Essays on Risk Management for Insurance Companies

Xiaohu Ping
Georgia State University

Follow this and additional works at: https://scholarworks.gsu.edu/rmi_diss

Recommended Citation
doi: https://doi.org/10.57709/7366031

This Dissertation is brought to you for free and open access by the Department of Risk Management and Insurance at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Risk Management and Insurance Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
PERMISSION TO BORROW

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or publish this dissertation may be granted by the author or, in his/her absence, the professor under whose direction it was written or, in his absence, by the Dean of the Robinson College of Business. Such quoting, copying, or publishing must be solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential gain will not be allowed without written permission of the author.

Xiaohu Ping
NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University Library must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:

Xiaohu Ping
Department of Risk Management and Insurance
35 Broad Street, Atlanta, GA 30303

The director of this dissertation is:

George Zanjani
Department of Risk Management and Insurance
Georgia State University
35 Broad Street, Atlanta, GA 30303
Essays on Risk Management for Insurance Companies

BY

Xiaohu Ping

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree

Of

Doctor of Philosophy

In the Robinson College of Business

Of

Georgia State University

GEORGIA STATE UNIVERSITY
ROBINSON COLLEGE OF BUSINESS
2015
ACCEPTANCE

This dissertation was prepared under the direction of the Xiaohu Ping’s Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

Richard Phillips, Dean

DISSERTATION COMMITTEE

Dr. George Zanjani
Dr. Ajay Subramanian
Dr. Daniel Bauer
Dr. Stephen Shore
Dr. Michael Powers
ABSTRACT

Essays on Risk Management for Insurance Companies

BY

Xiaohu Ping

07/20/2015

Committee Chair: George Zanjani

Major Academic Unit: Department of Risk Management and Insurance

The dissertation includes two essays on insurer's risk management. The first essay is about how insurers use contract structure to manage their underwriting risks. Existing research on insurance contract theory emphasizes information problems and demand side issues when explaining contract structure. Supply-side factors, especially risk considerations at the insurer, have received much less attention. In this paper, we extend the optimal contracting framework of Raviv, 1979 to explore how background risk at the insurer affects optimal contract structure. We confirm earlier findings that insurer background risk may reduce risk sharing in the optimal contract. We go further to show that positive correlation between the insurer's background risk and the insured's loss can yield contract forms ruled out by the standard model, such as upper limits on coverage, and explain patterns of risk sharing not addressed in the literature, such as large deductibles in catastrophe contracts.

The second essay studies the capital allocation rule in a general portfolio optimization problem with irreversible investments. I find that marginal cost pricing can still be connected to capital allocation in this setup, although the basis for allocation is different from that found in static problems. The investment decision for an opportunity presented today is made on the basis of an expected future marginal cost of risk associated with that opportunity. The capital allocated for today's opportunity is a probability-weighted average of the product of the marginal value of capital in future states of the world and the amount of capital consumed by today's opportunity in those future states. In addition, our numerical examples show that failure to explicitly model uncertainties regarding future opportunities can lead to understatement of the marginal cost of risk, and, as a result, over-investment in current opportunities.
I would like to express my special appreciation and thanks to my advisor Professor George Zanjani, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice on both research as well as on my career have been priceless. Without your persistent guidance and assistance, this dissertation would not have been possible. I would also like to thank my committee members, Professor Ajay Subramanian, Professor Daniel Bauer, Professor Stephen Shore, and Professor Michael Powers for serving as my committee members even at hardship. I also want to thank you for letting my defense be an enjoyable moment, and for your brilliant comments and suggestions, thanks to you.

It is a great honor to be a Ph.D. student in one of the best Risk Management and Insurance departments in the world at Georgia State University. I have learned and benefited more than I could ever imagine. I would like to thank all the faculty and staff in the department for their kindness. I would especially like to thank the dean of Robinson College: Professor Richard Phillips, the department chair: Professor Martin Grace, the Ph.D. coordinators: Professor George Zanjani and Professor Stephen Shore, and all the professors in the Department of Risk Management and Insurance. All of you have been there to support me without any hesitation when had problems with research and when I had questions about the job market. Possible errors and shortcomings in this dissertation are my responsibility.

A special thanks to my family. Words cannot express how grateful I am to my mother and father, for all of the sacrifices that you’ve made on my behalf. Your prayer for me was what sustained me thus far. I would also like to thank all of my friends who supported me in writing, and encouraged me to strive toward my goal.
## CONTENTS

ABSTRACT .............................................................................................................................. vi

ACKNOWLEDGMENTS ........................................................................................................ vii

CONTENTS ............................................................................................................................. viii

LIST OF FIGURES ................................................................................................................ ix

Chapter 1: Optimal Insurance Contracts with Insurer's Background Risk .................. 1

  1.1 Introduction ..................................................................................................................... 1

  1.2 Related Literature .......................................................................................................... 5

  1.3 Optimal Contracts with Insurer Background Risk ....................................................... 8

  1.4 Applications - CARA Utilities and Specific Loss Distributions .............................. 25

  1.5 Conclusion .................................................................................................................... 30

  1.6 Appendix ....................................................................................................................... 31

  1.7 References .................................................................................................................... 49

Chapter 2: Dynamic Capital Allocation with Irreversible Investments ....................... 59

  2.1 Introduction ..................................................................................................................... 59

  2.2 Dynamic Capital Allocation ......................................................................................... 61

  2.3 Numerical Example ....................................................................................................... 69

  2.4 Conclusion ..................................................................................................................... 78

  2.5 References .................................................................................................................... 79
LIST OF FIGURES

Figure 1.1 Optimal Indemnity Schedules when $c'(\cdot) = 0$ ........................................54
Figure 1.2 Optimal Indemnity Schedules when $c'(\cdot) > 0$ ........................................54
Figure 1.3 Nondecreasing Optimal Indemnity Schedules when $c'(\cdot) = 0$ ............55
Figure 1.4 Nondecreasing Optimal Indemnity Schedules when $c'(\cdot) > 0$ ............55
Figure 1.5 Optimal Indemnity Structure with Independent Background Risk and No
Indemnity Cost .................................................................56
Figure 1.6 Optimal Indemnity Structure with Independent Background Risk and
Increasing Indemnity Cost ..........................................................56
Figure 1.7 Optimal Indemnity Structure with Dependent Background Risk and No
Indemnity Cost .................................................................57
Figure 1.8 Optimal Indemnity Structure with Dependent Background Risk and Increasing
Indemnity Cost .................................................................57
Figure 1.9 Optimal Indemnity Structures with Different Dependence Structures with Cost
$(c = 0)$ ........................................................................58
Figure 1.10 Optimal Indemnity Structures with Different Dependence Structures with
Cost $(c = 0)$ ........................................................................58
Figure 2.1 $q_0$ depending on $\Phi_{\mu}$ in 2-period .....................................................72
Figure 2.2 $q_1$ depending on $\Phi_{\mu}$ in 2-period .....................................................72
Figure 2.3 $q_0$ depending on $\Phi_{\mu}$ in 3-period .....................................................73
Figure 2.4 $q_1$ depending on $\Phi_{\mu}$ in 3-period .....................................................73
Figure 2.5 $q_2$ depending on $\Phi_{\mu}$ in 3-period .....................................................74
Figure 2.6 $q_0$ depending on $\Phi_\sigma$ in 2-period ..................................................75

Figure 2.7 $q_1$ depending on $\Phi_\sigma$ in 2-period ..................................................76

Figure 2.8 $q_0$ depending on $\Phi_\sigma$ in 3-period ..................................................77

Figure 2.9 $q_1$ depending on $\Phi_\sigma$ in 3-period ..................................................77

Figure 2.10 $q_2$ depending on $\Phi_\sigma$ in 3-period ..................................................77
Chapter 1: Optimal Insurance Contracts with Insurer’s Background Risk

1 Introduction

Insurance contracts have well-known risk-sharing features, such as deductibles, upper limits, and coinsurance. Existing theoretical research emphasizes information problems and demand side issues when explaining these features and other aspects of insurance contract structure. In this paper, we argue that supply-side factors, especially risk management considerations at the level of the insurer, are also important for understanding contract structure and how risk-sharing features vary across markets.

Specifically, we extend the optimal contracting framework pioneered by Raviv (1979) to the case where the risk-averse insurer has random wealth and consider how correlation between the insurer’s wealth and the insured’s loss affects the structure of the optimal insurance contract. Negative correlation between the insurer’s wealth and the insured’s loss could emerge for a variety of reasons, but an important one concerns undiversifiable risk within the insurance market. When loss frequency and severity varies systematically within the insured population, a new contract will involve potential losses that are positively correlated with the insurer’s existing exposures.
We confirm earlier findings by showing that randomness in the insurer’s background wealth may reduce risk-sharing in the optimal contract when the randomness is negatively correlated with the insured’s loss (see Dana and Scarsini (2007) and Biffis and Millossovich (2012)). We go further to show that such correlation expands the set of possible indemnity schedules beyond those admitted under Raviv’s analysis: Namely, upper limits\(^1\) on coverage—which were ruled out in Raviv’s analysis in the absence of regulation—are possible in the presence of correlation. We also show that differences in the degree of correlation may explain differences in other contract features: Deductibles and coinsurance percentages, for example, are shown to increase with the degree of correlation under certain conditions.

These findings are important because existing theories do not fully explain observed patterns of risk sharing in insurance markets. To illustrate, consider deductibles and upper limits.

A typical deductible in a U.S. Homeowner’s contract is measured in hundreds of dollars and applies to many perils, including fire; the typical deductible for an Earthquake policy on the same home, however, is ten to fifteen percent of the dwelling limit—which, even for a modest home, easily pushes beyond $10,000. Common explanations for deductibles—administrative costs (Arrow (1963), Raviv (1979), and Townsend (1979)), adverse selection (Rothschild and Stiglitz (1976)), and moral hazard (e.g., Zeckhauser (1970))—seem to hold little promise in this case. For example, it is hard to imagine how insuring earthquake instead of fire would generate administrative cost differences capable of explaining such a huge difference in deductibles; nor is it obvious why asymmetric information issues would be any worse with earthquakes than with fire. On the other hand, one obvious difference between the two perils is the correlation in consumer losses: earthquake is widely regarded as a risk difficult

\(^{1}\)We use “upper limit” here in the sense of a maximum indemnity that 1) is less than the largest possible loss and 2) is awarded for at least some loss(es) that are less than the largest possible loss. This corresponds to the colloquial notion of a “policy limit” and Raviv’s description of a “ceiling.”
to diversify across consumers, and this situation can obviously lead to a correlation between the insured’s loss and the insurer’s other liabilities on similar contracts.

The upper limit is another ubiquitous contract characteristic that defies easy explanation. Arrow’s early analysis set full insurance beyond an initial deductible as a theoretical baseline for the indemnity schedule, and, though later refinements such as Raviv’s expanded the set of possible schedules, the upper limit remained elusive. Huberman, Mayers, and Smith (1983) observed that limited liability could induce consumers to demand policy limits. The seems an intuitive explanation in liability lines, where the loss exposure can exceed the insured’s wealth; however, policy limits are also common in property lines, where the the loss will typically fall within the insured’s wealth (if broadly defined to include property wealth). Doherty, Laux, and Muermann (2013) raise the possibility that upper limits arise from the incomplete nature of the insurance contract, with the insurer setting the policy limit at the level where it can credibly commit to paying—with the incentive for paying being the retention of the insured’s business. We raise here a third possibility, which is that the reason for the limit stems from the insurer’s need to manage its risk and limit exposure to correlated losses.

The rest of this paper is organized as follows. Section 2 reviews the literature. In Section 3, we introduce the basic model structure of Raviv (1979), extended to analyze randomness in the wealth of the insurer. When the insurer’s wealth is uncorrelated with the insured’s loss, we obtain results similar to those of Raviv: Upper limits are not possible, and variable production costs will lead the optimal contract to feature a deductible. We then 1) introduce stochastic dependence between the insurer’s background wealth and the insured’s loss and 2) constrain the indemnity schedule to be non-decreasing. These changes expand the set of possible optimal contract forms: In particular, upper limits now become possible.

In Section 4, we specialize to the case of CARA utility and examine several specific
loss distributions—the Normal, Bernoulli, and (truncated) Lognormal. In the case of the Normal and Bernoulli distributions, we can verify analytically that consumer retention increases with (negative) correlation between the consumer loss and the insurer’s wealth and specifically that the deductible increases with the strength of the correlation. In the case of the Normal distribution, the indemnity schedule takes a linear form, and, in addition to the deductible increasing, the slope of the indemnity decreases as the correlation between the insured loss and the insurer’s losses become greater. A similar result is shown, numerically, in examples using the truncated joint Lognormal distribution. Specifically, we show that greater correlation generates larger deductibles and greater coinsurance in the sense of a lower slope for the indemnity schedule.

Section 5 concludes with a discussion of the implications of the findings and a sketch of ideas for future research.
2 Related Literature

This paper is related to the literature on optimal insurance contracts originated by Arrow (1963) and later developed by Raviv (1979) and Townsend (1979). These papers derive optimal contracts with little restriction on the contract form and identify administrative cost as the driver of deductibles. In addition to fundamental models which endogenize contract form, a number of papers study optimization in environments with particular contractual forms, such as deductible, coinsurance, or upper limits. For example, Gould (1969), Pashigian, Schkade, and Menefee (1966), Moffet (1977), and Schlesinger (1981) have considered the optimal choice of a deductible level; Mossin (1968) and Mayers and Smith (1983) have discussed the demand for coinsurance; Smith (1968) and Cummins and Mahul (2004) have analyzed the demand for insurance with an upper limit.

Deductibles and coinsurance can arise endogenously from asymmetric information between the insured and the insurer. For examples, Zeckhauser (1970), Spence and Zeckhauser (1971), Rothschild and Stiglitz (1976), Harris and Raviv (1978), and Holmstrom (1979) find various forms of risk-sharing in insurance contracts under conditions of asymmetric information. The risk-sharing features can serve either as incentive devices (to restrain the moral hazard of the insured) or as sorting devices under conditions of adverse selection.

Fewer papers address the reasoning for upper limits, with exceptions noted above being Huberman, Mayers, and Smith (1983) and Doherty, Laux, and Muermann (2013). While the literature on insurance under state-dependent utility (see Arrow (1974) and Cook and Graham (1977)) has been aimed at overinsurance or underinsurance more generally and is typically motivated in the context of life and health insurance, it is likely that state-dependence or other preference alterations could imply upper limits: Indeed, the limited liability model of Huberman, Mayers, and Smith (1983) could be interpreted as a form of preference alteration.
Another strand of literature on optimal insurance contracts emphasizes incomplete markets. Various forms of incompleteness analyzed include background risk, insolvency risk, and counterparty default risk. The effect of background risk on insurance contracts has been analyzed in many papers. A number impose a deductible contract structure, including Turnbull (1983), Doherty and Schlesinger (1983b), and Doherty and Schlesinger (1983a). Gollier (1996), Mahul (2000), Vercammen (2001), and Dana and Scarsini (2007) investigate the effect of insured’s background risk in models with endogenous contract structures, with the common conclusion that the insured’s background risk may result in a disappearing deductible. Other papers about incompleteness in optimal insurance are Schlesinger and Doherty (1985), Doherty and Schlesinger (1990), Doherty and Dionne (1993), Kaplow (1994), Gollier and Schlesinger (1995), Hau (1999), and Cummins and Mahul (2003).

Although most of the existing literature has emphasized demand side explanations for insurance contract structure, some papers have considered the supply side. Even early work on contract structure (e.g., Arrow (1974)) noted that coinsurance (risk sharing between insured and insurer) could be explained by insurer risk aversion. Less attention, however, has been paid to the question of how background risk at the insurer affects contract structure, with two noteworthy exceptions—Dana and Scarsini (2007) and Biffis and Millosovich (2012). While their settings are different (and both mostly focus on the consequences of background risk at the insured), both observe that under certain conditions a negative relationship between the insurer’s wealth and the size of the loss can produce a nondecreasing retention schedule. That is, the amount the insured retains, either through deductible or through coinsurance of some kind, is rising in the amount of the loss.

