

NBER WORKING PAPER SERIES

MULTIDIMENSIONAL PRIVATE INFORMATION, MARKET STRUCTURE AND
INSURANCE MARKETS

Hanming Fang
Zenán Wu

Working Paper 22773
<http://www.nber.org/papers/w22773>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2016

We would like to thank Eduardo Azevedo, David de Meza, Ben Lester, Stephen Morris, Yeneng Sun, Venky Venkateswaran, Glen Weyl, and seminar participants at National University of Singapore, Monash University and Rice University for helpful discussions, suggestions and comments. Part of Fang's research on this project is funded by the generous financial support from NSF Grant SES-0844845. All remaining errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Hanming Fang and Zenán Wu. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Multidimensional Private Information, Market Structure and Insurance Markets
Hanming Fang and Zenan Wu
NBER Working Paper No. 22773
October 2016
JEL No. D82,G22,H11

ABSTRACT

A large empirical literature found that the correlation between insurance purchase and ex post realization of risk is often statistically insignificant or negative. This is inconsistent with the predictions from the classic models of insurance a la Akerlof (1970), Pauly (1974) and Rothschild and Stiglitz (1976) where consumers have one-dimensional heterogeneity in their risk types. It is suggested that selection based on multidimensional private information, e.g., risk and risk preference types, may be able to explain the empirical findings. In this paper, we systematically investigate whether selection based on multidimensional private information in risk and risk preferences, can, under different market structures, result in a negative correlation in equilibrium between insurance coverage and ex post realization of risk. We show that if the insurance market is perfectly competitive, selection based on multidimensional private information does not result in negative correlation property in equilibrium, unless there is a sufficiently high loading factor. If the insurance market is monopolistic or imperfectly competitive, however, we show that it is possible to generate negative correlation property in equilibrium when risk and risk preference types are sufficiently negative dependent, a notion we formalize using the concept of copula. We also clarify the connections between some of the important concepts such as adverse/advantageous selection and positive/negative correlation property.

Hanming Fang
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104
and NBER
hanming.fang@econ.upenn.edu

Zenan Wu
School of Economics
Peking University
Beijing, China
zenan@pku.edu.cn

1 Introduction

The classic asymmetric information models of insurance pioneered by Arrow (1963), Pauly (1974), Rothschild and Stiglitz (1976) and Wilson (1977) assume that potential insurance buyers have *one*-dimensional private information regarding their risk type. These models predict a positive correlation between insurance coverage and *ex post* realizations of losses. The reason is *ex ante* adverse selection, namely, that the “bad risks” (i.e., those relatively likely to suffer a loss) have a higher willingness to pay for insurance; and allowing for *ex post* moral hazard only strengthens the positive correlation between coverage and *ex post* losses. This “*positive correlation property*” of the classic asymmetric information models forms the basis for empirical tests of asymmetric information in several recent papers (see Chiappori and Salanié 2000).

However, the results from a growing empirical literature testing for the correlation between insurance coverage and *ex post* realization of risks are mixed and vary by market. In an auto insurance market, Chiappori and Salanié (2000) find that accident rates for young French drivers who choose comprehensive automobile insurance is *not statistically different* from those opting for the legal minimum coverage, after controlling for observable characteristics known to automobile insurers. In contrast, Cohen (2005), using data from an online Israeli insurer, finds that new auto insurance customers choosing a low deductible tend to have more accidents, leading to higher total losses for the insurer.¹ In a life insurance market, Cawley and Philipson (1999) find that the mortality rate of U.S. males who purchase life insurance is *below* that of the uninsured, even when controlling for many factors such as income that may be correlated with life expectancy.² For long term care (LTC) insurance market, Finkelstein and McGarry (2006), using panel data from a sample of Americans born before 1923 (the AHEAD study), find *no statistically significant* correlation between LTC coverage in 1995 and use of nursing home care between 1995-2000, even after controlling for insurers’ assessment of a person’s risk type; moreover, when they use whether respondents undertake various types of preventive health care as a proxy for risk aversion, they find that people who are more risk averse by this measure are both more likely to own LTC insurance and less likely to enter a nursing home. In an annuity insurance market, Finkelstein and Poterba (2004) find systematic relationships between *ex post* mortality and annuity characteristics, such as the timing of payments and the possibility of payments to the annuitants’ estate, but they do not find evidence of substantive mortality differences by annuity size. For Medigap insurance market, Fang, Keane and Silverman (2008) find that, conditional on controls for Medigap prices, those with Medigap spend on average \$4,000 *less* on medical care than those without, providing a strong evidence for *negative correlation* between Medigap purchase and *ex post* realization of risk.

These empirical findings fueled an emerging literature on the possibility that *multidimensional*

¹Others have examined the evidence of asymmetric information in the choice of insurance contracts such as deductibles and co-payments etc. For example, Puelz and Snow (1994) study automobile collision insurance and argue that, in an adverse selection equilibrium, individuals with lower risk will choose a contract with a higher deductible, and contracts with higher deductibles should be associated with lower average prices for coverage. They find evidence in support of each of these predictions using data from an automobile insurer in Georgia. However, see Chiappori and Salanié (2000) and Dionne et al. (2001) for critiques of the Puelz and Snow study.

²See He (2009) for a re-examination of the evidence.

private information may lead to what has been called “advantageous selection.”³ The formal theoretical literature is sparse. de Meza and Webb (2001) postulate a model in which individuals differ in their risk preferences, which they refer to as “timid” and “bold” types. They assume that more timid types may lower their risk exposure through increased insurance purchase and greater precautionary effort to reduce risks. They show that, in the presence of administrative costs in processing claims and issuing policies, there exists a pure-strategy, partial pooling, subgame-perfect Nash equilibrium in the insurance market that exhibits the negative correlation property. Thus, failure to condition on risk aversion may then mask the positive correlation between insurance coverage and *ex post* risk predicted by one-dimensional models. Following de Meza and Webb (2001), the existing literature points to risk preferences as the primary suspect behind advantageous selection. In general, however, *any* private information could function as a source of advantageous selection if it is positively correlated with insurance coverage and at the same time negatively correlated with risk. Finkelstein and McGarry (2006) argue that their findings on LTC insurance market is consistent with multidimensional private information and advantageous selection based on risk aversion; in fact, their findings suggest that, on net, adverse selection based on risk and advantageous selection based on risk aversion roughly cancel out in the LTC insurance market. Fang, Keane and Silverman (2008) find that, for Medigap insurance market, risk preferences do not appear as a source of advantageous selection, but cognitive ability is particularly important.

However, to the best of our knowledge, the precise conditions under which whether selection based on multidimensional private information may generate *in equilibrium* a positive or negative correlation between insurance purchase and *ex post* realization of risk is still unknown. Most of the existing papers that invoked the possibility of multidimensional private information as a possible explanation for the empirical findings discussed above rely on *partial* equilibrium intuition (much in the spirit of Hemenway, 1990).⁴ The goal of this paper is to fill in this gap. We present a simple model of insurance market where consumers have multidimensional private information in risk and risk preference types, and systematically investigate whether selection based on multidimensional private information can, under different market structures, result in negative correlation *in equilibrium* between insurance coverage and *ex post* realization of risk. We show that if the insurance market is perfectly competitive, selection based on multidimensional private information does *not* generate negative correlation property in equilibrium unless there is a sufficiently high loading factor. If the insurance market is monopolistic, however, we show that it is possible to generate negative correlation property in equilibrium when risk type and risk preference type are sufficiently negative dependent, a notion we formalize using the concept of copula. We further show that this result generalizes to imperfectly competitive market structure.

³The first description of this phenomenon in the economics literature appears to be Hemenway (1990), who used the term “propitious selection.”

⁴An important paper by Chiappori, Jullien, Salanié and Salanié (2006, henceforth CJSS) argue that in a competitive insurance market the positive correlation property is a general implication of insurance models with asymmetric information even when the private information is multidimensional in risk and risk preferences. The key assumptions are consumer rationality and a condition which they refer to as “nonincreasing profit” condition - that is, the per contract expected profit does not increase with the generosity of the contract. We will discuss the connection between our results with their results in Section 4.

The remainder of the paper is structured as follows. In Section 2 we provide a detailed discussion of the related literature. In Section 3, we describe our model environment in which consumers are heterogeneous in both risk and risk preference types. In Section 4, we consider the perfectly competitive market structure and show that, unless the loading factor is sufficiently high, selection based on multidimensional private information does *not* generate negative correlation property in equilibrium, regardless of the dependence structure between the two dimensions of private information. In Section 5, we analyze the other polar case of monopolistic market structure where we show that it is possible to generate negative correlation property in equilibrium when risk type and risk preference type are sufficiently negative dependent, a notion we formalize using the concept of copula. In Section 6, we clarify the confusions in this growing literature about the connections between some of the important concepts such as adverse/advantageous selection and positive/negative correlation property. In Section 7, we introduce a parameterization of the imperfect market structure and show that our results for the monopoly case derived in Section 5 are robust. In Section 8, we partially endogenize the contract space and again show that our results for the single contract case derived in Sections 4 and 5 continue to hold with natural and mild additional assumptions. In Section 9, we summarize our main findings and suggest directions for future research.

2 Related Literature

To the extent that our paper investigates on whether positive correlation property is robust to environments with multidimensional consumer heterogeneity, it is most related to CJSS (2006) and de Meza and Webb (2016). CJSS argue that, as long as consumers are rational and the per contract expected profit does not increase with the generosity of the contract (which they refer to as “nonincreasing profit” (NIP) condition), then the positive correlation property is robust to multidimensional private information in a competitive insurance market. This conclusion is similar to our results for the competitive insurance market presented in Propositions 1-2 and Proposition 8. Their results are proved using the revealed preference argument implied by the hypothesized consumer rationality and the nonincreasing profit condition assumed on the supply side, and as such they do not have to exploit the full set of equilibrium restrictions. In contrast, we exploit the full set of the equilibrium restrictions and as a result our positive correlation predictions for the competitive insurance market are sharper for the case of proportional loading factor case. We will provide more details of the comparison when we discuss Proposition 2. Also related, de Meza and Webb (2016) provide an insightful discussion that the positive correlation test is not a valid test to distinguish asymmetric information from symmetric information environment. The reason is that under symmetric information, the only insurance purchased by consumers will be full cover, unless the claim processing costs (or loading factors) are formally modeled. The results in our paper is more relevant in distinguishing multi- vs. one-dimensional private information models of insurance as opposed to symmetric vs. asymmetric information models of insurance.

Our paper is related to a recent literature that attempts to analyze the selection markets with potentially multidimensional private information. Einav, Finkelstein and Cullen (2010) propose an

approach to conduct empirical welfare analysis in insurance markets based on directly estimating the demand and average (and marginal) cost curves using exogenous variations in prices.⁵ Based on their graphical analysis of the demand and cost curves for the selection market, where the defining feature is that insurers’ costs depend on which consumers purchase their products and hence are endogenous to price, they also argue that slope of the estimated marginal cost curve provides a direct test of the existence and nature of selection, i.e., whether the selection is adverse or advantageous. Specifically, they argue that a rejection of the null hypothesis of a constant marginal cost curve is a rejection of the null hypothesis of no selection, whereas the selection is adverse if marginal cost is increasing in price or advantageous if marginal cost is decreasing in price. They also emphasize that an attractive feature of their approach of relying only on the estimated demand and cost curves is that “it does not require the researcher to make (often difficult-to-test) assumptions about consumers’ preferences or the nature of *ex ante* information” (Einav, Finkelstein and Cullen, 2010, p. 879). In Section 6, we show that an important limitation of their approach is that *generically* the marginal cost curve is *non-monotonic* when consumer heterogeneity is multidimensional.⁶ In fact, in our setting with two-dimensional private information in risk and risk aversion, it is monotonic only when the two dimensions are perfectly correlated. Therefore, depending on the range of the price variations available in the data that is used to estimate the demand and cost curves, it is likely the estimated cost curve only reflects the nature of the selection –adverse or advantageous– *locally*.⁷ We also provide an example (Example 4) in which the nature of the *local* selection being advantageous at the equilibrium price level does not imply a negative correlation between insurance purchase and *ex post* realization of risk in equilibrium.

We explicitly model consumers’ multidimensional heterogeneity in this paper. The intuition for why marginal cost curve is unlikely to be monotonically increasing in market size in our model can be easily explained. Consider an environment where consumers’ willingness to pay for insurance is increasing in their risk type m and risk aversion type λ . Suppose that risk type and risk aversion type are bounded in $[\underline{m}, \bar{m}]$ and $[\underline{\lambda}, \bar{\lambda}]$ respectively. Then the marginal cost of insurance when the market size is close to 0 will be close to \bar{m} and the marginal cost when the market size is close to 1 will be close to \underline{m} . That is, the marginal cost curve must always have at least one decreasing segment! Another benefit of modeling consumers’ multidimensional heterogeneity explicitly is that it allows us to examine how the market outcome changes when the dependence structure of individuals’ multidimensional heterogeneity varies. We use the concept of copula to parameterize the degree of dependence between the consumers’ multidimensional types.

Mahoney and Weyl (2016) build and analyze a model of imperfect competition in selection markets. They parameterize the degree of both market power and selection, and use graphical price-

⁵See also Einav, Finkelstein and Levin (2010), Einav and Finkelstein (2011) and Chetty and Finkelstein (2013) for related discussions of the demand and cost analysis of selection markets.

⁶Einav and Finkelstein (2011) are aware of the non-monotonicity issue of marginal cost curve, as they stated: “More generally, once we allow for preference heterogeneity, the marginal cost curve needs not be monotone. However, for simplicity and clarity we focus our discussion on the polar cases of monotone cost curves.” (footnote 7, p. 124).

⁷Einav, Finkelstein and Cullen (2010) have a total of six price levels (or three price levels if one only considers those with somewhat large number of consumers). Also, the range of the price variations is quite limited.

theoretic reasoning to analyze the interactions between selection and imperfect competition. Their parameterization of selection follows Einav, Finkelstein and Cullen (2010) by hypothesizing whether the marginal cost curve is upward or downward sloping, and as such it has the same limitation that globally increasing marginal cost curves are in fact non-generic in models of multidimensional heterogeneity.

Azevedo and Gottlieb (2016) propose an equilibrium concept for *competitive insurance market* where consumers may have multidimensional heterogeneity, and insurance companies compete for consumers by choosing contracts from a compact space and setting their corresponding prices. Their equilibrium concept, which relies on perturbations, guarantees existence. Veiga and Weyl (2016) instead study the incentives for a *monopolistic* insurer in its choice of insurance quality facing consumers with multidimensional heterogeneity. They derive a condition of the optimal insurance quality to emphasize that the sorting incentives of the monopolist is the ratio of two terms: the numerator is the covariance among marginal consumers between the marginal willingness to pay for quality and the cost for the firm, and the denominator is marginal consumer surplus, which measures market power. The analysis of Veiga and Weyl (2016) focus on the *marginal* consumers, and does not analyze correlation between insurance purchase and *ex post* realization of risk which is about the *average* insurance buyers and non-buyers; and they also focus on monopolistic market structure, while our paper highlights the interactions between multidimensional heterogeneity and market structure.⁸ In the basic model of our paper, the quality of insurance is assumed to be exogenously fixed, and we focus on the determination of premium; while Veiga and Weyl (2016) focuses on how the monopolistic insurer determines the profit-maximizing quality of insurance. Neither Azevedo and Gottlieb (2016) nor Veiga and Weyl (2016) analyze whether multidimensional consumer heterogeneity can generate negative correlation between insurance purchase and *ex post* realization of risk in equilibrium.⁹

3 The Model

Consumers. There is a continuum of consumers with heterogeneous types indexed by $\theta \equiv (m, \lambda)$, where $m \in [\underline{m}, \bar{m}]$ with $0 \leq \underline{m} < \bar{m} < \infty$ denotes consumer's risk type, and $\lambda \in [\underline{\lambda}, \bar{\lambda}]$ with $0 < \underline{\lambda} < \bar{\lambda} < \infty$ can be interpreted as any other characteristics of the consumer that may be related to his/her risk preference.¹⁰ As a notational convention, we use M and Λ respectively to denote the random variables for risk type and risk preference type, and their lowercase counterparts as their realizations. In the population, consumers' type, (m, λ) , is assumed to be independently

⁸Also related to our paper, Weyl and Veiga (2014) offer a quantitative strengthening of the notion of affiliation for multidimensional random vectors that is useful to relate dependence between risk type and risk preferences to the direction of selection.

⁹It should be noted that Azevedo and Gottlieb (2016) introduce a notion of *intensive margin selection coefficient* that measures the difference between the marginal changes of the premium and the cost of insuring the marginal consumers, both with respect to the insurance coverage. They suggest that this notion is related to the positive correlation test. We will discuss its connection with our results in Sections 6 and 8 below.

¹⁰We assume that m and λ are bounded above by \bar{m} and $\bar{\lambda}$ respectively for the simplicity in describing some of the intuitions for our results. All of our results are robust to unbounded supports.

drawn from joint CDF $H(m, \lambda)$, and we denote the marginal CDF of M and Λ by $F(m)$ and $G(\lambda)$ respectively. We use $h(m, \lambda)$, $f(m)$ and $g(\lambda)$ to denote the corresponding joint and marginal density functions of (M, Λ) , M and Λ , respectively. We assume that the marginal distribution of M is such that M has finite mean, denoted by $E[M]$. We allow *dependence* between M and Λ in this paper, and we will discuss the form of the dependence in details in Section 5.

Insurance Contract. Consumers decide whether or not to purchase insurance. In the basic model, we assume that insurance firms are regulated in the sense that they can *only* provide insurance with quality $x \in (0, 1]$ where a higher x indicating a contract with better coverage. Note that we assume that the insurance coverage quality x is *not* a choice variable for the firms. As such, our setup is in the spirit of Akerlof (1970) where insurance contract is exogenously given, rather than Rothschild and Stiglitz (1976) where the insurance quality x is endogenously chosen.

The cost to the insurance firms, *not inclusive of any loading cost*, for providing quality- x insurance to a type- $\theta \equiv (m, \lambda)$ consumer, denoted by $C(\theta; x)$, is increasing in consumer's risk type m and the coverage quality x . Note that it does *not* depend on the consumer's risk preference type λ . In particular, we let $C(\theta; x) = x \cdot c(m)$ where, without loss of generality, we assume that $c'(\cdot) > 0$.

Consumer Preference. Type- $\theta \equiv (m, \lambda)$ consumer's willingness to pay (WTP) for an insurance policy with quality x is denoted by $v(\theta; x)$. We make the following assumptions on $v(\cdot)$:

Assumption 1 $\partial v / \partial x > 0$, $\partial v / \partial m > 0$ and $\partial v / \partial \lambda > 0$.

Assumption 1 simply says that consumer's WTP for insurance is increasing in her risk type, in her risk preference type and the quality of the contract coverage.

Assumption 2 $v(\theta; x) > C(\theta; x) \equiv x \cdot c(m)$, for all θ .

Assumption 2 holds for many economic framework of insurance as long as individuals are risk-averse.¹¹ The difference between the WTP for insurance $v(\theta; x)$ and $C(\theta; x)$ is commonly referred to as the *risk premium* for type- (m, λ) consumer.

Facing a premium p for insurance coverage x , a type- (m, λ) consumer purchases insurance if and only if $p \leq v((m, \lambda); x)$. We use

$$\mathcal{B}(p) \equiv \{\theta : v(\theta; x) \geq p\} \tag{1}$$

to denote the set of consumers whose WTP for the insurance exceeds the premium and thus they are the set of *buyers*; and use

$$\mathcal{NB}(p) \equiv \{\theta : v(\theta; x) < p\} \tag{2}$$

to denote the set of *non-buyers* at price p .

¹¹In environments with heterogeneity in ex post moral hazard, such as that studied in Einav *et al.* (2013), it is possible that $v(\theta; x)$ does not always exceed $C(\theta; x)$, as shown in the Example 3 in Azevedo and Gottlieb (2016). Moral hazard can also indirectly lead to a violation of Assumption 1 if consumers face budget constraints.

Remark 1 *In practice, firm can charge premiums based on observable characteristics. In this paper we simplify our analysis by assuming the observed characteristics are the same across all consumers. It is useful to think of our analysis as being within the consumers of a particular risk classification class. This simplification allows us to focus on the comparison between multidimensional private information and one-dimensional private information.*

Example 1 (Binary States) *Each consumer has initial wealth y and is subject to a possible loss $\omega \in (0, y)$ with probability m . The consumer can purchase an insurance contract to cover a fraction x of the loss if it occurs. Let $u(\cdot; \lambda)$ be consumer's utility function where λ is the risk preference parameter. Then the expected cost to the insurance firm for insuring type- (m, λ) consumer is $C(m, \lambda; x) = xc(m) = xm\omega$. Consumers' WTP for insurance of coverage x is determined by:*

$$m \cdot u(y - v - (1 - x)\omega; \lambda) + (1 - m) \cdot u(y - v; \lambda) = m \cdot u(y - \omega; \lambda) + (1 - m) \cdot u(y; \lambda).$$

The concavity of $u(\cdot; \lambda)$ implies that $v(\theta; x) > xm\omega = C(\theta; x)$. By the implicit function theorem, it can be verified that $\partial v / \partial x > 0$ and $\partial v / \partial m > 0$. Finally, fixing $x = 1$, $\partial v / \partial \lambda > 0$ holds for both Constant Relative Risk Aversion (CRRA) and Constant Absolute Risk Aversion (CARA) utility functions when we interpret λ as relative risk aversion and absolute risk aversion respectively.¹²

Example 2 (CARA and Normal Shocks) *Consumers have initial wealth y and may experience a medical expenditure $Z \sim N(m, \sigma^2)$. A consumer has CARA utility $u(y) = -\exp(-\lambda y)$, where λ is consumer's constant absolute risk aversion. Type- (m, λ) consumers' WTP for insurance of coverage x is determined by:*

$$\mathbb{E} [u(y - (1 - x)Z - v)] = \mathbb{E} [u(y - Z)].$$

Solving for v yields,

$$v(\theta; x) = xm + \frac{x(2 - x)}{2} \sigma^2 \lambda = C(\theta; x) + \frac{x(2 - x)}{2} \sigma^2 \lambda. \quad (3)$$

It can be verified that both Assumptions 1-2 are satisfied in Examples 1 and 2.¹³

4 Competitive Insurance Market

In a competitive insurance market, insurance firms choose premium p for insurance coverage with the given quality x to compete for consumers. A firm's profit at premium p from offering

¹²Simulation shows that $\partial v / \partial \lambda > 0$ holds for $x \in (0, 1]$.

