Online Appendix

to

The Nature of Risk Preferences: Evidence from Insurance Choices

American Economic Review

Levon Barseghyan
Cornell University

Francesca Molinari
Cornell University

Ted O'Donoghue
Cornell University

Joshua C. Teitelbaum
Georgetown University

Contents

A Coverage Descriptions

Auto collision coverage pays for damage to the insured vehicle caused by a collision with another vehicle or object, without regard to fault. *Auto comprehensive* coverage pays for damage to the insured vehicle from all other causes (e.g., theft, fire, flood, windstorm, glass breakage, vandalism, hitting or being hit by an animal, or by falling or flying objects), without regard to fault. If the insured vehicle is stolen, auto comprehensive coverage also provides a certain amount per day for transportation expenses (e.g., rental car or public transportation). *Home all perils* coverage pays for damage to the insured home from all causes (e.g., fire, windstorm, hail, tornadoes, vandalism, or smoke damage), except those that are specifically excluded (e.g., flood, earthquake, or war). For simplicity, we often refer to home all perils merely as *home*.

B Identification

In this section, we prove Properties 1, 2, and 3 from Section II.C. Take any three deductible options $a, b, c \in \mathcal{D}$, with $a > b > c$, and consider a household with premium p_a for deductible a and claim probability μ. The household's r and $\Omega(\mu)$ determine the premium $\widetilde{p}_b$ that makes the household indifferent between deductibles a and b, as well as the premium $\widetilde{p}_c$ that makes the household indifferent between deductibles a and c. Notice that $\widetilde{p}_b - p_a$ reflects the household's maximum willingness to pay (WTP) to reduce its deductible from a to b, and $\widetilde{p}_c - \widetilde{p}_b$ reflects the household's additional WTP to reduce its deductible from b to c. In what follows, we simplify notation by suppressing the explicit dependence of $\widetilde{p}_b$ and $\widetilde{p}_c$ on p_a, μ, r, and $\Omega(\mu)$. We also suppress the argument of Ω. In addition, let L_a denote the deductible lottery associated with deductible a at premium p_a.

Recall equation (2) from Section II.A:

$$U(L_d) = -\left[p_d + \Omega d\right] - \frac{r}{2}\left[(1-\Omega)\left(p_d\right)^2 + \Omega\left(p_d + d\right)^2\right].$$

Define $p(x)$ as the premium for deductible x such that the household is indifferent between the resulting lottery and lottery L_a. Hence, $p(b) = \widetilde{p}_b$ and $p(c) = \widetilde{p}_c$. Applying equation (2), $p(x)$ is defined by each of the following equations (both of which we use below):[1]

$$-p(x) - \Omega x - \frac{r}{2}\left[(1-\Omega)(p(x))^2 + \Omega(p(x)+x)^2\right] = U(L_a) \qquad \text{(A.1)}$$

[1] These equations are equivalent, where the latter merely expands $U(L_a)$ and rearranges terms.

$$(p(x) - p_a) - \Omega\,(a - x) + \frac{r}{2}(p(x)^2 - p_a^2) + \Omega\frac{r}{2}\left\{(x^2 - a^2) + 2(p(x)x - p_a a)\right\} = 0. \qquad \text{(A.2)}$$

For the proofs below, it is useful to define W and V as

$$\begin{aligned} W(p,x,r,\Omega) &\equiv -p - \Omega x - \frac{r}{2}\left[(1-\Omega)(p)^2 + \Omega(p+x)^2\right] \\ V(p,x,r,\Omega) &\equiv (p - p_a) - \Omega\,(a-x) + \frac{r}{2}(p^2 - p_a^2) + \Omega\frac{r}{2}\left\{(x^2 - a^2) + 2(px - p_a a)\right\}. \end{aligned}$$

Our first lemma establishes that $p(x)$ is well behaved.

Lemma 1. *For any $r \geq 0$, $\Omega \in (0,1)$, $p_a > 0$, and $x \leq a$, $p(x)$ is a continuous and differentiable function with $dp/dx < 0$ (and thus $p(c) > p(b) > p_a$).*

It is straightforward to derive that W is twice differentiable and satisfies the conditions of the implicit function theorem. Thus $p(x)$ is a continuous and differentiable function, and

$$\frac{dp}{dx} = \frac{-\frac{\partial W}{\partial x}}{\frac{\partial W}{\partial p}} = \frac{-\Omega\left[1 + rp(x) + rx\right]}{1 + rp(x) + r\Omega x} < 0.$$

Our second lemma states that standard risk aversion implies that a household's WTP to reduce its deductible is strictly greater than the expected reduction in the deductible paid (evaluated at a claim probability of Ω).

Lemma 2. *For any $x' < x \leq a$, if $r = 0$ then $p(x') - p(x) = \Omega(x - x')$, and if $r > 0$ then $p(x') - p(x) > \Omega(x - x')$.*

Proof. The result for $r = 0$ is straightforward. For $r > 0$, define $\tilde{V}$ as

$$\tilde{V}(p,x',r,\Omega) \equiv [p - p(x)] - \Omega\,(x - x') + \frac{r}{2}(p^2 - p(x)^2) + \Omega\frac{r}{2}\left\{(x'^2 - x^2) + 2(px' - p(x)x)\right\},$$

in which case $p(x')$ is defined by $\tilde{V}(p(x'),x',r,\Omega) = 0$. Note that $p = \Omega(x - x') + p(x)$ implies

$$\begin{aligned} \tilde{V}(p,x',r,\Omega) &\equiv [(\Omega(x-x') + p(x)) - p(x)] - \Omega\,(x - x') + \frac{r}{2}((\Omega(x-x') + p(x))^2 - (p(x))^2) \\ &\quad + \Omega\frac{r}{2}\left\{(x'^2 - x^2) + 2((\Omega(x-x') + p(x))x' - p(x)x)\right\} \\ &= \frac{r}{2}(\Omega^2(x-x')^2 + 2p(x)\Omega(x-x')) \\ &\quad + \Omega\frac{r}{2}\left\{(x'^2 - x^2) + 2(\Omega(x-x')x' + p(x)(x'-x))\right\} \\ &= \frac{r}{2}\Omega(x-x')[\Omega(x+x') - (x+x')] < 0. \end{aligned}$$

Since $\partial\tilde{V}/\partial p = 1 + rp + \Omega r x' > 0$, it follows that $p(x') > \Omega(x - x') + p(x)$. □

Property 1 establishes the relationship between the magnitude of willingness to pay and risk preferences.

Property 1. *For any $x < a$, $p(x)$ is strictly increasing in r and Ω.*

Proof. By implicit function theorem:

$$\frac{\partial p(x)}{\partial r} = \frac{-\frac{\partial V}{\partial r}}{\frac{\partial V}{\partial p}} \qquad \text{and} \qquad \frac{\partial p(x)}{\partial \Omega} = \frac{-\frac{\partial V}{\partial \Omega}}{\frac{\partial V}{\partial p}}.$$

Note that

$$\frac{\partial V}{\partial r} = -\frac{1}{r}[(p(x) - p_a) - \Omega\,(a - x)] < 0,$$

where the equality uses equation (A.2) and the inequality follows from Lemma 2. Note further that

$$\frac{\partial V}{\partial \Omega} = -\frac{1}{\Omega}[(p(x) - p_a) + \frac{r}{2}(p(x)^2 - p_a^2)] < 0,$$

where the equality uses equation (A.2) and the inequality follows from Lemma 1. Finally, given that $\partial V/\partial p = 1 + rp + \Omega rx > 0$, it follows that $\frac{\partial p(x)}{\partial r} > 0$ and $\frac{\partial p(x)}{\partial \Omega} > 0$. □

We next establish Property 2, which shows that a risk averse household's WTP to avoid an incremental loss depends positively on the magnitude of the absolute loss.

Property 2. *For any fixed $\Omega(\mu)$, $r = 0$ implies $\frac{p(b)-p_a}{p(c)-p(b)} = \frac{a-b}{b-c}$, and the ratio $\frac{p(b)-p_a}{p(c)-p(b)}$ is strictly increasing in r.*

Proof. The result for $r = 0$ is straightforward. As a preliminary step in proving the second part, we first prove that, for any $r > 0$, $\frac{p(b)-p_a}{p(c)-p(b)} > \frac{a-b}{b-c}$. From Lemma 1, $p(x)$ is continuous and differentiable, and thus

$$p(b) - p_a = \int_b^a \left(-\frac{dp}{dx}\right) dx \qquad \text{and} \qquad p(c) - p(b) = \int_c^b \left(-\frac{dp}{dx}\right) dx.$$

From the proof of Lemma 1, $-\frac{dp}{dx} = \frac{\Omega[1+rp(x)+rx]}{1+rp(x)+r\Omega x} > 0$, and thus

$$\frac{d\left[-\frac{dp}{dx}\right]}{dx} = \Omega\frac{r(1-\Omega)(1+rp-rx\frac{dp}{dx})}{[1+rp(x)+r\Omega x]^2} > 0.$$

In words, $-\frac{dp}{dx}$ reflects the household's marginal willingness to pay to reduce its deductible, and $-\frac{dp}{dx} > 0$ reflects that a household is indeed willing to pay a higher premium to reduce its deductible. More importantly, $d\left[-\frac{dp}{dx}\right]/dx > 0$ reflects that the larger is its deductible, the larger is the household's marginal willingness to pay to reduce that deductible (or equivalently the smaller is its deductible, the smaller is the household's marginal willingness to pay to reduce that deductible). Finally, $d\left[-\frac{dp}{dx}\right]/dx > 0$ implies

$$\begin{aligned} p(b) - p_a &= \int_b^a \left(-\frac{dp}{dx}\right) dx > (a-b)\left(-\left.\frac{dp}{dx}\right|_{x=b}\right) \\ p(c) - p(b) &= \int_c^b \left(-\frac{dp}{dx}\right) dx < (b-c)\left(-\left.\frac{dp}{dx}\right|_{x=b}\right) \end{aligned}$$

which together imply $\frac{p(b)-p_a}{p(c)-p(b)} > \frac{a-b}{b-c}$. With this result in hand, we now prove that $\frac{p(b)-p_a}{p(c)-p(b)}$ is strictly increasing in r. Note that $d\left(\frac{p(b)-p_a}{p(c)-p(b)}\right)/dr > 0$ if and only if $\frac{1}{p(b)-p_a}\frac{\partial p(b)}{\partial r} > \frac{1}{p(c)-p_a}\frac{\partial p(c)}{\partial r}$. Applying $\frac{\partial p(x)}{\partial r}$ from the proof of Property 1,

$$\begin{aligned} \frac{1}{p(b)-p_a}\frac{\partial p(b)}{\partial r} &= \frac{1}{p(b)-p_a}\frac{\frac{1}{r}\left[(p(b)-p_a) - \Omega(a-b)\right]}{1+rp(b)+\Omega rb}, \\ \frac{1}{p(c)-p_a}\frac{\partial p(c)}{\partial r} &= \frac{1}{p(c)-p_a}\frac{\frac{1}{r}\left[(p(c)-p_a) - \Omega(a-c)\right]}{1+rp(c)+\Omega rc}. \end{aligned}$$

We have

$$\frac{1}{p(b)-p_a}\left[(p(b)-p_a) - \Omega(a-b)\right] > \frac{1}{p(c)-p_a}\left[(p(c)-p_a) - \Omega(a-c)\right],$$

because

$$(p(c)-p_a)\left[(p(b)-p_a) - \Omega(a-b)\right] > (p(b)-p_a)\left[(p(c)-p_a) - \Omega(a-c)\right]$$

$$\iff \frac{(p(b)-p_a)}{a-b} > \frac{(p(c)-p_a)}{a-c}$$

where the last inequality follows from the result above—because $\frac{(p(c)-p_a)}{a-c}$ is a convex combination of $\frac{(p(b)-p_a)}{a-b}$ and $\frac{(p(c)-p(b))}{b-c}$. Finally, we have $1 + rp(b) + \Omega rb < 1 + rp(c) + \Omega rc$, because

$$rp(b) + \Omega rb < rp(c) + \Omega rc$$

$$\iff \Omega(b-c) < p(c) - p(b),$$

where the last inequality follows from Lemma 2. The result follows. □

We conclude with the key property for identification, which establishes that different pairs of r and Ω have different implications for willingness to pay (provided there are at least three deductible options on the menu).