The foregoing characterization offers a useful starting point but of course admits a wide variety of possibilities—including full insurance after a deductible, coinsurance,
full insurance up to a policy limit, and so on. The purpose of this paper is to build on this idea, but in a more restricted setting in terms of contracting possibilities and, in some cases, preferences. Thus, we sacrifice generality in terms of characterizing the optimal contract in order to offer intuition on the impact of insurer background risk on certain features of real-world contracts, such as large deductibles and upper limits.
3 Optimal Contracts with Insurer Background Risk

3.1 Basic Model Structure

Following Raviv (1979), we consider the problem of finding the optimal insurance policy for a consumer subject to an insurer participation constraint, with our focus here being the case when an insurance buyer’s loss is correlated in some way with insurer’s wealth. Both the insurer and the insured are assumed to be risk averse. We start by imposing the standard constraints on contracting used by Raviv (1979)—that the indemnity must be no less than zero and no more than the amount of the loss. (In a later section we will introduce further constraints on contracting possibilities.)

Consider a risk averse insured who has an utility function $U : \mathbb{R} \rightarrow \mathbb{R}$, which is assumed to be increasing, strictly concave (i.e., $U' > 0$ and $U'' < 0$), and twice differentiable. The insured has a non-random initial wealth $W_1$ and faces the risk of a random loss $X$. Assume the random insurable loss $X$ has a cumulative distribution function $F(x)$ and the corresponding density function $f(x)$ on the domain $[0, \bar{x}]$.

The insurance contract is characterized by an indemnity function $I(X)$ and a premium $P$. The indemnity, or coverage, represents what the insurer will pay when loss $X$ occurs, and the premium $P$ is paid by insured for the coverage schedule $I(X)$. Following the literature, we assume the feasible insurance contract has a nonnegative premium and an indemnity that is nonnegative and cannot exceed the insured’s loss size, as in:

$$P \geq 0 \quad \text{and} \quad 0 \leq I \leq Id,$$

---

3 The insured’s initial wealth could be random when some risks are uninsurable. All the insurable risk is in $X$, while the insured’s initial wealth $W_1$ captures the uninsurable risks. The correlation between the insurable risk and uninsurable (background) risks is analyzed by Gollier (1996) and Dana and Scarsini (2007), among others.

4 Gollier (1987) analyzed the optimal insurance without the nonnegative indemnity constraint and showed the negative indemnity strategy is dominated by deductible. Nevertheless, if we consider risk transfer in a broader context, the indemnity may violate the standard constraints (e.g., in swaps the agent’s cash flows may go in both directions.).
where $Id$ denotes the identity function $Id(x) = x$. Due to the administrative expenses associated with the indemnity payment, we assume a nondecreasing and convex cost function $c(I)$ with

$$c(0) = 0, \quad c'(\cdot) \geq 0, \quad \text{and} \quad c''(\cdot) \geq 0.$$  

The insurance cost is borne by the insurer and depends on insurer's payout when insured loss $X$ is realized.

The insured wants a contract that maximizes her expected utility of ex post wealth $A = W_1 - X - P + I(X)$,\footnote{We assume that the insured has no default option. The case of a policyholder default option is considered by Huberman, Mayers, and Smith (1983).} which depends on the insurance contract $(P, I(\cdot))$ and the realization of the insured loss $X$. Thus, the insurance buyer is willing to purchase coverage $I(X)$ for a premium $P$ only if

$$\mathbb{E}[U(W_1 - X - P + I(X))] \geq \mathbb{E}[U(W_1 - X)].$$

On the supply side, we assume that the insurer is also risk averse, has an initial wealth $W_2$, and faces a random background loss $Y$ in addition to the random indemnity paid to the contracting insured. The insurer’s utility function $V : \mathbb{R} \mapsto \mathbb{R}$ is increasing, concave (i.e., $V' > 0$ and $V'' \leq 0$), and twice differentiable. Since we don’t allow the insurer to default,\footnote{The effects of incorporating insurer default are considered by Schlesinger and Schulenburg (1987), Doherty and Schlesinger (1990), and Biffis and Millossovich (2012).} its ex post wealth is $B = W_2 - Y + P - I(X) - c(I(X))$ which also depends on the insurance contract $(P, I(\cdot))$ and the realization of losses $X$ and $Y$. The insurer’s participation constraint is

$$\mathbb{E}[V(W_2 - Y + P - I(X) - c(I(X)))] \geq V \geq \mathbb{E}[V(W_2 - Y)],$$

in which $V$ denotes the insurer’s reservation utility level, which is no less than the insurer’s expected utility of its initial wealth (inclusive of the background risk).
The background loss $Y$ reflects the insurer’s background risk. We assume that random losses $X$ and $Y$ have a joint cumulative distribution function $H(X,Y)$ and a corresponding joint density $h(x,y)$ with the support $[0, x] \times [0, y]$. We denote the conditional density of $Y$ given $X = x$ by $g(y|x)$ and the corresponding cumulative distribution function by $G(Y|x)$ on the support $[0, y]$. Let $f(x)$ and $g(y)$ be the marginal densities of the insured and insurer’s losses respectively. If the insurer’s random loss $Y$ is independent of the insured’s loss $X$, then the marginal density is equal to the conditional density, i.e., $g(y|x) = g(y), \forall x \in [0, x]$.

To find the optimal insurance contract, we look for the insurance premium $P$ and the indemnity function $I(\cdot)$ that maximizes the insured’s expected utility of final wealth subject to participation constraints (for both the insured and the insurer) and a constraint restricting the indemnity to fall between zero and the amount of the loss:

$$
\max_{P,I(\cdot)} \mathbb{E}[U(W_1 - X - P + I(X))] 
$$

subject to

$$
\mathbb{E}[U(W_1 - X - P + I(X))] \geq \mathbb{E}[U(W_1 - X)].
$$

$$
\mathbb{E}[V(W_2 - Y + P - I(X) - c(I(X))))] \geq \underline{V} (3)
$$

$$
P \geq 0
$$

$$
0 \leq I(x) \leq x \quad \forall x \in [0, \overline{x}] (5)
$$

where $\underline{V}$ denotes the insurer’s reservation utility level and $\underline{V} \geq \mathbb{E}[V(W_2 - Y)]$. Note, following Gollier (1996), we use certain reserved utility level $\underline{V}$ instead of the autarky utility level $\mathbb{E}[V(W_2 - Y)]$, since it doesn’t affect the risk aversion behavior of the insurer.

We solve this problem in two steps: first, we solve the maximization problem for
a given premium $P$; second, we find the optimal $P^*$ and the corresponding optimal indemnification schedule $I^*$. If $(P^*, I^*)$ is an optimal contract, then there exists a multiplier $\lambda \geq 0$ such that $(P^*, I^*)$ solves

$$
\max_{P \geq 0, 0 \leq I \leq Id} \mathbb{E}[U(W_1 - X - P + I(X))] + \lambda \mathbb{E}[V(W_2 - Y + P - I(X) - c(I(X)))].
$$

Therefore, at the optimal premium $P^*$, for every $x \in [0, \overline{x}]$, $I^*(x)$ is the solution of a state-by-state maximization problem

$$
\max_{0 \leq I(x) \leq x} U(W_1 - x - P^* + I(x)) + \lambda \mathbb{E}[V(W_2 - Y + P^* - I(x) - c(I(x)))] | X = x.
$$

Before examining the solution, we now introduce some definitions of insurance contracts and commonly discussed indemnity schedules. With the feasible indemnity constraint (5), insurance coverage may be characterized according to three possibilities: no-insurance, full-insurance, and coinsurance. “No-insurance” refers to the portions of the indemnity schedule with zero payout; “full-insurance” refers to the portions of the indemnity schedule with full coverage for any corresponding loss; “coinsurance” refers to the portions of the indemnity schedule where part (but not all) of the corresponding loss is covered. Thus partitioning the loss interval $[0, \overline{x}]$ into three subsets, we have the following three possible coverage sets: no-insurance set, full-insurance set, and coinsurance set.

**Definition 1.** The subset $\Omega_1 = \{x \in [0, \overline{x}] | I(x) = 0\}$ denotes the no-insurance set; the subset $\Omega_2 = \{x \in [0, \overline{x}] | I(x) = x\}$ denotes the full-insurance set; and the subset $\Omega_3 = \{x \in [0, \overline{x}] | 0 < I(x) < x\}$ denotes the coinsurance set.

As mentioned in the introduction, deductibles and upper limits are well-known fea-

---

[7] Coinsurance in this paper refers to all situations except for full-insurance and no-insurance, while in other studies coinsurance is restricted to the contract where risk transfer $I$ and retention $Id - I$ are both nondecreasing, e.g., Dana and Scarsini (2007).
tures of insurance contracts. A deductible is the maximum level of the no-insurance set starting from zero loss, and an upper limit is the level above which indemnity stops increasing. In addition, there may be a “full insurance upper limit”, which is the maximum level of the full-insurance set starting from zero loss.\footnote{We inherit the terminology of Raviv (1979), using “full insurance upper limit” as the maximum level of full insurance, and use “upper limit” to define the level after which indemnity stops increasing.} The formal definitions of these critical loss thresholds in insurance contracts are given as follows.

**Definition 2.** A loss amount $x_d$ is called a **deductible** if $\forall x \in [0, x_d], I(x) = 0$, and $\forall \delta > 0, \exists \epsilon \in (0, \delta)$ such that $I(x_d + \epsilon) > 0$. A loss amount $x_u$ is called an **upper limit** if $\forall x \in [0, \bar{x}], I(x) \leq I(x_u)$ and $x_u < \bar{x}$. A loss amount $x_f$ is called a **full insurance upper limit** if $\forall x \in [0, x_f], I(x) = x$, and $\forall \delta > 0, \exists \epsilon \in (0, \delta)$ such that $I(x_f + \epsilon) < x + \epsilon$.

Moreover, there are several insurance contract structures widely studied in the literature—generalized deductible contracts, standard deductible contracts, disappearing deductible contracts, increasing-convex deductible contracts, generalized coinsurance contracts, and standard coinsurance contracts. We recall the definitions of these indemnity schedules as follows.

**Definition 3.** (i) A feasible contract $(P, I)$ has a **generalized deductible** if for some $x_d \in [0, \bar{x}]$ it satisfies $I(x) = 0$, $\forall x \in [0, x_d]$.

(ii) A feasible contract $(P, I)$ has a **standard deductible** if for some $x_d \in [0, \bar{x}]$ it satisfies $I(x) = \max (x - x_d, 0)$, $\forall x \in [0, \bar{x}]$.

(iii) A feasible contract $(P, I)$ has a **disappearing deductible** if there exists a nontrivial deductible level $x_d$, and the risk retention function $I_d - I$ is decreasing on $[x_d, \bar{x}]$.

(iv) A feasible contract $(P, I)$ has an **increasing-convex deductible** if there exists a nontrivial deductible level $x_d$, and the risk retention function $I_d - I$ is increasing and strictly convex on $[x_d, \bar{x}]$.

(v) A feasible contract $(P, I)$ is called a **generalized coinsurance** if it satisfies
A feasible contract \((P, I)\) is called a **standard coinsurance** if it satisfies 
\[0 < I(x) < x, \quad \forall x \in (0, \bar{x}],\]

(vi) The definition of the generalized deductible contract is different from that used in Dana and Scarsini (2007) in that we do not require nondecreasing indemnity \(I\) and retention \(Id - I\). In a disappearing deductible contract, the indemnity function is increasing after the deductible; for an additional unit of loss, insurer’s additional payout is more than one unit if it is positive; and a full-insurance interval may be included. In an increasing-convex deductible contract, the indemnity function may be decreasing and get to no-insurance in some intervals; for an additional unit of loss, insurer’s additional payout is less than one unit, and an upper limit may exist.

We now analyze how the nature of the relation between the losses \(X\) and \(Y\) affects the form of the optimal contract. We start in Section 3.2 with the case where the background loss of the insurer, \(Y\), is independent of the insured’s loss \(X\). We then move on in Section 3.3 to the case where the two losses exhibit dependence.

### 3.2 Independent Insurer Background Risk

Consider first the case with independent insurer background risk, i.e., \(Y\) and \(X\) are independent. If we define the insurer’s indirect utility function \(\hat{V} : \mathbb{R} \mapsto \mathbb{R}\) as \(\hat{V}(W) = \mathbb{E}[V(W - Y)]\) for all possible \(W\), \(\hat{V}\) inherits the properties of \(V\) and problem (1) turns out to be identical to the one studied by Raviv (1979), where the insurer’s utility function \(V\) is replaced by \(\hat{V}\). Define \(R_U(\cdot)\), \(R_V(\cdot)\), and \(R_{\hat{V}}(\cdot)\) as the Arrow-Pratt coefficients of absolute risk aversion associated with utility functions \(U\), \(V\), and \(\hat{V}\), separately. Specifically, \(R_{\hat{V}}(\cdot)\) is the insurer’s coefficient of absolute risk aversion, in Kihlstrom et al.’s sense,\(^9\) associated with utility function \(V\). Further, define \(\hat{B}^* = \)

---

\(^9\)Instead of the Arrow-Pratt measure of absolute risk aversion, the coefficient of absolute risk aversion in Kihlstrom et al.’s sense is commonly used in random initial wealth problems (see Kihlstrom,
\[ W_2 + P^* - I^*(x) - c(I^*(x)) \]
as the final wealth associated with the indirect utility function \( \hat{V} \) and the insurance contract \((P^*, I^*)\). Hence, the Arrow-Pratt coefficient of absolute risk aversion for the indirect utility function \( \hat{V} \) is

\[
R_{\hat{V}}(\hat{B}^*) = -\frac{\hat{V}''(\hat{B}^*)}{\hat{V}'(\hat{B}^*)} = -\frac{\mathbb{E}[V''(B^*)]}{\mathbb{E}[V'(B^*)]},
\]

where \( B^* = W_2 - Y + P^* - I^*(x) - c(I^*(x)) \) is the final wealth associated with utility function \( V \). Moreover, we can define \( c^* = c'(I^*(x)) \) evaluated at the optimal indemnity level \( I^* \), and \( A^* = W_1 - x - P^* + I^*(x) \) as the insured’s final wealth with the optimal insurance contract evaluated at insured’s loss \( x \). Note that the value of \( B^* \) is random due to the randomness of insurer’s loss \( Y \) even if insured’s loss \( X \) is known to be \( x \).

As described in the following theorem, the optimal contract has properties similar to that in the standard case without insurer background risk (i.e., Raviv (1979)):

**Theorem 1.** Assume \( X \) and \( Y \) are independent, then the optimal insurance contract exists and any optimal contract \((P^*, I^*)\) is such that \( I^* \) and \( Id - I^* \) are nondecreasing.

(a) No insurance \((P^* = 0, I^* = 0)\) is an optimal contract if and only if

\[
1 + c'(0) \geq \frac{U'(W_1 - \overline{X})}{\mathbb{E}[U'(W_1 - X)]}. \tag{6}
\]

(b) Full insurance \((P^* > 0, I^* = Id)\) is an optimal contract if and only if

\[
1 + c'(\overline{X}) \leq \frac{\mathbb{E}[V'(W_2 - Y + P^* - X - c(X))]}{\mathbb{E}[V'(W_2 - Y + P^* - \overline{X} - c(\overline{X}))]}, \tag{7}
\]

which is equivalent to \( c'(\cdot) = 0 \) and \( V''(\cdot) = 0 \).

(c) Suppose \((P^*, I^*)\) is an optimal contract, then for all \( x \in \Omega_3 \), we have

\[
I''(x) = \frac{R_U(A^*)}{R_U(A^*) + R_{\hat{V}}(\hat{B}^*)(1 + c^*) + c'''/(1 + c^*)}. \tag{8}
\]
(d) The optimal contract has no deductible, $x_d = 0$, if and only if $c'(\cdot) = 0$.

(e) If $V''(\cdot) = 0$ and constraint (6) is not satisfied, then we have:

1. If $c'(\cdot) = 0$, then $(P^*, I^*)$ is a full insurance contract.
2. If $c'(\cdot) = m > 0$, then $(P^*, I^*)$ is a standard deductible contract.
3. If $c'(\cdot) > 0$ and $c''(\cdot) > 0$, then $(P^*, I^*)$ is a generalized deductible contract.

(f) If $V''(\cdot) > 0$ and constraint (6) is not satisfied, then we have:

4. If $c'(\cdot) = 0$, then $(P^*, I^*)$ is a standard coinsurance contract.
5. If $c'(\cdot) > 0$, then $(P^*, I^*)$ is a generalized deductible contract.

Proof. See the Appendix 6.1.

As Raviv showed in the case without background risk, when the insurer’s background risk is independent, the optimal contract exists and makes the insurer’s and insured’s wealth comonotone. In other words, the optimal indemnity and risk retention are both nondecreasing with respect to the insured loss. In addition to comonotonicity, the form of the optimal contract remains the same as that in the case of no background risk. Of particular note are the effects of insurer risk aversion, which yield risk-sharing in the form of coinsurance even in the absence of administrative costs. In addition, deductibles appear exclusively as a consequence of an increasing administrative cost function.

