an unknown group of future profiteers such as, e.g., a generation of policyholders which participates in the liquidation or demutualization of the firm. However, such a violation of the no-arbitrage condition does not need to be recurring. Since most of the mutual insurance companies in our sample are rather old and well-capitalized firms (see Table 18), wealth transfers could have taken place in the past. Some of the affected individuals might have already left the company without adequate compensation. Current members benefit from this development as an accumulation of equity reserves through violations of the no-arbitrage condition in the past would imply that mutuals are now able to offer policies for a lower premium than stock insurers. Alternatively, wealth transfers could also persist between the policyholders and owners of the stock insurance companies, implying that the former overpay for their contracts. Finally, a combination of these sorts of wealth transfers within stock and mutual organizations is conceivable.

7   Conclusion

In this paper, we empirically and theoretically analyze the relationship between the insurance premium of stock and mutual companies. Evaluating panel data for the German motor liability insurance sector, we do not find evidence that mutuals charge significantly higher premiums than stock insurers. If at all, it seems that stock insurer policies are more expensive. Subsequently, we employ a comprehensive model framework

for the arbitrage-free pricing of stock and mutual insurance contracts. Based on a numerical implementation of our model, we then compare stock and mutual insurance companies with regard to the three central magnitudes premium size, safety level, and equity capital. Although we identify certain circumstances under which the mutual's premium should be equal to or smaller than the stock insurer's, these situations would generally require the mutual to hold less capital than the stock insurer. This being inconsistent with our empirical results, it appears that policies offered by stock insurers are overpriced relative to policies of mutuals.

Although various reasons for the observed deviation of our empirical and theoretical results are conceivable, we believe a violation of the no-arbitrage principle due to asymmetric information to be the most plausible explanation. Therefore, we argue that the documented discrepancies are an indicator for likely wealth transfers between different stakeholder groups of mutual and stock companies. A more detailed identification of the size and direction of these wealth transfers could be an interesting avenue for future research. Since such an analysis would need to be based on a separate consideration of the different stakes, our contingent claims model framework is well suited for an application in this context. On the empirical side, however, more detailed insurance company information would be required. Another interesting research question centers around the coexistence of stock and mutual insurance companies. Our normative results could be a starting point for a further consideration of this topic. As previously discussed, an arbitrage-free market implies that rational individuals would not be willing to pay for the non-realizable component of the equity stake. Hence, we suggested that mutual companies can only come into existence if, e.g., their initial members are granted the right to compensation payments for the non-realizable equity by future member generations or if a third party acts as founding capital provider. Since both alternatives are rarely observed in practice, it would be interesting to explore other possibilities which enable mutuals to coexist with stock companies.

References

Akgiray, V. and Booth, G. G. (1988). The Stable-Law Model of Stock Returns. *Journal of Business and Economic Statistics*, 6(1):51–57.

Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, 59(3):817–858.

Baltagi, B. H. (2005). *Econometric Analysis of Panel Data*. Wiley, 3rd edition.

Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654.

Breusch, T., Ward, M. B., Nguyen, H., and Kompas, T. (2010). On the Fixed-Effects Vector Decomposition. *Political Analysis*, (forthcoming).

Carson, J. M., Forster, M. D., and McNamara, M. J. (1998). Changes in Ownership Structure: Theory and Evidence from Life Insurer Demutualizations. *Journal of Insurance Issues*, 21(1):1–22.

Cass, D., Chichilnisky, G., and Wu, H.-M. (1996). Individual Risk and Mutual Insurance. *Econometrica*, 64(2):333 – 341.

Coase, R. H. (1960). The Problem of Social Cost. *Journal of Law and Economics*, 3(1):1–44.

Cox, J. C., Ingersoll, J. E. J., and Ross, S. A. (1981). A Re-Examination of Traditional Hypotheses about the Term Structure of Interest Rates. *Journal of Finance*, 36(4):769–799.

Cummins, J. D. and Danzon, P. M. (1997). Price, Financial Quality, and Capital Flows in Insurance Markets. *Journal of Financial Intermediation*, 6(1):3–38.

Cummins, J. D., Rubio-Misas, M., and Zi, H. (2004). The Effect of Organizational Structure on Efficiency: Evidence from the Spanish Insurance Industry. *Journal of Banking and Finance*, 28(12):3113–3150.

Cummins, J. D., Weiss, M. A., and Zi, H. (1999). Organizational Form and Efficiency: The Coexistence of Stock and Mutual Property-Liability Insurers. *Management Science*, 45(9):1254–1269.

Doherty, N. A. and Garven, J. R. (1986). Price Regulation in Property-Liability Insurance: A Contingent-Claims Approach. *Journal of Finance*, 41(5):1031–1050.

Driscoll, J. C. and Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics*, 80(4):549–560.

Erhemjamts, O. and Leverty, J. T. (2010). The Demise of the Mutual Organizational Form: An Investigation of the Life Insurance Industry. *Journal of Money, Credit and Banking*, (forthcoming).

Fama, E. F. and Jensen, M. C. (1983a). Agency Problems and Residual Claims. *Journal of Law and Economics*, 26(2):327–349.

Fama, E. F. and Jensen, M. C. (1983b). Separation of Ownership and Control. *Journal of Law and Economics*, 26(2):301–325.

Fitzgerald, J. F. (1973). Demutualization of Mutual Property and Liability Insurers. *Journal of Risk and Insurance*, 40(4):575–584.

Fletcher, L. P. (1966). Motivations Underlying the Mutualization of Stock Life Insurance Companies. *Journal of Risk and Insurance*, 33(1):19–32.

Friesen, G. (2007). On the Lack of Participating Policy Usage by Stock Insurance Companies. *Journal of Risk and Insurance*, 74(1):185–206.

Garven, J. R. (1987). On the Application of Finance Theory to the Insurance Firm. *Journal of Financial Services Research*, 1(1):57–76.

Garven, J. R. and Pottier, S. W. (1995). Incentive Contracting and the Role of Participation Rights in Stock Insurers. *Journal of Risk and Insurance*, 62(2):253–270.

Gouriéroux, C., Holly, A., and Monfort, A. (1982). Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters. *Econometrica*, 50(1):63–80.

Greene, M. R. and Johnson, R. E. (1980). Stocks vs Mutuals: Who Controls? *Journal of Risk and Insurance*, 47(1):165–174.

Greene, W. H. (2007). *Econometric Analysis*. Prentice Hall, 6th edition.

Greene, W. H. (2010). Fixed Effects Vector Decomposition: A Magical Solution to the Problem of Time Invariant Variables in Fixed Effects Models? *Working Paper, Stern School of Business New York University.*

Gron, A. (1994). Capacity Constraints and Cycles in Property-Casualty Insurance Markets. *RAND Journal of Economics*, 25(1):110–127.

Gropper, D. M. and Hudson, C. D. (2003). A Note on Savings and Loan Ownership Structure and Expense Preference: A Re-Examination. *Journal of Banking and Finance*, 27(10):2003–2014.

Hansmann, H. (1985). The Organization of Insurance Companies: Mutual versus Stock. *Journal of Law, Economics, and Organization*, 1(1):125–153.

Harrington, S. E. and Niehaus, G. (2002). Capital Structure Decisions in the Insurance Industry: Stocks versus Mutuals. *Journal of Financial Services Research*, 21(1):145–163.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6):1251–1271.

Hausman, J. A. and Taylor, W. E. (1981). Panel Data and Unobservable Individual Effects. *Econometrica*, 49(6):1377–1398.

He, E. and Sommer, D. W. (2010). Separation of Ownership and Control: Implications for Board Composition. *Journal of Risk and Insurance*, 77(2):265–295.

Hsiao, C. (2002). *Analysis of Panel Data.* Cambridge University Press, 2nd edition.

Iannotta, G., Nocera, G., and Sironi, A. (2007). Ownership Structure, Risk and Performance in the European Banking Industry. *Journal of Banking and Finance*, 31(7):2127–2149.

Jeng, V., Lai, G. C., and McNamara, M. J. (2007). Efficiency and Demutualization: Evidence From the U.S. Life Insurance Industry in the 1980s and 1990s. *Journal of Risk and Insurance*, 74(3):683–711.

Jensen, M. C. and Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics*, 3(4):305–360.

Lai, G. C., McNamara, M. J., and Yu, T. (2008). The Wealth Effect of Demutualization: Evidence from the U.S. Property-Liability and Life Insurance Industries. *Journal of Risk and Insurance*, 75(1):125–144.

Lamm-Tennant, J. and Starks, L. T. (1993). Stock Versus Mutual Ownership Structures: The Risk Implications. *Journal of Business*, 66(1):29–46.

Lau, A. H.-L., Lau, H.-S., and Wingender, J. R. (1990). The Distribution of Stock Returns: New Evidence against the Stable Model. *Journal of Business and Economic Statistics*, 8(2):217–223.

Laux, C. and Muermann, A. (2010). Financing Risk Transfer under Governance Problems: Mutual versus Stock Insurers. *Journal of Financial Intermediation*, 19(3):333–354.

Lee, S.-J., Mayers, D., and Smith, C. W. (1997). Guaranty Funds and Risk-Taking Evidence from the Insurance Industry. *Journal of Financial Economics*, 44(1):3–24.

Ligon, J. A. and Thistle, P. D. (2005). The Formation of Mutual Insurers in Markets with Adverse Selection. *Journal of Business*, 78(2):529–555.

Mayers, D. and Smith, C. W. (1981). Contractual Provisions, Organizational Structure, and Conflict Control in Insurance Markets. *Journal of Business*, 54(3):407–434.

Mayers, D. and Smith, C. W. (1986). Ownership Structure and Control: The Mutualization of Stock Life Insurance Companies. *Journal of Financial Economics*, 16(1):73–98.

Mayers, D. and Smith, C. W. (1988). Ownership Structure across Lines of Property-Casualty Insurance. *Journal of Law and Economics*, 31(2):351–378.

Mayers, D. and Smith, C. W. (1994). Managerial Discretion, Regulation, and Stock Insurer Ownership Structure. *Journal of Risk and Insurance*, 61(4):638–655.

Mayers, D. and Smith, C. W. (2005). Agency Problems and the Corporate Charter. *Journal of Law, Economics, and Organization*, 21(2):417–440.

McNamara, M. J. and Rhee, S. G. (1992). Ownership Structure and Performance: The Demutualization of Life Insurers. *Journal of Risk and Insurance*, 59(2):221–238.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, 29(2):449–470.

Officer, R. R. (1972). The Distribution of Stock Returns. *Journal of the American Statistical Association*, 67(340):807–812.

Petersen, M. A. (2008). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *Review of Financial Studies*, 22(1):435–480.

Phillips, R. D., Cummins, J. D., and Lin, Y. (2006). Capital Allocation and the Pricing of Financially Intermediated Risks: An Empirical Investigation. *Working Paper*.

Plumper, T. and Troeger, V. E. (2007). Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects. *Political Analysis*, 15(2):124–139.

Pottier, S. W. and Sommer, D. W. (1997). Agency Theory and Life Insurer Ownership Structure. *Journal of Risk and Insurance*, 64(3):529–543.

Rasmusen, E. (1988). Mutual Banks and Stock Banks. *Journal of Law and Economics*, 31(2):395–421.

Rubinstein, M. and Reiner, E. (1991). Unscrambling the Binary Code. *Risk*, 4(9):75–83.

Smith, B. D. and Stutzer, M. J. (1990). Adverse Selection, Aggregate Uncertainty, and the Role for Mutual Insurance Contracts. *Journal of Business*, 63(4):493–510.

Smith, B. D. and Stutzer, M. J. (1995). A Theory of Mutual Formation and Moral Hazard with Evidence from the History of the Insurance Industry. *Review of Financial Studies*, 8(2):545–577.

Spiller, R. (1972). Ownership and Performance: Stock and Mutual Life Insurance Companies. *Journal of Risk and Insurance*, 39(1):17–25.

Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5(2):177–188.

Verbeek, M. (2008). *A Guide to Modern Econometrics.* Wiley, 3rd edition.

Viswanathan, K. S. (2006). The Pricing of Insurer Demutualization Initial Public Offerings. *Journal of Risk and Insurance*, 73(3):439–468.

Viswanathan, K. S. and Cummins, J. D. (2003). Ownership Structure Changes in the Insurance Industry: An Analysis of Demutualization. *Journal of Risk and Insurance*, 70(3):401–437.

Wells, B. P., Cox, L. A., and Gaver, K. M. (1995). Free Cash Flow in the Life Insurance Industry. *Journal of Risk and Insurance*, 62(1):50–66.

Winter, R. A. (1994). The Dynamics of Competitive Insurance Markets. *Journal of Financial Intermediation*, 3(4):379–415.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data.* MIT Press, 2nd edition.

Zanjani, G. (2007). Regulation, Capital, and the Evolution of Organizational Form in US Life Insurance. *American Economic Review*, 97(3):973–983.

Zou, H., Yang, C., Wang, M., and Zhu, M. (2009). Dividend Decisions in the Property and Liability Insurance Industry: Mutual versus Stock Companies. *Review of Quantitative Finance and Accounting*, 33(2):113–139.

Part V

How Risky Are Interest Rate Guarantees Embedded in Participating Life Insurance Contracts? The Case of Germany

Abstract

This paper analyses the risk resulting from interest rate guarantes offered within participating life insurance contracts. To do so, a Monte Carlo simulation study is carried out which is calibrated using empirical data from the German bond market. As life insurance companies tend to invest the majority of their assets in bonds, the analysis is based on a term structure model. The results show that the interest rate guarantees offered in the German insurance market can be fulfilled to a very high probability with simple investment strategies using government bonds. One particular factor explaining this phenomenon is the existence of diversification effects between different investments which occur due to the pooling of the undistributed surplus.

1   Introduction

Insurance companies, in particular those offering life insurance products, generally face more challenges in relation to their investment strategies than ordinary non-financial firms. As the contract periods of many of the life insurance policies offered span a significant proportion of an insured's life, life insurers' liabilities are in general of a long-term nature, frequently reaching up to 40 years. However, there are very few liquid financial assets with a similar maturity, which can result in significant reinvestment and liquidity risk. Hence, insurance companies often need to put much effort into managing their assets and liabilities, in particular with respect to so-called duration and cash flow matching. Unmatched assets and liabilities can potentially be a source of significant risks for life insurance companies, even though dealing with term transformation can be seen as one of the key competences of a financial intermediary. In general, interest rate guarantees included in many of the participating life insurance contracts can additionally increase an insurer's risk exposure. These are particularly important, as the level of an interest rate guarantee usually is assigned to a contract at its time of signing and remains the same for its entire term.

In order to reduce the risk of a life insurer falling short of the promised guarantee, the level of interest rate guarantees is often regulated. In Germany, it is aligned with the interest rate of long-term government bonds and is thus estimated in a fairly transparent way. Despite this, each decrease in the threshold for interest rate guarantees seems to give rise to protests from insurance companies, which fear for the attractiveness of life insurance products in comparison to other investment opportunities.