Property 3. *If $p(b) - p_a$ is the same for $(r, \Omega(\mu))$ and $(r', \Omega(\mu)')$ with $r < r'$ (and thus $\Omega(\mu) > \Omega(\mu)'$), then $p(c) - p(b)$ is larger for $(r, \Omega(\mu))$ than for $(r', \Omega(\mu)')$.*

Proof. For a fixed $p(b)$, define $\Omega^b(r)$ by $V(p(b), b, r, \Omega^b(r)) = 0$, so that any pair $(r, \Omega^b(r))$ yields the same $p(b)$ and thus the same $p(b) - p_a$. Then

$$\frac{d\Omega^b}{dr} = \frac{-\frac{\partial V}{\partial r}}{\frac{\partial V}{\partial \Omega}} = -\frac{\frac{1}{r}\left[(p(b) - p_a) - \Omega^b(r)(a-b)\right]}{\frac{1}{\Omega^b(r)}\left[(p(b) - p_a) + \frac{r}{2}\left(p(b)^2 - p_a^2\right)\right]}.$$

Next, define $\check{p}_c(r)$ by $V(\check{p}_c(r), c, r, \Omega^b(r)) = 0$, so that $\check{p}_c(r)$ is the $p(c)$ associated with pair $(r, \Omega^b(r))$. The goal is to show that $d\check{p}_c(r)/dr < 0$, from which the result follows. Differentiating $V(\check{p}_c(r), c, r, \Omega^b(r)) = 0$ yields

$$\frac{d\left[V(\check{p}_c(r), c, r, \Omega^b(r))\right]}{dr} = \frac{\partial V}{\partial p}\frac{d\check{p}_c(r)}{dr} + \frac{\partial V}{\partial r} + \frac{\partial V}{\partial \Omega}\frac{d\Omega^b}{dr} = 0.$$

Note that

$$\begin{aligned}\frac{\partial V}{\partial r} + \frac{\partial V}{\partial \Omega}\frac{d\Omega^b}{dr} &= -\frac{1}{r}\left[(\check{p}_c(r) - p_a) - \Omega^b(r)(a-c)\right] \\ &\quad + \frac{1}{\Omega^b(r)}\left[(\check{p}_c(r) - p_a) + \frac{r}{2}(\check{p}_c(r)^2 - p_a^2)\right]\frac{\frac{1}{r}\left[(p(b) - p_a) - \Omega^b(r)(a-b)\right]}{\frac{1}{\Omega^b(r)}\left[(p(b) - p_a) + \frac{r}{2}\left(p(b)^2 - p_a^2\right)\right]}.\end{aligned}$$

We have

$$\left[(\check{p}_c(r) - p_a) - \Omega^b(r)(a-c)\right] < \left(\frac{\check{p}_c(r) - p_a}{p(b) - p_a}\right)\left[(p(b) - p_a) - \Omega^b(r)(a-b)\right]$$

as in the proof of Property 2. In addition,

$$\frac{\left[(\check{p}_c(r) - p_a) + \frac{r}{2}(\check{p}_c(r)^2 - p_a^2)\right]}{\left[(p(b) - p_a) + \frac{r}{2}\left(p(b)^2 - p_a^2\right)\right]} > \left(\frac{\check{p}_c(r) - p_a}{p(b) - p_a}\right)$$

because

$$\begin{aligned}&(p(b) - p_a)(\check{p}_c(r)^2 - p_a^2) > (\check{p}_c(r) - p_a)\left(p(b)^2 - p_a^2\right) \\ \Longleftrightarrow\ &(p(b) - p_a)(\check{p}_c(r) - p_a)(\check{p}_c(r) + p_a) > (\check{p}_c(r) - p_a)(p(b) - p_a)\left(p(b) + p_a\right) \\ \Longleftrightarrow\ &\check{p}_c(r) > p(b),\end{aligned}$$

where the last inequality follows from Lemma 1. Together, these imply $\frac{\partial V}{\partial r} + \frac{\partial V}{\partial \Omega}\frac{d\Omega^b}{dr} > 0$, and therefore $\frac{\partial V}{\partial p}\frac{d\check{p}_c(r)}{dr} < 0$. Hence, $\frac{\partial V}{\partial p} > 0$ implies $\frac{d\check{p}_c(r)}{dr} < 0$, and the result follows. □

C Sources of Probability Distortions

Throughout our main analysis, we assume that utility of deductible lottery L_d is given by

$$U(L_d) = (1 - \Omega(\mu))u(w - p_d) + \Omega(\mu)u(w - p_d - d), \quad \text{(A.3)}$$

where $\Omega(\mu)$ reflects probability distortions. As we discuss in Section III.C, there are a number of possible sources of probability distortions. In this section, we describe the details of models that can generate the probability distortions in equation (A.3).

C.1 Probability Weighting

One potential source of probability distortions is probability weighting, whereby probabilities are transformed into decision weights. Under a probability weighting model, and adopting the rank-dependent approach of Quiggin (1982), the utility of deductible lottery L_d is

$$U(L_d) = (1 - \pi(\mu))u(w - p_d) + \pi(\mu)u(w - p_d - d), \quad \text{(A.4)}$$

where $\pi(\mu)$ is the probability weighting function. Clearly, equation (A.4) is equivalent to equation (A.3) with $\Omega(\mu) = \pi(\mu)$.

Over the years, several functional forms for $\pi(\mu)$ have been proposed. A seminal paper in the literature is Kahneman and Tversky (1979), which proposes that the probability weighting function should exhibit (i) overweighting of small probabilities, (ii) underweighting of large probabilities, (iii) some insensitivity to probability changes (slope less than one), and (iv) discontinuities at $\mu = 0$ and $\mu = 1$. Later papers in the literature suggest functional forms that eliminate feature (iv) (e.g., Tversky and Kahneman 1992; Lattimore et al. 1992; Prelec 1998). These all entail an oversensitivity to probability changes (slope greater than one) for μ close to zero or one, where "close" typically (for the authors' suggested parameter values) includes the bulk of the claim probabilities in our data—e.g., using the one-parameter functional form proposed by Prelec (1998) and his suggested parameter value of $\alpha = 0.65$, $\pi'(\mu) > 1.0$ for $\mu < 0.069$, $\pi'(\mu) > 1.5$ for $\mu < 0.028$, and $\pi'(\mu) > 2.0$ for $\mu < 0.015$. As we discuss in Section III.C, insofar as our estimated $\Omega(\mu)$ reflects probability weighting, it is more in line with the function originally posited by Kahneman and Tversky (1979).

C.2 Kőszegi-Rabin Loss Aversion

Another possible source of probability distortions is loss aversion. The original, "status quo" loss aversion proposed by Kahneman and Tversky (1979)—wherein gains and losses are defined relative to initial wealth—cannot explain aversion to risk in the context of insurance deductible choices because all outcomes are losses relative to initial wealth. More recently, however, Kőszegi and Rabin (2007) and Sydnor (2010) have suggested that a form of "rational expectations" loss aversion proposed by Kőszegi and Rabin (2006)—wherein gains and losses are defined relative to expectations about outcomes—can explain the aversion to risk manifested in insurance deductible choices.

In the Kőszegi-Rabin (KR) model, the utility from choosing lottery $Y \equiv (y_n, q_n)_{n=1}^N$ given a reference lottery $\tilde{Y} \equiv (\tilde{y}_m, \tilde{q}_m)_{m=1}^M$ is

$$V(Y|\tilde{Y}) \equiv \sum_{n=1}^{N} \sum_{m=1}^{M} q_n \tilde{q}_m \left[u(y_n) + v(y_n|\tilde{y}_m) \right].$$

The function u represents standard "intrinsic" utility defined over final wealth states, just as in the expected utility model. The function v represents "gain-loss" utility that results from experiencing gains or losses relative to the reference point. For v, KR use

$$v(y|\tilde{y}) = \begin{cases} \eta \left[u(y) - u(\tilde{y}) \right] & \text{if } u(y) > u(\tilde{y}) \\ \eta\lambda \left[u(y) - u(\tilde{y}) \right] & \text{if } u(y) \leq u(\tilde{y}) \end{cases}.$$

In this formulation, the magnitude of gain-loss utility is determined by the intrinsic utility gain or loss relative to consuming the reference point. Moreover, gain-loss utility takes a two-part linear form, where $\eta \geq 0$ captures the importance of gain-loss utility relative to intrinsic utility and $\lambda \geq 1$ captures loss aversion. The model reduces to expected utility when $\eta = 0$ or $\lambda = 1$.

KR propose that the reference lottery equals recent expectations about outcomes—i.e., if a household expects to face lottery $\tilde{Y}$, then its reference lottery becomes $\tilde{Y}$. However, because situations vary in terms of when a household deliberates about its choices and when it commits to its choices, KR offer a number of solution concepts for the determination of the reference lottery. For insurance applications, KR suggest a "choice-acclimating personal equilibrium" (CPE). Formally:

Definition (CPE)**.** *Given a choice set $\mathcal{Y}$, a lottery $Y \in \mathcal{Y}$ is a* choice-acclimating personal equilibrium *if for all $Y' \in \mathcal{Y}$, $V(Y|Y) \geq V(Y'|Y')$.*

In a CPE, a household's reference lottery corresponds to its choice. KR argue that CPE is appropriate in situations where the household commits to a choice well in advance of the resolution of uncertainty, and thus it knows that by the time the uncertainty is resolved and it experiences utility, it will have become accustomed to its choice and hence expect the lottery induced by its choice. In particular, KR suggest that CPE is the appropriate solution concept for insurance applications.

Under the KR model with CPE, the utility to the household from choosing deductible lottery $L_d = (-p_d, 1-\mu; -p_d - d, \mu)$ is

$$\begin{aligned} U(L_d) = V(L_d|L_d) &= (1-\mu)u(w-p_d) + \mu u(w-p_d-d) \\ &\quad -\eta(\lambda-1)(1-\mu)\mu[u(w-p_d) - u(w-p_d-d)]. \end{aligned} \tag{A.5}$$

From equation (A.5), it is clear that one can not separately identify the parameters η and λ, and thus we focus on the product $\eta(\lambda-1) \equiv \Lambda$. Substituting Λ into equation (A.5) yields equation (5) in Section III.C.

C.3 Gul Disappointment Aversion

Probability distortions also can arise from disappointment aversion. The concept of disappointment aversion was proposed by Bell (1985) and further developed by Loomes and Sugden (1986) and Gul (1991). The basic idea is that a person is disappointed (or elated) if the outcome of a lottery is worse (or better) than "expected." The approaches differ in terms of the nature of the disutility from disappointment, and in terms of the definition of what is "expected."

Here, we follow the approach of Gul (1991), in which disutility arises when the outcome of a lottery is less than the certainty equivalent for the lottery. For deductible lotteries, an intuitive way to express this model is that $U(L_d)$ is the $u(z)$ such that

$$u(z) = [(1-\mu)u(w-p_d) + \mu u(w-p_d-d)] - \beta\mu\,[u(z) - u(w-p_d-d)]\,.$$

In this formulation, z represents the certainty equivalent. The first bracketed term is the standard expected utility of L_d. The second term reflects the expected disutility from disappointment that arises when the outcome is less than the certainty equivalent (which occurs in the event of a claim). The parameter β captures the magnitude of disappointment aversion, where the model reduces to expected utility for $\beta = 0$. We can rearrange this equation to yield

$$U(L_d) = \left(1 - \frac{\mu(1+\beta)}{1+\beta\mu}\right) u(w-p_d) + \left(\frac{\mu(1+\beta)}{1+\beta\mu}\right) u(w-p_d-d), \tag{A.6}$$

which is equivalent to equation (6) in Section III.C.[2]

As illustrated in Figure 3 in Section III.C, Gul disappointment aversion and KR loss aversion with CPE generate qualitatively similar probability distortions for claim probabilities in the range of our data. This follows from the fact that these models are quite similar to each other. Both models assume that a household experiences a form of "gain-loss utility" that depends on how its realized outcome compares to a reference point which is determined by the household's choice. The models differ only in the nature of the reference point associated with each choice—in Gul disappointment aversion the reference point is the certainty equivalent of the chosen lottery, whereas in KR loss aversion it is the chosen lottery itself. Bell (1985) and Loomes and Sugden (1986) differ from Gul (1991) in (effectively) assuming that the reference point is the standard expected utility of the chosen lottery. Combining that formulation with a two-part linear disappointment/elation function would yield a model that is equivalent to KR loss aversion with CPE.

