Supplement to "Inference under stability of risk preferences"

(*Quantitative Economics*, Vol. 7, No. 2, July 2016, 367–409)

Levon Barseghyan
Department of Economics, Cornell University

Francesca Molinari
Department of Economics, Cornell University

Joshua C. Teitelbaum
Law Center, Georgetown University

A. Deductible choices, pricing menus, and claim probabilities

Tables S1, S2, and S3 summarize the deductible choices, pricing menus, and claim probabilities, respectively, of the households in the core sample.

B. CRRA and NTD utility

In this section, we show that our results are very similar if, instead of assuming CARA utility, we assume either (i) CRRA for reasonable levels of wealth or (ii) NTD utility.

We begin by assuming CRRA utility. That is, we assume $\rho_i \equiv w_i \times r_i = -w_i u_i''(w_i)/u_i'(w_i)$ is a constant function of w_i. Following BMOT, we assume $w_i = \$33{,}000$, which corresponds to 2010 U.S. per capita disposable personal income. Figure S1 displays the percentage of rationalizable households that satisfy each shape restriction as we increase the upper bound on r_i from 0 to 0.0108 (which, given our wealth assumption, corresponds to increasing the upper bound on ρ_i from 0 to 356). The patterns displayed in Figure S1 are remarkably similar to the patterns displayed in Figure 2. We note that the patterns are essentially the same if we double wealth or cut it in half.[1]

Next, we assume NTD utility. That is, we consider a second-order Taylor expansion of $u_i(w_i)$ around w_i (Cohen and Einav (2007), Barseghyan, Prince, and Teitelbaum (2011), Barseghyan, Molinari, O'Donoghue, and Teitelbaum (2013)). Figure S2 displays the percentage of rationalizable households that satisfy each shape restriction as we increase the upper bound on r_i from 0 to 0.0108. Again, the patterns displayed are very similar to the patterns displayed in Figure 2.

Levon Barseghyan: lb247@cornell.edu
Francesca Molinari: fm72@cornell.edu
Joshua C. Teitelbaum: jct48@law.georgetown.edu

[1] Figures are available upon request.

DOI: 10.3982/QE511

Table S1. Summary of deductible choices. Core sample (4170 households).

Deductible	Collision	Comp.	Home
$50		5.2	
$100	1.0	4.1	0.9
$200	13.4	33.5	
$250	11.2	10.6	29.7
$500	67.7	43.0	51.9
$1000	6.7	3.6	15.9
$2500			1.2
$5000			0.4

Note: Values are percent of households. Comp. stands for comprehensive.

Table S2. Summary of pricing menus. Core sample (4170 households).

Coverage	Mean	Std. Dev.	1st Pctl.	99th Pctl.
Auto collision premium for $500 deductible	180	100	50	555
Auto comprehensive premium for $500 deductible	115	81	26	403
Home all perils premium for $500 deductible	679	519	216	2511
Cost of decreasing deductible from $500 to $250:				
Auto collision	54	31	14	169
Auto comprehensive	30	22	6	107
Home all perils	56	43	11	220
Savings from increasing deductible from $500 to $1000:				
Auto collision	41	23	11	127
Auto comprehensive	23	16	5	80
Home all perils	74	58	15	294

Note: Annual amounts in dollars.

Table S3. Claim probabilities (annual). Core sample (4170 households).

	Collision	Comp.	Home
Mean	0.069	0.021	0.084
Standard deviation	0.024	0.011	0.044
1st percentile	0.026	0.004	0.024
5th percentile	0.035	0.007	0.034
25th percentile	0.052	0.013	0.053
Median	0.066	0.019	0.076
75th percentile	0.083	0.027	0.104
95th percentile	0.114	0.041	0.163
99th percentile	0.139	0.054	0.233

Note: Comp. stands for comprehensive.

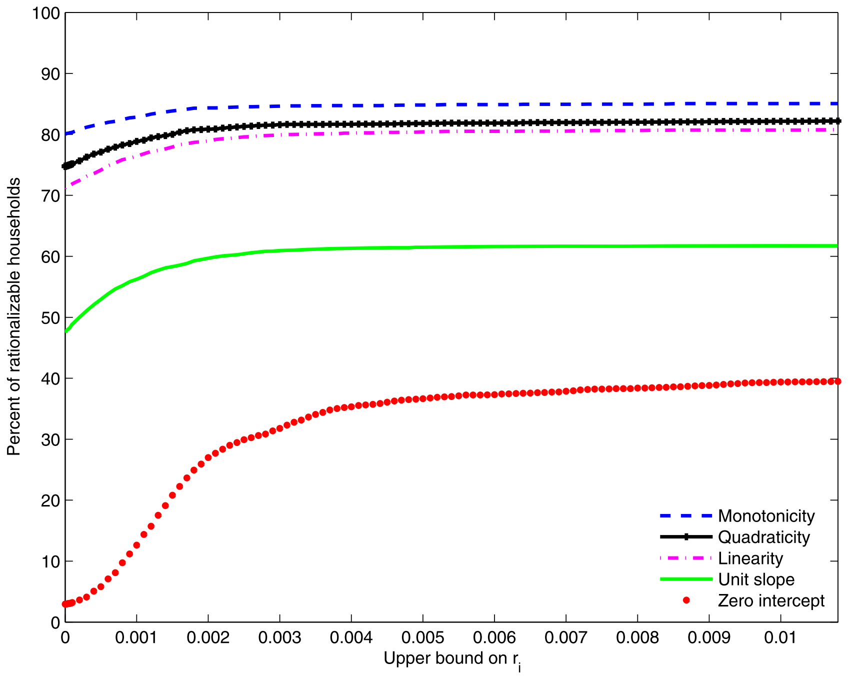

FIGURE S1. CRRA utility.

