Summer 2018

Essays on investment and risk management

Jie Ying
University of Iowa

ESSAYS ON INVESTMENT AND RISK MANAGEMENT

by

Jie Ying

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Business Administration
in the Graduate College of
The University of Iowa

August 2018

Thesis Supervisor: Professor David S. Bates

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Jie Ying

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Business Administration at the August 2018 graduation.

Thesis Committee: ______________________________
David S. Bates, Thesis Supervisor

____
Thomas R. Berry-Stölzle

____
Wei Li

____
Richard Peter

____
Tong Yao

To my parents

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my advisor and dissertation committee chair, David S. Bates, who gave me the freedom to pursue my own research interests while providing tremendous support throughout my graduate study at University of Iowa. Without his advice and persistent help this dissertation would not have been possible.

In addition, I would like to thank the other members of my dissertation committee: Thomas R. Berry-Stölzle, Wei Li, Richard Peter, and Tong Yao. They have been great teachers and coauthors, and have provided me invaluable guidance and encouragement for my research, teaching, and career development in academia. I am also grateful to all the faculty members and Ph.D. students at the Department of Finance, who have been very helpful during my time in Iowa.

Last but not least, I owe my sincere gratitude to my beloved parents, Linhan and Jin, who have always been there for me.

ABSTRACT

In this dissertation, I consider a range of topics in investment and risk management. I seek to understand several existing yet puzzling phenomena from a theoretical perspective.

In the first chapter, we study the optimal insurance demand of a risk- and ambiguity-averse consumer if contract nonperformance risk is perceived as ambiguous. Ambiguity lowers optimal insurance demand and the consumer's degree of ambiguity aversion is negatively associated with the optimal level of coverage. Biased beliefs and greater ambiguity may increase or decrease the optimal demand for insurance, and we determine sufficient conditions for a negative effect. We also discuss wealth effects and evaluate the robustness of our results by considering several alternative models of ambiguity aversion. Our findings show how ambiguous nonperformance risk can undermine the functioning of insurance markets, making it a concern for regulators. Caution is required though because, as we show, demand reactions are only imperfectly informative about the welfare effects of nonperformance risk.

In the second chapter, we address an ongoing debate on pension investment policy: should defined-benefit corporate pension plans invest aggressively in risky securities or completely de-risk their assets? In our model, firms maximize shareholder value subject to the participation constraint of employees, who are wealth-constrained and are partially exposed to pension investment risk via a corporate bankruptcy channel and a pension surplus sharing channel. For a reasonable set of parameter values, the model-suggested optimal

pension allocation to risky assets exceeds 50%. The level of pension risk-taking predicted by the model, and its relation with a firm's bankruptcy probability and pension funding ratio, match with empirical observations. We show that due to limited sharing of the investment risk by employees. Defined-benefit pensions may take on even more risk than what employees choose in the defined contribution plans. Further, firms may substantially reduce their overall pension funding costs under an alternative arrangement in which employees bear all the systematic pension investment risk. This is consistent with the secular trend of firms switching from defined benefit plans to defined contribution plans.

In the third chapter, I model the shadow banking mechanism and discusses its functionality of risk sharing and its impact on financial instability. In equilibrium, the shadow banking becomes more active when investors perceive higher expected returns from the capital market. The shadow banking yield arises when the capital market gets more volatile. Lower interest rates from regulated banks encourage the shadow banking and the magnitude of impacts depends on investors' aggregate risk preference. Overactive shadow banking activities can "cool down" themselves. The shadow banking's influence over the economy is twofold: it improves the overall welfare of heterogeneous agents by risk sharing, but it spreads the risk through the financing channel, which makes the savers more vulnerable to the negative shocks in financial markets.

PUBLIC ABSTRACT

My dissertation consists of three essays on investment and risk management. It answers three questions from a theoretical perspective: how customers choose the insurance coverage when the insurance contract may fail to perform as intended; how firms manage their pension plans considering the interests of both shareholders and employees; how the shadow banking system facilitates the risk sharing between lenders and borrowers.

In the first essay, our model shows that the ambiguity of the insurance contract's nonperformance risk lowers customers' insurance demand and the negative impact is stronger for customers who are more averse to ambiguity. Besides, biased beliefs and greater ambiguity can reduce insurance demand under certain conditions.

In the second essay, we answer the question "Should defined-benefit corporate pension plans invest in risky assets at all?" The answer is yes, and the level of risky investment depends on a variety of factors such as the firm's bankruptcy probability, the internal and external financing costs, the way a pension surplus is shared between the firm and employees, etc. Our model confirms that the DC plan is a more cost-efficient option for firms than the DB plan.

In the third essay, my model predicts a more active shadow banking system when a higher return is anticipated in the capital market and a higher shadow banking yield caused by the market's higher volatility. The model prediction also suggests that a lower interest rate from regulated banks encourages the shadow banking and overactive shadow banking activities can "cool down" themselves.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1
OPTIMAL INSURANCE DEMAND WHEN CONTRACT NONPERFORMANCE RISK IS PERCEIVED AS AMBIGUOUS

1.1 Introduction

Insurance allows risk-averse consumers to protect themselves against financial risks. In a frictionless world with perfect information, full insurance is optimal if the price is actuarially fair. At an unfair price, the consumer trades off expected wealth for a lower riskiness of his wealth distribution and chooses partial insurance (Mossin, 1968; Smith, 1968). Over the past decades, the theory of insurance demand has flourished and some claim it to be *"the purest example of economic behavior under uncertainty"* (Schlesinger, 2013).

It has been recognized though that insurance contracts may fail to perform (Doherty and Schlesinger, 1990), and much of insurance regulation centers around issues of nonperformance. The most important reason is insurer insolvency, which can invalidate the policyholders' claims. Even with guarantee funds, insurer insolvency results in uncertainty for the consumer as to the extent and timing of indemnification. Furthermore, the possibility of claims being contested in front of the courts may result in lack of coverage, and similarly, delays in the insurer's claims handling process, subtleties in the contractual language of the insurance policy or probationary periods can leave consumers uncovered despite having paid a premium to the insurance company.[1] Sometimes, the failure to perform is even considered as a specific feature of some insurance contracts such as the index insurance, or rainfall insurance, in which case the contract fails to make payment due to some basis risk (Morduch, 2006).

As stressed by Schlesinger (2013), it is the consumer's perception of nonperformance risk that causes a deviation of his insurance demand behavior relative to a situation without nonperformance risk. We argue in this paper that, due to information asymmetry, consumers will hardly know the odds of a contract not working as intended at the time of purchase. Indeed, all the aforementioned examples can create significant uncertainty for consumers when it comes to the performance of their insurance contracts. Contract nonperformance risk is likely to be perceived as ambiguous, and it is the implications of this ambiguity that we study in this paper.[2]

To achieve this goal we employ the model of smooth ambiguity aversion developed by Klibanoff et al. (2005) in the main part of the analysis. It disentangles ambiguity from ambiguity attitude and is particularly amenable for the type of comparative statics analysis conducted in this paper. Following the definition by Camerer and Weber (1992), *"ambi-*

---

[1]Crocker and Tennyson (2002) show that certain claims that are viewed as being easy to falsify, are less certain to be paid in full; also Tennyson and Warfel (2009) and Asmat and Tennyson (2014) provide evidence of claims underpayment and discuss the effect of the legal framework on the insurers' settlement and verification practices. Bourgeon and Picard (2014) develop an economic model of insurer "nitpicking", which reduces the efficiency of insurance contracts and can lead to substantial uncertainty about the performance of the contract. A somewhat related issue is that of nonveriable losses, see Doherty et al. (2013).

[2]Consistent with most of the literature we understand risk as a condition in which the event to be realized is unknown, but the odds of all possible events are perfectly known, either subjectively or objectively. Ambiguity refers to a condition where the probabilities of possible events are not uniquely assigned, see, for example, Iwaki and Osaki (2014).

guity is uncertainty about probability created by missing information that is relevant and could be known". This applies well to nonperformance risk in insurance contracts because at the time consumers enter into the agreement, they do not know the precise probability of the contract's performance in the future. That individuals are indeed averse to such ambiguity was first demonstrated in Ellsberg's (1961) seminal paper. Since then, ambiguity aversion has been documented in numerous laboratory experiments (e.g., Einhorn and Hogarth, 1986; Chow and Sarin, 2001), in market settings with educated individuals (see Sarin and Weber, 1993), and in surveys of business owners and managers (see Viscusi and Chesson, 1999; Chesson and Viscusi, 2003).[3] Recent survey evidence of U.S. households confirms the relevance of ambiguity aversion for financial decisions in the field (see Dimmock et al., 2016). Whereas the welfare consequences of ambiguity for ambiguity-averse agents are immediate, its behavioral implications are not, which opens up an interesting area of study (see, for example Gollier, 2011). Specifically, there is a growing literature on the effects of ambiguity in insurance (e.g., Alary et al., 2013; Huang et al., 2013; Gollier, 2014; Bajtelsmit et al., 2015) as part of the field of behavioral insurance (see Richter et al., 2014), to which this paper contributes.

Existing studies on nonperformance risk in insurance have largely focused on situations without ambiguity. On the theory side, Doherty and Schlesinger (1990, DS henceforth) provide the first expected utility analysis and find that most comparative statics of insurance demand are not upheld under nonperformance risk. Mossin's Theorem is violated

[3]For a recent summary of the literature and the boundary conditions of ambiguity aversion in the laboratory, see Kocher et al. (2018). For the prevalence of ambiguity attitudes in a large representative sample, see Dimmock et al. (2015b,a).

(see also Mahul and Wright, 2007), an increase in risk aversion does not necessarily raise the optimal level of coverage, and insurance may not be an inferior good under decreasing absolute risk aversion. One might thus expect ambiguity to further obfuscate the economic trade-offs associated with the insurance decision. The analysis of nonperformance risk has also been extended to the insurer-reinsurer relationship (Bernard and Ludkovski, 2012), to risk management instruments other than insurance (Briys et al., 1991; Schlesinger, 1993), and to include heterogeneous beliefs about nonperformance (Cummins and Mahul, 2003).[4]

There is a small but growing literature that studies the effects of nonperformance risk on insurance in an environment with ambiguity. Biener et al. (2017) show that the introduction of ambiguity associated with nonperformance risk raises the marginal willingness to pay at a given level of coverage (their Lemma 3). In a field experiment, ambiguity is found to reduce take-up by 14.5%. Bryan (2018) shows that mandatory rainfall insurance reduces the adoption rates of a new farming technology for ambiguity-averse farmers (his prediction 1). He also presents evidence from two RCTs and documents substantial income losses associated with ambiguity. We complement and extend these papers along several dimensions contributing to this strand of research with a more systematic analysis and discussion. First, we allow decision-makers to select a level of coverage to provide a direct answer to the question how optimal demand is affected by ambiguous nonperformance

[4]Other theoretical works incorporate taxation (Huang and Tzeng, 2007) or allow default risk to arise endogenously (Biffis and Millossovich, 2012). Recently, Liu and Myers (2016) develop a dynamic model of nonperformance risk and show that it reduces the demand for microinsurance.

risk.[5] Second, we provide comparative statics with respect to the degree of ambiguity aversion, the level of ambiguity and initial wealth. As we will show, some of the existing conclusions may not generalize or only under appropriate qualifications. This reveals new economic trade-offs and new testable hypothesis that are not apparent when comparing the ambiguous to the non-ambiguous scenario. Third, we assess the validity of our results by considering alternative models of ambiguity aversion.

That nonperformance risk does indeed affect behavior has been well established in the literature. Kahneman and Tversky (1979) introduce the concept of 'probabilistic insurance', which incorporates an element of nonperformance risk but also exhibits probabilistic repayment of the premium, which alters the economic trade-offs. Subjects appear to dislike both probabilistic insurance as well as pure nonperformance risk and adjust insurance demand accordingly, as suggested by various surveys (Kahneman and Tversky, 1979; Wakker et al., 1997; Zimmer et al., 2009) and evidenced in incentive-compatible laboratory experiments (Herrero et al., 2006; Zimmer et al., 2018). What's more, nonperformance risk reduces welfare of risk-averse individuals directly, because contracts are less efficient, and indirectly, because the compound nature of nonperformance induces violations of the reduction of compound lotteries axiom (Harrison and Ng, 2018). Biener et al. (2017) conduct a field experiment on insurance with nonperformance risk and find similar effects.

[5] Marginal willingness to pay, as studied in Biener et al. (2017), does not inform about optimal demand. Jaspersen (2016) provides an explicit example in Fn. 6 where the individual with lower willingness to pay prefers a contract with higher coverage than the individual with higher willingness to pay. This is also why Alary et al. (2013) study both concepts separately and find conditions specific to each when signing the comparative statics. In Bryan (2018), insurance is mandatory precluding questions of optimal demand. What's more, his set-up is one of basis risk so there is a chance of receiving an insurance indemnity even if no loss happened. This is a departure from nonperformance risk in the DS sense.

From a theoretical perspective, the behavioral implications of an ambiguous perception of nonperformance risk on insurance demand are not immediate. Snow (2011) and Alary et al. (2013) show that ambiguity aversion raises the optimal level of coverage for a binary loss distribution that is perceived as ambiguous. The case of nonperformance risk is different; under ambiguity, the consumers' perception will include probabilistic scenarios in which the insurance contract performs well, suggesting to buy more coverage, and scenarios in which the insurance contract performs poorly, suggesting to reduce coverage. The equilibrium effect is indeterminate *a priori*, which warrants our analysis. We are, however, able to establish a definitive negative net effect, which can be rationalized by a distorted probability distribution, that induces the consumer to behave as if he was more pessimistic about contract nonperformance risk. We complement our analysis by several comparative statics results and also show that our main findings do not depend on how ambiguity preferences are implemented. This allows us to conclude that the identified demand effects can be directly attributed to consumers' ambiguity aversion.

The main implication of our theory is that an ambiguous perception of nonperformance risk can be a significant behavioral impediment to insurance demand and can therefore undermine the functioning of insurance markets in the economy. This adds to the literature that shows how ambiguity compromises the effectiveness of financial instruments and markets (see Mukerji and Tallon, 2001). Regulations targeting the demand side, for example, in the form of providing accurate information to consumers, and on the supply side by tailoring insurer capital provisions accordingly should aim at mitigating uncertainty associated with nonperformance risk. Our results also show that demand reactions may not be indicative of welfare changes, making them only partially informative for policymakers.

1.2 Model and Notations

We consider a consumer with initial wealth W which is subject to a loss of $L \in (0, W)$ with probability $p \in (0, 1)$. Risk preferences are represented by a vNM utility function u of final wealth, which is increasing and concave, $u' > 0$ and $u'' < 0$, to reflect non-satiation and risk aversion. The consumer can insure against the risk: $\alpha \in [0, 1]$ denotes the level of coverage and αL the associated indemnity payment in case of a loss. Insurance requires payment of an insurance premium $P(\alpha)$ which is assumed to be an increasing and non-concave function of the level of coverage, $P' > 0$ and $P'' \geq 0$.

As in DS, insurance contracts are imperfect because they may not perform. This occurs with probability $(1 - q) \in (0, 1)$ whereas with probability q the contract works as intended and the indemnity is paid. We focus on the consumer's perception of nonperformance risk. More specifically, we assume that policyholders do not know the probability of nonperformance for sure and rather perceive nonperformance as ambiguous. This is modeled with an idiosyncratic distribution over q, so that we rewrite $(1 - \widetilde{q})$ to indicate probabilistic uncertainty over nonperformance risk. If $\overline{q}$ denotes the expected value of $\widetilde{q}$, we can define the zero-mean risk $\widetilde{\varepsilon} = \overline{q} - \widetilde{q}$. Without loss of generality, we can then decompose the consumer's perception into an unbiased belief and an additive noise term, $(1 - \widetilde{q}) = (1 - \overline{q}) + \widetilde{\varepsilon}$. This has the advantage that $\widetilde{\varepsilon}$ can be interpreted as the level of ambiguity. By construction, it holds that $\mathbb{E}\widetilde{\varepsilon} = 0$, and the support of $\widetilde{\varepsilon}$ is contained in $[\underline{\varepsilon}, \overline{\varepsilon}] \subseteq [-(1 - \overline{q}), \overline{q}]$.

Three different states of the world are relevant for the analysis. $W_1 = W - P$ is the consumer's final wealth if he does not suffer a loss, in which case contract nonperformance risk is irrelevant. $W_2 = W - P - L + \alpha L$ is the consumer's final wealth when a loss occurs and the insurance contract performs, and $W_3 = W - P - L$ is the consumer's final wealth when he suffers a loss but the contract does not perform. As noted by DS, in hindsight the consumer would have been better off in this last state had he not purchased insurance to save on premium money. For simplicity, we assume that there is no recovery in case of nonperformance so that no indemnity is paid at all.[6]

With these specifications, the different levels of the consumer's expected utility can be represented as follows:

$$U(\alpha, \widetilde{\varepsilon}) = (1-p)u(W_1) + p(\overline{q} - \widetilde{\varepsilon})u(W_2) + p(1 - \overline{q} + \widetilde{\varepsilon})u(W_3). \tag{1.1}$$

Due to ambiguity the consumer's expected utility is a random variable for a given level of coverage with outcomes depending on the realization of $\widetilde{\varepsilon}$. To incorporate the consumer's attitude towards ambiguity, we adopt the approach by Klibanoff et al. (2005) and derive the consumer's objective function as a ϕ-weighted expectation over different expected utilities based on second-order beliefs (see also Neilson, 2010). ϕ is the consumer's ambiguity function, which is increasing and concave, $\phi' > 0$ and $\phi'' < 0$, to reflect non-satiation (in expected utility) and ambiguity aversion. The consumer's ex-ante welfare as a function of

[6] This assumption is without loss of generality because the ordering of the different levels of final wealth is given by $W_3 < W_2 \leq W_1$ for any recovery rate. The last inequality is strict for partial insurance coverage.

the level of coverage is then given by

$$V(\alpha) = \mathbb{E}\phi\left[U(\alpha, \widetilde{\varepsilon})\right]. \tag{1.2}$$

The expectation is taken with respect to the cumulative distribution function F, which describes the consumer's second-order beliefs over the probability of contract nonperformance.

1.3 The Optimal Level of Coverage

The optimal level of insurance coverage is the one that maximizes the consumer's ex-ante welfare. For an interior solution, it is characterized by the associated first-order condition,

$$V'(\alpha) = \mathbb{E}\left\{\phi'\left[U(\alpha, \widetilde{\varepsilon})\right] U_\alpha(\alpha, \widetilde{\varepsilon})\right\} = 0, \tag{1.3}$$

where subscript α denotes the partial derivative with respect to the level of coverage. The consumer's objective function is globally concave in α,

$$V''(\alpha) = \mathbb{E}\left\{\phi''\left[U(\alpha, \widetilde{\varepsilon})\right] U_\alpha(\alpha, \widetilde{\varepsilon})^2 + \phi'\left[U(\alpha, \widetilde{\varepsilon})\right] U_{\alpha\alpha}(\alpha, \widetilde{\varepsilon})\right\} < 0. \tag{1.4}$$

Due to ambiguity aversion, $\phi''\left[U(\alpha, \varepsilon)\right] < 0$ for every ε, the realization of $\widetilde{\varepsilon}$, and due to risk aversion and the non-concavity of the premium schedule we obtain that $U_{\alpha\alpha}(\alpha, \varepsilon) < 0$ for every ε. Concavity of the objective function ensures that the first-order condition characterizes a maximum. In the sequel, we assume that it is optimal to purchase at least some insurance and denote the optimal level of coverage by $\alpha^* > 0$.[7]

[7] This is the case as long as $P'(0)$ is less than a threshold value, that can be derived by rearranging $V'(0) > 0$. This threshold value depends on the consumer's risk and ambiguity preferences.

To disentangle the effects of ambiguity, we decompose $U_\alpha(\alpha, \widetilde{\varepsilon})$ into the marginal utility cost and the marginal utility benefit:

$$\begin{aligned} MC(\alpha, \widetilde{\varepsilon}) &= P'(\alpha)\left[(1-p)u'(W_1) + p(\overline{q} - \widetilde{\varepsilon})u'(W_2) + p(1 - \overline{q} + \widetilde{\varepsilon})u'(W_3)\right], && (1.5) \\ MB(\alpha, \widetilde{\varepsilon}) &= Lp(\overline{q} - \widetilde{\varepsilon})u'(W_2). && (1.6) \end{aligned}$$

Insurance obligates the consumer to pay a premium which reduces his final wealth in any state. However, the indemnity increases final wealth if the loss occurs and the contract performs as intended. With these notations, the first-order condition can be rewritten as follows:

$$\mathbb{E}\left\{\phi'\left[U(\alpha, \widetilde{\varepsilon})\right] MC(\alpha, \widetilde{\varepsilon})\right\} = \mathbb{E}\left\{\phi'\left[U(\alpha, \widetilde{\varepsilon})\right] MB(\alpha, \widetilde{\varepsilon})\right\}. \quad (1.7)$$

In the absence of ambiguity ($\widetilde{\varepsilon} \equiv 0$), the consumer's expected utility is non-random and the optimal level of coverage is characterized by:

$$U_\alpha(\alpha, 0) = -MC(\alpha, 0) + MB(\alpha, 0) = 0. \quad (1.8)$$

We denote it by α^0, and it is the level of coverage studied in DS for the special case of a linear premium schedule. To develop some intuition, we investigate how the ambiguity associated with the risk of contract nonperformance affects the marginal utility cost and the marginal utility benefit of insurance. For the former, we derive

$$MC\left(\alpha^0, \widetilde{\varepsilon}\right) - MC\left(\alpha^0, 0\right) = P'(\alpha^0)p\widetilde{\varepsilon}\left(u'(W_3) - u'(W_2)\right). \quad (1.9)$$

The expression in round brackets is positive due to risk aversion. As a result, the overall sign coincides with the sign of the realization of $\widetilde{\varepsilon}$. If $\widetilde{\varepsilon} = \varepsilon > 0$, the consumer perceives contract nonperformance as more likely relative to a situation without ambiguity. This increases the marginal utility cost of insurance because incurring the insurance premium is most painful in precisely that state of the world where the insurance contract does not perform. Conversely, if $\widetilde{\varepsilon} = \varepsilon < 0$, the consumer perceives contract nonperformance to be less likely relative to a situation without ambiguity, which decreases the marginal utility cost of insurance.

For the marginal utility benefit, we obtain

$$MB\left(\alpha^0, \widetilde{\varepsilon}\right) - MB\left(\alpha^0, 0\right) = -pL\widetilde{\varepsilon}u'(W_2). \tag{1.10}$$

The sign is the opposite of the sign of the realization of $\widetilde{\varepsilon}$. If $\widetilde{\varepsilon} = \varepsilon > 0$, the consumer perceives contract nonperformance as more likely relative to a situation without ambiguity. This reduces the marginal utility benefit of insurance because the consumer is less likely to receive the indemnity in case of a loss. Conversely, if $\widetilde{\varepsilon} = \varepsilon < 0$, the consumer perceives contract nonperformance to be less likely relative to a situation without ambiguity, which increases the marginal utility benefit of insurance. It becomes clear that any positive realization of $\widetilde{\varepsilon}$ increases the marginal utility cost and decreases the marginal utility benefit of insurance, inducing the consumer to lower his demand relative to α^0, whereas any negative realization of $\widetilde{\varepsilon}$ has exactly the oppositve effects. Will these opposing effects cancel each other out in equilibrium? The following proposition establishes that the net effect has a definitive negative sign. All proofs are gathered in the appendix.

Proposition 1.3.1: *An ambiguous perception of nonperformance risk lowers the optimal demand for insurance for a risk- and ambiguity-averse consumer.*

To develop some intuition for this result, we return to the first-order condition (1.3) and evaluate it at the optimal level of coverage without ambiguity, which yields

$$\begin{aligned} V'(\alpha^0) =& \ \mathbb{E}\left\{\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]U_\alpha(\alpha^0,\widetilde{\varepsilon})\right\} \\ =& \ -\underbrace{\mathbb{E}\left\{\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]MC(\alpha^0,\widetilde{\varepsilon})\right\}}_{\text{unconditional marginal cost}}+\underbrace{\mathbb{E}\left\{\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]MB(\alpha^0,\widetilde{\varepsilon})\right\}}_{\text{unconditional marginal benefit}}. \end{aligned} \tag{1.11}$$

From the proof of Proposition 1.3.1 we know that $\phi'\left[U(\alpha^0,\varepsilon)\right]$ is increasing in ε if the consumer is ambiguity-averse. Scenarios with $\varepsilon < 0$ receive relatively less weight than scenarios with $\varepsilon > 0$. Said differently, an ambiguity-averse consumer behaves as if he was an expected utility maximizer who overweights scenarios in which the marginal utility cost is higher than the marginal utility benefit of insurance and underweights scenarios in which the marginal utility cost is lower than the marginal utility benefit of insurance. In the spirit of Gollier (2011), we can rewrite the consumer's first-order expression, evaluated at α^0, as follows:

$$\begin{aligned} V'(\alpha^0) =& \ \mathbb{E}\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]\mathbb{E}\left\{\frac{\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]}{\mathbb{E}\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]}U_\alpha(\alpha^0,\widetilde{\varepsilon})\right\} = \\ =& \ \mathbb{E}\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]\mathbb{E}^{Q^0}\left\{U_\alpha(\alpha^0,\widetilde{\varepsilon})\right\}, \end{aligned} \tag{1.12}$$

where Q^0 denotes a distorted probability measure. The ambiguity-averse consumer behaves in the same way as an expected-utility maximizing consumer who has distorted his second-order beliefs.[8] The associated distortion factor $\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]/\mathbb{E}\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]$ is a Radon-Nikodym derivative that describes the change of measure, and $\mathbb{E}^{Q^0}$ is the ambiguity-neutral expectation which corresponds to the risk-neutral expectation in finance. Our chosen notation is suggestive of the fact that this distorted measure is endogenous because it depends on the level of coverage selected by the consumer. Under Q^0, the consumer perceives contract

[8]Gollier (2011) refers to this interpretation as "observational equivalence".

nonperformance as more likely than under the physical probability. The reason is that

$$\mathbb{E}^{Q^0}\left\{1-\overline{q}+\widetilde{\varepsilon}\right\}=\mathbb{E}\left\{\frac{\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]}{\mathbb{E}\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]}(1-\overline{q}+\widetilde{\varepsilon})\right\}=(1-\overline{q})+\frac{\mathrm{Cov}\left\{\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right],\widetilde{\varepsilon}\right\}}{\mathbb{E}\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]},\tag{1.13}$$

with the second summand being positive because $\phi'\left[U(\alpha^0,\varepsilon)\right]$ is increasing in ε under ambiguity aversion. The behavior of an ambiguity-averse consumer can be interpreted as that of an expected utility maximizer who perceives the probability of the contract operating as intended as less likely relative to a situation without ambiguity. As a result, the optimal insurance demand for a risk- and ambiguity-averse consumer is lower when he perceives contract nonperformance risk as ambiguous than when he does not.

Before we proceed, we make some remarks on the optimality of full insurance ($\alpha^* = 1$), a topic that has received some attention on its own account. When nonperformance risk is absent, full coverage is optimal if and only if the price of insurance is actuarially fair (Mossin, 1968). Under nonperformance risk with total default, DS show that less than full coverage is optimal if the price is actuarially fair.[9] As a result, partial insurance is optimal *a fortiori* when nonperformance risk is perceived as ambiguous per Proposition 1.3.1. So Mossin's Theorem is not upheld in this case. As soon as nonperformance is partial with a sufficiently high recovery rate, full or more-than-full insurance can be optimal in the absence of ambiguity (Mahul and Wright, 2007). With the "right" combination of recovery rate, degree of ambiguity and intensity of ambiguity aversion, optimal insurance

[9]Direct computation shows that this result is readily extended to actuarially unfair premiums when the decision-maker is not imprudent ($u''' \geq 0$). How the levels of coverage with and without nonperformance risk compare to each other with a positive premium loading is, however, an unsettled question.

demand can then be at full coverage for the ambiguity-averse agent when the premium is actuarially fair.[10] Admittedly, this case is rather contrived although theoretically possible.

1.4 Some Comparative Statics Results

DS find that it is unclear under nonperformance risk whether a more risk-averse consumer will buy more coverage, even when using a strong increase in risk aversion in the sense of Ross (1981). The reason is that an increase in coverage reduces the spread between W_1 and W_2 but widens the spread between W_2 and W_3.[11] Are better results obtained when it comes to ambiguity aversion? Heterogeneity in the strength of ambiguity aversion is commonly observed in experiments, with some people being more sensitive to ambiguity than others. This is also the case for small probabilities (see Baillon and Emirmahmutoglu, 2018) as in the case of nonperformance risk, which motivates our next result.

Proposition 1.4.1: *Assume that nonperformance risk is perceived as ambiguous; then, an increase in the consumer's degree of ambiguity aversion will lead to a decrease in the optimal level of insurance coverage.*

Proposition 1.3.1 can be interpreted as a special case of Proposition 1.4.1. The reason is that the optimal insurance demand of an ambiguity-averse consumer in the absence of ambiguity coincides with that of an ambiguity-neutral consumer in the presence of am-

[10] Still, whether the "only if" part of Mossin's Theorem would hold too is yet another story because DS show that insurance demand under nonperformance risk is not necessarily lower with a positive loading compared to the actuarially fair case.

[11] Not until recently, Eeckhoudt et al. (2017) have shown that only the notion of a restricted increase in risk aversion, which is even more restrictive than comparative risk aversion in the sense of Ross (1981), yields a clear comparative static result.

biguity, at least under smooth ambiguity preferences. Using this analogy, Proposition 1.3.1 states that the optimal level of coverage for an ambiguity-averse consumer is lower than that for an ambiguity-neutral one, which also follows from Proposition 1.4.1 because the degree of ambiguity aversion of the ambiguity-neutral consumer is lower than that of the ambiguity-averse one.

To develop the underlying intuition for Proposition 1.4.1, we investigate how the consumer's unconditional marginal cost and unconditional marginal benefit of insurance are affected by an increase in his degree of ambiguity aversion. The former is given by

$$\mathbb{E}\left\{\psi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]MC(\alpha^*,\widetilde{\varepsilon})\right\}=\mathbb{E}\left\{k'\left[U(\alpha^*,\widetilde{\varepsilon})\right]\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]MC(\alpha^*,\widetilde{\varepsilon})\right\}, \tag{1.14}$$

whereas the latter is obtained as

$$\mathbb{E}\left\{\psi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]MB(\alpha^*,\widetilde{\varepsilon})\right\}=\mathbb{E}\left\{k'\left[U(\alpha^*,\widetilde{\varepsilon})\right]\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]MB(\alpha^*,\widetilde{\varepsilon})\right\}. \tag{1.15}$$

The difference to the less ambiguity-averse consumer is captured by $k'\left[U(\alpha^*,\varepsilon)\right]$, which is increasing in ε. This reinforces scenarios where ε is high so that the consumer perceives contract nonperformance as more likely relative to scenarios where ε is low and the consumer perceives contract nonperformance as less likely. Consequently, for a more ambiguity-averse consumer scenarios where the marginal cost of insurance is high and its marginal benefit is low are reinforced whereas scenarios where the marginal cost of insurance is low and its marginal benefit is high are attenuated. It is thus optimal to reduce the level of coverage.