We now examine the relation between the optimal contract and the size of insurer’s background risk. Compare two cases when the insurer has different independent background risks: $Y$ and $\tilde{Y}$, where $Y \sim G(y)$ second-order stochastically dominates $\tilde{Y} \sim \tilde{G}(y)$, or, roughly speaking, $Y$ involves less “risk” than $\tilde{Y}$.\(^{10}\) Of course, the size of the independent background risk does not alter the form of the optimal insur-

\(^{10}\)By the definition of second order stochastic dominance, we know that $\mathbb{E}[U(\tilde{Y})] \leq \mathbb{E}[U(Y)]$ for any nondecreasing and concave function $U(\cdot)$.
Corollary 1. Assume the independent insurer’s background risk is $\tilde{Y} \sim \tilde{G}(y)$, which is second-order stochastically dominated by $Y$. Then the optimal contract $(\tilde{P}^*, \tilde{I}^*)$ has the same properties as those in Theorem 1. Moreover,

(a) If $V''(\cdot) = 0$, then $(\tilde{P}^*, \tilde{I}^*) = (P^*, I^*)$.

(b) If the insurer exhibits a constant absolute risk aversion, then $(\tilde{P}^*, \tilde{I}^*) = (P^*, I^*)$.

(c) If the insurer exhibits a decreasing absolute risk aversion (DARA), then we have:

1. If $c'(\cdot) = 0$, then insurer’s risk vulnerability leads to $\tilde{I}^*(x) \leq I^*(x)$, $\forall x \in [0, \bar{x}]$.

2. If $c'(\cdot) = m > 0$, the deductible level and risk sharing fraction are ambiguous.

Proof. See Appendix 6.2.

To summarize, if the insurer is risk neutral or has constant absolute risk aversion, then variation in the size of the insurer’s independent background risk will not affect the optimal indemnity schedule. If the insurer is risk vulnerable and the indemnity cost is constant, then the optimal contract form remains a standard coinsurance, but has a lower indemnity for any insured loss. The situation is more complicated when indemnity costs are increasing: although the optimal contract form remains a generalized deductible, the directions of change in the deductible level and risk sharing fraction are ambiguous.

11 Analogous results are found for the insured’s background risk in Doherty and Schlesinger (1983b), Gollier (1996), and Mahul (2000), who compare the optimal insurance contract when the insured has zero-mean background risk to that obtained when the insured has no background risk.
3.3 Dependent Insurer Background Risk

In general, the insurer’s background risk will exhibit some form of dependence with the insured’s loss. One can imagine a number of scenarios that could lead to different relationships between the insured’s loss and the random portion of the insurer’s wealth derived from sources other than the insured’s loss. For example, the insurer’s background loss could be negatively correlated with insured’s loss in the situations where insurer has reinsured the underwritten risk or (over)invested in hedging instruments (e.g., catastrophe bonds, contingent surplus notes, and loss equity puts), all of which could make the insurer’s other wealth correlate positively with the specific insured loss. On the other hand, the insurer’s background loss could be positively correlated with insured’s loss if the insured’s loss is positively correlated with the losses on other contracts, which might happen in a variety of different lines of insurance, with the most extreme positive correlation coming in lines with catastrophe exposure such as earthquake insurance and flood insurance.

We now consider different possible dependence structures and find that risk sharing and the deductible depend on the specific nature of the dependence between $Y$ and $X$. Some dependence concepts are recalled in the following definitions.

A random vector $Y$ is associated with another vector $X$ if any two nondecreasing functions of them have a positive covariance.

Definition 4. Two random variables $X$ and $Y$ are associated, if $\text{Cov}(g(X,Y),k(X,Y)) \geq 0$ for all pair of nondecreasing functions $g$ and $k$ such that the covariance exists.

A random vector $Y$ is stochastically increasing in another vector $X$ if the conditional distribution of $Y$ becomes larger in the sense of first order stochastic dominance when conditioning on higher values of $X$.

---

12 The insured’s loss could even raise the insurer’s wealth in the case when the insurer is invested in hedging investments or has access to post-loss financing tools. This situation is investigated by Biffis and Millossovich (2012).
Definition 5. A random vector $Y$ is (strictly) **stochastically increasing** in random vector $X$, if the map $x \mapsto \mathbb{E}[z(Y)|X = x]$ is nondecreasing (increasing) for every nondecreasing function $z(\cdot)$ such that $\mathbb{E}[z(Y)]$ exists.

In addition, there is a stronger form of positive dependence for joint random variables called **affiliation** (monotone likelihood ratio property for bivariables). High values of some of the components of a random vector make the other components more likely to be high than small.

Definition 6. Two random variables $X$ and $Y$ with a twice differentiable joint density function $h(x, y)$ are **affiliated** if $h(x, y)$ is log-supermodular, i.e., if $h(y, z)$ satisfies

$$
\frac{\partial^2 \ln h(x, y)}{\partial x \partial y} \geq 0 \quad \text{for all } x, y \in \mathbb{R}^2.
$$

(9)

Affiliation implies that the semi-elasticity of the conditional density $g(y|x)$ with respect to one component $x$ is nondecreasing in the other component $y$. It is a stronger property than the stochastically increasing property, since it restricts not only the change in the average of the conditional distribution but also its shape. Moreover, we know that stochastic affiliation implies that each vector component is stochastically increasing with respect to the others, which in turn implies association.

3.3.1 Optimal Contracting in General

First, we consider the circumstance without constraining the indemnity schedule to be nondecreasing. Suppose there is a nontrivial optimal insurance contract, $(P^*, I^*) \neq (0, 0)$, then there exists a $\lambda \geq 0$ such that

$$
\mathbb{E}[U'(W_1 - X - P^* + I^*(X))] = \lambda \mathbb{E}[V'(W_2 - Y + P^* - I^*(X) - c(I^*(X)))]
$$

(10)
for all \( x \in \Omega_1 \),

\[
J(x) \equiv U'(W_1 - x - P^*) - \lambda \mathbb{E}[V'(W_2 - Y + P^*)|X = x](1 + c'(0)) \leq 0,
\]

(11)

for all \( x \in \Omega_2 \),

\[
K(x) \equiv U''(W_1 - P^*) - \lambda \mathbb{E}[V'(W_2 - Y + P^* - x - c(x))|X = x](1 + c'(x)) \geq 0,
\]

(12)

and for all \( x \in \Omega_3 \),

\[
M(x) \equiv U'(A^*(x)) - \lambda \mathbb{E}[V'(B^*(Y, x))|X = x](1 + c^*) = 0,
\]

(13)

where \( c^* = c'(I^*(x)) \) is evaluated at the optimal indemnity level \( I^* \), and

\[
A^*(x) = W_1 - x - P^* + I^*(x)
\]

\[
B^*(Y, x) = W_2 - Y + P^* - I^*(x) - c(I^*(x))
\]

are the insured’s and the insurer’s final wealth under the optimal contract evaluated at loss \( x \in \Omega_3 \). Note that \( B^* \) is random due to the randomness of the insurer’s loss \( Y \) even if the insured’s loss \( X \) is known to be \( x \). Similarly, we recall the indirect utility function \( \hat{V} \) and the associated coefficient of risk aversion \( \hat{R}_{\hat{V}}(\hat{B}^*(x)) \). The optimal contract in the presence of insurer background risk has the following properties.

**Theorem 2.** Assume \( Y \) and \( X \) are affiliated, then for any optimal insurance contract \( (P^*, I^*) \), we have

(a) The optimal risk retention \( Id - I^* \) is nondecreasing.

(b) No insurance \( (P^* = 0, I^* = 0) \) is optimal if and only if

\[
1 + c'(0) \geq \sup_{x \in [0, x]} \left\{ \frac{\mathbb{E}[V'(W_2 - Y)]}{\mathbb{E}[U'(W_1 - X)]} \cdot \frac{U'(W_1 - x)}{\mathbb{E}[V'(W_2 - Y)|X = x]} \right\}.
\]
(c) Full insurance \((P^* > 0, I^* = \text{Id})\) is optimal if and only if
\[
1 + c'(\bar{x}) \leq \frac{\mathbb{E}[V'(W_2 - Y + P^* - X - c(X))] - \mathbb{E}[V'(W_2 - Y + P^* - \bar{x} - c(\bar{x}))|X = \bar{x}]}{\mathbb{E}[V'(W_2 - Y + P^* - \bar{x} - c(\bar{x}))|X = \bar{x}]}.
\]

(d) For all \(x \in \Omega_3\),
\[
I''(x) = \frac{R_U(A^*) - \Psi(x)}{R_U(A^*) + R_V(\hat{B}^*)(1 + c''/c'/1)}.
\]  

(14) 

where
\[
\Psi(x) = \frac{\mathbb{E}[V'(B^*)(\partial \ln g(Y|x)/\partial x)]}{\mathbb{E}[V'(B^*)]} = \frac{\text{Cov}[V'(B^*), (\partial \ln g(Y|x)/\partial x)]}{\mathbb{E}[V'(B^*)]} \geq 0
\]

and
\[
R_V(\hat{B}^*) = -\frac{\mathbb{E}[V''(B^*)]}{\mathbb{E}[V'(B^*)]}.
\]

(e) If \(V''(\cdot) = 0\), then the optimal contract is the same as that in Theorem 1.

(f) If \(V''(\cdot) > 0\), then the optimal contract may have a full insurance interval if no insurance interval exists, i.e., \(\exists \Omega_2 \Rightarrow \exists \Omega_1\).

Proof. See Appendix 6.3.

In other words, when the insurer’s background loss is affiliated with the insured loss, the form of optimal insurance contract differs from that in the case without dependent background risk. The optimal risk retention is nondecreasing, while the indemnity is not necessarily nondecreasing. In comparison to the case with independent background risk in Theorem 1, we can see that the affiliation is another influence on the degree of risk sharing, in addition to the nature of risk aversion and the nature of the administrative cost function. The possibilities are illustrated in Figures 1 and 2, which display the possible indemnity schedules when \(c'(\cdot) = 0\) and \(c'(\cdot) > 0\), respectively.
In contrast to the independent case, the optimal indemnity may have a full insurance interval for small losses since the insurer suffers more when the insured loss is large due to the effect of the affiliated background risk. In order to smooth its utility, the insurer wants to raise more premium by selling more insurance at the lower loss levels while avoiding high payouts at higher insured loss levels because of the elevated background losses. Indeed, upper limits may be found in the optimal contract since affiliation drives the insurer to restrict payouts for larger losses.

Put differently, if $Y$ and $X$ are positively dependent, then larger insured losses are connected to larger losses at the insurer; this raises the shadow cost of supplying insurance, so the insurer tends to cover less than in the case without the background risk. When $x$ increases, the shift in the conditional distribution of the insurer’s loss $Y$ makes the insurer more risk averse with respect to the indemnity amount, i.e., an additional payout is more painful for the insurer. This may cause the optimal contract to have a decreasing marginal indemnity, which in turn may generate an optimal upper limit—an outcome which is not possible in the absence of dependent insurer background risk.

Deductibles may also be influenced by insurer background risk. Although an administrative cost is still necessary for deductibles, the amount of the deductible will vary since it is jointly determined by the nature of dependence and the insurer’s risk aversion.

In regions of coinsurance, the optimal form of insurance contract will differ from the benchmark case of independent background risk due to the effect of $\Psi(x)$. Since the term $\Psi(x)$ depends on the joint distribution function of random variable $X$ and $Y$, the key determinant is the joint density $h(x,y)$. Affiliation evidently lowers the slope of the indemnity schedule in coinsurance regions relative to the case without affiliation, as can be seen in (14).
In general, various forms of indemnity schedule can be optimal. We cannot eliminate the possibility of a decreasing indemnity schedule without additional restrictions on the schedule. We now consider such a restriction by introducing a nondecreasing indemnity constraint.

3.3.2 Optimal Contracting with a Nondecreasing Indemnity Schedule

In the previous subsection, it was shown that the optimal insurance contract could feature a decreasing indemnity schedule. However, in the real world, insurance contracts typically specify indemnities that are increasing in the amount of insured loss. There are a number of potential explanations for this, one being asymmetric information: If it is costly for the insurer to verify losses, policyholders could underreport losses when confronted with a decreasing indemnity schedule. We are of course not the first to use such reasoning. For example, Huberman, Mayers, and Smith (1983) justify a nondecreasing indemnity constraint as necessary to prevent downward misrepresentation of the damage by the insured. We follow this logic by introducing a nondecreasing constraint for the optimal indemnity in this section:

\[ I(x) \text{ is nondecreasing with respect to } x \text{ on } [0, \bar{x}], \quad (15) \]

which is equivalent to

\[ I'(x) \geq 0 \quad \forall x \in \Omega_3 \]

if \( I(x) \) is continuous on \([0, \bar{x}]\) and is differentiable on \( \Omega_3 \). Therefore, similar to the condition in subsection 2.2.1, the \( \lambda \) satisfies \( J(x) \leq 0, \forall x \in \Omega_1 \), \( K(x) \geq 0, \forall x \in \Omega_2 \), and

\[ M(x) \leq 0, \forall x \in \Omega_3, \quad (16) \]

in which the inequality exists for those \( x \in \Omega_3 \) making new constraint (15) binding.
The following theorem characterizes optimal indemnity schedules in the presence of the nondecreasing indemnity constraint:

**Theorem 3.** Assume that $Y$ and $X$ are affiliated and that the optimal indemnity exists with the nondecreasing indemnity constraint 15. Then the optimal indemnity schedule can take the following forms (See Figure 3 and 4.):

(a) If $c'(\cdot) = 0$, then the optimal indemnity schedule can take the form of
   1. full insurance, coinsurance, and then an upper limit.
   2. full insurance, and then an upper limit.
   3. and (4) standard coinsurance.

(b) If $c'(\cdot) > 0$, then the optimal indemnity schedule can take the form of
   1. full insurance, coinsurance, and then an upper limit.
   2. coinsurance, and then an upper limit.
   3. deductible and coinsurance.
   4. deductible, coinsurance, and then an upper limit.

Proof. See Appendix 6.4. $\square$

[ Insert Figure 3 and 4 ]

As is shown in Figures 3 and 4, the optimal indemnity may take multiple possible schedules, including forms that differ from Raviv (1979) and Dana and Scarsini (2007). The optimal risk retention remains nondecreasing. However, a decreasing indemnity schedule no longer exists due to the additional constraint. As a result, an upper limit in the optimal indemnity may appear, with indemnity schedule holding the upper limit after the point at which it is reached. Full insurance at low losses is possible if the nondecreasing constraint is binding at high losses. The reasoning is that the
insurer gains by transferring wealth into high loss states, where affiliated background risk has made consumption especially dear, by selling additional coverage in low loss states. In the optimal contract, an increasing indemnity cost may not generate a deductible, but it is still the necessary condition for a deductible. However, the size of the deductible may vary according to the nature of the dependence between $Y$ and $X$. 
4 Applications - CARA Utilities and Specific Loss Distributions

We now specialize the problem to build intuition on how affiliation influences the properties of the upper limit and the deductible. Specifically, we assume both insured and insurer have constant absolute risk aversion (CARA) utility functions, with coefficients of absolute risk aversion denoted by $a$ and $b$, respectively. Under this assumption, if $(P^*, I^*)$ solves the problem (1), then for any $x \in \Omega_3$, where $0 < I^*(x) < x$, the optimal marginal coverage satisfies the following equation:

$$I^*(x) = \frac{a - \Psi(x)}{a + b(1 + c'') + c''/(1 + c''')}.$$

where $c'' = c'(I^*(x))$ and $c''' = c''(I^*(x))$ are evaluated at the optimal indemnity level $I^*$ and

$$\Psi(x) = \frac{\partial \ln \mathbb{E}[e^{bY}|x]}{\partial x},$$

which is equivalent to the covariance between $e^{bY}$ and $\frac{\partial \ln g(Y|x)}{\partial x}$. By definition of the upper limit, $x_u$ makes marginal coverage $I^*(x_u)$ to be zero, that is

$$\Psi(x_u) = \frac{\partial \ln \mathbb{E}[e^{bY}|x]}{\partial x} |_{x_u} = a.$$

**Theorem 4.** Suppose $Y$ is affiliated with $X$, $c(I) = mI$, $R_U(\cdot) = a$, and $R_V(\cdot) = b$, then the optimal contract $(P^*, I^*)$ has the following properties:

(1) The optimal indemnity has an increasing-convex deductible if

$$\frac{\partial^2 \ln \mathbb{E}[e^{bY}|x]}{\partial x^2} > 0,$$

and an upper limit appears if and only if inequality (18) holds and $\bar{x} > x_u$.

