Estimating the Critical Parameter in Almost Stochastic Dominance from Insurance Deductibles

15 February, 2019

Abstract

Knowing how small violation individuals would accept against stochastic dominance rules is a prerequisite for applying almost stochastic dominance criteria. Different from previous results obtained by experiments, this paper estimates acceptable violation against stochastic dominance rules with 940,904 observations of real data on a deductible choice of automobile theft insurance. We find that for *all* policyholders in the sample who optimally chose a low deductible, the upper bound estimate of acceptable violation ratio is $8.198e^{-08}$ which is close to zero. On the other hand, considering *most* decision makers, such as 99% (95%) of the policyholders in the sample, who optimally chose the low deductible, the upper bound estimate of the acceptable violation ratio is 0.0399 (0.0727). Our results provide reference values of the acceptable violation ratio and justification for applying almost stochastic dominance rules.

JEL classification: D81, G22

Keywords: almost stochastic dominance; generalized almost second-degree stochastic dominance; preference parameter; automobile theft insurance; deductible

1 Introduction

Stochastic dominance (SD) rules have become main tools for ranking distributions since Rothschild and Stiglitz (1970) proposed a definition of an increase in risk in terms of a change in distribution. Based on SD rules, the literature analyzed comparative statics of an increase in risk and examined efficiency in optimization.[1] However, there exists cases where people apparently prefer one distribution to the other, which cannot be revealed by SD rules. For example, when facing two prospects–one prospect yields -1 dollar with probability 0.01 and one million with probability 0.99, and the other yields zero dollar for certainty, most people prefer the former to the latter. However, SD rules fail to rank these two prospects.

A new criterion for ranking distributions, almost stochastic dominance (ASD), is accordingly proposed by Leshno and Levy (2002) to solve the above paradox. ASD rule can shows a dominance between two distributions with crosses for each other as long as the area violation under SD rules is small enough to be accepted by *most* (but not *all*) decision makers. The preferences not considered by ASD rule are believed to be extreme and economically unrelated.

Since ASD rule was proposed to overcome the difficulty in SD rules, more and more related theories and applications have been being developed. Some researchers devoted to improve Leshno and Levy's (2002) version by proposing alternative and more general definitions (e.g., Lizyayev and Ruszczyński, 2012; Tzeng et al., 2013; Denuit et al., 2014b; Tsetlin et al., 2015).[2] Others applied the concept of ASD to develop other decision rules in almost version such as almost marginal conditional stochastic dominance (Denuit et al., 2014a) and almost expectation (Denuit et al., 2014c). In the meantime, ASD rule has also been employed to empirical studies. Some papers found that ASD rule can explain common practice in investments unanswered by SD rules (e.g., Bali et al., 2009; Levy, 2009; Bali et al., 2013)[3]. Moreover, researchers showed that investment efficient sets can be further improved by ASD rule (e.g., Levy, 2012).

[1] E.g., Levy (1992, 1998) provided a detail survey regarding SD rules.

[2] Lizyayev and Ruszczyński (2012) proposed another definition of ASD rule which has an advantage of easily implementing computation. Tzeng et al. (2013) provided another version of almost second-degree stochastic dominance (ASSD) which fixes problems of Leshno and Levy's (2002) version, and they also extended it to higher orders. ASD rule is even extended to the bivariate case in which joint distribution functions of two random variables and the utility function with two attributes are considered (Denuit et al., 2014b). Recently, Tsetlin et al. (2015) developed a general definition of ASD rule, called generalized almost stochastic dominance (GASD), which includes other versions established by previous papers as special cases.

[3] Bali et al. (2009) found that ASD rule can explain investors' preferences for stocks than bonds in long-term investment periods while Levy (2009) obtained an opposite conclusion. Bali et al. (2013) empirically showed that some hedge funds dominate stocks and bonds by ASD.

Though relevant studies of ASD have gained popularity, the question of how small violation against SD rules is allowed by *most* decision makers has not been well explored. Knowing the value of the acceptable violation ratio against SD rules, which is a parameter that defines a set of *most* decision makers' choice preferences, is a prerequisite for applying ASD rules. Given the value of the acceptable violation ratio, ASD rules can then be used to determine whether one distribution is preferred to the other by *most* decision makers. So far, all we know about its value is only based on the results obtained by experiments (e.g., Levy et al., 2010; Huang et al., 2015). However, in reality, whether the acceptable violation ratio would have a value like the experimental results has not been verified.

The purpose of this paper is to specifically estimate the acceptable violation ratio, ε_1, in generalized almost second-degree stochastic dominance (ε_1,0)-GASSD (Tsetlin et al., 2015) with *real* data. We use data on a deductible choice of automobile theft insurance contract to estimate ε_1. It is attractive to estimate ε_1 by observing the deductible choice. Whether to purchase insurance and which deductible levels would choose when purchasing insurance are decisions commonly faced in our daily lives. Insurance data, especially automobile insurance data, has the advantage of easy availability and large quantity. Our data is provided by a leading non-life insurance company in Taiwan. It covers automobile theft insurance contracts sold during year 2002 to year 2008. Our sample for estimation comprises rich observations (940,904 observations).

We estimate ε_1 by the percentage of policyholders in the sample who made an *optimal*[4] decision on choosing a 10%-deductible contract, i.e., they should purchase and actually purchased 10%-deductible contract based on (ε_1,0)-GASSD rule. If a policyholder chose the 10% deductible optimally, then he/she has a violation ratio for 20%-deductible contract to dominate 10%-deductible contract via (ε_1,0)-GASSD larger than ε_1. Otherwise, he/she should choose the 20% deductible. Therefore, for each policyholder in the sample, we first estimate the above violation ratio whose estimate is denoted by $\widehat{R}$. We then use the minimum of $\widehat{R}$ in the whole sample as an estimate of ε_1 by assuming *all* the policyholders in the sample purchased the 10%-deductible contract optimally, denoted by $\widehat{R}$(100%). Based on (ε_1,0)-GASSD rule, $\widehat{R}$(100%) could be an upper bound estimate for ε_1. Note that the minimum of $\widehat{R}$ in the whole sample

[4]In this paper, we term a decision as an *optimal* one if a decision maker made the decision in accordance with (ε_1,0)-GASSD rule.

decreases as the size of the sample increases. The estimation of ε_1 for the whole sample, especially for a large whole sample, is like searching for the acceptable violation ratio against SD rules for *all* risk averters. Thus, for our large sample, we expect that the estimate of ε_1 would be close to zero.

Indeed, ASD rules are derived for *most*, but not *all*, risk averters. As pointed out by Leshno and Levy (2002), some risk averters with pathological preferences could be economically irrelevant. Thus, we report the quantiles rather than the minimum of $\widehat{R}$ for the whole sample to exclude policyholders with extreme preferences. We propose a quantile-based estimation by allowing only $m\%\,(m < 100)$ of the policyholders in the sample made an optimal decision and find the minimal $\widehat{R}$ among the $m\%$ of the policyholders (i.e., the (100-m)th percentile of $\widehat{R}$) as an upper bound estimate of ε_1, denoted by $\widehat{R}(m\%)$. In addition, we also report a 95% confidence interval estimate for each $\widehat{R}(m\%)$ obtained by bootstrap.