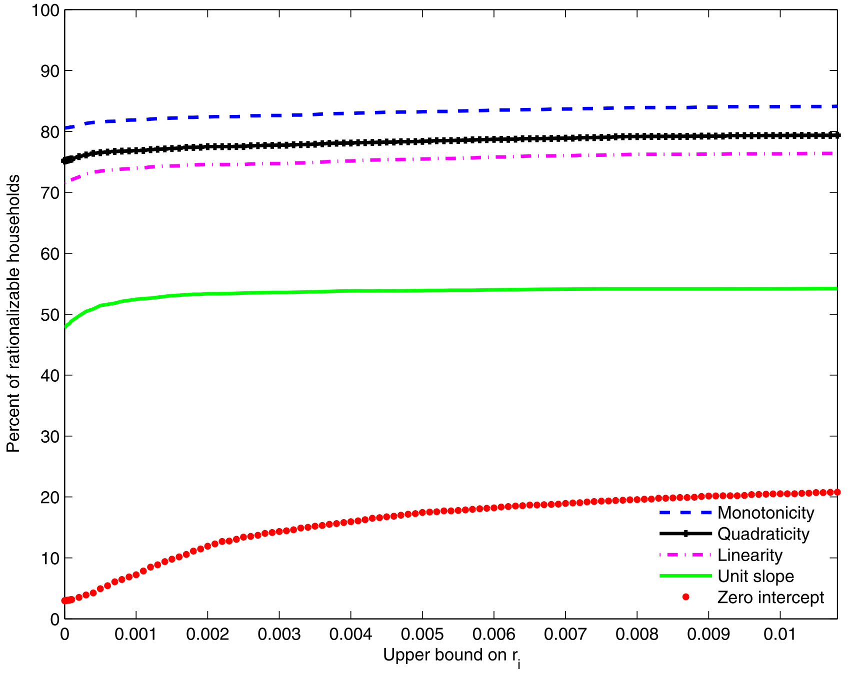

FIGURE S2. NTD utility.

C. Upper bound on r_i

In this section, we discuss the upper bound on r_i. For purposes of this discussion, let $\overline{r}_i$ denote this upper bound. Figure S3, panel (A) displays the percentage of all households in the core sample ($N = 4170$) that satisfy plausibility and each shape restriction on $\Omega_i(\cdot)$ as we increase $\overline{r}_i$ from 0 to 0.02. Panel (B) displays the percentage of rationalizable households that satisfy each shape restriction on $\Omega_i(\cdot)$ as we increase $\overline{r}_i$ from 0 to 0.02. In panel (B), the fraction of rationalizable households that satisfy a particular shape restriction at a given $\overline{r}_i$ is calculated dynamically: it is the number of households that satisfy the shape restriction at the given $\overline{r}_i$ divided by the number of households that satisfy plausibility at the given $\overline{r}_i$.[2]

The fraction of households that satisfy plausibility is roughly 87 percent for all $\overline{r}_i$ between zero and about 0.011. As we note in Section 4.1, virtually every household that violates plausibility in this range chose an auto collision deductible of \$200. For $\overline{r}_i$ greater than about 0.011, the fraction of households that satisfy plausibility steadily increases with $\overline{r}_i$, hitting roughly 97 percent at $\overline{r}_i = 0.02$. However, such levels of absolute risk aversion are absurdly high. Here they imply/require implausibly low values of $\Omega_i(\mu_{ij})$—close to zero for all μ_{ij}—so as to rationalize the deductible choices of these households (particularly their auto collision deductible choices). As a result, the zero intercept model (i.e., objective expected utility theory) cannot rationalize most of these households. This is why, once $\overline{r}_i$ surpasses about 0.011, the zero intercept curve levels off in panel (A) and declines in panel (B) (having achieved its maximum at 0.0108). The monotone probability distortions model, by contrast, can rationalize a greater number of these households. This is because $\Omega_i(\mu_{ij})$ can be increasing even if it is implausibly low for all μ_{ij}. This is why, once $\overline{r}_i$ surpasses about 0.011, the monotonicity curve increases along with the plausibility curve in panel (A) (though it also declines in panel (B), indicating that the monotone distortions model cannot rationalize the majority of these households).

D. Monotonicity as r_i increases

Figure 2 shows, inter alia, that the percentage of rationalizable households that satisfy monotonicity increases as we increase the upper bound on r_i. In this section, we discuss the intuition behind this result.

Consider a setting with two coverages, $j \in \{\mathrm{I}, \mathrm{II}\}$, and three deductible options in each coverage, $\mathcal{D}_j = \{250, 500, 1000\}$ for $j = \mathrm{I}, \mathrm{II}$, and suppose that $\mu_{i\mathrm{I}} < \mu_{i\mathrm{II}}$. Monotonicity fails if $\mathrm{LB}_{i\mathrm{I}} > \mathrm{UB}_{i\mathrm{II}}$.