¹³See Online Appendix A for the details of the proof that Example 1 satisfies Assumption 1-2. It is straightforward to see that Example 2 satisfies both assumptions.

insurance with coverage x , if there is no loading cost of offering health insurance, is given by:¹⁴

$$\pi(p) = \int_{\theta \in \mathcal{B}(p)} [p - xc(m)] dH(m, \lambda). \quad (4)$$

Denote p^* as the equilibrium price under perfect competition which, in the absence of loading costs, is simply determined by:

$$\pi(p^*) = 0. \quad (5)$$

Suppose that the market price of the insurance is p , then the average *ex post* realization of risk among those who purchase insurance is:¹⁵

$$E [c(M)|\mathcal{B}(p)] = \frac{\int_{\theta \in \mathcal{B}(p)} c(m)dH(m, \lambda)}{\int_{\theta \in \mathcal{B}(p)} dH(m, \lambda)}, \quad (6)$$

where the denominator is the measure of the insurance coverage penetration, and the numerator is the total cost realization of the insured. Similarly, the average *ex post* realization of risk among those who do not purchase insurance is

$$E [c(M)|\mathcal{NB}(p)] = \frac{\int_{\theta \in \mathcal{NB}(p)} c(m)dH(m, \lambda)}{\int_{\theta \in \mathcal{NB}(p)} dH(m, \lambda)}. \quad (7)$$

It is also useful to define the average *ex post* realization of the risk for the entire population:

$$\begin{aligned} E [c(M)] &= \int_{\underline{\lambda}}^{\bar{\lambda}} \int_{\underline{m}}^{\bar{m}} c(m)dH(m, \lambda) \\ &= \int_{\theta \in \mathcal{B}(p)} dH(m, \lambda)E [c(m)|\mathcal{B}(p)] + \int_{\theta \in \mathcal{NB}(p)} dH(m, \lambda)E [c(m)|\mathcal{NB}(p)]. \end{aligned} \quad (8)$$

Definition 1 (*Positive and Negative Correlation Property*) *The insurance market exhibits positive correlation property in equilibrium if $E [c(M)|\mathcal{B}(p^*)] > E [c(M)|\mathcal{NB}(p^*)]$, and it exhibits negative correlation property if $E [c(M)|\mathcal{B}(p^*)] < E [c(M)|\mathcal{NB}(p^*)]$, where the two terms are defined in (6) and (7) respectively.*¹⁶

Proposition 1 (*Positive Correlation Property Always Holds in Competitive Equilibrium without Loadings*) *Suppose Assumption 1 and 2 are satisfied and that the equilibrium price p^* is such that the measure of buyers and non-buyers are both strictly positive. Then positive correlation property always holds in equilibrium if the insurance market is perfectly competitive and there are no loadings.*

¹⁴We will consider how loading costs affect our results below.

¹⁵We assume that the the sets $\mathcal{B}(p)$ and $\mathcal{NB}(p)$ are of positive measures to ensure the conditional expectations are well defined.

¹⁶We will generalize the notion of positive and negative correlation property to equilibrium prices under any market structure in Section 6 and to the case of multiple contracts in Section 8.

Proof. Let m^\dagger be the solution to

$$c(m^\dagger) = \mathbb{E} [c(M)]. \quad (9)$$

In words, m^\dagger is the average risk type in the population of consumers in the market. Note that m^\dagger only depends on $c(\cdot)$ and the distribution of M in the population, and does not depend on the market equilibrium price or the distribution of risk preference type Λ .

We consider two cases depending on the level of the equilibrium premium relative to $v\left(\left(m^\dagger, \underline{\lambda}\right); x\right)$, the WTP for insurance of type- $\left(m^\dagger, \underline{\lambda}\right)$ consumer.

Case I: $p^* \leq v\left(\left(m^\dagger, \underline{\lambda}\right); x\right)$. For any $\lambda = \tilde{\lambda} \in [\underline{\lambda}, \bar{\lambda}]$, we have

$$\begin{aligned} \mathbb{E} \left[c(M) \mid \left(m, \tilde{\lambda}\right) \in \mathcal{NB}(p^*) \right] &= \mathbb{E} \left[c(M) \mid v\left(\left(m, \tilde{\lambda}\right); x\right) < p^* \right] \\ &\leq \mathbb{E} \left[c(M) \mid v\left(\left(m, \tilde{\lambda}\right); x\right) \leq v\left(\left(m^\dagger, \underline{\lambda}\right); x\right) \right] \\ &\leq \mathbb{E} \left[c(M) \mid v\left(\left(m, \tilde{\lambda}\right); x\right) \leq v\left(\left(m^\dagger, \tilde{\lambda}\right); x\right) \right] \\ &= \mathbb{E} \left[c(M) \mid \Lambda = \tilde{\lambda}, M \leq m^\dagger \right] < c(m^\dagger), \end{aligned}$$

where the first inequality follows from the assumption that $\partial v / \partial m > 0$; the second inequality follows from the assumptions that $\partial v / \partial \lambda > 0$ and $\partial v / \partial m > 0$. Therefore the average risk conditional on no insurance purchase can be bounded from above by

$$\begin{aligned} \mathbb{E} \left[c(M) \mid \mathcal{NB}(p^*) \right] &= \int_{\underline{\lambda}}^{\bar{\lambda}} \mathbb{E} \left[c(M) \mid \left(m, \tilde{\lambda}\right) \in \mathcal{NB}(p^*) \right] dG\left(\tilde{\lambda} \mid \mathcal{NB}(p^*)\right) \\ &< c(m^\dagger) = \mathbb{E} [c(M)]. \end{aligned}$$

From (6)-(8), it is clear that $\mathbb{E} [c(M) \mid \mathcal{B}(p^*)] > \mathbb{E} [c(M) \mid \mathcal{NB}(p^*)]$ follows from $\mathbb{E} [c(M) \mid \mathcal{NB}(p^*)] < \mathbb{E} [c(M)]$.

Case II: $p^* > v\left(\left(m^\dagger, \underline{\lambda}\right); x\right)$. Since p^* is determined by (5) in a competitive insurance market without loadings, we have:

$$\mathbb{E} \left[c(M) \mid \mathcal{B}(p^*) \right] = \frac{p^*}{x} > \frac{v\left(\left(m^\dagger, \underline{\lambda}\right); x\right)}{x} \geq c(m^\dagger) = \mathbb{E} [c(M)],$$

where the second inequality follows from Assumption 2. ■

Note that Proposition 1 states that negative correlation property will not emerge in a competitive insurance market without loadings, regardless of the dependence structure between risk type and risk preference type. The intuition for the result is in fact very simple. If there were a

negative correlation between insurance purchase and *ex post* risk realizations, then in a competitive insurance market, the equilibrium premium must be equal to the expected risk realization of the insured, which is lower than that of the uninsured under negative correlation property. But if the equilibrium premium were indeed lower than the expected risk realization of the uninsured, it must also be lower than their average WTP under Assumption 2. This in turn implies that at the equilibrium premium, some of the uninsured must prefer to purchase insurance as well, a contradiction.

Now we consider the role of loadings. Denote the loading factor by $\ell > 0$, and by $p^*(\ell)$ the competitive equilibrium price in a market with loading factor ℓ , which is determined by:

$$p^*(\ell) = (1 + \ell) x \mathbb{E} \left[c(M) | \mathcal{B}(p^*(\ell)) \right]. \quad (10)$$

We have the following result:¹⁷

Proposition 2 (Positive Correlation Property Holds in Competitive Equilibrium with Low Loadings) *Suppose Assumption 1 and 2 are satisfied. A sufficient condition for the positive correlation property to hold in equilibrium if the insurance market is perfectly competitive is*

$$\ell \leq \frac{v\left(m^\dagger, \underline{\lambda}\right); x}{xc(m^\dagger)} - 1, \quad (11)$$

where m^\dagger is defined in (9).

Proof. It is clear that the desired result still applies in Case I in the proof of Proposition 1. Under Case II, i.e., if $p^*(\ell) > v\left(m^\dagger, \underline{\lambda}\right); x$, then we have

$$\mathbb{E} \left[c(m) | \mathcal{B}(p^*(\ell)) \right] = \frac{p^*(\ell)}{(1 + \ell)x} > \frac{v\left(m^\dagger, \underline{\lambda}\right); x}{(1 + \ell)x} \geq c(m^\dagger) = \mathbb{E} \left[c(M) \right],$$

where the second inequality follows from the postulated restriction (11) on the loading factor ℓ . ■

Note that the upper bound on the loading factor specified by (11) depends on the ratio of the WTP for the consumer with *average risk* and lower-bound risk preference relative to its expected claim. Proposition 2 shows that if the loading factor is bounded by (11), then competitive insurance market will always exhibit positive correlation property in equilibrium even in the presence of loading factors. The intuition is again quite simple. In order for the equilibrium that low risk types

¹⁷If the loading factor is additive instead of multiplicative, i.e., if the equilibrium premium satisfies

$$p^*(\ell) = x \mathbb{E} \left[c(M) | \mathcal{B}(p^*(\ell)) \right] + \ell,$$

then the corresponding sufficient condition for Proposition 2 is

$$\ell \leq v\left(m^\dagger, \underline{\lambda}\right); x - xc(m^\dagger).$$

purchase insurance while the high types do not to exist, it must be the case that the premium is higher than the WTP for the high risk type. But the only way for the such levels of premium when the insured is actually of low risk type is that the loading factor must be very high, which is ruled out by the upper bound (11) on the loading factor.

Remark 2 *We could have redefined consumers’ “risk type” to be inclusive of the insurance loading. Such a redefinition of risk type will make Assumption 2 a more stringent assumption on all consumers’ risk premium. The sufficient condition (11) stated in Proposition 2 requires that the risk premium for the average risk type is sufficiently high.*

The following (extremely simple) example illustrates why the stated upper bound on loading factors is sufficient to rule out negative correlation property in equilibrium.

Example 3 *Suppose that there are three types of consumers in the population. For simplicity, we will describe their types by the combinations of their cost and WTP for insurance: (c_1, v_1) , (c_2, v_2) and (c_3, v_3) . Let q_j denote the probability that a consumer is of type (c_j, v_j) , $j \in \{1, 2, 3\}$, in the population. Suppose that $c_1 < c_2 < c_3$ and $c_2 = \sum_{j \in \{1, 2, 3\}} q_j c_j = E[c(M)]$ so that c_2 is exactly the average risk type m^\dagger as defined in (9).*

Now suppose that the market equilibrium price p^ is such that only type 1 consumers are purchasing insurance, i.e., that the market equilibrium exhibits negative correlation property. For this to be an equilibrium, it must be the case that the equilibrium price $p^* = (1 + \ell) c_1$ and it satisfies*

$$\max \{v_2, v_3\} < p^* \leq v_1. \tag{12}$$

However, this inequality is ruled out by the assumption (11), which for this example is reduced to

$$1 + \ell \leq \frac{v_2}{c_2}. \tag{13}$$

To see this, note that (13) implies that

$$p^* = (1 + \ell) c_1 \leq \frac{v_2}{c_2} c_1 < v_2, \tag{14}$$

which is a contradiction against (12), implying that type 2 consumers would have preferred to purchase insurance as well at p^ .*

The logic underlying Propositions 1 and 2 suggests that, in contrast to the view expressed by Chetty and Finkelstein (2013), whether multidimensional private information, particularly private information related to risk preferences, can explain the observed negative correlation between insurance purchase and *ex post* risk realization must be related to large loading factors in the insurance market, or non-competitive features of the insurance market.¹⁸

¹⁸Chetty and Finkelstein (2013) stated that “... if preferences are sufficiently important determinants of demand for insurance and sufficiently negatively correlated with risk type, the market can exhibit what has come to be called advantageous selection.” This view was shared by many others in the literature.

It is also useful to point out that, to the extent that the loading factor in the insurance market reflects the costs of administrating insurance sales and processing claims, competitive pressure in the insurance market will drive down such costs. As such, the loading factor in a perfectly competitive insurance market is likely to be low, suggesting the applicability of Proposition 2.¹⁹ The different findings in Chiappori and Salanié (2000) and Cohen (2005) we mentioned in the introduction can potentially result from the differential loading factors in the two markets. Recall that Cohen’s (2005) data is from an online Israeli insurer, while Chiappori and Salanié’s (2005) is from a traditional French insurance company. The online insurer is likely to have a much lower loading than the traditional insurer. Therefore Proposition 2 suggests that positive correlation property is more likely to hold in the Israeli data.

4.1 Relationship to CJSS (2006)

Let us now discuss in details the connections between our Propositions 1-2 and the results in CJSS (2006). Formally, using the notation in CJSS, a contract C_i reimburses the insured an amount $R_i(L)$ when loss L occurs, and they say that contract C_2 covers more than contract C_1 if $R_2(L) - R_1(L)$ is nondecreasing in L . Let

$$\pi(C_i) = p_i - \int R_i(L) dF_i(L) - \Gamma$$

denote the per contract profit from contract C_i where $F_i(L)$ is the distribution of loss among consumers purchasing contract C_i , and Γ is the fixed costs associated with the contract, and it is assumed to be the same across the contracts. The *nonincreasing profit (NIP) condition* states that $\pi(C_2) \leq \pi(C_1)$ if contract C_2 covers more than C_1 . CJSS’s main result, Proposition 2 (p. 789), states that under assumptions on consumer rationality (which we also assume) and the nonincreasing profit condition, it must be true that

$$\int R_2(L) dF_2(L) \geq \int R_2(L) dF_1(L). \tag{15}$$

In the notation of our paper, the case we considered in Proposition 1 is the case with $\Gamma = 0$, $R_2(L) = xL$ and $R_1(L) = 0$ where $L = C(M)$.²⁰ The NIP condition is automatically satisfied in our competitive market setting because of the zero profit condition. Under this interpretation, CJSS’s inequality (15) is equivalent to

$$E \left[c(M) | \mathcal{B}(p^*) \right] \geq E \left[c(M) | \mathcal{NB}(p^*) \right],$$

¹⁹However, de Meza and Webb (2001, p. 250) note that “Between 1985 and 1995 for U.K. insurers, expenses as a percentage of premium income averaged 25% for motor insurance and 37% for property damage insurance.” Note that such levels of *average* expense/premium ratio can still be consistent with the sufficient condition stipulated in Proposition 2, which is a condition imposed *only on the average risk type*.

²⁰When $\Gamma > 0$, CJSS’s Proposition 1 can not be used to show that the average *ex post* realization of risk for the insured is higher than the uninsured, as CJSS pointed out in their Footnote 5 (p. 789).

which is exactly the same as our result stated in Proposition 1.

When there is a proportional loading factor $\ell > 0$, the per contract profit from contract C_i is then

$$\pi(C_i) = p_i - (1 + \ell) \int R_i(L) dF_i(L),$$

and CJSS show that their testable implication is given by (their inequality (7) on p. 790, with tax rate $t = 0$):²¹

$$\int R_2(L) dF_2(L) - \int R_2(L) dF_1(L) \geq \ell \left[\int R_1(L) dF_1(L) - \int R_2(L) dF_2(L) \right]. \quad (16)$$

Again, if we let $R_2(L) = xL$ and $R_1(L) = 0$ where $L = C(M)$ to match the setting considered in our Proposition 2, the inequality (16) can be simplified as

$$E[c(M)|\mathcal{B}(p)] - E[c(M)|\mathcal{N}\mathcal{B}(p)] \geq -\ell E[c(M)|\mathcal{B}(p)],$$

which does not correspond to the positive correlation property even if the loading factor ℓ is sufficiently small. Indeed, CJSS comment that (p. 791), “... we can test some well-defined implication of asymmetric information (*which may not look like a positive correlation property any more*)” [italics added]. We thus consider our Proposition 2 a new result to the literature.

5 Monopolistic Insurance Market

In this section, we focus on the other extreme of the insurance market structure, assuming that there is a monopolistic insurance firm. We ask whether correlated multidimensional private information can lead to the emergence of negative correlation property (see Definition 1) in a monopolistic insurance market.

We first provide some background on how we will model dependence of the two dimensional private information M and Λ . In Section 3, we stated that, in the population, consumers’ type, (m, λ) , is independently drawn from joint CDF $H(m, \lambda)$, with marginal CDFs for M and Λ respectively denoted by $F(\cdot)$ and $G(\cdot)$. It turns out to be easier to parameterize the dependence structure of the two random variables M and Λ using the concept of *copula*.²² By Sklar’s Theorem, for every joint distribution $H(m, \lambda)$, there exists a unique copula $\mathcal{C}(\cdot, \cdot)$ such that $H(m, \lambda) = \mathcal{C}(F(m), G(\lambda))$. That is, the dependence structure between M and Λ can be represented by a copula and remains unchanged under strictly increasing transformations of the random variables.

5.1 Positive Dependence

We first consider the case of *positive dependence* between risk type M and risk preference type Λ . Although intuition suggests that positive correlation between risk and risk preference would

²¹CJSS used the notation $E_i[L]$ to denote the expected claims under contract C_i , and they wrote $E_i[L] = \int L dF_i(L)$. We believe that it is a typo and should be $E_i[L] = \int R_i(L) dF_i(L)$.

²²See Nelson (2006) for an excellent introduction to copulas.

exacerbate adverse selection and thus strengthen positive correlation between insurance coverage and *ex post* realization of risk, here we provide a precise sufficient condition for such a conclusion.

Definition 2 (*Positive Stochastic Monotonicity Dependence*) Λ is stochastically increasing in M if $\Pr(\Lambda > \lambda | M = m)$ is a nondecreasing function of m for all λ .

Nelson (2006, Corollary 5.2.11, p. 160) proved that Definition 2 is equivalent to $\mathcal{C}_{11}(z_1, z_2) \leq 0$ for all $(z_1, z_2) \in [0, 1]^2$ in the language of copula. Positive stochastic dependence means that a high realization of z_1 shifts the conditional distribution of z_2 according to first-order stochastic dominance. Because marginal distribution functions are monotonic, this property of copula translate directly into corresponding dependence property of the joint distribution of (M, Λ) .

Proposition 3 (*Positive Stochastic Monotonicity Dependence Implies Positive Correlation Property*) Suppose Assumption 1 is satisfied. If Λ is stochastically increasing in M , then positive correlation property holds under any market structure.

For the ease of exposition, we define the function $\tilde{\lambda}(m; p, x)$ so that $v\left(\left(m, \tilde{\lambda}(m; p, x)\right); x\right) = p$. In words, $\left(m, \tilde{\lambda}(m; p, x)\right)$ is the point on the iso-WTP curve valued at p in the (m, λ) space; or equivalently, $\tilde{\lambda}(m; p, x)$ is the threshold risk preference type for risk type m who is indifferent between purchasing the insurance of quality x at premium p . Accounting for the lower and upper bounds of λ , we define

$$\hat{\lambda}(m; p, x) \equiv \min\{\max\{\tilde{\lambda}(m; p, x), \underline{\lambda}\}, \bar{\lambda}\}.$$

With slight abuse of notation, we drop x in $\tilde{\lambda}(\cdot)$ and $\hat{\lambda}(\cdot)$ in what follows. Assumption 1 implies that $\hat{\lambda}(m; p)$ is nonincreasing in m for all $p \in \left(v((\underline{m}, \underline{\lambda}); x), v\left(\left(\bar{m}, \bar{\lambda}\right); x\right)\right)$.

Proof. Fix any price $p \in \left(v((\underline{m}, \underline{\lambda}); x), v\left(\left(\bar{m}, \bar{\lambda}\right); x\right)\right)$. The marginal density of M conditional on purchasing insurance is,

$$f(m|\mathcal{B}(p)) = \frac{\int_{\hat{\lambda}(m;p)}^{\bar{\lambda}} h(m, \lambda) d\lambda}{\int_{\underline{m}}^{\bar{m}} \int_{\hat{\lambda}(m;p)}^{\bar{\lambda}} h(m, \lambda) d\lambda dm}.$$

For $m'' > m'$, we must have:

$$\begin{aligned}
\frac{f(m''|\mathcal{B}(p))}{f(m'')} &= \frac{\int_{\widehat{\lambda}(m'';p)}^{\bar{\lambda}} h(m'', \lambda) d\lambda}{\int_{\underline{m}}^{\bar{m}} \int_{\widehat{\lambda}(m;p)}^{\bar{\lambda}} h(m, \lambda) d\lambda dm} \bigg/ f(m'') \\
&\geq \frac{1}{\int_{\underline{m}}^{\bar{m}} \int_{\widehat{\lambda}(m;p)}^{\bar{\lambda}} h(m, \lambda) d\lambda dm} \frac{\int_{\widehat{\lambda}(m'';p)}^{\bar{\lambda}} h(m'', \lambda) d\lambda}{f(m'')} = \frac{\Pr\left(\Lambda \geq \widehat{\lambda}(m'';p) | M = m''\right)}{\int_{\underline{m}}^{\bar{m}} \int_{\widehat{\lambda}(m;p)}^{\bar{\lambda}} h(m, \lambda) d\lambda dm} \\
&\geq \frac{\Pr\left(\Lambda \geq \widehat{\lambda}(m';p) | M = m'\right)}{\int_{\underline{m}}^{\bar{m}} \int_{\widehat{\lambda}(m;p)}^{\bar{\lambda}} h(m, \lambda) d\lambda dm} \\
&= \frac{\int_{\widehat{\lambda}(m';p)}^{\bar{\lambda}} h(m', \lambda) d\lambda}{\int_{\underline{m}}^{\bar{m}} \int_{\widehat{\lambda}(m';p)}^{\bar{\lambda}} h(m, \lambda) d\lambda dm} \bigg/ f(m') = \frac{f(m'|\mathcal{B}(p))}{f(m')},
\end{aligned}$$

where the first inequality follows from the fact that $\widehat{\lambda}(m;p)$ is nonincreasing in m for all p ; and the second inequality follows from the definition of positive stochastic monotonicity dependence. Hence $\frac{f(m|\mathcal{B}(p))}{f(m)}$ satisfies monotone likelihood ratio property, which implies that $F(m|\mathcal{B}(p))$ first-order stochastically dominates $F(m)$. Moreover, Assumption 1 implies that the first-order stochastic dominance is strict for at least a positive measure of values of m . Therefore $E[c(M)|\mathcal{B}(p)] > E[c(M)]$, which is equivalent to the positive correlation property as stated in Definition 1. ■

Proposition 3 is general in the sense that it does not rely on the functional form of consumer's WTP. As long as consumers' WTP is increasing in both m and λ , Proposition 3 holds.

5.2 Negative Dependence

Now we consider the case in which the risk type M and the risk preference type Λ exhibit *negative* dependence (to be made precise below) and investigate whether negative dependence between M and Λ may lead to the emergence of negative correlation property under a monopolistic market structure.

In this section, we consider the tractable case of CARA utility function and normally distributed shocks as described in Example 2. Note that for this CARA-normal specification, it is without loss of generality to assume that $x = 1$;²³ thus from (3), we have $C(\theta) = m$, and

$$v(m, \lambda) = m + k\lambda \text{ where } k \equiv \frac{\sigma^2}{2}. \quad (17)$$

We will interpret the parameter $k \equiv \sigma^2/2$, where σ^2 is the variance of the health expenditure shock, as the *relative importance of risk aversion* as a determinant of the consumer's WTP for insurance: a higher σ^2 means that consumers are subject to more volatility in health expenditure, and as a

²³For $x \in (0, 1)$, we can redefine $\hat{v}((m, \lambda); 1) \equiv v((m, \lambda); x)/x$ and $\hat{C}(\theta; 1) \equiv C(\theta; x)/x$.

result risk aversion becomes more important in determining the WTP for insurance.

5.2.1 The Role of Preferences

We study the effect of the relative importance between risk and preferences, i.e., the magnitude of k defined in (17), holding fixed the joint distribution $H(\cdot, \cdot)$.

Proposition 4 *Suppose that consumers have CARA utility functions and experience normally distributed risks as described in Example 2. For every $H(\cdot, \cdot)$, there exists a threshold $k^\dagger > 0$ such that for all $k < k^\dagger$, $E[M|\mathcal{B}(p)] > E[M]$ for all $p \in (\underline{m} + k\underline{\lambda}, \bar{m} + k\bar{\lambda})$.*

Proof. For notational convenience, denote the range of the risk preference type Λ by $R_\Lambda \equiv \bar{\lambda} - \underline{\lambda}$.