This intuition extends to the consideration of the observationally equivalent probability distortion of the more ambiguity-averse consumer. Due to the fact that $k'\left(\phi[U(\alpha^*,\varepsilon)]\right)$

is increasing in ε, the covariance between $k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right)$ and $\widetilde{\varepsilon}$ is positive. This holds under the physical measure but also with respect to the distorted measure Q^* with Radon-Nikodym derivative $\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]/\mathbb{E}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]$. Expanding the covariance between $k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right)$ and $\widetilde{\varepsilon}$ under Q^* yields

$$\begin{aligned}
& \mathrm{Cov}^{Q^*}\left\{k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right),\widetilde{\varepsilon}\right\} \\
= \; & \mathbb{E}^{Q^*}\left\{k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right)\widetilde{\varepsilon}\right\} - \mathbb{E}^{Q^*}\left\{k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right)\right\}\cdot\mathbb{E}^{Q^*}\left\{\widetilde{\varepsilon}\right\} \\
= \; & \mathbb{E}\left\{\frac{\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]}{\mathbb{E}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]}k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right)\widetilde{\varepsilon}\right\} \\
& -\mathbb{E}\left\{\frac{\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]}{\mathbb{E}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]}k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right)\right\}\cdot\mathbb{E}\left\{\frac{\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]}{\mathbb{E}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]}\widetilde{\varepsilon}\right\} \qquad (1.16) \\
= \; & \frac{1}{\mathbb{E}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]}\mathrm{Cov}\left\{k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right)\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right],\widetilde{\varepsilon}\right\} \\
& -\left(\frac{1}{\mathbb{E}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]}\right)^2\mathbb{E}\left\{k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right)\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]\right\}\mathrm{Cov}\left\{\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right],\widetilde{\varepsilon}\right\}.
\end{aligned}$$

This is positive if and only if

$$\frac{\mathrm{Cov}\left\{k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right)\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right],\widetilde{\varepsilon}\right\}}{\mathbb{E}\left\{k'\left(\phi[U(\alpha^*,\widetilde{\varepsilon})]\right)\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]\right\}} > \frac{\mathrm{Cov}\left\{\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right],\widetilde{\varepsilon}\right\}}{\mathbb{E}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]}, \qquad (1.17)$$

which, according to the intuition for Proposition 1.3.1, states that the behavior of a more ambiguity-averse consumer can be interpreted as that of an expected utility maximizer who distorts the probability of contract nonperformance more severely compared to a less ambiguity-averse consumer. The distorted beliefs consistent with the behavior of the former attribute a higher probability to the event that the contract does not perform relative to the distorted beliefs consistent with the behavior of the latter. Therefore, it is optimal for a more ambiguity-averse consumer to reduce his optimal demand for insurance.

The question dual to the one examined in the previous proposition is that of a change in ambiguity. As a benchmark, we briefly revisit the comparative statics of risk. It is well known that a first-order stochastically dominated shift or an increase in risk in the sense of Rothschild and Stiglitz (1970) do not necessarily reduce the demand for the risky asset in the standard portfolio problem (Rothschild and Stiglitz, 1971). Exploiting the equivalence with the problem of the optimal demand for coinsurance, these changes do not necessarily raise the demand for insurance either. Sufficient conditions that have been proposed in the literature are that relative risk aversion be bounded by unity or that relative prudence be bounded by 2, respectively (Hadar and Seo, 1990; Chiu et al., 2012).

For nonperformance risk, DS show that the relationship between the consumer's optimal insurance demand and the probability of contract nonperformance is indeterminate. They only obtain a monotonically decreasing relationship when default is total, the premium is actuarially fair and preferences are restriced either to quadratic or exponential utility. Again, this appears to leave little hope for the case of ambiguity. To fix ideas, we first provide two definitions.

Definition 1.4.1:

a) The consumer's beliefs are upward biased *if his second-order beliefs undergo a first-order stochastic improvement.*

b) The consumer perceives greater ambiguity *if his second-order beliefs undergo an increase in risk in the sense of Rothschild and Stiglitz (1970).*

Under the first definition, the consumer overestimates the probability of contract nonperformance on average and his beliefs are no longer unbiased because the first-order stochastic shift relative to F implies a positive mean. The second definition has been used to study the effects of greater ambiguity on optimal self-insurance and self-protection (Snow, 2011), on the value of information (Snow, 2010), and on the incentives for genetic testing (Hoy et al., 2014). The following proposition analyzes changes in ambiguity according to Definition 1.4.1 and presents sufficient conditions for sign-definite comparative statics. They involve intensity measures of relative ambiguity aversion and relative prudence in ambiguity preferences, $-z\phi''(z)/\phi'(z)$ and $-z\phi'''(z)/\phi''(z)$, respectively.[12]

Proposition 1.4.2:

a) Upward biased beliefs reduce optimal insurance demand if relative ambiguity aversion is bounded by unity.

b) Under non-negative ambiguity prudence, greater ambiguity reduces the optimal insurance demand if relative prudence in ambiguity preferences is bounded by 2*.*

The comparative statics of changes in ambiguity are structurally isomorphic to the comparative statics of risk in the coinsurance problem. Not all ambiguity-averse consumers with upward biased beliefs or who perceive greater ambiguity reduce their optimal insurance demand. They do so if their preferences meet certain restrictions, but might react in the opposite direction otherwise. Conversely, an increase in insurance demand is not necessarily indicative of a downward shift in beliefs or a decrease in ambiguity. This is important for empirical and experimental inference because relative ambiguity aversion and relative

[12] For recent studies on prudence in ambiguity preferences, see Berger (2014, 2016) and Baillon (2017).

prudence in ambiguity preferences emerge as essential control variables.[13]

We develop some intuition for the second result.[14] A threshold value of 2 for an intensity measure of relative prudence is well known in the risk literature, for example when it comes to the effect of an increase in interest rate risk on saving behavior (see Eeckhoudt and Schlesinger, 2008; Chiu et al., 2012). The consumer experiences a substitution effect and a precautionary effect, which are also operative in our setup. The substitution effect is negative because greater ambiguity about nonperformance risk compromises the effectiveness of the insurance contract and makes it less attractive. The precautionary effect may be positive or negative. Greater ambiguity induces an increase in risk in expected utility at a given level of insurance demand. Under non-negative prudence in ambiguity preference, the consumer thus experiences an incentive to adjust the level of coverage in such a way as to increase expected utility for precautionary purposes.[15] If the consumer perceives contract nonperformance to be likely ($\varepsilon > \hat{\varepsilon}$), this is achieved by reducing the level of coverage, consistent with the substitution effect. However, if the consumer perceives contract nonperformance to be unlikely ($\varepsilon < \hat{\varepsilon}$), this is achieved by increasing the level of cover-

[13]It is easy to show that upward biased beliefs lower insurance demand for an ambiguity-neutral consumer (i.e., if $\phi'' = 0$). This does not contradict with ambiguity-neutral agents being insensitive towards ambiguity because the behavioral change originates from a higher subjective probability of nonperformance.

[14]As for the first result, a constant relative ambiguity aversion of about 0.5 has been estimated by Berger and Bosetti (2016). Thus it will not be a surprise if the majority of consumers have a relative ambiguity aversion that is less than 1.

[15]In consumption/saving models under ambiguity, the link between $\phi''' > 0$ and precautionary saving breaks down, see Berger (2014). So whereas $\phi''' > 0$ may or may not induce precautionary savings in those types of models, it always generates an incentive to adjust behavior such as to increase expected utility in our model. Similarly though, the direction of this adjustment depends on second-order beliefs, making a further restriction on preferences necessary.

age resulting in a potentially conflicting *positive* precautionary effect. The restriction on ambiguity preferences developed in Proposition 1.4.2 ensures that the precautionary effect never dominates in those cases where it is positive.

Further intuition can be developed by analyzing how greater ambiguity affects the distorted beliefs that rationalize the optimal behavior of an ambiguity-averse consumer. The corresponding probability is given by

$$\mathbb{E}^{Q^*}\left\{1-\overline{q}+\widetilde{\varepsilon}\right\}=(1-\overline{q})+\frac{\mathbb{E}\left\{\widetilde{\varepsilon}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]\right\}}{\mathbb{E}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]}. \tag{1.18}$$

If the consumer has non-negative prudence in ambiguity preferences (i.e., $\phi''' \geq 0$), it holds that $\mathbb{E}\phi'\left[U(\alpha^*,\widetilde{\kappa})\right] \geq \mathbb{E}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]$ if $\widetilde{\kappa}$ denotes an increase in risk of $\widetilde{\varepsilon}$ in the sense of Rothschild and Stiglitz (1970) so that the denominator in (1.18) increases. The reason is that the second derivative of $\phi'\left[U(\alpha^*,\varepsilon)\right]$ with respect to ε is given by $\phi'''\left[U(\alpha^*,\varepsilon)\right]U_\varepsilon(\alpha^*,\varepsilon)^2$, which is non-negative. But also the numerator in (1.18) increases with greater ambiguity. The second derivative of $\varepsilon\phi'\left[U(\alpha^*,\varepsilon)\right]$ with respect to ε is given by

$$\begin{aligned}&2\phi''\left[U(\alpha^*,\varepsilon)\right]U_\varepsilon(\alpha^*,\varepsilon)+\varepsilon\phi'''\left[U(\alpha^*,\varepsilon)\right]U_\varepsilon(\alpha^*,\varepsilon)^2\\ =&-U_\varepsilon(\alpha^*,\varepsilon)\phi''\left[U(\alpha^*,\varepsilon)\right]\left\{-U(\alpha^*,\varepsilon)\frac{\phi'''\left[U(\alpha^*,\varepsilon)\right]}{\phi''\left[U(\alpha^*,\varepsilon)\right]}\cdot\frac{\varepsilon U_\varepsilon(\alpha^*,\varepsilon)}{U(\alpha^*,\varepsilon)}-2\right\}.\end{aligned} \tag{1.19}$$

The two terms outside the curly bracket are both negative. Drawing on $U(\alpha^*,\varepsilon)=U(\alpha^*,0)+\varepsilon U_\varepsilon(\alpha^*,\varepsilon)$, which is a restatement of the fact that expected utility is linear in ε, allows us to rewrite

$$\frac{\varepsilon U_\varepsilon(\alpha^*,\varepsilon)}{U(\alpha^*,\varepsilon)}=\frac{U(\alpha^*,\varepsilon)-U(\alpha^*,0)}{U(\alpha^*,\varepsilon)}<1, \tag{1.20}$$

for any value of ε. As a result, the curly bracket in (1.19) is negative as soon as the consumer's relative prudence in ambiguity preferences is bounded by 2. Together with the previous observation, the overall sign of (1.19) is positive so that $\mathbb{E}\left\{\widetilde{\kappa}\phi'\left[U(\alpha^*,\widetilde{\kappa})\right]\right\} > \mathbb{E}\left\{\widetilde{\varepsilon}\phi'\left[U(\alpha^*,\widetilde{\varepsilon})\right]\right\}$. From Proposition 1.4.2 we know, however, that the increase in the numerator must outweigh the increase in the denominator for greater ambiguity to induce a lower level of coverage to be optimal. In such a case, the consumer's behavior with greater ambiguity can be rationalized as that of an expected utility maximizer whose distorted belief is that nonperformance is more likely compared to the case with less ambiguity.

Finally, we come to the consideration of initial wealth. In the coinsurance problem, the comparative statics of risk aversion allow to derive clear wealth effects. If the consumer's absolute risk aversion is decreasing (constant, increasing) in wealth, then the optimal level of coverage decreases (stays constant, increases) when the consumer's level of initial wealth increases, see Mossin (1968) and Schlesinger (1981). In the face of nonperformance risk, these clear comparative statics results are not recouped (see DS). For example, when absolute risk aversion decreases with wealth, insurance may or may not be an inferior good. This indeterminacy will only be exacerbated if we introduce ambiguity. We can still investigate whether ambiguity reinforces or attenuates the wealth effect.

According to the implicit function rule we differentiate the consumer's first-order condition (1.3) with respect to wealth:

$$\mathbb{E}\left\{\phi''[U(\alpha^*,\widetilde{\varepsilon})]U_W(\alpha^*,\widetilde{\varepsilon})U_\alpha(\alpha^*,\widetilde{\varepsilon})\right\} + \mathbb{E}\left\{\phi'[U(\alpha^*,\widetilde{\varepsilon})]U_{\alpha W}(\alpha^*,\widetilde{\varepsilon})\right\}. \tag{1.21}$$

The second term, which is governed by the sign of $U_{\alpha W}$, measures the effect of a change in initial wealth on the consumer's expected utility trade-off. This effect is known to be indeterminate from DS. The first term measures how a change in initial wealth affects the consumer's behavioral response to ambiguity and, more specifically, how an increase in wealth induces him to adjust his level of coverage in order to react to ambiguity. We will focus on this effect to answer the question whether ambiguity has a positive or negative effect on the comparative statics of wealth. Our result involves the consumer's index of absolute ambiguity aversion, denoted by $A_\phi(z) = -\phi''(z)/\phi'(z)$, and his relative ambiguity aversion as defined previously.

Proposition 1.4.3: *Consider a consumer with non-increasing absolute ambiguity aversion and relative ambiguity aversion bounded by unity. Then, ambiguity reinforces the wealth effect.*

The last proposition shows that insurance is less likely to be an inferior good when contract nonperformance risk is perceived as ambiguous, the consumer's absolute ambiguity aversion is non-increasing and his relative ambiguity aversion is bounded by unity. The underlying intuition is simple: If the consumer's initial wealth increases, his expected utility is higher for any level of insurance coverage. With non-increasing absolute ambiguity aversion the consumer is less ambiguity-averse when expected utility is high than when it is low, and Proposition 1.4.1 informs us that it is optimal for the consumer to increase his level of coverage.

1.5 Extensions and Robustness

1.5.1 Weighted Maxmin Expected Utility

The model developed by Klibanoff et al. (2005) has the advantage that it disentangles ambiguity and ambiguity attitude. Nevertheless, it is by far not the only model of decision-making under ambiguity. In this subsection, we use the model developed by Ghirardato et al. (2004), in which the consumer's objective function is formed as a weighted average of the worst case and the best case. If all weight is attached to the worst case, maxmin expected utility emerges as a special case (Gilboa and Schmeidler, 1989).[16]

Let $\beta \in [0, 1]$ denote the consumer's degree of pessimism with respect to the perceived ambiguity of contract nonperformance $\widetilde{\varepsilon}$. Recall that the support of the consumer's beliefs is given by $[\underline{\varepsilon}, \overline{\varepsilon}]$ so that his objective function becomes

$$V(\alpha) = \beta \min_{\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]} U(\alpha, \varepsilon) + (1 - \beta) \max_{\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]} U(\alpha, \varepsilon) = \beta U(\alpha, \overline{\varepsilon}) + (1 - \beta) U(\alpha, \underline{\varepsilon}) = U(\alpha, \theta). \tag{1.22}$$

$\theta = \beta \overline{\varepsilon} + (1 - \beta)\underline{\varepsilon}$ combines the effects of ambiguity and ambiguity aversion, and it is immediate that the interpretation of the consumer's optimal behavior based on distorted beliefs is still valid. V is concave, and for ambiguity to have no effect on the consumer, it must be that $\theta = 0$ (ambiguity neutrality). Then, the optimal level of insurance coverage is given by α^0 as in equation (1.8). Ambiguity aversion obtains for $\theta > 0$ and renders decisions that are behaviorally equivalent to those made by an expected utility maximizer who perceives the probability of nonperformance as elevated. It follows from our previous results that $\alpha^* < \alpha^0$ in this case and Proposition 1.3.1 remains valid. Furthermore, an increase in β leads to an increase in θ, and insurance demand decreases, which is the weighted maxmin version of Proposition 1.4.1. When it comes to the comparative statics of ambiguity, the weighted maxmin model does not produce clear predictions. The reason is that a change in risk can leave the support of the distribution unchanged, expand it on

[16]Maxmin utility is also a special case of the Klibanoff et al. (2005) model for an exponential ambiguity function as the degree of ambiguity aversion tends to infinity. The weighted maxmin model is, however, not directly implied by our previous analysis, which warrants its separate investigation.

one end, or expand it on both ends. Adapting the terminology developed by Meyer and Ormiston (1985), a change in risk can be weak, "semi-strong", or strong depending on how it affects the support of the associated distribution. Conditional on β, the effect of such stochastic changes on insurance demand is nil in the first case, positive (negative) in the second case if it is the lower (upper) end of the support which is expanded, or indeterminate in the third case.[17] Wealth effects are those derived in DS, but under a distorted probability distribution. So except for the comparative statics of ambiguity, all our results continue to hold in the weighted maxmin expected utility model.

1.5.2 A Probability Weighting Model

Snow and Warren (2005) and Snow (2011) use a probability weighting model to analyze the effects of ambiguity on optimal decision making. Probability weighting models include rank dependent expected utility (Quiggin, 1982), the decision weighting model of Kahn and Sarin (1988), and cumulative prospect theory (Tversky and Kahneman, 1992). Let ξ denote the consumer's probability weighting function, which is an increasing function of cumulative probabilities with $\xi(0) = 0$ and $\xi(1) = 1$. We rule out overinsurance (i.e., $\alpha \leq 1$) so that $W_3 < W_2 \leq W_1$ with the last inequality strict for partial insurance (see Footnote 6). Conditional on $\widetilde{\varepsilon} = \varepsilon$, the consumer's welfare is

$$\xi(p(1 - \overline{q} + \varepsilon))u(W_3) + [\xi(p) - \xi(p(1 - \overline{q} + \varepsilon))]\, u(W_2) + [\xi(1) - \xi(p)]\, u(W_1), \quad (1.23)$$

[17]Notice that the interpretation of β is not independent of the support of the distribution of beliefs. See Huang and Tzeng (2017) for a recent application to portfolio choice.

so that his expected ex-ante welfare is given by the following utility objective:

$$V(\alpha) = u(W_1) + \xi(p)\left[u(W_2) - u(W_1)\right] + \mathbb{E}\xi\left(p(1 - \overline{q} + \widetilde{\varepsilon})\right)\left[u(W_3) - u(W_2)\right]. \quad (1.24)$$

Ambiguity neutrality is obtained if ambiguity does not affect the consumer's objective function. For this to hold for any level of ambiguity, ξ must be linear. Now consider an ambiguity-averse consumer; for any level of ambiguity to reduce his objective function, ξ must be convex so that $\mathbb{E}\xi\left(p(1 - \overline{q} + \widetilde{\varepsilon})\right) > \xi(p(1 - \overline{q}))$.[18] Then, it is straightforward to show that, if the consumer is ambiguity-averse, ambiguous nonperformance risk reduces the optimal insurance demand relative to the case without ambiguity. Furthermore, an increase in the consumer's degree of ambiguity aversion, operationalized as a convex transformation of ξ, lowers the optimal demand for insurance. The changes in ambiguity defined in Definition 1.4.1 always reduce the demand for insurance without any additional assumptions, and wealth effects correspond to those derived in DS when evaluated at a distorted probability distribution. So effectively all our results continue to hold under a probability weighting model of ambiguity aversion.

1.5.3 Some Remarks on the Supply Side of the Market

Our analysis assumes that the consumer faces an increasing and non-concave premium schedule, that is independent of his perception of ambiguity. This is quite general and includes the obvious case of a risk- and ambiguity-neutral insurer on a competitive

[18] Any twice continuously differentiable weighting function that is increasing, convex and satisfies $\xi(0) = 0$ and $\xi(1) = 1$ necessarily implies global underweighting of probabilities, that is $\xi(p) \leq p$ for all p. This disqualifies the descriptively attractive case of inverse S-shaped weighting functions. It also implies that the behavior of an ambiguity-averse agent in the absence of ambiguity will, in general, differ from the behavior of an ambiguity-neutral agent in the presence of ambiguity.

market. In such a situation, the premium would be based on the actuarial value of the policy such that $P(\alpha) = \alpha pqmL + K$ with $m \geq 1$ being a loading factor and $K \geq 0$ a fixed cost.

A natural question is whether such a premium schedule is still obtained if the insurer perceives nonperformance risk as ambiguous and exhibits aversion to this ambiguity (see Cabantous, 2007; Cabantous et al., 2011). We investigate this question based on the three criteria of decision-making under ambiguity that were presented thus far. The insurer is assumed to maximize profit as determined by the premium of the policy (P), the indemnity payment in case of a loss (αL), the probability of loss (p), the probability that the contract performs ($\overline{q} - \widetilde{\tau}$), the administrative costs as measured by the loading factor (m) and the fixed cost (K). For a risk-neutral insurer, the expected profit is given by

$$\widetilde{\Pi} = \Pi(\widetilde{\tau}) = P - \alpha p(\overline{q} - \widetilde{\tau})mL - K, \tag{1.25}$$

where we allow for the insurer's perception of nonperformance risk to deviate from the consumer's, as represented by $\widetilde{\tau}$ (instead of $\widetilde{\varepsilon}$). If the insurer is ambiguity-neutral and the market is perfectly competitive, the premium reduces to the expression at the end of the previous paragraph as long as beliefs are unbiased. To incorporate ambiguity aversion, let χ denote the insurer's ambiguity function with $\chi' > 0$ and $\chi'' < 0$. If the insurer has other assets in place, denoted by A, and the market is perfectly competitive, the premium is implicitly defined by

$$\mathbb{E}\chi(A + P^* - \alpha p(\overline{q} - \widetilde{\tau})mL - K) = \chi(A), \tag{1.26}$$

so that the insurer is just indifferent between offering and not offering the policy. In other words, the insurer's participation constraint must be binding under perfect competition. Then, the premium can be written as $P^* = \alpha p \overline{q} mL + K + \rho$, where $\rho > 0$ is a strictly positive ambiguity premium: An ambiguity-averse insurer charges a higher premium for the same policy than an ambiguity-neutral insurer would do.

To determine slope and curvature of the premium schedule, we apply the implicit function rule to (1.26). This yields

$$\frac{\mathrm{d}P^*}{\mathrm{d}\alpha} = pmL \left\{ \overline{q} - \frac{\mathbb{E}\widetilde{\tau}\chi'}{\mathbb{E}\chi'} \right\} > 0, \tag{1.27}$$

where the argument of χ' is suppressed to simplify notation. The premium schedule is increasing. Its second derivative is given by

$$\frac{\mathrm{d}^2 P^*}{\mathrm{d}\alpha^2} = \left(\frac{pmL}{\mathbb{E}\chi'}\right)^2 \left\{ -\mathbb{E}\widetilde{\tau}^2\chi''\mathbb{E}\chi' + 2\mathbb{E}\widetilde{\tau}\chi''\mathbb{E}\widetilde{\tau}\chi' - \frac{\mathbb{E}\chi''\left(\mathbb{E}\widetilde{\tau}\chi'\right)^2}{\mathbb{E}\chi'} \right\}. \tag{1.28}$$

The first and third term in the curly bracket are positive. The middle term is non-negative if the insurer is not prudent in ambiguity preferences ($\chi''' \leq 0$), in which case all our results go through. Another way to analyze the curly bracket is to rewrite it as follows:

$$\mathbb{E}\widetilde{\tau}\chi''\mathbb{E}\widetilde{\tau}\chi' - \mathbb{E}\widetilde{\tau}^2\chi''\mathbb{E}\chi' + \frac{\mathbb{E}\widetilde{\tau}\chi'}{\mathbb{E}\chi'}\left(\mathbb{E}\chi'\mathbb{E}\widetilde{\tau}\chi'' - \mathbb{E}\chi''\mathbb{E}\widetilde{\tau}\chi'\right). \tag{1.29}$$

Similar techniques as in Appendices 1 and 2 of Peter et al. (2017) show that the sign of the sum of the first and second term is governed by the slope of relative ambiguity aversion whereas the sign of the round bracket is governed by the slope of absolute ambiguity aversion.[19] If absolute ambiguity aversion is non-increasing and relative ambiguity aversion is non-decreasing, the overall sign is negative resulting in a concave premium schedule.

[19] $\mathbb{E}\widetilde{\tau}\chi' < 0$ follows from ambiguity aversion.

As such, a necessary condition for a non-concave premium schedule is that either absolute ambiguity aversion is increasing, in which case relative ambiguity aversion is too, or that relative ambiguity aversion is decreasing, in which case absolute ambiguity aversion is too.

In the weighted maxmin model, we let $\gamma \in [0,1]$ denote the insurer's degree of pessimism and $[\underline{\tau}, \overline{\tau}]$ the support of the insurer's beliefs. The premium is then determined from

$$\gamma \min_{\tau \in [\underline{\tau}, \overline{\tau}]} (A + \Pi(\tau)) + (1-\gamma) \max_{\tau \in [\underline{\tau}, \overline{\tau}]} (A + \Pi(\tau)) = A + \Pi(\eta) = A, \tag{1.30}$$

with $\eta = \gamma \underline{\tau} + (1-\gamma)\overline{\tau}$, so that $P^* = \alpha p(\overline{q} - \eta)mL + K$. From the insurer's perspective, $\eta < 0$ represents ambiguity aversion because scenarios with a higher probability of having to pay the indemnity reduce expected profits. The premium is higher than the premium charged by an ambiguity-neutral insurer but it is still linearly increasing in the level of coverage so that all our results go through.[20]

Finally, we can apply the probability weighting model to the insurer's pricing decision. For the insurer there are only two relevant states of the world, the one where it pays the indemnity and the one where it does not. The insurer's profit is lower in the first one, which occurs with probability $p(\overline{q} - \widetilde{\tau})$, than in the second one, which occurs with probability $(1 + p) + p(1 - \overline{q} + \widetilde{\tau})$. If ζ denotes the insurer's probability weighting function, the

[20]Depending on their informational endowment, consumers might be able to back out the insurer's perceived level of ambiguity from the observed market prices. Then, to be consistent the consumer's perception of ambiguity should be such that it is compatible with the one implied by the insurer's pricing.

premium is determined by requiring that

$$\begin{aligned}&\mathbb{E}\zeta\left(p(\overline{q}-\widetilde{\tau})\right)(A+P-\alpha mL-K)+\left(1-\mathbb{E}\zeta\left(p(\overline{q}-\widetilde{\tau})\right)\right)(A+P-K)\\ =\;& A+P-\alpha mL\cdot\mathbb{E}\zeta\left(p(\overline{q}-\widetilde{\tau})\right)-K=A,\end{aligned} \tag{1.31}$$

so that $P^* = \alpha mL \cdot \mathbb{E}\zeta\left(p(\overline{q}-\widetilde{\tau})\right) + K$. For ambiguity to reduce the insurer's expected profit, ζ needs to be convex. Then, ambiguity raises the premium charged by an ambiguity-averse insurer. The premium is, however, still a linear function of the level of coverage and our results on the effects of the consumer's perceived ambiguity about nonperformance risk continue to hold. The same qualification as in Footnote 20 applies.

Overall, ambiguity aversion of the insurer can result in concavity of the premium schedule under the smooth model but never under the other two models. Notice that non-concavity of the premium schedule was only a sufficient condition for the validity of the first-order approach but not necessary. Loosely speaking, as long as the premium schedule is not too concave, all our results are upheld.

1.5.4 Ambiguity Associated with the Risk of Loss

To isolate the effects of an ambiguous perception of nonperformance risk, we assumed the probability of loss to be known by the consumer. A natural question is to wonder to what extent our results carry over to situations in which both the risk of loss and the risk of nonperformance are perceived as ambiguous. When the probability of contract nonperformance is known, ambiguity associated with the risk of loss no longer has clear effects on insurance demand, contrary to the findings in Alary et al. (2013). Intuitively, the behavior of an ambiguity-averse consumer under ambiguity is observationally equivalent to that of an expected utility maximizer who is more pessimistic about the probability of loss. This increases the marginal benefit of insurance because the loss state in which the insurance contract performs is perceived as more likely. There is a negative effect on the marginal cost of insurance because the no-loss state becomes less likely but there are also two positive effects on the marginal cost because both the state in which the loss happens and insurance performs and in which the loss happens and insurance does not perform are perceived as more likely. The sum of the two positive effects preponderates the negative effect because marginal utility in the latter two cases is higher than in the first one so that overall the marginal cost of insurance increases. As a consequence, the net effect of an ambiguous perception of the risk of loss on insurance demand is indeterminate when contract nonperformance is possible.

However, we are still able to say something about the effects of an ambiguous perception of nonperformance risk. Conditional on the risk of loss being perceived as ambiguous, ambiguity associated with the risk of nonperformance lowers the demand for insurance when both sources of ambiguity are independent. Also, the sufficient conditions for upward biased beliefs and for greater ambiguity to lower the optimal demand for insurance and for a positive partial wealth effect remain unaltered. This shows that Propositions 1.3.1, 1.4.2 and 1.4.3 are robust to the consideration of ambiguity associated with the risk of loss as long as both sources of ambiguity are unrelated. Only Proposition 1.4.1 does no longer hold because an increase in ambiguity aversion results in a tension between a negative effect of ambiguity associated with nonperformance risk and an indeterminate and potentially positive effect of ambiguity associated with the risk of loss on insurance demand. This resembles the same tension that impedes signing the comparative statics of risk aversion in the DS model.

1.6 Conclusion

In this paper, we analyze optimal insurance demand if consumers face contract nonperformance risk. Insurance offers protection against financial risks, but sources of nonperformance abound including insurer default, contested claims, procedural delays, contractual uncertainty and probationary periods. Whereas existing literature assumes the level of nonperformance risk to be known by the consumer, we study the case of ambiguity and show how this ambiguity affects optimal demand. Despite the fact that most of the comparative statics of nonperformance risk are indeterminate (see DS), we are able to show that the comparative statics of ambiguity are much better behaved. Using the smooth model of ambiguity aversion, we find that ambiguity lowers the optimal demand for insurance. Furthermore, greater ambiguity aversion lowers the demand for insurance and so do upward biased beliefs and greater ambiguity if the consumer's relative ambiguity is bounded by unity or his relative prudence in ambiguity preferences is bounded by 2, respectively. Finally, we analyze a partial wealth effect and determine sufficient conditions under which insurance is less likely to be an inferior good when contract nonperformance risk is perceived as ambiguous. Our main results are robust to the specification of ambiguity preferences, which we scrutinize by investigating a weighted maxmin expected utility model and a probability weighting model. Furthermore, depending on the implementation of ambiguity preferences, our findings continue to hold when the insurer is ambiguity-averse, or when the risk of loss is perceived as ambiguous.

The reason why ambiguity associated with contract nonperformance risk has clear comparative statics whereas nonperformance risk itself does not is related to the mechanism through which both operate. In DS, nonperformance risk is reflected in the odds of the contract performing but also reduces the premium of the insurance contract. The latter is a wealth effect which undermines the comparative statics and renders most of the results indeterminate. In our paper, the focus is on the consumer's perception of contract nonperformance risk so that the insurer's pricing rule is taken as given. As a consequence our model does not contain the conflicting wealth effect that the original DS model does. This assumption appears natural in our context since most insurers are unlikely to be aware of their consumers' perception of nonperformance risk.

Our results show that ambiguity associated with contract nonperformance risk has clear demand effects, which are not obtained by a mere focus on nonperformance risk and risk aversion. The fact that optimal demand is lower allows us to identify ambiguous nonperformance risk is a potential cause of uninsurability. Some have argued that this might be particularly relevant in less developed insurance markets where trust in financial institutions can be limited and insurers are less stringently regulated (e.g., Biener et al., 2017). The findings in this paper suggest that informational policies that eliminate ambiguity associated with contract nonperformance risk increase the ex-ante welfare of ambiguity-averse consumers and stimulate demand on the market. However, for the more likely case of a partial reduction in ambiguity, ex-ante welfare still increases whereas demand might rise or fall, depending on the consumer's relative prudence in ambiguity preferences. This result is particularly important because it shows that demand reactions are only partially informative about underlying welfare changes, see also Harrison and Ng (2018) on this point. Furthermore, if one believes that ambiguity aversion is a psychological tendency that inhibits "good" decision-making and that expected utility represents the appropriate normative framework, our results show that ambiguity-averse consumers purchase less insurance than they should as soon as they perceive nonperformance risk as ambiguous. In such a situation policy interventions that reduce the consumer's sensitivity to ambiguity enhance welfare as do policy interventions that eliminate ambiguity altogether. This could be accomplished by the provision of information to consumers or by solvency regulations on the supply side of insurance to guarantee sufficient capital provisioning.