C.4 Combinations of Sources

Finally, we consider combinations of sources. In particular, we focus on combining a probability weighting function $\pi(\mu)$ with either KR loss aversion or Gul disappointment aversion.

When we combine probability weighting and KR loss aversion, an issue arises: Given a reference lottery $\tilde{Y} \equiv (\tilde{y}_m, \tilde{q}_m)_{m=1}^{M}$, should the utility comparisons in $V(Y|\tilde{Y})$ use the probabilities $\tilde{q}_m$ for the comparison weights, or should they use the decision weights $\pi(\tilde{q}_m)$? KR offer no guidance on this modeling choice, as they abstract from nonlinear decision weights. To our minds, it seems natural to assume that households treat the chosen lottery and the reference lottery symmetrically. Accordingly, we assume that the decision weights are the same for the chosen lottery and the reference lottery. Under this assumption, we can rewrite equation (A.5) as

$$\begin{aligned} U(L_d) = V(L_d|L_d) &= (1-\pi(\mu))u(w-p_d) + \pi(\mu)u(w-p_d-d) \\ &\quad -\Lambda(1-\pi(\mu))\pi(\mu)[u(w-p_d) - u(w-p_d-d)], \end{aligned}$$

which is equivalent to equation (A.3) with $\Omega(\mu) = \pi(\mu)[1+\Lambda(1-\pi(\mu))]$. From this equation, it is clear that, unless we impose strong functional form assumptions on $\pi(\mu)$, we cannot separately identify Λ and $\pi(\mu)$. Rather, the best we can do is to derive, for various values of Λ, an implied probability weighting function $\pi(\mu)$. Panel (c) of Figure 3 in Section III.C performs this exercise.

[2] Note that equation (A.6) is equivalent to the equation at the top of page 677 in Gul (1991).

Combining probability weighting and Gul disappointment aversion is more straightforward, as we merely replace μ with $\pi(\mu)$ in equation (A.6). With this substitution, equation (A.6) is equivalent to equation (A.3) with $\Omega(\mu) = \pi(\mu)(1+\beta)/(1+\beta\pi(\mu))$. From this it is clear that, unless we impose strong functional form assumptions on $\pi(\mu)$, we cannot separately identify β and $\pi(\mu)$. Rather, the best we can do is to derive, for various values of β, an implied probability weighting function $\pi(\mu)$. The resulting figure would be very similar to panel (c) of Figure 3 in Section III.C.

D MCMC Procedure: Hierarchical Bayes for Mixed Logit

In Models 3 and 4, we allow for unobserved heterogeneity in r_i and $\Omega_i(\mu)$. In particular, we assume

$$\ln r_i = \beta_r Z_i + \xi_{r,i} \qquad \text{and} \qquad \ln \Omega_i(\mu) = \beta_{\Omega,1} Z_i + \left(\beta_{\Omega,2} Z_i\right)\mu + \left(\beta_{\Omega,3} Z_i\right)\mu^2 + \xi_{\Omega,i},$$

where $\left(\xi_{r,i}, \xi_{\Omega,i}\right) \overset{iid}{\sim} N(0,\Phi)$. The utility from deductible $d \in \mathcal{D}$ is given by

$$\mathcal{U}(d) \equiv U(L_d) + \varepsilon_d,$$

where ε_d follows a type 1 extreme value distribution with scale parameter σ. Assuming $U(L_d)$ is specified by equation (2) in Section II.A, we have

$$\mathcal{U}(d) = -\left[p_d + \Omega(\mu)d\right] - \frac{r}{2}\left[\left(1-\Omega(\mu)\right)(p_d)^2 + \Omega(\mu)\left(p_d+d\right)^2\right] + \varepsilon_d,$$

which can be re-written as

$$\mathcal{U}(d) = -p_d - \Omega(\mu)d - \frac{r}{2}(p_d)^2 - \frac{r}{2}\Omega(\mu)\left[(p_d+d)^2 - (p_d)^2\right] + \varepsilon_d. \tag{A.7}$$

Hence, a household i chooses deductible d in coverage j when $\mathcal{U}_{ij}(d) > \mathcal{U}_{ij}(d')$ for all $d' \neq d$, and thus the probability that household i chooses deductible d in coverage j conditional on the observables *and* conditional on $\left(\xi_{r,i}, \xi_{\Omega,i}\right)$ is

$$\begin{aligned} \mathcal{P}(D_{ij}|\xi_{r,i},\xi_{\Omega,i}) &\equiv \Pr\left(D_{ij} = d | P_{ij}, \widehat{\mu}_{ij}, Z_i, \xi_{r,i}, \xi_{\Omega,i}\right) \qquad \text{(A.8)} \\ &= \Pr\left(\varepsilon_{d'} - \varepsilon_d < U(L_d) - U(L_{d'}) \text{ for all } d' \neq d | P_{ij}, \widehat{\mu}_{ij}, Z_i, \xi_{r,i}, \xi_{\Omega,i}\right) \\ &= \frac{\exp\left(U(L_d)/\sigma\right)}{\sum_{d'\in\mathcal{D}} \exp\left(U(L_{d'})/\sigma\right)}. \end{aligned}$$

The unconditional probability that household i chooses a triple (d_L, d_M, d_H) is given by the integral of the product of equation (A.8) across contexts, against the distribution of $(\xi_{r,i}, \xi_{\Omega,i})$:

$$\int \prod_{k=L,M,H} \mathcal{P}(D_{ik} = d_k | \xi_{r,i}, \xi_{\Omega,i}) \phi(\xi_{r,i}, \xi_{\Omega,i}; 0, \Phi) d\xi_{r,i} d\xi_{\Omega,i}, \tag{A.9}$$

where $\phi(\xi_{r,i}, \xi_{\Omega,i}; 0, \Phi)$ is the bivariate normal density. We observe data $\{D_{ij}, \Gamma_{ij}\}$, where D_{ij} is household i's deductible choice for coverage j and $\Gamma_{ij} \equiv (Z_i, \widehat{\mu}_{ij}, P_{ij})$. In Γ_{ij}, Z_i is a vector of household characteristics, $\widehat{\mu}_{ij}$ is household i's predicted claim probability for coverage j, and P_{ij} denotes household i's menu of premium-deductible pairs for coverage j. The set of fixed parameters (fixed coefficients) to be estimated is $\theta \equiv (\beta_r, \beta_{\Omega,1}, \beta_{\Omega,2}, \beta_{\Omega,3}, \sigma_L, \sigma_M, \sigma_H)$. Additionally, we estimate Φ, the covariance matrix of the heterogeneity terms (random coefficients) $\vec{\xi} \equiv (\xi_{r,i}, \xi_{\Omega,i})$. For a given claim probability μ, equation (A.7) can be seen as a linear function in the explanatory variables $\{p_d, d, (p_d)^2, [(p_d + d)^2 - (p_d)^2]\}$ and their corresponding coefficients, with the coefficient on the last variable constrained to be the product of the second and third coefficients. The first coefficient being one is a scale normalization, since the variance of the choice noise is unconstrained. Hence, identification follows from standard arguments for mixed logit models and our discussion in Section II.C, provided there is sufficient variation in p and d.[3]

Notice that maximum likelihood estimation cannot be directly carried out, because the integral in equation (A.9) cannot be calculated analytically. Hence, we employ Markov Chain Monte Carlo (MCMC) methods, which avoid integration entirely. Intuitively, in MCMC the integration over unobserved heterogeneity terms is substituted by augmenting the data with draws of the unobserved heterogeneity terms, and updating the distribution from which the draws are taken based on likelihood improvement criteria. The result of the MCMC procedure is the joint posterior distribution of (θ, Φ). By the Bernstein-von Mises theorem, the mean of the posterior is an estimator that, in frequentist terms, is asymptotically equivalent to the (computationally infeasible) maximum likelihood estimator (Train 2009, ch. 12).

The MCMC procedure that we use is built on the Gibbs sampler and Metropolis-Hastings algorithm. It requires the choice of priors, which then are combined with the data, and in particular the observed choices, to obtain the posterior distribution of (θ, Φ). We set:

[3] Furthermore, when we specify $\Omega(\mu) = a + b\mu \exp(\xi_b)$ in Section V.A, one can immediately see that the model continues to be linear in the explanatory variables $\{p_d, d, (p_d)^2, [(p_d+d)^2 - (p_d)^2], \mu[(p_d+d)^2 - (p_d)^2]\}$ and their corresponding coefficients, with some constraints between the coefficients. Hence, for this case identification also follows from standard arguments for mixed logit models and our discussion in Section II.C, provided there is sufficient variation in p and d and μ.

Priors

1. The fixed coefficients β_r, $\beta_{\Omega,1}$, $\beta_{\Omega,2}$, and $\beta_{\Omega,3}$ are assumed to have a normal distribution with a diffused prior. The fixed coefficients σ_L, σ_M, and σ_H are assumed to have a lognormal distribution with a diffused prior.

2. The prior on the matrix Φ is inverted Wishart with 2 degrees of freedom and scale matrix $I(2)$.

Initial Values

1. All unobserved heterogeneity terms (random coefficients) are set to zero.

2. The initial value of θ is set to that estimated under the assumption of no unobserved heterogeneity.[4]

3. The initial draw of Φ is $2 \cdot I(2)$.

The posterior distribution of (θ, Φ) conditional on Γ_{ij} is obtained via simulations, using Gibbs sampling and Metropolis-Hastings algorithm, as detailed below.

Simulation: Gibbs Sampling

1. Draws of $\Phi \mid \vec{\xi}_i, \forall i$: The posterior for Φ is inverted Wishart with $2 + N$ degrees of freedom and scale matrix $(2 \cdot I(2) + N \cdot S)/(2 + N)$, where N is the sample size and S is the sample variance of $\vec{\xi}$. A draw from inverted Wishart is obtained according to standard techniques (Train 2009, § 12.5.2).

2. Draws of $\vec{\xi}_i, \forall i \mid \theta, \Phi$: The posterior for each household's $\vec{\xi}$, conditional on θ, on its choices D, and on other observables Γ is

$$K(\vec{\xi}|\theta, \Phi, D, \Gamma) \propto L(D|\vec{\xi}, \theta, \Gamma) \cdot \phi(\vec{\xi}, 0, \Phi),$$

where

$$L(D|\vec{\xi}, \theta, \Gamma) = \prod_{k=L,M,H} \mathcal{P}(D_{ik} = d_k|\xi_{r,i}, \xi_{\Omega,i})$$

is the product of the three choice probabilities (across coverages) and $\phi(\vec{\xi}, 0, \Phi)$ is the bivariate normal density. Draws from this posterior are obtained with one step of the Metropolis-Hastings algorithm described below.

[4] For Model 3, the simulation was also run with different starting values for θ, including the one that mimics the case when the model is estimated with no probability distortions.

3. Draws of $\theta \mid \vec{\xi}_i, \forall i$: The posterior for the fixed coefficients θ, conditional on the draws of the random coefficients, $\vec{\xi}_i, \forall i$, on choices D, and on other observables Γ is

$$K(\theta|, \vec{\xi}_i, \forall i, \Phi, D, \Gamma) \propto \prod_i L(D|\vec{\xi}_i, \theta, \Gamma).$$

   Draws from this posterior are obtained with one step of the Metropolis Hastings algorithm on the pooled data, as described below.

Metropolis-Hastings Algorithm for Step 2

For each household i, and a given an initial draw $\vec{\xi}_i^0$:

- Draw a bivariate standard normal vector $\vec{\eta}$.
- Create a trial vector $\vec{\xi}_i^1 = \vec{\xi}_i^0 + \rho L \vec{\eta}$, where ρ is a positive scalar and L is the (lower triangular) Choleski factor of Φ.
- Draw a standard uniform variable κ.
- Calculate the ratio

$$F = \frac{L(D|\vec{\xi}^1, \theta, \Gamma) \cdot \phi(\vec{\xi}^1, 0, \Phi)}{L(D|\vec{\xi}^0, \theta, \Gamma) \cdot \phi(\vec{\xi}^0, 0, \Phi)}.$$

- If $\kappa \leq F$, accept the new value of $\vec{\xi}_i^1$. Otherwise, reset $\vec{\xi}_i^1 = \vec{\xi}_i^0$.
- Repeat.