When $p \in (\underline{m} + k\underline{\lambda}, E[M] + k\underline{\lambda})$, if a consumer's risk type m is such that $m \geq E[M]$, then her WTP $v(m, \lambda) \geq m + k\underline{\lambda} \geq E[M] + k\underline{\lambda} > p$, thus such a consumer will for sure purchase insurance at price p . Thus, all the consumers who do not purchase insurance must have risk type $m < E[M]$, which implies that $E[M|\mathcal{N}\mathcal{B}(p)] < E[M]$, or equivalently $E[M|\mathcal{B}(p)] > E[M|\mathcal{N}\mathcal{B}(p)]$.

When $p \in (E[M] + k\bar{\lambda}, \bar{m} + k\bar{\lambda})$, if a consumer's risk type m is such that $m \leq E[M]$, then her WTP for insurance $v(m, \lambda) = m + k\lambda \leq E[M] + k\bar{\lambda} < p$ for all $\lambda \in [\underline{\lambda}, \bar{\lambda}]$. Thus, the risk type of consumers who buy insurance at price p must be higher than $E[M]$. This implies that $E[M|\mathcal{B}(p)] > E[M]$, or equivalently, $E[M|\mathcal{B}(p)] > E[M|\mathcal{N}\mathcal{B}(p)]$.

Now we consider the case when $p \in [E[M] + k\underline{\lambda}, E[M] + k\bar{\lambda}]$. Suppose $k < k_1$ where

$$k_1 \equiv \min \left\{ (\bar{m} - E[M]) / R_\Lambda, (E[M] - \underline{m}) / R_\Lambda \right\}.$$

Then the iso-WTP curve of the marginal types of consumers, namely those consumers with $v(m, \lambda) = m + k\lambda = p$, will intersect the upper and lower bounds of the range of Λ (see Figure 1). For $p \in [E[M] + k\underline{\lambda}, E[M] + k\bar{\lambda}]$, those and only those consumers to the right of the iso-WTP line will purchase insurance. Thus, we have:

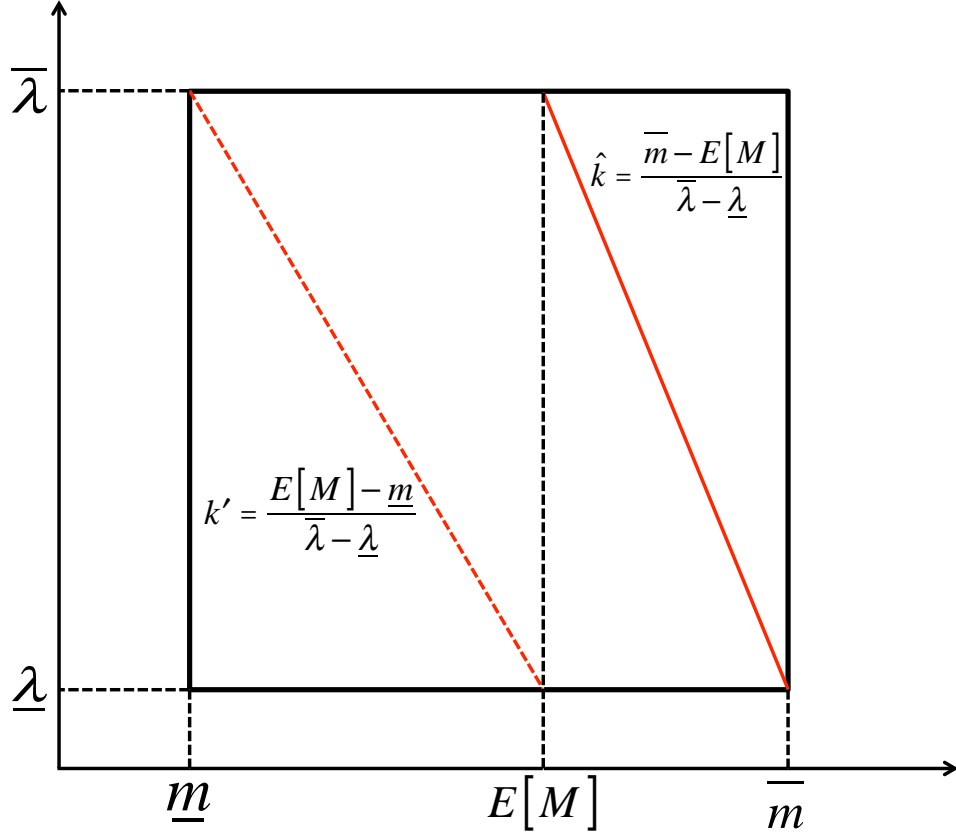


Figure 1: Proof of Proposition 4 when $p \in [E[M] + k\underline{\lambda}, E[M] + k\bar{\lambda}]$ and $k < k_1 \equiv \min\{k', \hat{k}\}$.

$$\begin{aligned}
E[M|\mathcal{NB}(p)] &= \frac{\int_{\underline{m}}^{p-k\underline{\lambda}} \int_{\underline{\lambda}}^{\min\{\frac{p-m}{k}, \bar{\lambda}\}} mh(m, \lambda) d\lambda dm}{\int_{\underline{m}}^{p-k\underline{\lambda}} \int_{\underline{\lambda}}^{\min\{\frac{p-m}{k}, \bar{\lambda}\}} h(m, \lambda) d\lambda dm} \\
&\leq \frac{\int_{\underline{m}}^{p-k\underline{\lambda}} \int_{\underline{\lambda}}^{\min\{\frac{p-m}{k}, \bar{\lambda}\}} mh(m, \lambda) d\lambda dm + \int_{p-k\underline{\lambda}}^{p-k\bar{\lambda}} \int_{\frac{p-m}{k}}^{\bar{\lambda}} (p - k\underline{\lambda}) h(m, \lambda) d\lambda dm}{\int_{\underline{m}}^{p-k\underline{\lambda}} \int_{\underline{\lambda}}^{\min\{\frac{p-m}{k}, \bar{\lambda}\}} h(m, \lambda) d\lambda dm + \int_{p-k\underline{\lambda}}^{p-k\bar{\lambda}} \int_{\frac{p-m}{k}}^{\bar{\lambda}} h(m, \lambda) d\lambda dm} \\
&= \frac{\int_{\underline{m}}^{p-k\underline{\lambda}} \int_{\underline{\lambda}}^{\bar{\lambda}} mh(m, \lambda) d\lambda dm}{\int_{\underline{m}}^{p-k\underline{\lambda}} \int_{\underline{\lambda}}^{\bar{\lambda}} h(m, \lambda) d\lambda dm} + \frac{\int_{p-k\underline{\lambda}}^{p-k\bar{\lambda}} \int_{\frac{p-m}{k}}^{\bar{\lambda}} (p - k\underline{\lambda} - m) h(m, \lambda) d\lambda dm}{\int_{\underline{m}}^{p-k\underline{\lambda}} \int_{\underline{\lambda}}^{\bar{\lambda}} h(m, \lambda) d\lambda dm} \\
&= \frac{\int_{\underline{m}}^{p-k\underline{\lambda}} mf(m) dm}{\int_{\underline{m}}^{p-k\underline{\lambda}} f(m) dm} + \frac{\int_{p-k\underline{\lambda}}^{p-k\bar{\lambda}} \int_{\frac{p-m}{k}}^{\bar{\lambda}} (p - k\underline{\lambda} - m) h(m, \lambda) d\lambda dm}{\int_{\underline{m}}^{p-k\underline{\lambda}} \int_{\underline{\lambda}}^{\bar{\lambda}} h(m, \lambda) d\lambda dm} \\
&< E[M|M \leq p - k\underline{\lambda}] + kR_{\Lambda} \\
&\leq E[M|M \leq E[M] + kR_{\Lambda}] + kR_{\Lambda},
\end{aligned}$$

where the first inequality follows from the fact that we are adding the set of consumers whose risk type m is such that $m \in [p - k\bar{\lambda}, p - k\underline{\lambda}]$ and whose WTP is above the premium, while assuming that their risk types were all $p - k\underline{\lambda}$; and the second inequality follows from the fact that in the integrand of the numerator in the second term, $p - k\underline{\lambda} - m \leq kR_\Lambda$ because $m \geq p - k\bar{\lambda}$.

Now consider $\eta(k) \equiv \mathbb{E}[M|M \leq \mathbb{E}[M] + kR_\Lambda] + kR_\Lambda$. It is clear that $\eta(k)$ is increasing in k ; moreover,

$$\begin{aligned}\eta(0) &= \mathbb{E}[M|M \leq \mathbb{E}[M]] < \mathbb{E}[M]; \\ \eta\left(\frac{\bar{m} - \mathbb{E}[M]}{R_\Lambda}\right) &= \bar{m} > \mathbb{E}[M].\end{aligned}$$

Let $k_2 \in \left(0, \frac{\bar{m} - \mathbb{E}[M]}{R_\Lambda}\right)$ be the unique solution to the equation $\eta(k) = \mathbb{E}[M]$. Then $\eta(k) < \mathbb{E}[M]$ for all $k < k^\dagger \equiv \min\{k_1, k_2\}$. This completes the proof. ■

Proposition 4 shows that if risk preference is not a sufficiently important determinant of the demand for insurance, then negative correlation property will not emerge under monopolistic market structure, regardless of the joint distribution, $H(\cdot, \cdot)$, of risk and risk preferences. Notice that this holds true even when the risk type M and the risk preference type Λ exhibit strong negative dependence. Figure 2 illustrates why this is so for the extreme case when M and Λ exhibit perfect negative dependence. Since M and Λ are perfectly negative dependent, there exists a one-to-one monotonic mapping between m and λ , which is shown in the dashed line in Figure 2. A sufficiently small k yields a steep iso-WTP curve, which would imply that for any price in $(\underline{m} + k\underline{\lambda}, \bar{m} + k\bar{\lambda})$, the iso-WTP curve that separates purchasers and no-purchasers of insurance at that price would intersect the dashed line as depicted: the higher risks (the darker segment of the dashed line) always purchase insurance before the lower risks (the lighter segment of the dashed line). This results in positive correlation property.

Intuitively, when k is sufficiently small, a change in the price charged by the monopolist will have a stronger influence on the support of risk type M rather than that of Λ ; as a result, the market is more susceptible to the risk-based adverse selection problem as in the case of one-dimensional private information in risk, and the potential countervailing effect of selection based on risk preferences is too weak to override the positive correlation property.

Now we consider the other limiting case, and show that if risk preference is a sufficiently important determinant factor for the demand of insurance, then when risk M and risk preference Λ are negatively dependent (to be defined more precisely below), a monopolistic market may exhibit negative correlation property in equilibrium. To this end, we first introduce the notion of *negative quadrant dependence*:

Definition 3 (Strict Negative Quadrant Dependence) *M and Λ are strictly negatively quadrant dependent if for all $(m, \lambda) \in [\underline{m}, \bar{m}] \times [\underline{\lambda}, \bar{\lambda}]$,*

$$H(m, \lambda) < F(m)G(\lambda).$$

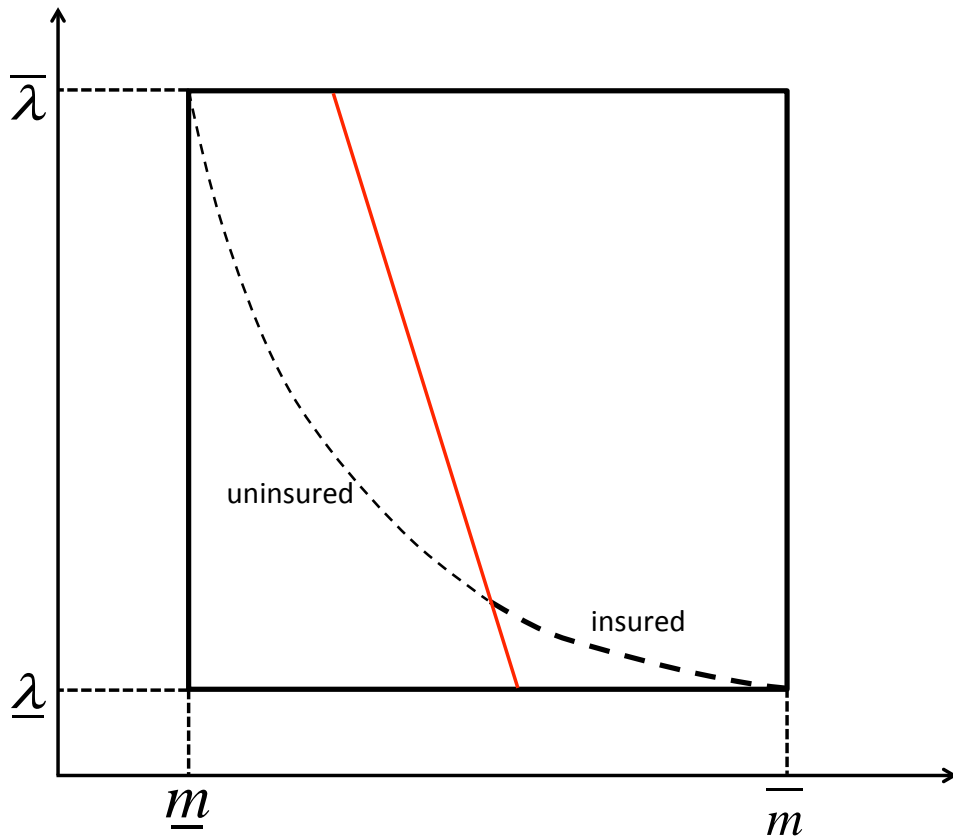


Figure 2: Selection Based on Risk Type: The Case of Low k

Strict negative quadrant dependence formalizes the notion that two random variables are negatively dependent if greater values of M are more likely to appear with smaller values of Λ and vice versa. In the language of copula, it is equivalent to $\mathcal{C}(z_1, z_2) < z_1 z_2$.²⁴ Notice that independent random variables does not satisfy strict negative quadrant dependence.²⁵

Proposition 5 *Suppose that consumers have CARA utility functions and experience normally distributed risks as described in Example 2. If M and Λ are strictly negatively quadrant dependent,*

²⁴Many families of copulas that include $\mathcal{C}(z_1, z_2) = z_1 z_2$ have subfamilies of positively quadrant dependent and negatively quadrant dependent copulas. For instance, the Mardia family, the FGM family, the Ali-Mikhail-Haq family and the Frank family (see Nelson 2006).

²⁵It is useful to point out that negative quadrant dependence is a weaker notion of negative dependence than *negative stochastic monotonicity dependence* (also known as negative regression dependence, see Nelson 2006), which is the opposite of the positive stochastic monotonicity dependence we introduced in Definition 2. Formally, we say that two random variables M and Λ have negative stochastic monotonicity dependence if one is stochastically decreasing in the other; for example, Λ is stochastically decreasing in M if $\Pr(\Lambda > \lambda | M = m)$ is a nonincreasing function of m for all λ .

Negative stochastic monotonicity dependence is in turn implied by and thus weaker than *negative likelihood ratio dependence* (see Bromek and Pleszczyńska (1991, p.115)). We say that M and Λ are negatively likelihood ratio dependent if

$$h(m, \lambda)h(m', \lambda') \leq h(m', \lambda)h(m, \lambda')$$

for all $m \leq m'$ and $\lambda \leq \lambda'$. Likelihood ratio dependence is equivalent to affiliation.

then there exists a threshold $k^{\dagger\dagger}$ such that negative correlation property emerges under monopoly when $k > k^{\dagger\dagger}$.

We first pin down the monopoly price in the limit. Note that the monopoly firm will choose price p^m to:

$$\max_{\{p\}} \pi(p; k) \equiv \int_{\theta \in \mathcal{B}(p) \equiv \{(m, \lambda) : m + k\lambda \geq p\}} (p - m) dH(m, \lambda). \quad (18)$$

Since we are interested in how the parameter k affects the emergence of the negative correlation property in equilibrium, we use $p^m(k)$ to denote the monopolist's profit maximizing premium when k is the parameter measuring the importance of risk preference as a determinant of insurance demand. For what follows, it is useful to define

$$\lambda^* \equiv \arg \max \lambda [1 - G(\lambda)], \quad (19)$$

where $G(\cdot)$ is the marginal CDF of Λ . We assume that λ^* is unique and $\lambda^* \in (\underline{\lambda}, \bar{\lambda})$.

Lemma 1 $\lim_{k \rightarrow \infty} \frac{p^m(k)}{k} = \lambda^*$.

Proof. It is equivalent to prove that for any arbitrarily small $\epsilon > 0$, there exists a threshold k_3 such that $\frac{p^m(k)}{k} \in (\lambda^* - \epsilon, \lambda^* + \epsilon)$ for $k > k_3$. Without loss of generality, we assume $\epsilon < \min\{\bar{\lambda} - \lambda^*, \lambda^* - \underline{\lambda}\}$.

First, notice that:

$$\begin{aligned} \left(\lambda^* - \frac{\bar{m}}{k} \right) \left[1 - G \left(\lambda^* - \frac{m}{k} \right) \right] &\leq \frac{1}{k} \pi(k\lambda^*; k) = \int_{\lambda^* \leq \frac{m}{k} + \lambda} \left(\lambda^* - \frac{m}{k} \right) dH(m, \lambda) \\ &\leq \lambda^* \left[1 - G \left(\lambda^* - \frac{\bar{m}}{k} \right) \right]. \end{aligned}$$

Taking limit of the above inequality yields:

$$\lim_{k \rightarrow \infty} \frac{1}{k} \pi(k\lambda^*; k) = \lambda^* [1 - G(\lambda^*)].$$

Second, when $k > m/(\lambda^* - \underline{\lambda} - \epsilon)$, $\lambda^* - \epsilon > \frac{m}{k} + \underline{\lambda}$ and $\frac{\bar{m}}{k} + \bar{\lambda} > \lambda^* + \epsilon$ hold. For $\frac{p}{k}$ that is outside the neighborhood of λ^* , i.e., $\frac{p}{k} \in [\frac{m}{k} + \underline{\lambda}, \lambda^* - \epsilon] \cup [\lambda^* + \epsilon, \frac{\bar{m}}{k} + \bar{\lambda}]$, we have

$$\frac{1}{k} \pi(p; k) = \int_{\frac{p}{k} \leq \frac{m}{k} + \lambda} \left(\frac{p}{k} - \frac{m}{k} \right) dH(m, \lambda) \leq \frac{p}{k} \left[1 - G \left(\frac{p}{k} - \frac{\bar{m}}{k} \right) \right].$$

Now let $\delta \equiv \lambda^* [1 - G(\lambda^*)] - \max_{[\frac{m}{k} + \underline{\lambda}, \lambda^* - \epsilon] \cup [\lambda^* + \epsilon, \frac{\bar{m}}{k} + \bar{\lambda}]} \lambda [1 - G(\lambda)] > 0$. By continuity, there exists a threshold \check{k} such that $\frac{p}{k} [1 - G(\frac{p}{k} - \frac{\bar{m}}{k})] \leq \lambda^* [1 - G(\lambda^*)] - \frac{\delta}{2}$ for $k > \check{k}$. Therefore, for $k > k_3 \equiv \max\{m/(\lambda^* - \underline{\lambda} - \epsilon), \check{k}\}$, the profit at price $p = k\lambda^*$ is greater than the profit at any price p such that $\frac{p}{k} \in [\frac{m}{k} + \underline{\lambda}, \lambda^* - \epsilon] \cup [\lambda^* + \epsilon, \frac{\bar{m}}{k} + \bar{\lambda}]$. This completes the proof. ■

Notice that Lemma 1 does not depend on the assumption that M and Λ exhibit negative quadrant dependence.

Proof of Proposition 5. The average risk of the uninsured is given by:

$$\begin{aligned}
\mathbb{E} \left[M | \mathcal{NB} (p^m(k)) \right] &= \frac{\int_{\frac{p^m(k)}{k} \geq \frac{\underline{m}}{k} + \lambda} m dH(m, \lambda)}{\int_{\frac{p^m(k)}{k} \geq \frac{\underline{m}}{k} + \lambda} dH(m, \lambda)} \\
&> \frac{\int_{\underline{m}}^{\overline{m}} \int_{\underline{\lambda}}^{\frac{p^m(k)-m}{k}} mh(m, \lambda) d\lambda dm + \int_{\underline{m}}^{\overline{m}} \int_{\frac{p^m(k)-\underline{m}}{k}}^{\frac{p^m(k)-m}{k}} \underline{m} h(m, \lambda) d\lambda dm}{\int_{\underline{m}}^{\overline{m}} \int_{\underline{\lambda}}^{\frac{p^m(k)-m}{k}} h(m, \lambda) d\lambda dm + \int_{\underline{m}}^{\overline{m}} \int_{\frac{p^m(k)-\underline{m}}{k}}^{\frac{p^m(k)-m}{k}} h(m, \lambda) d\lambda dm} \\
&> \frac{\int_{\underline{\lambda}}^{\frac{p^m(k)-\underline{m}}{k}} mg(\lambda) d\lambda}{G\left(\frac{p^m(k)-\underline{m}}{k}\right)} - (\overline{m} - \underline{m}) \left[1 - \frac{G\left(\frac{p^m(k)-\overline{m}}{k}\right)}{G\left(\frac{p^m(k)-\underline{m}}{k}\right)} \right] \\
&= \mathbb{E} \left[M | \lambda \leq \frac{p^m(k) - \underline{m}}{k} \right] - (\overline{m} - \underline{m}) \left[1 - \frac{G\left(\frac{p^m(k)-\overline{m}}{k}\right)}{G\left(\frac{p^m(k)-\underline{m}}{k}\right)} \right],
\end{aligned}$$

where the first inequality follows from the fact that we are adding the set of consumers whose risk aversion type λ is such that $\lambda \in \left[\frac{p^m(k) - \overline{m}}{k}, \frac{p^m(k) - \underline{m}}{k} \right]$ and whose WTP is above the premium, while assuming that their risk types were all \underline{m} ; and the second inequality follows from the fact that $m - \underline{m} < \overline{m} - \underline{m}$ and $\int_{\underline{m}}^{\overline{m}} \int_{\frac{p^m(k)-\underline{m}}{k}}^{\frac{p^m(k)-m}{k}} h(m, \lambda) d\lambda dm < G\left(\frac{p^m(k)-\underline{m}}{k}\right) - G\left(\frac{p^m(k)-\overline{m}}{k}\right)$.

By Lemma 1, $\lim_{k \rightarrow \infty} \frac{p^m(k) - \underline{m}}{k} = \lim_{k \rightarrow \infty} \frac{p^m(k) - \overline{m}}{k} = \lambda^*$. Hence,

$$\lim_{k \rightarrow \infty} \left\{ \mathbb{E} \left[M | \Lambda \leq \frac{p^m(k) - \underline{m}}{k} \right] - (\overline{m} - \underline{m}) \left[1 - \frac{G\left(\frac{p^m(k)-\overline{m}}{k}\right)}{G\left(\frac{p^m(k)-\underline{m}}{k}\right)} \right] \right\} = \mathbb{E} [M | \Lambda \leq \lambda^*].$$

From Definition 3, we must have

$$\Pr(M \leq m | \Lambda \leq \lambda) < \Pr(M \leq m) \text{ for all } (m, \lambda),$$

which implies that $\mathbb{E} [M | \Lambda \leq \lambda^*] > \mathbb{E} [M]$. Hence, there exists a threshold $k^{\dagger\dagger}$ such that $\mathbb{E} [M | \mathcal{NB} (p^m(k))] > \mathbb{E} [M]$ for $k > k^{\dagger\dagger}$. ■

The intuition for Proposition 5 is as follows. When risk aversion is a sufficiently important determinant of the demand for insurance, the iso-WTP curve is sufficiently flat in the (m, λ) space. Thus, a change in the price by the monopolistic firm will have a stronger impact on the distribution of the risk aversion type than the distribution of risk types among the set of the purchasers; that is, the selection of the consumers are more based on risk aversion type λ than on risk type m . In the limit when k is sufficiently large, it is profit maximizing for the monopolist to price in a way to select *only* consumers whose risk aversion is above λ^* as defined by (19). Because higher risk aversion is associated with lower risk type by the assumption of quadrant negative dependence,

consumers who purchase insurance have lower average risk than the entire population.