In this paper, the risk resulting from offering interest rate guarantees within participating life insurance products is measured. In particular, the aim is to investigate whether interest rate guarantees can cause significant risk for life insurance companies in the German insurance market, given an endogenous regulatory regime which has the power to assign the maximum interest rate guarantee. To the best of our knowledge, no analysis including an endogenous rule for setting the maximum level for interest rate guarantees has yet been conducted. Due to the signifi-

cant amount of bonds held by German insurers[105] as well as the current design of the regulatory framework[106], it is considered appropriate to model the insurer assets using a term structure model. The few previous authors who explicitly model stochastic interest rates in the context of cliquet-style guarantees use affine frameworks based either on Vasicek (1977) or Cox et al. (1985) (see Gerstner et al., 2008; Zaglauer and Bauer, 2008), the focus of which is on the modeling of the short-term end of the term structure. As several studies emphasize the poor performance of affine interest rate models for out of sample forecasting (see, e.g., Duffee, 2002), in particular with respect to long-term interest rates, this paper adopts a different approach. The analysis uses the forecasting methodology developed by Diebold and Li (2006) which applies the Nelson and Siegel (1987) framework. It allows a Monte Carlo simulation study to be designed, which is calibrated with market data for term structure development provided by the Deutsche Bundesbank. In this way, density estimates for the entire term structure are obtained. An analysis is then carried out of an insurer investing in a government bond portfolio. Thus, the reinvestment risk linked to investments in bonds is brought into focus. The results show that even though the insurance company in the analysis is not provided with any equity capital at the outset, if contracts of several policyholder collectives are in force the interest rate guarantees can be earned to a very high probability using simple investment strategies based on investments in government bonds. One particular factor explaining this phenomenon is the existence of diversification effects between different investments which occur due to the pooling of the undistributed surplus.

The remainder of this paper is structured in the following way: Section 2 gives an overview of the relevant literature. The current legal framework regulating participating life insurance contracts in Germany is described in Section 3. As this is essential to the final results, Section 4 gives a detailed description how the term structure has been modeled. The analysis of risk resulting from interest rate guarantees is conducted

---

[105]On average, bonds tend to constitute about 70 percent of German life insurers' investment portfolio; see GDV (2010).

[106]See Section 3, for a description of the regulation with respect to the level of the maximum level of interest rate guarantees.

in Section 5. Firstly, to explain the applied approach, Section 5.1 investigates a setting with only one policyholder collective. Secondly, in Section 5.2, a portfolio of two policyholder groups entering the company at different times is analyzed. Section 6 provides the conclusions.

2 Literature overview

Since the publication of the work of Brennan and Schwartz (1976), much attention has been paid in the academic literature covering participating life insurance products to their pricing, often in the contingent claims context and hence based on risk-neutral valuation. Briys and de Varenne (1997) investigate the impact of minimum interest rate guarantees and a bonus participation mechanism on the duration and convexity of insurer liabilities and show that their presence may have a significant impact on the interest rate sensitivity of such contracts. Grosen and Jørgensen (2000) develop a pricing model for contracts with surrender options and a reserve-smoothing bonus participation based on a cliquet-style interest rate guarantee and demonstrate that the fair value of such policies exhibits significant sensitivity to the level of interest rates. Bacinello (2001) and Haberman et al. (2003) introduce pricing models covering return-based bonus participation. A further emphasis on surrender options within contingent claims pricing is put in Albizzati and Geman (1994), Bacinello (2003), and Siu (2005). Jensen et al. (2001) and Tanskanen and Lukkarinen (2003) provide general frameworks for fair numerical pricing of contracts with surrender options and bonus participation if the path-dependence does not allow the development of closed-form solutions. Grosen and Jørgensen (2002) analyze the fair value of insurer equity and liabilities if there is a regulatory authority continuously monitoring the solvency level of the company. Ballotta (2005) introduces a pricing model for participating life insurance contracts providing interest rate guarantees which allows for possible jumps in the asset process.[107] Ballotta et al. (2006) extend the previous valuation method of Haberman et al. (2003) by additionally accounting for the default put option of the insurer. Bauer et al. (2006) present a model for the valuation of partic-

---

[107]This is also done by Siu (2005).

ipating life insurance contracts and adapt it to the German regulatory framework. Kleinow (2009) develops a fair pricing model for participating life insurance contracts under the assumption that the insurance company has the right to modify the investment strategy of the underlying portfolio at any time to hedge its shortfall risk. Hedging of interest rate guarantees with different investment strategies is also analyzed in Prieul et al. (2001) and Kleinow and Willder (2007). Several works propose pricing methods for life insurance contracts under stochastic development of interest rates. Miltersen and Persson (1999) and Bernard et al. (2005) apply the Heath et al. (1992) framework. As they do not consider either cliquet-style interest rate guarantees, typical distribution schemes or option features embedded in many life insurance contracts, Zaglauer and Bauer (2008) value the participating life insurance contracts and the embedded cliquet-style options under stochastic development of short-term interest rates. In their approach, the short-term interest rate is modeled in line with two different affine term structure models developed by Vasicek (1977) and Cox et al. (1985).[108] These affine term structure frameworks are also applied by Gerstner et al. (2008) who additionally allow for investments in long-term bonds.

Another strand of the literature focuses on risk comparison of participating life insurance contracts and bonus participation schemes. Kling et al. (2007b) investigate the influence of interest rate guarantees on the insurer's "real-world" shortfall probability and study how the risks depend on characteristics of the policy, on an insurer's reserve situation and asset allocation, on management decisions, and on regulatory parameters. In the subsequent work, Kling et al. (2007a) analyze the impact of different surplus distribution mechanisms on the risk exposure of the insurance company. In addition, Gatzert (2008) analyzes the impact of different asset management strategies on shortfall risk and risk-neutral pricing. Gatzert and Kling (2007) compare the risk according to the real-world measure of diverse contracts with the same value according to the risk-neutral measure. They show that the real-world risk of fair contracts may differ if the assumption of perfect hedging implied by the

---

[108] Also see Persson and Aase (1997), Hansen and Miltersen (2002), and Barbarin and Devolter (2005).

risk-neutral valuation is not met. Barbarin and Devolter (2005) combine both the financial and the actuarial approaches and develop a framework suitable for pricing and risk measurement in cases where perfect hedging assumptions are not realistic. An international comparison of different bonus distribution schemes based on future payoff distributions is provided by Cummins et al. (2004). Zemp (2010) develops a framework which allows a comparison of the risks of different bonus participation models previously presented in the literature.

Some of the literature also focuses on the performance analysis of participating life insurance contracts; see, e.g., Gründl et al. (2003); Waldow (2003); Faust et al. (2010). Furthermore, Hansen and Miltersen (2002) investigate the consequences of pooling the undistributed surplus over two groups of customers with different guaranteed interest levels. However, the latter are exogenous in the model.

3   The legal system in Germany

This section describes the most important aspects of the current regulatory framework covering participating life insurance products in Germany. In particular, it focuses on interest rate guarantees, investment control, and surplus participation mechanisms.

The establishment of a new European internal market in 1994 triggered significant changes in insurance regulation and supervision within the European Union. Since this time, products offered on the German insurance market are no longer subject to ex ante inspection and authorization by the supervisory authority (see Farny, 2011). In general, interest rate guarantees embedded in life insurance contracts offered in Germany are not directly regulated. Theoretically, a life insurer is free to offer their customers an arbitrary level of guaranteed return, as long as they are able to fulfill the general requirements of prudential insurance regulation and supervision. However, an indirect threshold is set by the regulator by prescribing the maximum actuarial interest rate which is the upper bound for discounting the technical liabilities resulting from participating life insurance contracts on an insurer's balance sheet. Hence, all guaranteed interest rates in excess of the maximum actuarial interest

rate decrease the book value of an insurer's equity. Thus, an insurance company has a considerable incentive to restrict the risk resulting from granting interest rate guarantees which are too high. Nevertheless, due to market competition it is not likely that an insurer will be able without difficulty to sell products with an interest rate guarantee much lower than the current maximum actuarial interest rate. Therefore, we can expect the overall level of granted interest to be close to this level.

The current level of the maximum actuarial interest rate in Germany is governed by the Actuarial Reserve Ordinance[109] and the Insurance Supervisory Act[110]. It is determined by the Federal Ministry of Finance[111] and generally not permitted to be higher than 60 percent of the long-term average yield on 10-year government bonds (see § 65,1 VAG).[112] The maximum actuarial interest rate assigned to a contract at its inception holds, in general, for its entire term (see § 2,2 DeckRV). In Figure 25,[113] the development of the maximum actuarial interest rate is plotted against the estimated yield on 10-year German government bonds as well as 60 percent of its 10-year rolling average. In particular, for the time period starting in 1994 we observe a forward-looking response of the regulator to the changing development of the term structure.[114] This indicates that the decision makers seem to put much effort into restricting the risk resulting from excessively high interest rate guarantees.

In a similar way to many other countries, the system of prudential regulation and supervision of the German insurance sector contains particular restrictions on the asset allocation of life insurance companies. These

---

[109]In German: *Deckungsrückstellungsverordnung* (DeckRV). See § 2 DeckRV, version dated July 22, 2009.

[110]In German: *Versicherungsaufsichtsgesetz* (VAG). See § 116 VAG, version dated January 1, 2011.

[111]It is done after consultations with the Federal Financial Supervisory Authority (*BaFin*) and the German Actuarial Association (*DAV*).

[112]In some cases, company- or product-specific maximum actuarial interest rates are allowed. Company specific maximum actuarial interest rates are to be aligned with the earnings on assets currently held by the insurer plus the expected earnings on future investments minus an appropriate safety margin; see § 65,2 VAG.

[113]See § 2 DeckRV and GDV, 2005, Geschäftsmodell Rentenversicherung, at `http://www.gdv.de`, accessed January, 15 2011.

[114]On February 22, 2011, the Federal Ministry of Finance decided to lower the maximum actuarial interest rate to 1.75 percent from January 1, 2012.
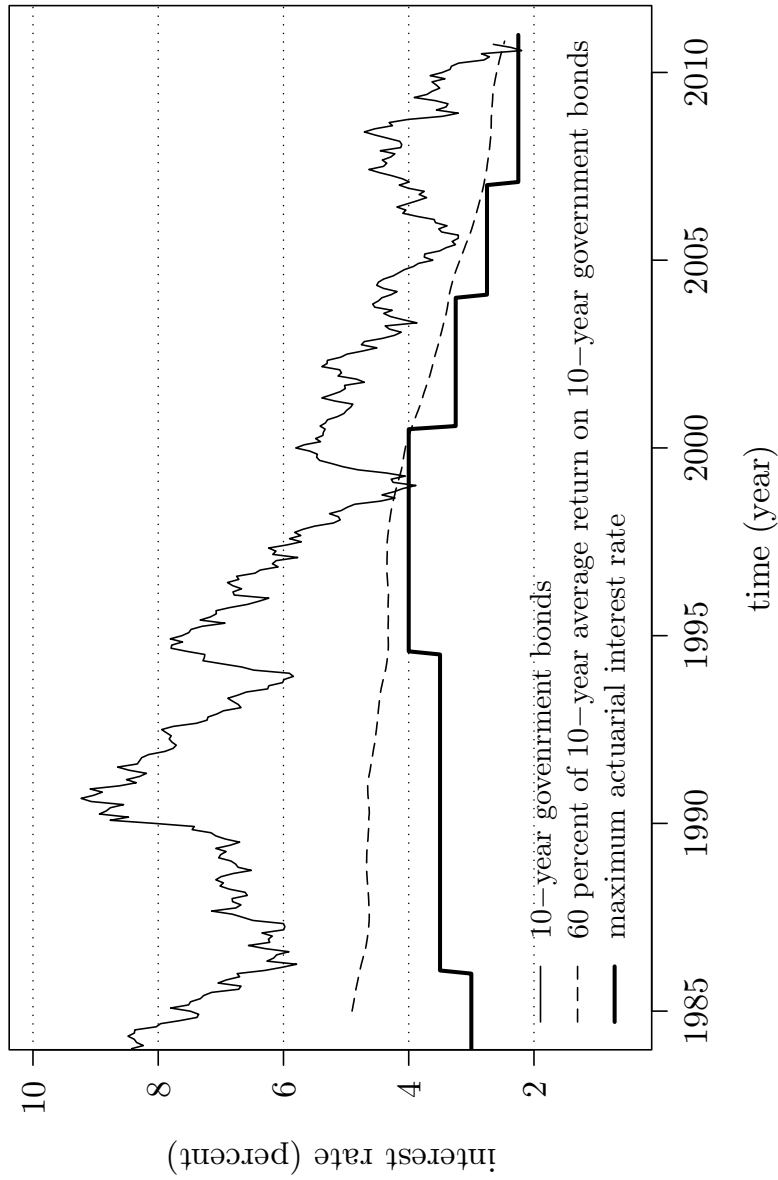
Figure 25: Development of the maximum actuarial interest rate for participating life insurance contracts and yield on 10-year government bonds in Germany for the time period 1983/01-2010/10, estimated using the Nelson-Siegel framework. The dashed line depicts 60 percent of the rolling 10-year average of yields on 10-year government bonds.

are primarily governed by the Investment Ordinance[115]. According to these regulations, a life insurance company faces investment restrictions with regard to the so-called tied-up capital which mainly contains of the assets corresponding to the insurance liabilities. Limits include those on investments in certain asset classes and on financial instruments with an increased level of default risk. In general, life insurance firms tend to invest the majority of their asset portfolio in bonds.

Surplus participation in the German life insurance sector is mainly governed by the Insurance Supervisory Act[116] and the Minimum Funding Ordinance[117]. In particular, the latter prescribes in which way the surpluses in excess of the actuarial interest rate and insurer's periodic operational costs are allocated between the policyholders and company's shareholders. In general, at least ninety percent of the surplus has to be either directly—e.g., via direct assignment to customer accounts or a premium reduction—or indirectly—via the so-called "provisions for premium refunds"—credited to the policyholders.[118] Undistributed provisions for premium refunds, which are owned collectively by all policyholders, are usually assigned to policyholder accounts either in the subsequent periods or at the contract expiration. The decisions in this regard are to some extent a matter of management discretion.

4   Simulation approach

In order to measure the risk resulting from offering investment guarantees, the term structure of German government bonds is simulated. This paper follows the approach introduced by Diebold and Li (2006), who reinterpret the widely applied Nelson and Siegel model. Nelson and Siegel (1987) propose a parsimonious three-component exponential approximation which is widely used—often in the extended version presented by Svensson (1994)—for fitting a parametric curve to the yields estimated from the current bond prices (see also Siegel and Nelson, 1988).

---

[115]In German: *Anlageverordnung* (AnlV).
[116]See § 56a and § 81c VAG.
[117]In German: *Mindestzuführungsverordnung* (MindZV).
[118]See § 4,3 MindZV, version dated April 4, 2008.