There are several avenues for further research. Our model has clear predictions and experimental or empirical follow-ups appear promising. Biener et al. (2017) conducted the first artefactual field experiment with subjects in the Philippines and find that an ambiguous perception of contract nonperformance risk reduces insurance take-up for ambiguity-averse individuals. This is consistent with our Proposition 1.3.1 but obviously insurance take-up is a crude measure of insurance demand, and a more detailed analysis in terms of the degree of ambiguity aversion, the level of ambiguity or baseline income is lacking to date. Empirically, it is not trivial to come up with convincing proxies for an ambiguous perception of nonperformance risk but contract length or the degree of claims verifiability based on data availability might be interesting starting points.[21]

On the theory side, a natural question is whether similar demand effects arise when the reliability of other risk management instruments like self-insurance or self-protection

[21] We would like to thank Alexander Mürmann for these suggestions.

is perceived ambiguous. Whereas self-insurance and (market) insurance appear to behave similarly in most contexts, suggesting the generalizability of our results to self-insurance, the contrary is true for self-protection and insurance (see Ehrlich and Becker, 1972), leaving this an interesting topic for future research. Furthermore, we formalized the consumer's loss exposure with a binary distribution so that insurance demand is collapsed into one single variable, the level of coverage. More generally, researchers have used less restrictive assumptions on the loss distribution to investigate optimal insurance design (see Gollier, 2013, for a survey), and it would be interesting to see how an ambiguous perception of nonperformance risk affects the optimality of deductible insurance in such a context. Our paper assumes ambiguity as a primitive of the consumer's decision-making environment. An important source of nonperformance risk is correlation between the risks of different insurees. In such a situation, nonperformance risk and the perception thereof are endogenously determined by insurance demand, which might generate further interesting implications. Lastly, applications of nonperformance risk in asymmetric information environments with and without ambiguity are largely unexplored.

CHAPTER 2
SHOULD CORPORATE PENSION FUNDS INVEST IN RISKY ASSETS?

2.1 Introduction

Defined-benefit corporate pension plans (hereafter simply referred to as pensions or DB plans) are legal entities set up by companies to provide a stable stream of incomes to employees upon their retirement. Despite a long history of pension evolution, how pension plans should invest remains an unsettled issue. At one extreme, many corporate pensions consider themselves as patient long-term investors, and act as major holders of illiquid, risky assets such as real estate, hedge funds, and private equities. At the other extreme, some pensions plans have subscribed to a complete "de-risking" strategy, holding only safe fixed-income securities or annuities. Collectively, U.S. corporate pensions' allocation to stocks and alternative risky assets hovers above 50% in recent years. Aggressive pension risk-taking, coupled with prevalent pension underfunding, has often been the subject of concern in news media and by policymakers.

Researchers have identified at least two reasons that defined-benefit pension plans should avoid risky securities and invest safely. The first, as pointed out by Black (1980) and Tepper (1981), is a corporate tax advantage for pensions to invest in fixed income securities. The second, and an even stronger reason, can be understood in the context of the corporate risk management model of Froot and Stein (1998). Any risky corporate investment without a positive alpha – including pension investment – does not create value for shareholders; meanwhile, risky investment may reduce firm value by forcing firms to raise costly financing. Therefore, to maximize shareholder value, pensions should only invest in riskfree assets that match the horizon of pension obligations.[1] From this perspective, the risky asset allocation by corporate pensions at the aggregate level poses a puzzle.[2]

Several studies (e.g., Sharpe, 1976; Sharpe and Treynor, 1977) link the risky investments by corporate pensions to a moral hazard problem, in a way similar to the risk-shifting problem for firms with high financial leverage. But moral hazard is unlikely the sole driver of corporate pensions' risky asset allocation decisions. Studies such as Bodie, Light, Morck and Taggart Jr (1985), Rauh (2008), and An, Huang and Zhang (2013) find that pensions sponsored by firms with higher bankruptcy risk – thus stronger risk-shifting incentives – take on lower investment risk. In addition, Lucas and Zeldes (2006) and Sundaresan and Zapatero (1997) point out that pensions may invest in stocks to hedge against the future growth of pension obligations. However, Lucas and Zeldes (2006) note that equity allocation by pension plans with low future wage growth remains quite high and cannot be explained by the hedging demand. It suffices to say that the search is still on for a better understanding of the risky investment policies adopted by most corporate pensions.

In this study, we take a stakeholder approach to model key pension decisions by firms, including the investment decision as well as the choices of pension benefit level and

---

[1] The intuitive version of this argument is suggested in the early pension literature; e.g., Bodie (1988). The risk-management based argument serves as a rationale for the pension de-risking strategy proposed by practitioners (e.g., Cooper and Bianco, 2003).

[2] Possibly, some firms invest pension money in risky assets with a hope to generate positive alphas. But evidence provided by a long stream of academic studies, from Lakonishok et al. (1992) to Busse et al. (2010), suggests that actively equity portfolios managed on behalf of pension funds on average fail to generate positive alphas. Further, the alpha-seeking motivation does not explain why pensions invest substantially in index portfolios. According to French (2008), during the period of 2000-2006, about 30% of DB plans' equity investments are passively managed.

pension funding. In our model, firms maximize shareholder value, but are additionally subject to the participation of wealth-constrained employees. As a consequence, firms must balance the risk management concern of shareholders with employees' preference for systematic risk exposure. The model predicts substantial risky assets in optimal pension portfolios, and thus offers a perspective for understanding the observed pension risk-taking behavior. For a reasonable set of parameters, the model-suggested optimal level of pension investment risk and its relation with the bankruptcy risk and pension funding ratio are consistent with those reported by empirical studies. We also show that the typical defined benefit pensions offer inefficient risk sharing between shareholders and employees, which exacerbates pension risk taking.

A key element of our model is the difference between a firm's shareholders and its employees in their capacity for bearing pension investment risk. Notably, the investment risk ensuing from a diversified pension portfolio is systematic. And as it turns out, employees may have a stronger capacity or appetite than shareholders for such systematic risk. To understand this, think about a standard assumption in the existing literature (e.g., Froot and Stein, 1998) that shareholders hold an optimally diversified portfolio to maximize their utility. By already holding an optimal portfolio, shareholders are indifferent to a small change in the level of systematic risk brought about by an incremental investment, as long as the investment is fairly valued (i.e., having a zero alpha). Employees, on the other hand, have a substantial part of their wealth tied up in safe wages. Because of this wealth constraint, employees may be under-exposed to systematic risk and find the market risk premium very attractive.[3] As a consequence, they may desire systematic risk exposure in their pension payoffs, even when such exposure does not deliver a positive alpha.[4] In sum, the different capacities for systematic risk are not because shareholders and employees have different utility functions, but rather because of the different levels of systematic risk they are already exposed to.

Another important element of our model is that despite the fixed pension benefits promised by firms, employees are at least partially exposed to pension investment risk. Employees' share of pension investment risk intuitively comes from two channels. The first is the bankruptcy channel – when a firm goes bankrupt, employees' pension payoff depends on pension assets value relative to pension liabilities, and the payoff becomes risky when the pension is underfunded. The second is a pension surplus sharing channel – when the value of pension assets exceeds the value of promised benefits, employees are entitled to at least a fraction of the excess. This pension surplus sharing mechanism, in short, can be understood as a direct outcome of the current tax code. Pension assets are held in a separate legal entity from the sponsoring firm, and there is a 50% punitive excise tax when a firm converts pension surplus back into firm assets. However, the tax rate is reduced to 20% if employees receive at least 20% of the surplus, e.g., in the form of increased pension

[3]Classic portfolio theory dictates that the risk of a marginal investment perceived by an individual hinges on the covariance of the investment performance with the person's overall wealth. A constrained employee's overall wealth has a relatively low correlation with the stock market's day-to-day fluctuation, compared against a typically wealthier shareholder of the firm. For this reason, an investment in market index is perceived as less risky in the eyes of a typical employee than in the eyes a typical shareholder.

[4]This is certainly different from the case of bearing firm-specific risk – unconstrained shareholders can more effectively diversify away firm-specific risk than wealth constrained employees do.

benefits. This creates an incentive for firms to share pension surplus with employees.[5] Note that the bankruptcy channel mainly exposes employees to the downside of pension investment risk, while the surplus sharing channel mainly exposes employees to the upside of pension investment risk.

The difference in their risk-bearing capacities and the sharing of pension investment risk between shareholders and employees jointly drive pension investment decisions. Suppose a pension chooses between a fairly-valued stock index fund and a riskfree asset. Although shareholders are indifferent to a small systematic risk increase in terms of their utility, an increase in risk reduces their value by increasing the firm's expected financing cost. Therefore, if shareholders are to make pension decisions solely on their own, as predicted in the standard corporate risk management setting, they would strictly prefer the riskfree investment. However, pension investment risk is attractive to employees. Under reasonable assumptions for the financing cost function and employees' utility function, there is an optimal level of pension investment risk, born by both shareholders and employees. Further, we show that the decisions on the level of pension benefits and on pension contribution are also affected by pension risk sharing. For example, for firms with low probabilities of bankruptcy, employee's exposure to pension investment risk is mainly

[5]The Employee Retirement Income Security Act (ERISA) of 1974 requires that pensions assets are managed to the exclusive benefit of beneficiaries, which can be construed as that beneficiaries have claims on pension surplus. However, the laws in this aspect are incomplete, and firms' effective control rights on pensions matter. Firms have various means to reclaim part of pension surplus, from reducing pension contributions, terminating overfunded pension plans, merging overfunded pensions with underfunded ones, to diverting pension assets to cover operating costs and restructuring costs. Meanwhile, various restrictions are in place to prevent firms from recapturing all the pension surplus and encourage firms to share pension surplus with employees – for example, the punitive excise tax for pension surplus reversion. Section II provides further details on pension surplus sharing. The fact that employees enjoy a fraction of pension surplus has been noted in existing literature; see, e.g., Miller and Scholes (1981) and Bulow and Scholes (1983).

through the pension surplus sharing channel. For such firms, it is optimal to maintain a high funding ratio (i.e., high pension contribution relative to promised pension benefits) and at the same time invest aggressively, to increase the chance of pension surplus. By contrast, for firms with high bankruptcy probabilities, the model predicts a low funding ratio.

We show that with reasonable parameter values, the model outcome matches several observed patterns on pension investments. For example, in a baseline calibration analysis, we consider a 30-year retirement horizon and an annual bankruptcy probability of 0.5% for the firm (typical for BBB-rated firms), with employees' share of pension surplus at a modest 20% and with other parameters such as market risk premium, market volatility, and financing costs either calibrated to historical data or taken from estimates of existing studies. The optimal weight on the risky assets in this case is around 55%, matching well with the observed level of risky allocation by corporate pensions. By varying model parameters, we also find that the optimal portfolio weight on risky assets tends to decrease with the bankruptcy probability, and increase with the pension funding ratio. These two patterns are consistent with empirical findings reported in existing studies (e.g., Bodie et al., 1985; Rauh, 2008; An et al., 2013).

Across various parameter choices, the relation between pension investment risk and employees' share of pension risk – through either the bankruptcy channel or the surplus sharing channel – tends to be substitutive in our model. That is, when employees' share of pension risk is lower, the pension invests more aggressively.[6] This relation, although

[6] An exception is that when employees' share of pension surplus is very low (close to zero), a small increase in the surplus sharing leads to an increase in pension investment risk.

surprising at first, can be intuitively understood as an effect to maintain a desirable level of systematic risk exposure in employees' pension payoffs. To further understand this relation, we explore a hypothetical pension arrangement where employees bear all the pension investment risk, which is similar to a defined-contribution plan (referred to as a "variable benefit" plan). We find that the optimal level of investment risk is lower under the "variable benefit" plan than under the typical defined-benefit arrangement. That is, investment decisions by defined benefit pension plans are more aggressive than what employees would choose for themselves in defined contribution plans. This gives rise to an interesting policy implication – if pensions' risk-taking behavior is a cause for concern, then designing a better pension risk-sharing mechanism for employees could be part of the redress.

Ultimately, our analysis highlights that risk sharing between shareholders and employees in typical defined benefit plans is inefficient. Given employees' preference for systematic risk and shareholders' concern for risk management, letting employees bear all the pension investment risk could be a better arrangement. To showcase the potential welfare improvement, we compare firms' total pension funding cost under the typical defined benefit plan with that under the "variable benefit" plan. For most parameter choices, firms' pension funding cost is substantially lower under the latter arrangement. It is worthwhile noting that in the recent decades, driven by the desire to reduce pension funding cost, many firms have switched away from defined benefit plans to defined contribution plans. DB plans' inefficiency in investment risk sharing might be one of the reasons for this secular trend.

Overall, our paper provides a new perspective to understand the risky asset allocation policies pursued by corporate pensions. The insight of our model is that pension risk taking could be driven by employees' preference for systematic risk exposure. In this regard, our paper is related to Love, Smith and Wilcox (2011), yet with different objectives and implications. The main purpose of their paper is to study the effect of PBGC insurance on pension risk taking. In their model, employees do not face wealth constraints and they price pension investment risk the same way as shareholders. As a result, their model has a strong prediction that pensions should invest riskfree as long as PBGC insurance is fairly priced.

The rest of the paper is organized as follows. Section II provides a discussion on the evolution of defined benefit pension plans and pension investments, and several legal aspects relevant for understanding pension investment risk sharing. In Section III, we introduce a one-period model of pension investments and provide analytical results under certain model specifications. Section IV reports the numeric solutions of the one-period model under more general model specifications, and examine various extensions of the model. Section V examines a dynamic model of pension investments and pension funding. Finally, Section VI provides concluding remarks.

2.2 Background on Corporate Pensions and Pension Investments

2.2.1 Evolution of Corporate Pensions and Pension Investments

The origin of U.S. corporate-sponsored pensions could be traced to the 1880s, when companies in the booming railroad industry used pension benefits to recruit workers. In 1875, American Express – a railway company at the time – established the first pension plan with defined benefit features. By 1929, there were about four hundred corporate pension plans in operation, sponsored by many large corporations of the time (Munnell, 1982). Corporate pensions took a hit during the Great Depression, but recovered afterwards and grew rapidly after World War II, covering 25%, 41%, 45%, 46%, and 43% of all private-sector workers in 1950, 1960, 1970, 1980, and 1990, respectively (McDonnell, 1998). In the recent decades, however, with the rise of defined contribution plans and individual retirement accounts (since late 1970s and 1980s), the importance of defined benefit pensions in overall retirement savings has declined. According to recent statistics, in 2014, DB plan assets stand at \$3.96 trillion, compared with \$5.32 trillion for defined contribution plans and \$6.23 trillion for individual retirement accounts (IRAs).[7] The waning popularity of corporate DB plans relative to DC plans and IRAs has been well noted and discussed in existing studies; see, e.g., Munnell and Soto (2007), and Rauh and Stefanescu (2009).

Prior to the stock market boom of the 1950s, corporate pensions invested only in safe assets such as bank deposits, government bonds and corporate bonds. In 1950, a DB plan of General Motors became the first to invest in the stock market (McDonnell, 1998). Over time, pensions have shifted toward substantial risky investments. By mid-1960s, the aggregate corporate pension allocation to stocks exceeded allocation to fixed income assets.[8] According to various recent statistics and surveys, corporate pensions allocate

---

[7]Statistics for DB and DC plans are from Private Pension Plan Bulletin, Abstract of 2014 Form 5500 Annual Reports, by Employee Benefits Security Administration, Depart of Labor. Available at: https://www.dol.gov/agencies/ebsa/researchers/statistics/retirement-bulletins/private-pension-plan. Statistics for IRAs are from The IRA Investor Profile: Traditional IRA Investor's Activity 2007-2015. Available at: https://www.ici.org/pdf/rpt_17_ira_traditional.pdf.

[8]Based on statistics provided by Federal Reserve Board Flow of Funds data (Z.1 Statistical Release).

between 50% to 60% of the investments to risky assets such as stocks, hedge funds, private equities, and real estates (e.g., Stockton, 2012; Andonov et al., 2012; Panis and Brien, 2015; and Watson, 2016). For example, based on corporate filings (Department of Labor Form 5500) in 2013, Panis and Brien (2015) find that on average 50.6% of pension assets are invested in stocks, with another 12% in real estate and alternative assets. They also note that pension allocation to risky assets is trending down in recent years.

Corporate pensions' wage into the stock market in the 1950s was motivated by the market boom during that period. In the 1980s, corporate pensions were among the early investors (together with public pensions and insurers) in junk bonds. Due to their long-term nature of pension obligations, DB plans are considered long-term investors who can afford to invest in illiquid and risky assets (e.g., Campbell and Viceira, 2005). Starting from the 1990s, corporate pensions also become pioneering investors, and remain a major force today, in alternative assets such as hedge funds and private equities. Meanwhile, their relative importance in the public equity market sees a peak in the mid 1990s and has since been on a decline (French, 2008).

The notion of liability-driven investing (LDI) by pensions was developed in the 1980s (Leibowitz, 1986; Ang, 2014). LDI takes into account pension liabilities when making investment decisions. An extreme version of this approach is to completely de-risk pension assets, i.e., investing in annuities and avoiding any form of market risk.[9] A more flexible version of LDI is to set the pension investment objective to be a concave function of pension surplus ratio (i.e., the ratio of pension assets to pension liabilities), thus

[9]Pension de-risking may also involve removal of uncertainty in future pension obligations.

introducing a hedging component in the optimal portfolio against the interest rate risk of pension liabilities. LDI strategies have gained traction among pension managers in recent years (e.g., Cooper and Bianco, 2003; Leibowitz and Ilmanen, 2016).

2.2.2 Ownership and Control of Pension Plans: Legal Aspects

Corporate DB pension plans are typically set up as trusts, with trust beneficiaries being qualified (current and retired) employees. The trustees are in charge of pension administration, supposedly in the interest of the beneficiaries. In practice, trustees are appointed by firms sponsoring the pension plans. This means that plan sponsors have effective control of pension decisions.

Prior to the Employee Retirement Income Security Act (ERISA) of 1974, corporate pension plans are governed by the general trust laws. ERISA sets a comprehensive list of standards for pension vesting, funding, termination, and disclosure. It subjects pension trustees and pension asset managers to an explicit set of fiduciary duties, and establishes the Pension Benefit Guaranty Corporation (PBGC) to provide insurance to beneficiaries against pension failures. ERISA is amended by subsequent legislations, such as the Pension Protection Act (PPA) of 2006 (which tightens pension funding and reporting requirements); but its major framework remains intact.

Several aspects of pension laws are of particular relevance to this study. First, Section 403(a) of ERISA explicitly requires pension assets to be held in a trust. This separates pension assets from the rest of a plan sponsor's assets. The same section further sets forth that "a fiduciary shall discharge his duties with respect to a plan solely in the interest of the participants and beneficiaries...". And Section 403(c) of ERISA requires that "the assets of a plan shall never inure to the benefit of any employer and shall be held for the exclusive purposes of providing benefits to participants in the plan ...". Combined, they suggest that pension decisions, including investment decisions, should be made in the interest of plan beneficiaries instead of plan sponsors. However, ERISA falls short of making specific requirements on how plan trustees should be appointed or how pension assets should be invested. With the exception of multi-employer plans (where labor unions have some controls over pension plans), corporate plan sponsors have effective control rights on pensions via the ability to appoint trustees.

Second, when a plan sponsor is in bankruptcy (either Chapter 7 or Chapter 11), because pension assets are held in a separate trust, they are protected from the claims of the plan sponsor's creditors. However, the plan sponsor in bankruptcy may choose to terminate an underfunded pension plan, leaving insufficient pension assets to cover pension liabilities. This distress termination (Section 4041(c) of ERISA) triggers PBGC to take over the pension plan and provide insurance to beneficiaries. Section 4062 of ERISA sets forth that in distress termination, PBGC holds a claim against a plan sponsor for unfunded pension liabilities, which is treated as general unsecured debt in bankruptcy proceedings. Also note that PBGC's pension benefit coverage has limits. In 2017, the maximum insured benefit for a 65-year old retiree in a single-employer plan is $64,000 annually. Further, certain types of vested benefits are not covered by PBGC.[10]

[10]According to a PBGC (2008) analysis of 125 (still healthy) pension plans with 525,000 participants, 16% of the participants would see their benefits reduced if PBGC takes over, and their benefits on average would be reduced by 28%.

Third, the most uncertain legal aspect about pension plans is perhaps who have ownership of pension surplus, i.e., the part of pension assets in excess of pension liabilities. Although ERISA requires pension assets to be managed in the exclusive benefit of plan participants and beneficiaries, firms have several ways to claw back pension surplus into corporate assets, which is known as "surplus reversion". Section 4041(b) of ERISA provides a standard termination procedure, allowing a plan sponsor to terminate a pension plan and claim pension assets under the condition that the sponsor makes alternative arrangements to meet all plan liabilities (e.g., paying a lump sum of cash to beneficiaries or replacing pensions with annuities). Outright terminations of overfunded pensions for surplus reversion were quite popular in the 1980s (e.g., VanDerhei, 1987; Cather, Cooperman and Wolfe, 1991), and generally survived court challenges. But in the Omnibus Budget Reconciliation Act of 1990, the congress imposed a 50% excise tax on surplus reversion. This effectively stopped the practice of surplus reversion through standard pension termination.[11] The excise tax rate is reduced to 20% if at least 25% of the surplus reversion is transferred into a qualified replacement plan (e.g., a defined contribution plan) or at least 20% of the surplus is used to increase the pension benefits of qualified participants. This provides an incentive for pension surplus sharing between plan sponsors and plan beneficiaries.

Finally, because plan sponsors have effective control over pension decisions, there are alternative ways for firms to recapture pension surplus. For example, a firm experienc-

[11]The Tax Reform Act of 1986 initially introduced an excise tax rate of 10% on pension surplus reversion, which was increased to 15% by the Technical and Miscellaneous Revenue Act of 1988, before the 50% rate imposed in 1990.

ing growth in pension liabilities can reduce surplus by simply reducing ongoing pension contributions. The excise tax can be avoided by merging an over-funded plan with an under-funded plan, either within a firm or through a merger of two firms. Finally, pension regulations have some vague parts in treating health benefits and severance benefits, allowing firms to dip into pension surplus to pay for what are otherwise considered normal operating costs or restructuring costs.

2.3 One-period Model

2.3.1 Model Setup

Our baseline one-period model has two dates: time 0 and time T. At time 0, a firm hires a representative employee, contributes to a pension plan, and makes an investment decision for the pension. At time T, the employee retires, the pension investment return is realized, and the employee receives a one-time pension payment.[12]

At time 0, the firm has a certain amount of discretionary cash, denoted by H_0, to be contributed to the pension. The firm can make incremental (non-negative) contributions to the pension fund at time 0 and T, denoted by h_0 and h_T respectively. We assume that the firm additionally incurs a financing cost (or opportunity cost) for pension contribution, denoted as $C_0(h_0)$ and $C_T(h_T)$, respectively. The cost functions $C_0(.)$ and $C_T(.)$ are increasing and convex.

At time 0, the pension can invest in a risk free asset and a risky asset (e.g., a stock).

[12] DB plans cover the longevity risk of retired employees, in that retirees receive annuities as pension payment until death. For simplicity we abstract away from the longevity risk coverage and assume a one-time pension payment.

The gross riskfree rate for the period is R_f, and the gross return on the risky asset for the period is R_m. The pension asset at time T (before time-T pension contribution) is then

$$W_T = W_0(wR_m + (1-w)R_f) \tag{2.1}$$

where w is the portfolio weight of the pension fund on the risky asset, and W_0 and W_T are the values of pension assets at time 0 and T respectively.

The firm's total pension related cash flow at time 0 is the contribution plus the financing cost, i.e., $CF_0 = -W_0 - C_0(h_0)$, where $h_0 = Max(W_0 - H_0, 0)$. Let S denote the pension payment to the employee at time T. Then the firm's time-T total cash flow is $CF_T = W_T - S - C_T(h_T)$ where $h_T = \max(S - W_T, 0)$. S depends on the pension assets W_T and other conditions, e.g., whether the firm is in bankruptcy and how pension surplus is shared. We will provide further details below.

The firm's objective is to maximize shareholder value subject to the employee's participation constraint. We assume that shareholders do not have wealth constraints or portfolio constraints. Thus they value the firm's pension-related cash flows CF_0 and CF_T via the risk-neutral probability approach. Further, let D be the random variable indicating firm bankruptcy. We assume that the firm's bankruptcy event is idiosyncratic, and thus shareholders do not demand a risk premium for the bankruptcy risk. That is, $E(D) = E^Q(D) = p$, where Q denote the risk neutral probability, and the objective probability of bankruptcy is the same as the risk neutral probability of bankruptcy, p.

The employee is wealth constrained. For simplicity, we assume that at time T, the employee does not have any wealth other than the pension payment S she receives. They value the pension payment using an increasing and concave utility function $U(S)$.

The firm's objective is to maximize shareholder value subject to the employee's participation constraint:

$$\max_{w,W_0,S} CF_0 + \frac{1}{R_f} E^Q(CF_T) \tag{2.2}$$

$$\text{subject to: } \; E(U(S)) = \mathbb{U} \tag{2.3}$$

where Q denotes the risk neutral probability. $\mathbb{U}$ is the employee's reservation utility. We omit from the above specification the pension asset process (Equation (2.1)) and several additional constraints imposed by the pension contract, which will become explicit once we specify the pension payment schedule.

2.3.2 Analytical Results for Some Payment Specifications

In the basic model setup above we leave out the details for the pension payment S, which depends on the contractual benefit of the pension plan, pension funding status, firm bankruptcy, and pension surplus sharing arrangements. Here we consider a few simple specifications of S that are susceptible to analytical examination. They serve as illustrations of the basic intuition and highlight some common underlying themes. We later examine more general specifications.

First, consider the following case. Let F denote the promised pension payment at time T. When the firm is not in bankruptcy, we assume that the employee receives the promised benefit F regardless of the pension funding status, while the firm gets the pension surplus $W_T - F$ and bears the pension deficit. In the case of bankruptcy, since the pension assets serve as collateral for pension obligations, we assume that the employee receives the entire pension assets. That is, the pension payment S can be specified as:

$$S_B = F(1 - D) + W_T D. \tag{2.4}$$

We use the subscript B to indicate this base case. As introduced earlier, D is a random variable indicating the firm's bankruptcy event, with $E(D) = E^Q(D) = p$.

We have the following results:

Proposition 2.3.1: *Consider the case of the pension payment specified in Equation (2.4)*

(1) In the case that the firm's default probability is zero (i.e., $p = 0$), the pension fund will invest only in risk free asset (i.e., $w = 0$)

(2) In the case of a positive default probability (i.e., $p > 0$), the pension fund's risky asset allocation is strictly positive (i.e., $w > 0$), and is lower than the optimal allocation of a stand-alone investor who has the same utility as the employee, with the sole exception that the pension is fully funded and the cost function's slope at 0 is higher than some threshold (i.e., $C_T'(0) > C$, where C is specified in the appendix), in which case, the pension fund's stock allocation may be zero.[13]

Proof. See Appendix.

From the firm's stand point, investing in the risky asset generates zero net present value to begin with, and further, the convexity in the financing cost function makes the firm averse to the investment risk. When there is zero probability of bankruptcy, the pension payment is fixed at F with no uncertainty. Therefore, the employee's utility is not affected

[13] We refer to the pension fund as over, under, or fully funded at time zero by comparing the fund asset W_0 with the PBO (i.e., the present value of promised payment F discounted at the risk free rate).

by the pension's investment decision. The firm's aversion to risk drives to the pension to invest completely in the risk free asset.

When a firm defaults, the employee's payoff is the value of pension assets. From the employee's perspective, her utility, conditional on the event of the firm's default, will be maximized if the pension asset is allocated between the risk free asset and the risky asset in line with her risk aversion to achieve a desirable risk and return trade-off. Thus, with a positive probability of default, the employee's preference for the risk-return trade-off is balanced with the firm's aversion to risk, leading to a compromised solution: the allocation on the risky asset will be strictly positive but lower than what is optimal for a stand-alone investor.[14]

To see the intuition, it helps to pay attention to the sensitivity of the employee's pension payment to the pension asset value, which implies how the investment risk is shared between the firm and the employee. We refer to this sensitivity as the delta of S to W_T, or briefly, delta. Formally, in the base case,

$$\Delta_B = (\partial S_B)/(\partial W_T) = D.^{15} \tag{2.5}$$

The systematic risk contained in the employee's pension payment thus depends on

[14]The sole exception in the proposition is when the pension fund is fully funded and the financing cost function has a high slope at zero. In this case, investing in the risk free will result in the time T fund value from the investment being equal to the committed payment. Consider increasing the risky asset allocation by an infinitesimal amount from 0. On the positive side, the expected return for the pension fund investment portfolio will increase proportional to w. On the negative side, the firm faces the chance of fund deficit, which incurs financing cost. Whether it is desirable to move away from the pure risk free investment depends on which of the above two effects dominates, the precise condition of which is given in the proposition.

[15]The partial derivative, conditional on D, is to be understood in the general sense, with zero measured indifferentiable points being ignored.

the firm's bankruptcy risk and asset allocation. When the firm defaults for sure, $\Delta_B = 1$, the employee is exposed to the same amount of systematic risk as in the pension portfolio. When the firm never defaults, $\Delta_B = 0$, the employee has zero exposure to the systematic risk regardless of the asset allocation. Therefore, to achieve the employee's desired risk-return trade-off, the firm adjusts the asset allocation based on its bankruptcy probability which determines the delta.

Proposition 2.3.1 can be extended to cover additional cases. We report results regarding two cases in the following corollary.

Corollary 2.3.1: *(1) In the case the payment is specified as*

$$S = F(1 - D) + \min(W_T, F)D. \tag{2.6}$$

and the default probability is positive, if the pension is underfunded at time 0, the optimal stock market allocation is strictly positive (i.e. $w > 0$*).*

(2) In the case that there is zero probability of default and the payment is specified as

$$S = F + (1 - \alpha)\max(W_T - F, 0), \tag{2.7}$$

where $0 \leq \alpha < 1$*, if the pension is over-funded at time 0, the optimal stock market allocation is strictly positive.*

Proof. See Appendix.

The corollary examines very different cases. It highlight the fact that our basic intuition can apply in a straightforward way to very different scenarios. In (1) of the corollary, the employee receives F when the firm does not default, the same as our baseline case.

Different from the baseline case, when the firm defaults, the employee gets paid by the pension asset up to the level of the promised payment. In (2) of the corollary, we take out the consideration of firm default, and assume that, in the case of pension surplus, the firm keeps a fraction α of the surplus, and the employee gets the remaining fraction $1 - \alpha$. It is worth noting in the first case of the corollary, the employee never receives a payment from the pension higher than the promised payment F. Yet, when the pension investment return is low, she receives less payment in the case of firm default. In effect, she bears the downside risk of the pension investment in the case of firm default. In the second case, there is no default possibility. The pension payment is never below the promised payment F. Yet, the employee shares the upside of pension investment outcome. Despite their apparent differences, in both cases, the optimal investment strategy involves positive allocation to the risky asset. So, as long as the employee is exposed to the investment risk, being it through upside surplus sharing or downside risk bearing, it can be optimal for the pension fund to invest some portion of its asset in the stock market.

It is worth noting that the two cases show two different channels of risk sharing between the firm and the employee. The risk is shared through bankruptcy in the first case while through surplus sharing in the second. Furthermore, the delta of the employee's pension payment is determined by the firm's bankruptcy probability and the employee's share of surplus respectively.[16] To further examine the richness of different variations, we resort to numerical analysis.

---

[16]In the first case, $\Delta = (W_T < F)D$. In the second case $\Delta = (1 - \alpha)(W_T > F)$. $(W_T < F)$ and $(W_T > F)$ indicate two different regions.

2.4 Numerical Analysis

In this section, we numerically calibrate the one-period model. We consider a variety of specifications of the pension payment schedule. We focus on the optimal risk allocation for the pension asset, and examine how different parameters affect this solution.

2.4.1 Calibration

We calibrate the model parameters in the following way. First, we set the risk premium on the risky asset, the riskfree rate, and the volatility of the risky asset to the historical averages of market risk premium, riskfree rate, and market return volatility from the Fama-French data on the market portfolio from 1926 to 2016. Specifically, the annual risk premium for the risky asset is 6.1%, the annual risk free rate is 3.3%, and the annual volatility of the risky asset return is 18.5%.

We assume that the employee retires in 30 years (i.e., T=30). We assume that the firm has a BBB credit rating, with an annual bankruptcy probability of 0.5% (from Berk and DeMarzo, 2016). This means that over a 30-year horizon, the probability of firm bankruptcy is 13.93%. We assume a CRRA utility function for the employee with the relative risk averse coefficient $\gamma = 6$, based on Constantinides (1990).