Proposition 5 represents a striking difference from the result reported in Proposition 1 for the case of perfectly competitive market, where we show that negative correlation property will not emerge under any joint distribution of M and Λ . To better explain the intuition why market structure plays such an important role in whether or not negative dependence between M and Λ can lead to negative correlation property in equilibrium, it is useful to further examine the difference, when k gets large, between the competitive equilibrium price p^* and the monopolistic price $p^m(k)$. In Lemma 1 we showed that under monopoly, $p^m(k)/k$ converges to λ^* when k gets large. The following claim shows that $p^*(k)/k$ – we use $p^*(k)$ to indicate the competitive equilibrium price when the relative importance parameter of the risk preference in WTP is k – will be converging to a smaller value than λ^* .

Claim 1 *Suppose that consumers have CARA utility functions and experience normally distributed risks as described in Example 2. Under perfect competition, there exists a threshold \tilde{k}^* such that $p^*(k) \leq E[M] + k\underline{\lambda}$ for $k > \tilde{k}^*$.*

Proof. Let $\tilde{k}^* = \frac{\bar{m} - E[M]}{\underline{\lambda}}$. If $p^*(k) > E[M] + k\underline{\lambda}$, firm's expected profit is

$$\begin{aligned} \pi(p^*(k); k) &= \int_{\theta \in \mathcal{B}(p^*(k))} (p - m) dH(m, \lambda) \\ &> \int_{\theta \in \mathcal{B}(p^*(k))} [E[M] + k\underline{\lambda} - m] dH(m, \lambda) \\ &> \int_{\theta \in \mathcal{B}(p^*(k))} [E[M] + k\underline{\lambda} - \bar{m}] dH(m, \lambda) \geq 0, \end{aligned}$$

which is a contradiction to the zero profit condition required for competitive equilibrium. ■

Claim 1 implies that as k gets sufficiently large, $\frac{p^*(k)}{k}$ will converge to something smaller than $\underline{\lambda}$; in contrast, $\frac{p^m(k)}{k}$ converges to $\lambda^* > \underline{\lambda}$. The difference between competitive and monopolistic market structure can be understood as follows. As k gets larger, it becomes less costly for the insurance firms to offer insurance due to the increase in the WTP for insurance for consumers of all risk types. This is true for both the competitive and the monopolistic market structure. However, in a perfectly competitive market, the competitive pressure will force insurance companies to reduce prices and cover more consumers, leading to a low price level in equilibrium. Claim 1 establishes that the competitive equilibrium price is lower than $E[M] + k\underline{\lambda}$ when k is sufficiently large. As a result, only consumers with low risk (lower than the unconditional expectation) and low degree of risk aversion choose to opt out (the lower dashed line in Figure 3). This implies directly that consumers with no insurance have lower average risk than the entire population.²⁶ In contrast, a monopolistic firm recognizes that, when M and Λ exhibit strong negative dependence, most densities concentrate on the diagonal as indicated in Figure 3. To maximize profit, a monopolist will choose a higher price

²⁶In the extreme case of perfect negative dependence, every consumer will buy insurance for any $k > 0$, hence $E[M|\mathcal{B}(p^*(k))] = E[M]$.

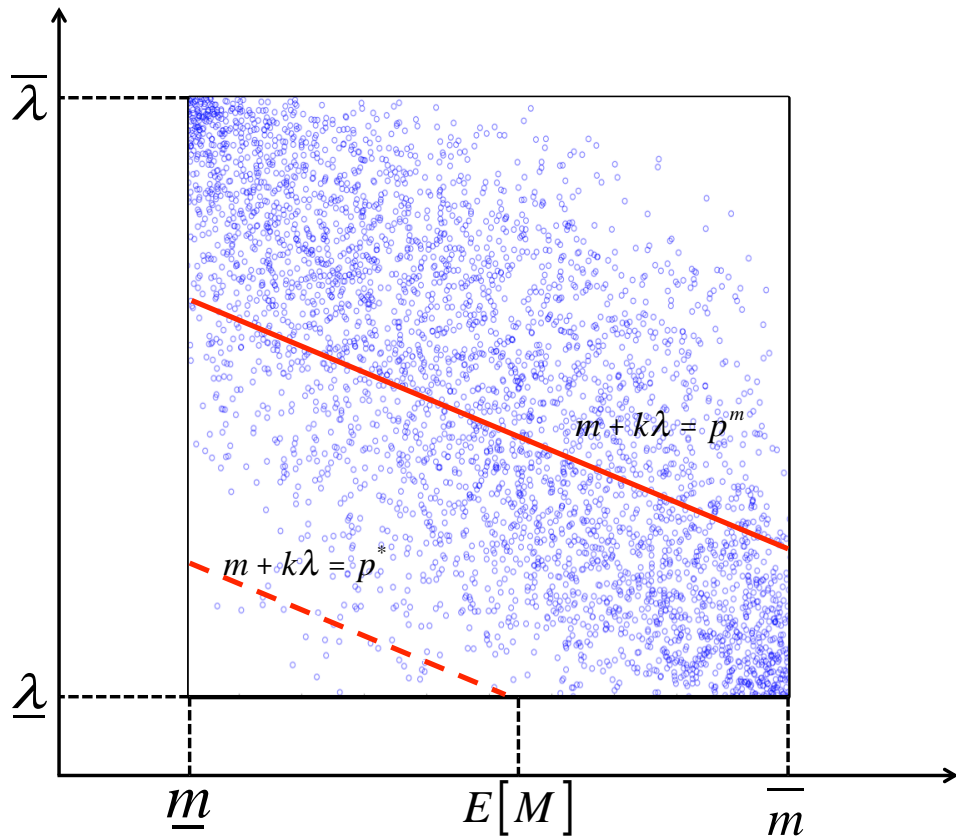


Figure 3: Perfect Competition v.s. Monopoly: the Case of High k

relative to the competitive equilibrium price p^* so as to exclude the higher risk consumers (the upper solid line in Figure 3).

5.2.2 Comparative Statics with Respect to the Degree of Negative Dependence

Proposition 5 is a *limiting result* under monopoly as the role of preference as a determinant of the demand for insurance becomes sufficiently important. In this subsection, we further parameterize the nature of the negative dependence between M and Λ , and examine its impact on the equilibrium outcomes of the monopolistic market, including premium and market size and the correlation between insurance purchase and *ex post* realization of risk.

To this end, we assume that both M and Λ are uniformly distributed between 0 and 1, and the joint distribution of M and Λ can be represented by a Bivariate Fréchet copula parameterized by $\mu \in [0, 1]$:

$$\mathcal{C}(z_1, z_2; \mu) = \mu \mathcal{W}(z_1, z_2) + (1 - \mu) \Pi(z_1, z_2), \quad (20)$$

where $\Pi(z_1, z_2) \equiv z_1 z_2$ is the product copula that exhibits *independence* and $\mathcal{W}(z_1, z_2) \equiv \max\{z_1 + z_2 - 1, 0\}$ is the Fréchet lower bound that shows *perfect negative dependence* between u and v .

Hence, the parameter μ measures the *degree of negative dependence* between z_1 and z_2 . With a slight abuse of notation, we use $p^m(\mu)$ to denote the monopolist's profit maximizing premium, and use $D(p^m(\mu))$ to denote the measure of consumers who will purchase insurance at the monopolist's price $p^m(\mu)$. We know from Proposition 1 that positive correlation property emerges for all k and μ if the market structure is perfect competition. For the case of monopolistic market structure, we have the following result:

Proposition 6 *Suppose that consumers have CARA utility functions and experience normally distributed risks as described in Example 2, and the joint distribution is $H(m, \lambda; \mu) = \mu\mathcal{W}(m, \lambda) + (1 - \mu)\Pi(m, \lambda)$. Assume $M \sim U[0, 1]$ and $\Lambda \sim U[0, 1]$. Suppose the market structure is monopoly. Then,*

1. if $k < 1$, positive correlation property emerges for all μ ;
2. if $k > 1$, there exists a threshold μ^\dagger such that:
 - (a) negative correlation property emerges if $\mu > \mu^\dagger$; and positive correlation property emerges if $\mu < \mu^\dagger$;
 - (b) $dp^m(\mu)/d\mu < 0$ and $dD(p^m(\mu))/d\mu > 0$.

Proof. (1). Suppose that $k < 1$. First consider $\Pi(m, \lambda)$, which is the joint distribution when M and Λ are independent, which trivially satisfies the definition of positive stochastic monotonicity dependence (see Definition 2). Thus Proposition 3 applies; and thus, for all $p \in (0, k + 1)$, we must have:²⁷

$$\begin{aligned} & \frac{\int_{\theta \in \mathcal{B}(p)} m d\Pi(m, \lambda)}{\int_{\theta \in \mathcal{B}(p)} d\Pi(m, \lambda)} > \mathbb{E}[M] \\ \Leftrightarrow & \int_{\theta \in \mathcal{B}(p)} m d\Pi(m, \lambda) > \mathbb{E}[M] \int_{\theta \in \mathcal{B}(p)} d\Pi(m, \lambda). \end{aligned} \quad (21)$$

Now consider $\mathcal{W}(m, \lambda)$. For $p \in (0, k]$, we have,

$$\frac{\int_{\theta \in \mathcal{B}(p)} m d\mathcal{W}(m, \lambda)}{\int_{\theta \in \mathcal{B}(p)} d\mathcal{W}(m, \lambda)} = \frac{\int_{\max\{\frac{p-k}{1-k}, 0\}}^1 m dm}{\int_{\max\{\frac{p-k}{1-k}, 0\}}^1 dm} = \frac{1}{2} \left[\max\left\{\frac{p-k}{1-k}, 0\right\} + 1 \right] \geq \frac{1}{2} = \mathbb{E}[M].$$

For $p \in [k, k + 1)$, we have $\int_{\theta \in \mathcal{B}(p)} m d\mathcal{W}(m, \lambda) = \int_{\theta \in \mathcal{B}(p)} d\mathcal{W}(m, \lambda) = 0$. Combining these, we conclude that for all $p \in (0, k + 1)$:

$$\int_{\theta \in \mathcal{B}(p)} m d\mathcal{W}(m, \lambda) \geq \mathbb{E}[M] \int_{\theta \in \mathcal{B}(p)} d\mathcal{W}(m, \lambda). \quad (22)$$

²⁷Note under the assumed support of M and Λ , the maximum WTP for insurance $\bar{m} + k\bar{\lambda}$ is $k + 1$.

Now since $H(m, \lambda; \mu) = \mu\mathcal{W}(m, \lambda) + (1 - \mu)\Pi(m, \lambda)$, the average risk of the insured can be bounded from below by

$$\begin{aligned} \mathbb{E} [c(M)|\mathcal{B}(p)] &= \frac{\int_{\theta \in \mathcal{B}(p)} c(m) dH(m, \lambda)}{\int_{\theta \in \mathcal{B}(p)} dH(m, \lambda)} \\ &= \frac{(1 - \mu) \int_{\theta \in \mathcal{B}(p)} m d\Pi(m, \lambda) + \mu \int_{\theta \in \mathcal{B}(p)} m d\mathcal{W}(m, \lambda)}{(1 - \mu) \int_{\theta \in \mathcal{B}(p)} d\Pi(m, \lambda) + \mu \int_{\theta \in \mathcal{B}(p)} d\mathcal{W}(m, \lambda)} \\ &> \frac{(1 - \mu) \mathbb{E} [M] \int_{\theta \in \mathcal{B}(p^m)} d\Pi(m, \lambda) + \mu \mathbb{E} [M] \int_{\theta \in \mathcal{B}(p^m)} d\mathcal{W}(m, \lambda)}{(1 - \mu) \int_{\theta \in \mathcal{B}(p^m)} d\Pi(m, \lambda) + \mu \int_{\theta \in \mathcal{B}(p^m)} d\mathcal{W}(m, \lambda)} = \mathbb{E} [M], \end{aligned}$$

where the second equality follows from the definition of $H(\cdot, \cdot)$, and the inequality follows from (21) and (22).

(2). Suppose $k > 1$. The profit function can be derived as:²⁸

$$\pi(p; \mu) = \begin{cases} (p - \frac{1}{2}) - (1 - \mu) \frac{1}{3k} p^3 & \text{for } p \in [0, 1] \\ (1 - \mu) \left[(p - \frac{1}{2}) + \frac{-3p^2 + 3p - 1}{3k} \right] + \frac{1}{2} \mu \frac{k-p}{k-1} \left(2p - \frac{k-p}{k-1} \right) & \text{for } p \in [1, k] \\ (1 - \mu) \left[\frac{(p-1)^3}{3k} - \frac{(p-1)^2}{2} + \frac{k^2}{6} \right] & \text{for } p \in [k, k+1] \end{cases}. \quad (23)$$

Notice that $\partial\pi/\partial p = 1 - (1 - \mu)p^2/k > 0$ for $p \in [0, 1]$ and $\partial\pi/\partial p = (1 - \mu)(p-1)[(p-1)/k - 1] < 0$ for $p \in [k, k+1]$. Moreover, $\partial^2\pi/\partial p^2 < 0$ for $p \in [1, k]$. Therefore, $p^m(\mu)$ is the solution to $\partial\pi/\partial p = 0$ for $p \in [1, k]$. Solving for p^m yields,

$$p^m(\mu) = \frac{\mu \frac{k^2}{k-1} + (1 - \mu)(k - \frac{1}{k})}{\mu(2 + \frac{1}{k-1}) + 2(1 - \mu)(1 - \frac{1}{k})}.$$

The market size is,

$$D(p^m(\mu)) = \mu \left(\frac{k - p^m(\mu)}{k - 1} \right) + (1 - \mu) \left\{ 1 + \frac{1}{k} \left[\frac{1}{2} - p^m(\mu) \right] \right\}.$$

It can be verified that $dp^m(\mu)/d\mu < 0$ and $dD(p^m(\mu))/d\mu > 0$. Finally, the average risk of the insured can be derived as,

$$\mathbb{E} [c(M)|\mathcal{B}(p^m(\mu))] = \frac{\frac{1}{2} \mu \left[\frac{k - p^m(\mu)}{k-1} \right]^2 + (1 - \mu) \left\{ \frac{1}{2} + \frac{1}{k} \left[\frac{1}{3} - \frac{1}{2} p^m(\mu) \right] \right\}}{\mu \left[\frac{k - p^m(\mu)}{k-1} \right] + (1 - \mu) \left\{ 1 + \frac{1}{k} \left[\frac{1}{2} - p^m(\mu) \right] \right\}}.$$

²⁸See Online Appendix B for details of the derivation of the profit function, the demand curve, and the cost curves.

Carrying out the algebra, it can be verified that $E [c(M)|\mathcal{B} (p^m (\mu))] < E [M]$ is equivalent to

$$[k - p^m(\mu)][p^m(\mu) - 1] - \frac{1}{6} \frac{1 - \mu}{\mu} \frac{(k - 1)^2}{k} > 0.$$

Denote the left hand side of the inequality as $\chi(\mu)$. Then we have,

$$\frac{d\chi(\mu)}{d\mu} = 2 \left(\frac{k + 1}{2} - p^m(\mu) \right) \frac{dp^m(\mu)}{d\mu} + \frac{1}{6\mu^2} \frac{(k - 1)^2}{k} > 0,$$

where the inequality follows from the fact that $dp^m(\mu)/d\mu < 0$ and $p^m(\mu) > p^m(1) = k^2/(2k - 1) > (k + 1)/2$. Therefore $\chi(\mu)$ is strictly increasing in μ ; moreover,

$$\begin{aligned} \lim_{\mu \rightarrow 0} \chi(\mu) &= -\infty < 0; \\ \chi(1) &= \frac{k(k - 1)^3}{2k - 1} > 0. \end{aligned}$$

Let μ^\dagger be the unique solution to the equation $\chi(\mu) = 0$. Then $E [c(M)|\mathcal{B} (p^m (\mu))] > E [M]$ for all $\mu < \mu^\dagger$ and $E [c(M)|\mathcal{B} (p^m (\mu))] < E [M]$ for all $\mu > \mu^\dagger$. This completes the proof. ■

The monopolist takes advantage of the increasing negative dependence between risk and risk aversion. In particular, when these two consumer characteristics become more negatively dependent, the monopolistic insurer has an incentive to set a high premium to rule out the “bad risks.” As a result, the monopolist optimally chooses a price in the region where negative correlation property emerges.

6 Local Adverse/Advantageous Selection and Positive/Negative Correlation Property

In this section, we use the example we analyzed in the previous section to clarify the connections between some important concepts related to selection markets, such as adverse/advantageous selection and positive/negative correlation property. In the classic models of insurance with one-dimensional consumer heterogeneity in risk types (e.g., Arrow 1963, Pauly 1974, Rothschild and Stiglitz 1976 and Wilson 1977), *adverse selection* refers to the idea that the average risk of those who choose to purchase insurance worsens as the insurance premium rises. The reason is very simple: as premium rises, the marginal consumer to drop out of coverage is the least risky among the insured in a *one*-dimensional model. To the extent that the average risks will translate into average cost of coverage for the insurance company, this translates into *average cost curve (and the marginal cost curve) being an increasing function of price*.²⁹

Einav, Finkelstein and Cullen (2010) generalize this insight to models of potential multidimen-

²⁹Or equivalently, the average (and the marginal) cost curve is a decreasing function of the fraction insured because the fraction insured is a monotonically decreasing function of the price.

sional consumer heterogeneity. Their approach to test for the nature of selection in the insurance market treats the marginal cost curve as a *sufficient statistics* for the distribution of consumers' potentially multidimensional heterogeneity, and to the extent that researchers have access to exogenous price variations to estimate the marginal cost curve, it would indeed be a very attractive approach relative to a fully structural alternative.

The first clarification we would like to make in this section is the following: in models with multidimensional consumer heterogeneity, *average cost and marginal cost curves are generically non-monotonic functions of premium.*³⁰ Thus, it is useful to define a *local* notion of selection based on how the marginal cost curve, which we will denote by $MC(p)$, is *locally* related to the premium change:

Definition 4 (Local Adverse/Advantageous Selection) *The market is said to be subject to local adverse selection at price p if $MC'(p) > 0$. Similarly, the market is said to be subject to local advantageous selection at price p if $MC'(p) < 0$.*

Remark 3 *If consumers have one-dimensional heterogeneity in risk, local selection is always adverse, as the marginal insurer buyer is always riskier as the premium increases; thus there is no distinction in the notion of local selection and global selection. Moreover, in the one-dimensional heterogeneity model, the average cost curve, which we denote by $AC(p)$, is always above the marginal cost curve except at the price levels that exceed the maximum WTP for insurance (where the two curves coincide). Thus $AC'(p) > 0$.*

Remark 3 also implies that local advantageous selection is a phenomenon that may arise only when consumer heterogeneity is multidimensional.

Remark 4 *When consumers have multidimensional heterogeneity, whether $AC(p)$ is increasing (or decreasing) at p depends on whether $AC(p)$ is larger (or smaller) than $MC(p)$. Thus, marginal cost curve locally increasing at p does not imply, and is not implied by the average cost locally increasing at p .*

In contrast, the concept of *positive or negative correlation* property (see Definition 1 for the case of competitive market structure, but can be obviously generalized to any market structure) refers to whether the *equilibrium* correlation between insurance purchase and *ex post* realization or risk is positive or negative. It is a property of the equilibrium, which depends on the *complete* distribution of consumer heterogeneity, while the local adverse/advantageous selection is determined by the *local* properties of the distribution of consumer heterogeneity on the iso-WTP sub-space. The market

³⁰As we pointed out in Footnote 6, Einav and Finkelstein (2011) note this as well in a footnote in their paper, but proceed assuming monotonicity. Mahoney and Weyl (2016) also realize that importance of studying non-monotone cost curves (footnote 5, p. 4): "It is possible that these slopes have different signs over different ranges or that the two have slopes of different signs over a particular range. All of these cases do not fall cleanly into one category or the other and are not our focus in what follows. It would be interesting to extend our analysis to such cases." Our discussion below is to emphasize that the non-monotonicity is a generic property of models of multidimensional heterogeneity.

equilibrium price depends on the market structure, of course; in the price-cost graph as in Figure 4 below, the competitive market equilibrium price is determined by the intersection of the 45 degree line with $AC(p)$ curve, while the monopolistic equilibrium price is determined by the intersection of the marginal revenue curve, which is $p + D(p)/D'(p)$, with $MC(p)$ curve.

Now we illustrate the notions of local adverse/advantageous selection in the context of the example environment we studied in Section 5.2.2. We define, as a function of premium p , the demand for insurance, the total (expected) cost, the average (expected) cost and the marginal (expected) cost at price p are equal to:

$$\begin{aligned}
D(p) &= \int_{\theta \in \mathcal{B}(p)} dH(m, \lambda), \\
TC(p) &= \int_{\theta \in \mathcal{B}(p)} c(m) dH(m, \lambda), \\
AC(p) &= E[c(M)|\mathcal{B}(p)] = \frac{TC(p)}{D(p)}, \\
MC(p) &= E[c(M)|v(\theta) = p] = \frac{\int_{\theta \in \{\theta: v(\theta; 1) = p\}} c(m) dH(m, \lambda)}{\int_{\theta \in \{\theta: v(\theta; 1) = p\}} dH(m, \lambda)} \equiv \frac{dTC(p)}{dp} \bigg/ \frac{dD(p)}{dp}.
\end{aligned}$$

Let us focus on the case where $k > 1$.

When $\mu = 0$, i.e., when the risk type and risk preference type are independent and uniformly distributed, $MC(p; \mu = 0)$ is given by:

$$MC(p; \mu = 0) = \begin{cases} \frac{1}{2}p, & \text{if } p \in [0, 1] \\ \frac{1}{2}, & \text{if } p \in [1, k] \\ \frac{1}{2}[(p - k) + 1], & \text{if } p \in [k, k + 1]. \end{cases}$$

It can be verified that $MC(p; \mu = 0)$ is nondecreasing in p .³¹ Hence, the AC curve lies above the MC curve and is increasing in p for $p \in [0, k + 1]$. This corresponds to the graphical representation of *adverse selection* in Einav, Finkelstein and Cullen (2010) and Mahoney and Weyl (2016). It is worth noting that the monotonicity of MC and AC in this example is a sufficient but not a necessary condition for the emergence of positive correlation property under any market structure. Indeed, from Proposition 3 we must have $AC(p) > AC(\underline{m} + k\underline{\lambda}) = E[M]$ for $p \in (\underline{m} + k\underline{\lambda}, \overline{m} + k\overline{\lambda})$ when $\mu = 0$ regardless the shape of cost curves.

When $\mu = 1$, i.e., when M and Λ are perfectly negative dependent, $MC(p; \mu = 1)$ is given by,

$$MC(p; \mu = 1) = \frac{k - p}{k - 1} \text{ for } p \in [1, k].$$

The joint distribution $H(\cdot, \cdot)$ degenerates to a one-dimensional distribution and there exists a one-to-one mapping between m and λ . In particular $\lambda = 1 - m$. Consumer's WTP is $k + (1 - k)m$,

³¹If M and Λ are not uniformly distributed as in this example, the independence of M and Λ does not generally imply that the MC curve in terms of p is monotonic. See Example 4 below for a construction.