The yield curve equation formulated by Diebold and Li (2006) is of the following form

$$y_{t,m} = \beta_{0t} + \beta_{1t}\left(\frac{1 - e^{-\lambda_t m}}{\lambda_t m}\right) + \beta_{2t}\left(\frac{1 - e^{-\lambda_t m}}{\lambda_t m} - e^{-\lambda_t m}\right), \quad (76)$$

where $y_{t,m}$ denotes the respective yield at time t with a maturity m.[119] β_{0t}, β_{1t}, and β_{2t} are latent dynamic factors with long-term, short-term, and middle-term impact, respectively. Therefore, they can be interpreted as the level, slope, and curvature of the term structure. The parameter λ_t influences the exponential decay rate which determines whether the curve better fits the data at short—i.e., for high values of λ_t—or long— i.e., for small values of λ_t—maturities. Furthermore, it determines the maturity at which the loading on β_{2t} reaches its maximum. In their work, Diebold and Li (2006) estimate the three β-factors for U.S. Treasuries and find them to be highly autocorrelated. They model the respective factors as AR(1) processes and use their results for forecasting purposes with very good results for long-term forecasts.

In the same manner, the parameter estimates provided by the Deutsche Bundesbank[120] are analyzed and used to calibrate the simulation of the term structure of German government bonds at several time points out of sample. As in Diebold and Li (2006), it is assumed that the respective parameters of the Nelson-Siegel curve follow lower order AR processes. The density estimates of the term structure are then used to analyze the shortfall risk on the interest rate guarantees of German life insurers.

Figure 26 shows the development of the estimated term structure for the time period between January 1983 and October 2010 at the end of each month, which we use for calibration purposes.[121] In most cases,

---

[119]Diebold and Li (2006) prove the identity of Equation (76) with the Nelson and Siegel (1987) model of the form: $y_t(m) = b_{0t} + b_{1t}\frac{1-e^{-\lambda_t m}}{\lambda_t m} - b_{2t}e^{-\lambda_t m}$ for $b_{0t} = \beta_{0t}$, $b_{1t} = \beta_{1t} + \beta_{2t}$, and $b_{2t} = \beta_{2t}$, see Diebold and Li (2006, p. 343).

[120]See Schich (1997, p. 17), for details regarding the estimation.

[121]Parameter estimates prior to this period were obtained based on a much lower amount of available government bonds, in particular, at the short and long end of the yield curve. Low number of financial instruments in connection with the easy overdetermination of the model led to fairly instable parameter estimates. Hence, these are excluded from the study. Nevertheless, the analysis is calibrated on a similar data frame to that of Diebold and Li (2006).
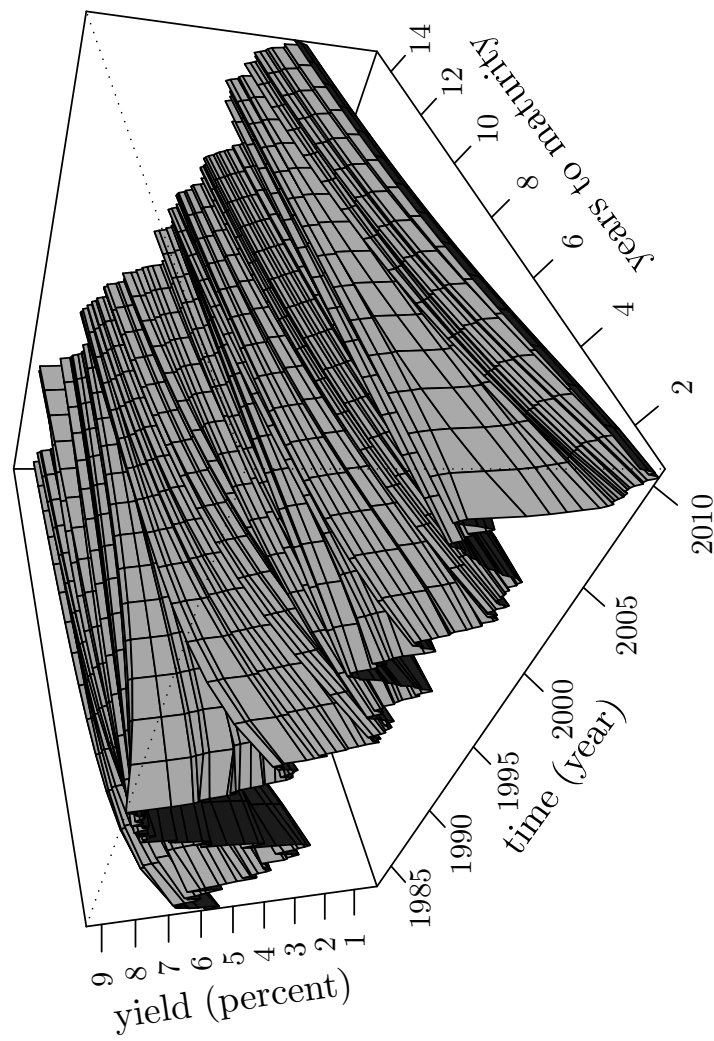
Figure 26: Estimates of the German government bond yields term structure in the time period 1983/01-2010/10, calculated according to the Nelson and Siegel (1987) method as described in Schich (1997).

the yield curve is increasing with maturity and its dynamics seems to be persistent. The volatility (persistence) of yields over time seems to decrease (increase) with increasing maturity.

Table 18 confirms those conjectures. It presents the descriptive statistics of the underlying Nelson-Siegel estimates. We observe that $\hat{\beta}_{2t}$, a parameter that determines the curvature of the term structure, shows the highest standard deviation. $\hat{\beta}_{0t}$, governing the level of the term structure, is the least volatile parameter and is characterized by a very strong persistence over time. Both the augmented Dickey-Fuller test[122] and the Phillips-Perron test suggest that $\hat{\beta}_{2t}$ does not have a unit root. They arrive at mixed results for $\hat{\beta}_{0t}$ and $\hat{\beta}_{1t}$.[123]

Figures 27, 28, and 29 show the level of persistence of the three parameters describing the level, slope, and curvature of the term structure as well as the goodness of fit of the estimated AR models. The respective autocorrelation functions in Figures 27(a), 28(a), and 29(a) suggest the strongest persistence in $\hat{\beta}_{0t}$ associated with the level of long-term interest rates. The lowest persistence can be seen in the estimate $\hat{\beta}_{2t}$ governing the curvature.[124] The remaining plots in Figure 27 show the respective residual autocorrelation functions. It should be noted that the residual autocorrelations for AR(2) models are in general lower than those for AR(1) models, which suggests that the AR(2) models seem to more precise describe the conditional means of the parameter estimates. Hence, in contrast to Diebold and Li (2006), this paper conducts its analysis based on AR(2) processes.

The respective parameters are simulated on a monthly basis—level, slope, and curvature—of Equation (76) as AR(2) processes calibrated on the entire sample—from January 1983 to October 2010. This allows the distribution of the entire term structure at the end of each year at several points out of sample to be forecasted. In line with Diebold and Li (2006), the parameter λ_t is set to 1.3941. This is the respective value of

---

[122]The lags in the augmented Dickey-Fuller unit root test are chosen based on the SIC (see Schwarz, 1978).

[123]Similar pattern can be observed in the data analyzed by Diebold and Li (2006, p. 350).

[124]Even though, a different data set is analyzed here—German not U.S. government bonds and over a slightly longer time period—the results are very similar to those of Diebold and Li (2006, p. 352).

Parameter	Mean	Std. Dev	Minimum	Maximum	$\hat{\rho}(1)$[a]	$\hat{\rho}(12)$[a]	$\hat{\rho}(30)$[a]	ADF[b][*]	PP[c][*]
$\hat{\beta}_{0t}$	6.5343	1.4789	1.9133	10.2481	0.8708	0.6985	0.4577	-2.6344	-3.4003
$\hat{\beta}_{1t}$	-2.0654	1.9381	-5.4252	3.8656	0.9057	0.6001	0.0047	-2.8932	-2.5189
$\hat{\beta}_{2t}$	-1.9816	2.9548	-11.9133	10.2485	0.6932	0.1746	0.0078	-5.7227	-4.2604

[a] Sample autocorrelation at a displacement of 1, 12, and 30 months, respectively.
[b] Augmented Dickey-Fuller (ADF) unit root test statistics (see Dickey and Fuller, 1979; Said and Dickey, 1984).
[c] Phillips-Perron (PP) unit root test statistics (see Phillips and Perron, 1988).
[*] MacKinnon (1996) critical values for the rejection of hypothesis of a unit root are -3.4498, -2.8700, and -2.5713, at a one, five, and ten percent level, respectively.

Table 18: Descriptive statistics of the parameter estimates
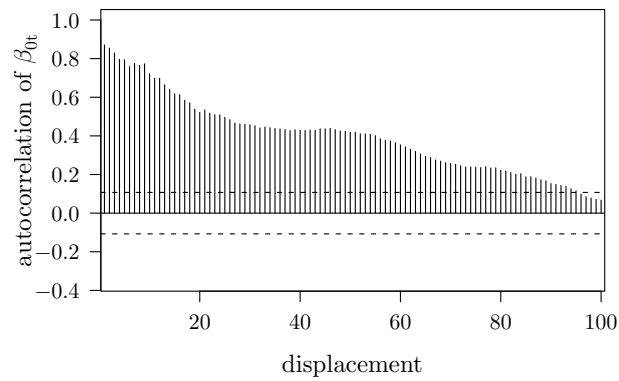
(a) Autocorrelations of $\hat{\beta}_{0t}$
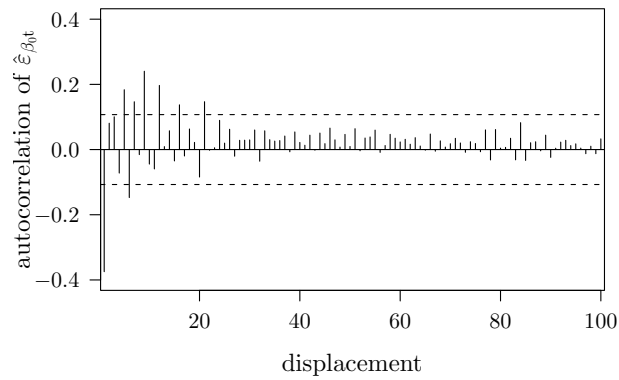
(b) Autocorrelations of $\hat{\beta}_{0t}$ AR(1) residuals
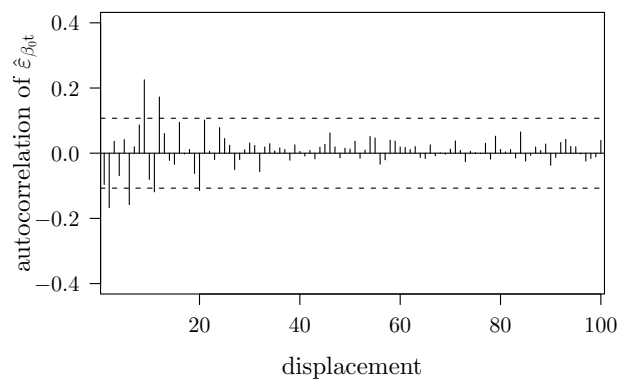
(c) Autocorrelations of $\hat{\beta}_{0t}$ AR(2) residuals

Figure 27: Autocorrelations and residual autocorrelations of the term structure parameters (1). Sample autocorrelations of $\hat{\beta}_{0t}$ and sample autocorrelations of AR(1) and AR(2) models fit to $\hat{\beta}_{0t}$ are presented. Dashed lines: 95 percent confidence bands.
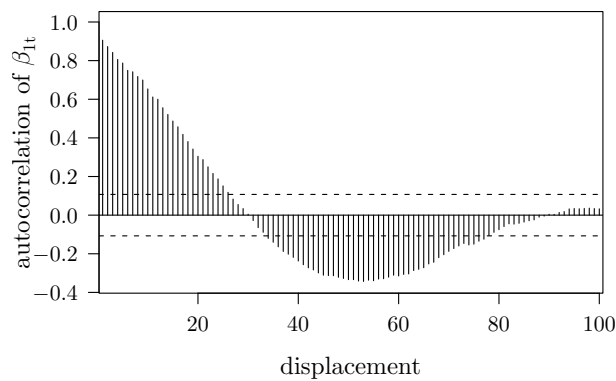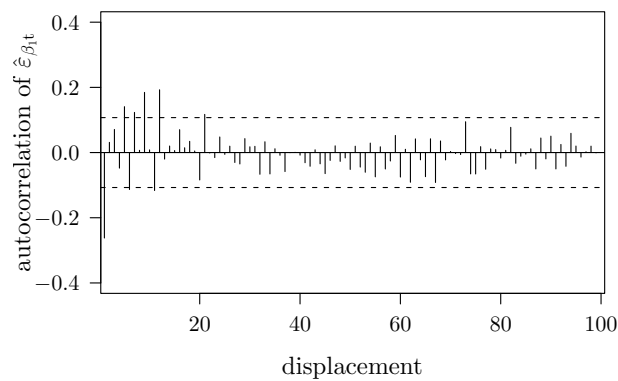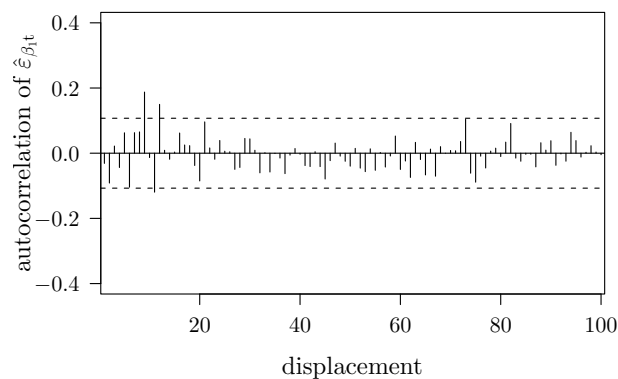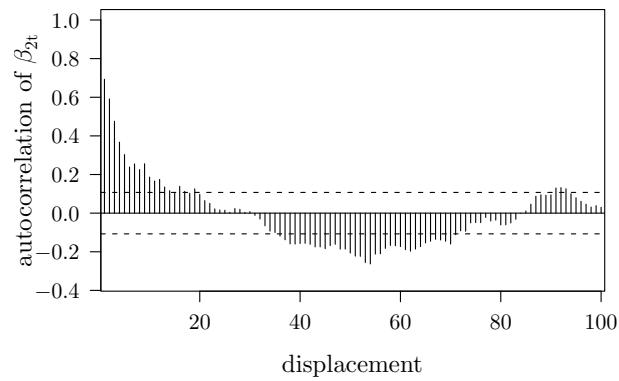
(a) Autocorrelations of $\hat{\beta}_{1t}$