Recall that

$$\mathrm{LB}_{ij} \equiv \max\Big\{0, \max_{d>d^*} \Delta_{ij}\Big\} \quad \text{and} \quad \mathrm{UB}_{ij} \equiv \min\Big\{1, \min_{d<d^*} \Delta_{ij}\Big\},$$

[2] In Figure 2, by contrast, the fraction of rationalizable households that satisfy a particular shape restriction at a given $\overline{r}_i$ is calculated statically: it is the number of households that satisfy the shape restriction at the given $\overline{r}_i$ divided by the number of households that satisfy plausibility with $\overline{r}_i$ fixed at 0.0108 ($N = 3629$).

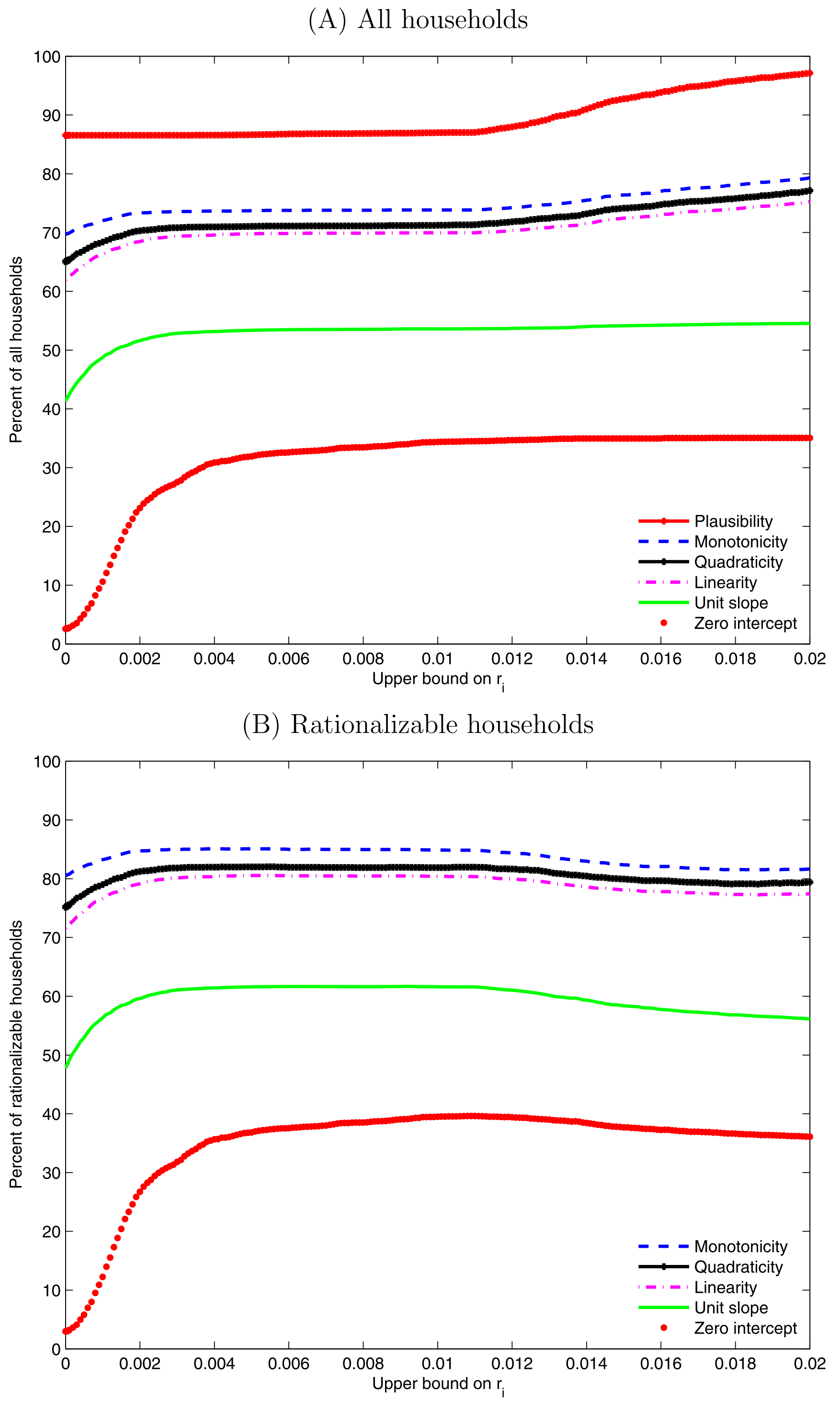

FIGURE S3. Increasing the upper bound on r_i.

where

$$\begin{aligned}\Delta_{ij} &= \Delta_{ij}(d)\\ &\equiv \big(u_i\big(w_i - p_{ij}(d)\big) - u_i\big(w_i - p_{ij}\big(d^*\big)\big)\big)\\ &\quad /\big\{\big[u_i\big(w_i - p_{ij}(d)\big) - u_i\big(w_i - p_{ij}(d) - d\big)\big]\\ &\quad - \big[u_i\big(w_i - p_{ij}\big(d^*\big)\big) - u_i\big(w_i - p_{ij}\big(d^*\big) - d^*\big)\big]\big\}.\end{aligned}$$

Assuming interior solutions, note that for each j, the deductible $d > d^*$ that defines LB_{ij} is necessarily higher than the deductible $d < d^*$ that defines UB_{ij}. For example, assuming the household chooses $d^* = 500$, then $\text{LB}_{ij} = \Delta_{ij}(1000)$ and $\text{UB}_{ij} = \Delta_{ij}(250)$.