Our results show that when assuming that *all* the policyholders in the sample made an optimal decision on purchasing the 10%-deductible contract based on (ε_1,0)-GASSD rule, the upper bound estimate of ε_1 is $8.198e^{-08}$ with a 95% confidence interval estimate of [$2.29e^{-16}$, 0.0026]. Our findings support the rationality that the acceptable violation ratio against SD rules for a large sample of risk averters would be close to zero. Compared with the estimates of ε_1 for *all* decision makers reported by previous works, our estimates obtained by real data with a large sample (940,904 observations) are much smaller than Levy et al.'s (2010) estimate of 0.059 obtained by lab experiments with a small sample (180 subjects).

On the other hand, our estimates of ε_1 for *most* risk averters are as follows. When assuming that 99% (95%) of the policyholders in the sample purchased the 10%-deductible contract optimally, the upper bound estimate of ε_1 is 0.0399 (0.0727) with a 95% confidence interval estimate of [0.0298, 0.0463] ([0.0635, 0.0791]). Our upper bound estimate of 0.0554 (with a 95% confidence interval of [0.0456, 0.0621]) for 97.5% of the policyholders in the sample purchased optimally is closest and slightly smaller than Levy et al.'s (2010) estimate of 0.059 for *all* subjects in the sample. Our quantile-based estimates therefore provide reasonable reference values of ε_1 for employing ASD rules.

To cross-check with the literature on estimating risk aversion indices (e.g., Gertner, 1993; Metrick, 1995; Holt and Laury, 2002; Bliss and Panigirtzoglou, 2004; Chetty, 2006; Cohen and

Einav, 2007; Andersen et al., 2008; Bollerslev et al., 2011; Bucciol and Miniaci, 2011; Brenner, 2015), by a derived relation between ε_1 and absolute risk aversion (ARA) coefficient, we further obtain upper bound estimates of the ARA coefficient on basis of our upper bound estimates of ε_1. Take the policyholder with an insured car valued at the mean NT$380,070 of the sample as an example. We find that when *all* the 10%-deductible decisions of the sample are optimal, the upper bound estimate of the ARA coefficient for such the policyholder is 0.0002 with a 95% confidence interval estimate of [$7.827e^{-05}$, 0.0005]. Compared with Cohen and Einav's (2007) ARA coefficient estimates also obtained by the data of the deductible choices in automobile insurance, our estimate is smaller than their mean estimate of 0.0067, but it is very close to and slightly smaller than their 75th percentile estimate of 0.00029.[5]

On the other hand, we find that when 99% (97.5% and 95%, respectively) of the 10%-deductible decisions of the sample are optimal, the upper bound estimate of the ARA coefficient at the mean insured car value is $4.184e^{-05}$ ($3.731e^{-05}$ and $3.349e^{-05}$, respectively) with a 95% confidence interval estimate of [$3.980e^{-05}$, $4.582e^{-05}$] ([$3.572e^{-05}$, $4.001e^{-05}$] and [$3.229e^{-05}$, $3.540e^{-05}$], respectively). All the above our estimates are smaller than Cohen and Einav's (2007) mean and 75th percentile estimates, but ours are close to and rather larger than their median estimate of $2.6e^{-05}$. Accordingly, our estimates of ARA coefficient obtained via our estimates of ε_1 are comparable to previous results.

Our paper contributes the literature on ASD in several aspects. As far as we know, this is the first paper to provide the information of the acceptable violation ratio against SD rules in (ε_1,0)-GASSD (also ε_1-AFSD) with real data, which is different from previous papers that estimated it by experimental data. We propose a quantile-based estimation which sheds light on the property of ASD and could be implemented with a large sample. Therefore, our estimates of ε_1 are closer to reality and could provide reference values of ε_1 for relevant studies. Furthermore, this is also the first paper link risk aversion intensity to ASD rules. Our ARA estimates directly linked to the acceptable violation ratio against SD rules are reasonable and comparable to the literature.

The remainders are as follows. Section 2 reviews (ε_1,0)-GASSD rule. Section 3 applies (ε_1,0)-GASSD to the decision of insurance deductibles and develops a condition for empirical

[5]We mainly compare our results with those in Cohen and Einav (2007) because both the two papers employed the data on insurance deductibles to obtain the results.

estimation. We propose a quantile-based estimation in Section 4. Section 5 describes our data. The results are presented in Section 6. Finally, Section 7 concludes the paper.

2 (ε_1,0)-GASSD

We review (ε_1,0)-GASSD (Tsetlin et al., 2015) in this section. Specifically, the definition of (ε_1, 0)-GASSD is as follows. First define utility function:
$\underline{U}_2(\varepsilon_1, 0) = \left\{ u'(x) > 0, u''(x) \le 0, u'(x) \le inf\,\{u'(x)\} \left(\frac{1}{\varepsilon_1} - 1 \right) \forall x, \varepsilon_1 \in \left(0, \frac{1}{2} \right) \right\}.$
Denote $F^{(2)}(x) = \int_{\underline{x}}^{x} F(t)\ dt$ and $G^{(2)}(x) = \int_{\underline{x}}^{x} G(t)\ dt$. Tsetlin et al. (2015) defined (ε_1, 0)-GASSD as follows:

Definition 1 *For $0 < \varepsilon_1 < \frac{1}{2}$, F dominates G by (ε_1, 0)-GASSD if and only if $F^{(2)}(\bar{x}) - G^{(2)}(\bar{x}) \le 0$ and*

$$\max_{x \in [\underline{x}, \bar{x}]} \left[F^{(2)}(x) - G^{(2)}(x) \right] \le \frac{\varepsilon_1}{1 - 2\varepsilon_1} \left[G^{(2)}(\bar{x}) - F^{(2)}(\bar{x}) \right]. \tag{1}$$

Note that $F^{(2)}(\bar{x}) \le G^{(2)}(\bar{x})$ is equivalent to $\mathrm{E}_F(x) \ge \mathrm{E}_G(x)$, where $\mathrm{E}_F(x)$ and $\mathrm{E}_G(x)$ denote the mean under F and G, respectively. Definition 1 states that F dominates G by (ε_1, 0)-GASSD means that the mean under F is larger than or equal to the mean under G and the maximum difference between $F^{(2)}$ and $G^{(2)}$ which violates SSD rule does not exceed the difference between F's mean and G's mean weighted by $\frac{\varepsilon_1}{1 - 2\varepsilon_1}$, where $\varepsilon_1 \in (0, \frac{1}{2})$.

Tsetlin et al. (2015) then showed (ε_1,0)-GASSD rule as the following theorem:

Theorem 1 *F dominates G by (ε_1, 0)-GASSD if and only if for all $u \in \underline{U}_2(\varepsilon_1, 0)$, $E_F(u) \ge E_G(u)$.*

Since we consider a decision of insurance deductibles, (ε_1,0)-GASSD rule (Tsetlin et al., 2015) under which decision makers are assumed to be risk averse, is adequate for being the studied decision rule.[6]

[6] Note that if $F(x)$ crosses $G(x)$ only once and from the bottom, then (ε_1,0)-GASSD rule happens to be ε_1-almost first-degree stochastic dominance (ε_1-AFSD) rule proposed by Leshno and Levy (2002).

3 A Choice of Insurance Deductibles

In this section, we apply (ε_1,0)-GASSD rule to the choice of deductible level when purchasing insurance and derive a condition for empirically estimation of ε_1. Note that (ε_1, 0)-GASSD rule assumes that decision makers are risk averse in accordance with the choice of insurance deductibles we discuss in this paper.