Metropolis-Hastings Algorithm for Step 3

For a given initial draw θ^0:

- Draw a standard normal vector $\vec{\eta}$ of the same dimensionality as θ.
- Create a trial vector $\theta^1 = \theta^0 + \delta\vec{\eta}$, where δ is a positive scalar.
- Draw a standard uniform variable κ.
- Calculate the ratio

$$F = \frac{\prod_i L(D|\vec{\xi}_i, \theta^1, \Gamma)}{\prod_i L(D|\vec{\xi}_i, \theta^0, \Gamma)}.$$

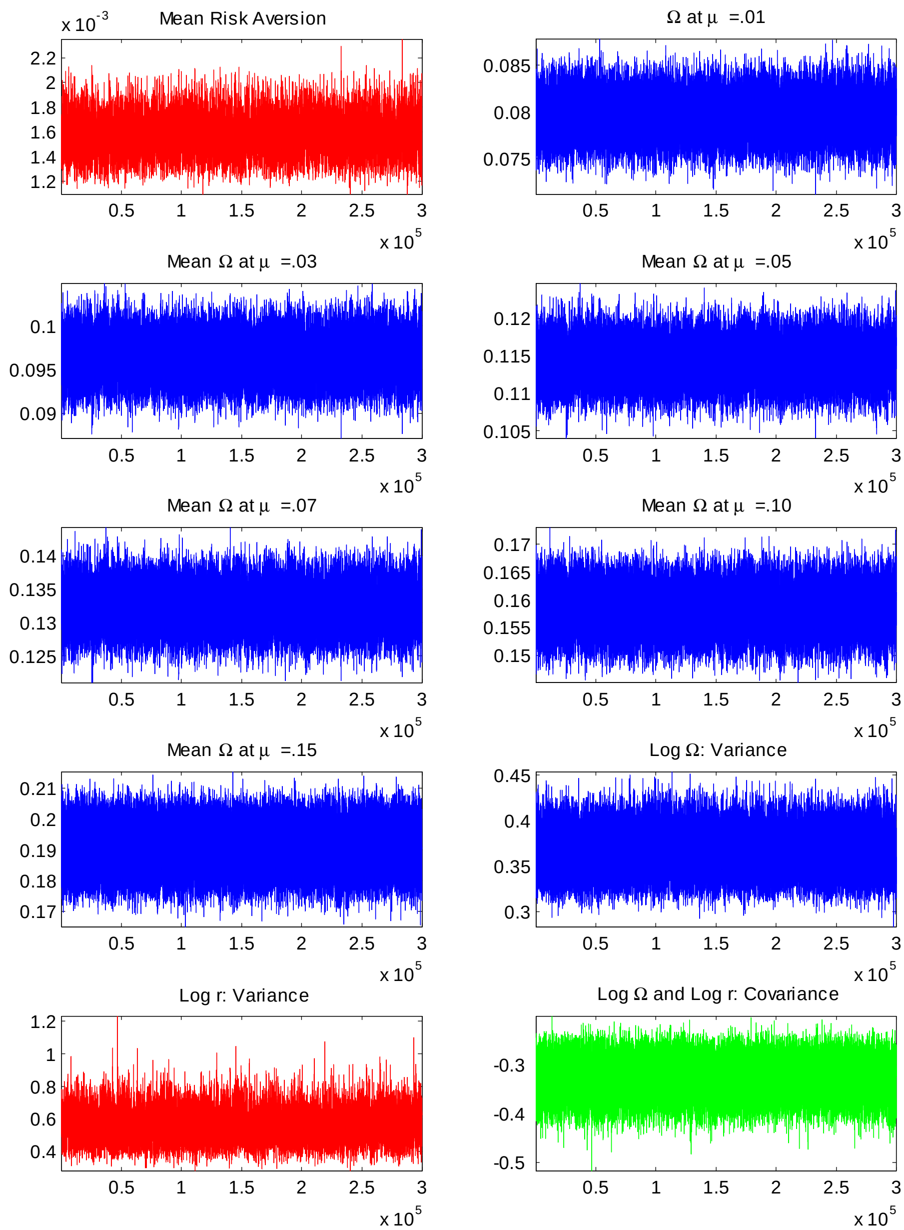
- If $\kappa \leq F$, accept new value of θ^1. Otherwise, reset $\theta^1 = \theta^0$.

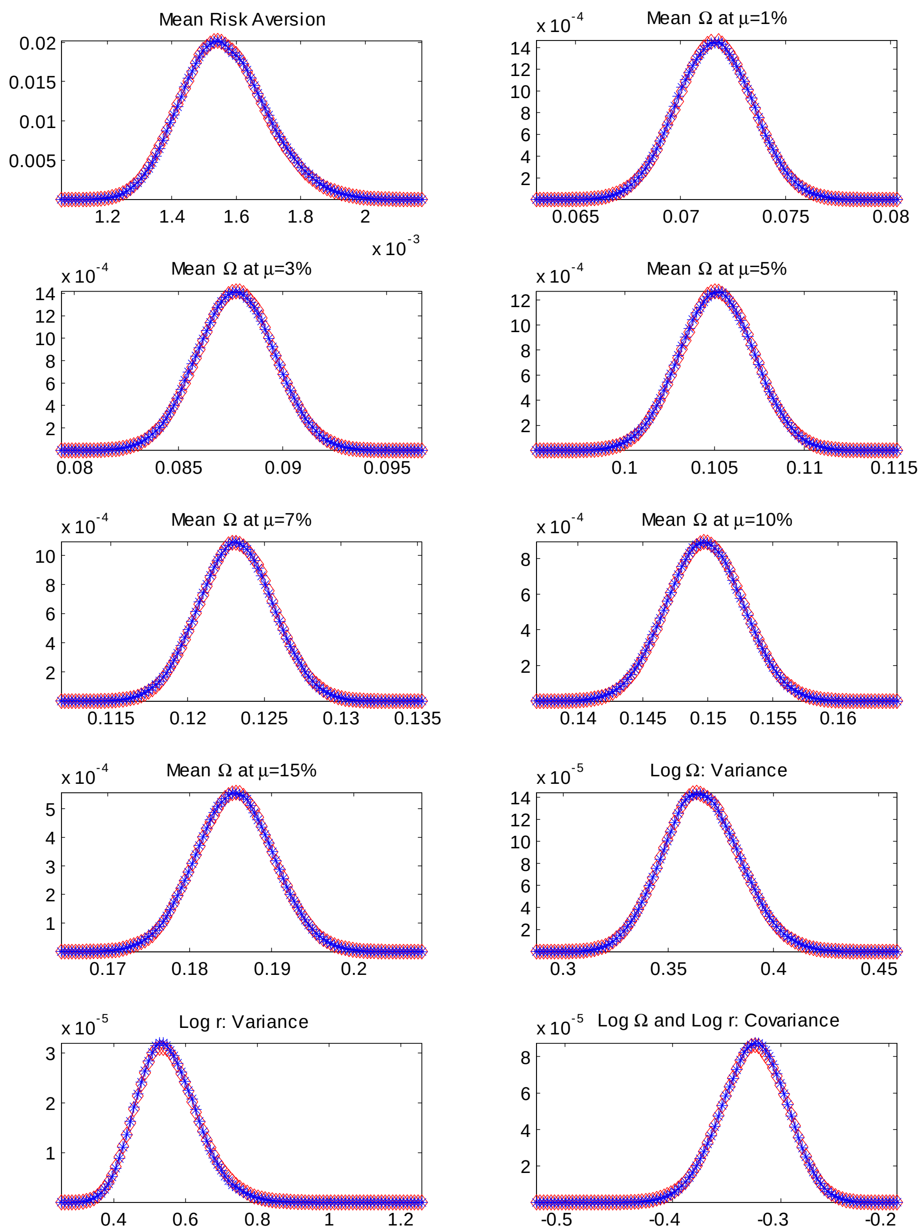
- Repeat.

The Run and Convergence

When running an MCMC procedure, one has to worry that the chain has run sufficiently long to achieve convergence. For Model 3 we run the chain 3,100,000 times.[5] We drop the first 100,000 draws as "burn-in" and retain only every 10th draw as "thinning." The first figure below shows the trace plot for selected objects of interest. The second figure compares the densities of the same objects for the first half of the chain versus the second half.

[5] For Model 4, we run the chain 10,100,000 times.


Mean Risk Aversion
x 10^-3
Ω at μ =.01
Mean Ω at μ =.03
Mean Ω at μ =.05
Mean Ω at μ =.07
Mean Ω at μ =.10
Mean Ω at μ =.15
Log Ω: Variance
Log r: Variance
Log Ω and Log r: Covariance
x 10^5


Mean Risk Aversion
x 10^-3
Mean Ω at μ=1%
x 10^-4
Mean Ω at μ=3%
x 10^-4
Mean Ω at μ=5%
x 10^-4
Mean Ω at μ=7%
x 10^-4
Mean Ω at μ=10%
x 10^-4
Mean Ω at μ=15%
x 10^-4
Log Ω: Variance
x 10^-5
Log r: Variance
x 10^-5
Log Ω and Log r: Covariance
x 10^-5

In order to more formally assess convergence of our chains, we employ a battery of statistical tests.[6] We note that in our simulation, as is commonly the case with Metropolis-Hastings based simulations, the draws are autocorrelated. To this end, we use the Raftery-Lewis diagnostics on each chain (i.e., for all parameters) to determine the "burn-in" and "thinning."[7] We then use the second half of the chain, further thinned by a factor of ten, that conforms with the Raftery-Lewis diagnostics and passes the Geweke's chi-squared test under the assumption of iid draws in the chain for each variable.[8]

Core Sample (4170 Households)

	Geweke test (iid)			**Autocorrelations**			
	Mean	**NSE**	**χ^2 pr**	**Lag 1**	**Lag 5**	**Lag 10**	**Lag 50**
Log r	-6.74	0.00	0.21	0.83	0.47	0.24	-0.01
Log $\Omega(\mu)$: constant	-2.82	0.00	0.24	0.59	0.23	0.10	-0.01
Log $\Omega(\mu)$: linear	10.72	0.01	0.52	0.75	0.25	0.06	-0.04
Log $\Omega(\mu)$: quadratic	-28.07	0.03	0.32	0.68	0.24	0.05	-0.03
σ_L	17.14	0.01	0.65	0.32	0.15	0.05	-0.01
σ_M	10.16	0.00	0.38	0.21	0.08	0.03	0.02
σ_H	95.64	0.07	0.44	0.67	0.38	0.19	-0.01
Φ_r	0.55	0.00	0.77	0.72	0.33	0.17	-0.01
Φ_Ω	0.37	0.00	0.11	0.25	0.13	0.06	0.00
$\Phi_{r,\Omega}$	-0.33	0.00	0.49	0.36	0.09	0.05	-0.01

Raftery-Lewis diagnostics for each parameter chain: I-stat = 3.86.

[6] We use a version of the CODA package adapted for MATLAB by James P. LeSage.

[7] The Raftery-Lewis diagnostic is a run length control diagnostic based on a criterion of accuracy of estimation of the quantile $q = 0.025$. The number of iterations required to estimate the quantile q to within an accuracy of $\pm\ r = 0.01$ with probability $s = 0.95$ is calculated. Separate calculations are performed for each variable within each chain.

[8] The Geweke test is based on the idea that if the sample of draws has attained an equilibrium state, the means of the first 20 percent of the sample of draws versus the last 50 percent of the sample should be roughly the same.

E Additional Sensitivity Checks

In this section, we report the results of several additional sensitivity checks.

E.1 CARA Utility

In our analysis, we consider a second-order Taylor expansion of the utility function, and also CRRA utility. Here we take yet another approach: we assume constant absolute risk aversion (CARA) utility, $u(w) = -\exp(-rw)$. That is, we specify utility as

$$U(L_d) = \frac{EU(L_d)}{u'(w)} = (1 - \Omega(\mu)) \frac{-\exp(rp_d)}{r} + \Omega(\mu) \frac{-\exp(r(p_d + d))}{r},$$

which we note is independent of wealth w. When we estimate Model 2 with CARA utility, the main message is the same. The estimates for $\Omega(\mu)$ indicate similar probability distortions, albeit somewhat less pronounced than the benchmark, while the estimates for r are higher than the benchmark. See Table A.17.