(2) Let $(\tilde{P}^*, \tilde{I}^*)$ be the optimal contract under independent background risk. Thus,
\[ x_d \geq \bar{x}_d, \]
\[ \text{if } \bar{\pi} < x_u. \]

Proof. See Appendix 6.5.

Thus, if \( \Psi(x) \) increases when \( x \) increases, the optimal marginal coverage will decrease with respect to \( x \), and the optimal contract will take the form of an increasing-convex deductible, and an indemnity ceiling will exist if and only if the maximum loss \( \bar{\pi} \) is large enough. If we assume that \( \Psi(x) \) decreases with insured loss \( x \), the optimal contract may take the form of an increasing-concave deductible contract, which will never have an upper limit.

Now consider the impact of affiliation on the deductible. With deductible insurance, the premium and indemnity schedule depends on the deductible amount and the optimal marginal coverage on \( \Omega_3 \), that is \( P^* = P(x_d) \) and \( I^*(x) = I(x_d, x) \). An increase in \( x_d \) will bring the benefit of a reduction in \( P \), which is offset by a drop in \( I(x) \). Increasing affiliation essentially makes insurance coverage more expensive on a per unit basis, which creates an incentive for the consumer to buy less coverage and bear more risk. One manifestation is a reduction in the slope of the indemnity schedule, and another manifestation is an increased deductible. After all, for the insured the least painful losses are the smallest ones, so it is easiest to sacrifice coverage in the realm of small losses when confronted with an effective increase in the cost of insurance.

To illustrate this theorem more clearly, we investigate some specific joint loss distributions: bivariate normal, binary insured loss, and censored joint log-normal.

**Bivariate Normal Distribution** Suppose the joint distribution \( H(X, Y) \) is bivariate normal \( N(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho) \),\(^{13}\) in which \( \mu_x \) and \( \mu_y \) are the means, \( \sigma_x \) and \( \sigma_y \) are the variances, and \( \rho \) is the correlation coefficient.

\(^{13}\)Particularly, a truncated bivariate normal distribution with a support of \([0, \bar{x}] \times [0, \bar{y}]\) should be considered. We use the bivariate normal with a support of \( \mathbb{R}^2 \) to make the proof easy, without altering the results.
are the standard deviations of $X$ and $Y$, and $\rho$ is the correlation. Then we have the following result:

**Corollary 2.** Suppose $X$ and $Y$ have a joint normal distribution, $c(I) = mI$, $R_U(\cdot) = a$, and $R_V(\cdot) = b$, then for any $x \in \Omega_3$, the optimal marginal coverage is

$$I^*(x) = \frac{a - b\rho \sigma_y / \sigma_x}{a + b(1 + m)}.$$

Moreover, suppose the optimal deductible is $x_d$, then

$$\frac{\partial x_d}{\partial \rho} > 0.$$

**Proof.** See Appendix 6.6. \qed

When the insurer’s loss and the insured’s loss are joint normal distributed, the optimal indemnity is linear for coinsurance intervals. The slope of the optimal indemnity is constant and depends on the risk aversion of both parties, the standard deviations and correlation of their losses, and the expense loading. An increase in the correlation will make indemnification more expensive for the insurer. As a result, a larger deductible and lower risk sharing, both in terms of absolute and marginal indemnification, appear as correlation increases. An upper limit will not appear in this circumstance.

**Binary Insured’s Loss** Suppose the insured’s loss is binary distributed as $X = \{0, L; 1 - \pi, \pi\}$ and the insurer’s background loss has a distribution based on the realized $X$. Due to the constraints on the indemnity, we know $I^*(0) = 0$, so the only choice variables in this problem are the premium $P$ and the coverage $I = I^*(L)$. Moreover, we let

$$\rho = \frac{\mathbb{E}[e^{bY}|X = L]}{\mathbb{E}[e^{bY}|X = 0]}.$$
From definition 5, we know that if $Y$ is stochastically increasing in $X$, then $\rho > 1$. The variable $\rho$ thus represents the intensity of the correlation in some sense. Consequently, we have the following corollary:

**Corollary 3.** Suppose the insured’s loss is binary distributed as $X = \{0, L; 1 - \pi, \pi\}$, the distribution of $Y$ depends on the value of $X$, $c(I) = mI$, $R_U(.) = a$, and $R_V(.) = b$, then

$$\frac{\partial I^*}{\partial \rho} < 0,$$

where $I^*$ is the optimal coverage when the insured has a positive loss.

*Proof.* See Appendix 6.7. \qed

In this setting, optimal coverage is decreasing in the intensity of the correlation. Intuitively, when the insurer is likely to suffer a bigger loss in the insured’s loss state, coverage is more costly to supply. Given the binary loss distribution, the decrease in indemnity can be interpreted either as a decrease in the upper limit or as an increase in the deductible.

**Truncated Joint Log-normal Distribution—A Numerical Example** Now we assume that $X$ and $Y$ follow a joint lognormal distribution with the following parameters: $\mu_x = 0.5, \mu_y = 1, \sigma_x = 0.4, \sigma_y = 0.2$, and the correlation of log($X$) and log($Y$) is $\rho$. In addition, the support of insured loss $X$ is $[0, 5]$, and the support of loss $Y$ is $[0, 10]$. We set the insured’s wealth at $W_1 = 3$, insurer’s wealth at $W_2 = 4$, and define utilities as $U(w) = 1 - \exp(-aw)$ and $V(w) = 1 - \exp(-bw)$, with the coefficient of absolute risk aversions $a = 2.5$ and $b = 1.5$. We consider cases with different correlations $\rho$ and and two different cost structures: no cost ($c(I) = 0$) and linear cost ($c(I) = .05 \times I$).

The following figures shows optimal indemnity schedules in four cases: 1) independent background risk and no administrative cost, 2) independent background...
risk but with linear administrative cost, 3) dependent background loss (with $\rho = 0.9$) without administrative cost, and 4) dependent background loss (with $\rho = 0.9$) with linear administrative cost. From the figures, we can see by comparing cases 1) and 2) that administrative cost leads to a positive deductible and a lower risk transfer, which is the same result as in the standard models; comparing cases 1) and 3) show that positive dependence in background risk leads to an overall lower degree of risk transfer; comparing cases 2) and 4) show that positive dependence in background risk leads to a larger deductible than in the case of independence.

[ Insert Figure 5, 6, 7, 8 ]

If we consider the effect of the correlation on the optimal indemnity structure, we find that, for the case without administrative cost, an increasing correlation will generate lower risk sharing (i.e., the slope of the indemnity schedule decreases). For the case with an administrative cost, increasing correlation generates a larger deductible and lower risk sharing simultaneously. These results are shown in Figures 9 and 10.

[ Insert Figure 9 and 10. ]

In summary, the examples verify that affiliation between insurer’s loss and insured loss is another important influence on the degree of risk sharing in the contract. Specifically, the examples illustrate how higher correlation leads to larger deductibles, lower upper limits, and a higher degree of coinsurance.
5 Conclusion

In this paper, we study optimal insurance contracts when the insurance buyer’s loss is correlated with the insurer’s background wealth, especially when the insurer’s background wealth is negatively correlated to a specific insurance buyer’s loss. This situation arises when risk exposures are correlated across policyholders—a situation that arises, to varying degrees, in many, if not all, insurance markets. We show that this correlation can produce optimal contract forms that are not possible in the standard model: Namely, upper limits on coverage can appear. We show further how the degree of correlation can reduce risk sharing and, in particular, increase the size of the optimal deductible—which is a finding that may explain contracting patterns in certain property catastrophe markets.

More generally, what we hope to accomplish in this paper is to bring attention to the importance of supply side risk considerations in understanding optimal contract design. Supply side risk considerations have received scant attention in comparison with asymmetric information considerations and demand side issues in the scholarly literature. Yet, the fundamental problems of insurance production are ones of diversification and risk management, and the task is typically performed by the supplier: Upon reflection, it is hard to believe that this risk management process would not have significant implications for the structure of contracts. Indeed, we have shown in the context of an extended blackboard model that risk considerations at the insurer can have a profound influence on the shadow costs of insurance coverage, and we hope that future scholarship will take this into account when analyzing contract forms.
6 Appendix

6.1 Proof of Theorem 1

Let \( \hat{V} : \mathbb{R} \mapsto \mathbb{R} \) be defined by \( \hat{V}(W) = \mathbb{E}[V(W - Y)] \) for all possible \( W \). Then \( \hat{V} \) inherits the properties of \( V \) and problem (1) turns out to be the same as the one studied by Raviv (1979). Since \( X \) and \( Y \) are independent, the optimization problem may be rewritten as a standard Pareto problem without background risk:

\[
EU(P, I) \equiv \max_{P \geq 0, 0 \leq I \leq Id} \mathbb{E}[U(W_1 - X - P + I(X))]
\]

subject to

\[
EV(P, I) \equiv \mathbb{E}[\hat{V}(W_2 + P - I(X) - c(I(X)))] \geq \hat{V}
\]

where \( \hat{V} = \hat{V}(W_2) = \mathbb{E}[V(W_2 - Y)] \). Following Arrow (1963), Raviv (1979), and Dana and Scarsini (2007), we know an optimal contract \((P^*, I^*)\) exists and is such that \( I^* \) and \( Id - I^* \) are nondecreasing.

(a) Consider a no insurance schedule \( I = 0 \), the corresponding premium has to be binding at zero, \( P = 0 \), since \( EU(P, I) \) is decreasing with respect to \( P \). Thus, the no insurance contract \((P = 0, I = 0)\) is an optimal solution of the above problem iff there exists \( \lambda \geq 0 \) such that

\[
-\mathbb{E}[U'(W_1 - X)] + \lambda \hat{V}'(W_2) \leq 0
\]

\[
U'(W_1 - x) - \lambda \hat{V}'(W_2) \leq 0 \quad \forall x \in [0, \varpi],
\]

or equivalently,

\[
-\mathbb{E}[U'(W_1 - X)] + \lambda \mathbb{E}[V'(W_2 - Y)] \leq 0
\]

\[
U'(W_1 - x) - \lambda \mathbb{E}[V'(W_2 - Y)] \leq 0 \quad \forall x \in [0, \varpi].
\]
Therefore, \( \forall x \in [0, \bar{x}] \), we have

\[
U'(W_1 - x) \leq \lambda \mathbb{E}[V'(W_2 - Y)](1 + c'(0)) \leq \mathbb{E}[U'(W_1 - X)](1 + c'(0)).
\]

Thus, no insurance contract is optimal iff

\[
1 + c'(0) \geq \frac{U'(W_1 - \bar{x})}{\mathbb{E}[U'(W_1 - X)]}.
\]

(b) Consider full insurance \( I(x) = x \), the corresponding premium has to be positive, \( P > 0 \), since \( EV(P, I) \) is increasing with respect to \( P \). Thus, the full insurance contract \( (P > 0, I = Id) \) is an optimal solution of the above problem if and only if there exists \( \lambda \geq 0 \) such that

\[
-U'(W_1 - P) + \lambda \mathbb{E}[\widehat{V}''(W_2 + P - X - c(X))]] = 0
\]

\[
U'(W_1 - P) - \lambda \widehat{V}''(W_2 + P - x - c(x))(1 + c'(x)) \geq 0 \quad \forall x \in [0, \bar{x}],
\]

or equivalently,

\[
-U'(W_1 - P) + \lambda \mathbb{E}[V''(W_2 - Y + P - X - c(X))]] = 0
\]

\[
U'(W_1 - P) - \lambda \mathbb{E}[V''(W_2 - Y + P - x - c(x))](1 + c'(x)) \geq 0 \quad \forall x \in [0, \bar{x}].
\]

Hence for all \( x \in [0, \bar{x}] \), we have

\[
1 + c'(x) \leq \frac{\mathbb{E}[V'(W_2 - Y + P - X - c(X))]}{\mathbb{E}[V'(W_2 - Y + P - x - c(x))]}.
\]

In the above inequality, since \( V''(\cdot) \leq 0 \) and \( c''(\cdot) \geq 0 \), left hand side is nondecreasing and right hand side is nonincreasing with respect to \( x \). Therefore, full insurance
contract is optimal iff
\[ 1 + c'(\bar{x}) \leq \frac{\mathbb{E}[V'(W_2 - Y + P - X - c(X))]}{\mathbb{E}[V'(W_2 - Y + P - \bar{x} - c(\bar{x}))]} . \]

Obviously, this condition is satisfied if and only if both sides of the equation are equal to 1, which in turn implies \( c'(\cdot) = 0 \) and \( V''(\cdot) = 0 \).

(c) Suppose \((P^*, I^*)\) is an optimal contract, similar to Raviv (1979), we have for all \( x \in \Omega_3 \), i.e., \( 0 < I(x) < x \),
\[ U'(A^*) - \lambda \mathbb{E}[V'(B^*)](1 + c'') = 0, \]
where \( c'' \) is evaluated at the indemnity level \( I^* \), and \( A^* = W_1 - x - P^* + I^*(x) \) and \( B^* = W_2 - Y + P^* - I^*(x) - c(I^*(x)) \) are insured and insurer’s final wealth with the insurance contract evaluated at insured loss \( x \in \Omega_3 \). Differentiating with respect to \( x \) and using the definition of \( A^* \) and \( B^* \), we obtain for all \( x \in \Omega_3 \),
\[ U''(A^*)(I''(x) - 1) + \lambda \mathbb{E}[V''(B^*)](1 + c'')^2 I''(x) - \lambda \mathbb{E}[V'(B^*)]c''I''(x) = 0. \]
Substituting \( \lambda \) from the first equation and solving for \( I''(x) \), we obtain
\[ I''(x) = \frac{R_U(A^*)}{R_U(A^*) + R_{\hat{v}}(\hat{B}^*)(1 + c'') + c''/(1 + c'')} \quad \text{for all } x \in \Omega_3. \]

Note that \( R_{\hat{v}}(\hat{B}^*) \) is the Arrow-Pratt coefficient of absolute risk aversion with the utility function \( \hat{V} \) and wealth \( \hat{B}^* \), which also implies the Kihlstrom et al. coefficient of absolute risk aversion with the utility function \( V \) and wealth \( B^* \),
\[ R_{\hat{v}}(\hat{B}^*) = -\frac{\hat{V}''(\hat{B}^*)}{\hat{V}'(\hat{B}^*)} = -\frac{\mathbb{E}[V''(\hat{B}^* - Y)]}{\mathbb{E}[V'(\hat{B}^* - Y)]} = -\frac{\mathbb{E}[V''(B^*)]}{\mathbb{E}[V'(B^*)]}. \]
(d) Since optimal contract \((P^*, I^*)\) is such that \(I^*\) and \(Id - I^*\) are nondecreasing, the feasible insurance indemnity include five cases: full insurance \((x_d = 0, x_f = \bar{x})\), full-insurance followed by coinsurance \((x_d = 0, x_f \in (0, \bar{x}))\), standard coinsurance contract \((x_d = x_f = 0)\), no insurance followed by coinsurance \((x_d \in (0, \bar{x}), x_f = 0)\), and no insurance \((x_d = \bar{x}, x_f = 0)\).

If \(c'(\cdot) = 0\), suppose \(x_d > 0\), the optimal contract can only be no insurance followed by coinsurance \((x_d \in (0, \bar{x}), x_f = 0)\) since no insurance condition (6) is not satisfied. Hence, there exists a \(\lambda \geq 0\) such that

\[
E[U'(W_1 - X - P^* + I^*(X))] = \lambda E[V'(W_2 - Y + P^* - I^*(X) - c(I^*(X)))]
\]

for all \(x \in [0, x_d]\), i.e., \(x \in \Omega_1\),

\[
U'(W_1 - x - P^*) - \lambda E[V'(W_2 - Y + P^*)] \leq 0,
\]

and for all \(x \in [x_d, \bar{x}]\), i.e., \(x \in \Omega_3\),

\[
U'(A^*) - \lambda E[V'(B^*)] = 0.
\]

Integrating the last two equations over \([0, \bar{x}]\), we find a contradiction with the first one due to the strictly concavity of \(U\). Hence, \(x_d\) has to be equal to zero.