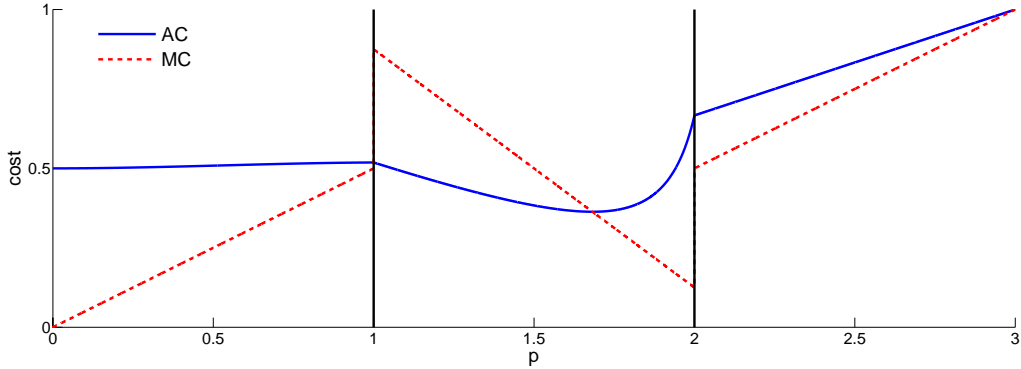


Figure 4: Cost Curves for $\mu \in (0, 1)$

which is decreasing in m when $k > 1$. Therefore, the lower risk types have higher WTP for a given price. This implies directly that the risk of the marginal consumer is decreasing as price increases and $AC(p)$ is decreasing in p . This corresponds to the graphical representation of *advantageous selection* in Einav, Finkelstein and Cullen (2010) and Mahoney and Weyl (2016).

When $\mu \in (0, 1)$, the marginal cost $MC(p; \mu \in (0, 1))$ curve is given by:

$$MC(p; \mu \in (0, 1)) = \begin{cases} \frac{1}{2}p & \text{if } p \in [0, 1] \\ \left[\frac{\mu}{k-1} \frac{k-p}{k-1} + \frac{1-\mu}{2k} \right] / \left[\frac{\mu}{k-1} + \frac{1-\mu}{k} \right] & \text{if } p \in [1, k] \\ \frac{1}{2}[(p-k)+1] & \text{if } p \in [k, k+1] \end{cases} .$$

It can be verified that MC is *non-monotone* in p . As depicted in Figure 4, the MC curve is increasing in p for $p \in [0, 1]$ and $p \in [k, k+1]$, and decreasing for $p \in [1, k]$.

The *non-monotonicity* of the MC curve is not a unique property of this example. In fact, when we model dependence on the joint distribution with positive density almost everywhere on $[\underline{m}, \bar{m}] \times [\underline{\lambda}, \bar{\lambda}]$, it is impossible to obtain a globally decreasing MC curve. To see this, notice that only consumers with high risks purchase insurance when the price is below but sufficiently close to the maximum price $\bar{m} + k\bar{\lambda}$. Hence, the MC is high when price is close to \bar{m} . By the

same token, when the price decreases gradually and is approaching to the minimum price $\underline{m} + k\lambda$, the additional consumers are those with low risks. In other words, MC is low (close to \underline{m}) near the lowest price. Therefore, MC always starts up low and ends up high. For negative correlation property to emerge, MC needs to change its monotonicity at least twice as Figure 4 illustrates. Only under joint distribution functions that puts zero density on certain combinations of (m, λ) is it possible to obtain a globally decreasing MC curve. For example, when M and Λ exhibit perfectly negative dependence (i.e. $\mu = 1$), the MC curve is globally decreasing as in Einav, Finkelstein and Cullen (2010) and Mahoney and Weyl (2016).

Having clarified the distinction between the local notion of adverse/advantageous selection in models of multidimensional heterogeneity and the equilibrium (and thus global) notion of positive/negative correlation property, the following claim discusses their connections:

Claim 2 (Connection between Local Advantageous/Adverse Selection and Equilibrium Negative/Positive Correlation Property)

1. If the market exhibits negative correlation property in equilibrium, then the market is subject to local advantageous selection at some prices.
2. However, the reverse does not hold.
3. Moreover, it is possible that selection is locally advantageous at the equilibrium price p^* , yet the market exhibits positive correlation property in equilibrium.

The first part of Claim 2 is obvious. Suppose that the market is subject to local adverse selection for all prices. Then $MC(p)$ is increasing in p for all p by definition. It follows immediately that $AC(p)$ is increasing in p globally, which implies that $E[c(M)|\mathcal{B}(p^*)] = AC(p^*) > AC(0) = E[c(M)]$. We provide an example to illustrate the second and third part of Claim 2.

Example 4 Suppose M and λ are independent and are both drawn from a mixture of left-right truncated normal distributions between 0 and 1 with the following marginal density functions,

$$f(m) = \frac{1}{3\sigma_m} \frac{\phi\left(\frac{m-m_1}{\sigma_m}\right)}{\Phi\left(\frac{1-m_1}{\sigma_m}\right) - \Phi\left(\frac{0-m_1}{\sigma_m}\right)} + \frac{1}{3\sigma_m} \frac{\phi\left(\frac{m-m_2}{\sigma_m}\right)}{\Phi\left(\frac{1-m_2}{\sigma_m}\right) - \Phi\left(\frac{0-m_2}{\sigma_m}\right)} + \frac{1}{3\sigma_m} \frac{\phi\left(\frac{m-m_3}{\sigma_m}\right)}{\Phi\left(\frac{1-m_3}{\sigma_m}\right) - \Phi\left(\frac{0-m_3}{\sigma_m}\right)},$$

and

$$g(\lambda) = \frac{1}{2\sigma_\lambda} \frac{\phi\left(\frac{\lambda-\lambda_1}{\sigma_\lambda}\right)}{\Phi\left(\frac{1-\lambda_1}{\sigma_\lambda}\right) - \Phi\left(\frac{0-\lambda_1}{\sigma_\lambda}\right)} + \frac{1}{2\sigma_\lambda} \frac{\phi\left(\frac{\lambda-\lambda_2}{\sigma_\lambda}\right)}{\Phi\left(\frac{1-\lambda_2}{\sigma_\lambda}\right) - \Phi\left(\frac{0-\lambda_2}{\sigma_\lambda}\right)}.$$

The joint density function is $h(m, \lambda) = f(m)g(\lambda)$. Suppose that $v((m, \lambda); x) = m+2\lambda$, $(m_1, m_2, m_3) = (0.1, 0.2, 0.9)$, $(\lambda_1, \lambda_2) = (0.1, 0.3)$, and $\sigma_m = \sigma_\lambda = 0.05$. The joint density function is illustrated in Figure 5. The profit curve and marginal cost curve are illustrated in Figure 6 and Figure 7 respectively. Figure 6 indicates that the equilibrium price is unique and lies between 0.5 and 0.6. While

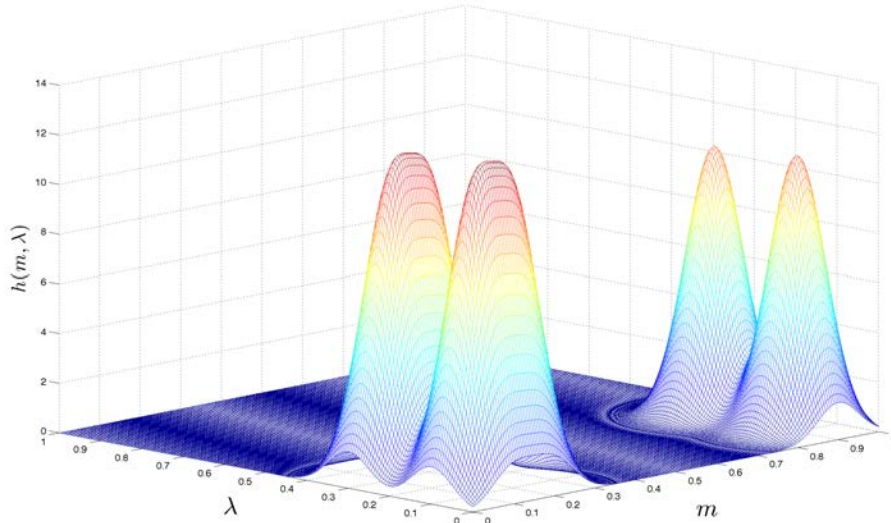


Figure 5: The Joint Density Function $h(m, \lambda)$ in Example 4

Figure 7 indicates that the market is subject to local advantageous selection for $p \in [0.5, 0.6]$. From Proposition 1, the market always exhibit positive correlation property in competitive equilibrium. Therefore, from this example we know the market can still be subject to local advantageous selection at the equilibrium price even though the positive correlation property holds.

6.1 Discussions

Clarifying the distinction between *local* selection (adverse or advantageous) and the *equilibrium* positive/negative correlation is important. As noted by de Meza and Webb (2016), this distinction is sometimes not clearly made. Einav, Finkelstein and Levin (2010, p.316) state that, “contract j is adversely selected if the expected cost of insuring j ’s enrollees under contract j is greater than the expected cost of insuring the population I under contract j ” and is advantageously selected otherwise. This is in fact referring to the equilibrium positive or negative correlation property. In contrast, Einav, Finkelstein and Cullen (2010, p. 879) state that “... the sign of the slope of the marginal cost curve tells us whether the resultant selection is adverse (or the marginal cost is increasing in price) or advantageous (if marginal cost is decreasing in price.” In our terminology, this is a local definition of adverse or advantageous selection. Since the marginal cost curve is generically non-monotonic in price, a property Einav, Finkelstein and Cullen (2010) do not emphasize but Einav and Finkelstein (2011, footnote 7, p. 124) do acknowledge, the two notions will often reach

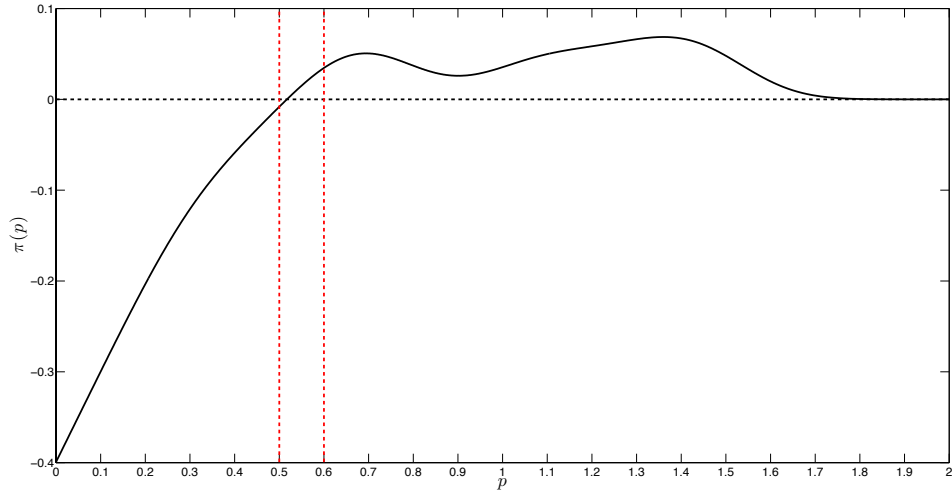


Figure 6: The Profit Curve $\pi(p)$ in Example 4

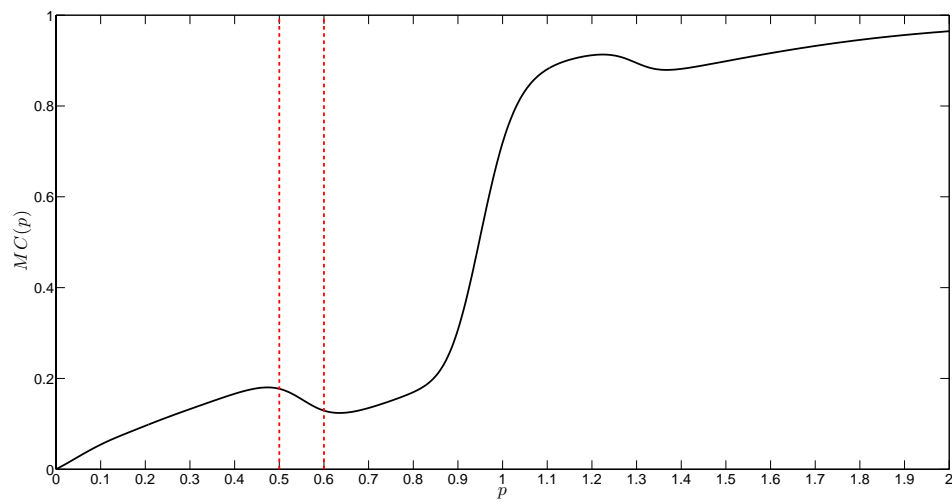


Figure 7: The Marginal Cost Curve $MC(p)$ in Example 4

different conclusions under multiple consumer heterogeneity,, as explicitly shown in our Claim 2 above. This is a point that was already made by de Meza and Webb (2016) in their two-type (bad and good risks) example, where they also argue that the local sign of the slope of the *average* cost with respect to *quantity*, which may not be monotonic, can be a useful measure of selection as well.

Azevedo and Gottlieb (2016) introduce a notion of intensive margin selection coefficient that measures the difference between the marginal changes of the premium and the marginal increase in the cost of insuring the consumers choosing a particular level of coverage, if they were to switch to a contract with an infinitesimally better coverage. Notice that their notion is with respect to the local changes in coverage, as supposed to the local changes in the premium used in our definition of local adverse/advantageous selection. They suggest that this notion is related to the positive correlation test. Indeed if the intensive margin selection coefficient is positive (negative), it means that, *locally*, there is a positive (negative) correlation between the *ex post* risk realization and the insurance coverage generosity. More importantly, the intensive margin selection coefficient not only provides the sign of the correlation but also the magnitude of the positive correlation. They state that “It is possible that there is adverse selection in one region of the contract space, and advantageous selection in another region.” In Section 8 below, we show in Proposition 8 that under a set of mild conditions, the multiple contract competitive equilibrium exhibits positive correlation property throughout the contract space. We will provide further connections between their intensive margin selection coefficient and the standard correlation test after we present Proposition 8.

7 Imperfectly Competitive Insurance Market

In this section, we generalize the market structure and study imperfect competition. We model imperfectly competitive insurance market as follows.³² Suppose that there are two insurance firms on the market where they engage in a modified “Bertrand competition” by setting a price for insurance of quality x . Different from the standard Bertrand model, we assume that consumers cannot compare prices perfectly; instead, a consumer receives a *noisy* signal regarding which of the two firms has a lower price, and then he/she inspects the actual price of the firm indicated by the noisy signal, and finally he/she decides whether to buy the product accordingly.^{33, 34} The noisy signal regarding which firm has the lower price creates spurious product differentiation and gives rise to market power to firms and thus induces imperfect competition.

Specifically, after firm $i \in \{1, 2\}$ posts price p_i , each consumer receives a signal $s \in \{1, 2\}$ about

³²Our approach of modeling imperfect competition is in the same spirit as Fisher and Plan (2015).

³³It is assumed that the consumer cannot switch to the other firm once the price of the chosen firm fully reveals. This specification of the consumer behavior can be rationalized by a sufficiently high switching cost.

³⁴Note that our approach to parameterize the imperfect competition is related to but distinct from from Lester et al. (2016). In their paper they assume that a buyer samples one offer with probability $p_1 \in (0, 1)$ and two offers with probability $p_2 = 1 - p_1$. The higher p_2 is, the more fierce the competition is.

which firm has the lower price as follows: given (p_1, p_2) ,

$$s = \begin{cases} 1 & \text{if } p_1 - p_2 + \epsilon \leq 0 \\ 2 & \text{otherwise,} \end{cases}$$

where $\epsilon \sim \mathcal{N}(0, \sigma_s^2)$.³⁵ It is clear that a consumer always follows the signal: if $s = i$, she will find out the actual price p_i and decide between purchasing insurance at price p_i and staying uninsured. Hence, conditional on the price vector (p_1, p_2) , the probability that a consumer considers purchasing from firm i is $\Phi\left(\frac{p_j - p_i}{\sigma_s}\right)$.³⁶ Conditional on observing firm i 's price p_i , the purchase decision of type- θ consumer remains the same as before: she will purchase insurance at price p_i if and only if $v(\theta; x) \geq p_i$, i.e., if and only if $\theta \in \mathcal{B}(p_i)$ where $\mathcal{B}(\cdot)$ is defined in (1). Note that σ_s^2 represents a measure of the market competitiveness. When $\sigma_s^2 = 0$, consumers always buy from the firm with the lower price, which indicates fierce price competition. When $\sigma_s^2 = \infty$, the signal is completely uninformative and the consumer randomly chooses between the two offers and each firm behaves as if they were a monopoly.

Now we can set up the strategic pricing game between the two firms. Fixing p_j , firm i chooses p_i to maximize the expected profit:

$$\Pi_i(p_i, p_j) = \Phi\left(\frac{p_j - p_i}{\sigma_s}\right) \int_{\theta \in \mathcal{B}(p_i)} [p - xc(m)] dH(m, \lambda) \equiv \Phi\left(\frac{p_j - p_i}{\sigma_s}\right) \pi(p_i), \quad (24)$$

where $\pi(\cdot)$ is defined in (4). Notice that a nice feature of our formulation of the imperfect competition is that the price competition between the two firms only affects which firm is in the consideration set of the consumer, not the subsequent decision of whether to purchase the insurance.

Assumption 3 (i) $\pi(p)$ is strictly single-peaked in p for $p \in \left(v(\underline{m}, \underline{\lambda}); x, v(\overline{m}, \overline{\lambda}); x\right)$; (ii) $\max_{p \geq 0} \{\pi(p)\} > 0$; (iii) $\pi(p)$ is differentiable and log-concave in $p \in (p^*, p^m]$, and $\pi'(p^m) = 0$ where p^* is the competitive equilibrium price and p^m is the monopolist's profit maximizing price.³⁷

The differentiability and single-peakedness of the profit function guarantees that the profit function under monopoly is well-behaved and the first-order condition is sufficient to pin down the optimal price; and $\max_{p \geq 0} \{\pi(p)\} > 0$ ensures that a monopolistic insurer will not exit the market. Finally, as it will be clear later, log-concavity ensures the symmetry and uniqueness of the equilibrium.

³⁵The assumption of Gaussian noise can be easily relaxed. All the results obtained in this section can be generalized to a noise Z parameterized by $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ with support $-\infty \leq \underline{z} \leq \overline{z} \leq \infty$, whose CDF $\Psi(z; \alpha)$ and PDF $\psi(z; \alpha)$ satisfy: (i) $\psi(z; \alpha) = \psi(-z; \alpha)$; (ii) $\partial \frac{\psi(z; \alpha)}{\Psi(z; \alpha)} / \partial z < 0$; (iii) $\partial \psi(0; \alpha) / \partial \alpha > 0$; (iv) $\lim_{\alpha \rightarrow \underline{\alpha}} \psi(0; \alpha) = 0$ and $\lim_{\alpha \rightarrow \overline{\alpha}} \psi(0; \alpha) = \infty$.

³⁶An alternative way of modeling is to assume that each potential consumer receives two signals $s_i = p_i + \epsilon_i$ and examine the price of the firm with the lower signal, where ϵ_i is assumed to be independently identically distributed according to a normal distribution with mean 0 and variance $\sigma_s^2/2$.

³⁷The profit function of the example in Section 5.2.2 for $k > 1$ satisfies Assumption 3. See Online Appendix C for the proof.

Lemma 2 (Equilibrium Prices under Imperfect Competition) Suppose Assumption 3 is satisfied. For any $\sigma_s^2 > 0$, there exists a unique equilibrium $(\hat{p}_1^*, \hat{p}_2^*)$ where $\hat{p}_1^* = \hat{p}_2^* = p^e \in (p^*, p^m)$ and p^e is the solution to

$$\frac{\pi'(p^e)}{\pi(p^e)} = \frac{1}{\sigma_s} \sqrt{\frac{2}{\pi}}. \quad (25)$$

Proof. First, notice that fixing p_j , we must have $p_i \in [p^*, p^m]$. For a price below p^* , the corresponding profit is negative, which is strictly dominated by $p_i = p^*$. For a price above p^m , firm i 's profit is less than $\Pi_i(p^m, p_j)$ due to the fact that $\Phi\left(\frac{p_j - p_i}{\sigma_s}\right)$ is strictly decreasing in p_i and $\pi(p_i) \leq \pi(p^m)$. The first order condition with respect to p_i yields,

$$\frac{1}{\sigma_s} \frac{\phi\left(\frac{p_j - p_i}{\sigma_s}\right)}{\Phi\left(\frac{p_j - p_i}{\sigma_s}\right)} = \frac{\pi'(p_i)}{\pi(p_i)}. \quad (26)$$

Next, we show that if an equilibrium exists, it must be symmetric. Suppose to the contrary that $\hat{p}_1^* > \hat{p}_2^*$ without loss of generality. It follows directly that $(\hat{p}_1^* - \hat{p}_2^*)/\sigma_s > 0$; together with the first order condition (26), we must have

$$\frac{\pi'(\hat{p}_1^*)}{\pi(\hat{p}_1^*)} = \frac{1}{\sigma_s} \frac{\phi\left(\frac{\hat{p}_2^* - \hat{p}_1^*}{\sigma_s}\right)}{\Phi\left(\frac{\hat{p}_2^* - \hat{p}_1^*}{\sigma_s}\right)} = \frac{1}{\sigma_s} \frac{\phi\left(\frac{\hat{p}_1^* - \hat{p}_2^*}{\sigma_s}\right)}{1 - \Phi\left(\frac{\hat{p}_1^* - \hat{p}_2^*}{\sigma_s}\right)} > \frac{1}{\sigma_s} \frac{\phi\left(\frac{\hat{p}_1^* - \hat{p}_2^*}{\sigma_s}\right)}{\Phi\left(\frac{\hat{p}_1^* - \hat{p}_2^*}{\sigma_s}\right)} = \frac{\pi'(\hat{p}_2^*)}{\pi(\hat{p}_2^*)}.$$

Therefore, $\hat{p}_1^* < \hat{p}_2^*$ from Assumption 3, a contradiction. Imposing the symmetry condition $\hat{p}_1^* = \hat{p}_2^* = p^e$ to the above first order condition yields equation (25).

Lastly, we prove the existence and uniqueness of the equilibrium. From the definition of p^* and Assumption 3, $\pi'(p^m) = 0$ and $\pi(p^*) = 0$, which implies directly that $\lim_{p^e \rightarrow p^*} \pi'(p^e)/\pi(p^e) = \infty$ and $\lim_{p^e \rightarrow p^m} \pi'(p^e)/\pi(p^e) = 0$. Therefore, for any $\sigma_s^2 > 0$, there exists a unique solution to equation (25). This completes the proof. ■

The following lemma derives the comparative statics of the equilibrium price with the market competitiveness.

Lemma 3 Suppose Assumption 3 is satisfied. The equilibrium price p^e is strictly increasing in σ_s^2 . Moreover, $\lim_{\sigma_s^2 \rightarrow \infty} p^e = p^m$ and $\lim_{\sigma_s^2 \rightarrow 0} p^e = p^*$.

The proof follows immediately from Assumption 3 that $\pi(p)$ is log-concave. This result is intuitive. When the market is imperfectly competitive, the equilibrium price lies between the competitive equilibrium price and the monopoly price. The equilibrium price becomes lower as the insurance market becomes more competitive.