Consider the marginal monotone household, for which $\text{LB}_{i\text{I}} = \text{UB}_{i\text{II}}$ at some $r_i \geq 0$. The key insight is that increasing r_i decreases both $\text{LB}_{i\text{I}}$ and $\text{UB}_{i\text{II}}$, but it decreases $\text{LB}_{i\text{I}}$ more. Intuitively, this is because the larger are the stakes (the deductibles), the larger is the decline in $\Omega_i(\mu_{ij})$ that is required to explain/preserve the household's choice (i.e., to keep the household from choosing a higher deductible). It follows that increasing r_i yields $\text{LB}_{i\text{I}} < \text{UB}_{i\text{II}}$.

This is best illustrated in the case of NTD utility. With NTD utility, the choice $d^* = 500$ implies the Ω-intervals

$$\frac{p_{i\text{I}}(500) - p_{i\text{I}}(1000)}{500 + \frac{1}{2}r_i\big(1000^2 - 500^2\big)} = \text{LB}_{i\text{I}} \leq \Omega_i(\mu_{i\text{I}}) \leq \text{UB}_{i\text{I}} = \frac{p_{i\text{I}}(250) - p_{i\text{I}}(500)}{250 + \frac{1}{2}r_i\big(500^2 - 250^2\big)}$$

and

$$\frac{p_{i\text{II}}(500) - p_{i\text{II}}(1000)}{500 + \frac{1}{2}r_i\big(1000^2 - 500^2\big)} = \text{LB}_{i\text{II}} \leq \Omega_i(\mu_{i\text{II}}) \leq \text{UB}_{i\text{II}} = \frac{p_{i\text{II}}(250) - p_{i\text{II}}(500)}{250 + \frac{1}{2}r_i\big(500^2 - 250^2\big)}.$$

Let $\text{LB}_{i\text{I}} = \text{UB}_{i\text{II}}$ at some $r_i \geq 0$. Note that

$$\frac{\partial}{\partial r_i}\text{LB}_{i\text{I}} = -\frac{1}{2}\text{LB}_{i\text{I}}\left[\frac{\big(1000^2 - 500^2\big)}{500 + \frac{1}{2}r_i\big(1000^2 - 500^2\big)}\right] < 0$$

and

$$\begin{aligned}\frac{\partial}{\partial r_i}\text{UB}_{i\text{II}} &= -\frac{1}{2}\text{UB}_{i\text{II}}\left[\frac{\big(500^2 - 250^2\big)}{250 + \frac{1}{2}r_i\big(500^2 - 250^2\big)}\right]\\ &= -\frac{1}{2}\text{LB}_{i\text{I}}\left[\frac{\big(500^2 - 250^2\big)}{250 + \frac{1}{2}r_i\big(500^2 - 250^2\big)}\right] < 0.\end{aligned}$$

Note further that

$$\frac{\big(1000^2 - 500^2\big)}{500 + \frac{1}{2}r_i\big(1000^2 - 500^2\big)} > \frac{\big(500^2 - 250^2\big)}{250 + \frac{1}{2}r_i\big(500^2 - 250^2\big)}.$$

To see this,

$$
\begin{aligned}
&\frac{(1000^2 - 500^2)}{500 + \frac{1}{2}r_i(1000^2 - 500^2)} > \frac{(500^2 - 250^2)}{250 + \frac{1}{2}r_i(500^2 - 250^2)},\\
&250(1000^2 - 500^2) > 500(500^2 - 250^2),\\
&(1000 - 500)(1000 + 500) > 2(500 - 250)(500 + 250),\\
&(1000 - 500)(1000 + 500) > (500 - 250)(1000 + 500),\\
&(1000 - 500) > (500 - 250).
\end{aligned}
$$

It follows that $\frac{\partial}{\partial r_i}\text{LB}_{i\text{I}} < \frac{\partial}{\partial r_i}\text{UB}_{i\text{I}} < 0$. Thus, increasing r_i yields $\text{LB}_{i\text{I}} < \text{UB}_{i\text{II}}$.

E. Power of revealed preference test

As explained in footnote 52, Dean and Martin (forthcoming) propose a modification of Beatty and Crawford's (2011) success measure that, in our application, calls for replacing Bronars' alternative of uniform random choice in each coverage with an alternative of random choice according to the marginal empirical distribution of choices in each coverage. Table S4 reports the Beatty–Crawford success measure, under Dean and Martin's alternative, for monotonicity, unit slope, KR loss aversion, Gul disappointment aversion, and zero intercept. Naturally, under Dean and Martin's alternative, which is closer to the null, the pass rates are higher and the Beatty–Crawford statistics are lower for each shape restriction. Nevertheless, the results continue to favor a model with unit slope probability distortions and a model with monotone probability distortions over the other models considered.

Table S4. Power of revealed preference test (Dean–Martin alternative). Rationalizable subsample (3629 households).

	(a)	(b)	(c)		(d)
	Percentage of Households Satisfying Restriction				
Shape Restriction	Actual	Empirically-Weighted Random Choice	95 Percent Confidence Interval		Beatty–Crawford Success Measure
Monotonicity	84.8	72.8	71.6	73.9	12.1
Unit slope	61.6	44.3	42.9	45.7	17.3
KR loss aversion	42.2	31.7	30.4	32.8	10.5
Gul disappointment aversion	43.0	31.6	30.3	32.7	11.4
Zero intercept	39.6	29.4	28.1	30.5	10.3

Note: Column (a) reports results for the actual data. Column (b) reports means across 200 simulated data sets, each comprising 3629 observations of three deductible choices (one for each coverage), where each choice is drawn randomly from the coverage-specific empirical distribution of observed choices. Column (c) reports 95 percent confidence intervals for the means reported in column (b). The Beatty–Crawford success measure is the difference between columns (a) and (b).