To calibrate the employee's reservation utility, we take the following approach. We consider a hypothetical pension commitment F and use the risk free rate to discount F to estimate the Projected Benefit Obligation (PBO), i.e., the present value of promised pension payment. Note that all the variables calibrated so far are scale invariant. So, it is without lose of generality to unify this hypothetical PBO to 1. Then we assume that the investor is handed this amount (PBO) to invest for herself for the 30-year period. We use the resulting utility at year 30 as the reservation utility. According to Chen, Yao, Yu and Zhang (2014), the average pension PBO is \$1,016.25 million. Therefore, 1 unit of PBO in our model represents approximately \$1 billion.

Following Hennessy and Whited (2007), we assume a quadratic form for the financing cost function. Specifically, $C_t(h) = c_a + c_b h + c_c h^2$, where $t = 0, T$. We take their estimated values with the following adjustment. In Hennessy and Whited (2007), the financing amount h is in millions. Given that our unit is in billions, we make the scale adjustment accordingly. We set the cost function parameters as $c_a = 5.98 \times 10^{-5}$, $c_b = 0.091$, $c_c = 0.4$. We assume that the firm's initial discretionary cash H_0 is zero.

We consider a more general form of the pension payment schedule S:

$$S = F(1 - D) + \min(W_T, F)D + (1 - \alpha) \times \max(W_T - F, 0). \tag{2.8}$$

In this specification, the employee gets the promised payment F when the firm does not default, represented by the term $F(1 - D)$. In the case of a default and if the pension asset falls below the promised payment, the employee only gets the pension asset, as represented by the term $\min(W_T, F)D$. Finally, if there is a surplus, the firm keeps a fraction α of the surplus and the employee gets the remainder fraction $(1 - \alpha)$, as represented by the term $(1 - \alpha) \times \max(W_T - F, 0)$. For the baseline setup, we assume that the firm keeps 80% of the surplus while the employee keeps the remaining 20%. That is, $\alpha = 0.8$. The baseline parameter choices are summarized in the following table.

Table 2.4.1: Benchmark Parameters

μ	r_f	σ	c_a	c_b	c_c	p	γ	H_0	α
0.094	0.033	0.185	5.98×10^{-5}	0.091	0.4	0.005	6	0	0.8

2.4.2 Comparative Results

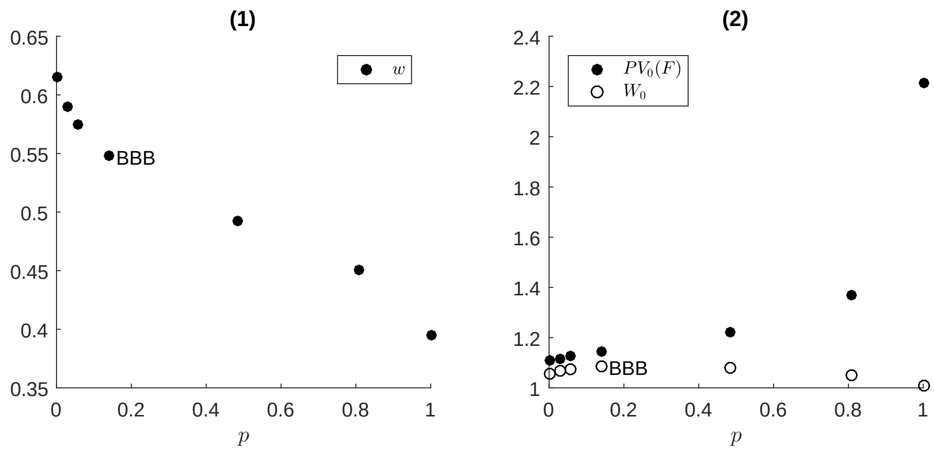
Figure 2.1 plots the portfolio's risk allocation (w), the promised pension payment (F) and the initial contribution (W_0), when we vary the default probability while keeping other parameters constant. Panel (1) of the figure shows that for the baseline firm with a credit rating of BBB, the portfolio weight on the risky asset is about 55%, similar to the observed level of allocation to risky assets by corporate pensions as a whole. Further, in Panel (1), we see that the portfolio's risk allocation decreases as the default probability increases. This is consistent with the empirical findings in existing studies; e.g., Bodie et al. (1985), Rauh (2008), and An et al. (2013).

With the general form of pension payment schedule, the sensitivity of the employee's pension payment to the pension asset value, delta, becomes

$$\Delta = (\partial S)/(\partial W_T) = (W_T < F)D + (1-\alpha)(W_T > F). \tag{2.9}$$

Note that the firm is averse to the market risk exposure due to the concern of external financing cost. The employee on the other hand finds the market risk premium attractive due to her under exposure to the stock market investment on her personal wealth. From a benevolent social planner's perspective, a high delta is desirable because it increases the employee's exposure to the investment risk while reducing the firm's exposure, ceteris paribus. Given that Δ in Equation (2.9) is affected by the firm's decisions, in particular,

Figure 2.1: Impact of Bankruptcy Probability on Pension Investment and Pension Funding Decisions



This figure illustrates the impact of bankruptcy probability (p) on pension investment and funding decisions. Panel (1) shows the pension allocation to the risky asset (w). Panel (2) shows the optimal pension PBO ($PV_0(F)$) and optimal pension contribution (W_0). The bankruptcy probabilities are calibrated to those for firms with credit ratings of AAA, AA, A, BBB, BB, and B, respectively.

in F and W_0, the firm makes its choices in an attempt to enlarge the effective delta while balancing other considerations.

Consider the case of default probability being 0. The pension payment takes an option-like form: $S = F + (1 - \alpha) \times \max(W_T - F, 0)$. That is, the employee gets the fixed payment F in the case that the pension's investment asset falls below F. In the case the investment asset is higher, the employee gets $1 - \alpha$ share of the surplus. Thus, $\Delta = 0$ in the region $W_T < F$ and $\Delta = 1 - \alpha$ in the region $W_T > F$. The average delta is thus somewhere in between 0 and $1 - \alpha$. With a relative low delta, it takes a high allocation of the pension fund to the risky asset for the employee to achieve desirable risk exposure from the pension payment. The firm invests thus aggressively in order to get closer to the employee's desired level. Furthermore, as shown in Panel (2) of the figure, the firm chooses a relatively low promised payment, which helps enlarge the relatively high delta region (i.e. the region of $\Delta = 1 - \alpha$ or equivalently $W_T > F$) and reduce the 0 delta region (i.e, the region of $W_T < F$), with the effect of a higher average delta. The higher average delta is desirable because it improves the effectiveness of passing the risk exposure of the pension fund investment to the pension payment that the employee gets. Furthermore, as part of the optimal strategy that involves aggressive risk allocation, the firm chooses to fund aggressively at time 0 which helps reduce the financing concern at time 1 and thus reduce the potentially high impact of the convex financing cost from the aggressively risky investment.

On the other extreme, consider the case that the firm defaults for sure. The pension payment is $S = \min(W_T, F) + (1 - \alpha) \times \max(W_T - F, 0)$. The pension payment's delta to the pension investment asset is 1 in the region of $W_T < F$, and the delta is $1 - \alpha$ in the region of $W_T > F$. The average delta is thus somewhere in between $1 - \alpha$ and 1. The overall delta in the case is in a relative high comparing with the first case discussed above. The riskiness of the pension asset investment can be effectively transferred to the employee through the highly variable pension payment. Therefore, the firm needs only to allocate a modest amount of the pension asset in the risky asset to achieve optimal. The firm would choose a high promised payment F in order to enlarge the average delta by enlarging the region of delta being 1 (i.e., the region of $W_T < F$). There will be severe underfunding at time zero, in order to balance the effect of the high level of the promised payment F, so as to keep the employee's utility at the reservation level. Doing so is optimal for two reasons. First, the future financing cost is a not a concern as it incurs only in the nondefault case. Therefore, underfunding carries no penalty. Second, low W_0 further carries the effect of increasing the probability of $W_T < F$, the high Δ region. This is desirable as it increases the average Δ. Once we understand the two extreme cases of the default probability, the firm's decisions for intermediary cases fall naturally somewhere in between the two extremes.

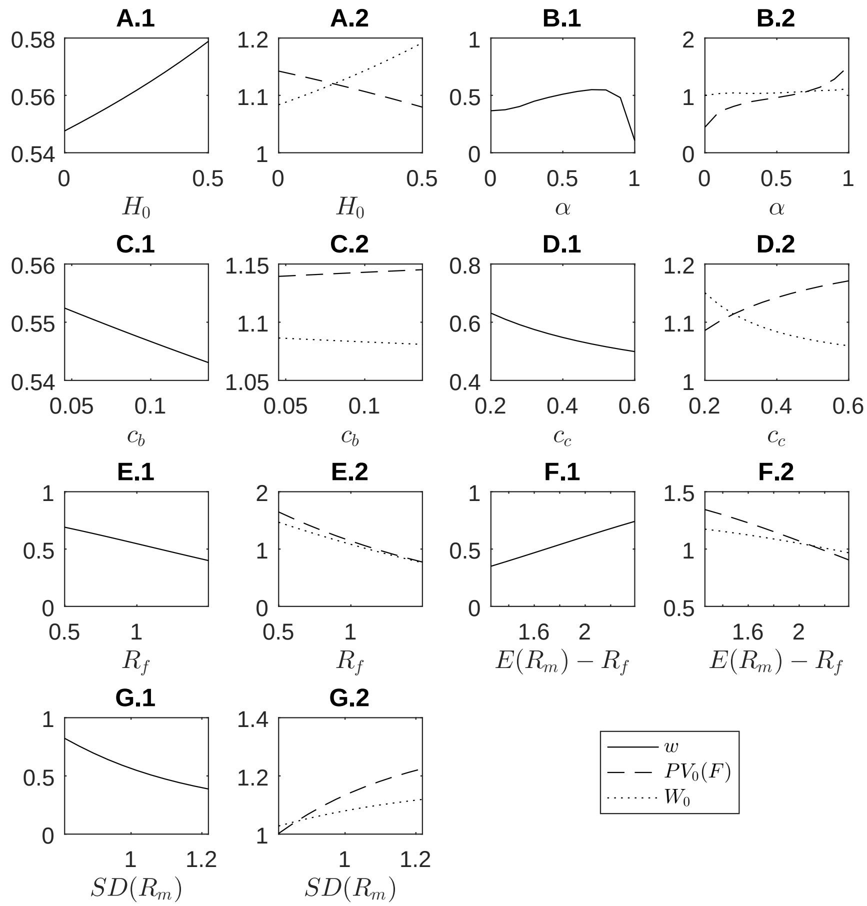
In Figure 2.2, we examine the effects of various factors on the optimal asset allocation and the initial funding. In Panel A, we look at the effect of the initial free cash flow. The intuition for this analysis is straightforward. With high free cash flow at time 0, naturally the firm contributes more at time 0, as shown on the right side chart of Panel A. Such contribution reduces the concern the potential deficit at time T. As a consequence, as shown on the left chart of the panel, the firm makes more aggressive risk allocation for the pension fund to be in line with the employee's preference on the risk-return trade-off. Consequently, the firm can promise a lower payment F while still keeping the employee at the reservation utility, as shown on the right chart of the panel.

In Panel B of Figure 2.2, we examine the effect of varying the ratio of surplus that the firm keeps. The left chart of Panel B shows a non-monotonic relation. As the firm's share of surplus (i.e., α) increases, the risk allocation initially increases and then decreases. As the firm keeps some surplus, the pension payment's delta in the region of $W_T > F$ decreases. Consequently, the overall average delta decreases as well. Naturally, it would take a more aggressive risk allocation in the pension fund to offset the decrease in delta, in order to keep the riskiness of the pension payment leveled. Indeed, the firm acts accordingly, as shown in the chart. As the firm's share of surplus further increases, the delta in the no-default case starts to approach zero. As a result, any reasonable level of risk allocation in the pension fund leads to only meager stock market exposure of the pension payment to the employee. That is, with extreme low values of delta, the channel of passing the risk exposure of the pension investment to the employee through the varying pension payment is by and large shut down, making it ineffective for the pension fund to provide the risk exposure to the stock market for the employee. Overall, when the firm takes large enough share of the surplus, further increasing in the firm's share of the surplus and thus reducing the employee's share of the surplus, the optimal risk allocation decreases.

As the firm's share of the surplus increases, the delta of the pension payment to the pension fund investment value decreases in the surplus region (i.e., the region of $W_T > F$). As shown in Plot B.2 , the firm chooses a high level of the PBO and accordingly the low level of the the initial funding ratio so as to reduce the probability of the surplus region, the region of the reduced delta. Overall, the firm chooses the funding ratio in an attempt to maintain a relatively high average delta.

The cases discussed above help establish the basic intuition of our model. We further examine other comparative statistics, including the effect of varying financing cost (Panels C and D of Figure 2.2), the effect of varying stock market condition (Panels E, F, and G). The results are largely consistent with our model's basic intuition: the optimal risky allocation decision is reached by negotiating between the demand of risky exposure from the employee and the convex financing cost.

Figure 2.2: Impact of Additional Factors on Pension Investment and Pension Funding Decisions



This figure illustrates the impact of an additional set of factors on pension investment and funding decision. Panel A shows the impact of firm's initial free cash (H_0) on pension allocation (w) to risky assets and optimal PBO ($PV_0(F)$) and the firm's pension contribution (W_0). Panel B shows the impact of pension surplus sharing (percentage of surplus to firm, α). Panel C shows the impact of the linear component of financing cost (c_b). Panel D shows the impact of the quadratic component of financing cost (c_c). Panel E shows the impact of risk free rate (R_f). Panel F shows the impact of the risk premium of the risky asset ($E(R_m) - R_f$). Panel G shows the impact of the volatility of the risky asset ($SD(R_m)$).

2.4.3 Cross-section of Pension Investments

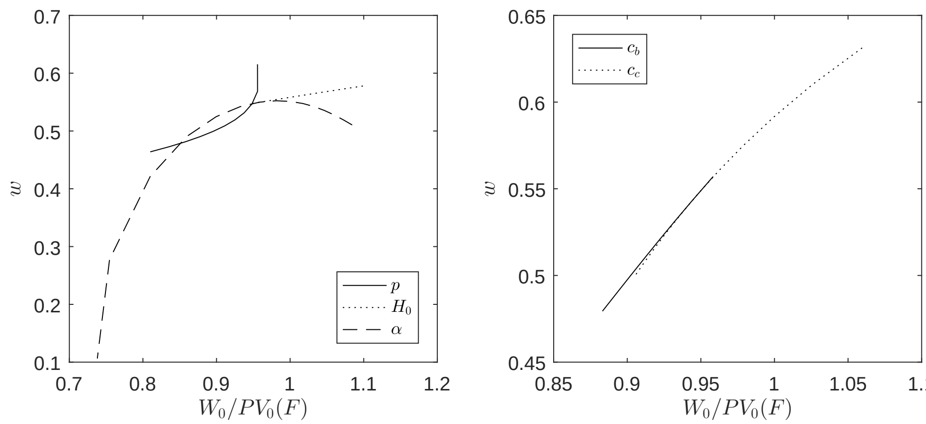
In our model, both the initial funding ratio and the asset allocation are endogenous. Thus, the empirically observed cross-sectional relation between the funding ratio and the risk allocation should be interpreted with care. In this subsection, we examine various cross sectional relations that can result from the model. The results of the analysis are presented in Figure 2.3.

In the model, a simple way to vary the initial funding ratio is to vary the initial free cash. If a firm has less free cash at time 0, it is naturally costly for the firm to fund the pension generously. The pension is thus likely to be underfunded, ceteris paribus. How the initial free cash affects the firm's initial funding, PBO, and the asset allocation is examined in Panel A of Figure 2.2, in Subsection 2.4.2. Here, we present the cross-sectional relation of the initial funding ratio and the risky allocation in Plot 1 of Figure 2.3 (the dot-line). Basically, as the initial funding ratio due to higher initial free cash flow, the fund's risky allocation is also higher.

In the plot, we also examine the cross-section by letting the bankruptcy probability vary. Some properties of this case are examined in Figure 2.1, in Subsection 2.4.2. As the bankruptcy probability increases, the promised obligation increases and the initial funding decreases. At the same time, the risky asset allocation decreases. So, if we simply relate the funding ratio and the asset allocation, we again see a positive relation.

Finally, in the plot, we present the cross-section by varying the surplus sharing. Some properties of the case are examined in Figure 2.2, Panel B, in Subsection 2.4.2. As the firm's share in the upside increases, the employee's share in the upside decreases. The

Figure 2.3: Relation Between Optimal Investment and Optimal Funding Ratio



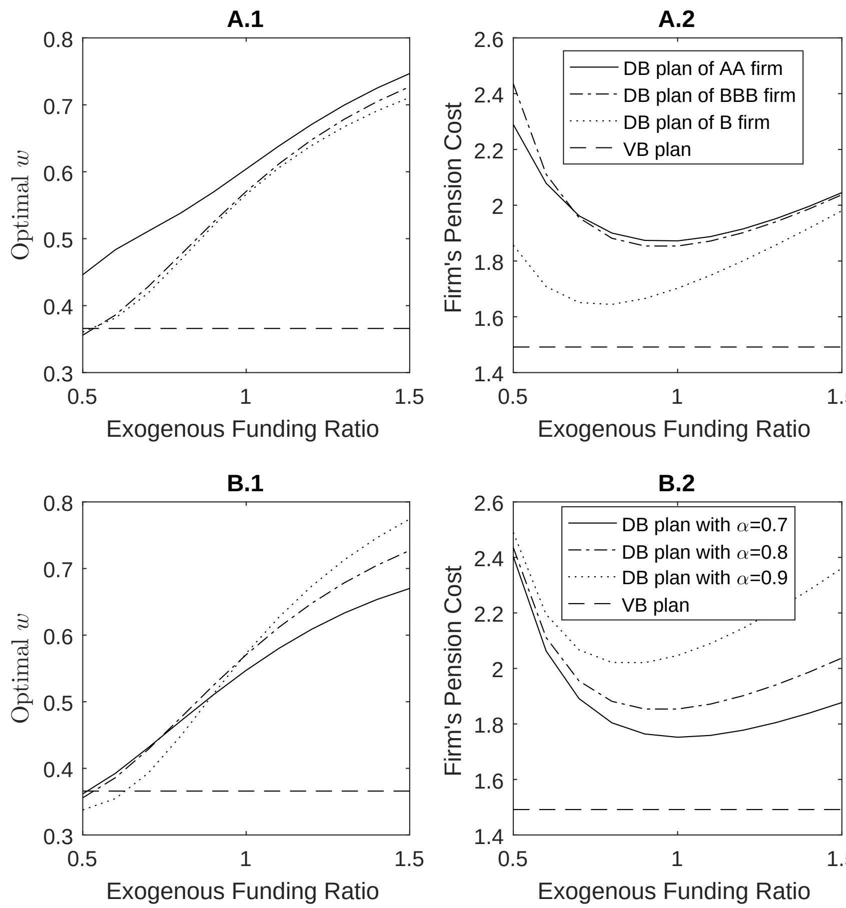
This figure illustrates the relation between optimal pension investment (w) and optimal funding ratio ($W_0/PV_0(F)$) under varying parameter values. Plot (1) shows the relation under various parameter values of bankruptcy probability (p), initial free cash (H_0), and surplus sharing (α). Plot (2) shows the relationship under varying values of linear financing cost (c_b) and quadratic financing cost (c_c).

funding ratio decreases and the asset allocation increases initially and decreases at the end. The intuition is discussed in detail in Subsection 2.4.2. Overall, there is a negative relation between the funding ratio and the risky allocation in some range of the two variables when the firm's share of the surplus is low, but in other range of the two variables when the firm's share of the surplus is high.

In summary, we see that, with different choice of the underlying force that causes the cross-sectional variation, the relation between the funding ratio and the risky allocation is mostly positive, but can be negative.

2.4.4 A "Variable Benefit" Plan

Figure 2.4: Pension Investment and Pension Funding Cost Under Exogenous Funding Ratios



This figure illustrates pension investment (w) and funding cost under exogenous funding ratios. The pension funding cost is the firm's cash outflow at time 0 plus the present value of expected cash outflow at time T. Panel A shows pension investment and funding cost with varying exogenous funding ratios where the bankruptcy probabilities are calibrated to those of AA-rated firms, BBB-rated firms, and B-rated firms. Panel B shows the impact of varying surplus sharing schemes, where employees receive 30% ($\alpha = 0.7$), 20% ($\alpha = 0.8$), and 10% ($\alpha = 0.9$) of pension surplus. For comparison, both panels include pension investment and pension funding cost of variable benefit plans, in which the employee receives 100% of pension assets and no firm guarantee on pension benefits (essentially a defined contribution plan).

Our analysis so far suggests a general result that it is optimal for the employee to bear a large fraction of pension investment risk, i.e., having a high delta. The defined benefit feature, however, severely limits employee's exposure to investment risk by having a low value of delta. This leads naturally to the question we explore in this part: what if we let the employee bear the full risk of the investment? That is, let $S = W_T$. This makes the employee's payoff similar to that in a defined contribution plan, although we still assume that the investment asset is pooled among all the employees and is managed uniformly, and the firm insures the longevity risk (thus, strictly speaking, a hybrid of the defined benefit plan and the defined contribution plan). We call it the variable benefit plan (the VB plan). Such a plan eliminates the firm's concern of financing cost at time T. More important, such a plan has the highest possible delta in our model, namely 1.

We compare the outcome of the defined benefit plan with that of the variable benefit plan. In particular, we examine how asset allocation differs across the two plans, and how much more the defined benefit plan costs the firm than the variable benefit plan while keeping the employee's reservation utility fixed. In Panel A of Figure 2.4, we study the comparison for firms with different bankruptcy probabilities. For this part of the analysis we keep the funding ratio exogenous. One can also interpret the result as if there is an exogenous shock to funding ratio. Alternatively, we can generate essentially the same chart by varying the time 0 free cash flow. The plot shows that as the funding ratio increases, similar to the results in Panel A of Figure 2.2, the risky allocation increases.

Interestingly, the DB plan in general invests much more aggressively than the VB plan, as shown in Panel A.1. In Panel A.2, we see that the total funding cost of the DB plan to the firm is much higher than that of a VB plan (about 40% higher for an investment-grade firm). The total funding cost is calculated as $CF_0 + \frac{1}{R_f}E^Q(CF_T)$. It may come as a surprise that the funding cost of the DB plan is the lowest for a low credit-rating firm. Thus, the results highlight that the importance of risk sharing in pension design. To minimize pension funding cost, it is optimal for the employee to bear all the (systematic) investment risk.

In recent decades, there is a trend of firms shifting away from defined benefit plans into defined contribution plans. One reason often cited by firms making this shift is that defined benefit plans have much higher funding cost (Munnell and Soto, 2007). The analysis performed here supports this notion.

2.4.5 The Effect of PBGC Insurance

In this section, we analyze the effect of PBGC pension insurance. PBGC provides insurance coverage on employees' promised pension benefits up to a ceiling — in 2017, \$64,000 a year for a retiree at age of 65. For highly paid employees such as pilots, PBGC insurance provides only a partial coverage. To capture the partial coverage feature of PBGC insurance, we assume that the employee receives θF, with $\theta \in (0,1)$, from PBGC when the firm is bankrupt and when the pension is underfunded. We initially set $\theta = 0.7$ and investigate variations of θ in later analysis. The employee's expected pension payment becomes:

$$S = F(1-D) + \min(W_T, 0.7 \times F)D + 0.2 \times \max(W_T - F, 0). \tag{2.10}$$

We also assume that PBGC has access to the optimal risky portfolio, but is thinly capitalized and faces a convexly shaped financing cost function (with lower parameter values relative to the firm's). We set $c_{ins,a} = 0$, $c_{ins,b} = 0.0353$, and $c_{ins,c} = 0.1333$. The insurance premium by PBGC is considered fairly priced if PBGC breaks even on its insurance. Too low a premium (undervalued premium) will result in PBGC subsidizing the firm, and too high a premium will result in PBGC gains at the expense of the firm. To model the over-/under-pricing of PBGC insurance, we follow the insurance literature (i.e., Doherty and Schlesinger 1990) and set the PBGC insurance premium as the product of the risk-neutral expectation of cash flow to PBGC in case of bankruptcy and a loading factor. The insurance is overpriced if the loading factor m is greater than 1, underpriced if m is smaller than 1, and fairly priced if m equals to 1. Thus PBGC insurance premium can expressed as

$$I = mD\frac{1}{R_f}E^Q\left[\max(0.7 \times F - W_T, 0) + C_B(\max(0.7 \times F - W_T, 0))\right] \tag{2.11}$$

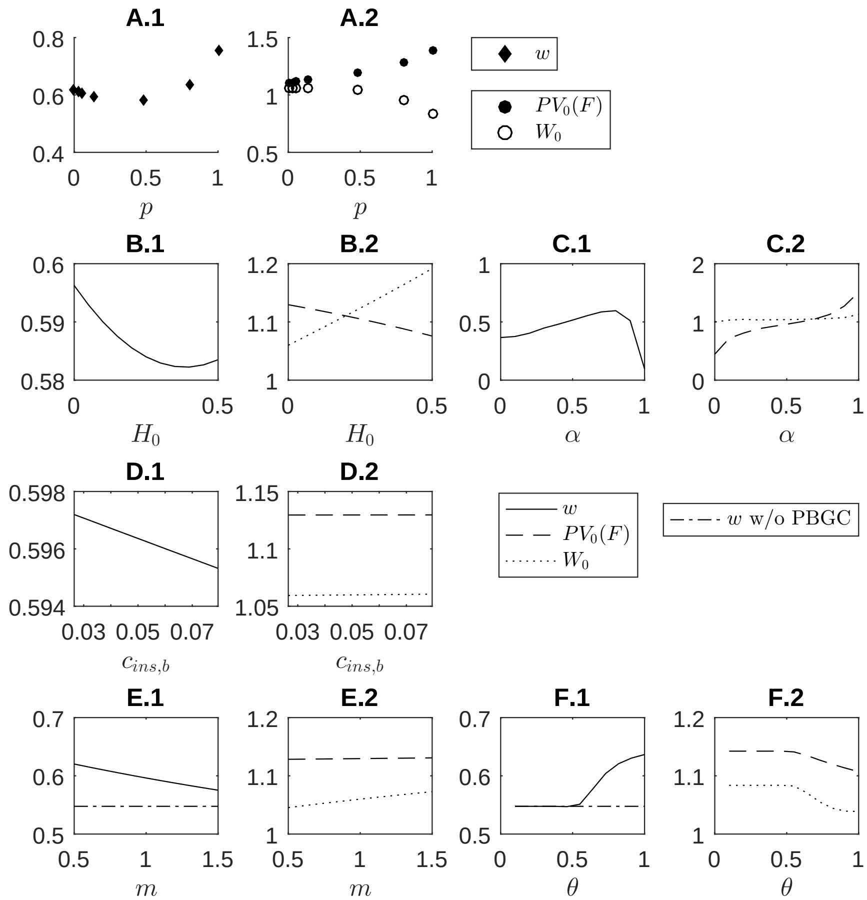
where C_B is the financing cost function of PBGC. Firms need to purchase PBGC insurance when forming the pension portfolio, which changes firm's cash flow at period 0 to $CF_0 = H_0 + h_0 - W_0 - I$. Again, initial contribution h_0 needs to be chosen such that CF_0 is non-negative.

Figure 2.5 displays the pension plan's risky allocation and funding decision in the presence of PBGC. In Panel A, we look at the effect of bankruptcy risk. When the bankruptcy risk is relatively low (less than 0.5), the relation between risky allocation and bankruptcy risk is similar to the case without PBGC insurance – pension investment becomes more aggressive when the bankruptcy risk is higher. However, as the bankruptcy probability increases above 0.5, the relationship is reversed. Two effects are driving the results here. First, insurance coverage reduces the employee's downside risk exposure in bankruptcy, therefore the employee desires more risky allocation as the perceived bankruptcy risk increases. Second, higher bankruptcy risk increases insurance premium, which discourages the firm from investing in the risky asset. The pattern shown in Panel A.1 results from the trade-off of these two effects. Finally, the relation between pension funding and bankruptcy risk in Panel A.2 is similar to that without PBGC – the pension becomes more underfunded when bankruptcy risk increases.

In Panel B, we investigate the effect of initial free cash. When firm has little initial free cash to fund pension plan, it takes advantage of pension insurance and the lower financing cost of PBGC by investing aggressively in the risky asset and contributing as little as possible. Essentially, the firm bets on its bankruptcy and PBGC insurance to meet the employee's reservation utility. As the firm's initial free cash increases, the above strategy becomes less attractive. The firm starts to contribute more, promise less, and allocate more to the risky asset.

The effect of surplus sharing, as shown in Panel C, is similar to the case without PBGC. Further, Panel D shows that the effect of PBGC's financing cost basically replicates the pattern of firm's own financing cost at time T.

Figure 2.5: Pension Investment and Pension Funding Decisions with PBGC Pension Insurance



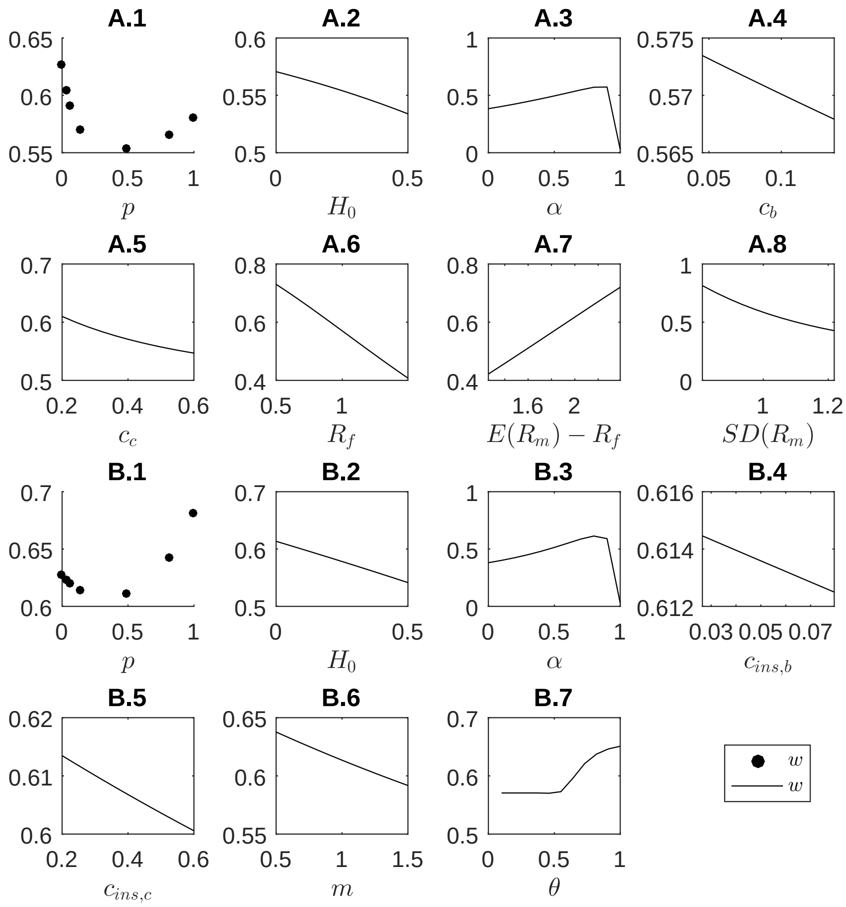
This figure illustrates the impact of various factors on pension investment and funding decisions under PBGC pension insurance coverage. Panel A shows the impact of bankruptcy probability (p) on allocation to risky assets (w), and the optimal PBO ($PV_0(F)$) and pension contribution (W_0). Panel B shows the impact of firm's initial free cash (H_0). Panel C shows the impact of pension surplus sharing between the firm and the employee (α). Panel D shows the impact of PBGC's linear component of financing cost ($c_{ins,b}$). Panel E shows the impact of underpriced and overpriced PBGC pension insurance (m). Panel F shows the impact of PBGC pension insurance coverage (θ).