On the other hand, if \(c'(\cdot) > 0\), suppose \(x_d = 0\), the optimal contract can be full-insurance followed by coinsurance \((x_d = 0, x_f \in (0, \bar{x}))\) or standard coinsurance contract \((x_d = x_f = 0)\) since full insurance condition (7) is not satisfied. For the case of full-insurance followed by coinsurance \((x_d = 0, x_f \in (0, \bar{x}))\), we have

\[
E[U'(W_1 - X - P^* + I^*(X))] = \lambda E[V'(W_2 - Y + P^* - I^*(X) - c(I^*(X)))]
\]

for all \(x \in [0, x_d]\), i.e., \(x \in \Omega_1\),

\[
U'(W_1 - x - P^*) - \lambda E[V'(W_2 - Y + P^*)] \leq 0,
\]

and for all \(x \in [x_d, \bar{x}]\), i.e., \(x \in \Omega_3\),

\[
U'(A^*) - \lambda E[V'(B^*)] = 0.
\]
for all $x \in [0, x_f]$, i.e., $x \in \Omega_2$,

$$U'(W_1 - P^*) \geq \lambda \mathbb{E}[V'(W_2 - Y + P^* - x - c(x))](1 + c'(x))$$

$$> \lambda \mathbb{E}[V'(W_2 - Y + P^* - x - c(x))] ,$$

and for all $x \in [x_f, x]$, i.e., $x \in \Omega_3$,

$$U'(A^*) = \lambda \mathbb{E}[V'(B^*)](1 + c^*) > \lambda \mathbb{E}[V'(B^*)].$$

Intergrading the last two equations over $[0, x]$, we find a contradiction with the first one. Similarly, for the case of standard coinsurance ($x_d = x_f = 0$), we have

$$\mathbb{E}[U'(W_1 - X - P^* + I^*(X))] = \lambda \mathbb{E}[V'(W_2 - Y + P^* - I^*(X) - c(I^*(X)))] ,$$

and for all $x \in [0, \bar{x}]$, i.e., $x \in \Omega_3$,

$$U'(A^*) = \lambda \mathbb{E}[V'(B^*)](1 + c^*) > \lambda \mathbb{E}[V'(B^*)].$$

Intergrading the second equation over $[0, \bar{x}]$, we find a contradiction with the first one. Hence, $x_d$ has to be positive.

Therefore, a necessary and sufficient condition for the optimal deductible to be zero, $x_d = 0$, is constant indemnity cost, $c'(\cdot) = 0$.

(e) and (f) Since condition (6) is not satisfied, we rule out the case of no insurance from the possible optimal contracts.

(1) If $V''(\cdot) = 0$ and $c'(\cdot) = 0$, then obviously full insurance condition (7) is satisfied, so full insurance contract is optimal.

(2) If $V''(\cdot) = 0$ and $c'(\cdot) = m > 0$, then we know $R_{\varphi}(\hat{B}^*) = 0$. Thus for all $x \in \Omega_3$, optimal marginal coverage equals to 1, $I^*(x) = 1$, which implies any standard
coinsurance or full-insurance followed by coinsurance cannot be optimal. Hence, the possible optimal contract can be full insurance or standard deductible contract. Obviously, the full insurance condition (7) is not satisfied since \( c'(\cdot) = m > 0 \), thus, the optimal contract is a standard deductible contract.

(3) If \( V''(\cdot) = 0 \), \( c'(\cdot) > 0 \), and \( c''(\cdot) > 0 \), then we can obtain the optimal marginal coverage is between 0 and 1, that is \( 0 < I^*(x) < 1 \), for any \( x \in \Omega_3 \). Obviously, from (d), the feasible contract can only be generalized deductible contract.

(4) If \( V''(\cdot) > 0 \) and \( c'(\cdot) = 0 \), then we can obtain the optimal marginal coverage is between 0 and 1, that is \( 0 < I^*(x) < 1 \), for any \( x \in \Omega_3 \). Obviously, from (d), the feasible contracts include three possible cases: full insurance, full-insurance followed by coinsurance, and standard coinsurance. Consider the full insurance contract, since \( V''(\cdot) > 0 \) and \( Y \) is independent of \( X \), we know \( \mathbb{E}[V'(W_2 - Y + P - x - c(x))] \) is increasing with respect to \( x \). Hence,

\[
\mathbb{E}[V'(W_2 - Y + P - X - c(X))] < \mathbb{E}[V'(W_2 - Y + P - \bar{x} - c(\bar{x}))],
\]

which in turn violate the full insurance condition (7). Consider the full insurance followed by coinsurance, then there exists a \( \lambda \geq 0 \) such that

\[
\mathbb{E}[U'(W_1 - X - P^* + I^*(X))] = \lambda \mathbb{E}[V'(W_2 - Y + P^* - I^*(X) - c(I^*(X)))] ,
\]

for all \( x \in [0, x_f] \), i.e., \( x \in \Omega_2 \),

\[
U'(W_1 - P^*) - \lambda \mathbb{E}[V'(W_2 - Y + P^* - x - c(x))] \geq 0,
\]

and for all \( x \in [x_d, \bar{x}] \), i.e., \( x \in \Omega_3 \),

\[
U'(A^*) - \lambda \mathbb{E}[V'(B^*)] = 0.
\]
Integrating the last two equations over $[0, \bar{x}]$, we find a contradiction with the first one due to the strictly concavity of $U$. Therefore, the optimal contract in this case can only be standard coinsurance.

(5) If $V''(\cdot) > 0$ and $c'(\cdot) > 0$, then we can obtain the optimal marginal coverage is between 0 and 1, that is $0 < I^*(x) < 1$, for any $x \in \Omega_3$. Obviously, the feasible insurance contracts include four possible cases: full insurance, full-insurance followed by coinsurance, standard coinsurance, and generalized deductible contracts. Similar to the proof in (e.3), due to the increasing cost $c'(\cdot) > 0$, full insurance, full-insurance followed by coinsurance, and standard coinsurance contracts are not optimal. The only possible optimal contract is generalized deductible.

6.2 Proof of Corollary 1

(a) If $V''(\cdot) = 0$, then the optimization problem is equivalent to

$$\max_{P \geq 0, 0 \leq I \leq I_d} \mathbb{E}[U(W_1 - X - P + I(X))]$$

subject to

$$P \geq I(X) + c(I(X)).$$

The probability distribution of background loss $Y$ cannot affect the optimal solution, thus $(\tilde{P}^*, \tilde{I}^*) = (P^*, I^*)$.

(b) We can assume $V(w) = 1 - e^{-bw}$ without loss of generality since $V(\cdot)$ is constant absolute risk aversion (CARA). Thus, the constraint (3) is

$$\mathbb{E}[1 - e^{W_2-Y+P-I(X)-c(I(X))}] \geq \mathbb{E}[1 - e^{W_2-Y}],$$

which is equal to

$$P \geq I(X) + c(I(X)).$$
The probability distribution of background loss $Y$ cannot affect the optimal solution, thus $(\tilde{P}^*, \tilde{I}^*) = (P^*, I^*)$.

(c) From Theorem (1), we know that

$$\tilde{I}^*(x) = \frac{R_U(\tilde{A}^*)}{R_U(A^*) + R_\tilde{V}(\tilde{B}^*)(1 + \tilde{c}^*) + \tilde{c}^*/(1 + \tilde{c}^*)}$$

for all $x \in \Omega$, where

$$\tilde{A}^* = W_1 - x - P^* + \tilde{I}^*(x), \quad \tilde{B}^* = W_2 - Y + P^* - \tilde{I}^*(x) - c(\tilde{I}^*(x)).$$

and

$$R_\tilde{V}(\tilde{B}^*) = -\mathbb{E}[V''(\tilde{B}^*)]/\mathbb{E}[V'(\tilde{B}^*)],$$

Assume there exist some particular $x \in [0, \bar{x}]$ such that $\tilde{I}^*(x) = I^*(x)$ in addition to the given $\tilde{P}^* = P^*$, therefore $R_U(\tilde{A}^*) = R_U(A^*)$, $\tilde{c}^* = c^*$, and $\tilde{c}^{**} = c^{**}$. Moreover, since $\tilde{Y}$ is riskier than $Y$ and insurer is risk vulnerable, we have $R_\tilde{V}(\tilde{B}^*) > R_\tilde{V}(\tilde{B}^*)$ based on the result in Gollier and Pratt (1996). Consequently, $\tilde{I}^*(x) < I^*(x)$ at those particular $x$.

If $c'(\cdot) = 0$, then we know $\tilde{x}_d = x_d = 0$, $\tilde{I}^*(0) = I^*(0) = 0$, $\tilde{I}(x) = \int_0^x \tilde{I}(t)dt$, and $I(x) = \int_0^x I(t)dt$. Obviously, if the insurer is risk vulnerable, $\tilde{I}^*(x) \leq I^*(x)$ for all $x \in [0, \bar{x}]$ and vice versa.

If $c'(\cdot) = m > 0$, the optimal marginal indemnity depends on both risk attitudes, the optimal premium and indemnity, which is ambiguous even for risk vulnerable insurer. Since the optimal indemnity has to be general deductible, then given a deductible level $x_d$, we obtain an indemnity schedule from the marginal indemnity function $8$. In addition, from the first order condition, we obtain the implicit function...
for optimal deductible $x_d$,

$$(1+m) \int_0^{x_d} \frac{U'[W_1 - x - P^*(x_d)]}{U'[W_1 - x_d - P^*(x_d)]} dF(x) + m \int_{x_d}^\infty \frac{U'[W_1 - x - P^*(x_d) + I^*(x_d, x)]}{U'[W_1 - x_d - P^*(x_d)]} dF(x) = F(x_d),$$

which is ambiguous since the premium can vary.

### 6.3 Proof of Theorem 2

(a) First, we define a function $\xi : \mathbb{R}^2 \mapsto \mathbb{R}$

$$\xi(m, n) = \mathbb{E}[V(W_2 + P + m - Y - n - c(n - m))|X = n].$$

Find the joint derivative of the new function

$$\frac{\partial^2 \xi(m, n)}{\partial n \partial m} = (1 + c')\mathbb{E}[V'(B) \frac{\partial \ln g(Y|n)}{\partial n}|n] + (1 + c')^2 \mathbb{E}[-V''(B)|n] + c''\mathbb{E}[V'(B)|n].$$

From the previous proof, we know that $\mathbb{E}[V'(B)\partial \ln g(Y|n)/\partial n|n] \geq 0$, hence,

$$\frac{\partial^2 \xi(m, n)}{\partial n \partial m} \geq 0.$$

Therefore, the function $\xi(m, n)$ is proved to be supermodular. If $\tilde{m}$ is a nondecreasing rearrangement (with the same distribution) of $m$ with respect to $n$, then

$$\mathbb{E}[\xi(\tilde{m}(n), n)] \geq \mathbb{E}[\xi(m(n), n)].$$

Let $X - \bar{I}(X)$ be a nondecreasing rearrangement of $X - I(X)$. Since the distribution of $X - \bar{I}(X)$ and $X - I(X)$ is unchanged,

$$\mathbb{E}[U(W_1 - X - P + \bar{I}(X))] = \mathbb{E}[U(W_1 - X - P + I(X))].$$
Consider $m = X - I(X)$, $n = X$, we obtain

$$
\mathbb{E}[V(W_2 - y(X) + P - \bar{I}(X) - c(\bar{I}(X))]] \geq \mathbb{E}[V(W_2 - y(X) + P - I(X) - c(I(X)))]
$$

Therefore, any contract $I(X)$ is dominated by $\bar{I}(X)$, which makes $X - \bar{I}(X)$ nondecreasing. In other words, the optimal contract is such that $Id - I^*$ is nondecreasing.

(b) Consider a no insurance schedule $I = 0$, the corresponding premium has to be binding or zero, $P = 0$, since $EU(P, I)$ is decreasing with respect to $P$. Thus no insurance contract $(P = 0, I = 0)$ is optimal if and only if there exists $\lambda \leq 0$ such that

$$
-U'(W_1 - X) + \lambda \mathbb{E}[V'(W_2 - Y)] \leq 0
$$

$$
U'(W_1 - x) - \lambda \mathbb{E}[V'(W_2 - Y)|X = x](1 + c'(0)) \leq 0 \quad \forall x \in [0, \bar{x}].
$$

Hence for all $x \in [0, \bar{x}]$, we have

$$
1 + c'(0) \geq \frac{U'(W_1 - x) \mathbb{E}[V'(W_2 - Y)]}{\mathbb{E}[V'(W_2 - Y)|X = x] \mathbb{E}[U'(W_1 - X)]}.
$$

Both the nominator and the denominator are increasing with respect to $x$. Thus, no insurance contract is optimal iff

$$
1 + c'(0) \geq \sup_{x \in [0, \bar{x}]} \left\{ \mathbb{E}[V'(W_2 - Y)] \frac{U'(W_1 - x)}{\mathbb{E}[U'(W_1 - X)] \mathbb{E}[V'(W_2 - Y)|X = x]} \right\}.
$$

(c) Consider a full insurance $I = Id$, the corresponding premium has to be positive, $P > 0$, since $EV(P, I)$ is increasing with respect to $P$. Thus, the full insurance contract $(P > 0, I = Id)$ is an optimal solution of the above problem if and only if
there exists $\lambda \geq 0$ such that

\[-U'(W_1 - P) + \lambda \mathbb{E}[V'(W_2 - Y + P - X - c(X))] = 0\]

\[U'(W_1 - P) - \lambda \mathbb{E}[V'(W_2 - Y + P - x - c(x))|X = x](1 + c'(x)) \geq 0 \quad \forall x \in [0, \bar{x}].\]

Hence we have

\[1 + c'(x) \leq \frac{\mathbb{E}[V'(W_2 - Y + P - x - c(x))]|X = x}{\mathbb{E}[V'(W_2 - Y + P - x - c(x))]|X = x}] \quad \text{for all } x \in [0, \bar{x}].\]

Since $X$ and $Y$ are affiliated and $V''(\cdot) \leq 0$, we know $\mathbb{E}[V'(W_2 - Y + P - x - c(x))|X = x]$ is nondecreasing with respect to $x$. Moreover, $1 + c'(x)$ is increasing with $x$ due to $c''(\cdot) \geq 0$. Therefore, full insurance contract is optimal iff

\[1 + c'(\bar{x}) \leq \frac{\mathbb{E}[V'(W_2 - Y + P - x - c(x))]|X = x]}{\mathbb{E}[V'(W_2 - Y + P - \bar{x} - c(\bar{x}))]|X = \bar{x}]}.\]

(d) Suppose $(P^*, I^*)$ is an optimal contract, then similar to part (c) in Theorem 1, we have the marginal benefit of shifting the indemnity from the optimal level when the optimal indemnity is coinsurance, i.e., for all $x \in \Omega_3$,

\[U'(A^*) - \lambda \mathbb{E}[V'(B^*)|X = x](1 + c^{\prime\prime}) = 0,\]

where $A^* = W_1 - P^* - x + I^*(x)$, $B^* = W_2 - Y + P^* - I^*(x) - c(I^*(x))$ are the ex-post wealth of insured and insurer, and $c^{\prime\prime}$, $c^{\prime\prime\prime}$ are evaluated at $I^*(x)$. Differentiating it with respect to $x$, we obtain

\[U''(A^*)(I^{\prime\prime} - 1) + \lambda(1 + c^{\prime\prime})\mathbb{E}[V''(B^*)(1 + c^{\prime\prime})I^{\prime\prime} - V'(B^*)\frac{\partial \ln g(Y|X)}{\partial x}|X = x] \]

\[-\lambda \mathbb{E}[V'(B^*)|X = x]c''I^{\prime\prime} = 0]
By substituting $\lambda$ we can solve for $I^*(x)$ as in the follows

$$I^*(x) = \frac{R_U(A^*) - \Psi(x)}{R_U(A^*) + R_{\hat{V}}(B^*)(1 + c^*) + c^*/(1 + c^*)}$$

for all $x \in \Omega_3$,

where

$$\Psi(x) = \frac{\mathbb{E}[V'(B^*)(\partial \ln g(Y|x)/\partial x)|X = x]}{\mathbb{E}[V'(B^*)|X = x]}$$

From the affiliation of $X$ and $Y$, we know that $\partial \ln g(y|x)/\partial x$ is nondecreasing with respect to $y$

$$\frac{\partial^2 \ln g(y|x)}{\partial x \partial y} \geq 0.$$

Moreover, we know that

$$\mathbb{E} \left[ \frac{\partial \ln g(Y|x)}{\partial x} \bigg| X = x \right] = \int_0^y \frac{\partial g(y|x)}{\partial x} dx = 0,$$

and $V'(B^*)$ is always positive and nondecreasing with respect to $y$. Therefore, we obtain that

$$\Psi(x) = \frac{\mathbb{E}[V'(B^*)(\partial \ln g(Y|x)/\partial x)|X = x]}{\mathbb{E}[V'(B^*)|X = x]} \geq 0$$

(e) Since the optimal contract is nontrivial, we rule out the case of no insurance from the possible optimal contracts.