F. Minimum distance Ω

F.1 *Identification, consistency, and asymptotic normality*

In this section, we prove that under mild conditions (satisfied in our data) the parameter vector $\boldsymbol{\theta}$ is point identified, and we establish the consistency and asymptotic normality of our sample analog estimator. We also demonstrate that its critical values can be consistently approximated by nonparametric bootstrap.

We estimate a linear point predictor

$$\begin{aligned}
&\widetilde{\Omega}(\mu_{ij}) = \boldsymbol{\theta}'\mathbf{m}_{ij},\\
&\boldsymbol{\theta} \equiv (a, b, c, d, e, f),\\
&\mathbf{m}_{ij} \equiv \big(1, \mu_{ij}, (\mu_{ij})^2, (\mu_{ij})^3, (\mu_{ij})^4, (\mu_{ij})^5\big),
\end{aligned}$$

obtained by finding the value of $\boldsymbol{\theta}$ that minimizes the expected average Euclidean distance from $\widetilde{\Omega}(\mu_{ij})$ to the random intervals $\mathcal{I}_{ij} \equiv [\mathrm{LB}_{ij}, \mathrm{UB}_{ij}]$, which result from the revealed preference arguments and the stability, CARA, and plausibility restrictions as explained in the paper, and where the average is taken over $j \in \{L, M, H\}$. We restrict the analysis to the subsample of monotone households, for a sample of size $N = 3079$. In what follows we let the members of this subsample be denoted $i = 1, \dots, N$. We recall that with CARA utility, LB_{ij} and UB_{ij} do not depend on wealth. Moreover, we recall that the values of LB_{ij} and UB_{ij} do not depend on μ_{ij}: only the relative locations of a household's Ω-intervals depend on its claim probabilities.

For a given point $t \in \mathbb{R}$ and interval $T = [\tau_L, \tau_U]$, let

$$d(t, T) = \inf_{\tau \in T} |t - \tau| = \max\big\{(\tau_L - t)_+, (t - \tau_U)_+\big\},$$

where $(z)_+ = \max(0, z)$. Then our point predictor satisfies

$$\begin{aligned}
\boldsymbol{\theta}_0 &\in \arg\min_{\boldsymbol{\theta}\in\Theta} E\left[\frac{1}{3}\sum_j d\big(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij}\big)\right]\\
&= \arg\min_{\boldsymbol{\theta}\in\Theta} E\left[\frac{1}{3}\sum_j \max\big\{\big(\mathrm{LB}_{ij} - \boldsymbol{\theta}'\mathbf{m}_{ij}\big)_+, \big(\boldsymbol{\theta}'\mathbf{m}_{ij} - \mathrm{UB}_{ij}\big)_+\big\}\right],
\end{aligned}$$

where Θ is a compact and convex parameter space and the term inside square brackets is the average distance of the point predictor to the intervals in the three contexts.

For brevity, we denote by p_{ij} the premium that household i pays in context j for chosen deductible d_{ij}. We now show that $\boldsymbol{\theta}_0$ is the unique minimizer of $E[\frac{1}{3}\sum_j d(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij})]$, we propose a sample analog estimator of $\boldsymbol{\theta}_0$, and we establish its consistency.

Theorem 1. *Suppose that we observe an independent and identically distributed (i.i.d.) sample $\{(p_{ij}, d_{ij}, \mu_{ij})_{j=L,M,H}\}_{i=1}^N$ from the joint distribution of $\{(p_j, d_j, \mu_j)_{j=L,M,H}\}$, such that for each $j \in \{L, M, H\}$, $\Pr(\mathrm{LB}_j \le \mathrm{UB}_j) = 1$ and assume that $\sum_{j\in\{L,M,H\}} \Pr(0 < \mathrm{LB}_j, \mathrm{UB}_j < 1) > 0$. Assume that the support of each p_j, $j = L, M, H$, is $\mathbb{R}_{++}$ and conditional on $(d_j)_{j=L,M,H}$, $\{(\mu_j, p_j)_{j=L,M,H}\}$ have an absolutely continuous joint distribution* (*with respect to Lebesgue measure*). *Assume that the parameter space* Θ *is compact and convex. Let*

$$\boldsymbol{\theta}_0 \in \arg\min_{\boldsymbol{\theta}\in\Theta} E\left[\frac{1}{3}\sum_{j\in\{L,M,H\}} d\big(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij}\big)\right],$$

$$\boldsymbol{\theta}_N \in \arg\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{N}\frac{1}{3}\sum_{i=1}^{N}\sum_{j\in\{L,M,H\}} d\big(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij}\big).$$

Then $\boldsymbol{\theta}_0$ *is the unique minimizer of* $E[\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij})]$ *and*

$$\|\boldsymbol{\theta}_N - \boldsymbol{\theta}_0\| \xrightarrow{p} 0 \quad \textit{as } N \longrightarrow \infty.$$

Proof. We verify assumptions (i)–(iv) in Newey and McFadden (1994), from which the result follows.