(b) Autocorrelations of $\hat{\beta}_{1t}$ AR(1) residuals

(c) Autocorrelations of $\hat{\beta}_{1t}$ AR(2) residuals

Figure 28: Autocorrelations and residual autocorrelations of the term structure parameters (2). Sample autocorrelations of $\hat{\beta}_{1t}$ and sample autocorrelations of AR(1) and AR(2) models fit to $\hat{\beta}_{1t}$ are presented. Dashed lines: 95 percent confidence bands.
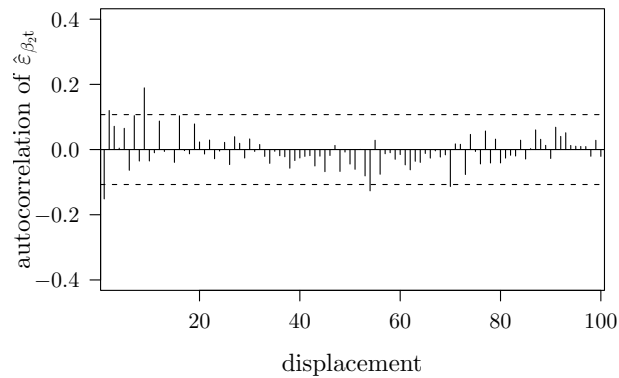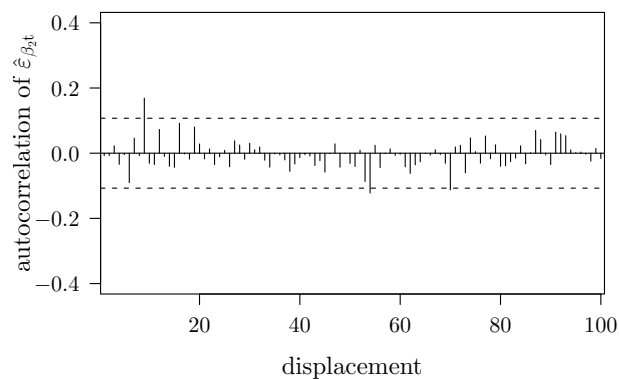
(a) Autocorrelations of $\hat{\beta}_{2t}$

(b) Autocorrelations of $\hat{\beta}_{2t}$ AR(1) residuals

(c) Autocorrelations of $\hat{\beta}_{2t}$ AR(2) residuals

Figure 29: Autocorrelations and residual autocorrelations of the term structure parameters (3). Sample autocorrelations of $\hat{\beta}_{2t}$ and sample autocorrelations of AR(1) and AR(2) models fit to $\hat{\beta}_{2t}$ are presented. Dashed lines: 95 percent confidence bands.

λ_t which maximizes the loading on the factor $\hat{\beta}_{2t}$ for the maturity of 2.5 years (30 months). A Monte Carlo simulation is carried out with 300 000 iterations for each β-parameter and the term structure distribution at the end of every year is estimated: up to 30 years seen from October 31, 2010. The results for the yields on German government (zero-)bonds with 5, 10, and 15 years to maturity five and ten years out of sample are presented in Figures 30 and 31.[125] As expected, in both cases the mean (volatility) of returns slightly increases (decreases) with increasing time to maturity. In the following, the simulation results are used to analyze the reinvestment risk resulting from granting the insureds interest rate guarantees.

5 Economic analysis of interest rate guarantees

In this section, the aim is to analyze the risk underlying interest rate guarantees of participating life insurance contracts. In the following, the magnitude of this risk for an insurer following simple investment strategies in risk-free financial instruments is investigated. Section 5.1 models an insurer with a contract portfolio consisting of policies signed at one point in time. Section 5.2 extends the analysis to multiple groups entering the company at different time points.

## 5.1 One policyholder group

Insurer assets

First, a model is provided for a group of insureds investing a certain amount in a participating life insurance policy with a contract period $T \in \mathbb{N}$ including an interest rate guarantee which, due to a high level of competition in the market, is set equal to the current level of the maximum actuarial interest rate g. $t \in \mathbb{Z}$ is the timing variable denominated in years. To simplify matters, it is assumed that the contract does not
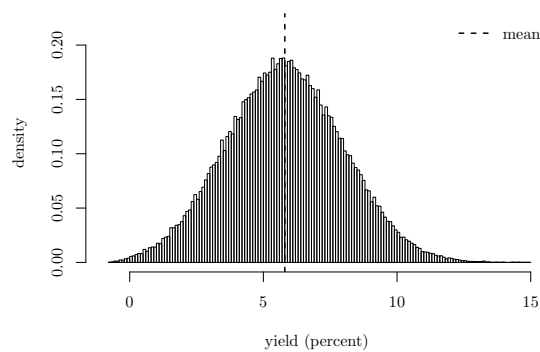
---

[125] As the Nelson-Siegel model does not restrict the yields to be positive, we set the negative yields to zero. Potentially negative interest rates are also a problem with the Vasicek (1977) model.
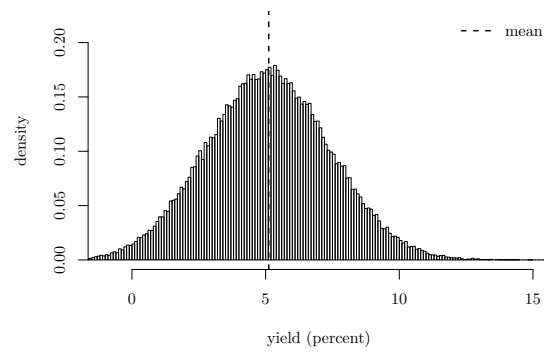
(a) Yield on a 5-year investment 5 years
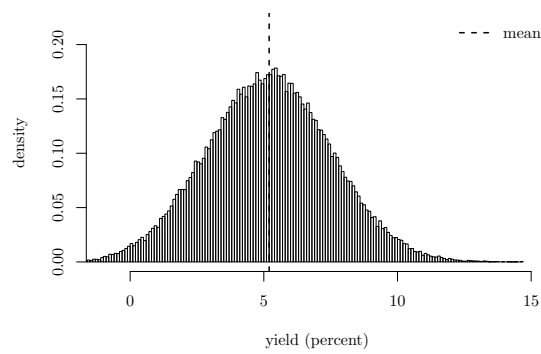out of sample
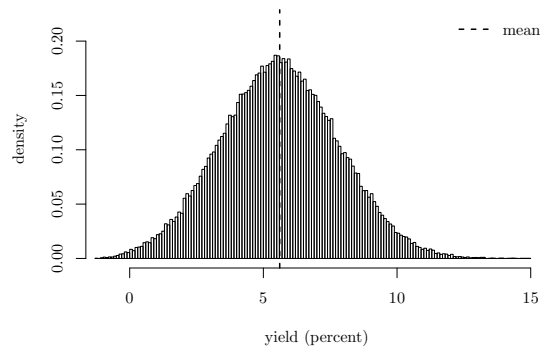
(b) Yield on a 10-year investment 5 years
out of sample

(c) Yield on a 15-year investment 5 years
out of sample

Figure 30: Simulated distribution of yields at diverse time points out of
sample (1). Histograms are presented for the yields on five, ten, and fif-
teen year investments in German government bonds, five years in excess
of the reference point.

(a) Yield on a 5-year investment 10 years
out of sample

(b) Yield on a 10-year investment 10 years
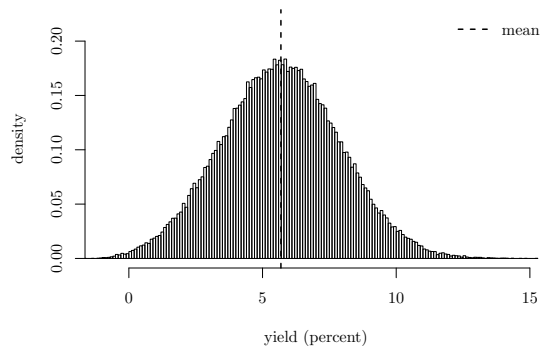out of sample

(c) Yield on a 15-year investment 10 years
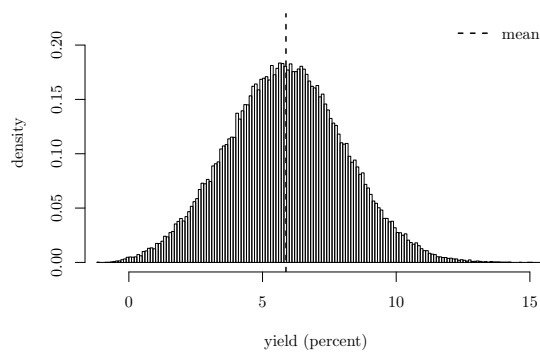out of sample

Figure 31: Simulated distribution of yields at diverse time points out
of sample (2). Histograms are presented for the yields on five, ten, and
fifteen year investments in German government bonds, ten years in excess
of the reference point.

include any death benefit payments or lapses.[126]   The insurance com-
pany is not provided with any equity capital at the outset; hence, the
amount paid in by the insureds constitutes its only assets. Its balance
sheet is presented in Figure 32.[127] Annual premiums, $\pi_\tau = \pi, \forall \tau$,[128] are
paid in advance, i.e., at times $\tau \in \{0, 1, \ldots, T-1\}$ and invested at the
time of their receipt exclusively in risk-free zero bonds which are held
to maturity.[129] As only zero bonds with a maturity of up to η years are
available, if $T > \eta$, the insurance company will certainly need to rein-
vest a part of its portfolio at least once over the term of the insurance
contract. Hence, it faces a certain degree of reinvestment risk.

If only one reinvestment is allowed, which is feasible only given $T \leq$
2η, the possible reinvestment points for the annual premium received at
$t = 0$ are equal to $t^* \in \mathbb{N} \cap [T - \eta, \eta]$. Subsequent to the choice of t^*, the
annual premium received at $t = 0$ is directly invested in a t^*-year bond
earning y_{0,t^*} (compare Equation (76)). At $t = t^*$, when the first bond
matures, the terminal value of the first investment, $\left[\pi (1 + y_{0,t^*})^{t^*} \right]$, is
invested in a $(T - t^*)$-year bond earning $y_{t^*,T}$. This yield is stochastic
from the perspective of $t = 0$, when the insurance policy is signed and
the first annual premium is invested.

---

[126]This is equivalent to an assumption that the death benefit payments and lapse
risk are either perfectly hedged or sufficiently diversified.

[127]This approach is based on the models widely applied in the modeling of companies
offering participating life insurance contracts; see, e.g., Hansen and Miltersen (2002);
Grosen and Jørgensen (2002); Kling et al. (2007a).

[128]In general, the amount of annual premium π depends on the price of the policy,
hence it is endogenous, i.e., determined by the contract parameters. However, we
can treat it as exogenous, as the subsequent analysis is based on investment returns
and calculate the risk as relative measures. In case of multiple insurance contracts,
where the relation of insurance prices might play a significant role, an appropriate
sensitivity analysis is conducted in Section 5.2.

[129]Zero bonds are often used in the modeling of life insurer assets; see, e.g., Albizzati
and Geman (1994); Barbarin and Devolter (2005). Even though zero bonds—in
particular, those with longer times to maturity—are relatively rare, the analysis is
based on those instruments. As the Deutsche Bundesbank allows for "stripping" of its
bonds (see Deutsche Bundesbank, 1997), this does not seem to be a strong assumption.
The stripping allows an investor to separate a single coupon payment of a bond from
its principal. In economic terms, buying a separated coupon or a separated principal
does not differ from buying a zero bond. If we exclude the possibility of trivial
arbitrage, prices of the respective elements of the stripped bonds should not differ
systematically from the prices of theoretical zero bonds which underlie the analyzed
term structure.

Assets	Liabilities
	P_t
A_t	
	B_t
A_t	A_t

Figure 32: Balance sheet of the insurer. One policyholder group.

A similar investment strategy is followed in case of the subsequent annual premiums. t^* is equal for all annuities. In general, the annual premium received at $t = \tau$ is initially invested in a zero bond with the maturity t^* earning $y_{\tau,\tau+t^*}$ and the subsequent (re-)investment at $t = \tau + t^*$ is conducted in a way allowing matching of the term of the underlying contract. Hence, at $t = \tau + t^*$ the insurer buys a bond with a maturity $T - (\tau + t^*)$, earning $y_{\tau+t^*,T}$. In cases when $\tau + t^* > T$, i.e., the initial investment at $t = \tau$ with maturity t^* would exceed the contract period, T, the insurer invests directly in a bond with maturity $T - \tau$ and no reinvestment is required. Respective revenues are booked on a pro rata temporis basis.

In this manner, depending on the chosen reinvestment point t^*, the terminal value at $t \in \{0, 1, 2, \ldots, T\}$ of the annual premium received at τ is equal to

$$
\begin{aligned}
a_{\tau,t} \;=\; & \pi \left[1 + y_{\tau,\min(\tau+t^*,T)}\right]^{\min(t-\tau,t^*)} \\
& \times \left[1 + y_{\min(\tau+t^*,T),T}\right]^{\max[t-(\tau+t^*),0]}.
\end{aligned}
\tag{77}
$$

Insurer assets at t can be calculated as a sum of respective terminal values of the annual premiums received prior to t. Hence,

$$
A_t = \sum_{\tau=0}^{t-1} a_{\tau,t}.
\tag{78}
$$

In general, the conditions under which the insurer will be able to fulfill the interest rate guarantee depend on the guarantee design. As German participating life insurance products usually include so-called cliquet-style guarantees, these are in the main focus of this study. A cliquet-style guarantee constitutes an obligation to credit the policyholder account with at least the guaranteed interest rate every year.[130] Hence, the accounting treatment of an insurer's assets, in particular with regard to the fluctuations in their market value, significantly influences whether an insurer is able to fulfill the guarantee in a given period. Under the German GAAP, specified in the German Commercial Code[131], an insurance company has an option—but not an obligation—to realize the losses resulting from non-permanent market value fluctuations of financial assets belonging to the so-called long-term (fixed) assets[132] in its financial statements.[133] As market value fluctuations of zero bonds which are held to maturity are in general not permanent, the development of book values described in Equation (78) will have the primary influence on the likelihood of fulfilling the guarantee in the analyzed setting. However, the current market values of the assets are also calculated as these will determine the severity of the potential loss, i.e., expected shortfall, faced by the policyholders.

Analogously to Equation (77), under the assumption of efficient financial markets, the market value at $t \in \{0, 1, 2, \ldots, T\}$ of the annual premium received at $t = \tau$ and directly invested in line with the introduced investment strategy is equal to

$$
m_{\tau,t} = \begin{cases} \dfrac{\pi\left[1+y_{\tau,\min(\tau+t^*,T)}\right]^{\min\left(t^*,T-\tau\right)}}{\left[1+y_{t,\min(\tau+t^*,T)}\right]^{\min(\tau+t^*-t,T-t)}} & \text{for } t \le \tau + t^* \\[2em] \dfrac{\pi\left(1+y_{\tau,\tau+t^*}\right)^{t^*}\left(1+y_{\tau+t^*,T}\right)^{T-t^*}}{(1+y_{t,T})^{T-t}} & \text{for } t > \tau + t^* \end{cases} . \qquad (79)
$$

---

[130]It differs from the so-called point-to-point guarantee which grants an insured a specified amount of money in excess of their initial investment at a certain point in the future, usually at contract expiration.