(1) If $V''(\cdot) = 0$ and $c'(\cdot) = 0$, then obviously full insurance condition is satisfied, so full insurance contract is optimal.

(2) If $V''(\cdot) = 0$ and $c'(\cdot) = m > 0$, then we know $R_{\hat{V}}(\hat{B}^*) = 0$. Moreover, $V'(B^*)$ is constant, thus $\Phi(x) = 0$. Hence, for all $x \in \Omega_3$, optimal marginal coverage equals to 1, $I^*(x) = 1$, which implies the possible optimal contract can be full insurance or standard deductible contract. Obviously, the full insurance condition is not satisfied since $c'(\cdot) = m > 0$, thus, the optimal contract is a standard deductible contract.
(3) If \( V''(\cdot) = 0, c'(\cdot) > 0, \) and \( c''(\cdot) > 0, \) then we can obtain the optimal marginal coverage is between 0 and 1, that is \( 0 < I''(x) < 1, \) for any \( x \in \Omega_3. \) Obviously, from (c), the feasible contract can only be generalized deductible contract.

(f) If \( V''(\cdot) > 0, \) consider an optimal contract with full insurance intervals, i.e., \( \exists \Omega_2. \) Suppose no insurance interval does not exist, then any \( x \in [0, \bar{x}] \) is in \( \Omega_2 \) or \( \Omega_3. \) Then, we can integrate inequalities (12) and (13) over \([0, \bar{x}],\)

\[
\mathbb{E}[U'(W_1 - X - P^* + I^*(X))] > \lambda \mathbb{E}[V'(W_2 - Y + P^* - I^*(X)) - c(I^*(X))](1 + c''),
\]

which obviously contradict with equation (10), even if \( c'(\cdot) = 0. \) Therefore, the optimal contract may have a full insurance interval only if no insurance interval exists.

### 6.4 Proof of Theorem 3

From Theorem 2 and the nondecreasing constraint (15), we know that the optimal marginal indemnity in \( \Omega_3 \) is between 0 and 1. Thus, the full insurance and no insurance intervals can only appear at the lowest losses. After integrating different combinations of inequalities (11), (12), and (16), and comparing with equation (10), we have the following results:

(a) If \( c'(\cdot) = 0, \) the optimal contract may have a full insurance interval if some flat indemnity intervals exist, such as (1), (3), and (4), for those larger losses not shown in the figures, both flat and increasing indemnities may appear; or the optimal contract is just a standard coinsurance as the result in literature.

(b) If \( c'(\cdot) > 0, \) either no insurance interval or binding of the nondecreasing indemnity has to occur to make those first order conditions consistent. Therefore, Deductibles can exist with or without any upper limit, such as (3) and (4); upper limits can exist with or without any full insurance interval, such as (1) and (2).
6.5 Proof of Theorem 4

Since $c'(\cdot) = m > 0$, we know the optimal contract will have an upper limit except for the case (3), which consists of a deductible and coinsurance.

(1) If condition 18 is satisfied, then $I''(x)$ is decreasing with respect to $x$, which in turn gives a concave indemnity and convex deductible. Moreover, if $\bar{x} > x_u$, then there is some $x$ such that $I'(x) < 0$, which violate the nondecreasing constraint. In addition, the amount of $x_u$ does not depend on the choice of the indemnity or the premium due to the CARA utilities. Obviously, we then have the upper limit above $x_u$. On the other hand, if $\bar{x} \leq x_u$, the $I'(x)$ is always nonnegative, thus the nondecreasing indemnity constraint is not binding. Therefore, no upper limit can be found.

(2) Consider the generalized deductible contracts (without upper limits) of the independent and dependent cases under the CARA and linear cost assumptions. For the dependent case, the generalized deductible contracts is shown in Figure 4 by (c) or (d) when $\bar{x} \leq x_u$. Thus the optimal indemnity is $I^*(x) = 0$ for all $x \in [0, x_d]$, and for $x \in [x_d, \bar{x}]$,

$$I''(x) = \int_{x_d}^{x} I''(t) dt = \frac{ax}{a+bm} - \frac{\ln E[e^{bY}|x]}{a+bm} - \frac{ax_d}{a+bm} + \frac{\ln E[e^{bY}|x_d]}{a+bm}.$$  

Moreover, we have the first order conditions (10), (11), and (13), and at the loss $x_d$,

$$U'(W_1 - x_d + P^*) = \lambda E[V'(W_2 - Y + P^*)|x_d](1 + m).$$

Therefore, we have the following implicit function of optimal deductible $x_d$

$$(1+m) \int_0^{x_d} e^{a(x-x_d)} dF(x) + m \int_{x_d}^{\bar{x}} e^{\frac{ah(1+m)}{a + b(1+m)}(x-x_d)} \left( \frac{E[e^{by}|x]}{E[e^{by}|x_d]} \right)^\frac{a}{a+b(1+m)} \frac{a}{a+b(1+m)} dF(x) = \int_0^{x_d} \frac{E[e^{by}|x]}{E[e^{by}|x_d]} dF(x).$$

44
Similarly, for the independent case, we have the implicit function for $x_d$

$$(1 + m) \int_0^{x_d} e^{a(x-x_d)} dF(x) + m \int_{x_d}^{\bar{x}} e^{\frac{ab(1+m)}{a+m}(x-x_d)} dF(x) = F(x_d)$$

Denote $L_i(x_d)$ and $R_i(x_d)$, where $i = 1, 2$ for the LHS and RHS of each equation. Let’s consider the LHS and RHS of both equations as functions of $x_d$ and analyze the properties of the functions. Obviously, $L_1(0) > L_2(0) > 0$, $L_1(\bar{x}) = L_2(\bar{x}) > 0$, $R_1(0) = R_2(0) = 0$, and $R_2(\bar{x}) = R_1(\bar{x}) > L(\bar{x}) > 0$. Since the optimal deductible amount is decided when the LHS and RHS equal, comparing the two equations we can see that the optimal deductible in the first equation if larger than that in the second equation.

6.6 Proof of Corollary 2

1) $X$ and $Y$ are affiliated. From the definition of bivariate normal, we have the conditional distribution

$$g(y|x) = \frac{1}{\sigma_y \sqrt{2\pi(1-\rho^2)}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \left( \frac{y - \mu_y}{\sigma_y} \right)^2 + \rho^2 \left( \frac{x - \mu_x}{\sigma_x} \right)^2 - 2\rho \frac{(x - \mu_x)(y - \mu_y)}{\sigma_x \sigma_y} \right] \right\},$$

and

$$\frac{\partial g(y|x)}{\partial x} = \frac{\rho}{\sigma_x (1-\rho^2)} \left[ \frac{y - \mu_y}{\sigma_y} - \rho \frac{x - \mu_x}{\sigma_x} \right] g(y|x).$$

Also, from the joint density function of bivariate normal, we obtain

$$\frac{\partial^2 \ln h(x,y)}{\partial x \partial y} = \frac{\rho}{\sigma_x \sigma_y (1-\rho^2)} \geq 0.$$ 

Therefore, the bivariate normal variables $X$ and $Y$ are affiliated.
2) From Theorem 2, we obtain that for any \( x \in \Omega_3 \), \( I'' = \frac{a - \Psi(x)}{a + b(1 + m)} \), where

\[
\Psi(x) = \frac{\partial \ln \mathbb{E}[e^{bY}|x]}{\partial x} = \frac{\mathbb{E}[e^{bY} \partial \ln g(Y|x)/\partial x|x]}{\mathbb{E}[e^{bY}|x]},
\]

in which the denominator is

\[
\mathbb{E}[e^{bY}|x] = \int_{-\infty}^{+\infty} e^{by} g(y|x) dy = e^n,
\]

where

\[
n = \frac{b^2(1-\rho^2)^2\sigma_y^2 + 2b(1-\rho^2)\mu_y + 2b^2(1-\rho^2)^2\sigma_x (x-\mu_x)}{2(1-\rho^2)}
\]

depends on the given parameters of the joint distribution and the amount of loss \( x \); while the numerator is

\[
\mathbb{E}[e^{bY} \partial \ln g(Y|x)/\partial x|x] = \int_{-\infty}^{+\infty} e^{by} \rho \frac{y-\mu_y}{\sigma_y} \left( \frac{y-\mu_y}{\sigma_y} - \rho \frac{x-\mu_x}{\sigma_x} \right) g(y|x) dy
\]

\[
= b\rho \frac{\sigma_y e^n}{\sigma_x} + \frac{\rho e^n}{\sqrt{2\pi(1+\rho^2)} \sigma_x (1-\rho^2)} \int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2(1-\rho^2)}} dt
\]

\[
= b\rho \frac{\sigma_y e^n}{\sigma_x},
\]

where \( t = \frac{y}{\sigma_y} - m \) and \( m = b(1-\rho^2)\sigma_y + \frac{\mu_y}{\sigma_y} + \rho \frac{\mu_y}{\sigma_x} \). Therefore, we have the optimal marginal indemnity

\[
I''(x) = \frac{a - b\rho \sigma_y / \sigma_x}{a + b(1 + m)}.
\]

3) We can write the optimal indemnity as

\[
I^* = \begin{cases} 
0 & \forall x \in (-\infty, x_d] \\
\frac{a - b\rho \sigma_y / \sigma_x}{a + b(1 + m)} (x - x_d) & \forall x \in (x_d, +\infty)
\end{cases}
\]

Hence, the original optimization problem has only one choice variable \( x_d \) now. Applying the first order conditions similar to (10) (11), and (13), we obtain

\[
(1 + m) \int_{-\infty}^{x_d} e^{a(x-x_d)} dF(x) + m \int_{x_d}^{+\infty} e^{a(1-\delta)(x-x_d)} dF(x) = \int_{-\infty}^{x_d} \frac{\mathbb{E}[e^{bY}|x]}{\mathbb{E}[e^{bY}|x_d]} dF(x),
\]
where $\delta = \frac{a - b \rho \sigma_y / \sigma_x}{a + b(1 + m)}$ is the slope of the indemnity function. Due to proof of part 2), the right hand side of the equation is equal to

$$\int_{-\infty}^{x_d} e^{b \rho \sigma_y / \sigma_x (x - x_d)} dF(x),$$

Now, we do derivative to both sides of the equation with respect to $\rho$ and get

$$\frac{\partial x_d}{\partial \rho} = \frac{abm \sigma_y}{(a + b(1 + m)) \sigma_x} \int_{x_d}^{\infty} e^{(1 - \delta)(x - x_d)} dF(x) + \frac{b \sigma_y}{\sigma_x} \int_{-\infty}^{x_d} (x_d - x) e^{b \rho \sigma_y / \sigma_x (x - x_d)} dF(x)$$

Since the assumptions make sure that $I'' = \delta \geq 0$, which means $a \geq b \rho \sigma_y / \sigma_x$, and thus

$$a(1 + m) e^a - b \rho \sigma_y / \sigma_x e^{b \rho \sigma_y / \sigma_x} \geq 0.$$

Other terms in the equation of $\frac{\partial x_d}{\partial \rho}$ is obviously positive. Therefore, we have a positive $\frac{\partial x_d}{\partial \rho}$ and the optimal deductible increases with correlation $\rho$.

### 6.7 Proof of Corollary 3

The optimization problem is written as

$$\max_{P, I} \pi U(W_1 - P - L + I) + (1 - \pi)U(W_1 - P)$$

subject to

$$\pi \mathbb{E}[V(W_2 + P - Y - (1 + m)I)|X = L] + (1 - \pi)\mathbb{E}[V(W_2 + P - Y)|X = 0] \geq V,$$
\( P \geq 0 \), and \( 0 \leq I \leq L \), where \( \bar{V} \) denotes the insurer's reservation utility level. By solving this problem, we have

\[
m\pi + (1 + m)(1 - \pi)e^{-a(L - I^*)} = (1 - \pi)\frac{1}{\rho}e^{-b(1+m)I^*},
\]

where \( \rho = \frac{\mathbb{E}[e^{bY|L}]}{\mathbb{E}[e^{bY|0}]} \). Thus, we have

\[
\frac{\partial I^*}{\partial \rho} = -\frac{(1 - \pi)e^{-b(1+m)I^*}}{(1 + m)(1 - \pi)\rho[a\rho e^{-a(L - I^*)} + be^{-b(1+m)I^*}]} < 0.
\]
References


Figure 1: Optimal Indemnity Schedules when $c'(\cdot) = 0$

(a) Standard Coinsurance

(b) Full Insurance - Coinsurance - No Insurance

(c) Full Insurance - Coinsurance - No Insurance

(d) Coinsurance - No Insurance

(e) Deductible - Coinsurance

(f) Deductible - Coinsurance - No Insurance

Figure 2: Optimal Indemnity Schedules when $c'(\cdot) > 0$
Figure 3: Nondecreasing Optimal Indemnity Schedules when $c'(\cdot) = 0$

(a) Full Insurance - Coinsurance - Upper Limit

(b) Standard Coinsurance

(c) Full Insurance - Upper Limit

(d) Full Insurance - Coinsurance

(e) Full Insurance - Coinsurance - Upper Limit

(f) Coinsurance - Upper Limit

(g) Deductible - Coinsurance

(h) Deductible - Coinsurance - Upper Limit

Figure 4: Nondecreasing Optimal Indemnity Schedules when $c'(\cdot) > 0$
Figure 5: Optimal Indemnity Structure with Independent Background Risk and No Indemnity Cost

Figure 6: Optimal Indemnity Structure with Independent Background Risk and Increasing Indemnity Cost
Figure 7: Optimal Indemnity Structure with Dependent Background Risk and No Indemnity Cost

Figure 8: Optimal Indemnity Structure with Dependent Background Risk and Increasing Indemnity Cost
Figure 9: Optimal Indemnity Structures with Different Dependence Structures with Cost \((c = 0)\);
Black: \(\rho = 0\); Blue: \(\rho = 0.3\); Green: \(\rho = 0.6\); Pink: \(\rho = 0.9\).

Figure 10: Optimal Indemnity Structures with Different Dependence Structures with Cost \((c(I) = 0.05 \times I)\);
Black: \(\rho = 0\); Blue: \(\rho = 0.3\); Green: \(\rho = 0.6\); Pink: \(\rho = 0.9\).
Chapter 2: Dynamic Capital Allocation with Irreversible Investments

1 Introduction

Practitioners have long wrestled with the problem of how to allocate a financial institution’s capital to the various risks within its portfolio for purposes of pricing and performance measurement. Over the past 15 years, a number of papers have explored the problem of allocating capital based on marginal risk contributions in different settings via different conceptual approaches (see Bauer and Zanjani (2013) for a review). However, these techniques are largely restricted to static portfolios, where the exposures are either set in advance or chosen simultaneously in an optimization problem.\footnote{While some papers do study “dynamic” approaches to allocation (Tsanakas (2004); Laevan and Goovaerts (2004)), the “dynamics” are introduced by evolving information about the components of a previously chosen portfolio rather than changes to the composition of the portfolio itself.}

Unfortunately, the real world application of these concepts is more complicated. Specifically, allocation for pricing purposes is typically done in an environment where the portfolio is not fixed in advance and is instead in the process of being constructed. Opportunities for underwriting or investment do not all arrive at the same time. If they did, or if positions in existing opportunities could be costlessly modified as new opportunities arrived, existing allocation techniques could be applied without modification. More typically, however, opportunities arrive sequentially in time, and, moreover, 1) decisions must be made about those opportunities when they arrive and 2) those decisions are typically irreversible, or reversible only with significant cost. This characterization applies to insurance companies, private equity firms, banks, and other institutions dealing in illiquid investments. How should risk be priced in such an environment? And can risk pricing be connected to capital
allocation?

This paper tackles these questions by studying the usual portfolio optimization problem—where expected returns are optimized subject to a risk measure constraint—with three complications motivated by the discussion above. First, investment opportunities arrive sequentially. Second, the risk and return properties of future opportunities are uncertain. Third, investment decisions are irreversible.