Assumption (i) of Newey and McFadden (1994) requires $E[\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij})]$ to be uniquely minimized at $\boldsymbol{\theta}_0$. Observe that the objective function is convex in $\boldsymbol{\theta}$ because $\mathcal{I}_{ij}$ is a convex set and the sum of convex functions yields a convex function. Hence its set of minimizers is convex. Suppose $\boldsymbol{\theta}_1$ is also a minimizer of $E[\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij})]$. For any $\gamma \in (0,1)$ let $\boldsymbol{\theta}_\gamma = \gamma\boldsymbol{\theta}_0 + (1-\gamma)\boldsymbol{\theta}_1$, and for $u \in \{+1,-1\}$ let $h(\mathcal{I}_{ij}, u) = \max\{u\mathrm{LB}_j, u\mathrm{UB}_j\}$. Then

$$\begin{aligned}
&E\left[\frac{1}{3}\sum_{j\in\{L,M,H\}} d\big(\boldsymbol{\theta}_\gamma'\mathbf{m}_{ij}, \mathcal{I}_{ij}\big)\right] \\
&\quad= \frac{1}{3}\sum_{j\in\{L,M,H\}} E\Big[\max_{u\in\{+1,-1\}}\big(u\boldsymbol{\theta}_\gamma'\mathbf{m}_{ij} - h(\mathcal{I}_{ij}, u)\big)_+\Big] \\
&\quad= \frac{1}{3}\sum_{j\in\{L,M,H\}} E\Big[\max_{u\in\{+1,-1\}}\big(\gamma\big(u\boldsymbol{\theta}_0'\mathbf{m}_{ij} - h(\mathcal{I}_{ij}, u)\big) \\
&\qquad + (1-\gamma)\big(u\boldsymbol{\theta}_1'\mathbf{m}_{ij} - h(\mathcal{I}_{ij}, u)\big)\big)_+\Big] \\
&\quad\leq \frac{1}{3}\sum_{j\in\{L,M,H\}} E\Big[\Big(\max_{u\in\{+1,-1\}}\gamma\big(u\boldsymbol{\theta}_0'\mathbf{m}_{ij} - h(\mathcal{I}_{ij}, u)\big) \\
&\qquad + \max_{u\in\{+1,-1\}}(1-\gamma)\big(u\boldsymbol{\theta}_1'\mathbf{m}_{ij} - h(\mathcal{I}_{ij}, u)\big)\Big)_+\Big].
\end{aligned}$$

Observe that

$$\begin{aligned}
&\Big(\max_{u\in\{+1,-1\}}\gamma\big(u\boldsymbol{\theta}_0'\mathbf{m}_{ij} - h(\mathcal{I}_{ij}, u)\big) + \max_{u\in\{+1,-1\}}(1-\gamma)\big(u\boldsymbol{\theta}_1'\mathbf{m}_{ij} - h(\mathcal{I}_{ij}, u)\big)\Big)_+ \\
&\quad\leq \gamma\Big(\max_{u\in\{+1,-1\}} u\boldsymbol{\theta}_0'\mathbf{m}_{ij} - h(\mathcal{I}_{ij}, u)\Big)_+ + (1-\gamma)\Big(\max_{u\in\{+1,-1\}} u\boldsymbol{\theta}_1'\mathbf{m}_{ij} - h(\mathcal{I}_{ij}, u)\Big)_+,
\end{aligned}$$

and a strict inequality holds if and only if

$$\Big(\max_{u\in\{+1,-1\}} u\boldsymbol{\theta}_0'\mathbf{m}_{ij} - h(\mathcal{I}_{ij},u)\Big)\Big(\max_{u\in\{+1,-1\}} u\boldsymbol{\theta}_1'\mathbf{m}_{ij} - h(\mathcal{I}_{ij},u)\Big) < 0.$$

This occurs if and only if $\boldsymbol{\theta}_0'\mathbf{m}_{ij} \notin \mathcal{I}_{ij}$ and $\boldsymbol{\theta}_1'\mathbf{m}_{ij} \in \mathcal{I}_{ij}$ or $\boldsymbol{\theta}_0'\mathbf{m}_{ij} \in \mathcal{I}_{ij}$ and $\boldsymbol{\theta}_1'\mathbf{m}_{ij} \notin \mathcal{I}_{ij}$. Hence, if for all $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_0$ such that $\boldsymbol{\theta}_1 \in \Theta$,

$$\sum_{j\in\{L,M,H\}} \big(\Pr\big(\boldsymbol{\theta}_0'\mathbf{m}_{ij} \notin \mathcal{I}_{ij}, \boldsymbol{\theta}_1'\mathbf{m}_{ij} \in \mathcal{I}_{ij}\big) + \Pr\big(\boldsymbol{\theta}_0'\mathbf{m}_{ij} \in \mathcal{I}_{ij}, \boldsymbol{\theta}_1'\mathbf{m}_{ij} \notin \mathcal{I}_{ij}\big)\big) > 0,$$

the objective function is strictly convex at $\boldsymbol{\theta}_0$, and therefore $\boldsymbol{\theta}_0$ is its unique minimizer. To see that this condition is satisfied, consider the event $(\mathrm{LB}_j \le \boldsymbol{\theta}_0'\mathbf{m}_{ij} \le \mathrm{UB}_j)$ and suppose it has positive probability for at least one $j \in \{H, L, M\}$. If $\mathrm{LB}_j = \mathrm{UB}_j$ it immediately follows that $\boldsymbol{\theta}_1'\mathbf{m}_{ij} \notin [\mathrm{LB}_j, \mathrm{UB}_j]$. Hence suppose $\mathrm{LB}_j < \mathrm{UB}_j$. Now we want to show that with positive probability $\boldsymbol{\theta}_0'\mathbf{m}_{ij} \in [\mathrm{LB}_j, \mathrm{UB}_j]$ and $\boldsymbol{\theta}_1'\mathbf{m}_{ij} \notin [\mathrm{LB}_j, \mathrm{UB}_j]$. Because p_j has full support on $\mathbb{R}_{++}$ and because LB_j and UB_j depend on p_j, we can find a set of (p_j, μ_j) of positive probability where either $\boldsymbol{\theta}_1'\mathbf{m}_{ij} < \mathrm{LB}_j \le \boldsymbol{\theta}_0'\mathbf{m}_{ij} \le \mathrm{UB}_j$ or $\mathrm{LB}_j \le \boldsymbol{\theta}_0'\mathbf{m}_{ij} \le \mathrm{UB}_j < \boldsymbol{\theta}_1'\mathbf{m}_{ij}$ holds. Hence, the result follows.