E.2 Alternative Samples

In the core sample, we restrict attention to households who hold both auto and home policies and who first purchased their auto and home policies from the company in the same year, in either 2005 or 2006. Here we estimate Model 2 using two less restrictive samples: (1) households who hold auto policies and who first purchased their auto policies from the company in the same year, in either 2005 or 2006; and (2) households who hold both auto and home policies and who first purchased their auto and home policies from the company in the same year, in either 2004, 2005, or 2006. Again, the main message is the same. In both samples, the estimates for $\Omega(\mu)$ indicate probability distortions that are very similar to the benchmark. As for standard risk aversion, in sample 1 the estimates for r are higher than the benchmark estimates, while in sample 2 they are lower than the benchmark estimates. See Tables A.18 and A.19.

E.3 Restricted Menus

In our main analysis, we use the full menu of deductible options for each coverage, up to \$1000. In each coverage, however, the vast majority of households choose one of three deductibles: 92.3 percent of households choose a deductible of \$200, \$250, or \$500 in auto collision; 87.1 percent of households choose a deductible of \$200, \$250, or \$500 in auto comprehensive; and 97.5 percent of households choose a deductible of \$250, \$500, or \$1000 in home. Given these choice patterns, one might worry that households do not really consider the other deductible options, which could bias our estimates.[9] To address this concern, we estimate Model 2 when we restrict the menu of deductible options to $\{\$200, \$250, \$500\}$ for each auto coverage and to $\{\$250, \$500, \$1000\}$ for home coverage.[10] The estimates for $\Omega(\mu)$ indicate probability distortions that are similar to the benchmark. Indeed, the overweighting is more pronounced at high claim probabilities. The estimates for r are lower than the benchmark estimates. See Table A.20.

E.4 Alternative Error Structures

In our main analysis, we assume that the utility from every deductible $d \in \mathcal{D}$ is given by $\mathcal{U}(d) = U(L_d) + \varepsilon_d$, where ε_d is an iid Gumbel random variable. Here, we estimate Model 2 under two alternative assumptions (and, for computational and theoretical reasons, using the restricted menus from Section E.3 above): (A) we assume (as before) that the utility from every deductible $d \in \mathcal{D}$ is given by $\mathcal{U}(d) = U(L_d) + \varepsilon_d$, but we assume that ε_d is an iid normal random variable; and (B) we assume that the utility from the maximum deductible, D, is given by $\mathcal{U}(D) = U(L_D) + \varepsilon_D$, where ε_D is an iid normal random variable, but that the utility from the other deductibles are given by $\mathcal{U}(d) = U(L_d) + \zeta_d$, where $\zeta_d = -\varepsilon_D$ for the minimum deductible and $\zeta_d = 0$ for the intermediate deductible. Alternative A provides a check of the Gumbel error assumption. Alternative B adds a check of the iid assumption. More specifically, we consider alternative B to address concerns arising from the fact that in principle the iid assumption allows for nonmonotonic ranking of deductibles. Once again, the main message is the same. Under both alternatives, the estimates for $\Omega(\mu)$ indicate similar probability distortions, though generally somewhat more pronounced. In addition, under alternative A the estimates for r are lower than the benchmark estimates, while under alternative B they are higher than the benchmark estimates. See Tables A.21 and A.22.

F Appendix Tables and Figures

On the ensuing pages, we report Tables A.1 through A.22 and Figures A.1 through A.4.

[9] For instance, when a household chooses a \$250 deductible in home, we are using the fact that it did not choose a \$100 deductible to infer an upper bound on its aversion to risk. But if the household in fact does not even consider the \$100 deductible as an option, our inference would be invalid. Similarly, when a household chooses a \$500 deductible in auto comprehensive, we are using the fact that it did not choose a \$1000 deductible to infer a lower bound on its aversion to risk. Again, if the household in fact does not even consider the \$1000 deductible as an option, our inference would be invalid.

[10] In each case, if a household's actual deductible choice is outside the restricted menu, we assign to the household the deductible option from the restricted menu that is closest to their actual deductible choice. In this respect, we follow Cohen and Einav (2007).

Table A.1: Summary of Premium Menus - Auto Collision

Core sample (4170 households)

	Deductible choice					
	\$100	**\$200**	**\$250**	**\$500**	**\$1000**	**All**
Mean annual premium for coverage with \$500 deductible	110	129	146	189	255	180
Standard deviation	54	54	66	96	168	100
Mean cost of decreasing deductible from \$500 to \$250	33	38	44	57	77	54
Standard deviation	17	17	20	29	52	31
Mean savings from increasing deductible from \$500 to \$1000	24	29	33	43	58	41
Standard deviation	12	12	15	22	39	23
Number of households	42	559	467	2822	280	4170

Note: All values in dollars, except number of households.

Table A.2: Summary of Premium Menus - Auto Comprehensive

Core sample (4170 households)

	Deductible choice						
	\$50	**\$100**	**\$200**	**\$250**	**\$500**	**\$1000**	**All**
Mean annual premium for coverage with \$500 deductible	61	70	92	98	136	258	115
Standard deviation	27	33	43	41	71	247	81
Mean cost of decreasing deductible from \$500 to \$250	16	18	24	26	36	68	30
Standard deviation	7	9	11	11	19	66	22
Mean savings from increasing deductible from \$500 to \$1000	12	14	18	19	27	51	23
Standard deviation	5	7	9	8	14	49	16
Number of households	216	171	1397	440	1795	151	4170

Note: All values in dollars, except number of households.

Table A.3: Summary of Premium Menus - Home

Core sample (4170 households)

	Deductible choice						
	\$100	**\$250**	**\$500**	**\$1000**	**\$2500**	**\$5000**	**All**
Mean annual premium for coverage with \$500 deductible	366	520	631	972	2218	3366	679
Standard deviation	113	218	308	593	2289	1808	519
Mean cost of decreasing deductible from \$500 to \$250	31	42	52	80	183	275	56
Standard deviation	6	18	26	48	201	140	43
Mean savings from increasing deductible from \$500 to \$1000	41	57	69	107	244	368	74
Standard deviation	8	23	34	64	268	188	58
Number of households	36	1239	2166	664	50	15	4170

Note: All values in dollars, except number of households.

Table A.4: Claim Rate Regressions - Auto

Poisson panel regression model with random effects

Full data set (1,348,020 household-year records)

	Collision			Comprehensive		
	Coefficient		Standard error	Coefficient		Standard error
Constant	-6.7646	**	0.0616	-7.9277	**	0.1057
Driver 2 Indicator	-0.0485		0.0593	-0.3542	**	0.1022
Driver 3+ Indicator	0.3215	**	0.0733	-0.1261		0.1201
Vehicle 2 Indicator	0.5991	**	0.0466	0.6502	**	0.0782
Vehicle 3+ Indicator	0.7312	**	0.0596	0.8766	**	0.0937
Young Driver	-0.0058		0.0296	0.0895	**	0.0453
Driver 1 Age	-0.0210	**	0.0015	0.0113	**	0.0029
Driver 1 Age Squared	0.0002	**	0.0000	-0.0002	**	0.0000
Driver 1 Female	0.1040	**	0.0093	-0.0672	**	0.0168
Driver 1 Married	0.0630	**	0.0111	0.0640	**	0.0201
Driver 1 Divorced	0.0186		0.0141	0.0914	**	0.0247
Driver 1 Separated	0.0392		0.0256	0.0791		0.0428
Driver 1 Single	.		.	.		.
Driver 1 Widowed	0.0031		0.0160	-0.0170		0.0335
Vehicle 1 Age	-0.0354	**	0.0019	-0.0286	**	0.0030
Vehicle 1 Age Squared	-0.0006	**	0.0001	0.0000		0.0002
Vehicle 1 Business	.		.	.		.
Vehicle 1 Farm	-0.2575	**	0.0872	0.0206		0.1194
Vehicle 1 Pleasure	-0.1094	**	0.0306	-0.1118	**	0.0526
Vehicle 1 Work	-0.0831	**	0.0304	-0.0620		0.0523
Vehicle 1 Passive Restraint	-0.1087	**	0.0239	-0.0858	**	0.0352
Vehicle 1 Anti-Theft	0.0754	**	0.0078	0.0735	**	0.0136
Vehicle 1 Anti-Lock	0.0581	**	0.0080	0.0729	**	0.0139
Driver 2 Age	0.0115	**	0.0024	0.0181	**	0.0042
Driver 2 Age Squared	-0.0001	**	0.0000	-0.0001	**	0.0000
Driver 2 Female	0.1204	**	0.0151	-0.0376		0.0257
Driver 2 Married	-0.0835	**	0.0191	-0.0408		0.0326
Driver 2 Divorced	-0.1579		0.1027	-0.1347		0.1636
Driver 2 Separated	0.0254		0.2130	0.1796		0.3226
Driver 2 Single	.		.	.		.
Driver 2 Widowed	-0.0802		0.1383	-1.1835	**	0.3864
Vehicle 2 Age	-0.0332	**	0.0016	-0.0229	**	0.0027
Vehicle 2 Age Squared	0.0004	**	0.0001	0.0002	**	0.0001
Vehicle 2 Business	.		.	.		.
Vehicle 2 Farm	-0.1703		0.1056	-0.1345		0.1500
Vehicle 2 Pleasure	-0.1805	**	0.0380	-0.0563		0.0663
Vehicle 2 Work	-0.1670	**	0.0381	0.0119		0.0664
Vehicle 2 Passive Restraint	-0.0428	**	0.0201	-0.0875	**	0.0294
Vehicle 2 Anti-Theft	0.0547	**	0.0103	0.0385	**	0.0171
Vehicle 2 Anti-Lock	0.0317	**	0.0105	0.0199		0.0170
Insurance Score	-0.0017	**	0.0000	-0.0013	**	0.0001
Previous Accident	0.0913	**	0.0156	0.0756	**	0.0277
Previous Convictions	0.1476		0.0888	0.0648		0.1670
Previous Reinstated	0.0170		0.0558	0.0003		0.0996
Previous Revocation	-0.0218		0.1456	0.3156		0.1967
Previous Suspension	0.0463		0.0564	0.0125		0.1026
Previous Violation	0.0827	**	0.0093	0.0577	**	0.0161
Year Dummies	Yes			Yes		
Territory Codes	Yes			Yes		
Variance (ϕ)	0.2242	**	0.0065	0.5661	**	0.0198
Loglikelihood	-399,318			-169,817		

Note: Territory codes indicate rating territories, which are based on actuarial risk factors such as traffic and weather patterns, population demographics, wildlife density, and the cost of goods and services.

** Significant at the 5 percent level.

Table A.5: Claim Rate Regression - Home
Poisson panel regression model with random effects
Full data set (1,265,229 household-year records)

	Coefficient		Standard error
Constant	-7.3642	**	0.0978
Dwelling Value	0.0000	**	0.0000
Home Age	0.0016	**	0.0006
Home Age Squared	0.0000	**	0.0000
Number of Families	-0.0021		0.0023
Distance to Hydrant	0.0000		0.0000
Alarm Discount	0.2463	**	0.0195
Protection Devices	-0.1852	**	0.0239
Farm/Business	0.1044	**	0.0242
Primary Home	0.4832	**	0.0819
Owner Occupied	0.2674	**	0.0419
Construction: Fire Resistant	0.1525		0.1342
Construction: Masonry	0.0751	**	0.0172
Construction: Masonry/Veneer	0.0755	**	0.0252
Construction: Frame	.		.
Insurance Score	-0.0026	**	0.0000
Year Dummies		Yes	
Protection Classes		Yes	
Territory Codes		Yes	
Variance (ϕ)	0.4514	**	0.0086
Loglikelihood		-347,278	

Notes: Territory codes indicate rating territories, which are based on actuarial risk factors such as traffic and weather patterns, population demographics, wildlife density, and the cost of goods and services. Protection classes gauge the effectiveness of local fire protection and building codes.