Proposition 7 Suppose Assumption 3 is satisfied and suppose $E[c(M)|\mathcal{B}(p^m)] < E[c(M)]$. Then there exists $\underline{\sigma}_s$ and $\bar{\sigma}_s$, with $\underline{\sigma}_s \leq \bar{\sigma}_s$, such that positive (respectively, negative) correlation property emerges in equilibrium for $\sigma_s^2 < \underline{\sigma}_s^2$ (respectively, $\sigma_s^2 > \bar{\sigma}_s^2$). Furthermore, if there exists a unique solution to $AC(p) = E[c(M)]$ for $p \in [p^*, p^m]$, then $\bar{\sigma}_s = \underline{\sigma}_s$.

Proof. From Proposition 1, we must have $E [c(M)|\mathcal{B}(p^*)] > E [c(M)]$. In addition, $E [c(M)|\mathcal{B}(p^m)] < E [c(M)]$ holds by assumption. From Lemma 3 and the continuity of $E [c(M)|\mathcal{B}(p)]$, we have

$$\begin{aligned} \lim_{\sigma_s^2 \rightarrow 0} E [c(M)|\mathcal{B}(p^e)] &= E [c(M)|\mathcal{B}(p^*)] > E [c(M)] ; \\ \lim_{\sigma_s^2 \rightarrow \infty} E [c(M)|\mathcal{B}(p^e)] &= E [c(M)|\mathcal{B}(p^m)] < E [c(M)] . \end{aligned}$$

Therefore, positive (respectively, negative) correlation property emerges in equilibrium when σ_s^2 is sufficiently low (respectively, high).

Denote the unique solution (if it exists) to $AC(p) = E [c(M)]$ for $p \in [p^*, p^m]$ as \tilde{p} . Because $E [c(M)|\mathcal{B}(p^*)] > E [c(M)]$, we must have $E [c(M)|\mathcal{B}(p)] > E [c(M)]$ for $p \in [p^*, \tilde{p}]$ and $E [c(M)|\mathcal{B}(p)] < E [c(M)]$ for $p \in [\tilde{p}, p^m]$. Define $\tilde{\sigma}_s^2$ as the solution to $p^e(\sigma_s^2) = \tilde{p}$. From Lemma 3, $\tilde{\sigma}_s^2$ exists and is unique. Moreover, $p^e(\sigma_s^2) \leq \tilde{p}$ for $\sigma_s^2 \leq \tilde{\sigma}_s^2$. Therefore, positive (respectively, negative) correlation property emerges in equilibrium for $\sigma_s^2 < \tilde{\sigma}_s^2$ (respectively, $\sigma_s^2 > \tilde{\sigma}_s^2$). This completes the proof. ■

8 Extension: Endogenizing the Contracts

In the basic model, the quality of the insurance is predetermined, and hence we are comparing no purchase with purchasing $x \in (0, 1]$. In this section we relax this assumption and show that the main results derived in both the competitive and monopolistic case are indeed robust.

8.1 Competitive Insurance Market: Endogenous Contract

Consider a perfectly competitive market. Instead of allowing the competing insurance firms to choose the quality of the contract arbitrarily, we assume that a firm can provide contracts from a set $X = \{x_0, x_1, \dots, x_N\}$ with $0 = x_0 < x_1 < \dots < x_N \leq 1$, where x_0 refers to the null contract costing nothing (i.e., $C(\theta; x_0) \equiv x_0 \cdot c(m) = 0$) and providing zero utility (i.e., $v(\theta; x_0) = 0$) to all consumers. This approach allows us to endogenize the insurance quality with a minimal departure from the basic model we considered in Section 3. We would like to study whether negative correlation or positive correlation property will emerge in equilibrium when comparing x_i versus x_j for $i < j$, as well as those who do not purchase. Denote the price of contract x_i by p_i for $i \in \{0, 1, \dots, N\}$. Fixing the premium vector $p = (p_0, p_1, \dots, p_N)$, denote the set and measure of consumers that selects x_i by $\mathcal{B}_i(p)$ and $D_i(p)$ respectively, that is,

$$\mathcal{B}_i(p) \equiv \left\{ \theta : v(\theta; x_i) - p_i \geq \max_{j < i} \{v(\theta; x_j) - p_j\}, v(\theta; x_i) - p_i > \max_{j > i} \{v(\theta; x_j) - p_j\} \right\},^{38} \quad (27)$$

and

$$D_i(p) \equiv \int_{\theta \in \mathcal{B}_i(p)} dH(m, \lambda). \quad (28)$$

³⁸It is assumed that whenever a consumer is indifferent between two contracts, she always prefers the one with the higher coverage.

The average *ex post* realization of risk among those who purchase insurance x_i is:

$$\mathbb{E} [c(M)|\mathcal{B}_i(p)] = \frac{\int_{\theta \in \mathcal{B}_i(p)} c(m) dH(m, \lambda)}{\int_{\theta \in \mathcal{B}_i(p)} dH(m, \lambda)}. \quad (29)$$

Next, we define the employed equilibrium notion, which borrows from Azevedo and Gottlieb (2016). This definition corresponds to the *weak equilibrium* in Azevedo and Gottlieb (2016).

Definition 5 *The price vector $p^* = (p_0^*, p_1^*, \dots, p_N^*)$ is a **competitive equilibrium** if*

- i. For each contract x_i , firms make no profits.*
- ii. Consumers make purchase decision and choose contracts optimally.*

This price-taking definition requires firms to earn zero profits on each contract (either with positive or zero demand) and hence rules out cross-subsidies between contracts.³⁹ The existence of equilibrium is guaranteed by Theorem 1 and Proposition 1 in Azevedo and Gottlieb (2016). It is worth noting that the equilibrium premium of the null contract is zero if the corresponding demand is strictly positive.

Next, we define positive and negative correlation property with multiple contracts. It is useful to denote the set of contracts that induces positive insurance demand by $\tilde{X}(p)$ and the corresponding price vector by \tilde{p} holding fixed p and X .

Definition 6 (Positive and Negative Correlation Property with Multiple Contracts)

Suppose $|\tilde{X}(p^)| \geq 2$. The insurance market exhibits positive correlation property in equilibrium if for every pair $x_i, x_j \in \tilde{X}(p^*)$ with $i > j$, $\mathbb{E} [c(M)|\mathcal{B}_i(p^*)] > \mathbb{E} [c(M)|\mathcal{B}_j(p^*)]$, and it exhibits negative correlation property if $\mathbb{E} [c(M)|\mathcal{B}_i(p^*)] < \mathbb{E} [c(M)|\mathcal{B}_j(p^*)]$.*

We assume that the insurance coverage x affects consumers' WTP in the natural way. Specifically, we impose the following assumptions on $v(\cdot)$:

Assumption 4 $\partial v / \partial x \geq c(m)$.

Assumption 4 simply says that consumer's WTP for an additional unit of insurance quality is no less than the cost to the insurance firms. Assumption 4 implies and is a natural extension of Assumption 2 for the multiple contract case.

Assumption 5 $\partial^2 v / \partial m \partial x > 0$ and $\partial^2 v / \partial \lambda \partial x > 0$.

Assumption 5 states that consumer's marginal WTP for insurance quality of the contract coverage (i.e. $\partial v / \partial x$) is increasing in her risk type and in her risk preference type.

³⁹As argued by Azevedo and Gottlieb (2016), this equilibrium concept can be justified as the limit of a strategic model with differentiated products. The intuition is simple: a firm that cross-subsidizes contracts has incentives to sell contracts with positive profits only.

Assumption 6

$$\left. \frac{\partial v / \partial m}{\partial v / \partial \lambda} \right|_{x=x_N} > \dots > \left. \frac{\partial v / \partial m}{\partial v / \partial \lambda} \right|_{x=x_1} .$$

Denote the iso-WTP curve $v(\theta; x_i) - p_i = 0$ by $\lambda = \mathcal{I}_i(m)$. Assumption 6 simply says that the single-crossing condition holds for the iso-WTP curves of insurance coverage $x_i > x_j > 0$. Specifically, $\mathcal{I}_i(m)$ and $\mathcal{I}_j(m)$ intersect at most once in the (m, λ) -space. Moreover, the iso-WTP curve $\mathcal{I}_i(m)$ is steeper than $\mathcal{I}_j(m)$ at the intersection if they intersect.

It is clear that Assumption 4, 5 and 6 are satisfied in Example 2. To see this, recall that consumer's WTP for insurance is given by $v(\theta; x) = xm + \frac{x(2-x)}{2}\sigma^2\lambda$. Therefore, $\partial v / \partial x = m + (1-x)\sigma^2\lambda \geq c(m)$, $\partial^2 v / \partial m \partial x = 1 > 0$, and $\partial^2 v / \partial \lambda \partial x = (1-x)\sigma^2 \geq 0$, and Assumption 5 is satisfied. Moreover, for $x_i > x_j$,

$$\left. \frac{\partial v / \partial m}{\partial v / \partial \lambda} \right|_{x=x_i} = \frac{2}{\sigma^2} \frac{1}{2-x_i} > \frac{2}{\sigma^2} \frac{1}{2-x_j} = \left. \frac{\partial v / \partial m}{\partial v / \partial \lambda} \right|_{x=x_j} ,$$

and hence Assumption 6 is satisfied.

Consider a pair of contracts (x_i, p_i) and (x_j, p_j) with $x_i \neq x_j$. Due to Assumption 1, the set of consumer types that are indifferent between the two contracts, namely, $\{\theta : v(\theta; x_i) - p_i = v(\theta; x_j) - p_j\}$, can be represented by an *indifference curve*, which we denote by $\lambda = \mathcal{I}_{ij}(m)$ for $i \neq j$. It is obvious that $\mathcal{I}_{ij}(m) = \mathcal{I}_{ji}(m)$. When one of the contract pair is $(x_0, 0)$, i.e., the null insurance, then the indifference curve between contract (x_i, p_i) and $(x_0, 0)$, namely, $\lambda = \mathcal{I}_{i0}(m)$, coincides exactly with the iso-WTP curve $\lambda = \mathcal{I}_i(m)$ defined previously. The following lemma illustrates some additional properties of the indifference curves.

Lemma 4 *Suppose Assumption 1 and 5 are satisfied. Consider a pair of contracts (x_i, p_i) and (x_j, p_j) with $x_i > x_j$. Then,*

1. $\mathcal{I}_{ij}(m)$ is strictly decreasing in m ;
2. If $(\dot{m}, \dot{\lambda}) > (m, \mathcal{I}_{ij}(m))$, then type- $(\dot{m}, \dot{\lambda})$ consumer prefers contract x_i over x_j .
3. If in addition Assumption 6 is satisfied and $x_j > 0$, then $\mathcal{I}_{ij}(m)$ and $\mathcal{I}_{i0}(m)$ obey single crossing condition and $\mathcal{I}_{ij}(m)$ is steeper than $\mathcal{I}_{i0}(m)$ at the intersection.

Proof. To prove (1), we note that, from the implicit function theorem, the slope of $\mathcal{I}_{ij}(m)$ is,

$$\mathcal{I}'_{ij}(m) = - \frac{\left. \partial v / \partial m \right|_{x=x_i} - \left. \partial v / \partial m \right|_{x=x_j}}{\left. \partial v / \partial \lambda \right|_{x=x_i} - \left. \partial v / \partial \lambda \right|_{x=x_j}} .$$

If $x_j = 0$, $\left. \partial v / \partial m \right|_{x=x_j} = \left. \partial v / \partial \lambda \right|_{x=x_j} = 0$ and $\mathcal{I}'_{ij}(m) < 0$ from Assumption 1. For $x_j > 0$, Assumption 5 implies directly that both the numerator and the denominator are positive. Therefore, $\mathcal{I}'_{ij}(m) < 0$.

To prove (2), note that because $\partial^2 v / \partial m \partial x > 0$ and $\partial^2 v / \partial \lambda \partial x > 0$, $v((m, \lambda); x_i) - v((m, \lambda); x_j)$ is strictly increasing in both m and λ . Therefore, if $\hat{m} > m$ and $\hat{\lambda} > \mathcal{I}_{ij}(m)$, we have

$$\begin{aligned} v\left(\left(\hat{m}, \hat{\lambda}\right); x_i\right) - v\left(\left(\hat{m}, \hat{\lambda}\right); x_j\right) &> v\left(\left(m, \hat{\lambda}\right); x_i\right) - v\left(\left(m, \hat{\lambda}\right); x_j\right) \\ &> v\left(\left(m, \mathcal{I}_{ij}(m)\right); x_i\right) - v\left(\left(m, \mathcal{I}_{ij}(m)\right); x_j\right) = 0 \end{aligned}$$

where the last equality follows from the definition of $\mathcal{I}_{ij}(m)$.

To prove (3), note that it can be verified, after some algebra, that $-\mathcal{I}'_{ij}(m) > -\mathcal{I}'_{i0}(m)$ is equivalent to

$$\left. \frac{\partial v / \partial m}{\partial v / \partial \lambda} \right|_{x=x_i} > \left. \frac{\partial v / \partial m}{\partial v / \partial \lambda} \right|_{x=x_j},$$

which is the postulated Assumption 6. This completes the proof. ■

Lemma 5 *Suppose Assumptions 1, 4, 5 and 6 are satisfied, $x_i > x_j > x_0$, and $x_i, x_j, x_0 \in \tilde{X}(p^*)$. Then $p_i^*/x_i > p_j^*/x_j$.*

Proof. Because $x_i, x_j, x_0 \in \tilde{X}(p^*)$, both $\mathcal{I}_{i0}(m)$ and $\mathcal{I}_{j0}(m)$ cut the support $[\underline{m}, \bar{m}] \times [\underline{\lambda}, \bar{\lambda}]$ into two pieces. Suppose to the contrary that $p_i^*/x_i \leq p_j^*/x_j$. We consider three cases depending on whether $\mathcal{I}_{i0}(m)$ and $\mathcal{I}_{j0}(m)$ intersect.

Case I: $\mathcal{I}_{i0}(m)$ and $\mathcal{I}_{j0}(m)$ intersect at $(\hat{m}, \hat{\lambda}) \in [\underline{m}, \bar{m}] \times [\underline{\lambda}, \bar{\lambda}]$. Note that if $\mathcal{I}_{i0}(m)$ and $\mathcal{I}_{j0}(m)$ intersect at $(\hat{m}, \hat{\lambda})$, then type- $(\hat{m}, \hat{\lambda})$ consumer must be indifferent between contracts (x_i, p_i) and (x_j, p_j) , thus $\mathcal{I}_{ij}(m)$ also intersects both $\mathcal{I}_{i0}(m)$ and $\mathcal{I}_{j0}(m)$ at $(\hat{m}, \hat{\lambda})$. This case is depicted by Figure 8(a). From Lemma 4, the set of consumers selecting x_j can be rewritten as ,

$$\begin{aligned} \mathcal{B}_j(p^*) &\equiv \left\{ \theta : v(\theta; x_j) - p_j^* \geq \max_{j < i} \{v(\theta; x_i) - p_i^*\}, v(\theta; x_j) - p_j^* > \max_{j > i} \{v(\theta; x_i) - p_i^*\} \right\} \\ &\subseteq \left\{ \theta : v(\theta; x_j) - p_j^* \geq v(\theta; x_0) - p_0^*, v(\theta; x_j) - p_j^* \geq \max \{v(\theta; x_i) - p_i^*\} \right\} \\ &= \left\{ (m, \lambda) : \lambda \geq \mathcal{I}_{j0}(m), \lambda \leq \mathcal{I}_{ij}(m), \underline{m} \leq m \leq \bar{m}, \lambda \leq \lambda \leq \bar{\lambda} \right\} \\ &= \left\{ (m, \lambda) : \underline{m} \leq m \leq \hat{m}, \mathcal{I}_{j0}(m) \leq \lambda \leq \min\{\bar{\lambda}, \mathcal{I}_{ij}(m)\} \right\} := \mathcal{B}'_j(p^*). \end{aligned}$$

The shaded region of Figure 8(a) illustrates $\mathcal{B}'_j(p^*)$, which can be decomposed into two groups: those who only prefer contract (x_j, p_j) over not purchasing (the dark dotted region in Figure 8(a)) and those who prefer both (x_i, p_i) and (x_j, p_j) over not purchasing, but prefer (x_j, p_j) over (x_i, p_i) (the light dotted region in Figure 8(a)).

Next, notice that $p_0^* = 0$ from the postulated $x_0 \in \tilde{X}(p^*)$ and hence $v(\theta, x_0) - p_0^* = 0$ for all θ . Moreover, we have

$$v\left(c^{-1}\left(\frac{p_i^*}{x_i}\right), \hat{\lambda}; x_i\right) - p_i^* > x_i \cdot c\left(c^{-1}\left(\frac{p_i^*}{x_i}\right)\right) - p_i^* = 0 = v(\hat{m}, \hat{\lambda}; x_i) - p_i^*,$$

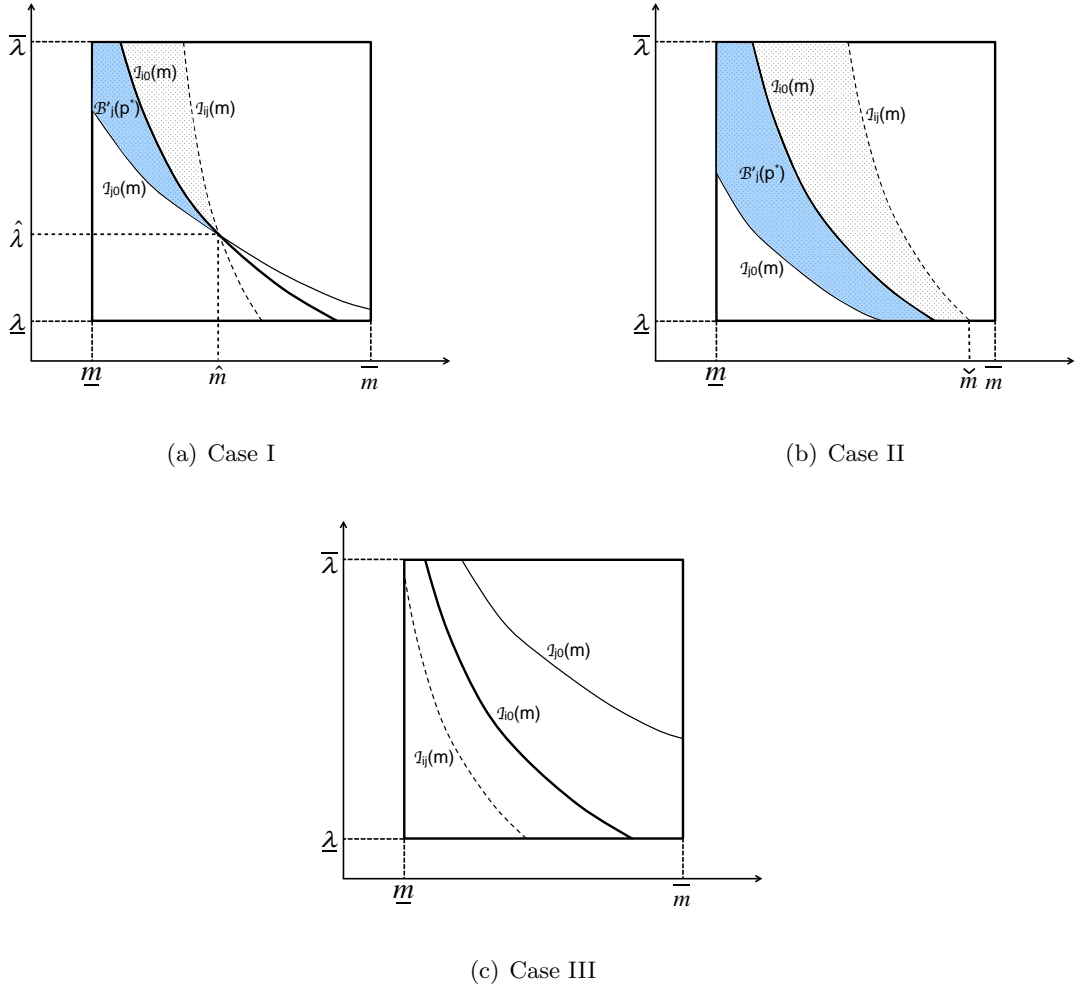


Figure 8: Proof of Proposition 8

where the inequality follows from Assumption 4. Therefore, $\hat{m} < c^{-1}(p_i^*/x_i)$ and

$$\mathbb{E} \left[c(M) | \mathcal{B}_j(p^*) \right] < c(\hat{m}) < c \left(c^{-1} \left(\frac{p_i^*}{x_i} \right) \right) = \frac{p_i^*}{x_i} \leq \frac{p_j^*}{x_j},$$

which is a contradiction to the zero profit condition required for contract x_j .

Case II: $\mathcal{I}_{i0}(m)$ lies above $\mathcal{I}_{j0}(m)$. This case is depicted by Figure 8(b). First, note that $\mathcal{I}_{ij}(m)$ must cut the support $[\underline{m}, \bar{m}] \times [\underline{\lambda}, \bar{\lambda}]$ into two pieces. Otherwise, either $D_i(p^*) = 0$ or $D_j(p^*) = 0$, a contradiction. Second, from the definition of indifference curve, $\mathcal{I}_{ij}(m)$ does not intersect with either $\mathcal{I}_{i0}(m)$ or $\mathcal{I}_{j0}(m)$. Together with Lemma 4, $\mathcal{I}_{ij}(m)$ must lie above $\mathcal{I}_{i0}(m)$ (see Figure 8(b)) and the risk type of consumers selecting x_j can be bounded from above by \check{m} , where $\check{m} = \mathcal{I}_{ij}^{-1}(\underline{\lambda})$ if $v(\bar{m}, \underline{\lambda}; x_i) - p_i^* > v(\bar{m}, \underline{\lambda}; x_j) - p_j^*$, and $\check{m} = \bar{m}$ otherwise. By definition, type- $(\check{m}, \underline{\lambda})$ consumer

weakly prefers contract (x_j, p_j^*) to (x_i, p_i^*) . Therefore,

$$p_i^* - p_j^* \geq v((\check{m}, \underline{\lambda}); x_i) - v((\check{m}, \underline{\lambda}); x_j) \geq c(\check{m})(x_i - x_j),$$

where the second inequality follows from Assumption 4. Therefore, $c(\check{m})$ can be bounded from above by,

$$c(\check{m}) \leq \frac{p_i^* - p_j^*}{x_i - x_j} \leq \frac{p_j^*}{x_j},$$

where the second inequality follows from the postulated $p_i^*/x_i \leq p_j^*/x_j$. Therefore,

$$\mathbb{E} \left[c(M) | \mathcal{B}_j(p^*) \right] < c(\check{m}) \leq \frac{p_j^*}{x_j},$$

which is a contradiction to the zero profit condition required for contract (x_j, p_j^*) .