As shown in Panel E, pension insurance encourages pension risk taking particularly when the insurance premium is under valued. Even with fair valued insurance, the risky allocation under PBGC insurance is still higher than the case without PBGC. This is because of the lower financing cost of PBGC relative to that of the firm. Meanwhile, overpriced PBGC insurance makes the firm promise more to compensate for the reduction in risk exposure. Pension plan becomes less underfunded as a result of increased insurance premium.

Finally, we look at the effect of partial PBGC coverage in Panel F. Higher PBGC coverage reduces the employee's downside risk exposure, which requires firm to make more risky investment. If the PBGC coverage is too low, the employee bears almost the same downside risk exposure as in the case without PBGC. And as a result, the risky investment in the presence of PBGC gets close to that without PBGC.

For robustness, we further examine the cross-sectional relations among key variables in a setting where the pension is fully funded at the initial stage. We thus focus only on the the firm's asset allocation decision, while the fund contribution decision is fixed by assumption. We repeat the above numerical exercises for both the cases of with and without the PBGC. Figure 2.6 shows that all the results obtained earlier in this section remain qualitatively robust to this variation. It is noticeable that even without PBGC insurance, firms with overly high bankruptcy probability when required to fully fund the pension plan will have a very aggressive investment policy. This is because for firms with high bankruptcy probability, it is optimal to underfund the pension plan, which, however, is not allowed in this setting. Thus firms lower the promised pension payment and compensate employees with higher risky allocation.

Figure 2.6: Investment Decisions by Initially Fully-funded Pensions



This figure illustrates the impact of various factors on pension asset allocation when a pension is initially fully-funded. Panel A shows the case without PBGC pension insurance. Panel B shows the case with PBGC pension insurance.

2.5 Dynamic Model

While our one-period model delivers a rich set of results, it assumes that firms cannot adjust pension investment allocations or adjust pension contribution in the interim. This might be restrictive. In this section, we consider a dynamic model that relaxes these restrictive assumptions.

2.5.1 Model Setup

The entire time period in the dynamic model is from time 0 to T. At time 0, the firm hires the employee. At time T, the employee retires and receives the pension payment. The investment opportunity again consists of two assets, the risk free asset with the instantaneous rate r_f, and the risky asset its value process following a geometrical Brownian motion:

$$\frac{dP_m}{P_m} = (\mu + r_f)dt + \sigma dB \tag{2.12}$$

where B is a standard Brownian motion, μ the risk premium, and σ the asset volatility. We also assume that the market is dynamically complete, and there is the unique the risk neutral probability Q.

$$\frac{dP_m}{P_m} = r_f dt + \sigma dB^Q, \tag{2.13}$$

where B^Q is a standard Brownian motion under the risk neutral probability Q.

Let w_t be the portfolio weight on the market portfolio. We denote the cumulative pension contribution by the firm up to time *t* by H_t, which is a non-decreasing stochastic process. Together, the pension value process is governed by

$$dW_t = W_t[(w_t\mu + r_f)dt + w_t\sigma dB] + dH_t. \tag{2.14}$$

To facilitate numerical solutions, we assume that the firm only adjusts its portfolio weight and makes pension contribution at T+1 discrete time points: at times $n = 0, 1, \ldots, T$. Specifically, the portfolio weight is kept constant between the two discrete decision time points, i.e., $w_t = w_n$ if $t \in [n, n+1)$ for integer n. Furthermore, the cumulative contribution process H is a stochastic process that increases only by steps at the discrete time points. That is, $H_t = H_n$ if $t \in [n, n+1)$, and $dH_n = h_n$ for some discrete process h_n.[17] At the time when $dH_n > 0$, a financing cost is incurred. We denote the cumulative cost process by C_t. We use $C(\cdot)$ to denote the cost function. Formally, we can write $dC_t = C(dH_t)(dH_t > 0)$ for $t \in [0, T]$. Of course, dC_t defined this way can be non-zero only at the discrete time points n and only when there is a positive pension contribution. Specifically, $dC_n = C(h_n)(h_n > 0)$.[18]

The firm's default indicator, D_t, is a dynamic process governed by the first arrival of a Poisson process with constant intensity p. We denote the firm default time by τ, a stopping time defined by $\tau = \min(t|D_t = 1)$. The pension will be liquidated and payment is made to the employee either when the firm defaults at time τ ($\tau \leq T$) or at the terminal time T. The payment schedule is in principle the same as in Equation (2.8) with some notation adjustments to accommodate the dynamic model structure. Specifically,

$$S_\tau = F(1 - D_\tau) + \min(W_{\tau-}, F)D_\tau + (1 - \alpha) \times \max(W_{\tau-} - F, 0) \tag{2.15}$$

The firm keeps the difference between the pension assets and pension payment. That is, $CF_\tau = W_\tau - S_\tau$. At the payment time, we further adopt the convention that $W_\tau \geq S_\tau$. That is, if the pension's asset falls short of the payment S_τ (i.e., $W_{\tau-} < S_\tau$), the firm's

[17] A more precise way of writing the dynamic of H_t is $dH_t = \left(\sum_{n=1}^{T} h_n \delta(t-n)\right) dt$, where $\delta(\cdot)$ here stands for the standard delta function in time. We choose a more heuristic way of writing over the mathematically more accurate way for notational simplicity. Furthermore, the notation δ is already occupied by a term of common financial meaning in our paper. All the following similar expressions in the paper needs to be understood under this light.

[18] We adopt de Finetti's notation convention and not distinguish a random event with its indicate variable.

time-τ contribution has to make up the shortfall. This ensures that CF_τ is non-negative.

The employee has a CRRA utility function on time-T wealth, $U(S) = \frac{S^{1-\gamma}}{1-\gamma}$. If at time τ, the employee receives the pension payment as the consequence of firm bankruptcy, she will invest optimally with the weight on the market portfolio $w = \frac{\mu}{\gamma\sigma^2}$ for the remainder time interval $[\tau, T]$. Thus,

$$S_T = S_\tau e^{(r_f + w\mu - \frac{1}{2}w^2\sigma^2)(T-\tau) + w\sigma(B_T - B_\tau)} \tag{2.16}$$

The firm values future cash flows (including the non-negative cash flows, the contributions, and the financing cost) using the risk neutral probability and the risk free discount rate. The firm's optimization problem is:

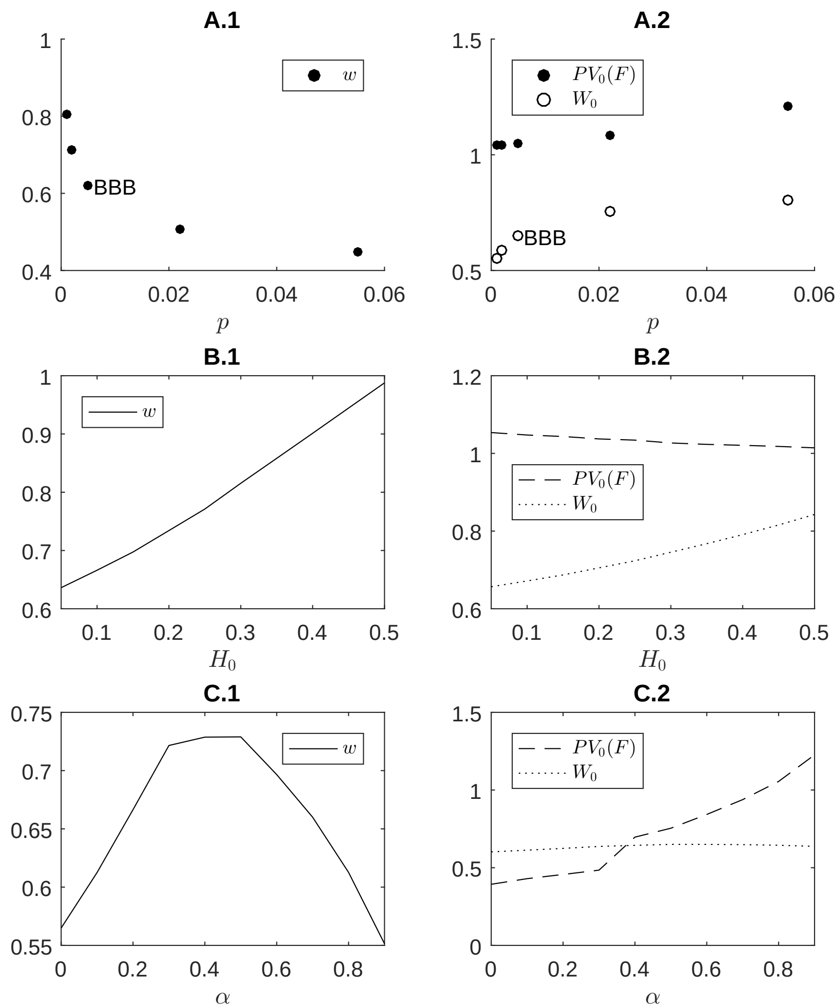
$$\max E^Q \left[e^{-r_f \tau} CF_\tau - \int_0^\tau e^{-r_f t}(dH_t + dC_t) \right]. \tag{2.17}$$

$$\text{subject to:} \;\; E(U(S_T)) = \mathbb{U}. \tag{2.18}$$

There are additional constraints, including CF_τ being non-negative, pension asset value process being governed by Equation (2.14), and the final pension payment to the employee being specified by Equations (2.15) and (2.16).

We numerically solve the dynamic model. Most of the economic parameters, including the stock market information, and parameters in the financing cost function, the employee's risk aversion, the number of years till retirement, and the annual default rate are the same as those in the one-period model. The only difference is that we assume the fixed cost of external financing as zero in the dynamic model to smooth the optimization solutions.

Figure 2.7: Initial Investment and Initial Cash Contribution Decisions in the Dynamic Model



This figure illustrates the firm's asset allocation (w) and cash contribution (W_0) at time 0 in the dynamic model. Panel A shows the optimal allocations and cash contributions at time 0 under different bankruptcy probabilities (p). Panel B shows the optimal allocation and cash contribution at time 0 under different amounts of initial free cash (H_0). Panel C shows the optimal allocation and cash contribution at time 0 under different surplus sharing ratios (α).

2.5.2 Initial Pension Decisions

Within the dynamic model setup, we revisit some of the cross-sectional relations observed for the one-period model. In particular, we examine how the firm's initial (time-0) pension funding decision and the initial (time-0) investment decision vary when we change model parameters. The results are reported in Figure 2.7. We have three sets of results regarding the cross-section of the firm's initial investment allocation (w_0), cash contribution (h_0) and PBO (F), as we vary the probability of bankruptcy (p), the firm's initial free cash (H_0), and the percentage of the pension surplus the firm keeps (α).

The cross-sectional patterns observed in Figure 2.7 is qualitatively similar to those seen in Figures 1 and 2.

In Panel A.1 of the figure, similar to what we find in Figure 2.1 for the one-period model, we see that a firm with higher default probability tends to choose less risky allocation, everything else the same. For an AAA-rated firm, the risky asset weight of the portfolio is above 80%. In comparison, for a BBB-rated firm, the weight drops to 62%. Also note that the in general the risky allocation in the dynamic model here is higher that that in the one-period model. Thus, having the ability to adjust asset allocation and make pension contribution at multiple time points does not reduce a pension's investment risk-taking.

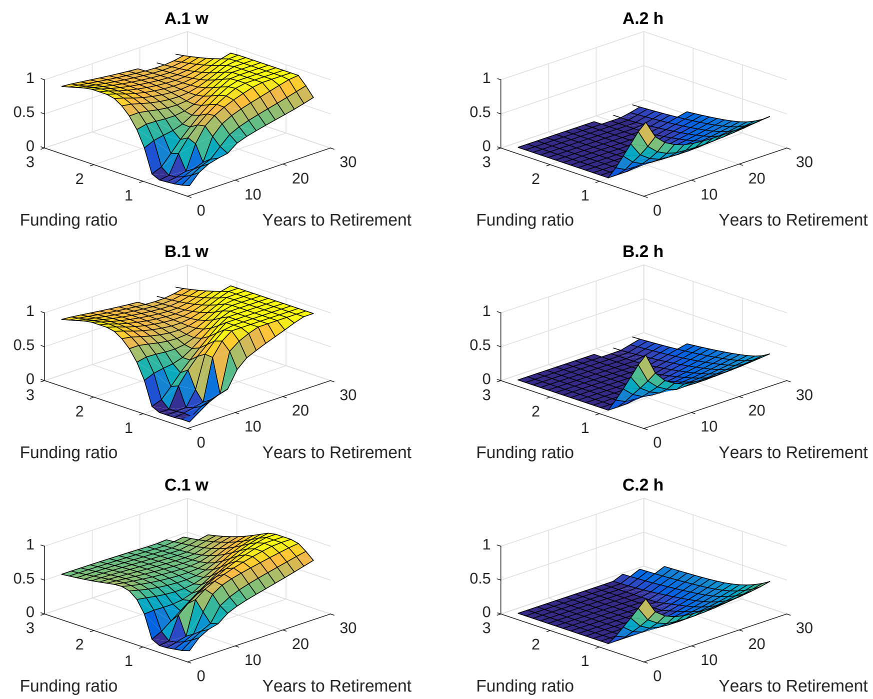
In Panels B and C, we vary the initial free cash (H_0) and the parameter of the surplus sharing (α), respectively. The results are qualitatively similar as those in Figure 2.2. Specifically, when the firm has more free cash at time 0, it contributes more to the pension fund. This further leads to more aggressive asset allocation because more initial contribution mitigates the concern of financing cost due to future funding shortfalls. The relation between investment risk-taking and the fraction of pension surplus kept by the firm is non-monotonic, with investment risk being the highest when the firm's share of the surplus is moderate. The investment risk is lower when the firm's share of surplus is either very high or very low. The intuition is the same as discussed in Section III.2. It is worth noting that the portfolio weight on the risky asset stays above 55% throughout the entire range of the surplus sharing parameter.

Overall, we find that in the dynamic model, corporate pension may continue to take substantial investment risk, and the relation of pension risk taking with some of the key parameters remain similar to what we find in the one-period model. Thus, the ability of the firm to dynamically adjust pension contribution over time does not significantly reduces pension risk taking.

2.5.3 The Dynamics of Pension Investments and Contributions

We further investigate how pension investments and pension contributions vary over time. Figure 2.8 displays the portfolio weight on the risky asset over time under different sets of parameters. In Panel A, we start with the benchmark case where the annual default probability is 0.05% (for BBB-rated firms) and the firm's fraction of pension surplus is 80%. We show the portfolio weights in the left plot and pension contributions in the right plot. In Panel B, we change the default probability to 0.01% (for AAA-rated firms). In Panel C, we change the surplus sharing parameter: the firm keeps 60% of the surplus and the employee keeps 40%.

Figure 2.8: Optimal Investment and Pension Funding Decisions Over Time



This figure illustrates pension investment (w) and pension contribution (h) decisions across time and states (funding ratio before cash contribution). Panel A shows the optimal allocations and cash contributions under the benchmark parameters, where the bankruptcy probability is calibrated to that of a BBB-rated firm and the firm keeps 80% of pension surplus. Panel B shows the optimal allocations and cash contributions under a lower probability of bankruptcy (for an AAA-rated firm). Panel C shows the optimal allocations and cash contributions when the firm keeps a lower (60%) fraction of the pension surplus.

To varying degrees, all the plots share some common features. In the early years, that is, when the retirement time (T) is far into the future, the risky allocation is relatively high. As the retirement time draws closer, for most range of the funding ratio, the risky allocation drops. The intuition in Proposition 2.3.1 can help us understand this pattern, but there is also important differences due to the added richness of the dynamic model, as we explain below.

Unlike the one-period model where the asset allocation decision is made only once, in the dynamic model, asset allocation varies over time in response to funding level. The two main counterbalancing factors identified in Proposition 2.3.1 for optimal asset allocation are: (1) the employee's desire to have access to the market risk premium; and (2) the firm's financing cost. In the dynamic model, the firm chooses to have high risk exposure at times and states that such investment risk exposure can be effectively passed on to the employee and that the cost of such risk exposure (due to the convex financing cost) is relatively low.

Along the time dimension, the firm would choose to have high risk exposure when the retirement time is far away. A long investment horizon means that the optimal risk exposure has a bigger impact on the employee's utility, ceteris paribus. Furthermore, when the time horizon is long, the firm has the time to make the periodic contributions, and this would substantially reduces the impact of the convexity in the financing cost. Consider this simple thought experiment: there is a shortfall of \$10. If the firm has to make a one time contribution, the convex component of the financing cost is proportion to $10^2 = 100$. In contrast, if the firm breaks down the contribution into 10 equal pieces. Then the convex component of the financing cost adds up to an amount proportion to only $10 \times 1^2 = 10$, which is only one tenth of the one time contribution cost. In addition, the firm has the option to choose whether and when to make the contributions, as well as to adjustment portfolio allocations when needed. Such option has a higher value when the uncertainty is higher, that is, when the pension's risk allocation is high. All this point to the direction that the pension plan should have a high risk allocation early on, ceteris paribus.

In addition to the time dimension, when consider the state dimension of the time-state space, the firm should also allocation the pension fund's risk exposure across the state in an optimal fashion. A number of factors influence the decision. One can enhance the overall risk exposure, while keeping the average risk exposure across the state space fixed, by take up more risky positions in the up and down sides of the funding status. This idea is simple, as we make both ends of the state-space more volatile, the total volatility increases. Therefore, high risk allocations in the both the highly over-funding and the highly under-funding states, which make the pension asset more risky, help to expose the employee to the stock market risk.

Second, from a cost-effective stand point, when the under-funding is severe, the ex ante average delta of the pension payment to the pension asset approaches 1, the upper bound of the delta in the model. Therefore, it is the most effective in this range to expose the employee to the investment risk. In the middle range, when the pension plan is close to the fully funded or slightly over-funded, given the surplus sharing, the investment risk can still be passed to the employee, though at a lower delta. Yet, the convex cost is still a concern. The combination result is that the pension fund's risk allocation is relatively low in this range. Yet on the other end, when the pension plan is substantially over-funded, the decrease in the average delta reaches a plateau, and yet the financing cost become a lesser concern because the likelihood of future fund shortfall is low. The joint effect is that the risk allocation is high.

The patterns we observe from asset allocations across firms with different default probabilities help us gain further intuition on the risk allocation problem in the dynamic setting. When we reduce the default probability from that of a BBB firm in our benchmark case in Panel A to a AAA firm in Panel B, we see that overall the risk allocation is higher. This is consistent with what we find in Panel A of Figure 2.1. That is, an AAA firm has a low pension payment to pension asset delta than a BBB firm, and therefore, the AAA firm should choose a higher risk allocation for the pension investment to make up the reduced delta. On surface, this statement seems to contradict what we stated in the paragraph above, that it is more effective for the firm to choose high investment risk allocation in the states where the payment delta is relative high, ceteris paribus. The key difference is that, when we are considering the dynamic problem for a single firm, there is a internal balancing across states and times. The firm decides to allocate more to the risky asset in some states and times to make up the low allocation in some other states and times. In contrast, there is no such balancing effect across firms, that is, the two firms are not dealing with each other to reach a mutually beneficial agreement to jointly solve the the pension optimization problem. Furthermore, we focus only on the comparison of the average allocation across firms.

As we reduce the firm's share of the surplus, we see the allocation varies from the charts of Panel A to the charts in C. The main result is that the allocation to the risky asset is reduced as the firm's share of the surplus reduces. This pattern is consistent with what we observe in Panel B of Figure 2.2 for the one-period model. As the firm's share of the surplus increases, the employee's share naturally decreases. Consequently, the pension payment's delta to the pension's investment value is lowered. To provide the desirable level of the risk-return trade-off, in other words, to provide the desirable exposure to the investment risk for the employee, the asset allocation has to increase accordingly to compensate for the decline in delta.

2.6 Conclusions

This paper provides a perspective to understand the risky asset allocation policies pursued by corporate pensions. In our model, pension risk taking is driven by employees' preference for systematic risk exposure, while the firm balances employees' preference with its concern of financing cost. The pension investment risk is shared between shareholders and employees. The firm's decisions on pension benefits and pension funding are endogenous to such risk sharing. For a reasonable set of parameter values, the optimal pension investment risk and its relations with a firms' bankruptcy probability and pension funding ratio predicted by the model are consistent with empirical observations.

The stakeholder approach we take combines two polarized views on the objective function of pension decisions – the shareholder value maximization view and the employee utility maximization view. The former has been a prevalent approach in recent academic studies to evaluate corporate pension decisions. The challenge for this view is the extreme implication that pension investment should be riskfree, which contrasts dramatically with the observed pension investment behavior. The employee utility maximization view is recognized by Bodie (1990) as an alternative interpretation of the risky pension investment behavior. Perhaps implicitly, under this view a DB plan should invest in a way similar to what employees were to manage their portfolios outside a DB plan – a mean-variance efficient portfolio, for example. However, this does not take into account the convoluted cash flow rights (in particular, the guarantee by the firm on pension benefits) a DB plan offers to employees, which may substantially alter the investment decisions. The stakeholder approach successfully combines the two views to deliver a reasonable interpretation of the observed investment policies by corporate pensions.

Our analysis highlights the inefficient risk sharing in typical DB plans and its important consequences. The defined benefit plans, out of a motive to protect employees from firm-specific risk, make employees' pension payoffs relatively insulated from systematic risk, despite the preference by employees for systematic risk. Therefore, defined benefit pension contracts are suboptimal in allocating the systematic component of pension investment risk between shareholders and employees. A more efficient contract would let employees to shoulder all the pension investment risk while keeping them insulated from firm-specific risks. Interestingly, this arrangement resembles what a defined contribution plan offers (although DC plans lack the longevity risk sharing feature of DB plans). Our analysis shows that such an arrangement may substantially reduce firms' pension funding costs. Perhaps this is one of the reasons that many firms have switched from DB plans to DC plans.

CHAPTER 3
SHADOW BANKING AND RISK SHARING

3.1 Introduction

Financial intermediaries improve resource allocation and market completeness in the economy by providing extra financing channels among agents with different endowments, preferences, and productivities. History shows that many financial crises were attributed to the failure of financial intermediaries such as the 1930s Bank Runs crisis and the 2007-2009 sub-prime crisis. With the implementation of deposit insurance and discount window, regulated banks now can provide safe claims to households, which efficiently prevent bank runs. However, the "safety" of regulated banks comes with costs. The capital requirements and other restrictions make regulated banks less appealing to investors than the other form of banking: shadow banking.

Shadow banking refers to a series of securitization and secured funding techniques such as money market mutual funds (MMMFs), collateralized debt obligations (CDOs) and repurchase agreements (repos) (Pozsar et al. (2010)). It provides safe claims to investors through collateralizing the deposits with liquid assets or securitized illiquid assets. Once borrowers fail to repay the loan and contracted interest, lenders are entitled to liquidate borrowers' asset for repayment. On one hand, shadow banking is more efficient in terms of credit and liquidity transformation because of its floating, customizable interest rates, and less restricted terms. On the other hand, compared to regulated banking, shadow banking is more fragile under unanticipated negative shocks. As argued by Gennaioli et al. (2013) and

Martin et al. (2014), the collapse of shadow banking played a critical role in the formation of the 2007-2009 financial crisis.

In this paper, I propose a new model in which shadow banking provides additional risk sharing between depositors and borrowers through collateralized lending. In my model, following the seminal work of Diamond and Dybvig (1983), each agent faces an idiosyncratic preference shock. I extend the static three-period model into a dynamic equilibrium model with overlapping generations as in Qi (1994).[1] The regulated banking provides intragenerational risk sharing by offering a predetermined interest rate. Once preferences are revealed, aggressive agents can lever up their security investments by borrowing from the new generation whose preferences are still uncertain. The lending contract is collateralized with aggressive agents' financial assets. This intergenerational funding channel is defined as shadow banking because (1) it connects the borrowers who demand funding and depositors whose saving demand is not satisfied by regulated banking, and (2) the interest rate is endogenized and unregulated. Since borrowing through regulated banking usually requires physical assets as collateral, aggressive agents who invest in financial assets are unable to lever up their position through regulated banking.

The model solution shows two important features of shadow banking: procyclicality and path-dependence. When the economy is prosperous, the high expected return in financial market encourages levered investment and the lower interest rate in regulated banking drives the savers to transfer deposits from regulated banking account to shadow

[1] One "generation" in this paper can be very short - overnight when agents make their deposits and during the day when they withdraw and consume, or invest and consume later.

banking account, which both increases shadow banking volume. Another phenomenon associated with a booming economy is high market participation. If more people turn to invest the saving in financial market instead of consuming it right away, this preference shock also drives up the volume of shadow banking activity. The path-dependence of shadow banking is reflected in the impact of shadow banking gains of last generation on the shadow banking activities of present generation. Overactive shadow banking in previous generation "cools down" the shadow banking activities in current generation, which results from increased collateral and decreased interest rate.

The results also show that shadow banking improves the welfare of both aggressive and conservative agents. As creditors both kinds of agents have the chance to receive higher interest returns by lending through shadow banking, and aggressive agents enjoy the profit from leverage. Shadow banking, however, adds instability to the economy by spreading capital market risk from aggressive agents to conservative agents. Once a negative shock is realized in the investment, aggressive agents default on the shadow banking debt. The liquidation value of the collateral can be even lower than the regulated banking yield, which makes creditors of the shadow banking system worse off.

There has been some emerging literature in shadow banking research. Most of the literature attempts to describe the shadow banking system, for example, Pozsar et al. (2010), Gorton et al. (2010), Gorton and Metrick (2012), and Acharya et al. (2013). The authors document the growth of shadow banking in the United States, describe the functions of all forms of shadow banking, and discuss the reason behind the growth of shadow banking. Gorton et al. (2010) argues that the growth of shadow banking is attributed to both supply and demand sides. Financial innovations allow intermediaries to create money-like instruments to overcome the comparative advantage of regulated banks, and the increasing financial transactions require the development of securitization and collateralization. All the papers believe the collapse of shadow banking is central to the most recent financial crisis.

This paper contributes to the class of papers which focuses on modeling the mechanism behind shadow banking. In contrast with Gennaioli et al. (2013), Hanson et al. (2015), and Sunderam (2014), all of which model shadow banking with static equilibrium, my model captures the dynamic feature of the development of shadow banking activities. Inspired by Hanson et al. (2015), shadow banking in my model creates safe claims and stable interest incomes when adverse shock is moderate, however, once the negative shock is large enough the default option of shadow banking debts allows borrowers of shadow banking to share the loss with its lenders. Among the dynamic models of shadow banking, Moreira and Savov (2014) use a jump process of a latent crash variable to derive the liquidity crisis in an economy with shadow banking. My approach to model the market crash is simpler: investors' dramatic loss during market crash is a consequence of deep leverage through shadow banking.

The model explored in this paper is also related to the literature that researches financial frictions. Benanke and Gertler (1989) and Bernanke et al. (1999) captures the amplification and propagation effect of friction assuming the convex adjustment cost. Bernanke et al. (1991) describes the process of credit crunch and its macroeconomic impact. The collateralization setting in my model is inspired by Kiyotaki and Moore (1997), who consider the collateral value of assets as the primary limit of external funding; and Shleifer and Vishny (1992) who argue that the liquidation value of collateral affects corporate debt capacity. Similar to Brunnermeier and Pedersen (2009) who show that the restriction of collateralized lending is linked to the volatility of the collateral assets, the volatility of risky investment in my model determines investors' expectations of a default in shadow banking and affect the shadow banking activities. Other research related to this paper includes Curdia and Woodford (2010), Adrian et al. (2012), and Brunnermeier et al. (2012).

Lastly, my paper adds to the literature on financial intermediaries. In recent years, the focus of research on financial crisis has shifted to financial intermediaries, such as Holmstrom and Tirole (1997), Gertler and Karadi (2011), Duffie and Strulovici (2012), He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014), Adrian and Boyarchenko (2012), Muir (2016), and Rampini and Viswanathan (2017). They model financial intermediaries as a special agent between households and entrepreneurs, and intermediaries have their own risk preference and consumption optimization problem. Here I do not explicitly model shadow banking as a stand-alone agent. Instead, I distinguish regulated banking and shadow banking as two financing channels. This definition simplifies the analysis and helps separate the roles of regulated banking and shadow banking as two types of intermediaries. Similar to the financial intermediary literature, my model provides an explanation to the financial crisis.

The rest of the paper is organized as follows. In Section 2, I introduce the model setup. In Section 3, I formalize the model solution in equilibrium and discuss the numerical results. In Section 4, I show the results from simulation. Section 5 concludes.

3.2 Model

In the model, there is an infinite number of overlapping generations of agents in the economy. Each generation lives three periods (T=0, 1, 2). There is a continuum of agents in every generation. When agents are born in period 0 with endowment of 1 unit, they are uncertain about their preference: aggressive or conservative. Each agent learns its privately observed preference in period 1 when they establish their utility function:

$$U(C_1, C_2) = \begin{cases} C_1, & \text{if conservative, with probability } e \\ \\ C_2, & \text{if aggressive, with probability } 1 - e \end{cases}$$

where C_i stands for agents' consumption at period i. Agents' preference in each generation is exogenously determined: e of one generation is conservative and $1 - e$ is aggressive. e may change over generations and is unknown to agents in period 0.

Agents' risk aversion changes over time as well. All agents in period 0 (uncertain agents) are risk-neutral. Conservative agents in period 1 are infinitely risk-averse so they stop investing or saving and consume everything. Aggressive agents in period 1 are risk neutral and will start to invest. To match reality, the uncertain agents in the model represent households who receive a fortune the first time and need time to figure out what to do with the fortune. After they figure out whether they are conservative or aggressive, the conservative households withdraw and consume the fortune immediately, and the aggressive households become active investors in financial market.

Investment bears aggregate risk in the capital market. For each time t, investment yields return $\tilde{R}_t$, where $\tilde{R}_t \sim N(\mu_t, \sigma_t^2)$. $R_t's$ are assumed to be independent across time, and μ_t and σ_t are exogenously determined.

In period 0, limited market participation is assumed as in Diamond (1997). Uncertain agents will either deposit the endowment in regulated banks, or lend to aggressive agents of previous generation through shadow banking. The rationale is that cash management decision is a trade-off between regulated banking and shadow banking, which are essentially different from equity or bond investment in terms of risk and liquidity. If one generation refers to one day, the generation starts at 3:00 PM when the market is closed, and investor can only choose to keep the cash either in regulated banking account or in shadow banking account such as purchasing overnight repo agreement.

Therefore, uncertain agents in generation n deposit w_n of their endowments in regulated bank and lend $1 - w_n$ to the aggressive agents of generation $n - 1$ through shadow banking. The deposits in regulated banks earn a fixed interest of R_b. The interest serving as the indemnity of deposit insurance is paid by the aggressive agents to the conservative agents within the same generation, which is related to the argument in Diamond and Dybvig (1983). I assume a fixed R_b because regulated banking interest rate is closely related to federal funds rate which is usually much less volatile than shadow banking interest rate such as repo rate.

The shadow banking lending has a contracted interest of $R_{s,n}$. However, if the aggressive agents in previous generation default on the shadow banking debts, agents in generation n are entitled to liquidate the aggressive agents' assets and have a return of

$R'_{s,n}$. Hence the payoff for agents in generation n at period 1 is

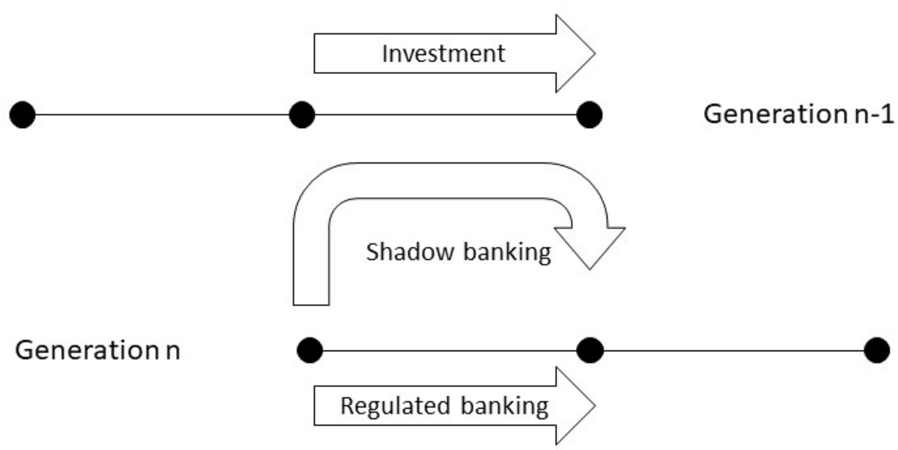
$$V_{n,1} = \begin{cases} w_n R_b + (1 - w_n) R_{s,n}, & \text{if shadow banking debt is not defaulted} \\ w_n R_b + R'_{s,n}, & \text{if shadow banking debt is defaulted} \end{cases}$$

and the optimization problem for uncertain agents in generation n at period 0 is

$$\max_{w_n, R_{s,n}} E(V_{n,1}) = \max_{w_n, R_{s,n}} w_n R_b + E\left[(1 - w_n) R_{s,n} | 1 - D\right] + E\left[R'_{s,n} | D\right]$$

The plot below illustrate the cash flows within generation and between generations.