The settings are as follows. A risk-averse individual with utility function $u \in \underline{U}_2(\varepsilon_1, 0)$ has an initial wealth w and a potential loss L with a probability of occurrence π. The individual would like to purchase insurance to transfer his/her risk to an insurer. There are two deductible levels: 10% and 20%. In other words, the individual faces a menu with premiums and coverage amounts provided by the insurer: ($P^{10\%}Q^{10\%}$, $Q^{10\%}$) and ($P^{20\%}Q^{20\%}$, $Q^{20\%}$), where P^k denotes a premium rate for per unit of coverage amount and Q^k denotes a coverage amount in the 10%- or 20%-deductible contract, $k = 10\%$ or 20%. Note that under the 10%-deductible contract, the individual receives a high coverage ($Q^{10\%} = 0.9L$) while under the 20%-deductible contract, the individual receives a low coverage ($Q^{20\%} = 0.8L$).

The individual has to decide which deductible level to purchase. He/She makes a choice by ranking the cumulative distributions of final wealth x under the 10%-deductible contract (denoted by $B(x)^{10\%}$) and the 20%-deductible contract (denoted by $B(x)^{20\%}$) based on (ε_1, 0)-GASSD rule. If choosing the 10%-deductible contract, the individual has probability π to obtain $w - L + (1 - P^{10\%})Q^{10\%}$ and probability $1 - \pi$ to obtain $w - P^{10\%}Q^{10\%}$. If choosing the 20%-deductible contract, he/she has probability π to obtain $w - L + (1 - P^{20\%})Q^{20\%}$ and probability $1 - \pi$ to obtain $w - P^{20\%}Q^{20\%}$. Under the above settings, the relation between $B(x)^{10\%}$ and $B(x)^{20\%}$ is presented in Figure 1.

[Insert Figure 1 here]

Figure 1 shows that the deductible choice is a single-cross case with $B(x)^{10\%}$ lying below $B(x)^{20\%}$ before the cross. Denote C the area where $B(x)^{20\%}$ is higher than $B(x)^{10\%}$ and D the area where $B(x)^{10\%}$ is higher than $B(x)^{20\%}$. Further, let $B^{(2)}(x)^{20\%}$ and $B^{(2)}(x)^{10\%}$ denote $\int_{\underline{x}}^{x} B(t)^{20\%}\,dt$ and $\int_{\underline{x}}^{x} B(t)^{10\%}\,dt$, respectively.

Since $B(x)^{20\%}$ crosses $B(x)^{10\%}$ only once, from Figure 1, $max\left[B^{(2)}(x)^{20\%} - B^{(2)}(x)^{10\%}\right] = C$ and $B^{(2)}(\bar{x})^{10\%} - B^{(2)}(\bar{x})^{20\%} = D - C$. Thus, according to Definition 1, $B(x)^{20\%}$ dominates

$B(x)^{10\%}$ by $(\varepsilon_1,\,0)$-GASSD for all $u \in \underline{U}_2(\varepsilon_1, 0)$ if and only if

$$C \leq \frac{\varepsilon_1}{1-2\varepsilon_1}\,(D-C)\,, \tag{2}$$

where $D - C \geq 0$. Let $R = \dfrac{C}{C+D}$ denote a violation ratio for $B(x)^{20\%}$ dominating $B(x)^{10\%}$. Equation (2) can be accordingly rewritten as

$$R = \frac{C}{C+D} \leq \varepsilon_1, \tag{3}$$

where $C = \pi\left[(1-P^{10\%})Q^{10\%} - (1-P^{20\%})Q^{20\%}\right]$ and $D = (1-\pi)\left[P^{10\%}Q^{10\%} - P^{20\%}Q^{20\%}\right]$.[7] R can be calculated as long as we have π, $P^{10\%}$, $P^{20\%}$, $Q^{10\%}$, and $Q^{20\%}$.

Note that, in this paper, we extract the information of ε_1 from the data on the 10%-deductible contracts rather than the 20%-deductible contracts.[8] Based on Equation (3) and Theorem 1, if we observe an individual i with $u \in \underline{U}_2(\varepsilon_1, 0)$ who prefers 10%-deductible contract to 20%-deductible contract, then R_i should satisfy

$$R_i > \varepsilon_1. \tag{4}$$

We then use Condition (4) to obtain estimates of ε_1.

4 Empirical Estimation

For *all* the policyholders in the sample to make an optimal choice on purchasing the 10%-deductible contract, i.e., Condition (4) holds for *all* the policyholders in the sample, the minimum of R_i would be an upper bound for ε_1, i.e.,

$$\min_i R_i > \varepsilon_1,$$

[7]The decision rule shown as Equation (3) is the same as that if $B(x)^{20\%}$ dominates $B(x)^{10\%}$ by ε_1-AFSD defined by Leshno and Levy (2002). We provide the formal proof as the proof of Result 1 in Appendix A.

[8]The same methodology could not be employed to estimate ε_1 if we focus on the 20%-deductible contracts instead. It is also shown in the proof of Result 2 in Appendix A that, $(\varepsilon_1,\,0)$-GASSD rule under which all $u \in \underline{U}_2(\varepsilon_1, 0)$ prefer the 10%-deductible contract to the 20%-deductible contract is exactly SSD rule. Thus, we could not infer ε_1 from the sample in which policyholders purchased the 20%-deductible contract.

where R_i denotes individual i's R. Thus, an upper bound estimate of ε_1 would be the minimum of the estimates of R_i in the whole sample, i.e.,

$$\widehat{R}(100\%) = \min_i \widehat{R}_i,$$

where $\widehat{R}(100\%)$ denotes the upper bound estimate of ε_1 when 100% of the policyholders in the sample who purchased the 10%-deductible contract made an optimal decision on basis of (ε_1, 0)-GASSD rule.

Note that typically, when the size of the sample gets larger and larger, we could push $\widehat{R}(100\%)$ to become smaller and smaller. Therefore, $\widehat{R}(100\%)$ estimated with a large whole sample would be expected to be close to zero and consistent with SD rules. On the other hand, when assuming that only a certain percentage (say $m\%$) of the policyholders in the sample made an optimal decision according to (ε_1, 0)-GASSD rule, which is called *correct rate* $m\%$ *based on (*ε_1*, 0)-GASSD rule* throughout the paper, $\widehat{R}(m\%)$ would be expected to be away from zero and consistent with ASD rules.

On basis of the above observation, we propose the following quantile-based estimation for the upper bound of ε_1, which would provide more meaningful information for empirical practice. Suppose that $R_i > R_j$. We know that if

$$R_j > \varepsilon_1,$$

then

$$R_i > \varepsilon_1.$$

In other words, given that $R_i > R_j$, if policyholder j made an optimal decision on purchasing the 10%-deductible contract, then policyholder i would choose to purchase the 10%-deductible one optimally. Thus, an $m\%$ upper bound estimate of ε_1 is obtained as

$$\widehat{R}(m\%) = \min_{i \in \mathsf{M}} \widehat{R}_i,$$

where M denotes the set of $m\%$ of the policyholders in the sample. $\widehat{R}(m\%)$ is the value of $\widehat{R}$

at the (100-m)th percentile of the sample distribution conditional on all policyholders in the sample who purchased the 10%-deductible contract. This is also an upper bound estimate of ε_1.