[131]In German: *Handelsgesetzbuch* (HGB).

[132]In German: *Anlagevermögen.*

[133]See § 253,3 HGB, version dated December 8, 2010, and, e.g., Wörner (2003, p. 125). Under the International Financial Reporting Standards (IFRS) market value fluctuations are irrelevant if a zero bond is classified as *held to maturity* under IFRS 39.

In general, as the insurance company invests an annual premium π at each time point τ in a zero bond held to maturity, the nominal value of the initial investment—identical to the cash-flow at the maturity—is equal to π times the return of the bond until its maturity. However, if the initial investment matures in the meanwhile $(t > \tau + t^*)$, instead of π, the firm reinvests the nominal value of the initial investment; hence the nominal value of the second investment is higher than that of the previous one. In both cases, to calculate the present value of the annual premium at t the respective nominal value of the investment has to be discounted in line with the current term structure. Obviously, book and market values are equal at the respective reinvestment points, $t = \tau + t^*$ and the maturity of the policy, $t = T$.

Along the lines of Equation (78), the respective market value of assets at t is equal to

$$M_t = \sum_{\tau=0}^{t-1} m_{\tau,t}. \tag{80}$$

Insurer liabilities

In this model, insurer liabilities at t consist of two separate accounts: the sum of individual policyholder accounts, P_t, and the bonus account, B_t, which belongs to the entire group of policyholders; see Figure 32. The latter can in general be used to build up reserves which are released in the event of the company not earning enough in a given period to fulfill the interest rate guarantee granted on the individual policyholder accounts.[134]

Similarly to in Equation (77), we can calculate the guaranteed terminal value at t of the annual premium π received at $t = \tau$ in the following way[135]

$$p_{\tau,t} = \pi \prod_{i=\tau+1}^{t} (1 + r_i^{\mathrm{p}}). \tag{81}$$

It is calculated based on the return on policyholders accounts r_t^{p} which in general depends on the reserves build up in the collective bonus account,

---

[134]It is equivalent to the provisions for premium refunds described in Section 3.
[135]See, e.g., Grosen and Jørgensen (2000).

B_{t-1}, and is at least equal to the guaranteed interest rate, g. In line with Equation (78), the overall value of policyholder accounts at t equals

$$P_t = \sum_{\tau=0}^{t-1} p_{\tau,t}. \tag{82}$$

The respective bonus account is a residual and equal to

$$B_t = A_t - P_t. \tag{83}$$

The interest rate granted to the policyholder accounts at t is based on the bonus distribution approach presented in Grosen and Jørgensen (2000). Hence,

$$r_t^{\mathrm{p}} = \max\left[g, \alpha\left(\frac{B_{t-1}}{P_{t-1}} - \gamma\right)\right]. \tag{84}$$

In this way, the policyholder accounts are credited at least the guaranteed interest rate, g. However, if the insurer was able to build up significant reserves in the previous periods, i.e., leading to a ratio of bonus account to the sum of policyholder accounts higher than the target buffer ratio γ, a fraction, α, of the excessive bonus reserve is distributed to the individual policyholder accounts.

Risk resulting from interest rate guarantees

Fulfilling a cliquet-style guarantee in the analyzed setting requires that the insurer is able to credit the policyholder accounts, P_t, every year with at least the guaranteed interest rate. In this way, the company remains solvent until a certain time point, $t = n$, if

$$A_t \geq P_t, \ \forall t \in \{1, ..., n\}. \tag{85}$$

In general, the insurer should seek initial investments with a return at least as high as the guaranteed interest rate, $y_{0,t^*} \geq g$. Otherwise, they would in every case face a bankruptcy in their first period of operation.[136]

---

[136]See Equations (77) and (81).

It is assumed that if the company becomes insolvent the policyholders are paid out and the policies are terminated.

Hence, we can calculate the probability that the insurer goes bankrupt at a certain time point, t, under the condition that they has not yet become insolvent by t, as

$$SP_t = \mathrm{P}\left(A_t < P_t \mid A_i \geq P_i\right), \ \forall i \in \{1, ..., t-1\}.$$ (86)

Thus, the probability that the insurer becomes insolvent during the entire term of the contract is equal to

$$SP_T = \sum_{t=1}^{T} SP_t.$$ (87)

In case of default, which is determined on the basis of the book value criteria, the policyholders are paid out the market value of the assets. Hence, the expected policyholder deficit ratio which is based on the unconditional mean of the shortfall is equal to

$$EPDR_t = \frac{\mathrm{E}\left[\max\left(P_t - M_t, 0\right) \mid \left(A_i \geq P_i\right)\right]}{P_t},$$ (88)

$$\forall i \in \{1, ..., t-1\}.$$

The respective expected shortfall ratio—which is based on a conditional mean and hence equals the expected shortfall given default—can be written as[137]

$$ESR_t = \frac{\mathrm{E}\left\{\max\left(P_t - M_t, 0\right) \mid \left[\left(A_t < P_t\right) \wedge \left(A_i \geq P_i\right)\right]\right\}}{P_t},$$ (89)

$$\forall i \in \{1, ..., t-1\}.$$

The analysis of the respective risk severity is based on ratios due to the fact that, as both the assets and policyholder accounts of the insurer

---

[137]To simplify matters, we implicitly make a conservative assumption that the insurer may go bankrupt due to the book value criteria even though in some cases the market value of the assets can be higher than the book value of policyholder accounts. In such a case, the incurred default is set to zero. In practice, an insurer in a similar position could sell some of the assets in order to avoid bankruptcy.
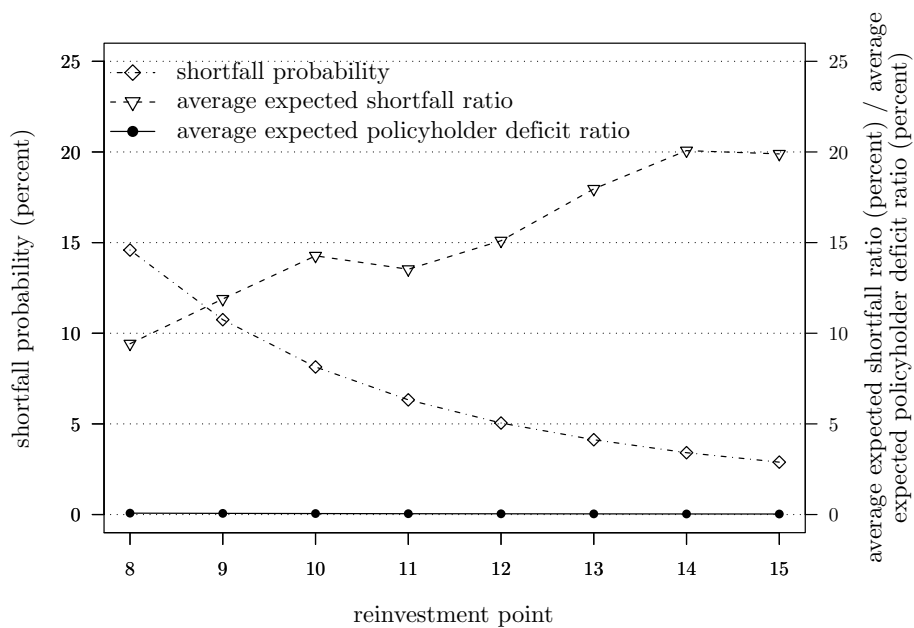
(a) Shortfall probabilities, SP_T, average expected shortfall ratios, $\frac{1}{T}\sum_{t=1}^{T} ESR_t$, and average expected policyholder deficit ratios, $\frac{1}{T}\sum_{t=1}^{T} EPDR_t$, dependent on the reinvestment point.
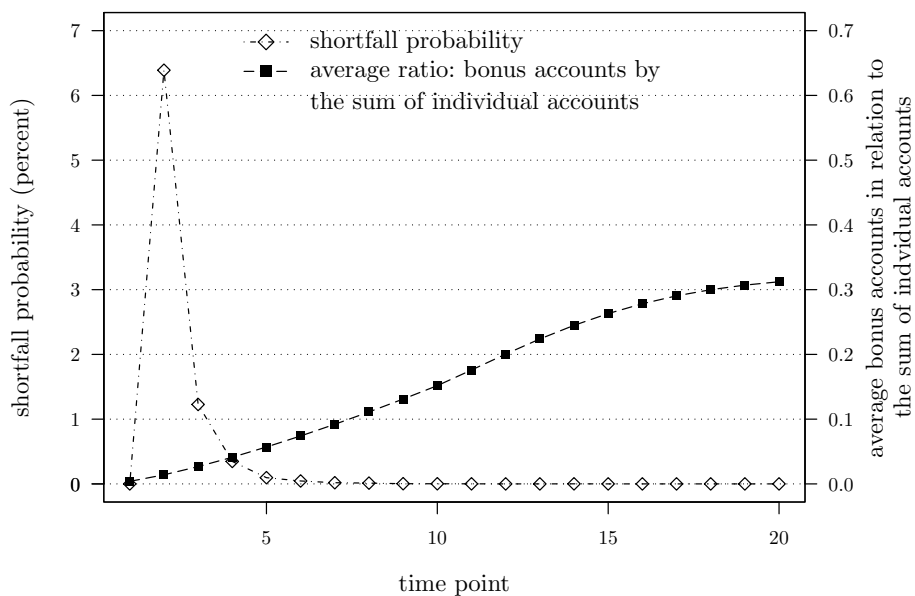
(b) Probability of shortfall, SP_t, and average bonus account development, average B_t/P_t, during the time of the contract given $t^* = 10$.

Figure 33: Results for a single policyholder collective. Annual premiums are invested in zero bonds which are held to maturity. Only one reinvestment is allowed.

increase continuously due to annual payments of the policyholders, the absolute shortfall values change over time by construction.

Figure 33(a) shows the shortfall probabilities (SP_T) as well as the average expected policyholder ratios $\left(\frac{1}{T} \sum_{t=1}^{T} EPDR_t \right)$ and the average expected shortfall ratios $\left(\frac{1}{T} \sum_{t=1}^{T} ESR_t \right)$ dependent on the reinvestment point t^* for a group of policyholders joining the company at $t = 0$, which is equivalent to October 31, 2010. The maximum actuarial interest on this day is $g = 2.25\%$; see Figure 25. Contract length is $T = 20$ years and zero bonds with maturities up to $\eta = 15$ are available. Hence, reinvestments at $t^* \in \{5, 6, ..., 15\}$ are possible. However, as $y_{0,t^*} < g$ for $t^* \in \{5, 6, 7\}$, those investment strategies are excluded.[138] The distributions of yields at the respective reinvestment points have been taken from the simulation described in Section 4. In line with Grosen and Jørgensen (2000), we set α and γ to 0.25 and 0.15, respectively.

It can be observed that the risk of failing to fulfill the interest rate guarantee depends on the assigned reinvestment point. In the chosen setting, the higher the maturity of the initial investment, the lower the resulting reinvestment risk in terms of both shortfall probability and the expected policyholder deficit ratio. In general, the level of the latter is very close to zero for all possible reinvestment points. However, with increasing length of the initial investment, the average expected shortfall ratio rises which indicates growing severity of risk.

Figure 33(b) allows us to investigate the shortfall risk depending on time. We observe that in the given setting it is the greatest in the first five periods after the contract inception.[139] This appears to be the most critical period in terms of risk. Subsequently, as the level of bonus accounts gradually increases, the shortfall risk decreases.

5.2 Two policyholder groups

This section investigates the diversification effects resulting from the fact that an insurance company normally signs contracts with more than one

---

[138] As previously stated, an insurer choosing such an investment strategy would face certain bankruptcy in the first period of operation.

[139] As $y_{0,10}$ is deterministic and $y_{0,10} > g$, the shortfall probability in the first year is equal to zero.

collective of insureds. Diverging times of contract inception suggest that insurers generally invest—and roll—their assets at different time points. If returns on diverse investments are pooled prior to assigning them to policyholders, this may lead to diversification.

The analyzed setting

Consider a setting in which in addition to the first group jointing the company at $t = 0$ (group 1) there is a second group of policyholders entering it at $t = \theta$ (group 2), where $\theta > 0$ and $\theta \in \mathbb{N}$.[140] Both groups sign contracts with identical periods, hence, $T_1 = T_2$. Contract conditions are similar to those from the previous setting. However, policies of group 1 and 2 differ with regard to the guaranteed interest rate. The guaranteed interest rate for group 1 is equal to g_1 and already known at $t = 0$. Group 2 is granted an interest rate guarantee g_2 which is stochastic from the perspective of $t = 0$ and dependent on the decision of the regulator at $t = \theta$. They decide upon the level of maximum actuarial interest rate based on the 60 percent of the 10-year arithmetic average of the yield on risk-free investments with a 10-year maturity. The regulator chooses values rounded downwards to quarters, e.g., 2.00, 2.25, 2.50, and 2.75. Hence,

$$
g_2 = \lfloor \delta \rfloor +
\begin{cases}
0.00 & \text{for } 0.00 \leq \delta - \lfloor \delta \rfloor < 0.25 \\
0.25 & \text{for } 0.25 \leq \delta - \lfloor \delta \rfloor < 0.50 \\
0.50 & \text{for } 0.50 \leq \delta - \lfloor \delta \rfloor < 0.75 \\
0.75 & \text{for } 0.75 \leq \delta - \lfloor \delta \rfloor < 1.00
\end{cases},
\tag{90}
$$

where

$$
\delta = 0.6 \frac{1}{10} \sum_{t=\theta-9}^{\theta} y_{t,10}
\tag{91}
$$

and $\lfloor . \rfloor$ is a floor function.[141] This approach is akin to the regulations in the German insurance market described in Section 3, see in particular

---

[140]This approach is similar to the approach proposed by Hansen and Miltersen (2002).

[141]For example, floors for 1.34, 1.50, and 1.67 are all equal to 1.

Figure 25. Investments of the annual premium payments of both groups, $\pi^{(1)}$ and $\pi^{(2)}$, are conducted in exactly the same manner as in the previous setting; however, t_1^* and t_2^*, the reinvestment points for group 1 and 2, are allowed to differ. For a given reinvestment point t_2^*, the respective reinvestment for group 2 takes place at $t = \theta + t_2^*$.