Figure 3.1: Fund Flows of Regulated Banking and Shadow Banking



This figure illustrates the fund flows within and between generations through regulated banking and shadow banking proposed in the model.

In period 1, $(1-e_n)$ agents turn to be aggressive and borrow $\frac{1-w_{n+1}}{1-e_n}$ from generation $n+1$ through shadow banking. So the asset in place for aggressive agents in period 1 is

$$A_n = w_n \frac{1 - R_b e_n}{1 - e_n} + \tilde{R}_{s,n} + \frac{1 - w_{n+1}}{1 - e_n} \tag{3.1}$$

where $\tilde{R}_{s,n} = (1 - w_n)R_{s,n}$ if shadow banking debt is not defaulted, $\tilde{R}_{s,n} = R'_{s,n}$ if shadow banking debt is defaulted.

The first term on the right hand side of (3.1) is $w_n \frac{1-R_b e_n}{1-e_n}$ instead of $w_n R_b$ because, as argued in Diamond and Dybvig (1983), regulated banking transforms the credit between different agents in the future, and R_b becomes the interest aggressive agents pay to conservative agents once preference is revealed. Also, if aggressive agents know their preference in period 0 they would have invested in the risky assets, so $w_n R_b - w_n \frac{1-R_b e_n}{1-e_n}$ reflects aggressive agents' opportunity cost brought by the preference shock.

Assuming the aggressive agents face quadratic adjustment cost for investment, the payoff in period 2 is

$$V_{n,2} = \max \left[A_n R_n - A_n^2 c - \frac{(1 - w_{n+1})R_{s,n+1}}{1 - e_n}, 0 \right]$$

where R_n is the realized return of the investment for generation n in period 2. c is the adjustment cost coefficient. The adjustment cost can be seen as the fee of managing assets in place which prevents aggressive agents levering up without limits. $R_{s,n+1}$ is the shadow banking interest contracted between generation n and $n+1$. The optimization problem for aggressive agents in generation n at period 1 is

$$\max_{w_{n+1}, R_{s,n+1}} E(V_{n,2}) = \max_{w_{n+1}, R_{s,n+1}} E \left[A_n R_n - A_n^2 c - \frac{(1 - w_{n+1})R_{s,n+1}}{1 - e_n} | 1 - D \right]$$

3.3 Equilibrium

Since uncertain agents in generation n can observe the historical mean and volatility of the investment's return for generation $n - 1$, they perceive the probability of shadow banking being defaulted as

$$Prob(\text{default}) = F\left(R_{n-1} \leq \frac{A_{n-1}^2 c + \frac{(1-w_n)R_{s,n}}{1-e_{n-1}}}{A_{n-1}}\right)$$

where $F(\cdot)$ is the c.d.f of normal distribution. Then the optimization problem for uncertain agents of generation n in period 0 can be rewritten as

$$\begin{aligned}
&\max_{w_n, R_{s,n}} w_n R_b + \int_{d_{n-1}}^{\infty} (1-w_n)R_{s,n} f(R_{n-1})dR_{n-1} + \int_{-\infty}^{d_{n-1}} R'_{s,n} f(R_{n-1})dR_{n-1} \\
&\text{s.t. } d_{n-1} = A_{n-1}c + \frac{(1-w_n)R_{s,n}}{(1-e_{n-1})A_{n-1}} \\
&\qquad R'_{s,n} = (A_{n-1}R_{t-1} - A_{n-1}^2 c)(1-e_{n-1})
\end{aligned}$$

where d_{n-1} is the critical return of the investment that makes aggressive agents of generation $n-1$ indifferent between default or not default. $R'_{s,n}$ is the liquidation return once default happens. The p.d.f $f(R_{n-1})$ is identified by the mean μ_{n-1} and standard deviation σ_{n-1}.

As the other side of the borrowing/lending relation, the aggressive agents also decide w_n and $R_{s,n}$ to maximize their expected utility as written out below:

$$\max_{w_n, R_{s,n}} \int_{d_{n-1}}^{\infty} A_{n-1}R_{n-1} - A_{n-1}^2 c - \frac{(1-w_n)R_{s,n}}{1-e_{n-1}} f(R_{n-1})dR_{n-1}$$

Thus we can solve for w_n and $R_{s,n}$ by maximizing borrowers' and lenders' expected utility simultaneously.

w_n^* and $R_{s,n}^*$ in equilibrium are solved numerically given the following exogenously valued parameters: (1) expected risky return μ_{n-1}, (2) observed market volatility σ_{n-1}, (3) regulated banking yield R_b, (4) generation $n-1$'s shadow banking activities $1-w_{n-1}$, (5) generation $n-1$'s yield from shadow banking $\tilde{R}_{s,n-1}$, (6) percentage of generation $n-1$ as aggressive agents $1-e_{n-1}$, and (7) adjustment cost c. The baseline parameter values are listed in the following table.

Table 3.3.1: Baseline Parameters

μ_{n-1}	σ_{n-1}	R_b	$1-w_{n-1}$	$\tilde{R}_{s,n-1}$	$1-e_{n-1}$	c
1.1	0.18	1.02	0.5	0.5105	0.5	0.02

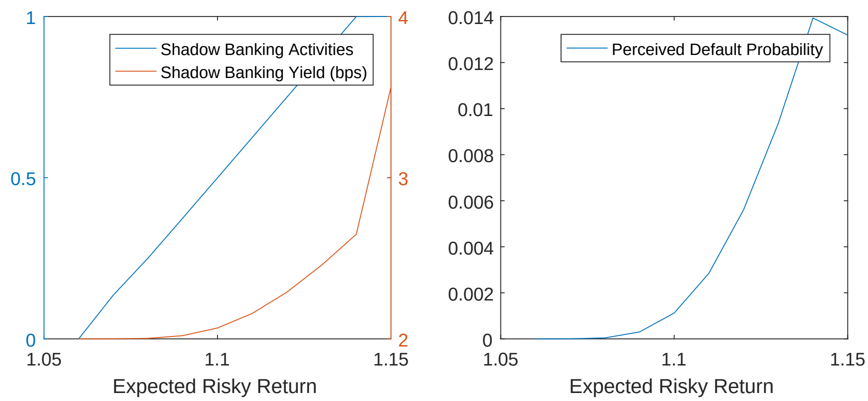
The distribution of risky return is calibrated using S&P 500 index from 1926 to 2016. The generation $n-1$'s yield from shadow banking is calculated assuming 50% shadow banking weight and 2.1% interest without defaults.

3.3.1 Shadow Banking and Expected Risky Return

When aggressive agents have a higher expectation about risky investment, they have an incentive to borrow more through shadow banking and offer relatively higher interest rate at the same time. According to figure 3.2, shadow banking activities $1-w_n$ is much more sensitive to the change of expected risky return μ_{n-1} than the shadow banking yield $R_{s,n}$ is. On one hand, with an increase of about 8% in expected risky return, uncertain agents in period 0 change from not funding through shadow banking at all to lending entire endowments through shadow banking. Shadow banking yield, on the other hand, changes marginally with the expected risky return.

The result is consistent with reality: when active investors observe a highly prof-

Figure 3.2: Shadow Banking and Expected Risky Return

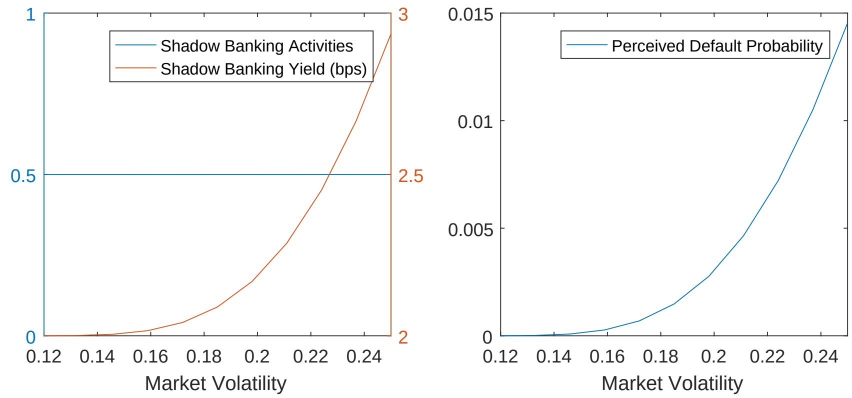


The left plot shows the change of shadow banking volume and yield due to the change of the expected risky return. The right plot shows how the perceived defaulting probability changes with the expected risky return.

itable investment opportunity, they tend to increase their leverage, and households who are optimistic about the underlying asset are willing to purchase more asset-backed securities. It is worth noting that even if no one wants to lend through shadow banking, $1 - w_n = 0$, there still exists a shadow banking interest which merely reflects the hypothetical marginal interest that aggressive agents would like to pay.

It will help to understand the increase of the shadow banking yield by investigating the impact of a higher expected return in capital market on lender's perceived probability of defaulting. As shown in figure 3.2, when borrowers and lenders are optimistic about the future return from the capital market, which drives up the demand and supply of shadow banking, the likelihood of borrowers defaulting on the "shadow debts" increases as well. The slight drop of defaulting probability occurs when the shadow lending is capped at 100%, in which case the equilibrium is facilitated through lifting shadow banking yield

Figure 3.3: Shadow Banking and Capital Market Volatility



The left plot shows the change of shadow banking volume and yield due to the change of the capital market volatility. The right plot shows how the perceived defaulting probability changes with the capital market volatility.

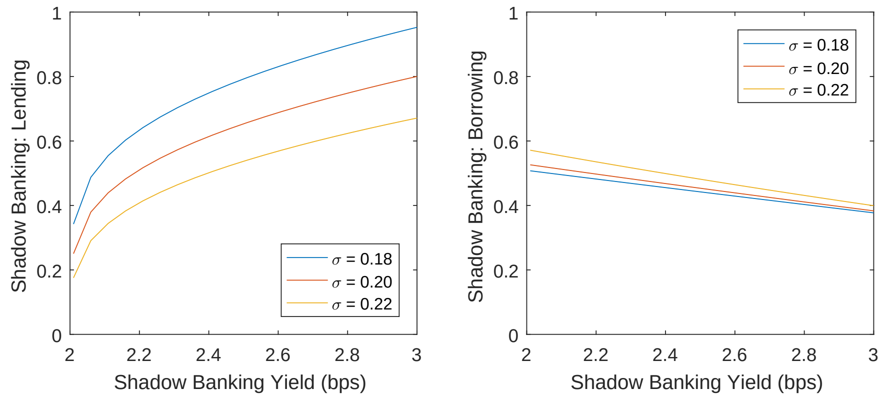
from both sides.

3.3.2 Shadow Banking and Market Volatility

Figure 3.3 shows that the level of shadow banking activeness is somewhat insensitive to capital market volatility. This result appears to be surprising yet consistent with the empirical observation. Shadow banking volume has been growing gradually over the past decades regardless of the up-and-down of the capital market except for the 2009 Financial Crisis, in which case the shadow banking collapses. The seemingly counter-intuitive finding implies one characteristic of the shadow banking system - the banking channel is quite robust against the volatile capital market.

As shown in the bottom of Figure 3.3, the higher market volatility increases the perceived defaulting probability because of the put option feature of the shadow banking

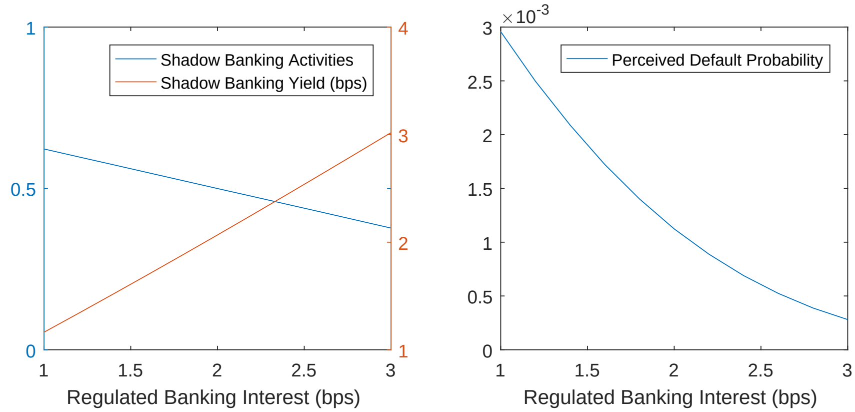
Figure 3.4: Demand and Supply Curves of Shadow Banking under Varying Capital Market Volatilities



The left plot shows the desired level of shadow lending corresponding to different shadow banking yield under varying capital market volatilities. The right plot shows the desired level of shadow borrowing corresponding to different shadow banking yield under varying capital market volatilities.

contract. The reduction of the lending, however, is canceled off by the attraction of the higher shadow banking yield the borrowers are willing to pay. To further investigate how the demand and supply of shadow banking are impacted by the market volatility, I show the demand and supply curves separately with different capital market volatilities in Figure 3.4. It is clear that the supply curve of shadow banking loans is much more sensitive to the change of market volatility compared to the demand curve, which implies that when the market becomes more volatile, lenders need to be compensated with a higher interest to make the same amount of deposit through shadow banking. The borrowers, on the other hand, have a much more stable willingness to borrow facing a change of market volatility. It is because the benefit of exploiting lenders via defaulting on shadow banking debt is second-order for the risk-neutral borrowers.

Figure 3.5: Shadow Banking and Regulated Banking



The left plot shows the change of shadow banking volume and yield due to the change of the regulated banking interest. The right plot shows how the perceived defaulting probability changes with the regulated banking interest.

3.3.3 Shadow Banking and Regulated Banking

As discussed in Hanson et al. (2015), the coexistence of regulated banks and shadow banks poses a new challenge to the government. Fire-sale is more likely to occur if shadow banks hold more risky asset. My model captures the incentive of shadow banking from the perspective of regulated banking. Higher regulated banking interest which is essentially risk-free due to deposit insurance and government guarantees makes the shadow banking which suffers default risk less attractive to the investors. Aggressive agents have to raise the shadow banking interest in order to borrow through shadow banking under increasing regulated banking interest.

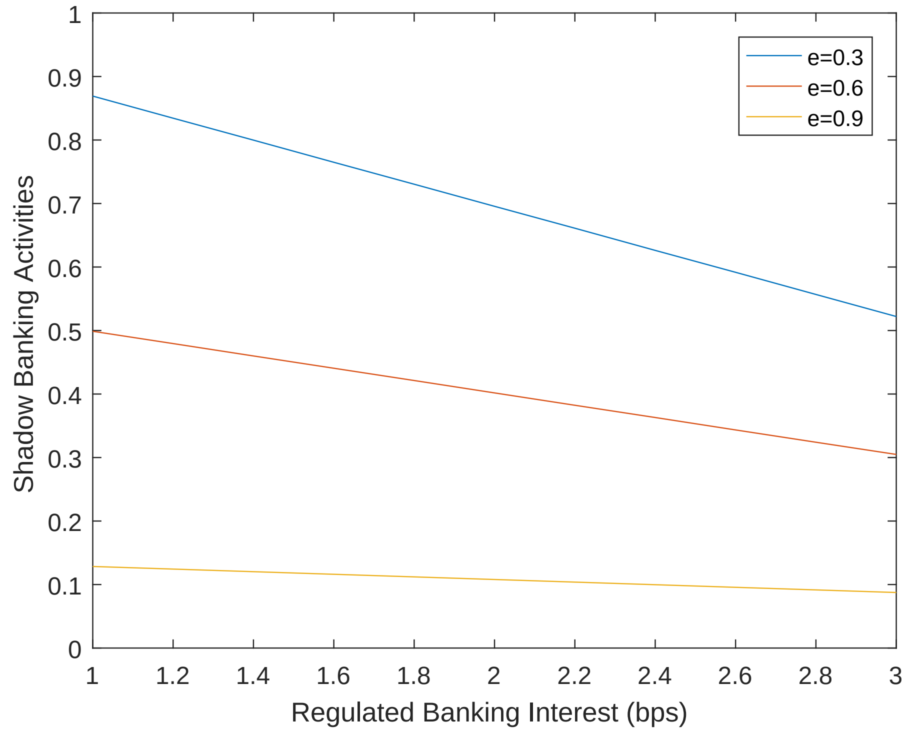
In practice, the regulated banking interest rate is often adjusted by government as an policy tool. When the economy shows a hint of slowing down, government lowers the interest rate to encourage investment. However, as shown in my model, the decrease of regulated banking interest drives the savers to have deeper position in shadow banking which is exposed to aggregate risk. It makes the savers more vulnerable when the market crashes. This implication is consistent with the federal funds rate downward adjustments before 2001 and 2008 market crashes. Lower "safe" interest rate leads to active investors' irrational enthusiasm about investment and leverage deeper through shadow banking. Meanwhile, as shown in the bottom of Figure 3.5, lenders perceive a higher probability of defaulting occurrence. When the bubble actually burst, a series of defaults hurt the creditors in the shadow banking.

The bubble in shadow banking not only depends the regulated banking interest, but also depends on the percentage of aggressive agents in previous generation. Figure 3.6 shows the relationships of regulated banking interest and shadow banking activities with different percentages of aggressive agents in previous generation. It illustrates that when e_{n-1} is lower – meaning there are higher number of aggressive agents – one unit decrease of regulated banking interest encourages more shadow banking activities.

3.3.4 Shadow Banking between Generations

My model demonstrates the cyclical feature of shadow banking activities over generations as shown in Figure 3.7. If the previous generation gains more wealth in period 0, aggressive agents of this generation in period 1 will have more collateral given no default in period 1. More collateral of debtors lowers the shadow banking rate that creditors require to compensate the possibility of default since debtors suffer a larger loss if defaulting at this time. The downward sloping curve of shadow banking yield illustrates this point.

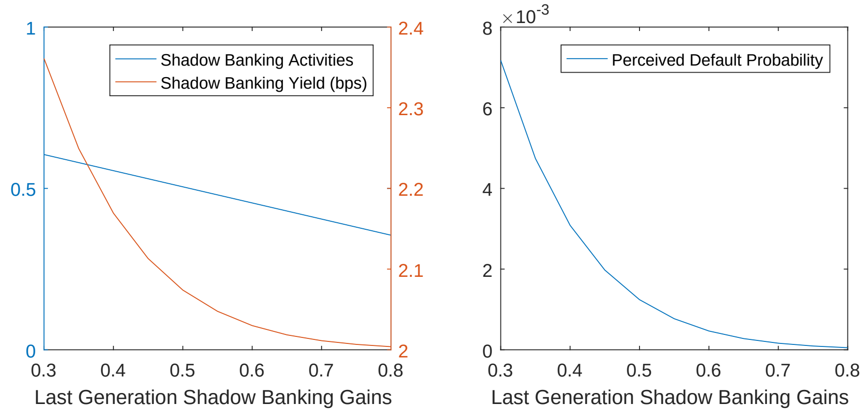
Figure 3.6: The Relationship of Shadow Banking Activities and Regulated Banking Interest with Varying Preference Shocks



This plot illustrates the relationships of regulated banking interest and shadow banking activities with different percentages of aggressive agents in previous generation.

Ceteris paribus, an overactive shadow banking in previous generation "cools down" the shadow banking activities in current generation, illustrated by Figure 3.7. This relationship suggests that shadow banking goes through a fluctuating process at beginning phase and will eventually become stable. In the long run, shadow banking will not fully replace the role of regulated banking.

Figure 3.7: Shadow Banking and Last Generation's Wealth Gain



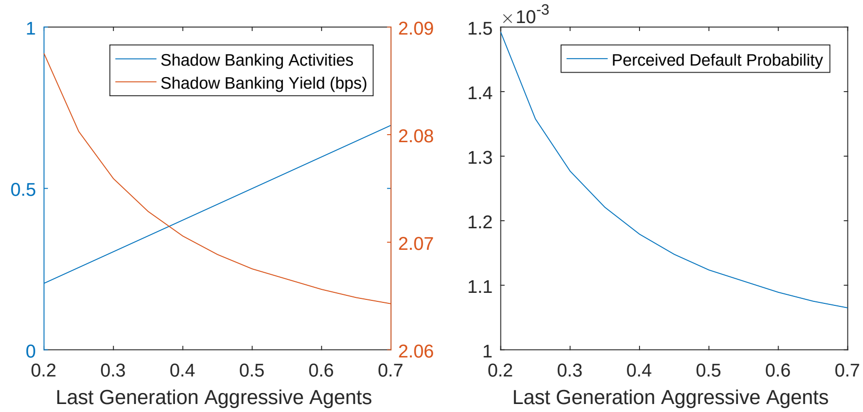
The left plot shows the change of shadow banking volume and yield due to the change of the last generation's wealth gain from shadow banking. The right plot shows how the perceived defaulting probability changes with the last generation's wealth gain from shadow banking.

3.3.5 Shadow Banking and Preference Shock

Since shadow banking is primarily initiated by aggressive agents who want to leverage their investment, the shadow banking activities are positively associated with the percentage of aggressive agents in their generation. In practice, the number of active investors depends on many factors – culture, economy, war, technology, etc.. For example, a predominant trend of entrepreneurship in the society may turn more people into aggressive investors, which increases the demand of shadow banking.

With very few aggressive agents, the shadow banking interest is high because there is little aggregate collateral. When there are more aggressive agents with more assets as collateral, the perceived defaulting probability drops, and the shadow banking interest sharply shrinks. It is consistent with people's fear of default towards the start-ups. But when the group of start-ups becomes larger forming the scale effect, like the IT start-ups in Silicon

Figure 3.8: Shadow Banking and Preference Shock



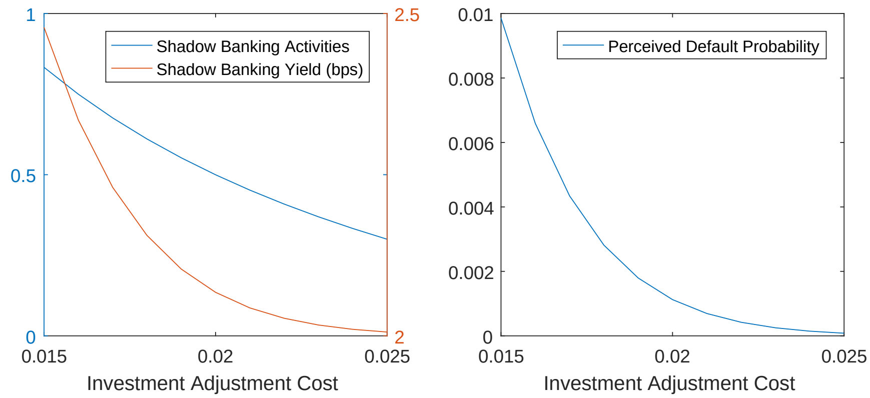
The left plot shows the change of shadow banking volume and yield due to the change of the percentage of aggressive agents revealed in the previous generation. The right plot shows how the perceived defaulting probability changes with the change of the percentage of aggressive agents revealed in the previous generation.

Valley, investors are more willing to fund even though they know the firms are very immature. Once the shadow banking debt supply is capped, additional demand drives up the shadow banking interest as illustrated in figure 3.8 when $1 - e_{n-1} > 0.7$.

3.3.6 Shadow Banking and Adjustment Cost

As shown in figure 3.9, shadow banking activities are very sensitive to the adjustment cost. Higher adjustment cost discourages aggressive agents' investment in the risky assets, which further reduces the demand of shadow banking debt. Also the higher adjustment cost makes aggressive agents more careful with selecting risky investment and lowers the chance of default. The lower chance of default reduces the interest required by the shadow banking creditors.

Figure 3.9: Shadow Banking and Investment Adjustment Cost



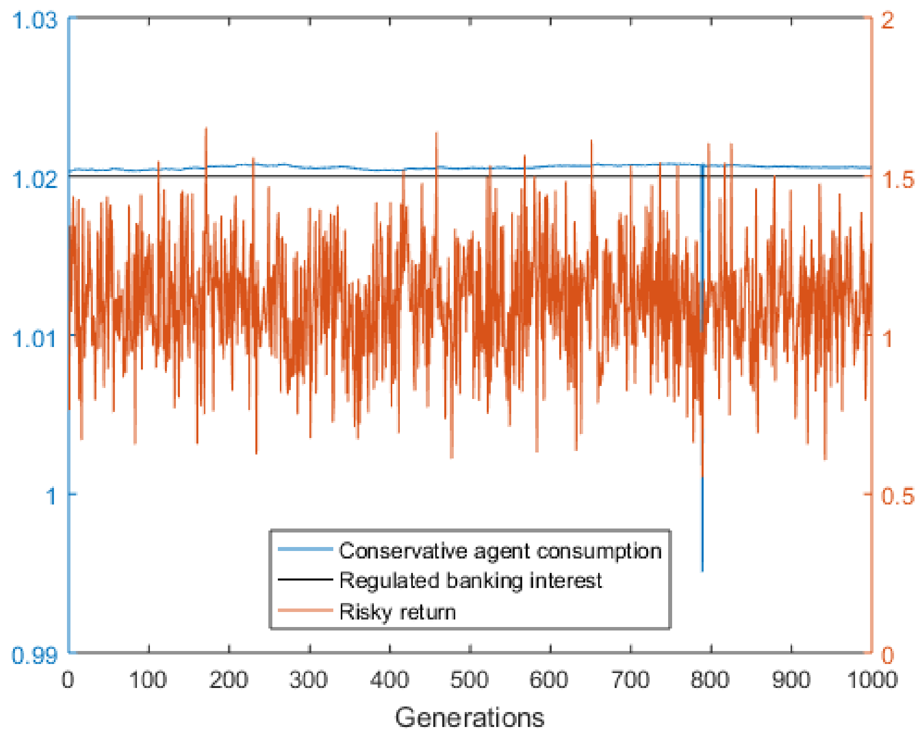
The left plot shows the change of shadow banking volume and yield due to the change of the investment adjustment cost. The right plot shows how the perceived defaulting probability changes with the change of the investment adjustment cost.

3.4 Simulation

To investigate the impact of shadow banking on both conservative agents and aggressive agents, I implement a simulation procedure. In period 1 of each generation, agents are split into two types: conservative agents who consume immediately and aggressive agents who invest in risky assets and consume in period 2. With regulated banking only, conservative agents always consume R_b because the interest payment is secured by the government, and aggressive agents consume $\frac{1-R_b e}{1-e}R$ where R is the realized investment return. With both regulated banking and shadow banking, agents' consumption depends on the state variables discussed above and the realized investment return.

If the realized risky return is high enough, aggressive agents do not default and conservative agents receive the contracted shadow banking yield. If realized risky return is low, aggressive agents default and lose everything while conservative agents consume the

Figure 3.10: Simulated Conservative Agent's Consumption Over Generations



liquidation value of aggressive agents' asset.

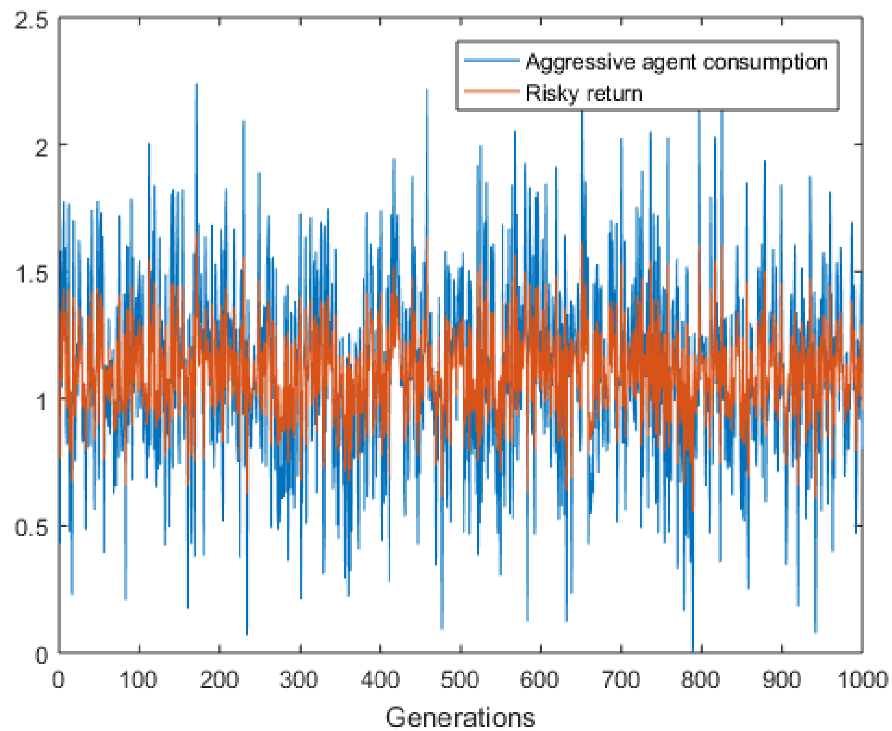
I let risky returns follow normal distribution with mean 1.12 and standard deviation 0.18. Since shadow banking activities and shadow banking yield are endogenously determined by the expected risky return and its volatility, I assume agents in the economy can only estimate the true mean and standard deviation by the historical average and sample standard deviation. As a result, the observed expected risky return and volatility vary over generations. The variability of the perceived expected risky return and volatility adds fluctuation to the simulated shadow banking yield.

Figure 3.10 shows the simulated conservative agents' consumption over generations. The red line indicates the constant regulated banking yield. The blue wavy line is conservative agents' consumption when shadow banking is available. As shown in the figure 3.10, conservative agents' consumption is higher with than without shadow banking for most generations. It suggests that shadow banking does improve conservative agents' welfare by risk sharing. However, once risky investment encounters large negative shock, conservative agents suffer dramatic loss. The scatter dramatic losses of conservative agents are caused by the aggregate risk brought in by the shadow banking channel. When a negative shock in risky investment is realized, shadow banking debts default and both agents share the loss.

Figure 3.11 illustrates the simulated aggressive agents' consumption over generations. As discussed in previous sections, shadow banking helps lever up aggressive agents' investment, so the gain and loss are both exaggerated. But sine the loss is capped at 0, the average consumption of aggressive agents is higher with than without shadow banking. Aggressive agents are risk neutral so shadow banking improves their welfare as well.

The simulated results suggest that shadow banking improves both agents' welfare, however, making the economy more instable. During 2007-2009 crisis, the unanticipated sudden drop of housing price made the households who invest passively and pursue the stable earning suffer the most. The risk-taking investors are voluntarily exposed to high risk so the financial crisis does not impact their utility that much.

Figure 3.11: Simulated Aggressive Agent's Consumption Over Generations



3.5 Conclusion

My model of shadow banking system answers two important questions: (1) what determines the level of shadow banking activities? (2) how does shadow banking influence the economy, especially the financial instability? For the first question, my model suggests that, in dynamic equilibrium, shadow banking is more active when investors expect higher return or less volatility from risky investment, lower regulated banking interest, higher demand of shadow banking debts from active investors, and lower adjustment cost of investments.

For the second question, the simulations show that shadow banking improves the welfare of both conservative agents and aggressive agents by providing extra risk sharing. Conservative agents consume more since they earn more from lending through shadow banking. Aggressive agents are risk neutral and better off with shadow banking because their gain from positive shock of risky investment is levered up while their loss is capped by defaulting on the shadow banking debts. However, when the economy experiences negative shock, conservative agents suffer dramatic loss which they will never have without shadow banking. The simulation result suggests that shadow banking makes the financial market more unstable and expose the savers who pursue stable interest income to the production risk.