The estimation of ε_1 requires the computation of $\widehat{R}$. To obtain $\widehat{R}$, we need the the probability of theft π, the premium rates $P^{10\%}$ and $P^{20\%}$, and the potential loss L for each policyholder. Below we explain how we obtain these information from our data.

Using the sample of the 10%-deductible contract, we compute $P^{10\%}$ by the premium actually paid by the policyholders divided by the coverage amount (i.e., the insured car value covered by the contract). However, we do not observe $P^{20\%}$ for these policyholders. To obtain an estimate of $P^{20\%}$, we construct a generic premium rate by using the insurance company's price schedule denoted by $\widehat{P}_i^{20\%}$. According to this price schedule, the generic premium rate is a function of characteristics of both the insured (e.g., age, gender, marital status, and claim record) and the insured vehicle (e.g., age, brand, value).

Note that a policyholder may obtain a discount for his/her premium rate due to some characteristics observed by the insurance company (or the insurance agent) but not available in the data. To account for this, we compute the discount rate that the policyholder actually obtained (d_i) by comparing his/her acutally-paid 10%-deductible premium rate $P_i^{10\%}$ with the one generated by the insurance company's price schedule (denoted by $\widehat{P}_i^{10\%}$), i.e., $d_i = P_i^{10\%}/\widehat{P}_i^{10\%}$. Since d_i is irrelevant to the policyholder's choice of contract, we can further estimate a 20%-deductible premium rate for individual i adjusted by d_i (denoted by $\widetilde{P}_i^{20\%}$) as follows.

$$\widetilde{P}_i^{20\%} = d_i \times \widehat{P}_i^{20\%}.$$

For the theft probability π, we estimate the probability for an insured vehicle to be stolen, denoted by $\widehat{\pi}$, by the following probit model:

$$\mathsf{stolen}_i = \begin{cases} 1 & \text{if } \boldsymbol{x}_i\boldsymbol{\beta} + e_i > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

where stolen_i is a binary indicator of whether insured i's vehicle was reported to the insurance company for being stolen, the error term e_i is assumed to follow a standard univariate normal distribution, $\boldsymbol{x}_i$ is a vector of characteristics of insured i and his/her vehicle, and $\boldsymbol{\beta}$ is a vector of parameters to be estimated. We then use the estimated coefficients $\widehat{\boldsymbol{\beta}}$ to compute a predicted theft probability of the insured car for policyholder i, denoted by $\widehat{\pi}_i$. That is, $\widehat{\pi}_i = \Phi(\boldsymbol{x}_i\widehat{\boldsymbol{\beta}})$, where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

According to the definition of R in Equation (3), our estimate of R for policyholder i is obtained as

$$\widehat{R}_i = \frac{1}{1 + \dfrac{\dfrac{1-\widehat{\pi}_i}{\widehat{\pi}_i}}{\dfrac{0.1}{0.9P_i^{10\%} - 0.8\widetilde{P}_i^{20\%}} - 1}}. \tag{6}$$

For statistical inference, we obtain confidence intervals for $\widehat{R}(m\%)$ by bootstrap. To obtain bootstrapped confidence intervals, we draw 1,000 bootstrapped samples. Each bootstrapped sample is drawn from our original research sample with replacement (i.e., an observation in the original sample may appear more than once in the bootstrapped sample) and is of the same size. We obtain 1,000 bootstrapped $\widehat{R}(m\%)$ by using these 1,000 bootstrapped samples. The upper and lower bounds of a bootstrapped 95% confidence interval are the values of these bootstrapped $\widehat{R}(m\%)$ at the 2.5th and 97.5th percentiles, respectively.

5 Data

Our research data is obtained from one large non-life insurance company in Taiwan. Its market share in automobile insurance market in Taiwan is over 20%. In addition to the compulsory automobile liability insurance, several kinds of voluntary automobile insurance contracts are sold in the market, such as: automobile liability insurance for bodily injury and property damage, automobile insurance for car damage, and automobile theft insurance.

Among the above contracts, it is easier for policyholders with automobile theft insurance to evaluate their final wealth distributions in the occurrence of automobile theft. First, automobile theft insurance contract possesses the characteristic of a valued contract. A valued contract means that the claim amount paid by the insurer in the event of total loss is agreed upon between the insurer and the policyholder when the policy is sold. In other words, under automobile theft insurance contract, the claim amount is a known constant. Furthermore, automobile theft insurance contract provides indemnity for only total theft loss but not for partial theft loss. Due to these advantages, we focus on investigating automobile theft insurance.

We collect data on automobile theft insurance contracts sold during year 2002 to year 2008 with total 1,045,487 observations. There are two kinds of deductible level in the contracts, 10%-deductible and 20% deductible. In the data, 93.5% of the policyholders who chose 10%-deductible contracts and only 6.5% of the policyholders who chose the 20%-deductible contracts. As pointed out in the previous section that we infer ε_1 by the policyholders who purchased the 10%-deductible contract rather than the 20%-deductible contract, we exclude the policyholders who purchased the 20%-deductible contract. After further excluding some observations with missing information, we have a large research sample consisting of 940,904 observations.

From the data, we observe the individual level information for each insurance contract, including demographic characteristics of the policyholder (gender, age, and marital status), characteristics of the insured vehicles (market value when it was insured, age, engine size, size, brand, whether it is registered in a city, and registered area), information regarding the contract (premiums and issue year) and claim records during the contract period. All the definitions of these information are provided in Appendix B.

As shown by Equation (6), when estimating ε_1, for each policyholder in the sample of the 10%-deductible contract, we also estimate the premium rate of the 20%-deductible contract by it in the tariff further adjusted for the discount rate. The estimated premium rate of the 20%-deductible contract and summary statistics of other variables are listed in Table 1.

[Insert Table 1 here]

As we can see, for those who actually purchased the 10%-deductible contract, the average estimated premium rate if purchasing the 20%-deductible contract instead, is 0.0052, which is smaller than the average paid premium rate in the 10%-deductible contract (0.0059). In addition, the mean estimated theft probability is 0.0041. A lower mean estimated theft probability than the mean premium rate in the 10%-deductible contract suggests that the pricing is adequate and pretty well controlled by the insurance company.

6 Empirical Results

This section reports our empirical results. We show quantiles of the upper bound estimates of ε_1 in Table 2. Then, by a derived relation between ε_1 and absolute risk aversion coefficient, we further infer quantile-based upper bound estimates of the ARA coefficient based on the obtained estimates of ε_1, which are shown in Table 3.

Table 2 shows that if the decisions of *all* policyholders in the sample on purchasing the 10%-deductible contract are optimal, $\widehat{R}(100\%) = 8.198e^{-08}$, which is very small and close to zero. The bootstrapped 95% confidence interval estimate is $[2.29e^{-16}, 0.00255]$, which indicates that there is a 2.5% of chance that the upper bound estimate of ε_1 at correct rate 100% is beyond 0.00255 though. For the whole sample, our estimates of ε_1 are close to zero and provide evidence supporting that it could be reasonable to choose a zero ε_1 for a large sample of risk averters. Our estimates are much smaller than those obtained by previous studies. For example, through lab experiments with 180 subjects, Levy et al. (2010) obtained an estimate of 0.059 for *all* subjects.[9] By the real data consists of 940,904 observations, we obtained a much smaller estimate for *all* policyholders ($8.198e^{-08}$). This difference could be attributed to the property of large sample that the obtained estimate gets smaller as the sample size increases, as noted in Section 4.