Case III: $\mathcal{I}_{i0}(m)$ lies below $\mathcal{I}_{j0}(m)$. This case is depicted by Figure 8(c). By the same argument as in Case II, $\mathcal{I}_{ij}(m)$ does not intersect with either $\mathcal{I}_{i0}(m)$ or $\mathcal{I}_{j0}(m)$, and $\mathcal{I}_{ij}(m)$ must lie below $\mathcal{I}_i(m)$ (See Figure 8(c)). This implies that a consumer whose type is above or on $\lambda = \mathcal{I}_{i0}(m)$ will purchase contract (x_i, p_i^*) , and will end up with no insurance otherwise. Therefore, there will be no insurance demand for contract (x_j, p_j^*) , a contradiction to $D_j(p^*) > 0$. ■

Lemma 5 states that the unit price of the high quality insurance must be higher than that of the low quality insurance in a competitive equilibrium. The intuition is as follows. Consumers can be roughly classified into four groups based on their risk type (high risk v.s. low risk) and risk preference type (high risk aversion v.s. low risk aversion). Single crossing condition of the indifference curves guarantees that consumers selecting the low quality insurance must have both lower risk and lower risk aversion, and consumers of higher risk types (independent of risk preference type) will select into the high quality insurance. This market segmentation is illustrated by Figure 8(a). Now suppose that the unit price of the high quality insurance is lower relative to that of the low quality insurance. Then providing the low quality insurance helps the insurance firms to maintain a high premium and to attract consumers of low risk type (and lower risk aversion). This results in a net profit and contradicts to the zero profit condition required by the definition of equilibrium. Therefore, the unit price of insurance has to be strictly increasing in the insurance quality in a competitive equilibrium.⁴⁰

Proposition 8 (*Positive Correlation Property Holds in Competitive Equilibrium with multiple contracts*) *Suppose Assumptions 1, 4, 5 and 6 are satisfied, $|\tilde{X}(p^*)| \geq 2$ and $x_0 \in \tilde{X}(p^*)$. Then positive correlation property always holds in a competitive equilibrium without loadings.*

Proof. Suppose $|\tilde{X}(p^*)| = 2$. Then Proposition 1 applies and positive correlation property must

⁴⁰We conjecture that we can use arguments analogous to those in the proof of Proposition 2 to show that positive correlation property holds in competitive equilibrium with multiple contracts under a positive proportional loading factor as long as it is sufficiently low.

hold. Suppose $|\tilde{X}(p^*)| \geq 3$, for $x_i > x_j > 0$, we must have,

$$\mathbb{E} \left[c(M) \mid (m, \lambda) \in \mathcal{B}_i(p^*) \right] = \frac{p_i^*}{x_i} > \frac{p_j^*}{x_j} = \mathbb{E} \left[c(M) \mid (m, \lambda) \in \mathcal{B}_j(p^*) \right],$$

where the two equalities follow from the zero profit condition required for contract x_i and x_j , and the strict inequality follows from Lemma 5.

Let $x_s \equiv \min\{\tilde{X}(p^*) \setminus \{x_0\}\}$. It remains to be shown that

$$\mathbb{E} \left[c(M) \mid (m, \lambda) \in \mathcal{B}_s(p^*) \right] > \mathbb{E} \left[c(M) \mid (m, \lambda) \in \mathcal{B}_0(p^*) \right].$$

To prove this, suppose there is only one contract of quality x_s available on the market where consumer characteristics are drawn from the set $\mathcal{B}_s(p^*) \cup \mathcal{B}_0(p^*)$. It is obvious that p_s^* is an equilibrium price in such a market. Moreover, the set of buyers and the set of non-buyers of contract x_s are $\mathcal{B}_s(p^*)$ and $\mathcal{B}_0(p^*)$ respectively. We can thus apply Proposition 1 to conclude $\mathbb{E} \left[c(M) \mid (m, \lambda) \in \mathcal{B}_s(p^*) \right] > \mathbb{E} \left[c(M) \mid (m, \lambda) \in \mathcal{B}_0(p^*) \right]$. This completes the proof. ■

It is useful to connect the result stated in Proposition 8 to the notion of the intensive margin selection coefficient introduced in Azevedo and Gottlieb (2016), which they denoted by $S_I(x)$. They defined $S_I(x)$ for continuous contract space, but it is easy to adapt it to our discrete environment setting. Specially, the intensive margin selection coefficient evaluated at contract $x_k, k = 0, 1, 2, \dots, N - 1$ can be written as:

$$S_I(x_k) = \overbrace{\frac{p_{k+1}^* - p_k^*}{x_{k+1} - x_k}}^{\text{Marginal Premium Change}} - \overbrace{\frac{1}{\int_{\theta \in \mathcal{B}_k(p^*)} dH(\theta)} \frac{\int_{\theta \in \mathcal{B}_k(p^*)} [C(\theta; x_{k+1}) - C(\theta; x_k)] dH(\theta)}{x_{k+1} - x_k}}^{\text{Average Marginal Cost Change}}, \quad (30)$$

where the first term reflects the marginal premium change per unit increase in coverage locally at contract x_k , and the second term reflects the average marginal cost increase of covering *those consumers who purchase contract x_k* , i.e., those with $\theta \in \mathcal{B}_k(p^*)$, if they were to switch to contract x_{k+1} . In a competitive equilibrium, we have, for $i = k$ and $i = k + 1$

$$p_i^* = \mathbb{E} \left[C(\theta; x_i) \mid \theta \in \mathcal{B}_i(p^*) \right] = \frac{\int_{\theta \in \mathcal{B}_i(p^*)} C(\theta; x_i) dH(\theta)}{\int_{\theta \in \mathcal{B}_i(p^*)} dH(\theta)}.$$

Thus we can rewrite (30) as

$$\begin{aligned}
S_I(x_k) &= \frac{1}{x_{k+1} - x_k} \left[\frac{\int_{\theta \in \mathcal{B}_{k+1}(p^*)} C(\theta; x_{k+1}) dH(\theta)}{\int_{\theta \in \mathcal{B}_{k+1}(p^*)} dH(\theta)} - \frac{\int_{\theta \in \mathcal{B}_k(p^*)} C(\theta; x_{k+1}) dH(\theta)}{\int_{\theta \in \mathcal{B}_k(p^*)} dH(\theta)} \right] \\
&= \frac{1}{x_{k+1} - x_k} \left[x_{k+1} \left(\frac{\int_{\theta \in \mathcal{B}_{k+1}(p^*)} c(m) dH(\theta)}{\int_{\theta \in \mathcal{B}_{k+1}(p^*)} dH(\theta)} - \frac{\int_{\theta \in \mathcal{B}_k(p^*)} c(m) dH(\theta)}{\int_{\theta \in \mathcal{B}_k(p^*)} dH(\theta)} \right) \right] \\
&= \frac{x_{k+1}}{x_{k+1} - x_k} \left\{ \mathbb{E} \left[c(M) \mid (m, \lambda) \in \mathcal{B}_{k+1}(p^*) \right] - \mathbb{E} \left[c(M) \mid (m, \lambda) \in \mathcal{B}_k(p^*) \right] \right\},
\end{aligned}$$

where the second equality follows from our assumption that $C(\theta; x_i) = x_i \cdot c(m)$. Proposition 8 thus implies that in our setting the intensive margin selection coefficient $S_I(x_k)$ defined in Azevedo and Gottlieb (2016) is always positive. Indeed, the examples in Azevedo and Gottlieb (2016) where they find changing signs of $S_I(\cdot)$ at different levels of x feature heterogeneity in ex post moral hazard, which may cause a violation of Assumption 1.

8.2 Monopolistic Insurance Market: Endogenous Contract

Now we consider a monopolistic insurance firm that chooses premium p^m and insurance coverage x^m to:⁴¹

$$\max_{\{p, x\}} \pi(p, x) \equiv \int_{\theta \in \mathcal{B}(p, x)} (p - xm) dH(m, \lambda) \quad (31)$$

where $\mathcal{B}(p, x)$ is defined as

$$\mathcal{B}(p, x) \equiv \{\theta : v(\theta; x) - p \geq 0\}. \quad (32)$$

For tractability, we employ the expression of WTP derived from the CARA-normal specification in Example 2, that is, $v(\theta; x) = xm + x(2 - x)k\lambda$. For this class of examples, we will use $\pi(p, x; k)$ and $\mathcal{B}(p, x; k)$ to indicate that both are related to the parameter $k \equiv \sigma^2/2$ as defined in (17), which measures the relative importance of risk aversion as a determinant of the consumer's WTP for insurance. The following lemma can be implied immediately from the linearity of consumers' WTP.

Lemma 6 *Suppose $v(\theta; x) = xm + x(2 - x)k\lambda$, then $\pi(p, x; k) = x\pi(p/x, 1; (2 - x)k)$.*

The proof follows directly from the fact that $\mathcal{B}(p, x; k) = \mathcal{B}(p/x, 1; (2 - x)k)$ and is omitted for brevity. Lemma 6 uncovers the trade-off between x and the degree of adverse selection in an intuitive way: holding fixed the per-unit price of the insurance (i.e., p/x), increasing insurance coverage x will directly increase the revenue received from each consumer that purchases insurance, at the cost of yielding steeper iso-WTP curves in the (m, λ) space, which indicates more severe adverse selection.

⁴¹Same as in Veiga and Weyl (2016), we assume the monopolist offers a single contract instead of a menu of contracts.

The Role of Preferences The next two propositions report results that are parallel to those in Proposition 4 and 5.

Proposition 9 *Suppose that consumers have CARA utility functions and experience normally distributed risks as described in Example 2. For every $H(\cdot, \cdot)$, there exists a threshold $\hat{k}^\dagger > 0$ such that for all $k < \hat{k}^\dagger$, $E[M|\mathcal{B}(p, x; k)] > E[M]$ for all $p \in (\underline{m} + k\underline{\lambda}, \bar{m} + k\bar{\lambda})$ and $x \in (0, 1]$.*

Proof. From Lemma 6, $\pi(p, x; k) = x\pi(p/x, 1; (2-x)k)$. Notice that $(2-x)k < 2k$. Applying Proposition 4 with $\hat{k}^\dagger = \frac{1}{2}k^\dagger$ completes the proof. ■

Proposition 10 *Suppose that consumers have CARA utility functions and experience normally distributed risks as described in Example 2. If M and Λ are negatively quadrant dependent, then there exists a threshold $\hat{k}^{\dagger\dagger}$ such that negative correlation property emerges under monopoly when $k > \hat{k}^{\dagger\dagger}$.*

Proof. From Lemma 6, $\pi(p, x; k) = x\pi(p/x, 1; (2-x)k)$. Notice that $(2-x)k \geq k$. Applying Proposition 5 with $\hat{k}^{\dagger\dagger} = k^{\dagger\dagger}$ completes the proof. ■

Proposition 9 and 10 show that the results in Proposition 4 and 5 are robust to the endogenizing of insurance quality in a monopolistic market. Before we explain the results, it is useful to discuss the sources of advantageous selection. Because both the joint distribution of M and Λ and the shape of the iso-WTP curves will influence firm's cost curves, there are *two sources of advantageous selection* in a model of multidimensional private information. The first source of advantageous selection comes from the joint distribution of M and Λ as emphasized in Section 5.2.2. Intuitively, negative dependence between M and Λ favors the monopolist and decreases firm's cost of providing insurance. The second source of advantageous selection is a result of the downward sloping nature of the iso-WTP curves. Fixing the risk type, consumers of a high risk aversion type are more willing to purchase insurance relative to those of a low risk aversion type. As a result, consumers with *low* risk type and *high* risk aversion type will purchase insurance, mitigating adverse selection compared to a model of one-dimensional private information in risk where the low risk type will opt not to purchase insurance.

Fixing the joint distribution of M and Λ , different from the model of exogenous insurance quality, the monopolist is able to choose insurance coverage x to change the composition of consumers and hence mitigate adverse selection it faces. However, the impact of x in influencing selection is limited. Formally, the absolute value of the slope of the iso-WTP curve is given by $(2-x)k$, which is bounded from below by k and above by $2k$. Therefore, the effect of k in determining selection takes over as k becomes sufficiently large (respectively, small) and the result in Proposition 4 (respectively, Proposition 5) remains.

Impact of Dependence between M and Λ In this subsection we report numerical results to shed some light on the role of dependence between M and Λ on the design of the optimal contract, especially on the correlation between insurance purchase and *ex post* realization of risk.

To proceed, we assume that $M \sim U[0, 1]$, $\Lambda \sim U[0, 1]$, and the joint distribution is $H(m, \lambda; \mu) = \mu\mathcal{W}(m, \lambda) + (1 - \mu)\Pi(m, \lambda)$ as in Section 5.2.2. With slight abuse of notation, we denote the monopolist's optimal contract by $(p^m(\mu, k), x^m(\mu, k))$.

We will first briefly describe the algorithm of searching for the optimal contract in the numerical analysis.⁴² We use $\hat{p}^m(\mu, k)$ to denote the monopolist's optimal premium fixing $x = 1$. We first completely solve for $\hat{p}^m(\mu, k)$ for any k .

Suppose $k > 1$, we have already shown in (23) that the profit function is,

$$\pi(p, 1; k) = \begin{cases} p - \frac{1}{2} & \text{for } p \in (-\infty, 0] \\ (p - \frac{1}{2}) - (1 - \mu)\frac{1}{3k}p^3 & \text{for } p \in [0, 1] \\ (1 - \mu) \left[(p - \frac{1}{2}) + \frac{-3p^2 + 3p - 1}{3k} \right] + \frac{1}{2}\mu\frac{k-p}{k-1} \left[2p - \frac{k-p}{k-1} \right] & \text{for } p \in [1, k] \\ (1 - \mu) \left[\frac{(p-1)^3}{3k} - \frac{(p-1)^2}{2} + \frac{k^2}{6} \right] & \text{for } p \in [k, k+1] \\ 0 & \text{for } p \in [k+1, \infty). \end{cases}$$

From the proof of Proposition 6, the optimal premium is,

$$\hat{p}^m(\mu, k) = \frac{\mu\frac{k^2}{k-1} + (1 - \mu)(k - \frac{1}{k})}{\mu(2 + \frac{1}{k-1}) + 2(1 - \mu)(1 - \frac{1}{k})}.$$

Similarly, suppose $k = 1$, the profit function is:

$$\pi(p, 1; 1) = \begin{cases} p - \frac{1}{2} & \text{for } p \in (-\infty, 0] \\ (p - \frac{1}{2}) - (1 - \mu)\frac{1}{3}p^3 & \text{for } p \in [0, 1] \\ (1 - \mu) \left[\frac{(p-1)^3}{3} - \frac{(p-1)^2}{2} + \frac{1}{6} \right] & \text{for } p \in (1, 2] \\ 0 & \text{for } p \in [k+1, \infty). \end{cases}$$

For this special case, the profit function is discontinuous at $p = 1$. It can be verified that $\hat{p}^m(\mu, 1) = 1$; and it can be verified that positive correlation property emerges in this case.

Lastly, suppose $k < 1$, the profit function is:⁴³

$$\pi(p, 1; k) = \begin{cases} p - \frac{1}{2} & \text{for } p \in (-\infty, 0] \\ (p - \frac{1}{2}) - (1 - \mu)\frac{1}{3k}p^3 & \text{for } p \in [0, k] \\ (1 - \mu) \left[\frac{1}{6}k^2 - \frac{1}{2}(p-1)^2 \right] + \mu \left(k - \frac{1}{2} \right) \frac{(p-1)^2}{(k-1)^2} & \text{for } p \in [k, 1] \\ (1 - \mu) \left[\frac{(p-1)^3}{3k} - \frac{(p-1)^2}{2} + \frac{k^2}{6} \right] & \text{for } p \in [1, k+1] \\ 0 & \text{for } p \in [k+1, \infty). \end{cases}$$

It can be verified that the profit is increasing for $p \in [0, k]$ and decreasing for $p \in [1, k+1]$. Therefore, $\hat{p}^m(\mu, k) \in [k, 1]$. Moreover, if $\mu \leq (1/k - 1)^2$, the profit is increasing in p for $p \in [k, 1]$, indicating $\hat{p}^m(\mu, k) = 1$. If $\mu > (1/k - 1)^2$, the profit is decreasing in p for $p \in [k, 1]$, indicating

⁴²The program used in the numerical analysis is available from the authors upon request.

⁴³See Online Appendix D for details of the derivations for the case $k < 1$.

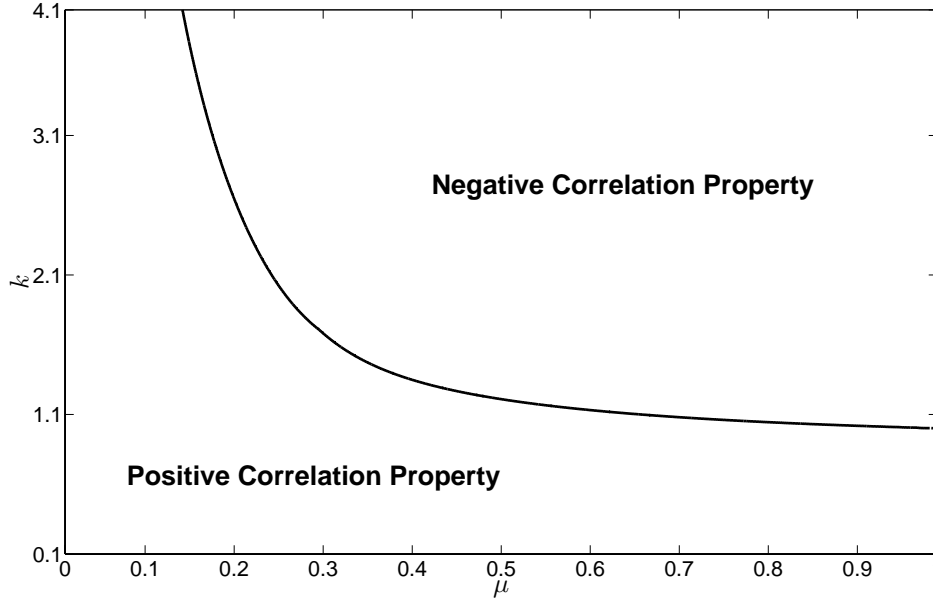


Figure 9: Positive or Negative Correlation Property under Monopoly with an Endogenous Contract

$\hat{p}^m(\mu, k) = k$. From Proposition 6, positive correlation property emerges.

The optimal coverage $x^m(\mu, k)$ solves the following one-dimensional optimization problem:

$$\max_{x \in [0,1]} x \cdot \pi \left(\hat{p}^m(\mu, (2-x)k), 1; (2-x)k \right). \quad (33)$$

After we numerically compute $x^m(\mu, k)$, the optimal premium is given by $p^m(\mu, k) = x^m(\mu, k) \cdot \hat{p}^m(\mu, [2 - x^m(\mu, k)]k)$ from Lemma 6.

Figure 9 graphically illustrates our numerical results. The solid curve is the combination of (μ, k) for which the expected risk conditional on purchase under optimal contract is equal to the unconditional expectation (i.e. the contour plot of $E[M|\mathcal{B}(p^m(\mu, k), x^m(\mu, k))] = E[M]$ in the (μ, k) space).⁴⁴ The region of (μ, k) to the right (respectively, to the left) of the solid curve depicts the combination of (μ, k) for which negative correlation property (respectively, positive correlation property) emerges under optimal contract. The first pattern to notice is that fixing the degree of negative dependence between M and Λ , negative correlation property (respectively, positive correlation property) emerges under optimal contract when k is sufficiently large (respectively, small). This confirms the results in Proposition 9 and 10. Second, the result in Proposition 6 is robust to endogenous insurance quality. Specifically, holding fixed the degree of relative importance of risk aversion, negative correlation property is more likely to appear when M and Λ are sufficiently negative dependent.⁴⁵

⁴⁴The contour plots are shown only for $(\mu, k) \in [0, 1] \times [0.1, 4.1]$.

⁴⁵Some readers may want to compare the contour plot of (μ, k) when $x = 1$ and that when x is endogenously chosen

9 Conclusion

A large empirical literature has found that the correlation between insurance purchase and ex post realization of risk is often *statistically insignificant* or *negative*, which is inconsistent with the predictions from the classic models of insurance a la Akerlof (1970), Pauly (1974) and Rothschild and Stiglitz (1976), where consumers differ only in their risk types. It is suggested that the selection based on multidimensional private information, e.g., risk type and risk preference type, may be able to reconcile the empirical findings. In this paper, we systematically investigate, under different market structures, whether selection based on multidimensional private information can result in negative correlation between insurance coverage and ex post realization of risk *in equilibrium*. We show that if the insurance market is perfectly competitive, selection based on multidimensional private information does *not* generate negative correlation property in equilibrium, unless there is a sufficiently high loading factor. If the insurance market is monopolistic, however, we show that it is possible to generate negative correlation property in equilibrium when risk type and risk preference type are sufficiently negative dependent, a notion we formalize using the concept of copula. We further show that this result generalizes to imperfectly competitive market structure, as well as when contracts are partially endogenized. We also clarify the confusions in this growing literature about the connections between some of the important concepts such as adverse/advantageous selection and positive/negative correlation property.

There are some interesting directions for future research. First, in this paper we studied the role of additional consumer heterogeneity in risk preference. It is important to model and examine whether other sources of heterogeneity, such as heterogeneity in moral hazard (Einav *et al.*, 2013) and heterogeneity in *imperfect rationality* (e.g., Fang, Keane and Silverman, 2008), will lead to different conclusions on the emergence of positive or negative correlation property. As we pointed out in Footnote 11, introducing moral hazard is likely to lead to the violations of Assumption 1. Heterogeneity in imperfect rationality will call for a plausible behavioral model of consumers' insurance purchase decisions. Second, in our paper we have identified the potential role of loading factors in possibly affecting the equilibrium of the insurance market under different market structures. In particular, Proposition 2 shows that a sufficiently small loading factor is sufficient to ensure that positive correlation property always holds under a competitive insurance market regardless of the dependence structure of the multidimensional heterogeneity. Loading factors can also drive a wedge between the WTP for insurance and marginal cost (inclusive of loading factors) of providing coverage. Further investigations, empirically about the magnitude of loading factors, and theoretically about how loading factors – potentially heterogeneous among insurance firms – may impact the equilibrium of the insurance market, are also an important avenue for future research.

by the monopolist. Simulation shows that endogenizing x slightly shifts the contour plot to the left and enlarges the region of negative correlation property under optimal contract. This result is intuitive: the monopolist can better take advantage of the negative dependence of the joint distribution if it is allowed to design x . Therefore, negative correlation property is more likely to emerge under optimal contract.

References

- [1] Akerlof, George A. (1970). "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics*, 84(3), 488-500.
- [2] Arrow, Kenneth J. (1963). "Uncertainty and the Welfare Economics of Medical Care." *The American Economic Review*, 53(5), 941-973.
- [3] Azevedo, Eduardo M. and Daniel Gottlieb (2016). "Perfect Competition in Markets with Adverse Selection." Forthcoming, *Econometrica*.
- [4] Bromek, Tadeusz and Elzbieta Pleszczyńska (1991). "Evaluation of Stochastic Dependence," in *Statistical Inference: Theory and Practice*, ed. Tadeusz Bromek and Elzbieta Pleszczyńska p. 106-136. Kluwer Academic Publishers.
- [5] Cawley, John and Tomas Philipson (1999). "An Empirical Examination of Information Barriers to Trade in Insurance." *American Economic Review*, 89(4), 827-846.
- [6] Chetty, Raj and Amy Finkelstein (2013). "Social Insurance: Connecting Theory to Data," in *Handbook of Public Economics*, Volume 5, edited by A. Auerbach, R. Chetty, M. Feldstein, and E. Saez, Elsevier, 111-193.
- [7] Chiappori, Pierre-André and Bernard Salanié (2000). "Testing for Asymmetric Information in Insurance Markets." *Journal of Political Economy*, 108(1), 56-78.
- [8] Chiappori, Pierre-André and Jullien, Bruno and Salanié, Bernard and Francois Salanié (2006). "Asymmetric Information in Insurance: General Testable Implications." *The RAND Journal of Economics*, 37(4), 783-798.
- [9] Cohen, Alma (2005). "Asymmetric Information and Learning: Evidence from the Automobile Insurance Market." *Review of Economics and Statistics*, 87(2), 197-207.
- [10] de Meza, David, and David C. Webb (2001). "Advantageous Selection in Insurance Markets." *The RAND Journal of Economics*, 32 (Summer), 249-262.
- [11] de Meza, David and David C. Webb (2016). "False Diagnoses: Pitfalls of Testing for Asymmetric Information in Insurance Markets." Forthcoming, *Economic Journal*, doi: 10.1111/eoj.12393.
- [12] Dionne, Georges and Gouriéroux, Christian and Charles Vanasse (2001). "Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment." *Journal of Political Economy*, 109(2), 444-453.
- [13] Einav, Liran and Amy Finkelstein (2011). "Selection in Insurance Markets: Theory and Empirics in Pictures." *The Journal of Economic Perspectives*, 25(1), 115.