APPENDIX A
APPENDIX TO CHAPTER 1

Proof of Proposition 1.3.1

We insert the optimal level of coverage without ambiguity into the consumer's first-order condition with ambiguity and determine the sign. We obtain that

$$\begin{aligned} V'(\alpha^0) = & \ \mathbb{E}\left\{\phi'\left[U(\alpha^0,\widetilde{\epsilon})\right]U_\alpha(\alpha^0,\widetilde{\epsilon})\right\} \\ = & \ \mathrm{Cov}\left\{\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right],U_\alpha(\alpha^0,\widetilde{\varepsilon})\right\}+\mathbb{E}\left\{\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right]\right\}\mathbb{E}\left\{U_\alpha(\alpha^0,\widetilde{\varepsilon})\right\} \end{aligned} \tag{A.1}$$

from the covariance rule. Expanding expected utility with ambiguity yields that

$$\begin{aligned} U_\alpha(\alpha^0,\widetilde{\varepsilon}) &= -(1-p)P'(\alpha^0)\cdot u'(W_1^0)+p(\overline{q}-\widetilde{\varepsilon})(L-P'(\alpha^0))\cdot u'(W_2^0) \\ & \quad -p(1-\overline{q}+\widetilde{\varepsilon})P'(\alpha^0)\cdot u'(W_3^0) \\ &= U_\alpha(\alpha^0,0)+\widetilde{\varepsilon}p\left[P'(\alpha^0)(u'(W_2^0)-u'(W_3^0))-Lu'(W_2^0)\right], \end{aligned} \tag{A.2}$$

where superscript 0 indicates final wealth levels when the level of coverage is given by α^0. Now $U_\alpha(\alpha^0,0)=0$ by definition of α^0 and $\mathbb{E}\widetilde{\varepsilon}=0$ by definition of $\widetilde{\varepsilon}$. As a result, we obtain that $\mathbb{E}\left\{U_\alpha(\alpha^0,\widetilde{\varepsilon})\right\}=0$ so that

$$V'(\alpha^0)=\mathrm{Cov}\left\{\phi'\left[U(\alpha^0,\widetilde{\varepsilon})\right],U_\alpha(\alpha^0,\widetilde{\varepsilon})\right\}. \tag{A.3}$$

To sign the covariance we investigate how $\phi'\left[U(\alpha^0,\varepsilon)\right]$ and $U_\alpha(\alpha^0,\varepsilon)$ vary in ε. The consumer's expected utility strictly decreases in the realization of $\widetilde{\varepsilon}$ because the higher ε, the higher the consumer's belief that the contract will not perform. Due to the concavity of ϕ, the term $\phi'\left[U(\alpha^0,\varepsilon)\right]$ is then strictly increasing in ε,

$$\frac{\partial\phi'\left[U(\alpha^0,\varepsilon)\right]}{\partial\varepsilon}=\phi''\left[U(\alpha^0,\varepsilon)\right]U_\varepsilon(\alpha^0,\varepsilon)=-\phi''\left[U(\alpha^0,\varepsilon)\right]p(u(W_2^0)-u(W_3^0))>0. \tag{A.4}$$

Per direct computation we obtain that

$$U_{\alpha\varepsilon}(\alpha^0, \varepsilon) = p\left[P'(\alpha^0)(u'(W_2^0) - u'(W_3^0)) - Lu'(W_2^0)\right]. \tag{A.5}$$

The square bracket is negative because marginal utility is positive and diminishing so that $U_\alpha(\alpha^0, \varepsilon)$ is decreasing in ε. Therefore, the covariance term in equation (A.3) is negative because its arguments are countermonotonic. Hence, $V'(\alpha^0) < 0$, and $\alpha^* < \alpha^0$ then follows from the concavity of the objective function.

Proof of Proposition 1.4.1

Following Klibanoff et al. (2005), we model an increase in the consumer's degree of ambiguity aversion by replacing his ambiguity function with ψ, where ψ is an increasing and concave transformation of ϕ, $\psi = k(\phi)$ with $k' > 0$ and $k'' < 0$. Let T denote the more ambiguity-averse consumer's objective function; his optimal level of coverage, α^{**}, is then characterized by the corresponding first-order condition,

$$T'(\alpha) = \mathbb{E}\left\{\psi'\left[U(\alpha, \widetilde{\varepsilon})\right] U_\alpha(\alpha, \widetilde{\varepsilon})\right\} = \mathbb{E}\left\{k'\left(\phi\left[U(\alpha, \widetilde{\varepsilon})\right]\right) \phi'\left[U(\alpha, \widetilde{\varepsilon})\right] U_\alpha(\alpha, \widetilde{\varepsilon})\right\} = 0. \tag{A.6}$$

The second-order condition is satisfied due to the concavity of ψ and u. To compare the optimal level of coverage of the less ambiguity-averse consumer (α^*) with that of the more ambiguity-averse consumer (α^{**}), we insert the former into the first-order condition for the latter and determine the sign. Notice that $k'\left(\phi\left[U(\alpha, \varepsilon)\right]\right)$ is increasing in ε,

$$\begin{aligned} \frac{\partial k'\left(\phi[U(\alpha, \varepsilon)]\right)}{\partial \varepsilon} &= k''\left(\phi[U(\alpha, \varepsilon)]\right) \phi'[U(\alpha, \varepsilon)] U_\varepsilon(\alpha, \varepsilon) \\ &= -k''\left(\phi[U(\alpha, \varepsilon)]\right) \phi'[U(\alpha, \varepsilon)] p(u(W_2) - u(W_3)) > 0. \end{aligned} \tag{A.7}$$

As a result, for any $\varepsilon \in (\underline{\varepsilon}, \overline{\varepsilon})$, it holds that

$$k'\left(\phi[U(\alpha, \underline{\varepsilon})]\right) < k'\left(\phi[U(\alpha, \varepsilon)]\right) < k'\left(\phi[U(\alpha, \overline{\epsilon})]\right). \tag{A.8}$$

The optimal level of coverage of the less ambiguity-averse consumer (α^*) is obtained from his first-order condition,

$$V'(\alpha^*) = \mathbb{E}\left\{\phi'\left[U(\alpha^*, \widetilde{\varepsilon})\right] U_\alpha(\alpha^*, \widetilde{\varepsilon})\right\} = \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \phi'\left[U(\alpha^*, \varepsilon)\right] U_\alpha(\alpha^*, \varepsilon) \mathrm{d}F(\varepsilon) = 0. \tag{A.9}$$

$\phi'\left[U(\alpha^*, \varepsilon)\right]$ is strictly positive for any $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$ and $U_\alpha(\alpha^*, \varepsilon)$ is strictly decreasing in ε by the proof of Proposition 1.3.1. For the integral to be zero, it follows that $U_\alpha(\alpha^*, \varepsilon)$ must change sign on $[\underline{\varepsilon}, \overline{\varepsilon}]$; due to strict monotonicity, this can only happen once. If $\hat{\varepsilon} \in (\underline{\varepsilon}, \overline{\varepsilon})$ denotes the null of $U_\alpha(\alpha^*, \varepsilon)$, we obtain that

$$\int_{\underline{\varepsilon}}^{\hat{\varepsilon}} \phi'\left[U(\alpha^*, \varepsilon)\right] U_\alpha(\alpha^*, \varepsilon) \mathrm{d}F(\varepsilon) > 0, \text{ and } \int_{\hat{\varepsilon}}^{\overline{\varepsilon}} \phi'\left[U(\alpha^*, \varepsilon)\right] U_\alpha(\alpha^*, \varepsilon) \mathrm{d}F(\varepsilon) < 0. \tag{A.10}$$

Combining this with (A.8) yields,

$$\begin{aligned}
T'(\alpha^*) &= \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} k'\left(\phi\left[U(\alpha^*, \varepsilon)\right]\right) \phi'\left[U(\alpha^*, \varepsilon)\right] U_\alpha(\alpha^*, \varepsilon) \mathrm{d}F(\varepsilon) \\
&= \int_{\underline{\varepsilon}}^{\hat{\varepsilon}} k'\left(\phi[U(\alpha^*, \varepsilon)]\right) \phi'\left[U(\alpha^*, \varepsilon)\right] U_\alpha(\alpha^*, \varepsilon) \mathrm{d}F(\varepsilon) \\
&\quad + \int_{\hat{\varepsilon}}^{\overline{\varepsilon}} k'\left(\phi[U(\alpha^*, \varepsilon)]\right) \phi'\left[U(\alpha^*, \varepsilon)\right] U_\alpha(\alpha^*, \varepsilon) \mathrm{d}F(\epsilon) \\
&< k'\left(\phi[U(\alpha^*, \hat{\varepsilon})]\right) \int_{\underline{\varepsilon}}^{\hat{\varepsilon}} \phi'\left[U(\alpha^*, \varepsilon)\right] U_\alpha(\alpha^*, \varepsilon) \mathrm{d}F(\varepsilon) \qquad (A.11) \\
&\quad + k'\left(\phi[U(\alpha^*, \hat{\varepsilon})]\right) \int_{\hat{\varepsilon}}^{\overline{\varepsilon}} \phi'\left[U(\alpha^*, \varepsilon)\right] U_\alpha(\alpha^*, \varepsilon) \mathrm{d}F(\epsilon) \\
&= k'\left(\phi[U(\alpha^*, \hat{\varepsilon})]\right) \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \phi'\left[U(\alpha^*, \varepsilon)\right] U_\alpha(\alpha^*, \varepsilon) \mathrm{d}F(\varepsilon) \\
&= k'\left(\phi[U(\alpha^*, \hat{\varepsilon})]\right) V'(\alpha^*) = 0,
\end{aligned}$$

so that $T'(\alpha^*)$ is strictly negative. As a result, it is optimal for the more ambiguity-averse consumer to decrease his level of insurance below α^* such that $\alpha^{**} < \alpha^*$.

Proof of Proposition 1.4.2

We define the consumer's conditional first-order expression evaluated at the optimal level of insurance coverage:

$$g(\varepsilon) = \phi' \left[U(\alpha^*, \varepsilon)\right] U_\alpha(\alpha^*, \varepsilon). \tag{A.12}$$

We recoup the consumer's first-order condition as $\mathbb{E}g(\widetilde{\varepsilon}) = 0$. As shown by Rothschild and Stiglitz (1970) and more generally by Ekern (1980), the sign of subsequent derivatives of $g(\varepsilon)$ informs about how changes in risk in the distribution of $\widetilde{\varepsilon}$ affect this expectation. We obtain the first and second derivative of $g(\varepsilon)$ as

$$g'(\varepsilon) = \phi'' \left[U(\alpha^*, \varepsilon)\right] U_\varepsilon(\alpha^*, \varepsilon) U_\alpha(\alpha^*, \varepsilon) + \phi' \left[U(\alpha^*, \varepsilon)\right] U_{\alpha\varepsilon}(\alpha^*, \varepsilon), \text{ and} \tag{A.13}$$

$$g''(\varepsilon) = \phi''' \left[U(\alpha^*, \varepsilon)\right] U_\varepsilon(\alpha^*, \varepsilon)^2 U_\alpha(\alpha^*, \varepsilon) + 2\phi'' \left[U(\alpha^*, \varepsilon)\right] U_\varepsilon(\alpha^*, \varepsilon) U_{\alpha\varepsilon}(\alpha^*, \varepsilon), \tag{A.14}$$

because $U_{\varepsilon\varepsilon}(\alpha^*, \varepsilon) = U_{\alpha\varepsilon\varepsilon}(\alpha^*, \varepsilon) = 0$. We can rewrite these derivatives as follows:

$$\begin{aligned} g'(\varepsilon) &= -\phi' \left[U(\alpha^*, \varepsilon)\right] U_{\alpha\varepsilon}(\alpha^*, \varepsilon) \left\{ -U(\alpha^*, \varepsilon) \frac{\phi'' \left[U(\alpha^*, \varepsilon)\right]}{\phi' \left[U(\alpha^*, \varepsilon)\right]} \cdot \frac{U_\varepsilon(\alpha^*, \varepsilon) U_\alpha(\alpha^*, \varepsilon)}{U(\alpha^*, \varepsilon) U_{\alpha\varepsilon}(\alpha^*, \varepsilon)} - 1 \right\}, \\ g''(\varepsilon) &= -\phi'' \left[U(\alpha^*, \varepsilon)\right] U_\varepsilon(\alpha^*, \varepsilon) U_{\alpha\varepsilon}(\alpha^*, \varepsilon) \\ &\quad \cdot \left\{ -U(\alpha^*, \varepsilon) \frac{\phi''' \left[U(\alpha^*, \varepsilon)\right]}{\phi'' \left[U(\alpha^*, \varepsilon)\right]} \cdot \frac{U_\varepsilon(\alpha^*, \varepsilon) U_\alpha(\alpha^*, \varepsilon)}{U(\alpha^*, \varepsilon) U_{\alpha\varepsilon}(\alpha^*, \varepsilon)} - 2 \right\}. \end{aligned} \tag{A.15}$$

Both the sign of $g'(\varepsilon)$ and of $g''(\varepsilon)$ are determined by the sign of the respective curly bracket. We define

$$h(\varepsilon) = \frac{U_\varepsilon(\alpha^*, \varepsilon) U_\alpha(\alpha^*, \varepsilon)}{U(\alpha^*, \varepsilon) U_{\alpha\varepsilon}(\alpha^*, \varepsilon)}, \tag{A.16}$$

which is positive for $\varepsilon < \hat{\varepsilon}$, zero for $\varepsilon = \hat{\varepsilon}$ and negative for $\varepsilon > \hat{\varepsilon}$. We can rewrite it as

$$h(\varepsilon) = \left(-\frac{U_\varepsilon(\alpha^*,\varepsilon)}{U(\alpha^*,\varepsilon)}\right) / \left(-\frac{U_{\alpha\varepsilon}(\alpha^*,\varepsilon)}{U_\alpha(\alpha^*,\varepsilon)}\right) \text{ for } \varepsilon \neq \hat{\varepsilon}; \tag{A.17}$$

as such, $h(\varepsilon)$ compares the decay rate of expected utility with respect to the subjective belief of contract nonperformance with the decay rate of the first-order expression with respect to the subjective belief of contract nonperformance. The derivative of $h(\varepsilon)$ is given by

$$h'(\varepsilon) = \frac{U_\varepsilon(\alpha^*,\varepsilon)U_{\alpha\varepsilon}(\alpha^*,\varepsilon)\Big[U(\alpha^*,\varepsilon)U_{\alpha\varepsilon}(\alpha^*,\varepsilon) - U_\varepsilon(\alpha^*,\varepsilon)U_\alpha(\alpha^*,\varepsilon)\Big]}{U(\alpha^*,\varepsilon)^2 U_{\alpha\varepsilon}(\alpha^*,\varepsilon)^2}. \tag{A.18}$$

Using the fact that the first two terms in the numerator are negative, the sign of $h'(\varepsilon)$ is determined by the square bracket in the numerator. For $\varepsilon < \hat{\varepsilon}$, we obtain the equivalence that $h'(\varepsilon) > 0$ if and only if $h(\varepsilon) > 1$. However, due to continuity of $h(\varepsilon)$, this contradicts with $h(\hat{\varepsilon}) = 0$ so that $h'(\varepsilon) \leq 0$ and $h(\varepsilon) \leq 1$ must be satisfied for $\varepsilon < \hat{\varepsilon}$. This property together with the assumptions stated in Proposition 1.4.2 imply that $g'(\varepsilon) \leq 0$ and $g''(\varepsilon) \leq 0$ for all ε, which completes the proof.

Proof of Proposition 1.4.3

We conclude that

$$\begin{aligned}
& \mathbb{E}\left\{\phi''[U(\alpha^*,\widetilde{\varepsilon})]U_W(\alpha^*,\widetilde{\varepsilon})U_\alpha(\alpha^*,\widetilde{\varepsilon})\right\} \\
= \quad & -\mathbb{E}\left\{-\frac{\phi''[U(\alpha^*,\widetilde{\varepsilon})]}{\phi'[U(\alpha^*,\widetilde{\varepsilon})]}U_W(\alpha^*,\widetilde{\varepsilon})\phi'[U(\alpha^*,\widetilde{\varepsilon})])U_\alpha(\alpha^*,\widetilde{\varepsilon})\right\} \\
= \quad & -\mathrm{Cov}\left\{A_\phi(U(\alpha^*,\widetilde{\varepsilon}))U_W(\alpha^*,\widetilde{\varepsilon}), \phi'[U(\alpha^*,\widetilde{\varepsilon})]U_\alpha(\alpha^*,\widetilde{\varepsilon})\right\},
\end{aligned} \tag{A.19}$$

where the second equality follows from the consumer's first-order condition (1.3). Under non-increasing absolute ambiguity aversion, the first term in the covariance is an increasing function of ε. To see this, note that

$$U_W(\alpha^*, \varepsilon) = (1-p)u'(W_1^*) + p(\overline{q} - \varepsilon)u'(W_2^*) + p(1 - \overline{q} + \varepsilon)u'(W_3^*) > 0, \tag{A.20}$$

and that

$$U_{W\varepsilon}(\alpha^*, \varepsilon) = p\left(u'(W_3^*) - u'(W_2^*)\right) > 0, \tag{A.21}$$

where the asterisk indicates terminal wealth levels for $\alpha = \alpha^*$. Consequently, we obtain that

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left(A_\phi(U(\alpha^*, \varepsilon))U_W(\alpha^*, \varepsilon)\right) \\ &= \underbrace{A'_\phi(U(\alpha^*, \varepsilon))}_{\leq 0}\underbrace{U_\varepsilon(\alpha^*, \varepsilon)}_{<0}\underbrace{U_W(\alpha^*, \varepsilon)}_{>0} + \underbrace{A_\phi(U(\alpha^*, \varepsilon))}_{>0}\underbrace{U_{W\varepsilon}(\alpha^*, \varepsilon)}_{>0} > 0. \end{aligned} \tag{A.22}$$

Whether the second term in the covariance is increasing or decreasing in ε depends on the sign of the function $g'(\varepsilon)$, which was introduced in the proof of Proposition 1.4.2 and shown to be non-positive if relative ambiguity aversion is bounded by unity. Consequently, the covariance is negative and the overall sign of the partial wealth effect is positive.

APPENDIX B
APPENDIX TO CHAPTER 2

Proof of Proposition 2.3.1

First we define pension plan funding ratio δ as the ratio of initial pension asset W_0 to the present value of the firm's pension liability F with discount rate R_f, thus we have $W_0 = FR_f^{-1}\delta$.

In the first case with no bankruptcy risk, the pension payment is $S = F$. Define function $\phi(x)$ as $\phi(x) = C_1(x)$ if $x > 0$, and $\phi(x) = 0$ otherwise. The firm's objective function is specified as

$$G(w, \delta) = -C_0((W_0 - H_0)^+) - FR_f^{-1} - R_f^{-1}E^Q\left(\phi(z)\right)$$

where $z = F - W_T$.

For the case of $\delta \geq 1$, we have $G(0, \delta) = -C_0((W_0 - H_0)^+) - FR_f^{-1}$. Since $\phi(z) \geq 0$, we have $G(w, \delta) \leq -C_0((W_0 - H_0)^+) - FR_f^{-1}$. Thus $(0, \delta) = \operatorname{argmax} G(w, \delta)$ for $\delta \geq 1$.

For the case of $\delta < 1$, $G(0, \delta) = -C_0((W_0 - H_0)^+) - FR_f^{-1} - R_f^{-1}\phi(z_0)$ where $z_0 = F(1-\delta)$. Under the assumption that $\phi(z)$ is convex, we have $\phi(z) \geq \phi(z_0) + c(z - z_0)$ for some constant c.[1] Therefore,

$$\begin{aligned} G(w, \delta) \leq & -C_0((W_0 - H_0)^+) - FR_f^{-1} - R_f^{-1}E^Q\left[(\phi(z_0) + c(z - z_0))\right] \\ =& G(0, \delta) - cR_f^{-1}E^Q[(z - z_0)] = G(0, \delta) - cR_f^{-1}E^Q\left[(F - W_T - F + F\delta)\right] \end{aligned}$$

---

[1]Where c is the slope of a support of the convex function. If $\phi(x)$ is differentiable at z_0, then the support is unique and $c = \phi'(z_0)$.

$$=G(0,\delta) - cW_0 R_f^{-1} E^Q \left[(R_f - R(w))\right] = G(0,\delta)$$

Given that $\phi(z)$ is strictly convex at least at one point (otherwise it would be constant 0), the above inequalities will be strict when $w \neq 0$. Hence $G(0,\delta)$ is the unique maximum for any δ.

Now turn to the case with bankruptcy risk. We need the following two lemmas for the proof.

Lemma B.0.1: *Given any promised retirement benefit fixed, more risky asset investment will raise the present value of external financing cost regardless of the initial funding status.*

Need to show that

$$\frac{\partial E^Q \left[C_T\left((F - W_T)^+\right)\right]}{\partial w} > 0, \text{ for any given } F$$

Taking the derivative with respect to w, we have

$$\frac{\partial E^Q \left[C_T\left((F - W_T)^+\right)\right]}{\partial w} = -\int_{-\infty}^{R_m^c} C_T'(F - W_T) W_0 R_e f^Q(R_m) dR_m$$

where R_m^c is the realized risky return such that $F = W_T$.

If the pension plan is initially overfunded, $\delta > 1$, $R_m^c < R_f$, the risky asset's excess return R_e is always negative in the integration. Thus the derivative is positive.

If the pension plan is initially underfunded, $\delta > 1$, $R_m^c > R_f$, the derivative can be written as

$$\frac{\partial E^Q \left[C_T\left((F - W_T)^+\right)\right]}{\partial w}$$

$$= -\int_{-\infty}^{R_f} C_T'(F - W_T)W_0 R_e f^Q(R_m)dR_m - \int_{R_f}^{R_m^c} C_T'(F - W_T)W_0 R_e f^Q(R_m)dR_m$$
$$> -\int_{-\infty}^{R_f} C_T'(F - F\delta)W_0 R_e f^Q(R_m)dR_m - \int_{R_f}^{R_m^c} C_T'(F - F\delta)W_0 R_e f^Q(R_m)dR_m$$
$$= -\int_{-\infty}^{R_m^c} C_T'(F - F\delta)W_0 R_e f^Q(R_m)dR_m > -C_T'(F - F\delta)W_0 \int_{-\infty}^{+\infty} R_e f^Q(R_m)dR_m = 0$$

The second inequality holds because $\frac{\partial C_T'(F-W_T)}{\partial R_m} = C_T''(F - W_T)(-W_0 w) < 0$ as long as $w > 0$. Therefore, $C_T'(F-W_T) > C_T'(F-F\delta)$ for $R_m < R_f$, and $C_T'(F-W_T) < C_T'(F - F\delta)$ for $R_m > R_f$.

Next, consider the employees' participation constraint

$$(1 - p)U(F) + pEU(W_T(w, \delta)) = \mathbb{U}$$

which implies $F = F(\mathbb{U}, w, \delta)$. We can prove the following lemma.

Lemma B.0.2: $\frac{\partial F(\mathbb{U},0,\delta)}{\partial w} < 0$, $\frac{\partial F(\mathbb{U},w_e,\delta)}{\partial w} = 0$, *and* $\frac{\partial F(\mathbb{U},\hat{w}_e,\delta)}{\partial w} > 0$, *for all* $\mathbb{U} > 0$ *and* $\delta > 0$. w_e *is employee's optimal investment allocation of self-management portfolio, and* $\hat{w}_e > w_e$.

With the other parameters exogenously given, we can obtain the general form of $\frac{dF}{dw}$ using implicit function theorem.

$$\frac{dF}{dw} = \frac{\partial F}{\partial w} = -\frac{pE[U'(W_T)W_0 R_e]}{(1-p)U'(F) + pE[U'(W_T)R_f^{-1}\delta R(w)]}$$

Now let's figure out the sign of $\frac{dF}{dw}$ evaluated at three critical w's: 0, w_e, and $\hat{w}_e$.

$$\left.\frac{dF}{dw}\right|_{w=0} = -\frac{pU'(F\delta)W_0 E(R_e)}{(1-p)U'(F) + pU'(F\delta)\delta} < 0$$

The CRRA employee chooses w_e such that $\frac{dEU(W_T)}{dw} = E[U'(F,w_e)W_0R_e] = 0$ for any F. Therefore,

$$\left.\frac{dF}{dw}\right|_{w=w_e} = -\frac{pE[U'(F,w_e)W_0R_e]}{(1-p)U'(F) + pE[U'(F,w_e)R_f^{-1}\delta R(w_e)]} = 0$$

For any given F,

$$\frac{dE[U'(F,w)R_e]}{dw} = E[U''(F,w)R_e^2]W_0 < 0$$

which implies $E[U'(F,\hat{w}_e)R_e] < E[U'(F,w_e)R_e] = 0$. Further we are able to show

$$\begin{aligned}\frac{dE[U'(F,w)R(w)]}{dw} =& E[U''(F,w)W_0R_eR(w) + U'(F,w)R_e] \\ =& E[U''(W_T)W_TR_e + U'(W_T)R_e] = E[(1-\gamma)U'(F,w)R_e] \geq 0, \forall w \geq w_e\end{aligned}$$

Hence $E[U'(F,\hat{w}_e)R(\hat{w}_e)] > E[U'(F,w_e)R(w_e)] = EU'(F,w_e)R_f > 0$. So we have

$$\left.\frac{dF}{dw}\right|_{w=\hat{w}_e} = -\frac{pE[U'(F,\hat{w}_e)W_0R_e]}{(1-p)U'(F) + pE[U'(F,\hat{w}_e)R_f^{-1}\delta R(\hat{w}_e)]} > 0$$

Now, to finish the proof of Proposition 1, we need to sign $G'(w)$ for $w = 0$, w_e, and $\hat{w}_e$. Let's start from $w = w_e$ and $w = \hat{w}_e$.

$$\begin{aligned}G'(w_e) =& -\left.\frac{dF}{dw}\right|_{w=w_e} R_f^{-1}(p\delta + (1-p)) - \left.\frac{dF}{dw}\right|_{w=w_e} C_0'(W_0 - H_0)R_f^{-1}\delta \\ & - (1-p)R_f^{-1}\int_{-\infty}^{R_m^c(w_e)} C_T'(F - W_T(w_e))(1 - R_f^{-1}\delta R(w_e)) \left.\frac{dF}{dw}\right|_{w=w_e} f^Q(R_m)dR_m \\ & + (1-p)R_f^{-1}\int_{-\infty}^{R_m^c(w_e)} C_T'(F - W_T(w_e))W_0R_ef^Q(R_m)dR_m \\ =& (1-p)R_f^{-1}\int_{-\infty}^{R_m^c(w_e)} C_T'(F - W_T(w_e))W_0R_ef^Q(R_m)dR_m < 0\end{aligned}$$

The last inequality comes from Lemma B.0.1.

$$\begin{aligned} G'(\hat{w}_e) = &- \left.\frac{dF}{dw}\right|_{w=\hat{w}_e} R_f^{-1}(p\delta + (1-p)) - \left.\frac{dF}{dw}\right|_{w=\hat{w}_e} C_0'(W_0 - H_0)R_f^{-1}\delta \\ &- (1-p)R_f^{-1}\int_{-\infty}^{R_m^c(\hat{w}_e)} C_T'(F - W_T(\hat{w}_e))(1 - R_f^{-1}\delta R(\hat{w}_e)) \left.\frac{dF}{dw}\right|_{w=\hat{w}_e} f^Q(R_m)dR_m \\ &+ (1-p)R_f^{-1}\int_{-\infty}^{R_m^c(\hat{w}_e)} C_T'(F - W_T(\hat{w}_e))FR_f^{-1}\delta R_e f^Q(R_m)dR_m < 0 \end{aligned}$$

For $w = 0$, we discuss the results separately with different δ:

1. If the pension plan is initially underfunded, $\delta < 1$, then $R_m^c|_{w=0} = +\infty$,

$$\begin{aligned} G'(0) = &- \left.\frac{dF}{dw}\right|_{w=0} R_f^{-1}(p\delta + (1-p)) - \left.\frac{dF}{dw}\right|_{w=0} C_0'(W_0 - H_0)R_f^{-1}\delta \\ &- (1-p)R_f^{-1}\int_{-\infty}^{+\infty} C_T'(F - F\delta)(1-\delta) \left.\frac{dF}{dw}\right|_{w=0} f^Q(R_m)dR_m \\ &+ (1-p)R_f^{-1}\int_{-\infty}^{+\infty} C_T'(F - F\delta)W_0 R_e f^Q(R_m)dR_m \\ = &- \left.\frac{dF}{dw}\right|_{w=0} R_f^{-1}\left[p\delta + C_0'(W_0 - H_0)\delta + (1-p)(1 + C_T'(F - F\delta)(1-\delta))\right] > 0 \end{aligned}$$

2. If the pension plan is initially overfunded, $\delta > 1$, then $R_m^c|_{w=0} = -\infty$,

$$\begin{aligned} G'(0) = &- \left.\frac{dF}{dw}\right|_{w=0} R_f^{-1}(p\delta + (1-p)) - \left.\frac{dF}{dw}\right|_{w=0} C_0'(W_0 - H_0)R_f^{-1}\delta \\ &- (1-p)R_f^{-1}\int_{-\infty}^{-\infty} C_T'(F - F\delta)(1-\delta) \left.\frac{dF}{dw}\right|_{w=0} f^Q(R_m)dR_m \\ &+ (1-p)R_f^{-1}\int_{-\infty}^{-\infty} C_T'(F - F\delta)V_0 R_e f^Q(R_m)dR_m \\ = &- \left.\frac{dF}{dw}\right|_{w=0} R_f^{-1}\left[(p\delta + (1-p)) + C_0'(W_0 - H_0)\delta\right] > 0 \end{aligned}$$

3. If the pension plan is initially fully funded, $\delta = 1$, then $R_m^c|_{w=0} = R_f$, $\frac{dF}{dw}\big|_{w=0} = -pFR_f^{-1}E[R_e]$,

$$G'(0) = - \left.\frac{dF}{dw}\right|_{w=0} R_f^{-1}(p + (1-p)) - \left.\frac{dF}{dw}\right|_{w=0} C_0'(W_0 - H_0)R_f^{-1}$$

$$
\begin{aligned}
&- (1-p)R_f^{-1}\int_{-\infty}^{R_f} C_T'(F-F)(1-1)\left.\frac{dF}{dw}\right|_{w=0} f^Q(R_m)dR_m \\
&+ (1-p)R_f^{-1}\int_{-\infty}^{R_f} C_T'(F-F)FR_f^{-1}R_e f^Q(R_m)dR_m \\
=&FR_f^{-1}\left\{pR_f^{-1}E[R_e]\left[1+C_0'(FR_f^{-1}-H_0)\right]-(1-p)R_f^{-1}C_T'(0)E^Q[R_e^+]\right\}
\end{aligned}
$$

If the cost function is in quadratic form, $C_{0,T}'(0) = 0$, and $G'(0) > 0$; otherwise $V'(0) > 0$ if and only if $C_T'(0) < \frac{pR_f^{-1}E[R_e]\left[1+C_0'(FR_f^{-1}-H_0)\right]}{(1-p)R_f^{-1}E^Q[R_e^+]} \equiv C$.

Given the sign of $G'(0)$, $G'(w_e)$, and $G'(\hat{w}_e)$, we can conclude that there exists global maximization solution $w^* \in (0, w_e)$ such that $G'(w^*) = 0$.

Proof of Corollary I

The first pension payment schedule implies a positive bankruptcy risk and that the firm takes all the surplus if any.