[Insert Table 2 here]

A quantile-based estimation further sheds light on how slight violation against SD rules that *most* individuals would accept. Table 2 presents that when 99% (97.5% and 95%, respectively) of the policyholders in our sample purchased the 10%-deductible contract optimally, the upper bound estimate of ε_1 is 0.0399 (0.0554 and 0.0727, respectively) with a 95% confidence interval estimate of [0.0298, 0.0463] ([0.0456, 0.0621] and [0.0635, 0.0791], respectively). The upper bound estimate of ε_1 becomes smaller as the percentage of the policyholders in the sample who made an optimal decision gets higher. Compared with the literature, our estimate of 0.0554 at correct rate 97.5% is closest and slightly smaller than Levy et al.'s (2010) estimate of 0.059 at correct rate 100% while our estimate of 0.0727 at correct rate 95% is larger than that of

[9] From the survey data with 223 subjects, Huang et al. (2015) obtained an upper bound estimate of 0.0527 for *all* subjects which is slightly smaller than that in Levy et al. (2010).

Levy et al. (2010). Our quantile-based estimates of ε_1 accordingly suggest reasonable violation ratios against SD rules for *most* risk averters.

We can further infer an estimate of ARA coefficient for the policyholders at different correct rates via our estimates of ε_1. According to the definition of $\underline{U}_2(\varepsilon_1, 0)$, for all $u \in \underline{U}_2(\varepsilon_1, 0)$, a relation between the ARA coefficient and ε_1 is derived as follows:

$$
\begin{aligned}
&\frac{\sup\{u'(x)\}}{\inf\{u'(x)\}} \leq \frac{1}{\varepsilon_1} - 1 \\
\Leftrightarrow &\int_{\underline{x}}^{\bar{x}} -\frac{u''(t)}{u'(t)}\, dt \leq ln\left(\frac{1}{\varepsilon_1} - 1\right) \\
\Leftrightarrow &\frac{\int_{\underline{x}}^{\bar{x}} -\frac{u''(t)}{u'(t)}\, dt}{\bar{x} - \underline{x}} \leq \frac{ln\left(\frac{1}{\varepsilon_1} - 1\right)}{\bar{x} - \underline{x}}.
\end{aligned} \tag{7}
$$

It is noted that $\frac{\int_{\underline{x}}^{\bar{x}} -\frac{u''(t)}{u'(t)}\, dt}{\bar{x} - \underline{x}}$ is the average ARA coefficient of the policyholder with his/her final wealth within the interval $[\underline{x}, \bar{x}]$; thus, we could consider it as a measure of the degree of the policyholder's ARA. In addition, inequality (7) shows that $\frac{ln\left(\frac{1}{\varepsilon_1} - 1\right)}{\bar{x} - \underline{x}} \equiv \gamma(\varepsilon_1; L)$ is an upper bound for the policyholder's ARA coefficient, where in our setting, $\bar{x} - \underline{x} = \left(w - P^{20\%}Q^{20\%}\right) - \left(w - L + \left(1 - P^{20\%}\right) Q^{20\%}\right) = 0.2L$. Therefore, assuming that there are $m\%$ of the policyholders in the sample who purchased the 10% deductible optimally, based on our upper estimate of ε_1 (i.e., $\widehat{R}(m\%)$) in Table 2, we obtain the upper bound estimate of the ARA coefficient for the policyholder denoted by $\widehat{\gamma}(\widehat{R}(m\%); L)$ is

$$
\widehat{\gamma}(\widehat{R}(m\%); L) = \frac{ln\left(\frac{1}{\widehat{R}(m\%)} - 1\right)}{0.2L}. \tag{8}
$$

For an illustration, we take the mean insured car value of the sample (denoted by $\bar{L}$), which is NT$380,070, as the value of L. The results are shown in Table 3.

[Insert Table 3 here]

Table 3 shows that when *all* the policyholders in the sample purchased the 10% deductible optimally, the policyholder with an insured car valued at NT$380,070 has an ARA coefficient no larger than 0.0002 (i.e., $\widehat{\gamma}(\widehat{R}(100\%); \bar{L}) = 0.0002$). On basis of the bootstrapped interval estimate in Table 2 and Condition (8), we also obtain a 95% confidence interval estimate of $\widehat{\gamma}(\widehat{R}(100\%); \bar{L})$ which is [$7.827e^{-05}$, 0.0005]. Compared to the estimates shown in the literature, e.g, Cohen and Einav (2007), our point estimate of 0.0002 at the mean insured car value is smaller than Cohen and Einav's (2007) mean estimate of 0.0067 but larger than their median estimate of $2.6e^{-05}$.[10] Our estimate is actually very close to and slightly smaller than their 75th percentile estimate of 0.00029. Both their estimates of median and the 75th percentile are included in our 95% confidence interval estimate.

On the other hand, Table 3 shows that when 99% (97.5% and 95%, respectively) of the policyholders in the sample purchased the 10% deductible optimally, the upper bound estimate of ARA coefficient at the mean insured car value is $4.184e^{-05}$ ($3.731e^{-05}$ and $3.349e^{-05}$, respectively) with a 95% confidence interval estimate of [$3.980e^{-05}$, $4.582e^{-05}$] ([$3.572e^{-05}$, $4.001e^{-05}$] and [$3.229e^{-05}$, $3.540e^{-05}$], respectively]). The above results reveal that the upper bound estimate of ARA coefficient increases with the percentage of the optimal choice in the sample, which can be seen from the negative relation between the estimates of ε_1 and ARA coefficient in Equation (8). Compared with Cohen and Einav's (2007) results, all of our m% upper bound estimates (m=99, 97.5, and 95) at the mean insured car value are smaller than their estimates of mean and the 75th percentile (0.0067 and 0.00029, respectively); however, ours are close to and rather larger than their median estimate of $2.6e^{-05}$. Our quantile-based estimates of ARA coefficient are accordingly reasonable and comparable to previous works.

7 Conclusions

In this paper, we estimate an upper bound for ε_1 in (ε_1,0)-GASSD with the data on 10%-deductible contract of automobile theft insurance. We propose a quantile-based methodology for empirical estimation. Our estimates of ε_1 obtained from real data are smaller than previous findings obtained by experiments, but ours are closer to reality and provide more meaningful information for empirical applications of ASD rules. Based on the estimates of ε_1, we further infer an upper bound for ARA coefficient. Our estimates of the ARA coefficient are not only comparable to the previous findings but also linked to ASD decision rules.

[10]By the similar data on deductible choices in automobile insurance, Cohen and Einav (2007) directly estimated the ARA coefficient and reported the quantiles of the estimate while we obtain our estimates via the estimates of the acceptable violation ratio against SD rules.

For future research, it would be interesting to investigate heterogeneities in our estimates of ε_1. Our results are preliminary research implemented with real data on the value of the acceptable violation ratio against SD rules. However, it is unknown that whether the upper bound estimate of ε_1 varies with characteristics of the policyholder or the insured car. It therefore could be a complement to the literature on ASD to empirically examine this issue.

Appendices

A Equivalent Conditions for (ε_1,0)-GASSD in (1) $B(x)^{20\%}$ dominating $B(x)^{10\%}$ and (2) $B(x)^{10\%}$ dominating $B(x)^{20\%}$

In this section, based on Figure 1, we show equivalent conditions for $B(x)^{20\%}$ dominating $B(x)^{10\%}$ by (ε_1,0)-GASSD and $B(x)^{10\%}$ dominating $B(x)^{20\%}$ by (ε_1,0)-GASSD, respectively.