Assumptions (ii) and (iii) in Newey and McFadden (1994) are immediately satisfied, because we have assumed Θ to be compact and because $E[\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij})]$ is convex in $\boldsymbol{\theta}$ and therefore continuous in $\boldsymbol{\theta}$.

Assumption (iv) in Newey and McFadden (1994) requires

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{3L}\sum_{i=1}^{N}\sum_{j\in\{L,M,H\}} d\big(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij}\big) - E\left[\frac{1}{3}\sum_{j\in\{L,M,H\}} d\big(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij}\big)\right]\right| \xrightarrow{p} 0$$
$$\text{as } N \longrightarrow \infty.$$

This uniform convergence obtains observing that for any $\boldsymbol{\theta} \in \Theta$ and for each $j \in \{L, M, H\}$,

$$\begin{aligned}\max_{u\in\{+1,-1\}} \big(u\boldsymbol{\theta}'\mathbf{m}_{ij} - h(\mathcal{I}_{ij},u)\big)_+ &\le \max_{u\in\{+1,-1\}} \big|u\boldsymbol{\theta}'\mathbf{m}_{ij} - h(\mathcal{I}_{ij},u)\big| \\ &\le \|\boldsymbol{\theta}\|\|\mathbf{m}_{ij}\| + \max_{u\in\{+1,-1\}}\big|h(\mathcal{I}_{ij},u)\big| \\ &\le \|\boldsymbol{\theta}\|\|\mathbf{m}_{ij}\| + \max\big\{|\mathrm{LB}_j|, |\mathrm{UB}_j|\big\}.\end{aligned}$$

Because Θ is a compact set, $\|\boldsymbol{\theta}\|$ is bounded for any $\boldsymbol{\theta} \in \Theta$; because $\mu_j \in [0,1]$, $\|\mathbf{m}_{ij}\|$ is bounded; and because $\Pr(0 \le \mathrm{LB}_j \le \mathrm{UB}_j \le 1) = 1$, also $\max\{|\mathrm{LB}_j|, |\mathrm{UB}_j|\}$ is almost surely bounded. Hence, $E[\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij})] < \infty$, and therefore for each $\boldsymbol{\theta} \in \Theta$, by the weak law of large numbers for i.i.d. random variables,

$$\left|\frac{1}{3L}\sum_{i=1}^{N}\sum_{j\in\{L,M,H\}} d\big(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij}\big) - E\left[\frac{1}{3}\sum_{j\in\{L,M,H\}} d\big(\boldsymbol{\theta}'\mathbf{m}_{ij}, \mathcal{I}_{ij}\big)\right]\right| \xrightarrow{p} 0$$
$$\text{as } N \longrightarrow \infty.$$

Recalling that $\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij})$ is a convex function of $\boldsymbol{\theta}$, uniform convergence follows from Pollard's convexity lemma (Pollard (1991)). □

Next, we show asymptotic normality of our estimator.

Theorem 2. *Let the assumptions of Theorem 1 hold, and assume that for each* d_j, LB_j, UB_j *have an absolutely continuous distribution* (*with respect to Lebesgue measure*) *with density function* $f_{s_j}(t)$, $s_j = \mathrm{LB}_j, \mathrm{UB}_j$ *such that for each* $j \in \{L, M, H\}$, $E[(f_{\mathrm{UB}_j}(\boldsymbol{\theta}_0'\mathbf{m}_{ij}) + f_{\mathrm{LB}_j}(\boldsymbol{\theta}_0'\mathbf{m}_{ij}))\mathbf{m}_{ij}\mathbf{m}_{ij}']$ *exists and is nonsingular. Then*

$$\sqrt{N}(\boldsymbol{\theta}_N - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(0, \Sigma),$$

where Σ *is nonsingular and provided in equation* (2) *below.*

Proof. We establish the result by verifying the conditions of Example 3.2.22 in van der Vaart and Wellner (1996), which in turn verify the conditions of their Theorem 3.2.16.

By the triangle inequality, $d(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij})$ is Lipschitz in $\boldsymbol{\theta}$, and in particular for any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta$,

$$\begin{aligned}
&\left|\frac{1}{3}\sum_{j\in\{L,M,H\}}\left[d\left(\boldsymbol{\theta}_1'\mathbf{m}_{ij},\mathcal{I}_{ij}\right) - d\left(\boldsymbol{\theta}_2'\mathbf{m}_{ij},\mathcal{I}_{ij}\right)\right]\right| \\
&\quad\le \frac{1}{3}\sum_{j\in\{L,M,H\}}\left|d\left(\boldsymbol{\theta}_1'\mathbf{m}_{ij},\mathcal{I}_{ij}\right) - d\left(\boldsymbol{\theta}_2'\mathbf{m}_{ij},\mathcal{I}_{ij}\right)\right| \\
&\quad\le \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|\frac{1}{3}\sum_{j\in\{L,M,H\}}\|\mathbf{m}_{ij}\|.
\end{aligned}$$

This verifies the first condition in Example 3.2.22 in van der Vaart and Wellner (1996).