** Significant at the 5 percent level.

Table A.6: Model 2

Core sample (4170 households)

	r			Log Ω(μ): constant			Log Ω(μ): linear			Log Ω(μ): quadratic		
	Estimate		**Standard error**	**Estimate**		**Standard error**	**Estimate**		**Standard error**	**Estimate**		**Standard error**
Constant	-7.39	**	0.09	-2.73	**	0.02	12.40	**	0.41	-35.61	**	2.47
Driver 1 age	-1.47	**	0.14	0.18	**	0.05	2.52	**	0.48	1.00	**	0.31
Driver 1 age squared	1.09	**	0.15	0.00		0.05	-4.94	**	0.43	9.92	**	1.82
Driver 1 female	0.15	**	0.04	-0.05	**	0.01	1.57	**	0.28	-12.29	**	1.66
Driver 1 single	0.08		0.05	-0.01		0.01	0.77	**	0.26	-5.93	**	1.91
Driver 1 married	0.09		0.06	-0.03		0.02	1.40	**	0.27	-9.34	**	1.40
Insurance score	-0.15	**	0.05	-0.02		0.01	2.31	**	0.21	-11.00	**	1.23
Driver 2 indicator	-0.04		0.06	0.00		0.02	-1.44	**	0.38	7.63	**	2.19
Mean fitted value	0.00073			-2.73			12.40			-35.61		
Median fitted value	0.00056			-2.73			12.42			-34.46		
σ_L	27.22	**	0.76									
σ_M	17.91	**	0.48									
σ_H	65.45	**	2.68									

Note: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at the 5 percent level.

Table A.7: Model 3

Core sample (4170 households)

	Estimate		Standard error
r	0.00118	***	0.00011
Log Ω(μ): constant	-2.82	***	0.03
Log Ω(μ): linear	10.72	***	0.52
Log Ω(μ): quadratic	-28.04	***	3.46
σ_L	17.14	***	0.61
σ_M	10.16	***	0.38
σ_H	95.69	***	6.74
Φ_r	0.55	***	0.09
Φ_Ω	0.37	***	0.02
$\Phi_{r,\Omega}$	-0.33	***	0.03
Mean fitted r	0.00156		
Implied corr($\xi_{r,i}$,$\xi_{\Omega,i}$)	-0.72		

*** Significant at the 1 percent level.

Table A.8: Model 4

Core sample (4170 households)

	r			Log Ω(μ): constant			Log Ω(μ): linear			Log Ω(μ): quadratic		
	Estimate		**Standard error**	**Estimate**		**Standard error**	**Estimate**		**Standard error**	**Estimate**		**Standard error**
Constant	-6.81	**	0.10	-2.83	**	0.03	11.22	**	0.50	-31.67	**	3.42
Driver 1 age	-0.41	**	0.06	0.18	**	0.02	-2.01	**	0.51	7.14	**	3.35
Driver 1 age squared	0.07		0.04	0.01		0.02	-0.66		0.44	1.96		2.80
Driver 1 female	0.06		0.05	-0.05		0.03	2.03	**	0.59	-14.48	**	3.77
Driver 1 single	-0.02		0.05	0.02		0.03	0.24		0.60	-3.65		3.87
Driver 1 married	-0.03		0.08	0.00		0.04	1.06		0.81	-6.68		5.07
Insurance score	-0.14	**	0.05	0.00		0.02	1.27	**	0.55	-5.43		3.43
Driver 2 indicator	-0.13		0.07	0.04		0.03	-1.84	**	0.72	7.59		4.59
Mean fitted value	0.00147			-2.83			11.22			-31.67		
Median fitted value	0.00145			-2.83			11.04			-28.53		
σ_L	17.09	**	0.61									
σ_M	10.46	**	0.40									
σ_H	90.34	**	6.23									
Φ_r	0.58	**	0.09									
Φ_Ω	0.35	**	0.02									
$\Phi_{r,\Omega}$	-0.33	**	0.03									
Implied corr($\xi_{r,i}$,$\xi_{\Omega,i}$)	-0.72											

Note: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at the 5 percent level.

Table A.9: Unobserved Heterogeneity in Risk – Model 2

Core sample (4170 households)

	r			**Log Ω(μ): constant**			**Log Ω(μ): linear**			**Log Ω(μ): quadratic**		
	Estimate		**Standard error**	**Estimate**		**Standard error**	**Estimate**		**Standard error**	**Estimate**		**Standard error**
Constant	-7.11	**	0.08	-2.73	**	0.02	11.04	**	0.41	-21.78	**	1.95
Driver 1 age	-1.22	**	0.19	0.15	**	0.07	3.31	**	0.57	-3.86	**	1.66
Driver 1 age squared	0.75	**	0.20	0.02		0.06	-4.92	**	0.61	7.64	**	1.10
Driver 1 female	0.05		0.05	-0.01		0.02	1.24	**	0.29	-9.90	**	1.32
Driver 1 single	0.03		0.07	0.00		0.02	0.56	**	0.21	-3.25	**	1.20
Driver 1 married	0.02		0.12	-0.01		0.05	1.24	**	0.46	-6.87	**	1.00
Insurance score	-0.12	**	0.05	-0.01		0.03	1.49	**	0.62	-3.64		2.73
Driver 2 indicator	-0.09		0.08	0.01		0.04	-1.40		0.75	5.27	**	1.99
Mean fitted value	0.00097			-2.73			11.04			-21.78		
Median fitted value	0.00076			-2.74			11.05			-19.49		
σ_L	25.93	**	0.80									
σ_M	18.02	**	0.56									
σ_H	66.01	**	2.78									

Note: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at the 5 percent level.

Table A.10: Unobserved Heterogeneity in Risk – Model 3

Core sample (4170 households)

	Estimate		Standard error
r	0.00134	***	0.00012
Log Ω(μ): constant	-2.77	***	0.09
Log Ω(μ): linear	8.76	***	0.03
Log Ω(μ): quadratic	-16.70	***	0.08
σ_L	17.00	***	0.04
σ_M	9.72	***	0.61
σ_H	105.49	***	0.38
Φ_r	0.42		7.99
Φ_Ω	0.37	***	0.07
$\Phi_{r,\Omega}$	-0.28	***	0.02
Mean fitted r	0.00166		
Implied corr($\xi_{r,i}$,$\xi_{\Omega,i}$)	-0.72		

*** Significant at the 1 percent level.

Table A.11: Unobserved Heterogeneity in Risk

Core sample (4170 households)

	Estimate		Standard error
r	0.00160	***	0.00020
Ω(μ): intercept	0.04	***	0.00
Ω(μ): slope	0.86	***	0.03
σ_L	21.83	***	0.97
σ_M	10.98	***	0.50
σ_H	89.33	***	7.01
Ψ_r	0.67	***	0.09
Ψ_L	0.76	***	0.05
Ψ_M	3.11	***	0.16
Ψ_H	0.88	***	0.07
$\Psi_{r,L}$	-0.33	***	0.06
$\Psi_{r,M}$	-0.69	***	0.11
$\Psi_{r,H}$	0.17	***	0.06
$\Psi_{L,H}$	0.40	***	0.05
$\Psi_{M,H}$	0.85	***	0.10
$\Psi_{L,M}$	1.44	***	0.08
Mean fitted r	0.00224		
Implied corr($\xi_{r,i}$,$\xi_{b,Li}$)	-0.45		
Implied corr($\xi_{r,i}$,$\xi_{b,Mi}$)	-0.47		
Implied corr($\xi_{r,i}$,$\xi_{b,Hi}$)	0.22		
Implied corr($\xi_{b,Li}$,$\xi_{b,Mi}$)	0.94		
Implied corr($\xi_{b,Li}$,$\xi_{b,Hi}$)	0.49		
Implied corr($\xi_{b,Mi}$,$\xi_{b,Hi}$)	0.51		

*** Significant at the 1 percent level.

Table A.12: Restricted Choice Noise – Model 2

Core sample (4170 households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Estimate	**Standard error**	**Estimate**	**Standard error**	**Estimate**	**Standard error**	**Estimate**	**Standard error**
Constant	-14.37 **	0.14	-2.52 **	0.00	12.07 **	0.07	-36.63 **	2.08
Driver 1 age	1.89 **	0.01	0.18 **	0.00	3.11 **	0.01	-1.65 **	0.03
Driver 1 age squared	-6.73 **	0.07	0.00	0.00	-4.91 **	0.02	11.08 **	1.07
Driver 1 female	0.00	0.01	-0.05 **	0.00	1.24 **	0.01	-9.66 **	0.78
Driver 1 single	0.20 **	0.00	-0.01 **	0.00	0.35 **	0.00	-3.36 **	0.03
Driver 1 married	0.21 **	0.00	-0.03 **	0.00	1.05 **	0.00	-7.29 **	0.30
Insurance score	-0.05 **	0.01	-0.02 **	0.00	2.07 **	0.01	-9.91 **	0.54
Driver 2 indicator	-0.27 **	0.01	0.00 **	0.00	-1.49 **	0.02	8.37 **	0.32
Mean fitted value	0.00005		-2.52		12.07		-36.63	
Median fitted value	0.00000		-2.53		12.13		-36.25	
σ_L	13.61							
σ_M	8.95							
σ_H	32.72							

Notes: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*). The variance terms σ_L, σ_M, and σ_H are by construction.

** Significant at the 5 percent level.

Table A.13: Restricted Choice Noise – Model 3

Core sample (4170 households)

	Estimate		Standard error
r	0.00000		0.00000
Log Ω(μ): constant	-2.40	***	0.01
Log Ω(μ): linear	10.17	***	0.20
Log Ω(μ): quadratic	-37.00	***	1.44
σ_L	8.57		
σ_M	5.08		
σ_H	47.84		
Φ_r	2.77		3.47
Φ_Ω	0.25	***	0.01
$\Phi_{r,\Omega}$	0.05		0.26
Mean fitted r	0.00000		
Implied corr($\xi_{r,i}$,$\xi_{\Omega,i}$)	0.07		

Note: The variance terms σ_L, σ_M, and σ_H are by construction.

*** Significant at the 1 percent level.

Table A.14: CRRA Utility – Model 2

Core sample (4170 households)

	ρ		Log Ω(μ): constant		Log Ω(μ): linear		Log Ω(μ): quadratic	
	Estimate	**Standard error**	**Estimate**	**Standard error**	**Estimate**	**Standard error**	**Estimate**	**Standard error**
Constant	-1.16 **	0.15	-2.51 **	0.02	12.31 **	0.46	-39.53 **	3.01
Driver 1 age	-0.44	1.00	-0.10	0.11	1.56	3.08	5.04	20.31
Driver 1 age squared	-0.05	0.99	0.20	0.11	-3.61	3.09	5.90	20.43
Driver 1 female	-0.16	0.15	-0.01	0.02	1.14 **	0.56	-8.48 **	3.66
Driver 1 single	-0.30	0.17	0.02	0.02	0.18	0.55	-1.23	3.65
Driver 1 married	-0.07	0.24	-0.01	0.03	1.01	0.83	-6.40	5.49
Insurance score	-0.11	0.13	-0.03	0.02	1.74 **	0.51	-8.36 **	3.33
Driver 2 indicator	-0.30	0.21	0.01	0.03	-1.32	0.71	7.36	4.71
Mean fitted value	0.37		-2.51		12.31		-39.53	
Median fitted value	0.31		-2.53		12.35		-39.31	
σ_L	22.79 **	0.68						
σ_M	15.77 **	0.52						
σ_H	40.24 **	1.21						

Note: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at the 5 percent level.