Result 1: *(ε_1,0)-GASSD and ε_1-AFSD require the same condition for $B(x)^{20\%}$ dominating $B(x)^{10\%}$.*

Proof: $B(x)^{20\%}$ dominates $B(x)^{10\%}$ by ε_1-AFSD (Leshno and Levy, 2002) is defined as:

Definition A1 *For $0 < \varepsilon_1 < \frac{1}{2}$, $B(x)^{20\%}$ dominates $B(x)^{10\%}$ by ε_1-almost FSD if and only if*

$$\int_{S_1} \left[B(x)^{20\%} - B(x)^{10\%}\right] dx \leq \varepsilon \parallel B(x)^{20\%} - B(x)^{10\%} \parallel ,$$

where $S_1 = \left\{x \in \left[w - L + (1 - P^{20\%})Q^{20\%}, w - P^{20\%}Q^{20\%}\right] : B(x)^{10\%} < B(x)^{20\%}\right\}$ *and* $\parallel B(x)^{20\%} - B(x)^{10\%} \parallel = \int_{w-L+(1-P^{20\%})Q^{20\%}}^{w-P^{20\%}Q^{20\%}} \mid B(x)^{20\%} - B(x)^{10\%} \mid dx.$

According to Definition A1, the condition for $B(x)^{20\%}$ to dominate $B(x)^{10\%}$ by ε_1-AFSD is

$$R_{AFSD}^{20\%\succ 10\%} = \frac{C}{C+D} \leq \varepsilon_1, \tag{A1}$$

which happens to be the same as Condition (3) for $B(x)^{20\%}$ dominating $B(x)^{10\%}$ by (ε_1,0)-GASSD.

Q.E.D

Result 2: *(ε_1,0)-GASSD is equivalent to SSD for $B(x)^{10\%}$ dominating $B(x)^{20\%}$.*

Proof: According to Definition 1, $B(x)^{10\%}$ dominants $B(x)^{20\%}$ by (ε_1,0)-GASSD for all $u \in \underline{U}_2(\varepsilon_1, 0)$ if and only if $B^{(2)}(w - P^{20\%}Q^{20\%})^{10\%} - B^{(2)}(w - P^{20\%}Q^{20\%})^{20\%} \leq 0$ and

$$\begin{aligned} &\max_{x\in[w-L+(1-P^{20\%})Q^{20\%}, w-P^{20\%}Q^{20\%}]} \left[B^{(2)}(x)^{10\%} - B^{(2)}(x)^{20\%}\right] \\ &\quad \leq \frac{\varepsilon_1}{1-2\varepsilon_1}\left[B^{(2)}(w - P^{20\%}Q^{20\%})^{20\%} - B^{(2)}(w - P^{20\%}Q^{20\%})^{10\%}\right]. \end{aligned} \tag{A2}$$

$B^{(2)}(w - P^{20\%}Q^{20\%})^{10\%} - B^{(2)}(w - P^{20\%}Q^{20\%})^{20\%} \le 0$, which means that $D - C \le 0$. Accordingly, $\max_{x \in [w-L+(1-P^{20\%})Q^{20\%}, w-P^{20\%}Q^{20\%}]} \left[B^{(2)}(x)^{10\%} - B^{(2)}(x)^{20\%}\right] = 0$ and Equation (A2) is rewritten as

$$0 \le \frac{\varepsilon_1}{1 - 2\varepsilon_1}(C - D) \Leftrightarrow \varepsilon_1 \ge 0, \tag{A3}$$

where $1 - 2\varepsilon_1 \ge 0$ and $C - D \ge 0$. In other words, $B(x)^{10\%}$ dominants $B(x)^{20\%}$ by SSD.

Q.E.D

B Variable Definitions

We describe part of the variables defined in this paper as follows:

- Age—The age of the policyholder (vehicle owner).
- Female—A dummy variable equals 1 when the insured is female.
- Marriage—A dummy variable equals 1 when the insured was in marital status.
- Car age—The age of the vehicle when it was insured, and four dummy variables further indicate the class of the insured vehicle's age from age 0 to age 4 (denoted by *Car age 0*, *Car age 1*, *Car age 2*, *Car age 3*, and *Car age 4*).
- Engine size—The volume of engine which is measured in cubic centimeters (cc).
- Size—Two dummy variables indicate the class of the engine size of the insured vehicle, including the size between 1,800 c.c. and 3,000 c.c. (denoted by *Median size*) and greater than 3,000 c.c. (denoted by *Large size*).
- Brand—13 dummy variables indicate brands of vehicles sold in Taiwan, which are denoted by *Nissan*, *Ford*, *Honda*, *Toyota*, *Mitsubishi*, *Mazda*, *French*, *VW* (Volkswagen), *Korean* (some Korean brands), *British* (some British brands), *US* (some other US brands), *Luxury* (some Luxury brands), *OJapan* (some old Japanese brands), respectively.
- City—A dummy variable equals 1 when the insured vehicle was registered in a city.

- Area—Three dummy variables indicate the insured vehicle's registered area in Taiwan, including the North (denoted by *Northern*), the South (denoted by *Southern*), and the midland (denoted by *Central*).

- Policy year—Six dummy variables indicate the year in which the policy was issued from year 2002 to year 2007 (denoted by *2002*, *2003*, *2004*, *2005*, *2006*, and *2007*).

References

[1] Anderson, S., G. W. Harrison, M. I. Lau, and E. E. Rutström, 2008, Eliciting Risk and Time Preferences, *Econometrica*, 76(3): 583–618.

[2] Bali, T. G., K. O. Demirtas, H. Levy, and A. Wolf, 2009, Bonds versus Stocks: Investors' Age and Risk Taking, *Journal of Monetary Economics*, 56(6): 817–830.

[3] Bali, T. G., S. J. Brown, and K. O. Demirtas, 2013, Do Hedge Funds Outperform Stocks and Bonds? *Management Science*, 59(8): 1887–1903.

[4] Bliss, P. R., and N. Panigirtzoglou, 2004, Option-Implied Risk Aversion Estimates, *Journal of Finance*, 59(1): 407–446.

[5] Bollerslev, T., M. Gibson, and H. Zhou, 2011, Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities, *Journal of Econometrics*, 160(1): 235–245.

[6] Brenner, S., 2015. The Risk Preferences of U.S. Executives, *Management Science*, 61(6): 1344–1361.

[7] Bucciol, A., and R. Miniaci, 2011, Household Portfolios and Implicit Risk Preference, *The Review of Economics and Statistics*, 93(4): 1235–1250.

[8] Chetty, R., 2006, A New Method of Estimating Risk Aversion, *American Economic Review*, 96(5): 1821–1834.

[9] Cohen A., and L. Einav, 2007, Estimating Risk Preferences from Deductible Choice, *American Economic Review*, 97(3): 745–788.

[10] Denuit, M. M., R. J. Huang, L. Y. Tzeng, and C. W. Wang, 2014, Almost Marginal Conditional Stochastic Dominance, *Journal of Banking and Finance*, 41: 57–66.

[11] Denuit, M. M., R. J. Huang, and L. Y. Tzeng, 2014, Bivariate Almost Stochastic Dominance, *Economic Theory*, 57(2):377–405.