- [14] Einav, Liran, Amy Finkelstein and Mark R. Cullen (2010). “Estimating Welfare in Insurance Markets Using Variation in Prices.” *The Quarterly Journal of Economics*, 125(3), 877-921.
- [15] Einav, Liran, Amy Finkelstein and Jonathan Levin (2010). “Beyond Testing: Empirical Models of Insurance Markets.” *Annual Review of Economics*, Vol. 2, 311-336.
- [16] Einav, Liran, Amy Finkelstein, Stephen P. Ryan, Paul Schrimpf, and Mark R. Cullen (2013). “Selection on Moral Hazard in Health Insurance.” *American Economic Review*, 103 (1), 178-219.
- [17] Fang, Hanming, Michael P. Keane and Dan Silverman (2008). “Sources of Advantageous Selection: Evidence from the Medigap Insurance Market.” *Journal of Political Economy*, Vol. 116, No. 2, 303-350.
- [18] Finkelstein, Amy and James Poterba (2004). “Adverse Selection in Insurance Markets: Policyholder Evidence from the UK Annuity Market.” *Journal of Political Economy*, 112(1), 183-208.
- [19] Finkelstein, Amy and Kathleen McGarry (2006). “Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market.” *American Economic Review*, 96(4), 938-958.
- [20] Fisher, James C.D. and Asaf Plan (2015). “Competition for an Imperfectly Informed Consumer.” Working Paper, University of Arizona.
- [21] He, Daifeng (2009). “The Life Insurance Market: Asymmetric Information Revisited.” *Journal of Public Economics*, Vol 93: 1090-1097.
- [22] Hemenway, David (1990). “Propitious Selection.” *The Quarterly Journal of Economics*, 105(4), 1063-1069.
- [23] Lester, Benjamin, Ali Shourideh, Venky Venkateswaran and Ariel Zetlin-Jones (2016). “Screening and Adverse Selection in Frictional Markets.” Working Paper, Research Department, Federal Reserve Bank of Philadelphia.
- [24] Mahoney, Neale and E. Glen Weyl (2016). “Imperfect Competition in Selection Markets.” Forthcoming, *Review of Economics and Statistics*.
- [25] Nelsen, Roger B. (2006). *An Introduction to Copulas*. 2nd edition (New York: Springer-Verlag).
- [26] Pauly, Mark V (1974). “Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection.” *The Quarterly Journal of Economics*, Vol. 88 (February), 44-62.
- [27] Puelz, Robert and Arthur Snow (1994). “Evidence on Adverse Selection: Equilibrium Signaling and Cross-Subsidization in the Insurance Market.” *Journal of Political Economy*, 102 (April), 236-257.

- [28] Rothschild, Michael and Joseph Stiglitz (1976). “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information.” *The Quarterly Journal of Economics*, 90(4), 629-649.
- [29] Veiga, André and E. Glen Weyl (2016). “Product Design in Selection Markets.” *Quarterly Journal of Economics*, 126(2), 1007-1056.
- [30] Weyl, E. Glen and André Veiga (2014). “The Leibniz Rule for Multidimensional Heterogeneity.” Working Paper, University of Chicago.
- [31] Wilson, Charles (1977). “A Model of Insurance Markets with Incomplete Information.” *Journal of Economic Theory*, 16(2), 167-207.

Online Appendix

In this online appendix, we collect the materials omitted from the main text of the paper. The appendices are ordered according to where they are first referenced in the main text. In Appendix A we show that Assumption 1-2 are satisfied in the binary state model of insurance in Example 1; in Appendix B we provide the details of the derivation of the profit function, the demand curve, and the cost curves in Proposition 6; in Appendix C we prove that the profit function of the example in Section 5.2.2 for $k > 1$ satisfies Assumption 3; in Appendix D we derive the profit function used in the numerical analysis in Section 8.2 for $k < 1$.

A Assumptions 1-2 are Satisfied in Example 1

First, we show that Assumption 2 is satisfied for binary states with strictly concave utility function. Suppose to the contrary that $v(\theta; x) \leq x \cdot c(m) = xm\omega$, then

$$m \cdot u(y - v - (1 - x)\omega) + (1 - m) \cdot u(y - v) \geq m \cdot u(y - \omega + (1 - m)x\omega) + (1 - m) \cdot u(y - xm\omega).$$

For notational convenience, denote y , $y - xm\omega$, $y - \omega + (1 - m)x\omega$, and $y - l$ as w_1 , w_2 , w_3 and w_4 respectively. It can be verified that $w_1 > w_2 > w_3 > w_4$ for $x > 0$, $m \in (0, 1)$ and $mw_4 + (1 - m)w_1 = mw_3 + (1 - m)w_2$. It remains to show that $m \cdot u(w_3) + (1 - m) \cdot u(w_2) > m \cdot u(w_4) + (1 - m) \cdot u(w_1)$. In fact,

$$\begin{aligned} & m \cdot u(w_3) + (1 - m) \cdot u(w_2) \\ & > m \left[\frac{w_1 - w_3}{w_1 - w_4} u(w_4) + \frac{w_3 - w_4}{w_1 - w_4} u(w_1) \right] + (1 - m) \left[\frac{w_1 - w_2}{w_1 - w_4} u(w_4) + \frac{w_2 - w_4}{w_1 - w_4} u(w_1) \right] \\ & = m \cdot u(w_4) + (1 - m) \cdot u(w_1), \end{aligned}$$

where the inequality follows from the concavity of $u(\cdot)$ and the equality follows from the fact that $mw_4 + (1 - m)w_1 = mw_3 + (1 - m)w_2$.

Next, we show that $\partial v / \partial x > 0$ and $\partial v / \partial m > 0$. Consumer's indifference condition can be rewritten as,

$$\mathcal{L}(v, x, m) := m \cdot u(y - v - (1 - x)\omega) + (1 - m) \cdot u(y - v) - m \cdot u(y - \omega) - (1 - m) \cdot u(y) = 0.$$

The implicit function theorem implies,

$$\frac{\partial v}{\partial x} = - \frac{\partial \mathcal{L} / \partial x}{\partial \mathcal{L} / \partial v} = \frac{m\omega \cdot u'(y - v - (1 - x)\omega)}{m \cdot u'(y - v - (1 - x)\omega) + (1 - m) \cdot u'(y - v)} > 0,$$

and

$$\begin{aligned}\frac{\partial v}{\partial m} &= -\frac{\partial \mathcal{L}/\partial m}{\partial \mathcal{L}/\partial v} = \frac{u(y-v-(1-x)\omega) - u(y-v) - u(y-\omega) + u(y)}{m \cdot u'(y-v-(1-x)\omega) + (1-m) \cdot u'(y-v)} \\ &= \frac{1}{m} \times \frac{u(y) - u(y-v)}{m \cdot u'(y-v-(1-x)\omega) + (1-m) \cdot u'(y-v)} > 0,\end{aligned}$$

where the last equality follows from the rearrangement of consumers' indifference condition.

Last, we show that $\partial v/\partial \lambda > 0$ holds for CRRA and CARA utility function when $x = 1$. For CARA utility function (i.e., $u(c; \lambda) = -e^{-\lambda c}$), consumer's indifference condition can be simplified as,

$$e^{\lambda v} = me^{\lambda \omega} + (1-m).$$

Carrying out the algebra and solving for v yields,

$$v = \frac{\ln [me^{\lambda \omega} + (1-m)]}{\lambda}.$$

Therefore,

$$\frac{\partial v}{\partial \lambda} = \left[\frac{\lambda \omega me^{\lambda \omega}}{me^{\lambda \omega} + (1-m)} - \ln [me^{\lambda \omega} + (1-m)] \right] / \lambda^2,$$

Denote $\lambda \omega$ by t . It remains to prove that,

$$b(t) := mte^t - [me^t + (1-m)] \times \ln [me^t + (1-m)] > 0, \text{ for } t > 0.$$

Notice that

$$b'(t) = me^t \left[t - \ln [me^t + (1-m)] \right] > me^t \left[t - \ln [me^t + (1-m)e^t] \right] = 0,$$

and $b(0) = 0$. Therefore, $b(t) > 0$ for $t > 0$.

For CRRA utility function (i.e. $u(c) = [c^{1-\lambda} - 1]/(1-\lambda)$), consumer's indifference condition can be derived as,

$$v = y - \left[(1-m)y^{1-\lambda} + m \cdot (y-\omega)^{1-\lambda} \right]^{\frac{1}{1-\lambda}}.$$

Define $\hat{v} := v/y$ and $\hat{\omega} := \omega/y$, then we have

$$\hat{v} = 1 - \left[(1-m) + m \cdot (1-\hat{\omega})^{1-\lambda} \right]^{\frac{1}{1-\lambda}}.$$

It is clear that $\partial v/\partial \lambda > 0$ is equivalent to $\partial \hat{v}/\partial \lambda > 0$, which can be simplified as,

$$\ln \left[(1-m) + m(1-\hat{\omega})^{1-\lambda} \right] \times \left[(1-m) + m(1-\hat{\omega})^{1-\lambda} \right] - m \ln \left[(1-\hat{\omega})^{1-\lambda} \right] (1-\hat{\omega})^{1-\lambda} < 0.$$

For notational convenience, denote $(1 - \hat{w})^{1-\lambda}$ by \hat{t} . It follows directly that $\hat{t} > 0$. Define

$$\hat{b}(\hat{t}) := \ln \left[(1 - m) + m\hat{t} \right] \times \left[(1 - m) + m\hat{t} \right] - m\hat{t} \ln \hat{t}.$$

It remains to show that $\hat{b}(\hat{t}) \leq 0$ for $\hat{t} > 0$. Taking derivative of $\hat{b}(\hat{t})$ with respect to \hat{t} yields,

$$\hat{b}'(\hat{t}) := m \times \ln \left[\frac{1 - m + m\hat{t}}{\hat{t}} \right].$$

Notice that $\hat{b}'(\hat{t}) \geq 0$ is equivalent to $\hat{t} \leq 1$. Therefore, $\hat{b}(\hat{t})$ is maximized at $\hat{t} = 1$, which in turn implies $\hat{b}(\hat{t}) \leq \hat{b}(1) = 0$. This completes the proof.

B Derivation of the Profit Function, the Demand Curve and the Cost Curves in Proposition 6

Profit function. For $p \leq 0$, all consumers purchase insurance and firm's expected profit is $\pi(p) = p - E[M] = p - \frac{1}{2}$. For $p \geq k + 1 \equiv \bar{m} + k\bar{\lambda}$, the price exceeds the highest WTP and no consumers purchase insurance. Therefore, $\pi(p) = 0$.

For $p \in [0, 1]$, the expected profit is,

$$\begin{aligned} \pi(p; \mu) &= \int_{\theta \in \mathcal{B}(p) \equiv \{\theta: v(\theta; x) \geq p\}} [p - xc(m)] dH(m, \lambda) \\ &= p - E[c(M)] - \left[\mu \int_{\theta \in \mathcal{N}\mathcal{B}(p)} [p - c(m)] d\mathcal{W}(m, \lambda) + (1 - \mu) \int_{\theta \in \mathcal{N}\mathcal{B}(p)} [p - c(m)] d\Pi(m, \lambda) \right] \\ &= \left(p - \frac{1}{2}\right) - (1 - \mu) \int_0^p \int_0^{\frac{p-m}{k}} (p - m) d\lambda dm \\ &= \left(p - \frac{1}{2}\right) - (1 - \mu) \frac{1}{k} \int_0^p (p - m)^2 dm = \left(p - \frac{1}{2}\right) - (1 - \mu) \frac{1}{3k} p^3. \end{aligned}$$

For $p \in [1, k]$, the expected profit is,

$$\begin{aligned} \pi(p; \mu) &= \mu \int_{\theta \in \mathcal{B}(p)} [p - m] d\mathcal{W}(m, \lambda) + (1 - \mu) \int_{\theta \in \mathcal{B}(p)} [p - m] d\Pi(m, \lambda) \\ &= (1 - \mu) \int_0^1 \int_{\frac{p-m}{k}}^1 [p - m] d\lambda dm + \mu \int_0^{\frac{k-p}{k-1}} [p - m] dm \\ &= (1 - \mu) \left[\left(p - \frac{1}{2}\right) + \frac{-3p^2 + 3p - 1}{3k} \right] + \frac{1}{2} \mu \frac{k-p}{k-1} \left[2p - \frac{k-p}{k-1} \right]. \end{aligned}$$

For $p \in [k, k + 1]$, the expected profit is,

$$\begin{aligned}
\pi(p; \mu) &= \mu \int_{\theta \in \mathcal{B}(p)} [p - m] d\mathcal{W}(m, \lambda) + (1 - \mu) \int_{\theta \in \mathcal{B}(p)} [p - m] d\Pi(m, \lambda) \\
&= (1 - \mu) \int_{p-k}^1 \int_{\frac{p-m}{k}}^1 [p - m] d\lambda dm \\
&= (1 - \mu) \left[\frac{(p-1)^3}{3k} - \frac{(p-1)^2}{2} + \frac{k^2}{6} \right].
\end{aligned}$$

To summarize,

$$\pi(p; \mu) = \begin{cases} p - \frac{1}{2} & \text{for } p \in (-\infty, 0] \\ (1 - \mu) \left[(p - \frac{1}{2}) - (1 - \mu) \frac{1}{3k} p^3 \right] & \text{for } p \in [0, 1] \\ (1 - \mu) \left[(p - \frac{1}{2}) + \frac{-3p^2 + 3p - 1}{3k} \right] + \frac{1}{2} \mu \frac{k-p}{k-1} \left[2p - \frac{k-p}{k-1} \right] & \text{for } p \in [1, k] \\ (1 - \mu) \left[\frac{(p-1)^3}{3k} - \frac{(p-1)^2}{2} + \frac{k^2}{6} \right] & \text{for } p \in [k, k + 1] \\ 0 & \text{for } p \in [k + 1, \infty). \end{cases}$$

Demand curve, total cost curve, and average cost curve. The derivation of the demand curve and the total cost curve is similar to the derivation of the profit function and is omitted. The total cost curve is given by,

$$TC(p; \mu) = \begin{cases} \frac{1}{2} & \text{for } p \in (-\infty, 0] \\ (1 - \mu) \left(\frac{1}{2} - \frac{1}{6k} p^3 \right) + \frac{1}{2} \mu & \text{for } p \in [0, 1] \\ \frac{1}{2} \mu \left(\frac{k-p}{k-1} \right)^2 + (1 - \mu) \left[\frac{1}{2} + \frac{1}{k} \left(\frac{1}{3} - \frac{1}{2} p \right) \right] & \text{for } p \in [1, k] \\ (1 - \mu) \left\{ \frac{1}{2} \left(1 - \frac{p}{k} \right) [1 - (p - k)^2] + \frac{1}{3k} [1 - (p - k)^3] \right\} & \text{for } p \in [k, k + 1] \\ 0 & \text{for } p \in [k + 1, \infty) \end{cases},$$

and the demand curve is given by,

$$D(p; \mu) = \begin{cases} 1 & \text{for } p \in (-\infty, 0] \\ \mu + (1 - \mu) \left(1 - \frac{1}{2k} p^2 \right) & \text{for } p \in [0, 1] \\ \mu \left(\frac{k-p}{k-1} \right) + (1 - \mu) \left[1 + \frac{1}{k} \left(\frac{1}{2} - p \right) \right] & \text{for } p \in [1, k] \\ (1 - \mu) \frac{1}{2k} (1 + k - p)^2 & \text{for } p \in [k, k + 1] \\ 0 & \text{for } p \in [k + 1, \infty) \end{cases}.$$

Finally, the average cost is given by,

$$AC(p; \mu) = \mathbb{E} [c(M) | \mathcal{B}(p)] = \frac{TC(p; \mu)}{D(p; \mu)}.$$

Marginal cost curve. The marginal cost is defined as,

$$MC(p; \mu) = \mathbb{E} [M|v(\theta) = p] = \frac{\int_{\theta \in \{\theta: v(\theta; 1) = p\}} m dH(m, \lambda)}{\int_{\theta \in \{\theta: v(\theta; 1) = p\}} dH(m, \lambda)} \equiv \frac{dTC(p)}{dp} \Big/ \frac{dD(p)}{dp}.$$

Notice that, for $MC(p)$ is well-defined only when there exists demand for insurance at a price. For $\mu = 1$, the price that induces positive demand lies between 1 to k . Therefore, $MC(p; \mu)$ is given by,

$$MC(p) = \frac{k-p}{k-1} \text{ for } p \in [1, k].$$

For $\mu \in [0, 1)$, the price range associated with positive demand is $[0, k+1]$. Therefore, $MC(p; \mu)$ is given by,

$$MC(p; \mu) = \begin{cases} \frac{1}{2}p & \text{for } p \in [0, 1] \\ \left[\frac{\mu}{k-1} \frac{k-p}{k-1} + \frac{1-\mu}{2k} \right] \Big/ \left[\frac{\mu}{k-1} + \frac{1-\mu}{k} \right] & \text{for } p \in [1, k] \\ \frac{1}{2}[(p-k)+1] & \text{for } p \in [k, k+1] \end{cases}.$$

C Assumption 3 is Satisfied in the Example in Section 5.2.2 for $k > 1$

In this example, $v(\underline{m}, \underline{\lambda}); x) = 0$ and $v(\overline{m}, \overline{\lambda}); x) = k+1$. It can be verified that $\pi(p)$ is single-peaked in p for $p \in [0, k+1]$. Moreover, $p^* \in (0, 1)$, $p^m \in (1, k)$ and $\pi'(p^m) = 0$.

For $p \in [p^*, 1)$, the partial derivative with respect to p is,

$$\frac{\partial \pi(p; \mu)}{\partial p} = 1 - \frac{1-\mu}{k} p^2 \geq 1 - \frac{1-\mu}{k} > 0.$$

Therefore, $\pi(p; \mu)$ is strictly increasing in p . Moreover, it is clear that $\partial \pi(p; \mu) / \partial p$ is strictly decreasing in p . Therefore, $\frac{\partial \pi(p; \mu)}{\partial p} \Big/ \pi(p; \mu)$ is strictly decreasing in p for $p \in [p^*, 1)$.

Similarly, for $p \in (1, p^m]$, the partial derivative with respect to p is,

$$\frac{\partial \pi(p; \mu)}{\partial p} = (1-\mu) \left[1 + \frac{1}{k} - \frac{2}{k} p \right] + \mu \frac{k^2 - (2k-1)p}{(k-1)^2}.$$

It can be verified that $\pi(p; \mu)$ is strictly increasing in p , and $\partial \pi(p; \mu) / \partial p$ is strictly decreasing in p . Therefore, $\frac{\partial \pi(p; \mu)}{\partial p} \Big/ \pi(p; \mu)$ is strictly decreasing in p for $p \in (1, p^m]$.

For $p = 1$, notice that the left limit of the partial derivative with respect to p is,

$$\lim_{p \uparrow 1} \frac{\partial \pi(p; \mu)}{\partial p} = 1 - \frac{1-\mu}{k}.$$

The right limit of the partial derivative with respect to p at $p = 1$ is,

$$\lim_{p \downarrow 1} \frac{\partial \pi(p; \mu)}{\partial p} = (1-\mu) \left[1 + \frac{1}{k} - \frac{2}{k} \right] + \mu \frac{k^2 - (2k-1)}{(k-1)^2} = 1 - \frac{1-\mu}{k} = \lim_{p \uparrow 1} \frac{\partial \pi(p; \mu)}{\partial p}.$$

Hence, the derivative at $p = 1$ is well-defined, and $\pi(p; \mu)$ is log-concave in p for $p \in (p^*, p^m]$.

D Derivation of the Profit Function in the Numerical Analysis in Section 8.2 for $k < 1$

For $p \leq 0$, all consumers purchase insurance and firm's expected profit is $\pi(p) = p - E[M] = p - \frac{1}{2}$. For $p \geq k + 1 \equiv \bar{m} + k\bar{\lambda}$, the price exceeds the highest WTP and no consumers purchase insurance. Therefore, $\pi(p) = 0$.

For $p \in [0, k]$, the expected profit is,

$$\begin{aligned}
\pi(p, 1; k) &= \int_{\theta \in \mathcal{B}(p) \equiv \{\theta: v(\theta; x) \geq p\}} [p - xc(m)] dH(m, \lambda) \\
&= p - E[c(M)] - \left[\mu \int_{\theta \in \mathcal{N}\mathcal{B}(p)} [p - c(m)] d\mathcal{W}(m, \lambda) + (1 - \mu) \int_{\theta \in \mathcal{N}\mathcal{B}(p)} [p - c(m)] d\Pi(m, \lambda) \right] \\
&= \left(p - \frac{1}{2}\right) - (1 - \mu) \int_0^p \int_0^{\frac{p-m}{k}} (p - m) d\lambda dm \\
&= \left(p - \frac{1}{2}\right) - (1 - \mu) \frac{1}{k} \int_0^p (p - m)^2 dm = \left(p - \frac{1}{2}\right) - (1 - \mu) \frac{1}{3k} p^3.
\end{aligned}$$

For $p \in [k, 1]$, the expected profit is,

$$\begin{aligned}
\pi(p, 1; k) &= \mu \int_{\theta \in \mathcal{B}(p)} (p - m) d\mathcal{W}(m, \lambda) + (1 - \mu) \int_{\theta \in \mathcal{B}(p)} (p - m) d\Pi(m, \lambda) \\
&= (1 - \mu) \int_0^1 \int_{p-k\lambda}^1 (p - m) dm d\lambda + \mu \int_{\frac{k-p}{k-1}}^1 (p - m) dm \\
&= (1 - \mu) \left[\frac{1}{6} k^2 - \frac{1}{2} (p - 1)^2 \right] + \frac{1}{2} \mu \left(1 - \frac{k-p}{k-1} \right) \left(2p - 1 - \frac{k-p}{k-1} \right) \\
&= (1 - \mu) \left[\frac{1}{6} k^2 - \frac{1}{2} (p - 1)^2 \right] + \mu \left(k - \frac{1}{2} \right) \frac{(p - 1)^2}{(k - 1)^2}.
\end{aligned}$$

For $p \in [1, k + 1]$, the expected profit is,

$$\begin{aligned}
\pi(p, 1; k) &= \mu \int_{\theta \in \mathcal{B}(p)} [p - m] d\mathcal{W}(m, \lambda) + (1 - \mu) \int_{\theta \in \mathcal{B}(p)} [p - m] d\Pi(m, \lambda) \\
&= (1 - \mu) \int_{p-k}^1 \int_{\frac{p-m}{k}}^1 [p - m] d\lambda dm \\
&= (1 - \mu) \left[\frac{(p - 1)^3}{3k} - \frac{(p - 1)^2}{2} + \frac{k^2}{6} \right].
\end{aligned}$$