We find that marginal cost pricing can indeed be connected to capital allocation in this setup, although the basis for allocation is different from that found in static problems. In static problems, the marginal value of capital is known with certainty, and capital is allocated to each risk based on how it consumes that capital at the margin. In this dynamic problem, the ultimate marginal value of a dollar of capital will not be known until all opportunities have been presented and decisions made; similarly, it is also not known what demands a given investment will make on the firm’s capital until the portfolio is set. What is known—assuming that the uncertainty in the risk and return structure of future opportunities is understood—is the future distribution of possibilities. To elaborate, we know the possible paths ahead and their probabilities, so we know what the value of marginal unit of capital will be in various future states of the world, and we also know in each of those states how that marginal unit of capital is consumed by the risks in the portfolio.

Thus, the investment decision for an opportunity presented today is made on the basis of an expected future marginal cost of risk associated with that opportunity. Expected future marginal costs can alternatively be expressed in terms an allocation of the firm’s capital to current investments and future opportunities. The basis of the allocation for today’s opportunity is a probability-weighted average of the product of the marginal value of capital in future states of the world and the amount of capital consumed by today’s opportunity in those future states of the world.

The logic and intuition of the model is related to the literature on real options (e.g., Dixit and Pindyck (1994)). The fundamental problem in this literature entails the impact of uncertainty on investments that are either partially or completely irreversible. The classic intuition is that simple internal rate of return calculations, using a firm’s cost of capital
as a hurdle rate, lead to overly aggressive investment because they fail to account for the option value of waiting for more information on the productivity of the investment. As a result, hurdle rates need to be adjusted upward, and greater uncertainty about productivity can be shown to lead to even higher hurdle rates (e.g., McDonald and Siegel (1986)). To the extent that the amount of investment is discretionary, current investment is delayed or reduced as there is value created by waiting.

The rest of this paper is organized as follows. We start in Section 2 by sketching basic theory, first in a two period model to build intuition, and next in a multiperiod generalization. We then use simple numerical examples in Section 3 to show how future uncertainties can affect optimal decisionmaking today.

2 Dynamic Capital Allocation

2.1 Two opportunities

The financial institution receives two opportunities in sequence. Upon receipt of the first opportunity, it must decide on the quantity \( q_0 \) to invest in the opportunity. The second opportunity is known to be coming, but there is uncertainty about its risk and properties. This uncertainty is denoted by the random variable \( \alpha \) which takes values in the set \( \{ \alpha_1, \cdots, \alpha_m \} \) with probabilities \( \{ p(\alpha_1), \cdots, p(\alpha_m) \} \). The uncertainty is resolved upon receipt of the second opportunity, at which point the institution decides on the quantity \( q_1 \) to invest in the second opportunity.

Let the return on the first opportunity be a random variable \( r_0 \) and the return on the second having a distribution that depends on the realization of \( \alpha \), as in \( r_1(\alpha_i) \). Thus, different realizations of \( \alpha \) will affect the prospective distribution of outcomes for the second opportunity. Let the firm’s capital be \( K \) and the frictional cost of capital be \( \tau \). If we consider the firm’s problem at decision time on the first opportunity, we can frame the optimization problem as choosing a quantity for the first investment and a set of plans—
contingent on different realizations of $\alpha$—for the second investment:

$$
\max_{q_0,\{q_1^i, \ldots, q_m^i\}} q_0 \mathbb{E}[r_0] + \sum_{i=1}^{m} p(\alpha_i) q_i^1 \mathbb{E}[r_1(\alpha_i)] - \tau K
$$

(1)

subject to a set of constraints based on a risk measure:

$$
\rho(q_0, q_i^1; \alpha_i) \leq K \quad \forall i = 1, \ldots, m
$$

(2)

Thus, in every future state of the world, the risk of the portfolio must not exceed the firm’s capital.

The first order conditions can be expressed as:

$$
\mathbb{E}[r_0] - \sum_{i=1}^{m} \lambda_i \frac{\partial \rho(q_0, q_i^1; \alpha_i)}{\partial q_0} = 0
$$

(3)

$$
p(\alpha_i) \mathbb{E}[r_1(\alpha_i)] - \lambda_i \frac{\partial \rho(q_0, q_i^1; \alpha_i)}{\partial q_i^1} = 0 \quad \forall i = 1, \ldots, m
$$

(4)

where $\lambda_i$ is the Lagrange multiplier associated with constraint $i$. Equation (3) balances the marginal benefit of exposure to the first risk (the expected return) with its marginal cost, which is a risk cost relating to the risk measure constraint, while equation (4) does the same for the second risk in each of the possible states of the world that could exist. Note in each state that $\lambda_i$ reflects the shadow price or marginal value of state-contingent capital (i.e., at time zero, what an additional dollar of capital in state $i$ (and only in state $i$) would be worth to the firm):

$$
\lambda_i = p(\alpha_i) \frac{\mathbb{E}[r_1(\alpha_i)]}{\frac{\partial \rho(q_0, q_i^1; \alpha_i)}{\partial q_i^1}} \quad \forall i = 1, \ldots, m
$$

(5)

Thus, the marginal cost of risk in each case is equal to the marginal amount of capital consumed by the risk (calculated as the partial derivative of the risk measure with respect to the quantity of the risk) times the value of that capital. In the case of the first risk, that capital is consumed in all possible states of the world since the firm must commit to the exposure, as seen in equation (3).
It is easily verified that this marginal cost can be interpreted as an allocation of capital if the risk measure is homogeneous. That is, the sum of marginal costs times exposures is:

\[ \sum_{i=1}^{m} \lambda_i \left[ \frac{\partial \rho(q_0, q_i; \alpha_i)}{\partial q_0} q_0 + \frac{\partial \rho(q_0, q_i; \alpha_i)}{\partial q_i} q_i \right] = \sum_{i=1}^{m} \lambda_i K \quad (6) \]

or total capital times its value at the margin. Thus, marginal cost has a capital allocation interpretation that “adds up.”

However, it is important to contemplate the nature of the allocation to the first risk. At the time of decision making, the marginal cost of the first risk, and the capital allocated to the first risk, may differ from the marginal cost of that risk after uncertainty associated with \( \alpha \) has been resolved and the portfolio completed. This an \textit{ex ante} allocation that is an average of possible final cost allocations in future states of the world. To illustrate, by using equations (5) and (6), we see that the \textit{ex ante} cost allocation to the first risk can be expressed as:

\[ \sum_{i=1}^{m} p(\alpha_i) \left[ \mathbb{E}[r_1(\alpha_i)] \right] \left[ \frac{\partial \rho(q_0, q_i; \alpha_i)}{\partial q_0} q_0 \right] \]

Thus, for each state, we multiply the probability of state \( i \) times the marginal value of capital in state \( i \) (the first bracketed term) times the capital allocated to the first risk in state \( i \) (the second bracketed term); the total cost allocation is then the sum of these figures over all states.

Remarks on the two-state case:

1. Note that if the uncertainty about the second opportunity concerns only the expected return and not the risk properties, then the allocation \( \frac{\partial \rho(q_0, q_1; \alpha_i)}{\partial q_1} q_1 \) will be the same in all states of the world. In other words, capital allocation will not be affected by the uncertainty. This is almost certainly an artifact of a two-opportunity setup; this is not likely to generalize to more than two opportunities.

2. If we allow capital to be adjusted costlessly after the second investment is made, the
problem changes to:

$$\max_{q_0, \{q_1, \ldots, q_m\}, \{K_i\}} \quad q_0 \mathbb{E}[r_0] + \sum_{i=1}^{m} p(\alpha_i) q_1 \mathbb{E}[r_1(\alpha_i)] - \tau \sum_{i=1}^{m} p(\alpha_i) K_i$$

(7)

subject to a set of constraints based on a risk measure:

$$\rho(q_0, q_1; \alpha_i) \leq K_i \quad \forall i = 1, \ldots, m$$

(8)

We then get

$$\mathbb{E}[r_0] = \tau \sum_{i=1}^{m} p(\alpha_i) \frac{\partial \rho(q_0, q_1; \alpha_i)}{\partial q_0} = 0$$

(9)

This expression can be rearranged to yield a RORAC interpretation,

$$\frac{\mathbb{E}[r_0]}{\sum_{i=1}^{m} p(\alpha_i) \frac{\partial \rho(q_0, q_1; \alpha_i)}{\partial q_0}} = \tau$$

(10)

which equates 1) the expected return divided by the EXPECTED allocated capital with 2) the cost of capital.

On the other hand, due to the homogeneous risk measure $\rho$, the solution to this problem will explode, i.e., investors will invest an infinite amount and keep infinite capital. Under the above RORAC expression, we get a local solution for investment, however, investors can easily double the capital $K$ and the investment $q$ in each state. As a result, the risk measure constraint remains binding while the net profit is doubled.

3. Without allowing costless adjustment of capital as described in the preceding remark, the marginal value of capital will not generally equal $\tau$ in every future state. If we allow capital to be set at the time of the first investment, we will be guaranteed that
the average future marginal value will be equal to $\tau$. That is,
$$\sum_{i=1}^{m} p(\alpha_i) \left[ \frac{E[r_1(\alpha_i)]}{\partial r(q_0, z_1; \alpha_i)} \right] = \sum_{i=1}^{m} \lambda_i = \tau.$$ But if capital is simply inherited and not optimally set, the firm could be undercapitalized or overcapitalized, leading the marginal value of capital to differ from its marginal cost.

2.2 Multiple opportunities

Suppose the financial institution receives $n + 1$ opportunities in sequence from period 0 to period $n$. Upon receipt of the period-zero opportunity, it must decide on the quantity $q_0$ to invest in this known opportunity. Subsequent to this choice, $n$ opportunities will arrive. The risk characteristics of the subsequent opportunities are uncertain, but the possible characteristics, as well as the probabilities associated with their appearance, is known to the decision maker. For the $i$th opportunity, uncertainty about the risk characteristics is resolved upon receipt of the $i$th opportunity, at which point the institution decides on the quantity $q_i$ to invest in that opportunity. We denote the pool of possibilities for the $i$th opportunity as the set $\Omega_i$, and denote the uncertainty in the $i$th period by a random variable $\alpha_i$. The random variable representing all possible states of the world, covering all possible opportunities, is denoted $\alpha = \{\alpha_1, \cdots, \alpha_n\}$.

As in the two-opportunity case, the institution has a capital $K$ with the frictional cost of capital $\tau K$. In every future state of the world, the measured risk of firm’s portfolio must not exceed the firm’s capital. The complete set of possible realizations for $\alpha$ is represented by the set $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, or equivalently, $\Omega^n$. The probability of any particular realization $\alpha \in \Omega^n$ is given by $p(\alpha)$.

If we consider decisionmaking at the time of the period-zero opportunity, the optimization problem is to find the optimal quantity for the period-zero investment and a set of plans—contingent on different states of the world—for all subsequent $n$ opportunities. Moreover, the quantity chosen for the $i$th opportunity must be adapted to the information available the arrival of the $i$th opportunity, since the $i$th decision only depends on the realized opportunity in that period along with the realizations of previous periods and the decisions made up to that point. Denote the random variable reflecting possible state of the
world at the arrival of the $i$th opportunity to be $\alpha^i = \{\alpha_1, \alpha_2, \ldots, \alpha_i\}$, which takes values in the set $\Omega^i = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_i$, each realization having the corresponding probability $p_i(\alpha^i)$. Thus, the $i$th decision $q_i$ must be contingency plan for all possible realizations of the random variable $\alpha^i$. Therefore, the choice variables at the decision time on the current opportunity are $\{q_0, q_i(\alpha^i), \forall i \in [1, \cdots, n], \forall \alpha^i \in \Omega^i\}$. So we can write the optimization problem as:

$$\max_{q_0, \{q_i(\alpha^i), \forall i \in [1,\cdots,n], \forall \alpha^i \in \Omega^i\}} q_0 E[r_0] + \sum_{i=1}^{n} \left[ \sum_{\alpha^i \in \Omega^i} p_i(\alpha^i)q_i(\alpha^i)E[r_i(\alpha^i)] \right] - \tau K \quad (11)$$

subject to a set of constraints that constrain risk for each possible realization of $\alpha$:

$$\rho(q_0, q_1(\alpha^1), \ldots, q_n(\alpha^n); \alpha) \leq K \quad \forall \alpha \in \Omega^n \quad (12)$$

The first order condition for the problem can be expressed as:

$$E[r_0] - \sum_{\alpha \in \Omega} \lambda(\alpha) \frac{\partial \rho(q_0, q_1(\alpha^1), \ldots, q_n(\alpha^n); \alpha)}{\partial q_0} = 0 \quad (13)$$

$$p_j(\alpha^j)E[r_j(\alpha^j)] - \sum_{\alpha \in \Omega^{-\alpha^j}} \lambda(\alpha) \frac{\partial \rho(q_0, q_1(\alpha^1), \ldots, q_n(\alpha^n); \alpha)}{\partial q_j(\alpha^j)} = 0 \quad \forall \alpha^j \in \Omega^j \quad \forall j \in [1, \cdots, n] \quad (14)$$

where $\lambda(\alpha)$ is the Lagrange multiplier associated with constraint for the particular state $\alpha$, and $\Omega^{-\alpha^j}$ denotes the set of all possible states conditional on the first $j$ opportunities having been realized, in other words, $\Omega^{-\alpha^j} \equiv \{\bar{\alpha} | \bar{\alpha}_1 = \alpha_1, \ldots, \bar{\alpha}_j = \alpha_j, \bar{\alpha}_{j+1} \in \Omega_{j+1}, \ldots, \bar{\alpha}_n \in \Omega_n\}$.

Particularly, we can get the first order condition for the last opportunity, when $j = n$, as follows:

$$p_n(\alpha)E[r_n(\alpha)] - \lambda(\alpha) \frac{\partial \rho(q_0, q_1(\alpha^1), \ldots, q_n(\alpha^n); \alpha)}{\partial q_n(\alpha)} = 0 \quad \forall \alpha \in \Omega^n$$

Equation (13) balances the marginal benefit of exposure to the zero-period risk (the expected return) with its marginal cost, which is a risk cost relating to the risk measure
constraint, while equation (14) does the same for the other risks in each of the possible states of the world given the previous path of realizations before the \( j \)th opportunity. Note in each state that \( \lambda(\alpha) \) reflects the shadow price or marginal value of state-contingent capital (i.e., at time zero, what an additional dollar of capital in state \( \alpha \) (and only in state \( \alpha \)) would be worth to the firm):

\[
\lambda(\alpha) = p_n(\alpha) \frac{E[r_n(\alpha)]}{\partial p(q_0, q_1(\alpha_1), ..., q_n(\alpha^n); \alpha)} \quad \forall \alpha \in \Omega^n \quad (15)
\]

Thus, the marginal cost of risk in each final state is equal to the marginal amount of capital consumed by the risk (calculated as the partial derivative of the risk measure with respect to the quantity of the risk) times the value of that capital. In the case of the period-zero risk, the marginal cost of risk at the time that \( q_0 \) is selected is an expected marginal cost, since both the amount of capital that the period-zero risk will ultimately consume and the value of a marginal unit of capital will depend on the course of future opportunity realizations, as revealed in equation (13).