[12] Denuit, M. M., R. J. Huang, and L. Y. Tzeng, 2014, Almost Expectation and Excess Dependence Notions, *Theory Decision*, 1–27.

[13] Gertner, R., 1993, Game Shows and Economic Behavior: Risk-Taking on "Card Sharks", *The Quarterly Journal of Economics*, 108(2): 507–521.

[14] Holt, C. A., and S. K. Laury, 2002, Risk Aversion and Incentive Effects, *American Economic Review*, 92(5): 1644–1655.

[15] Huang, R. J., Y.-H. Huang, L. Y. Tzeng, 2015, Experimental Estimation of the Preference Parameters in Almost Stochastic Dominance, *Working paper*.

[16] Leshno, M., and H. Levy, 2002, Preferred by "All" and Preferred by "Almost" Decision Makers: Almost Stochastic Dominance, *Management Science*, 48(8): 1074–1085.

[17] Levy, H., 1992, Stochastic Dominance and Expected Utility: Survey and Analysis, *Management Science*, 38(4): 555–593.

[18] Levy, H., 1998, Stochastic Dominance: Investment Decision Making under Uncertainty, *Kluwer Academic Publishers*, Boston, MA.

[19] Levy, H., M. Leshno, and B. Leibovitch, 2010, Economically Relevant Preferences for All Observed Epsilon, *Annals of Operations Research*, 176(1): 153–178.

[20] Levy, M., 2009, Almost Stochastic Dominance and Stocks for the long run, *European Journal of Operational Research*, 194(1): 250–257.

[21] Levy, M., 2012, Almost Stochastic Dominance and Efficient Investment Sets, *American Journal of Operational Research*, 2(3): 313–321.

[22] Lizyayev, A., and A. Ruszczyński, 2012, Tractable Almost Stochastic Dominance, *European Journal of Operational Research*, 218(2): 448–455.

[23] Metrick, A., 1995, A Natural Experiment in "Jeopardy!", *American Economic Review*, 85(1): 240–253.

[24] Rothschild, M., and J. E. Stiglitz, 1970, Increasing Risk: I. A definition, *Journal of Economic Theory*, 2(3): 225–243.

[25] Tsetlin, I., R. L. Winkler, R. J. Huang, and L. Y. Tzeng, 2015, Generalized Almost Stochastic Dominance, *Operations Research*, 63(2): 363–377.

[26] Tzeng, L. Y., R. J. Huang, and P.-T. Shih, 2013, Revisiting Almost Second-Degree Stochastic Dominance, *Management Science*, 59(5): 1250–1254.

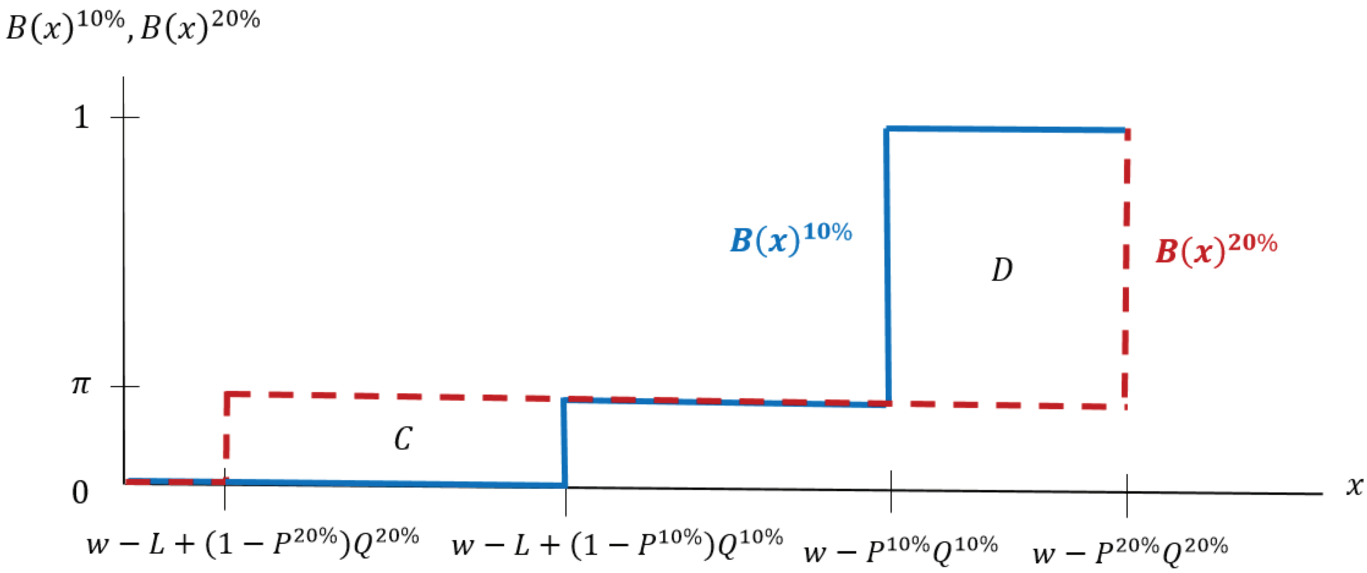

Figure 1: Cumulative Distribuion Functions under 10%-Deductible and 20%-Deductible Insurance Contracts

Notes: This figure presents the cumulative distribution functions (CDFs) of the final wealth x under a 10%-deductible insurance contract, denoted by $B(x)^{10\%}$, and a 20%-deductible insurance contract, denoted by $B(x)^{20\%}$. The solid line draws $B(x)^{10\%}$ and the dotted line draws $B(x)^{20\%}$. A risk-averse individual with an initial wealth w and a potential loss L with a probability of occurrence π considers which levels of insurance deductibles to purchase to transfer the risk by ranking $B(x)^{10\%}$ and $B(x)^{20\%}$. P^k denotes the premium rate for per unit of coverage amount and Q^k denotes the coverage amount in the 10%-or 20%-deductible contract, where $k = 10\%$ or 20%. Under the 10%-deductible contract, the individual obtains $w - L + (1 - P^{10\%})Q^{10\%}$ with probability π and $w - P^{10\%}Q^{10\%}$ with probability with $1 - \pi$. Under the 20%-deductible contract, the individual obtains $w - L + (1 - P^{20\%})Q^{20\%}$ with probability π and $w - P^{20\%}Q^{20\%}$ with probability with $1 - \pi$. In general, $(1 - P^{10\%})Q^{10\%} > (1 - P^{20\%})Q^{20\%}$. C denotes the area where $B(x)^{20\%}$ is higher than $B(x)^{10\%}$ and D denotes the area where $B(x)^{10\%}$ is higher than $B(x)^{20\%}$.

Table 1: SUMMARY STATISTICS—COVARIATES

Variable			Mean	Std. dev.	Min	Max
Premium rate:[a]			0.0059	0.0017	3.81E-07	0.0150
Estimated premium rate:[b] (if choosing 20%-deductible contract instead)			0.0052	0.0015	3.38E-07	0.0133
Estimated theft rate:			0.0041	0.0034	5.90E-10	0.0476
Demographics:	Age		42.5027	9.8939	18	93
	Female		0.6297	0.4829	0	1
	Marriage		0.8986	0.3019	0	1
Car attributes:	Value (NTD)[c]		3.8007	3.4045	0.0100	99.8000
	Car age (Interval)		3.5035	2.8220	0	89
	Car age (Dummy)	Car age 0	0.1218	0.3270	0	1
		Car age 1	0.1748	0.3798	0	1
		Car age 2	0.1524	0.3594	0	1
		Car age 3	0.1312	0.3376	0	1
		Car age 4[d]	0.0986	0.2981	0	1
	Engine size (cc)[e]		1.9126	0.4665	0.0030	22.1630
	Size[f]	Median size	0.3434	0.4748	0	1
		Large size	0.1533	0.3603	0	1

Note: This table reports summary statistics of variables used in empirical estimation, which are obtained with our full sample which consists of policyholders who purchased the 10%-deductible automobile theft insurance contract.