Figure 34 illustrates the new balance sheet of the insurer. Due to different levels of interest rate guarantees both collectives of policyholders have to be treated individually. Thus, $P_t^{(1)}$ and $P_t^{(2)}$ denote the sum of individual policyholder accounts of group 1 and 2, respectively, the sum of which is equal to P_t. The bonus account is owned jointly by both collectives of those insured and calculated in exactly the same way as in the previous setting; see Equation (83). In analogy to Equation (84), conditional on the financial position of the entire firm, the individual policyholder accounts are credited every period with an interest rate

$$r_t^{\mathrm{p},(1)} = \max\left[g_1, \alpha\left(\frac{B_{t-1}}{P_{t-1}} - \gamma \right) \right] \tag{92}$$

and

$$r_t^{\mathrm{p},(2)} = \max\left[g_2, \alpha\left(\frac{B_{t-1}}{P_{t-1}} - \gamma \right) \right], \tag{93}$$

for group 1 and 2, respectively. Hence, if the insurance company enjoys a good financial status and generates earnings higher than the interest rate guarantees for both groups, it is expected to credit all policyholders with an equal rate of return.

Insurer assets, both book and market values, are calculated as sums of current market and book values of the annual premiums paid in by both policyholder groups. Those are reckoned in line with Equations (78) and (80).

As contracts of both policyholder groups are simultaneously in force only during the time period $[\theta, T_1]$, the policyholders of group 2 do not face a loss in the event of this occurring prior to the inception of their policy. If the insurance company goes bankrupt while contracts of both policyholder collectives are valid, the loss is allocated to both policyholder collectives in proportion with their claims against the insurance company. Hence, in the event of bankruptcy of the insurer at t, group 1

	Assets		Liabilities

$$
\begin{array}{c|c}
\text{Assets} & \text{Liabilities} \\
\hline
& \left.\begin{array}{c} P_t^{(1)} \\[1em] P_t^{(2)} \end{array}\right\} P_t \\
A_t & \\
& B_t \\
\hline
A_t & A_t
\end{array}
$$

Figure 34: Balance sheet of the insurer. Two policyholder groups.

and 2 receive a payoff equal to $\frac{P_t^{(1)}}{P_t} M_t$ and $\frac{P_t^{(2)}}{P_t} M_t$, respectively. As in Hansen and Miltersen (2002), when leaving the company, policyholders belonging to group 1 are entitled to a fraction of the collectively owned bonus account. In this framework, it also is determined by the proportion of their individual accounts to the sum of individual accounts belonging to group 1 and 2, which is equal to $\frac{P_t^{(1)}}{P_t}$ of B_t at $t = T_1$.

Risk resulting from interest rate guarantees

In general, the respective risk measures in the case of two policyholder collectives have been calculated in line with those from the previous setting; see Section 5.1. However, it is important to mention that the respective policyholder collectives are only interested in whether the insurance company faces a default while their insurance policy is still in force. Thus, in terms of Equations (86) and (87), the respective shortfall probabilities for the entire contract term of group 1 and 2 are equal to

$$
SP_T^{(1)} = \sum_{t=1}^{T_1} SP_t \tag{94}
$$

and

$$
SP_T^{(2)} = \sum_{t=\theta+1}^{\theta+T_2} SP_t. \tag{95}
$$

As the intention is to measure both the expected policyholder deficit and the expected shortfall ratios from the perspective of the policyholders, Equations (88) and (89) are adjusted accordingly. The individual expected policyholder deficit and expected shortfall ratios for group 1 given $t \in \{1, \ldots, T_1\}$ are equal to

$$EPDR_t^{(1)} = \frac{\mathrm{E}\left[\max\left(P_t^{(1)} - \frac{P_t^{(1)}}{P_t}M_t, 0\right) \mid (A_i \geq P_i)\right]}{P_t^{(1)}}, \qquad (96)$$

$$\forall i \in \{1, ..., t-1\}$$

and

$$ESR_t^{(1)} = \frac{\mathrm{E}\left\{\max\left(P_t^{(1)} - \frac{P_t^{(1)}}{P_t}M_t, 0\right) \mid [(A_t < P_t) \wedge (A_i \geq P_i)]\right\}}{P_t^{(1)}},$$

$$(97)$$

$$\forall i \in \{1, ..., t-1\}.$$

The respective ratios for group 2 given $t \in \{\theta+1, \ldots, \theta+T_2\}$ are equal to

$$EPDR_t^{(2)} = \frac{\mathrm{E}\left[\max\left(P_t^{(2)} - \frac{P_t^{(2)}}{P_t}M_t, 0\right) \mid (A_i \geq P_i)\right]}{P_t^{(2)}}, \qquad (98)$$

$$\forall i \in \{1, ..., t-1\}$$

and

$$ESR_t^{(2)} = \frac{\mathrm{E}\left\{\max\left(P_t^{(2)} - \frac{P_t^{(2)}}{P_t}M_t, 0\right) \mid [(A_t < P_t) \wedge (A_i \geq P_i)]\right\}}{P_t^{(2)}},$$

$$(99)$$

$$\forall i \in \{1, ..., t-1\}.$$

Figures 35, 36, and 37 present the results for two policyholder collectives dependent on the choice of respective reinvestment points, t_1^* and t_2^*. In Figures 35(a) and 35(b), we can see that the shortfall probabilities for group 2 are significantly lower than the shortfall probabilities for group 1. We also observe a significant difference between both policyholder collectives with respect to the level of the average expected policyholder deficit ratio in Figures 36(a) and 36(b). It seems that overlapping portfolios, in connection with the pooling mechanism of the undistributed bonus account, allow a significant reduction of the shortfall risk for the policyholder group entering the company while the first contract is still in force. In this way, the shortfall risk, which for the chosen investment strategy is highest in the first few periods following a contract's inception,[142] can be significantly reduced. For group 2 the resulting shortfall risk can be decreased to a level acceptable from the regulatory perspective, e.g., under the value-at-risk restrictions of the Solvency II regime.[143] Furthermore, we observe that the shortfall probability for both policyholder collectives decreases with increasing lengths of the initial investments for both groups, t_1^* and t_2^*.[144] This is consistent with the outcomes presented in Section 5.1. Moreover, we observe a slight difference in terms of risk severity measured as average expected shortfall ratio; see Figures 37(a) and 37(b). It is lower for group 2.

Sensitivity analysis

To investigate the robustness of the results with respect to changes in several underlying parameters, some sensitivity analyses are conducted. The results are presented in Figure 38. Figures 38(a) and 38(b) show how the risk for group 1 and group 2 changes if the regulatory authority either anticipates or reacts too late to changes in interest rates. We observe that the regulatory rule based on the long-term interest rate progression generally leads to stable results with respect to shortfall risk.


[142]See the results presented in Figure 33(b).

[143]Solvency II, the new solvency framework for insurance companies operating in the European Union, is intended to require the insurance companies to hold an amount of solvency capital sufficient to establish a situation in which policyholders face a (partial) loss of their claims with a probability not exceeding 0.5 percent.

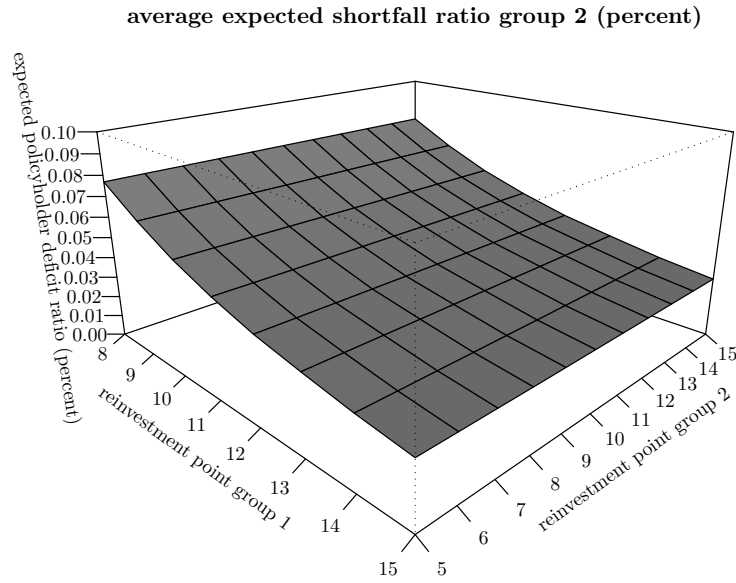[144]The effect for group 2 is less apparent due to the graph scaling.

shortfall probability group 1 (percent)

(a) Shortfall probability group 1
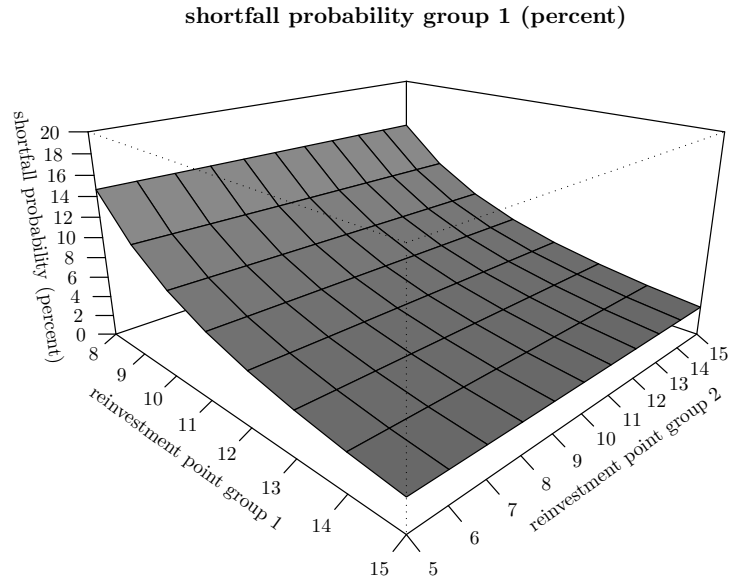
shortfall probability group 2 (percent)
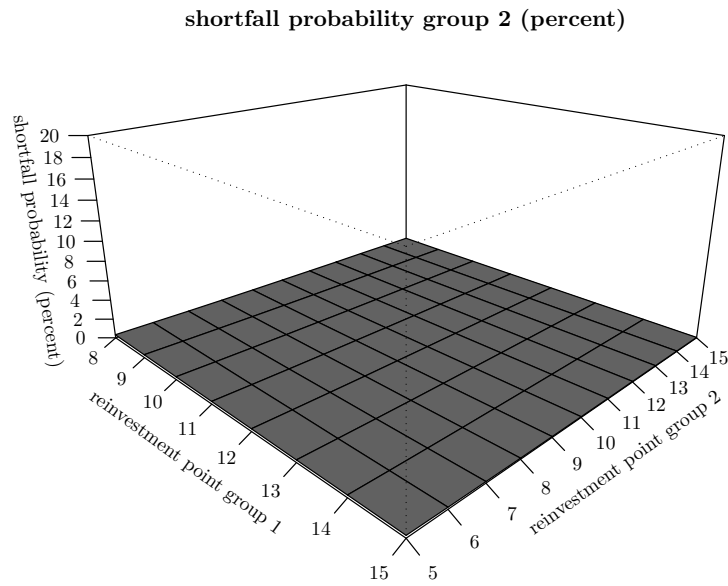
(b) Shortfall probability group 2

Figure 35: Shortfall probabilities, $SP_T^{(1)}$ and $SP_T^{(2)}$, given different investment strategies for group 1 and group 2. Group 2 enters the company at $\theta = 5$ and the annuities paid by both policyholder groups are equal, $\pi^{(1)}/\pi^{(2)} = 1$. Investment strategies leading to a certain default are excluded.
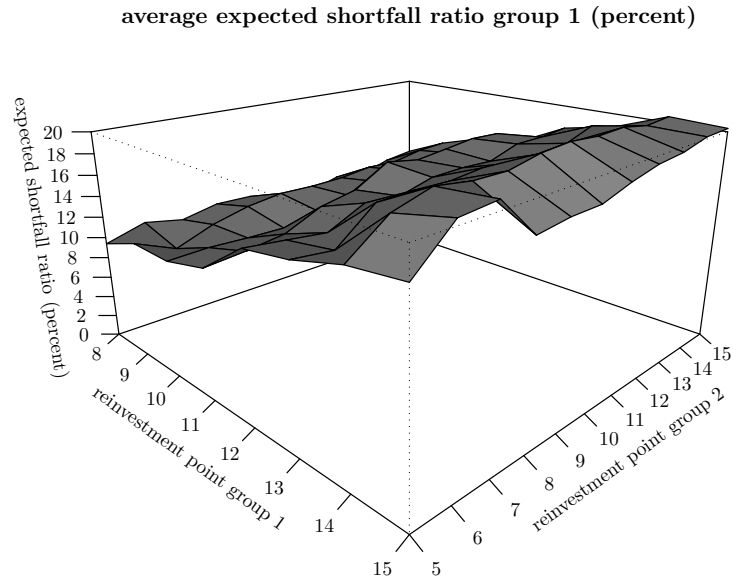
average expected shortfall ratio group 2 (percent)

(a) Average expected policyholder deficit ratio group 1

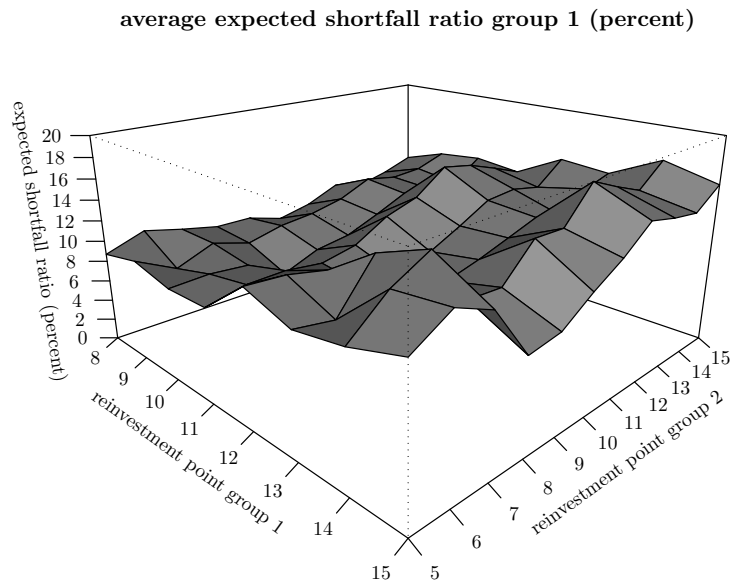**average expected policyholder deficit ratio group 2 (percent)**
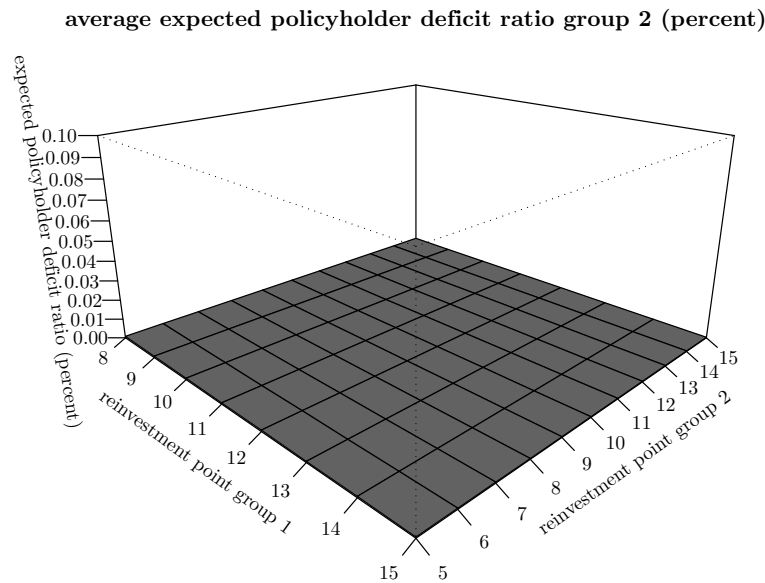
(b) Average expected policyholder deficit ratio group 2

Figure 36: Average expected shortfall ratios, $\frac{1}{T_1} \sum_{t=1}^{T_1} ESR_t^{(1)}$ and $\frac{1}{T_2} \sum_{t=\theta+1}^{\theta+T_2} ESR_t^{(2)}$, given different investment strategies for group 1 and group 2. Group 2 enters the company at $\theta = 5$ and the annuities paid by both policyholder groups are equal, $\pi^{(1)}/\pi^{(2)} = 1$. Investment strategies leading to a certain default are excluded.

average expected shortfall ratio group 1 (percent)

(a) Average expected shortfall ratio group 1

average expected shortfall ratio group 1 (percent)

(b) Average expected shortfall ratio group 2

Figure 37: Average expected policyholder deficit ratios, $\frac{1}{T_1} \sum_{t=1}^{T_1} EPDR_t^{(1)}$ and $\frac{1}{T_2} \sum_{t=\theta+1}^{\theta+T_2} EPDR_t^{(2)}$, given different investment strategies for group 1 and group 2. Group 2 enters the company at $\theta = 5$ and the annuities paid by both policyholder groups are equal, $\pi^{(1)}/\pi^{(2)} = 1$. Investment strategies leading to a certain default are excluded.