The firm's objective function $G(w,\delta)$ becomes,

$$
\begin{aligned}
G(w,\delta) =& - W_0(\delta) - C_0\left((W_0(\delta)-H_0)^+\right) + pR_f^{-1}E^Q\left[(W_T-F)^+\right] \\
&+ (1-p)R_f^{-1}E^Q\left[W_T(w,\delta) - F - C_T\left((F-W_T(w,\delta))^+\right)\right]
\end{aligned}
$$

Employee's participation constraint becomes,

$$
\mathbb{U} = (1-p)U(F) + pEU(F + (W_T(w,\delta)-F)^-) \equiv EU(w,\delta)
$$

The firm needs to maximize the following Lagrange function $\mathbb{L} = G(w,\delta) + \lambda(EU(w,\delta) - \mathbb{U})$. The first-order derivative w.r.t. w is

$$
\frac{\partial \mathbb{L}}{\partial w} = \frac{\partial G(w,\delta)}{\partial w} + \lambda \frac{\partial EU(w,\delta)}{\partial w}
$$

Look at the first term, the marginal utility of firm,

$$\frac{\partial G(w,\delta)}{\partial w} = R_f^{-1} \int_{-\infty}^{R_m^c} W_0 R_e \left[C_T'(F - W_0 R(w))(1-p) - p \right] f^Q(R_m) dR_m$$

where $R_m^c = \left(\frac{F}{W_0} - R_f \right) \frac{1}{w} + R_f$.

At $w = 0$, $R_m^c \to +\infty$ if underfunded; $R_m^c \to -\infty$ if overfunded, either way we have

$$\frac{\partial G(0,\delta)}{\partial w} = 0$$

Now look at the second term, the marginal utility of employee,

$$\frac{\partial EU}{\partial w} = p \int_{-\infty}^{R_m^c} U'(W_0 R(w)) W_0 R_e f(R_m) dR_m$$

If pension is underfunded, $W_0 R_f < F$, $R_m^c > R_f$,

$$\frac{\partial EU(0,\delta)}{\partial w} = p \int_{-\infty}^{\infty} U'(W_0 R_f) W_0 R_e f(R_m) dR_m = pU'(W_0 R_f) E(R_m - R_f) > 0$$

Therefore, when pension is underfunded,

$$\left. \frac{\partial \mathbb{L}}{\partial w} \right|_{w=0} = \frac{\partial G(0,\delta)}{\partial w} + \lambda \frac{\partial EU(0,\delta)}{\partial w} > 0$$

The second pension payment schedule implies no bankruptcy risk and that the firm and the employee shares the surplus if any. The firm's objective function $G(w,\delta)$ becomes

$$\begin{aligned} & G(w,\delta) \\ = & -W_0(\delta) - C_0\left((W_0(\delta) - H_0)^+ \right) + R_f^{-1} E^Q \left[(W_T - F)^+ \alpha + (W_T - F)^- - C_T((F - W_T)^+) \right] \end{aligned}$$

Employee's participation constraint becomes

$$\mathbb{U} = EU\left(F + (W_T - F)^+ (1-\alpha) \right)$$

Following the same procedure as above,

$$\frac{\partial G}{\partial w} = R_f^{-1} \int_{-\infty}^{R_m^c} W_0 R_e \left[(1-\alpha) + C_T'(F - W_T)\right] f^Q(R_m) dR_m$$

where $R_m^c = \left(\frac{F}{W_0} - R_f\right)\frac{1}{w} + R_f$.

At $w = 0$, $R_m^c \to +\infty$ if underfunded; $R_m^c \to -\infty$ if overfunded, either way we have

$$\frac{\partial G(0,\delta)}{\partial w} = 0$$

Similarly,

$$\frac{\partial EU}{\partial w} = \int_{R_m^c}^{+\infty} U'(F\alpha + W_T(1-\alpha))(1-\alpha)W_0 R_e f(R_m) dR_m$$

If fully funded,$R_m^c = R_f$,

$$\frac{\partial EU(0,1)}{\partial w} = \int_{R_f}^{+\infty} U'(F\alpha + W_0 R_f(1-\alpha))(1-\alpha)W_0 R_e f(R_m) dR_m > 0$$

If overfunded, $R_m^c \to -\infty$,

$$\frac{\partial EU(0,\delta)}{\partial w} = \int_{-\infty}^{+\infty} U'(F\alpha + W_0 R_f(1-\alpha))(1-\alpha)W_0 R_e f(R_m) dR_m > 0$$

Therefore, when pension is not underfunded,

$$\left.\frac{\partial \mathbb{L}}{\partial w}\right|_{w=0} = \frac{\partial G(0,\delta)}{\partial w} + \lambda \frac{\partial EU(0,\delta)}{\partial w} > 0$$

REFERENCES

Acharya, Viral V, Philipp Schnabl, and Gustavo Suarez, "Securitization without risk transfer," *Journal of Financial economics*, 2013, *107* (3), 515–536.

Adrian, Tobias and Adam B Ashcraft, "Shadow banking: a review of the literature," *FRB of New York Staff Report*, 2012, (580).

_ and Nina Boyarchenko, "Intermediary leverage cycles and financial stability," *Becker Friedman Institute for Research in Economics Working Paper*, 2012, (2012-010).

_ , Paolo Colla, and Hyun Song Shin, "Which financial frictions? Parsing the evidence from the financial crisis of 2007-9," Technical Report, National Bureau of Economic Research 2012.

Alary, David, Christian Gollier, and Nicolas Treich, "The Effect of Ambiguity Aversion on Insurance and Self-protection," *The Economic Journal*, 2013, *123* (573), 1188–1202.

An, Heng, Zhaodan Huang, and Ting Zhang, "What determines corporate pension fund risk-taking strategy?," *Journal of Banking & Finance*, 2013, *37* (2), 597–613.

Andonov, Aleksandar, Rob Bauer, and Martijn Cremers, "Can large pension funds beat the market? Asset allocation, market timing, security selection and the limits of liquidity," *Working Paper, Maastricht University*, 2012.

Ang, Andrew, *Asset management: A systematic approach to factor investing*, Oxford University Press, 2014.

Asmat, Daniel and Sharon Tennyson, "Does the Threat of Insurer Liability for "Bad Faith" Affect Insurance Settlements?," *Journal of Risk and Insurance*, 2014, *81* (1), 1–26.

Baillon, Aurélien, "Prudence with respect to ambiguity," *The Economic Journal*, 2017, *127* (604), 1731–1755.

_ and Aysil Emirmahmutoglu, "Zooming in on ambiguity attitudes," *International Economic Review (forthcoming)*, 2018.

Bajtelsmit, Vickie, Jennifer Coats, and Paul Thistle, "The effect of ambiguity on risk management choices: An experimental study," *Journal of Risk and Uncertainty*, 2015, *50* (3), 249–280.

Benanke, BS and M Gertler, "Agency Costs, Net Worth, and Business Fluctuation," *The American Economic Review*, 1989, *1989*, 14–31.

Berger, Loïc, "Precautionary saving and the notion of ambiguity prudence," *Economics Letters*, 2014, *123* (2), 248–251.

— , "The impact of ambiguity and prudence on prevention decisions," *Theory and Decision*, 2016, *80* (3), 389–409.

— **and Valentina Bosetti**, "Ellsberg re-revisited: An experiment disentangling model uncertainty and risk aversion," *Working Paper (IGIER and Bocconi University)*, 2016.

Berk, Jonathan B and Peter M DeMarzo, *Corporate finance (4th edition)*, Pearson Education, 2016.

Bernanke, Ben S, Cara S Lown, and Benjamin M Friedman, "The credit crunch," *Brookings papers on economic activity*, 1991, pp. 205–247.

— **, Mark Gertler, and Simon Gilchrist**, "The financial accelerator in a quantitative business cycle framework," *Handbook of macroeconomics*, 1999, *1*, 1341–1393.

Bernard, Carole and Mike Ludkovski, "Impact of counterparty risk on the reinsurance market," *North American Actuarial Journal*, 2012, *16* (1), 87–111.

Biener, Christian, Andreas Landmann, and Maria Santana, "Contract nonperformance risk and ambiguity in insurance markets," *University of St. Gallen (Working Paper)*, 2017.

Biffis, Enrico and Pietro Millossovich, "Optimal insurance with counterparty default risk," *Georgia State University (Working Paper)*, 2012.

Black, Fischer, "The tax consequences of long-run pension policy," *Financial Analysts Journal*, 1980, *36* (4), 21–28.

Bodie, Zvi, "Pension fund investment policy," 1988.

— , "The ABO, the PBO and pension investment policy," *Financial Analysts Journal*, 1990, *46* (5), 27–34.

— **, Jay O Light, Randall Morck, and Robert A Taggart Jr**, "Corporate pension policy: An empirical investigation," *Financial Analysts Journal*, 1985, *41* (5), 10–16.

Bourgeon, Jean-Marc and Pierre Picard, "Fraudulent claims and nitpicky insurers," *The American Economic Review*, 2014, *104* (9), 2900–2917.

Briys, Eric, Georges Dionne, and Louis Eeckhoudt, "More on insurance as a Giffen good," *Journal of Risk and Uncertainty*, 1989, *2* (4), 415–420.

— **, Harris Schlesinger, and J-Matthias Graf vd Schulenburg**, "Reliability of risk management: market insurance, self-insurance and self-protection reconsidered," *Geneva Papers on Risk and Insurance Theory*, 1991, *16* (1), 45–58.

Brunnermeier, Markus K and Lasse Heje Pedersen, "Market liquidity and funding liquidity," *Review of Financial studies*, 2009, *22* (6), 2201–2238.

__ **and Yuliy Sannikov**, "A Macroeconomic Model with a Financial Sector," *The American Economic Review*, 2014, *104* (2), 379–421.

__ **, Thomas M Eisenbach, and Yuliy Sannikov**, "Macroeconomics with financial frictions: A survey," Technical Report, National Bureau of Economic Research 2012.

Bryan, Gharad, "Ambiguity aversion decreases the impact of partial insurance: Evidence from African farmers," *Journal of the European Economic Association (forthcoming)*, 2018.

Bulow, Jeremy I and Myron S Scholes, "Who owns the assets in a defined-benefit pension plan?," in "Financial aspects of the United States pension system," University of Chicago Press, 1983, pp. 17–36.

Busse, Jeffrey A, Amit Goyal, and Sunil Wahal, "Performance and persistence in institutional investment management," *The Journal of Finance*, 2010, *65* (2), 765–790.

Caballé, Jordi and Alexey Pomansky, "Mixed risk aversion," *Journal of Economic Theory*, 1996, *71* (2), 485–513.

Cabantous, Laure, "Ambiguity aversion in the field of insurance: Insurers' attitude to imprecise and conflicting probability estimates," *Theory and Decision*, 2007, *62* (3), 219–240.

__ **, Denis Hilton, Howard Kunreuther, and Erwann Michel-Kerjan**, "Is imprecise knowledge better than conflicting expertise? Evidence from insurers' decisions in the United States," *Journal of Risk and Uncertainty*, 2011, *42* (3), 211–232.

Camerer, C. and M. Weber, "Recent developments in modeling preferences: Uncertainty and ambiguity," *Journal of Risk and Uncertainty*, 1992, *5* (4), 325–370.

Campbell, John Y and Luis M Viceira, "Strategic asset allocation for pension plans," in "The Oxford Handbook of Pensions and Retirement Income" 2005.

Cather, David A, Elizabeth S Cooperman, and Glenn A Wolfe, "Excess pension asset reversions and corporate acquisition activity," *Journal of Business Research*, 1991, *23* (4), 337–348.

Chen, Xuanjuan, Tong Yao, Tong Yu, and Ting Zhang, "Learning and incentive: A study on analyst response to pension underfunding," *Journal of Banking & Finance*, 2014, *45*, 26–42.

Chernenko, Sergey and Adi Sunderam, "Frictions in Shadow Banking: Evidence from the Lending Behavior of Money Market Mutual Funds," *Review of Financial Studies*, 2014, *27* (6), 1717–1750.

Chesson, H.W. and W.K. Viscusi, "Commonalities in Time and Ambiguity Aversion for Long-Term Risks," *Theory and Decision*, 2003, *54* (1), 57–71.

Chiu, W Henry, Louis Eeckhoudt, and Beatrice Rey, "On relative and partial risk attitudes: Theory and implications," *Economic Theory*, 2012, *50* (1), 151–167.

Chow, C.C. and R.K. Sarin, "Comparative ignorance and the Ellsberg paradox," *Journal of Risk and Uncertainty*, 2001, *22* (2), 129–139.

Condie, Scott and Jayant V Ganguli, "Ambiguity and rational expectations equilibria," *The Review of Economic Studies*, 2011, *78* (3), 821–845.

Constantinides, George M, "Habit formation: A resolution of the equity premium puzzle," *Journal of political Economy*, 1990, *98* (3), 519–543.

Cooper, Stephen and David Bianco, "Should pension funds invest in equities," *UBS Investment Research, Q-Series: Pension Fund Asset Allocation*, 2003.

Crocker, Keith and Sharon Tennyson, "Insurance fraud and optimal claims settlement strategies," *Journal of Law and Economics*, 2002, *45* (2), 469.

Cummins, David and Olivier Mahul, "Optimal insurance with divergent beliefs about insurer total default risk," *Journal of Risk and Uncertainty*, 2003, *27* (2), 121–138.

Curdia, Vasco and Michael Woodford, "Credit spreads and monetary policy," *Journal of Money, Credit and Banking*, 2010, *42* (s1), 3–35.

Diamond, Douglas W, "Liquidity, banks, and markets," *Journal of Political Economy*, 1997, *105* (5), 928–956.

— **and Philip H Dybvig**, "Bank runs, deposit insurance, and liquidity," *The journal of political economy*, 1983, pp. 401–419.

Dimmock, Stephen, Roy Kouwenberg, and Peter Wakker, "Ambiguity attitudes in a large representative sample," *Management Science*, 2015, *62* (5), 1363–1380.

— **,** — **, Olivia Mitchell, and Kim Peijnenburg**, "Estimating ambiguity preferences and perceptions in multiple prior models: Evidence from the field," *Journal of Risk and Uncertainty*, 2015, *51* (3), 219–244.

— **,** — **,** — **, and** — , "Ambiguity aversion and household portfolio choice puzzles: Empirical evidence," *Journal of Financial Economics*, 2016, *119* (3), 559–577.

Doherty, Neil A and Harris Schlesinger, "Rational Insurance Purchasing: Consideration of Contract Nonperformance," *The Quarterly Journal of Economics*, 1990, *105* (1), 243–253.

Doherty, Neil, Christian Laux, and Alexander Muermann, "Insuring Nonverifiable Losses," *Review of Finance*, 2013, *19* (1), 283–316.

Duffie, Darrell and Bruno Strulovici, "CAPITAL MOBILITY AND ASSET PRICING," *Econometrica*, 2012, *80* (6), 2469–2509.

Eeckhoudt, Louis and Harris Schlesinger, "Changes in risk and the demand for saving," *Journal of Monetary Economics*, 2008, *55* (7), 1329–1336.

_ , Christian Gollier, and Harris Schlesinger, *Economic and financial decisions under risk*, Princeton University Press, 2005.

_ , Liqun Liu, and Jack Meyer, "Restricted increases in risk aversion and their application," *Economic Theory*, 2017, *64* (1), 161–181.

_ , Ph Godfroid, and Christian Gollier, "Willingness to pay, the risk premium and risk aversion," *Economics Letters*, 1997, *55* (3), 355–360.

Ehrlich, Isaac and Gary S Becker, "Market insurance, self-insurance, and self-protection," *Journal of Political Economy*, 1972, *80* (4), 623–648.

Einhorn, H.J. and R.M. Hogarth, "Decision making under ambiguity," *Journal of Business*, 1986, *59* (4), 225–250.

Ekern, Steinar, "Increasing Nth degree risk," *Economics Letters*, 1980, *6* (4), 329–333.

Ellsberg, D., "Risk, ambiguity, and the Savage axioms," *The Quarterly Journal of Economics*, 1961, *75* (4), 643–669.

French, Kenneth R, "Presidential address: The cost of active investing," *The Journal of Finance*, 2008, *63* (4), 1537–1573.

Froot, Kenneth A and Jeremy C Stein, "Risk management, capital budgeting, and capital structure policy for financial institutions: an integrated approach," *Journal of financial economics*, 1998, *47* (1), 55–82.

Gennaioli, Nicola, Andrei Shleifer, and Robert W Vishny, "A model of shadow banking," *The Journal of Finance*, 2013, *68* (4), 1331–1363.

Gertler, Mark and Peter Karadi, "A model of unconventional monetary policy," *Journal of monetary Economics*, 2011, *58* (1), 17–34.

Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci, "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory*, 2004, *118* (2), 133–173.

Gilboa, Itzhak, "Expected utility with purely subjective non-additive probabilities," *Journal of Mathematical Economics*, 1987, *16* (1), 65–88.

_ and David Schmeidler, "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 1989, *18* (2), 141–153.

Gollier, Christian, "Portfolio choices and asset prices: The comparative statics of ambiguity aversion," *Review of Economic Studies*, 2011, *78* (4), 1329–1344.

__ , "The economics of optimal insurance design," in Georges Dionne, ed., *Handbook of Insurance*, 2 ed., Springer, 2013, chapter 4, pp. 107–122.

__ , "Optimal insurance design of ambiguous risks," *Economic Theory*, 2014, *57* (3), 555–576.

Gorton, Gary and Andrew Metrick, "Securitized banking and the run on repo," *Journal of Financial economics*, 2012, *104* (3), 425–451.

__ , __ , Andrei Shleifer, and Daniel K Tarullo, "Regulating the shadow banking system [with comments and discussion]," *Brookings Papers on Economic Activity*, 2010, pp. 261–312.

Hadar, Josef and Tae Kun Seo, "The effects of shifts in a return distribution on optimal portfolios," *International Economic Review*, 1990, *31* (3), 721–736.

Hanson, Samuel G, Andrei Shleifer, Jeremy C Stein, and Robert W Vishny, "Banks as patient fixed-income investors," *Journal of Financial Economics*, 2015, *117* (3), 449–469.

Harrison, Glenn and Jia Ng, "Welfare Effects of Insurance Contract Non-Performance," *Geneva Risk and Insurance Review (forthcoming)*, 2018.

Hau, Arthur, "When Is a Coinsurance-Type Insurance Policy Inferior or Even Giffen?," *The Journal of Risk and Insurance*, 2008, *75* (2), 343–364.

He, Zhiguo and Arvind Krishnamurthy, "A model of capital and crises," *The Review of Economic Studies*, 2011, p. rdr036.

__ **and** __ , "Intermediary Asset Pricing," *American Economic Review*, 2013, *103* (2), 732–70.

Hennessy, Christopher A and Toni M Whited, "How costly is external financing? Evidence from a structural estimation," *The Journal of Finance*, 2007, *62* (4), 1705–1745.

Herrero, Carmen, Josefa Tomás, and Antonio Villar, "Decision theories and probabilistic insurance: An experimental test," *Spanish Economic Review*, 2006, *8* (1), 35–52.

Holmstrom, Bengt and Jean Tirole, "Financial intermediation, loanable funds, and the real sector," *the Quarterly Journal of economics*, 1997, pp. 663–691.

Hoy, Michael and Arthur J Robson, "Insurance as a Giffen good," *Economics Letters*, 1981, *8* (1), 47–51.

__ , Richard Peter, and Andreas Richter, "Take-up for genetic tests and ambiguity," *Journal of Risk and Uncertainty*, 2014, *48* (2), 111–133.

Hu, Grace Xing, Jun Pan, and Jiang Wang, "Noise as information for illiquidity," *The Journal of Finance*, 2013, *68* (6), 2341–2382.

Huang, Rachel and Larry Tzeng, "Insurer's insolvency risk and tax deductions for the individual's net losses," *Geneva Risk and Insurance Review*, 2007, *32* (2), 129–145.

__ **, Yi-Chieh Huang, and Larry Tzeng**, "Insurance bargaining under ambiguity," *Insurance: Mathematics and Economics*, 2013, *53* (3), 812–820.

Huang, Yi-Chieh and Larry Tzeng, "A Mean-Preserving Increase in Ambiguity and Portfolio Choices," *Journal of Risk and Insurance (forthcoming)*, 2017.

Ingersoll, Jonathan E, *Theory of financial decision making*, Lanham, MD: Rowman & Littlefield, 1987.

Iwaki, Hideki and Yusuke Osaki, "The dual theory of the smooth ambiguity model," *Economic Theory*, 2014, *56* (2), 275–289.

Jaspersen, Johannes, "Hypothetical surveys and experimental studies of insurance demand: A review," *Journal of Risk and Insurance*, 2016, *83* (1), 217–255.

Jean, William H, "The geometric mean and stochastic dominance," *The Journal of Finance*, 1980, *35* (1), 151–158.

__ , "The harmonic mean and other necessary conditions for stochastic dominance," *The Journal of Finance*, 1984, *39* (2), 527–534.

Jindapon, Paan and William S Neilson, "Higher-order generalizations of Arrow–Pratt and Ross risk aversion: A comparative statics approach," *Journal of Economic Theory*, 2007, *136* (1), 719–728.

Kahn, Barbara and Rakesh Sarin, "Modeling ambiguity in decisions under uncertainty," *Journal of Consumer Research*, 1988, *15* (2), 265–272.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–292.

Kimball, Miles S, "Precautionary Saving in the Small and in the Large," *Econometrica*, 1990, *58* (1), 53–73.

__ , "Precautionary motives for holding assets," Technical Report, National Bureau of Economic Research 1991.

Kiyotaki, Nobuhiro and John Moore, "Credit Cycles," *The Journal of Political Economy*, 1997, *105* (2), 211–248.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, "A smooth model of decision making under ambiguity," *Econometrica*, 2005, *73* (6), 1849–1892.

Kocher, Martin, Amrei Lahno, and Stefan Trautmann, "Ambiguity aversion is not universal," *European Economic Review*, 2018, *101*, 268–283.

Koufopoulos, Kostas and Roman Kozhan, "Welfare-improving ambiguity in insurance markets with asymmetric information," *Journal of Economic Theory*, 2014, *151*, 551–560.

Lakonishok, Josef, Andrei Shleifer, Robert W Vishny, Oliver Hart, and George L Perry, "The structure and performance of the money management industry," *Brookings Papers on Economic Activity. Microeconomics*, 1992, *1992*, 339–391.

Lang, Matthias, "First-order and second-order ambiguity aversion," *Management Science*, 2016, *63* (4), 1254–1269.

Lee, Kangoh, "Risk aversion and self-insurance-cum-protection," *Journal of Risk and Uncertainty*, 1998, *17* (2), 139–151.

Leibowitz, Martin L, *Liability Returns: A New Perspective on Asset Allocation*, Salomon Bros., 1986.

_ **and Antti Ilmanen**, "US Corporate DB Pension Plans–Today's Challenges," *The Journal of Retirement*, 2016, *3* (4), 34.

Li, Hui, Seth Neumuller, and Casey Rothschild, "Optimal annuitization with imperfect information about insolvency risk," *Working Paper (Wellesley College)*, 2018.

Liu, Yanyan and Robert J Myers, "The dynamics of microinsurance demand in developing countries under liquidity constraints and insurer default risk," *The Journal of Risk and Insurance*, 2016, *83* (1), 121–138.

Love, David A, Paul A Smith, and David W Wilcox, "The effect of regulation on optimal corporate pension risk," *Journal of Financial Economics*, 2011, *101* (1), 18–35.

Lucas, Deborah and Stephen P Zeldes, "Valuing and hedging defined benefit pension obligations—The role of stocks revisited," *Northwestern University and Columbia University, working paper, September*, 2006.

Lusardi, Annamaria and Olivia Mitchelli, "Financial literacy and retirement preparedness: Evidence and implications for financial education," *Business Economics*, 2007, *42* (1), 35–44.

Mahul, Olivier and Brian Wright, "Optimal coverage for incompletely reliable insurance," *Economics Letters*, 2007, *95* (3), 456–461.

Martin, Antoine, David Skeie, and Ernst-Ludwig Von Thadden, "Repo runs," *Review of Financial Studies*, 2014, *27* (4), 957–989.

McDonnell, Ken, "History of Pension Plans," *online publication by Employee Benefit Research Institute*, 1998.

Menezes, Carmen, Charles Geiss, and John Tressler, "Increasing downside risk," *The American Economic Review*, 1980, *70* (5), 921–932.

Meyer, Jack and Michael B Ormiston, "Analyzing the demand for deductible insurance," *Journal of Risk and Uncertainty*, 1999, *18* (3), 223–230.

_ **and Michael Ormiston**, "Strong increases in risk and their comparative statics," *International Economic Review*, 1985, *26* (2), 425–437.

Miller, Merton and Myron Scholes, "Pension funding and corporate valuation," *University of Chicago*, 1981, *31.*

Morduch, Jonathan, "Microinsurance: The next revolution," *Understanding poverty*, 2006, pp. 337–356.

Moreira, Alan and Alexi Savov, "The macroeconomics of shadow banking," Technical Report, National Bureau of Economic Research 2014.

Mossin, Jan, "Aspects of rational insurance purchasing," *Journal of Political Economy*, 1968, *46* (4), 553–568.

Muir, Tyler, "Financial crises and risk premia," *Available at SSRN 2379608*, 2016.

Mukerji, Sujoy and Jean-Marc Tallon, "Ambiguity aversion and incompleteness of financial markets," *The Review of Economic Studies*, 2001, *68* (4), 883–904.

Munnell, Alicia and Mauricio Soto, "Why are companies freezing their pensions?," *Working Paper, Center for Retirement Research at Boston College*, 2007.

Munnell, Alicia Haydock, *The economics of private pensions*, Vol. 21, Brookings Inst Press, 1982.

Neilson, William, "A simplified axiomatic approach to ambiguity aversion," *Journal of Risk and Uncertainty*, 2010, *41* (2), 113–124.

Panis, Constantijn WA and Michael J Brien, "Asset allocation of defined benefit pension plans," *Deloitte Consulting Publications*, 2015.

PBGC, "PBGC's Guarantee Limits: An Update.," *Available at http://www.pbgc.gov/documents/guaranteelimits.pdf.*, 2008.

Peter, Richard, "The critical loading factor in insurance," *Working Paper (University of Iowa)*, 2016.

_ **, Andreas Richter, and Paul Thistle**, "Ambiguity aversion in competitive insurance markets: Adverse and advantageous selection," *University of Iowa (Working Paper)*, 2017.

Pozsar, Zoltan, Tobias Adrian, Adam B Ashcraft, and Haley Boesky, "Shadow Banking," *FRB of New York Staff Report*, 2010, (458).

Pratt, John W, "Risk aversion in the small and in the large," *Econometrica*, 1964, *32* (1–2), 122–136.

Qi, Jianping, "Bank liquidity and stability in an overlapping generations model," *Review of Financial Studies*, 1994, *7* (2), 389–417.

Quiggin, John, "A theory of anticipated utility," *Journal of Economic Behavior & Organization*, 1982, *3* (4), 323–343.

Rampini, Adriano A and S Viswanathan, "Financial intermediary capital," Technical Report, National Bureau of Economic Research 2017.

Rauh, Joshua D, "Risk shifting versus risk management: investment policy in corporate pension plans," *The Review of Financial Studies*, 2008, *22* (7), 2687–2733.

_ and Irina Stefanescu, "Why are firms in the United States abandoning defined benefit plans?," *Rotman International Journal of Pension Management*, 2009, *2* (2), 18–25.

Raviv, Artur, "The design of an optimal insurance policy," *American Economic Review*, 1979, *69* (1), 84–96.

Reinhart, Carmen M and Kenneth Rogoff, *This time is different: eight centuries of financial folly*, princeton university press, 2009.

Richter, Andreas, Jörg Schiller, and Harris Schlesinger, "Behavioral insurance: Theory and experiments," *Journal of Risk and Uncertainty*, 2014, *48* (2), 85–96.

Ross, Stephen, "Some stronger measures of risk aversion in the small and the large with applications," *Econometrica*, 1981, *49* (3), 621–638.

Rothschild, Michael and Joseph Stiglitz, "Increasing risk: I. A definition," *Journal of Economic Theory*, 1970, *2* (3), 225–243.

_ and _ , "Increasing risk II: Its economic consequences," *Journal of Economic Theory*, 1971, *3* (1), 66–84.

Sarin, R.K. and M. Weber, "Effects of ambiguity in market experiments," *Management Science*, 1993, *39* (5), 602–615.

Schlesinger, Harris, "The optimal level of deductibility in insurance contracts," *Journal of Risk and Insurance*, 1981, *48* (3), 465–481.

_ , "Risk management decisions when effectiveness is unreliable," in "Workers' Compensation Insurance: Claim Costs, Prices, and Regulation," Springer, 1993, pp. 307–328.

_ , "The theory of insurance demand," in Georges Dionne, ed., *Handbook of Insurance*, 2 ed., Springer, 2013, chapter 7, pp. 167–184.

Schmeidler, David, "Subjective probability and expected utility without additivity," *Econometrica*, 1989, *57* (3), 571–587.

Segal, Uzi and Avia Spivak, "First order versus second order risk aversion," *Journal of Economic Theory*, 1990, *51* (1), 111–125.

Sharpe, William F, "Corporate pension funding policy," *Journal of financial Economics*, 1976, *3* (3), 183–193.

_ **and Jack L Treynor**, "The principles of corporate pension finance," *The Journal of Finance*, 1977, *32* (2), 627–638.

Shleifer, Andrei and Robert W Vishny, "Liquidation values and debt capacity: A market equilibrium approach," *The Journal of Finance*, 1992, *47* (4), 1343–1366.

Smith, Vernon, "Optimal insurance coverage," *Journal of Political Economy*, 1968, *76* (1), 68–77.

Snow, Arthur, "Ambiguity and the value of information," *Journal of Risk and Uncertainty*, 2010, *40* (2), 133–145.

_ , "Ambiguity aversion and the propensities for self-insurance and self-protection," *Journal of Risk and Uncertainty*, 2011, *42* (1), 27–43.

_ **and Ronald Warren**, "Ambiguity about audit probability, tax compliance, and taxpayer welfare," *Economic Inquiry*, 2005, *43* (4), 865–871.

Stockton, Kimberly A, "Survey of defined benefit plan sponsors," *Available at Vanduard.com*, 2012.

Sundaresan, Suresh and Fernando Zapatero, "Valuation, optimal asset allocation and retirement incentives of pension plans," *The Review of Financial Studies*, 1997, *10* (3), 631–660.

Sunderam, Adi, "Money creation and the shadow banking system," *Review of Financial Studies*, 2014, p. hhu083.

Tennyson, Sharon, "Economic institutions and individual ethics: A study of consumer attitudes toward insurance fraud," *Journal of Economic Behavior & Organization*, 1997, *32* (2), 247–265.

_ **and William Warfel**, "Law and Economics of First-Party Insurance Bad Faith Liability," *Connecticut Insurance Law Journal*, 2009, *16* (1), 203–242.

Tepper, Irwin, "Taxation and corporate pension policy," *The Journal of Finance*, 1981, *36* (1), 1–13.

Tversky, Amos and Daniel Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, 1992, *5* (4), 297–323.

VanDerhei, Jack L, "The effect of voluntary termination of overfunded pension plans on shareholder wealth," *Journal of Risk and Insurance*, 1987, pp. 131–156.

Viscusi, W.K. and H. Chesson, "Hopes and fears: The conflicting effects of risk ambiguity," *Theory and Decision*, 1999, *47* (2), 157–184.

Wakker, Peter P, Daniëlle RM Timmermans, and Irma Machielse, "The effects of statistical information on risk and ambiguity attitudes, and on rational insurance decisions," *Management Science*, 2007, *53* (11), 1770–1784.

Wakker, Peter, Richard Thaler, and Amos Tversky, "Probabilistic insurance," *Journal of Risk and Uncertainty*, 1997, *15* (1), 7–28.

Watson, Willis Towers, "2015 asset allocation in Fortune 1000 pension plans," *Insider 26, Available at willistowerswatson.com.*, 2016.

Zimmer, Anja, Christian Schade, and Helmut Gründl, "Is default risk acceptable when purchasing insurance? Experimental evidence for different probability representations, reasons for default, and framings," *Journal of Economic Psychology*, 2009, *30* (1), 11–23.

_, **Helmut Gründl, Christian Schade, and Franca Glenzer**, "An Incentive-Compatible Experiment on Probabilistic Insurance and Implications for an Insurer's Solvency Level," *Journal of Risk and Insurance*, 2018, *85* (1), 245–273.