[a] *Premium rate* is calculated for each policyholder in the 10%-deductible sample by the premiums actually paid divided by the coverage amount.

[b] *Estimated premium rate* is the premium rate estimated for each policyholder in the 10%-deductible sample if choosing 20%-deductible contract instead.

[c] *Value* is measured in the unit of one hundred thousand NT dollars.

[d] The insured cars that are over than 4 years are the reference group.

[e] *Engine size* is measured in the unit of one thousand cubic capacity (cc).

[f] *Small size* which represents the insured cars with the engine size less than 1,800 c.c., are the reference group.

Table 1: SUMMARY STATISTICS—COVARIATES (Continued)

Variable			Mean	Std. dev.	Min	Max
Car attributes:	Brand[g]	Nissan	0.1237	0.3293	0	1
		Ford	0.1058	0.3075	0	1
		Honda	0.0661	0.2484	0	1
		Toyota	0.4025	0.4904	0	1
		Mitsubishi	0.0825	0.2751	0	1
		Mazda	0.0419	0.2004	0	1
		French	0.0079	0.0883	0	1
		VW	0.0349	0.1834	0	1
		Korean	0.0338	0.1806	0	1
		British	0.0011	0.0331	0	1
		US	0.0111	0.1048	0	1
		Luxury	0.0628	0.2427	0	1
		OJapan	0.0203	0.1409	0	1
Registered:	City		0.5231	0.4995	0	1
	Area[h]	Northern	0.4995	0.5000	0	1
		Southern	0.2714	0.4447	0	1
		Central	0.1715	0.3770	0	1
Policy year:[i]	2002		0.1189	0.3236	0	1
	2003		0.1301	0.3364	0	1
	2004		0.1456	0.3527	0	1
	2005		0.1555	0.3624	0	1
	2006		0.1323	0.3388	0	1
	2007		0.1632	0.3696	0	1
Obs.	940,904					

[g] The insured cars with brands not belonging to the above ones are the reference group.

[h] The insured cars registered in eastern part of Taiwan are the reference group.

[i] Polices sold in year 2008 are the reference group.

Table 2: Upper Bound Estimates of ε_1

Correct Rate $m\%$[a]	Point Estimate $(\widehat{R}(m\%))$[b]	Bootstrapped 95% Confidence Interval Estimate[c]
100%	$8.198e^{-08}$	$2.29e^{-16}$, 0.0026
99%	0.0399	0.0298, 0.0463
97.5%	0.0554	0.0456, 0.0621
95%	0.0727	0.0635, 0.0791
Obs.	940,904	
Comparable Estimates at Correct Rate 100%:		
Levy et al. (2010) (180 Obs.)	0.059	N.A.
Huang et al. (2015) (223 Obs.)	0.0527	N.A.

Note: This table reports upper bound point estimate of ε_1 (i.e., $\widehat{R}$) and bootstrapped 95% confidence interval estimate of ε_1 at correct rates 100%, 99%, 97.5%, and 95%, respectively, obtained with our full sample of the 10%-deductible automobile theft insurance contract.

[a] Assume that $m\%$ of the policyholders in the full sample made an optimal decision on purchasing the 10%-deductible automobile theft insurance contract based on (ε_1, 0)-GASSD rule.

[b] An upper bound estimate of ε_1 at correct rate $m\%$ is denoted by $\widehat{R}(m\%)$. It is a quantile-based upper bound estimate of ε_1 which denotes the $(100 - m)$th percentile of the empirical distribution of $\widehat{R}$.

[c] The 95% confidence interval estimate is constructed by bootstrap. 1,000 bootstrapped samples with an equal size were drawn from the original research sample with replacement. The lower and upper bounds of bootstrapped 95% confidence interval of $\widehat{R}$ are the 2.5th and 97.5th percentiles, respectively, of the empirical distribution of $\widehat{R}$ obtained by the 1,000 bootstrapped samples.

Table 3: Upper Bound Estimates of Absolute Risk Aversion Coefficient at The Mean Insured Car Value NT$380,070

Correct Rate $m\%$[a]	Point Estimate $\widehat{\gamma}(\widehat{R}(m\%);\bar{L})$[b]	95% Confidence Interval Estimate[c]
100%	0.0002	$7.827e^{-05}$, 0.0005
99%	$4.184e^{-05}$	$3.980e^{-05}$, $4.582e^{-05}$
97.5%	$3.731e^{-05}$	$3.572e^{-05}$, $4.001e^{-05}$
95%	$3.349e^{-05}$	$3.229e^{-05}$, $3.540e^{-05}$
Obs.	940,904	
Comparable Estimates (Cohen and Einav, 2007):		
Mean	0.0067	N.A.
Median	$2.6e^{-05}$	N.A.
75th Percentile	0.00029	N.A.

Note: This table reports upper bound estimates of absolute risk aversion (ARA) coefficient for the policyholder with an insured car valued at the mean of the sample (NT$380,070) at correct rates 100%, 99%, 97.5%, and 95%, respectively, obtained on basis of the upper bound estimates of ε_1 in Table 2.

[a] Assume that $m\%$ of the policyholders in the sample made an optimal decision on purchasing the 10%-deductible automobile theft insurance contract based on $(\varepsilon_1, 0)$-GASSD rule.

[b] The upper bound point estimate of the ARA coefficient is estimated by the derived relation that $\frac{\int_{\underline{x}}^{\bar{x}} -\frac{u''(t)}{u'(t)}\,dt}{\bar{x}-\underline{x}} \leq \frac{ln\left(\frac{1}{\varepsilon}-1\right)}{\bar{x}-\underline{x}}$, where $\bar{x}-\underline{x}=0.2L$. Let the mean insured car value ($\bar{L}$) as the value of L. Accordingly, we obtain the upper bound point estimate of the ARA coefficient at correct rate $m\%$ and the mean insured car value denoted by $\widehat{\gamma}(\widehat{R}(m\%);\bar{L})$ is $\frac{ln\left(\frac{1}{\widehat{R}(m\%)}-1\right)}{0.2\bar{L}}$, where $\bar{L}=380,070$. It is also a quantile-based upper bound estimate of the ARA coefficient at the mean insured car value which denotes the $(100-m)$th percentile of the empirical distribution of $\widehat{\gamma}(\widehat{R}(m\%);\bar{L})$.

[c] The 95% confidence interval estimate is estimated by the same derived relation on basis of the bootstrapped 95% confidence interval estimates in Table 2. The lower and upper bounds of 95% confidence interval of $\widehat{\gamma}(\widehat{R}(m\%);\bar{L})$ are the 2.5th and 97.5th percentiles, respectively, of the empirical distribution of $\widehat{\gamma}(\widehat{R}(m\%);\bar{L})$.