This phenomenon is explained by the general persistence of long-term interest rates; see Section 4. However, severity of risk increases if the maximum level of interest rate guarantees is not changed early enough. A premature reaction does not seem to have a major impact.

Figures 39(a) and 39(b) show how the results change for a slightly different ratio of annual premiums paid by policyholder collectives, $\pi^{(1)}/\pi^{(2)}$. Again, shortfall risk does not much change. Figures 40(a) and 40(b) provide sensitivity analysis with respect to the time point of the contract inception for group 2. On the one hand, as the main reason for significantly lower shortfall probabilities is the diversification effects between both groups in the critical initial phase of the contract, we observe that the closer the contract inception of group 2 to the contract inception of group 1, the higher the shortfall probability for both groups. On the other hand, the overall diversification effect for group 2 decreases with increasing time distance between the contract inceptions of both groups. Again, we can explain this phenomenon with diversification effects, which are generally expected to weaken if the time period when contracts of both policyholder collectives are still in force is shortened.

5.3   Economic implications

As portfolios of life insurance companies usually consist of many different policyholder collectives signing contracts at diverse time points, the situation of group 2 seems to be more informative in terms of the general economic implications. In this context, the analysis suggests that the shortfall risk resulting from cliquet-style interest rate guarantees embedded in life insurance products offered on the German market can be low if the insurer has a fairly diversified underwriting portfolio and invests in government bonds held to maturity. In general, a good mix of insurance contracts with diverse points of inception and guaranteed level of interest as well as accurately timed regulatory actions can additionally restrict the severity of risk given default. Those results hold despite the fact that the applied investment strategies are heuristic, i.e., not optimized in terms of risk.
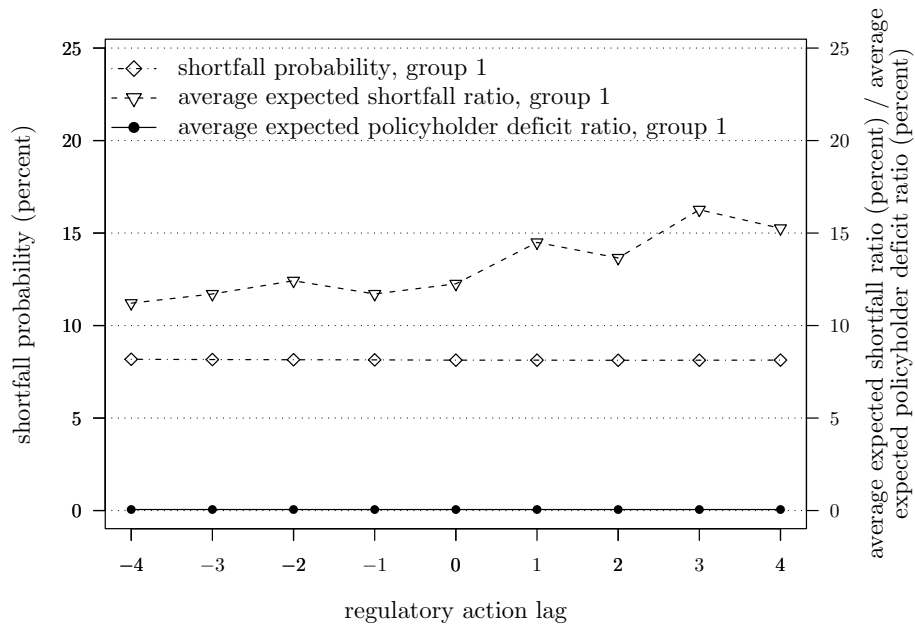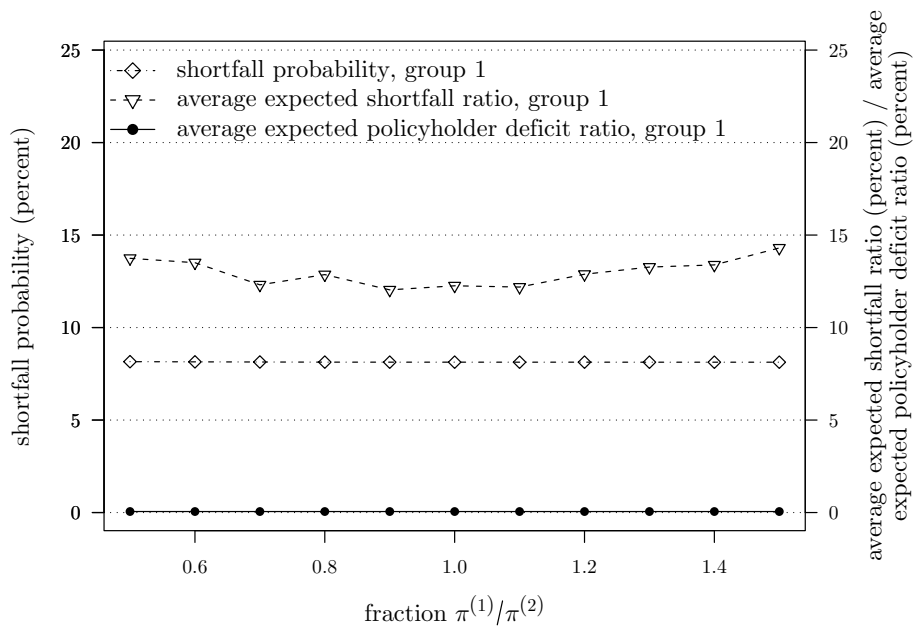
(a) Sensitivity analysis: lag of regulator's reaction, group 1
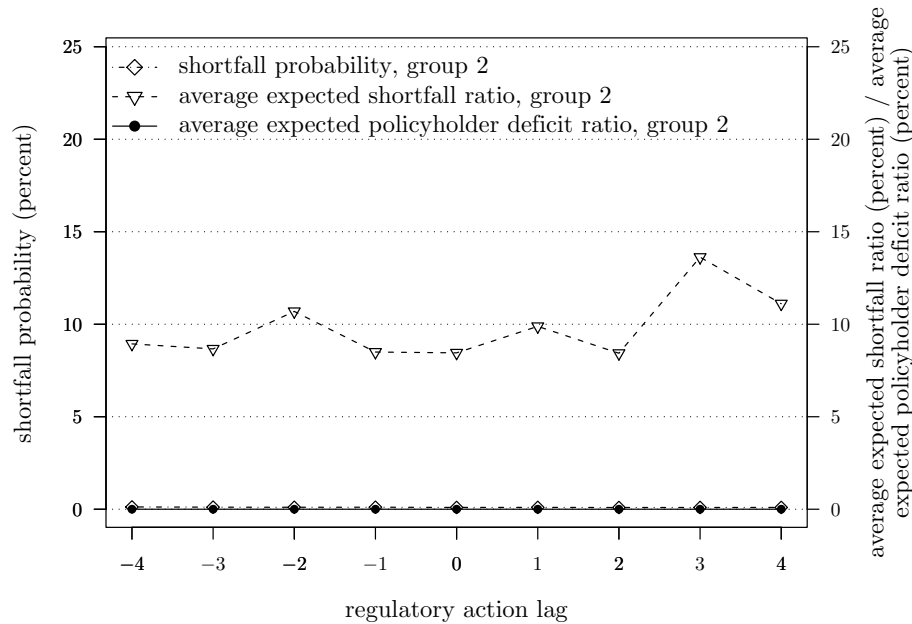
(b) Sensitivity analysis: lag of regulator's reaction, group 2

Figure 38: Sensitivity to the regulator's reaction. The initial investment for both collectives has the same length, $t_1^* = t_2^* = 10$, group 2 enters the company at $\theta = 5$, annual premiums of both collectives are equal, $\pi^{(1)}/\pi^{(2)} = 1$

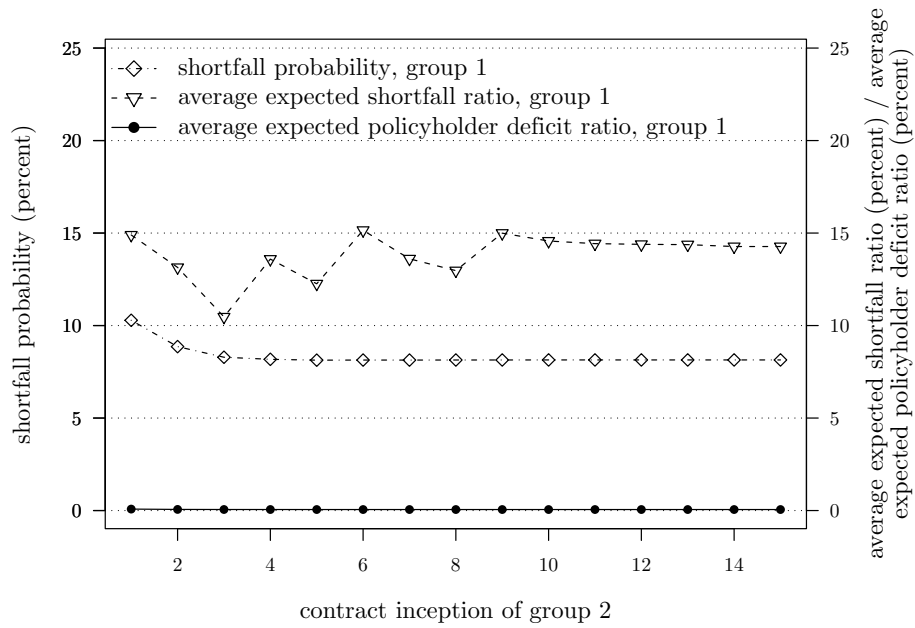(a) Sensitivity analysis: relation of annuities, $\pi^{(1)}/\pi^{(2)}$, group 1
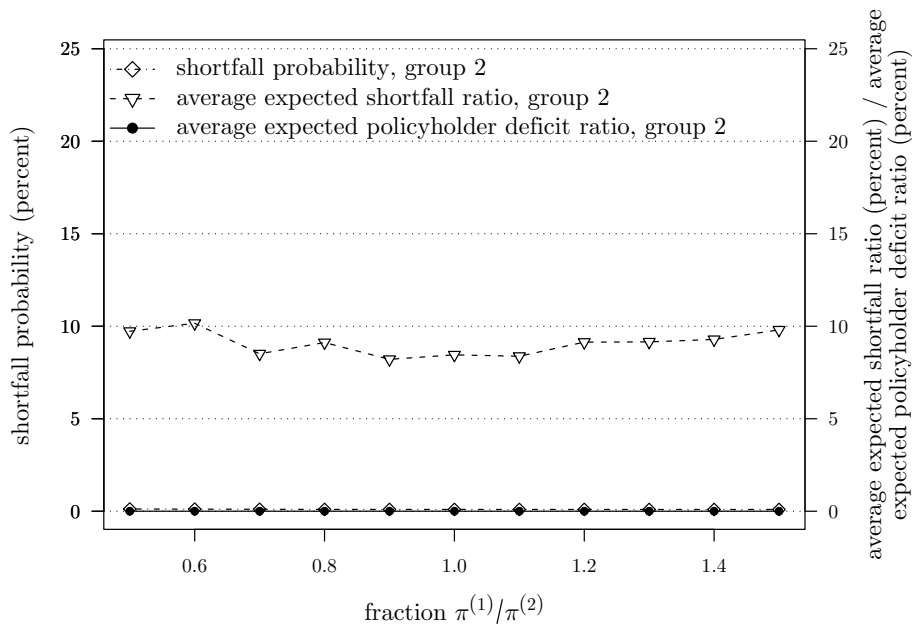
(b) Sensitivity analysis: relation of annuities, $\pi^{(1)}/\pi^{(2)}$, group 2

Figure 39: Sensitivity to the composition of the underwriting portfolio. The initial investment for both collectives has the same length, $t_1^* = t_2^* = 10$, group 2 enters the company at $\theta = 5$, and there is no regulatory lag.