Table A.15: CRRA Utility – Model 3

Core sample (4170 households)

	Estimate		Standard error
ρ	0.18	***	0.03
Log Ω(μ): constant	-2.47	***	0.02
Log Ω(μ): linear	11.39	***	0.47
Log Ω(μ): quadratic	-37.86	***	3.23
σ_L	13.21	***	0.38
σ_M	8.09	***	0.27
σ_H	42.41	***	1.37
Φ_r	0.27	***	0.07
Φ_Ω	0.18	***	0.01
$\Phi_{r,\Omega}$	-0.12	***	0.02
Mean fitted ρ	0.21		
Implied corr($\xi_{r,i}$,$\xi_{\Omega,i}$)	-0.56		

*** Significant at the 1 percent level.

Table A.16: All Claims $750 – Model 1a

Core sample (4170 households)

	Estimate		Standard error
r	0.00195	***	0.00000
Log Ω(μ): constant	-2.73	***	0.03
Log Ω(μ): linear	12.45	***	0.30
Log Ω(μ): quadratic	-32.19	***	2.17
σ_L	54.30	***	2.57
σ_M	28.08	***	0.47
σ_H	244.69	***	0.00

*** Significant at the 1 percent level.

Table A.17: CARA Utility – Model 2

Core sample (4170 households)

	r			Log Ω(μ): constant			Log Ω(μ): linear			Log Ω(μ): quadratic		
	Estimate		**Standard error**	**Estimate**		**Standard error**	**Estimate**		**Standard error**	**Estimate**		**Standard error**
Constant	-6.83	**	0.05	-3.01	**	0.03	11.94	**	0.65	-34.82	**	3.94
Driver 1 age	-0.75	**	0.34	0.33	**	0.15	8.43	**	1.33	-18.69	**	2.73
Driver 1 age squared	0.53		0.39	-0.08		0.16	-10.42	**	1.49	28.08	**	3.37
Driver 1 female	0.04		0.02	-0.02		0.03	0.90		0.75	-9.00	**	4.25
Driver 1 single	0.02		0.02	-0.01		0.02	0.34		0.23	-3.13		1.90
Driver 1 married	-0.02		0.03	0.02		0.04	0.67		0.89	-4.54		2.74
Insurance score	-0.04		0.03	-0.02		0.03	2.30	**	0.78	-10.59	**	3.83
Driver 2 indicator	-0.05		0.03	0.01		0.03	-0.95	**	0.46	6.21	**	1.57
Mean fitted value	0.00113			-3.01			11.94			-34.82		
Median fitted value	0.00103			-3.01			12.06			-34.67		
σ_L	33.31	**	0.98									
σ_M	21.30	**	0.64									
σ_H	104.30	**	5.29									

Note: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at the 5 percent level.

Table A.18: Alternative Sample 1 – Model 2

(20,662 households)

	r		Log Ω(μ): constant		Log Ω(μ): linear		Log Ω(μ): quadratic	
	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Constant	-6.83 **	0.04	-2.88 **	0.01	14.01 **	0.34	-43.10 **	2.78
Driver 1 age	1.14 **	0.34	-0.31 **	0.09	-1.14 **	0.54	-7.11 **	2.75
Driver 1 age squared	-1.28 **	0.36	0.41 **	0.08	-2.79 **	0.34	26.03 **	1.11
Driver 1 female	0.16 **	0.02	-0.01	0.01	-0.89 **	0.22	2.40 **	0.86
Driver 1 single	-0.08 **	0.04	0.03 **	0.02	0.14	0.21	-1.65	1.27
Driver 1 married	0.01	0.03	0.03 **	0.01	-0.97 **	0.14	2.68 **	0.32
Insurance score	-0.07 **	0.02	0.03 **	0.01	0.86 **	0.25	-4.51 **	1.21
Driver 2 indicator	0.10 **	0.03	0.01	0.01	-2.23 **	0.30	10.63 **	2.04
Mean fitted value	0.00113		-2.88		14.01		-43.10	
Median fitted value	0.00112		-2.92		14.26		-45.60	
σ_L	31.17 **	0.48						
σ_M	17.91 **	0.43						
σ_H	18.93 **	0.27						

Note: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at the 5 percent level.

Table A.19: Alternative Sample 2 – Model 2

(6824 households)

	r		Log Ω(μ): constant		Log Ω(μ): linear		Log Ω(μ): quadratic	
	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Constant	-7.57 **	0.08	-2.63 **	0.02	12.91 **	0.43	-42.12 **	2.97
Driver 1 age	-1.31 **	0.21	0.06	0.09	4.52	2.46	-11.77	15.32
Driver 1 age squared	0.91 **	0.22	0.08	0.09	-6.73 **	2.53	22.96	16.35
Driver 1 female	0.15 **	0.04	-0.03	0.02	0.94 **	0.29	-7.90 **	1.65
Driver 1 single	0.05	0.05	0.02	0.02	-0.10	0.33	-0.18	2.01
Driver 1 married	-0.13	0.08	0.04 **	0.02	0.42	0.45	-4.30	2.58
Insurance score	-0.08 **	0.04	-0.01	0.02	2.19 **	0.38	-12.41 **	2.31
Driver 2 indicator	0.13	0.07	-0.03	0.02	-1.32 **	0.41	8.32 **	2.86
Mean fitted value	0.00060		-2.63		12.91		-42.12	
Median fitted value	0.00048		-2.65		12.99		-42.40	
σ_L	28.61 **	0.71						
σ_M	18.78 **	0.47						
σ_H	58.74 **	2.00						

Note: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at the 5 percent level.

Table A.20: Restricted Menus – Model 2

Core sample (4170 households)

	r		Log Ω(μ): constant		Log Ω(μ): linear		Log Ω(μ): quadratic	
	Estimate	**Standard error**	**Estimate**	**Standard error**	**Estimate**	**Standard error**	**Estimate**	**Standard error**
Constant	-8.79 **	0.27	-2.81 **	0.04	12.88 **	0.65	-30.29 **	3.12
Driver 1 age	-4.22 **	0.85	0.03	0.17	6.47 **	2.37	-8.09	8.02
Driver 1 age squared	3.92 **	0.94	0.38 **	0.17	-10.94 **	2.28	23.24 **	7.44
Driver 1 female	0.14	0.16	0.03	0.04	0.91	1.00	-9.94 **	5.00
Driver 1 single	0.18	0.19	-0.01	0.04	0.96	0.99	-6.98	5.07
Driver 1 married	0.20	0.27	-0.02	0.04	1.31	0.76	-8.92 **	3.09
Insurance score	-0.54 **	0.17	-0.04	0.03	2.59 **	0.66	-10.26 **	2.93
Driver 2 indicator	-0.07	0.21	0.00	0.04	-1.19	0.79	5.52	4.05
Mean fitted value	0.00029		-2.81		12.88		-30.29	
Median fitted value	0.00013		-2.88		13.46		-30.35	
σ_L	26.26 **	0.84						
σ_M	19.83 **	0.89						
σ_H	76.56 **	4.49						

Note: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at the 5 percent level.

Table A.21: Alternative Error Structure A – Model 2

Core sample (4170 households)

	r		Log Ω(μ): constant		Log Ω(μ): linear		Log Ω(μ): quadratic	
	Estimate	**Standard error**	**Estimate**	**Standard error**	**Estimate**	**Standard error**	**Estimate**	**Standard error**
Constant	-9.22 **	0.20	-2.81 **	0.05	13.48 **	0.80	-32.52 **	3.46
Driver 1 age	-4.89 **	0.97	0.04	0.20	6.23 **	1.69	-6.01	5.40
Driver 1 age squared	4.61 **	1.22	0.35	0.24	-10.31 **	2.68	19.36 **	2.34
Driver 1 female	0.09	0.10	0.03	0.02	0.91 **	0.27	-9.67 **	2.15
Driver 1 single	0.30 **	0.13	-0.01	0.02	0.84 **	0.37	-5.96 **	1.87
Driver 1 married	0.44 **	0.13	-0.02	0.03	1.22 **	0.32	-8.34 **	1.26
Insurance score	-0.61 **	0.13	-0.04	0.03	2.57 **	0.76	-10.16 **	3.30
Driver 2 indicator	-0.07	0.18	0.00	0.03	-1.18 **	0.47	5.43 **	2.14
Mean fitted value	0.00022		-2.81		13.48		-32.52	
Median fitted value	0.00008		-2.87		13.96		-32.32	
σ_L	51.63 **	1.34						
σ_M	36.49 **	1.46						
σ_H	130.57 **	6.96						

Note: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at the 5 percent level.

Table A.22: Alternative Error Structure B – Model 2

Core sample (4170 households)

	r		Log Ω(μ): constant		Log Ω(μ): linear		Log Ω(μ): quadratic	
	Estimate	**Standard error**	**Estimate**	**Standard error**	**Estimate**	**Standard error**	**Estimate**	**Standard error**
Constant	-7.04 **	0.09	-2.04 **	0.03	3.21 **	0.64	-7.30 **	3.34
Driver 1 age	-1.43 **	0.18	0.42 **	0.11	2.89	1.90	-7.31	5.79
Driver 1 age squared	1.16 **	0.19	-0.28 **	0.12	-3.60	2.27	10.37	8.06
Driver 1 female	0.00	0.03	0.03	0.02	0.08	0.50	-2.96	1.98
Driver 1 single	-0.05	0.04	-0.09 **	0.04	2.03	1.11	-7.82	5.07
Driver 1 married	-0.11 **	0.06	0.01	0.05	1.32	1.01	-7.73 **	3.23
Insurance score	-0.26 **	0.03	-0.01	0.01	2.25 **	0.29	-8.80 **	1.26
Driver 2 indicator	-0.07	0.05	-0.02	0.02	0.14	0.22	2.25	2.11
Mean fitted value	0.00101		-2.04		3.21		-7.30	
Median fitted value	0.00081		-2.01		3.36		-7.54	
σ_L	62.75 **	2.22						
σ_M	53.33 **	3.23						
σ_H	65.80 **	2.94						

Note: Each independent variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at the 5 percent level.