If the risk measure \( \rho \) is homogeneous, we can get

\[
\frac{\partial \rho(q_0, q_1(\alpha_1), ..., q_n(\alpha^n); \alpha)}{\partial q_0} q_0 + \sum_{i=1}^{n} \frac{\partial \rho(q_0, q_1(\alpha_1), ..., q_n(\alpha^n); \alpha)}{\partial q_i(\alpha^i)} q_i(\alpha^i) = K \quad \forall \alpha \in \Omega^n \quad (16)
\]

and thus, we have \( \forall \alpha \in \Omega^n \)

\[
\sum_{\alpha \in \Omega^n} \lambda(\alpha) \left[ \frac{\partial \rho(q_0, q_1(\alpha_1), ..., q_n(\alpha^n); \alpha)}{\partial q_0} q_0 + \sum_{i=1}^{n} \frac{\partial \rho(q_0, q_1(\alpha_1), ..., q_n(\alpha^n); \alpha)}{\partial q_i(\alpha^i)} q_i(\alpha^i) \right] = \sum_{\alpha \in \Omega^n} \lambda(\alpha) K
\]

Therefore, marginal cost has a capital allocation interpretation that “adds up” in the multiple opportunity case as well. If we consider the capital allocation to the \( j \)th risk, we can have:

\[
\sum_{\alpha \in \Omega^n} p_n(\alpha) \left[ \frac{E[r_n(\alpha)]}{\partial p(q_0, q_1(\alpha_1), ..., q_n(\alpha^n); \alpha)} \right] \left[ \frac{\partial \rho(q_0, q_1(\alpha_1), ..., q_n(\alpha^n); \alpha)}{\partial q_j(\alpha^j)} q_j(\alpha^j) \right]
\]

For each final state of the world, we multiply the probability of the state \( \alpha \) times the marginal value of capital in state \( \alpha \) times the capital allocated to the \( j \)th risk in state \( \alpha \);
the total cost allocation is then the sum of these figures over all states in $\Omega^n$. Moreover, the capital allocation can be rewritten as the following:

$$\sum_{\alpha^j \in \Omega} p_j(\alpha^j) \left\{ \sum_{\alpha \in \Omega - \alpha^j} \Pr(\alpha|\alpha^j) \left[ \frac{\E[r_n(\alpha)]}{\frac{\partial \rho(q_0, q_1(\alpha^1), \ldots, q_n(\alpha^n); \alpha)}{\partial q_n(\alpha)}} \right] \left[ \frac{\partial \rho(q_0, q_1(\alpha^1), \ldots, q_n(\alpha^n); \alpha)}{\partial q_j(\alpha)} \right] q_j(\alpha^j) \right\}$$

For each state of the world, given a particular path until the period $j$, we multiply the conditional probability of the remaining possible states times the marginal value of capital in the state times the capital allocation to the $j$th risk in the state; the total cost allocation for the particular path $\alpha^j$ is the sum of all the figures over all sub-states; the sum of all the production of the path-contingent capital allocation times the probability of the path is the total cost allocation.

In particular, the capital allocation to the period-zero risk can be expressed as the following:

$$\sum_{\alpha \in \Omega^n} p_n(\alpha) \left[ \frac{\E[r_n(\alpha)]}{\frac{\partial \rho(q_0, q_1(\alpha^1), \ldots, q_n(\alpha^n); \alpha)}{\partial q_n(\alpha)}} \right] \left[ \frac{\partial \rho(q_0, q_1(\alpha^1), \ldots, q_n(\alpha^n); \alpha)}{\partial q_0} \right] q_0$$

and the capital allocation to the last period risk, in an *ex ante* sense, is:

$$\sum_{\alpha \in \Omega^n} p_n(\alpha) q_n(\alpha) \E[r_n(\alpha)]$$

As in the two opportunity case, we can recover a RORAC result if we allow capital to be adjusted costlessly at the terminal period. Capital then becomes a state contingent choice variable:

$$\max_{q_0, \{q_i(\alpha^i), \forall i \in [1, \ldots, n], \forall \alpha^i \in \Omega^i\}, \{K_\alpha\}} q_0 \E[r_0] + \sum_{i=1}^n \left[ \sum_{\alpha^i \in \Omega^i} p_i(\alpha^i) q_i(\alpha^i) \E[r_i(\alpha^i)] \right] - \tau \sum_{\alpha \in \Omega^n} p_n(\alpha) K_\alpha$$

(17)

again subject to a set of constraints that constrain risk for each possible realization of $\alpha$, with the difference that capital can be varied in each state.
\[ \rho(q_0, q_1(\alpha^1), ..., q_n(\alpha^n) ; \alpha) \leq K_{\alpha} \quad \forall \alpha \in \Omega^n \tag{18} \]

Assuming all constraints hold at equality, we again get the result that

\[ \frac{\mathbb{E}[r_0]}{\sum_{\alpha \in \Omega} P_n(\alpha) \frac{\partial \rho(q_0, q_1(\alpha^1), ..., q_n(\alpha^n) ; \alpha)}{\partial q_0}} = \tau \]

or that the ratio of the expected return to EXPECTED allocated capital is equated to the hurdle rate, which is the cost of capital.

Similar to the two period case, this particular problem with a flexible capital is explosive due to the same reason. As long as we find a certain solution to the above RORAC equation, we can easily double the capital amount and the investment for every state, so that we have the same binding constraints but with a double profit than the previous local optimal profit. To avoid such explosive issue, we can either maximize the ratio of return and risk measure, i.e.,

\[ \max_{q_0, \{q_i(\alpha^i) \mid i \in [1, \cdots, n] \}, \{K_\alpha\}} \frac{q_0 \mathbb{E}[r_0] + \sum_{i=1}^{n} \left[ \sum_{\alpha^i \in \Omega^i} p_i(\alpha^i) q_i(\alpha^i) \mathbb{E}[r_i(\alpha^i)] \right]}{\tau K} \]

or use the risk measure instead of capital as base for calculating capital cost, i.e.,

\[ \max_{q_0, \{q_i(\alpha^i) \mid i \in [1, \cdots, n] \}, \{K_\alpha\}} q_0 \mathbb{E}[r_0] + \sum_{i=1}^{n} \left[ \sum_{\alpha^i \in \Omega^i} p_i(\alpha^i) q_i(\alpha^i) \mathbb{E}[r_i(\alpha^i)] \right] - \tau \rho \]

such that

\[ \rho(q_0, q_1(\alpha^1), ..., q_n(\alpha^n) ; \alpha) \leq K \quad \forall \alpha \in \Omega^n \]

3 Numerical Example

Based on the model setting above, we use a simple numerical example to investigate how uncertainty regarding future opportunities affects the optimal choice for current opportunity. In other words, we are looking into how the period-zero uncertainty about the distributions of future \( r_i \) affects the optimal \( q_0 \). After all the decisions have been made, the
final portfolio return is the random variable:

\[ R(\alpha) \equiv \sum_{i=0}^{n} q_i(\alpha) r_i(\alpha^i) \quad \forall \alpha \in \Omega \]

For the risk measure, we use value-at-risk adjusted by the mean. Note that the left tail of the portfolio distribution represents catastrophic loss and the confidence level is \( 1 - \theta \). Thus the risk measure constraint is

\[ \rho(q_0, q_1(\alpha^1), ..., q_n(\alpha^n); \alpha) \equiv E[R(\alpha)] - \text{VaR}_\theta(R(\alpha)) \leq K \quad \forall \alpha \in \Omega. \]

The example consists of two parts. In Section 3.1, we analyze effects of uncertainty in the mean return of future opportunities while we holding the variance constant. In Section 3.2, we consider the effects of uncertainty in the variance of future opportunities while holding the mean return constant.

Throughout, we assume that all opportunities follow independent normal distributions, i.e., \( r_i \sim N(\mu, \sigma) \) for all \( i \in [0, m] \). Moreover, we define \( \Phi(\cdot) \) as the cumulative distribution function of standard normal.

### 3.1 Uncertainty in Expected Returns

For simplicity, we model the uncertainty of future opportunities with a binary distribution, with a high state and low state. That is, \( \mu_0 = \mu \) and \( \mu_i = \mu \pm \phi \mu \) for any \( i \in [1, m] \). We will study the effect of \( \phi \mu \) on optimal \( q_0 \).

First, we consider the two-period problem,

\[
\max_{q_0, q_1^H, q_1^L} E[r_0 q_0 + \frac{1}{2} r_1^H q_1^H + \frac{1}{2} r_1^L q_1^L - \tau K]
\]

subject to risk measure constraints

\[
E[R^H] - \text{VaR}_\theta(R^H) \leq K \\
E[R^L] - \text{VaR}_\theta(R^L) \leq K
\]
Since we assume opportunities in different periods are independent, thus the risk measure of final portfolio only depends on their variance. Simplifying, we have

$$\max_{q_0, q_1^H, q_1^L} \mu q_0 + \frac{1}{2} q_1^H (\mu + \phi \mu) + \frac{1}{2} q_1^L (\mu - \phi \mu) - \tau K$$

subject to

$$-\sigma \sqrt{q_0^2 + q_1^H^2} \Phi(\theta) \leq K$$
$$-\sigma \sqrt{q_0^2 + q_1^L^2} \Phi(\theta) \leq K,$$

where $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution. (We further restrict $q_0$ and $q_1$ to be nonnegative.)

Since both constraints should be binding, we know that $q_1^H = q_1^L$. Thus $\phi \mu$ disappears in the objective function. Therefore, $\phi \mu$ doesn’t affect the optimal $q_0$ in the two-period case, confirming Remark 1 in Section 2.1.

We next consider a three-period case, rewriting the problem as

$$\max_{q_0, q_1^H, q_1^L, q_2^H, q_2^L, q_3^H, q_3^L} \mu q_0 + \frac{1}{2} q_1^H (\mu^H) + \frac{1}{2} q_1^L (\mu^L) + \frac{1}{4} q_2^H (\mu^H) + \frac{1}{4} q_2^L (\mu^L) + \frac{1}{4} q_2^H (\mu^H) + \frac{1}{4} q_2^L (\mu^L) - \tau K$$

subject to

$$-\sigma \sqrt{q_0^2 + q_1^H^2 + q_2^H^2} \Phi(\theta) \leq K$$
$$-\sigma \sqrt{q_0^2 + q_1^L^2 + q_2^L^2} \Phi(\theta) \leq K$$
$$-\sigma \sqrt{q_0^2 + q_1^L^2 + q_2^L^2} \Phi(\theta) \leq K$$
$$-\sigma \sqrt{q_0^2 + q_1^L^2 + q_2^L^2} \Phi(\theta) \leq K,$$

where $\mu^H = \mu + \phi \mu$, and $\mu^L = \mu - \phi \mu$, and all choice variables are again restricted to be nonnegative.

We now look at results for the case where $\mu = 10$, $\sigma = 20$, $\tau = 0.2$, $K = 18$, and $\theta = 5\%$. 

71
After optimizing the above problem with different $\phi_\mu$, we get the following results.

First, in two-period example, the optimal $q_0$ is constant when varying $\phi_\mu$ (as expected).

Similarly, the optimal $q_1$ also doesn’t depend on $\phi_\mu$ either. As noted earlier, this follows analytically from the fact that the risk measure, by design, is unaffected by the mean return (since the expected profit is removed).

Next, let’s consider the numerical results in the 3-period case. The optimal $q_0$ is decreasing with $\phi_\mu$ as shown in Figure 3. The reason is that when the uncertainty in the future mean return is enlarged in following periods, it becomes optimal to “leave room” for the possible good opportunities in the future.
Similar logic guides decision-making in the first period, as seen in Figure 4. As uncertainty increases, the optimal quantity for period-1 risk also rises for the period-1 realization. It decreases for bad realizations. The intuition is that when the uncertainty is large, a good realization in period-1 is so attractive that the individual prefers to invest more in period-1 rather than take the risk of being stuck with a bad realization in period-2. On the other hand, if the period-1 realization is bad, the decisionmaker has little to lose and much to gain by waiting, since the future could bring a favorable opportunity.

Figure 5 shows how the optimal $q_2$ varies with expected return uncertainty. Here we see that the key determinant is the period 1 realization. In this example, period 2 is the final period, so remaining capacity is dedicated to the final opportunity regardless of whether it is a high or low realization. The remaining capacity depends on what happened in period 1—where, as we saw in Figure 4, increasing uncertainty led to increasing investment, with
the obvious consequence that less capacity was left for the final opportunity, regardless of outcome. Conversely, increasing uncertainty leads to higher investment in the final period when the period 1 realization is unfavorable—since capacity is increasingly deferred to the final period when the period 1 opportunity is unattractive.

![Figure 5: $q_2$ depending on $\phi_\mu$ in 3-period](image)

In summary, this example illustrates how increasing uncertainty about expected returns associated with future opportunities can lead companies to optimally reduce investment in current opportunities when those opportunities are unattractive in relation to future opportunities, should favorable conditions materialize.

### 3.2 Uncertainty in Variance of Risk

Now we explore the effects of uncertainty about the variance of future opportunities while keeping the mean constant. We continue to assume that all the returns in different periods follow independent normal distributions, i.e., $r_i \sim N(\mu, \sigma_i)$ for all $i \in [0, m]$. Now there is uncertainty in the standard deviation of each future opportunity, with two equally likely possibilities—a high state and low state. That is, $\sigma_0 = \sigma$ and $\sigma_i = \sigma \pm \phi_\sigma, \phi_\sigma \in (0, \sigma)$ for any $i \in [1, m]$. We want to understand the effect of $\phi_\sigma$ on optimal $q_0$.

First, we consider the two-period problem,

$$
\max_{q_0, q_1^H, q_1^L} \mu q_0 + \frac{1}{2} q_1^H \mu + \frac{1}{2} q_1^L \mu - \tau K
$$
subject to

\[-\sqrt{q_0^2 \sigma^2 + q_1^{H^2} \sigma^{H^2} \Phi(\theta)} \leq K\]
\[-\sqrt{q_0^2 \sigma^2 + q_1^{L^2} \sigma^{L^2} \Phi(\theta)} \leq K,\]

where $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution and $\sigma^H = \sigma + \phi_{\sigma}$, and $\sigma^L = \sigma - \phi_{\sigma}$. It is worth mentioning that the returns of different opportunities are independent.

Figure 6 shows that the optimal quantity of the period zero investment is decreasing with uncertainty about the variance of future risk. When the uncertainty becomes larger, the optimal quantity for current investment becomes lower. The intuition is that, as uncertainty rises, the low variance state becomes increasingly profitable, and it becomes imperative for the decisionmaker to “leave room” for the low variance opportunity should it materialize.

![Figure 6: $q_0$ depending on $\phi_{\sigma}$ in 2-period](image)

Naturally, it follows that the increasing deferral of capacity to the second period leads to greater investment in the second opportunity when the low variance state is realized. This is shown in Figure 7, which also shows that investment in the second opportunity falls with uncertainty when high variance state is realized. Evidently, the positive influence of capacity deferral is offset by the negative influence of the higher variance in the high state.
Now, consider a three-period case. The problem becomes:

$$\max_{q_0,q_1,q_2} \mu q_0 + \frac{1}{2} q_1^H \mu + \frac{1}{2} q_1^L \mu + \frac{1}{4} q_2^{HH} \mu + \frac{1}{4} q_2^{HL} \mu + \frac{1}{4} q_2^{LH} \mu + \frac{1}{4} q_2^{LL} \mu - \tau K$$

subject to

$$-\sqrt{q_0^2 \sigma^2 + q_1^H \sigma^2 + q_2^{HH} \sigma^2} \phi(\theta) \leq K$$

$$-\sqrt{q_0^2 \sigma^2 + q_1^L \sigma^2 + q_2^{HL} \sigma^2} \phi(\theta) \leq K$$

$$-\sqrt{q_0^2 \sigma^2 + q_1^L \sigma^2 + q_2^{LH} \sigma^2} \phi(\theta) \leq K$$

$$-\sqrt{q_0^2 \sigma^2 + q_1^L \sigma^2 + q_2^{LL} \sigma^2} \phi(\theta) \leq K,$$

where $q_0$, $q_1$, and $q_2$ are nonnegative. The results are similar to those in the two period case, as shown in Figures 8, 9 and 10.
Figure 8: \( q_0 \) depending on \( \phi_\sigma \) in 3-period

Figure 9: \( q_1 \) depending on \( \phi_\sigma \) in 3-period

Figure 10: \( q_2 \) depending on \( \phi_\sigma \) in 3-period
In summary, in a sequential portfolio optimization problem, the optimal investment depends on uncertainty in the variance of future opportunities. If the future opportunities have higher variances, the individual will invest less in period-0 when no information is given. Deferral effects similar to those in Section 3.1 are also seen: when the uncertainty of variance becomes larger, optimal period one investment in a high variance opportunity decreases—as it becomes optimal to defer capacity to the future in hope of a better draw. On the other hand, low variance outcomes in period 1 are met with increasing investment as uncertainty rises.

4 Conclusion

Reinsurers, private equity firms, and other financial institutions investing in illiquid opportunities must take into account the expected risk-return profile of future opportunities when making current investments. Current capital allocation methods, which are largely developed in static full information settings, must be adjusted to this reality. RORAC evaluation of investments in this setting must consider an average of possible future capital allocations, rather than a single capital allocation based on a pro forma portfolio.

We have shown that such adjustment is feasible if the uncertainties are understood. Capital allocation can proceed in a generalized form, with risk pricing reflecting averages of capital allocations in the future. Failure to explicitly model uncertainties about future opportunities can lead to misstatement of the marginal cost of risk, and, as a result, misallocation of capacity in current opportunities. At least in some cases, we have shown that rising uncertainty should lead firms to reduce investment in current opportunities and leave capacity for attractive future opportunities should they appear, which mirrors the intuition of the impact of uncertainty offered by the real options literature.
References