(a) Sensitivity analysis: contract inception of group 2, θ, group 1
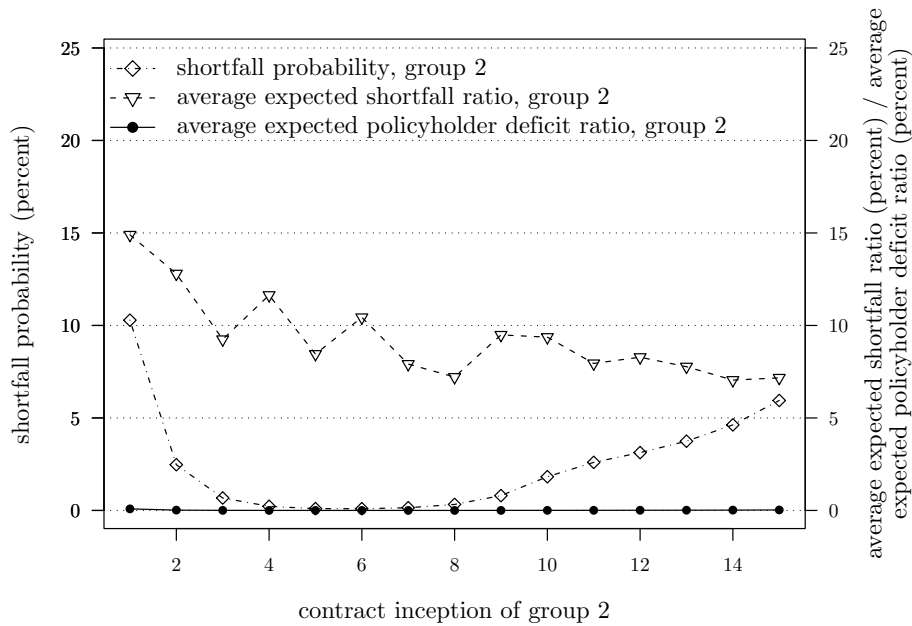
(b) Sensitivity analysis: contract inception of group 2, θ, group 2

Figure 40: Sensitivity to the time point of contract inception for group 2. The initial investment for both collectives has the same length, $t_1^* = t_2^* = 10$, annual premiums of both collectives are equal, $\pi^{(1)}/\pi^{(2)} = 1$, and there is no regulatory lag.

The question of whether similar conclusions can be derived for a real life insurance company which does not invest its entire asset portfolio in government bonds and additionally holds a prescribed amount of solvency capital primarily depends on two issues. First, it is dependent on the actual structure of the asset portfolio of the insurer. As German life insurance companies in general tend to invest the majority of their assets in long-term—both commercial and government—bonds (see GDV, 2010), it is likely that such investments will underlie a significant proportion of the assets corresponding to insurance liabilities, including among others liabilities resulting from interest rate guarantees. Nevertheless, the results may be less relevant if the tied-up assets of an insurer consist at least partially of financial instruments other than bonds. This is also the case if the insurer follows different bond investment strategies to those analyzed. Due to the fact that the level of counterparty default risk of the investments conducted by the companies offering traditional life insurance products is, in general, fairly low (see CEIOPS, 2008), we might expect that the results will also hold for a portfolio with a considerable fraction of corporate bonds.

Second, this analysis is based on a zero bond portfolio. In principle, this simplifying assumption could underestimate the shortfall risk resulting from the reinvestment of the respective coupons. However, as this study includes annual premium payments which are directly invested in line with the current market conditions, it is reasonably to assume that a part of the coupon reinvestment risk is captured. The results are also supported by the fact that the underwriting portfolios of life insurers usually consist of more than two policyholder groups which is likely to further improve diversification.

However, this approach also faces some limitations. As in the analyzed setting the insurer invests solely in government bonds which are used as a proxy for risk-free financial instruments, counterparty default risk is by assumption excluded. Thus, as the value of all assets in the analyzed approach depends on the financial condition of only a small number of debtors, the risk resulting from investments in such a non-diversified portfolio could be significantly underestimated. This will be the case in particular if the assumption that government bonds are nearly risk-free

is not met. Furthermore, no consideration is given to mortality and lapse risk or the risk resulting from variable operating costs, which could also influence the results. In addition, the possibility cannot be excluded of the insurance company following a different investment strategy to the one analyzed, which would certainly have different risk as well as return implications for the policyholders.

Despite the fact that the results of this analysis indicate a significant potential for risk reduction, it would be wrong to claim that limiting reinvestment risk—and hence also the upside potential of an investment—should be seen as a main goal of a financial intermediary. As risk transfer and transformation generally constitute the core functions of insurance companies, the major aim of a life insurer should be to follow investment strategies which comply with policyholders' preferences in terms of both risk and return. This is crucial, as in an imperfect market setting those insured might not be able to achieve—at all or with the same transaction costs—the desired risk-return positions on their own. This analysis is intended instead to show that under certain circumstances a significant reduction of risk resulting from interest rate guarantees is in principle feasible.

6   Conclusion

In this paper, the goal is to analyze the risk implied by the interest rate guarantees embedded in participating life insurance contracts offered in Germany under the assumption of an endogenous regulatory regime and a mutually owned undistributed bonus account. Thus, a Monte Carlo simulation study is implemented based on the widely applied Nelson and Siegel model, which is calibrated with data provided by the Deutsche Bundesbank in line with the methodology developed by Diebold and Li (2006). In this way, it is possible to estimate the empirical development of the German term structure of government bonds.

It is shown that if the insurer's portfolio consists of overlapping policyholder collectives, the interest rate guarantees can in most cases be easily satisfied at a security level compliant with leading regulatory standards, e.g., Solvency II. This is despite the fact that the life insurance

company in the model is not provided with any equity capital at the outset and is only allowed to follow simple investment strategies based on risk-free financial instruments. Due to the structure of the asset and underwriting portfolio of German life insurance companies, these results can under certain conditions be applied to a significant part of the German life insurance market.

References

Albizzati, M.-O. and Geman, H. (1994). Interest Rate Risk Management and Valuation of the Surrender Option in Life Insurance Policies. *Journal of Risk and Insurance*, 61(4):616.

Bacinello, A. R. (2001). Fair Pricing of Life Insurance Participating Policies with a Minimum Interest Rate Guaranteed. *Astin Bulletin*, 31(2):275–297.

Bacinello, A. R. (2003). Fair Valuation of a Guaranteed Life Insurance Participating Contract Embedding a Surrender Option. *Journal of Risk and Insurance*, 70(3):461–487.

Ballotta, L. (2005). A Lévy Process-Based Framework for the Fair Valuation of Participating Life Insurance Contracts. *Insurance: Mathematics and Economics*, 37(2):173–196.

Ballotta, L., Haberman, S., and Wang, N. (2006). Guarantees in With-Profit and Unitized With-Profit Life Insurance Contracts: Fair Valuation Problem in Presence of the Default Option. *Journal of Risk and Insurance*, 73(1):97–121.

Barbarin, J. and Devolter, P. (2005). Risk Measure and Fair Valuation of an Investment Guarantee in Life Insurance. *Insurance: Mathematics and Economics*, 37(2):297–323.

Bauer, D., Kiesel, R., Kling, A., and Ruß, J. (2006). Risk-Neutral Valuation of Participating Life Insurance Contracts. *Insurance: Mathematics and Economics*, 39(2):171–183.

Bernard, C., Courtois, O. L., and Quittard-Pinon, F. (2005). Market Value of Life Insurance Contracts under Stochastic Interest Rates and Default Risk. *Insurance: Mathematics and Economics*, 36(3):499–516.

Brennan, M. J. and Schwartz, E. S. (1976). The Pricing of Equity-Linked Life Insurance Policies with an Asset Value Guarantee. *Journal of Financial Economics*, 3(3):195–213.

Briys, E. and de Varenne, F. (1997). On the Risk of Insurance Liabilities: Debunking Some Common Pitfalls. *Journal of Risk and Insurance*, 64(4):673–694.

Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS) (2008). CEIOPS' Report on its Fourth Quantitative Impact Study (QIS4) for Solvency II. *URL http://www.ceiops.eu.*

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385–407.

Cummins, J. D., Miltersen, K. R., and Persson, S.-A. (2004). International Comparison of Interest Rate Guarantees in Life Insurance. *Working Paper, Norwegian School of Economics and Business Administration.*

Deutsche Bundesbank (1997). Stripping of Bunds. *Deutsche Bundesbank Monthly Report*, 7:17–22.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, 74(366):427–431.

Diebold, F. X. and Li, C. (2006). Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics*, 130(2):337–364.

Duffee, G. R. (2002). Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance*, 57(1):405–443.

Farny, D. (2011). *Versicherungsbetriebslehre.* VVW Karlsruhe, 5th edition.

Faust, R., Schmeiser, H., and Zemp, A. (2010). A Performance Analysis of Participating Life Insurance Contracts. *Working Paper, University of St. Gallen.*

Gatzert, N. (2008). Asset Management and Surplus Distribution Strategies in Life Insurance: An Examination with Respect to Risk Pricing and Risk Measurement. *Insurance: Mathematics and Economics*, 42(2):839–849.

Gatzert, N. and Kling, A. (2007). Analysis of Participating Life Insurance Contracts: A Unification Approach. *Journal of Risk and Insurance*, 74(3):547–570.

Gerstner, T., Griebel, M., Holz, M., Goschnick, R., and Haep, M. (2008). A General AssetLiability Management Model for the Efficient Simulation of Portfolios of Life Insurance Policies. *Insurance: Mathematics and Economics*, 42(2):704–716.

Gesamtverband der Deutschen Versicherungswirtschaft (GDV) (2010). Statistisches Taschenbuch der Versicherungswirtschaft 2010.

Grosen, A. and Jørgensen, P. L. (2000). Fair Valuation of Life Insurance Liabilities: the Impact of Interest Rate Guarantees, Surrender Options, and Bonus Policies. *Insurance: Mathematics and Economics*, 26(1):37–57.

Grosen, A. and Jørgensen, P. L. (2002). Life Insurance Liabilities at Market Value: An Analysis of Insolvency Risk, Bonus Policy, and Regulatory Intervention Rules in a Barrier Option Framework. *Journal of Risk and Insurance*, 69(1):63–91.

Gründl, H., Stehle, R., and Waldow, T. (2003). Zur Vorteilhaftigkeit von Kapitallebensversicherungen gegenüber alternativen Anlageformen - Eine Analyse aus Anlegersicht. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 55(6):549–577.

Haberman, S., Ballotta, L., and Wang, N. (2003). Modelling and Valuation of Guarantees in With-Profit and Unitised With Profit Life Insurance Contracts. *Working Paper, Cass Business School.*

Hansen, M. and Miltersen, K. (2002). Minimum Rate of Return Guarantees: The Danish Case. *Scandinavian Actuarial Journal*, 2002(4):280–318.

Heath, D., Jarrow, R., and Morton, A. (1992). Bond Pricing and the Term Structure of Interest Rates : A New Methodology for Contingent Claims Valuation. *Econometrica*, 60(1):77–105.

Jensen, B., Jørgensen, P. L., and Grosen, A. (2001). A Finite Difference Approach to the Valuation of Path Dependent Life Insurance Liabilities. *The Geneva Papers on Risk and Insurance Theory*, 26(1):57–84.

Kleinow, T. (2009). Valuation and Hedging of Participating Life-Insurance Policies under Management Discretion. *Insurance: Mathematics and Economics*, 44(1):78–87.

Kleinow, T. and Willder, M. (2007). The Effect of Management Discretion on Hedging and Fair Valuation of Participating Policies with Maturity Guarantees. *Insurance: Mathematics and Economics*, 40(3):445–458.

Kling, A., Richter, A., and Ruß, J. (2007a). The Impact of Surplus Distribution on the Risk Exposure of With Profit Life Insurance Policies Including Interest Rate Guarantees. *Journal of Risk and Insurance*, 74(3):571–589.

Kling, A., Richter, A., and Ruß, J. (2007b). The Interaction of Guarantees, Surplus Distribution, and Asset Allocation in With-Profit Life Insurance Policies. *Insurance: Mathematics and Economics*, 40(1):164–178.

MacKinnon, J. G. (1996). Numerical Distribution Functions for Unit Root and Cointegration Tests. *Journal of Applied Econometrics*, 11(6):601–618.

Miltersen, K. R. and Persson, S.-A. (1999). Pricing rate of return guarantees in a HeathJarrowMorton framework. *Insurance: Mathematics and Economics*, 25(3):307–325.

Nelson, C. R. and Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves. *Journal of Business*, 60(4):473–489.

Persson, S.-A. and Aase, K. K. (1997). Valuation of the Minimum Guaranteed Return Embedded in Life Insurance Products. *Journal of Risk and Insurance*, 64(4):599–617.

Phillips, P. C. B. and Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika*, 75(2):335–346.

Prieul, D., Putyatin, V., and Nassar, T. (2001). On Pricing and Reserving With-Profits Life Insurance Contracts. *Applied Mathematical Finance*, 2001(8):145–166.

Said, S. E. and Dickey, D. A. (1984). Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika*, 71(3):599–607.

Schich, S. T. (1997). Estimating the German Term Structure. *Economic Research Group of the Deutsche Bundesbank, Discussion Paper*, 4/1997.

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, 6(2):461–464.

Siegel, A. F. and Nelson, C. R. (1988). Long-Term Behavior of Yield Curves. *Journal of Financial and Quantitative Analysis*, 23(1):105–110.

Siu, T. K. (2005). Fair Valuation of Participating Policies with Surrender Options and Regime Switching. *Insurance: Mathematics and Economics*, 37(3):533–552.

Svensson, L. E. (1994). Estimating and Interpreting Forward Interest Rates: Sweden 1992 - 1994. *NBER Working Paper Series*, 4871.

Tanskanen, A. J. and Lukkarinen, J. (2003). Fair Valuation of Path-Dependent Participating Life Insurance Contracts. *Insurance: Mathematics and Economics*, 33(3):595–609.

Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5(2):177–188.

Waldow, T. (2003). *Die Bewertung der gemischten Kapitallebensversicherung*. VVW Karlsruhe.

Wörner, G. (2003). *Handels- und Steuerbilanz nach neuem Recht*. Oldenbourg, 8th edition.

Zaglauer, K. and Bauer, D. (2008). Risk-Neutral Valuation of Participating Life Insurance Contracts in a Stochastic Interest Rate Environment. *Insurance: Mathematics and Economics*, 43(1):29–40.

Zemp, A. (2010). Risk Comparison of Different Bonus Distribution Approaches in Participating Life Insurance. *Working Paper, University of St. Gallen*.

Curriculum Vitae

Personal Information

Name	Przemysław Zbigniew Rymaszewski
Date of Birth	March 13, 1985
Place of Birth	Koszalin, Poland
Nationality	Polish

Education

07/2009 − present	**University of St. Gallen**, St. Gallen, Switzerland Doctoral Candidate in Management (Specialization: Finance)
07/2010 − 08/2010	**University of Michigan**, Ann Arbor, USA Summer Program in Quantitative Research Methods
10/2004 − 06/2009	**University of Cologne**, Cologne, Germany Diploma in Business Administration
08/2008 − 12/2008	**Stockholm School of Economics**, Stockholm, Sweden Student Exchange
09/1999 − 06/2004	**Dubois Liceum**, Koszalin, Poland General Qualification for University Entrance

Work Experience

07/2009 − present	**Institute of Insurance Economics**, **University of St. Gallen**, St. Gallen, Switzerland Project Manager and Research Associate
02/2008 − 04/2008	**Boston Consulting Group**, Cologne, Germany Visiting Associate
08/2007 − 10/2007	**Bear Stearns International**, London, UK Intern
04/2006 − 06/2008	**University of Cologne**, Cologne, Germany Tutor in Finance and Insurance
03/2006 − 04/2008	**Bank PeKaO SA**, Koszalin, Poland Intern