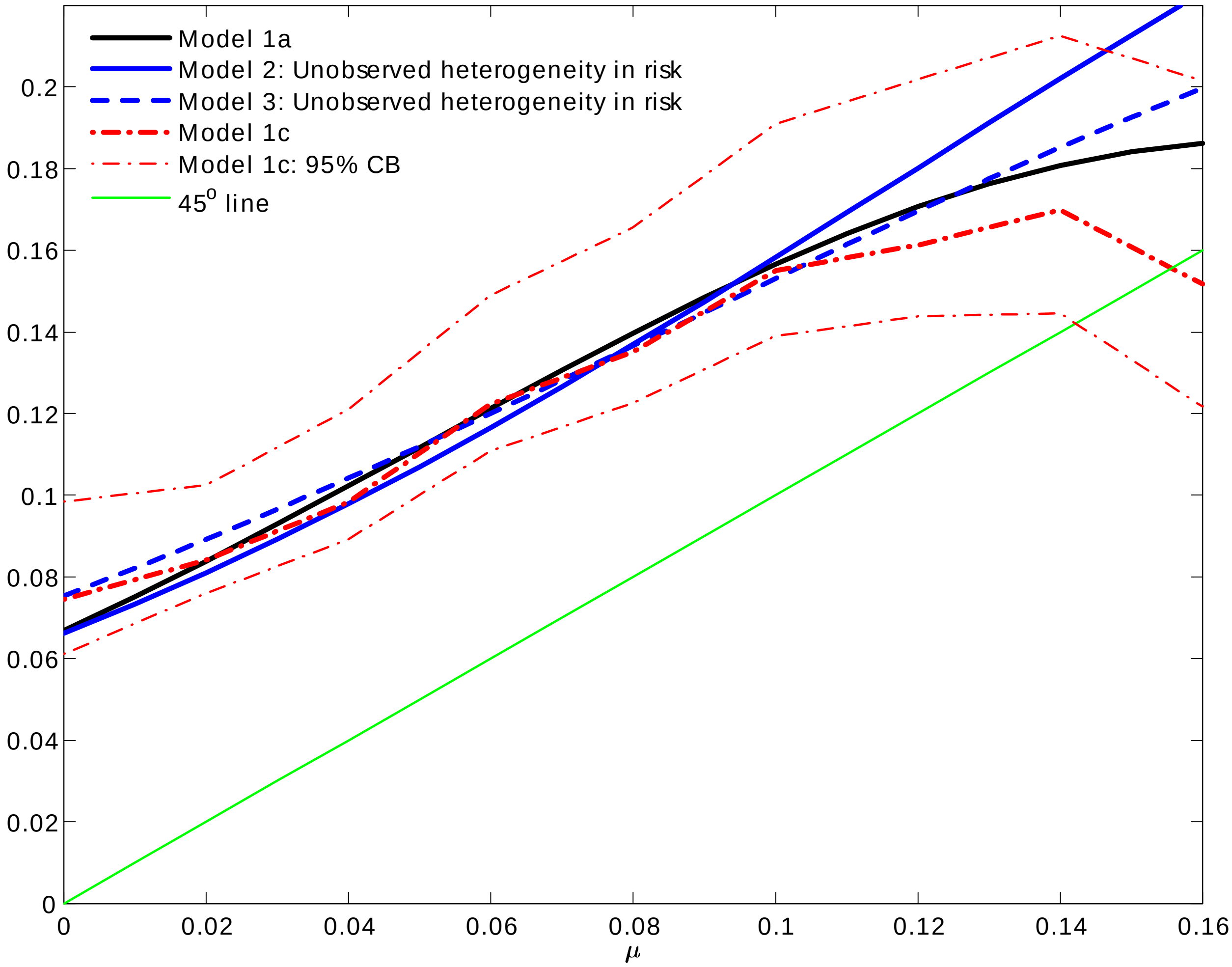


Figure A.1: Mean Estimated $\Omega(\mu)$ – Unobserved Heterogeneity in Risk

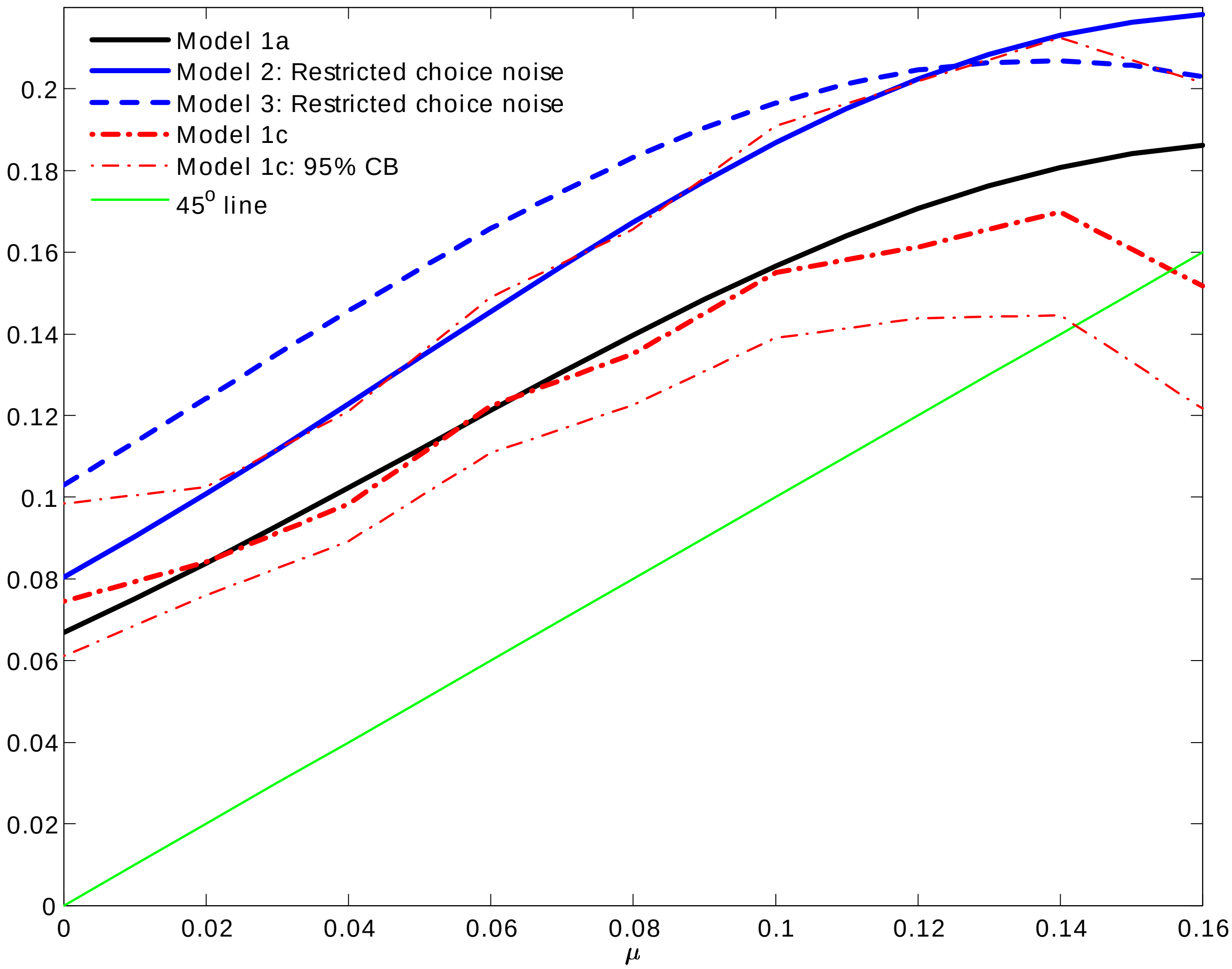


Figure A.2: Mean Estimated $\Omega(\mu)$ – Restricted Choice Noise

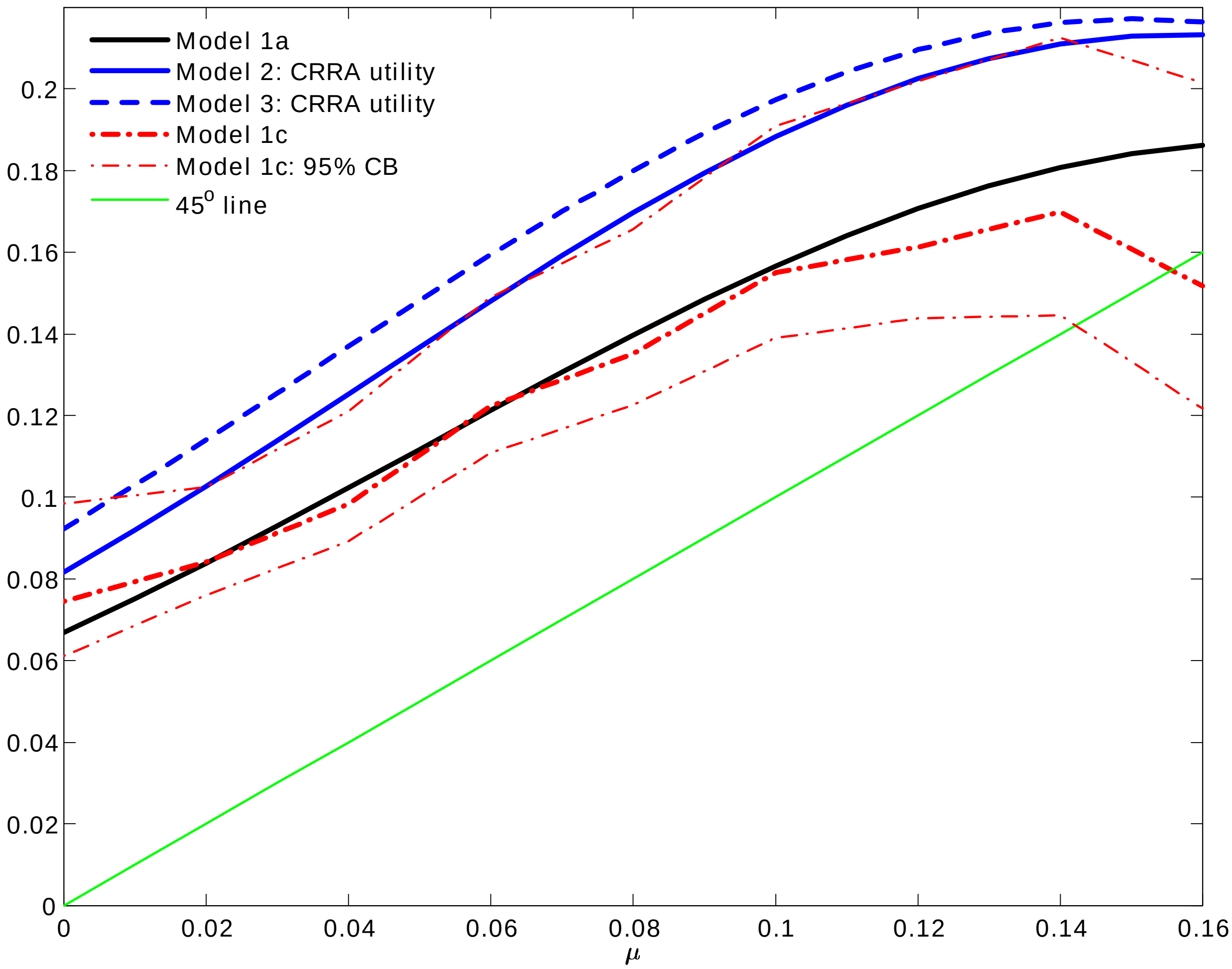


Figure A.3: Mean Estimated $\Omega(\mu)$ – CRRA Utility

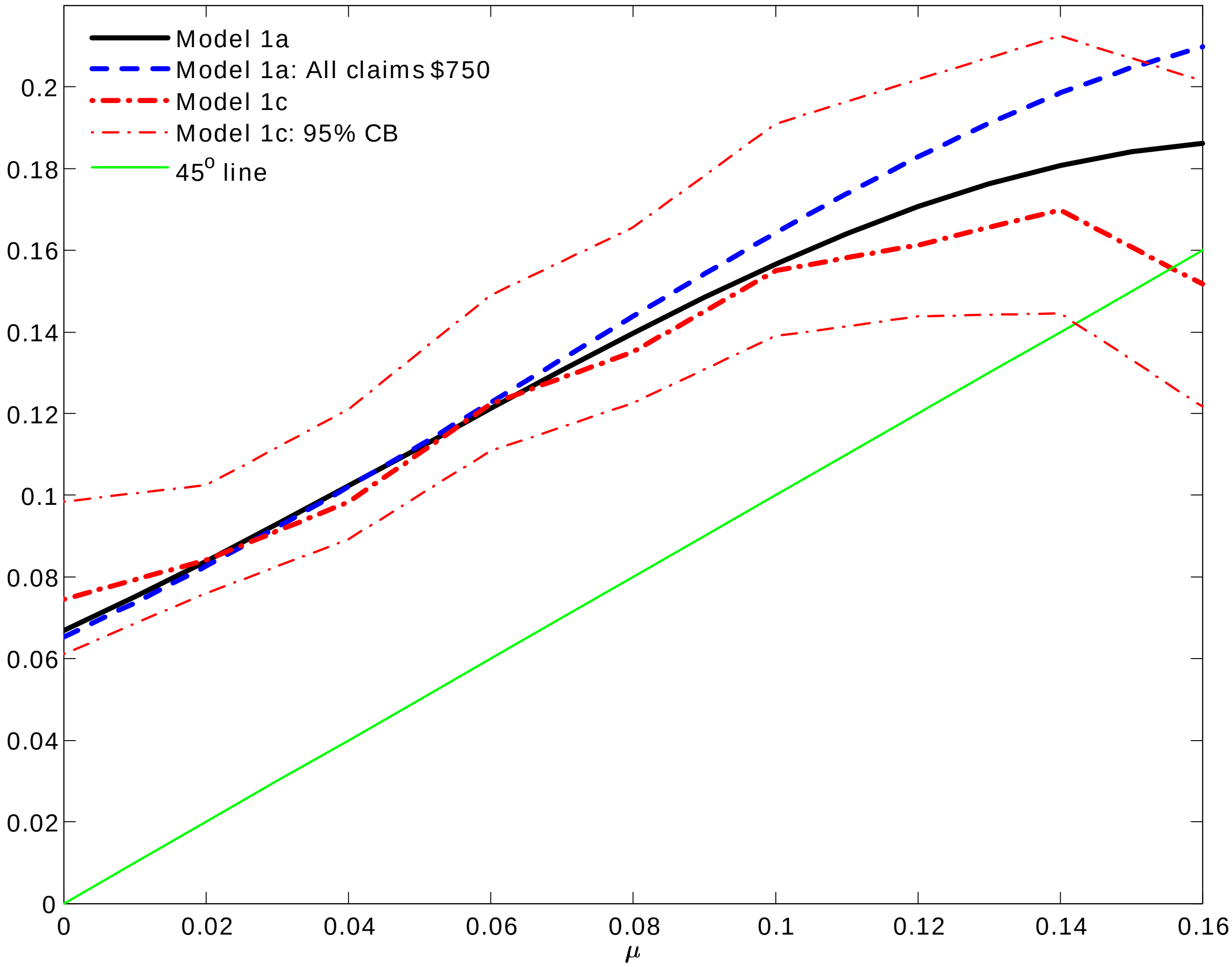


Figure A.4: Estimated $\Omega(\mu)$ – All Claims $750