Next, observe that for any $\boldsymbol{\theta} \in \Theta$ such that $\boldsymbol{\theta}'\mathbf{m}_{ij}$ is in the interior of $\mathcal{I}_{ij}$, the gradient of $d(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij})$ with respect to $\boldsymbol{\theta}$ exists and is equal to 0. For any $\boldsymbol{\theta} \in \Theta$ such that $\boldsymbol{\theta}'\mathbf{m}_{ij} \notin \mathcal{I}_{ij}$,

$$\begin{aligned}
&\nabla_{\boldsymbol{\theta}}\left(\frac{1}{3}\sum_{j\in\{L,M,H\}} d\left(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij}\right)\right) \\
&\quad= \frac{1}{3}\sum_{j\in\{L,M,H\}}\mathbf{m}_{ij}\left[-1\left(\mathrm{LB}_{ij} - \boldsymbol{\theta}'\mathbf{m}_{ij} > 0\right) + 1\left(\boldsymbol{\theta}'\mathbf{m}_{ij} - \mathrm{UB}_{ij} > 0\right)\right].
\end{aligned} \tag{1}$$

For any $\boldsymbol{\theta} \in \Theta$ such that $\boldsymbol{\theta}'\mathbf{m}_{ij} = \mathrm{LB}_{ij}$ or $\boldsymbol{\theta}'\mathbf{m}_{ij} = \mathrm{UB}_{ij}$, the directional derivatives do not coincide. However, under our assumptions of full support for p (and w) on $\mathbb{R}_{++}$, this happens with probability 0. On the other hand, $\frac{1}{3}\sum_{j\in\{L,M,H\}}\Pr(\boldsymbol{\theta}_0'\mathbf{m}_{ij} \notin \mathcal{I}_{ij}(m_j)) > 0$. Hence, observing that each element on the right hand side of equation (1) is bounded by 1 in absolute value, we obtain

$$
\begin{aligned}
&E\Bigg[\bigg(\frac{1}{3}\sum_{j\in\{L,M,H\}}\big(d(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij}) - d(\boldsymbol{\theta}'\mathbf{m}_{ij0},\mathcal{I}_{ij})\big) \\
&\qquad - (\boldsymbol{\theta}-\boldsymbol{\theta}_0)'\nabla_{\boldsymbol{\theta}}\bigg(\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij})\bigg)\bigg)^2\Bigg] \\
&\quad = o\big(\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|^2\big).
\end{aligned}
$$

This verifies the second condition in Example 3.2.22 in van der Vaart and Wellner (1996).

Consistency of $\boldsymbol{\theta}_N$ for $\boldsymbol{\theta}_0$ is established in Theorem 1. We are left to show that the map $\boldsymbol{\theta}\longmapsto E[\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij})]$ is twice continuously differentiable at $\boldsymbol{\theta}_0$ with nonsingular second derivative matrix V. Observe that

$$
\begin{aligned}
V &= \frac{\partial}{\partial\boldsymbol{\theta}}E\bigg[\nabla_{\boldsymbol{\theta}}\bigg(\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij})\bigg)\bigg]\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\
&= \frac{1}{3}\sum_{j\in\{L,M,H\}}\frac{\partial}{\partial\boldsymbol{\theta}}E\big[\mathbf{m}_{ij}\{\Pr(\boldsymbol{\theta}'\mathbf{m}_{ij}-\mathrm{UB}_j>0|\mu_j) \\
&\qquad - \Pr(\mathrm{LB}_j-\boldsymbol{\theta}'\mathbf{m}_{ij}>0|\mu_j)\}\big]\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\
&= \frac{1}{3}\sum_{j\in\{L,M,H\}} E\big[\big(f_{\mathrm{UB}_j}(\boldsymbol{\theta}_0'\mathbf{m}_{ij}) + f_{\mathrm{LB}_j}(\boldsymbol{\theta}_0'\mathbf{m}_{ij})\big)\mathbf{m}_{ij}\mathbf{m}_{ij}'\big].
\end{aligned}
$$

It follows that V is nonsingular.

Finally, using the result in Example 3.2.22 in van der Vaart and Wellner (1996), we obtain

$$
\begin{aligned}
\Sigma &= V^{-1}E\bigg[\nabla_{\boldsymbol{\theta}}\bigg(\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij})\bigg)'\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\
&\qquad \times\nabla_{\boldsymbol{\theta}}\bigg(\frac{1}{3}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij})\bigg)'\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\bigg]V^{-1}.
\end{aligned} \tag{2}
$$

□

Last, we show that the critical values of the asymptotic distribution in Theorem 2 can be consistently approximated by nonparametric bootstrap.

Corollary 3. *Let the assumptions of Theorem 2 hold. Let* $F_N(t)=P(\sqrt{N}(\boldsymbol{\theta}_N-\boldsymbol{\theta}_0)\le t)$ *and* $F_B(t)=P_B(\sigma_N^{-1}\sqrt{N}(\boldsymbol{\theta}_N^B-\boldsymbol{\theta}_N)\le t)$, *where* P_B *is the probability conditional on the data,* $\sigma_N^2=\frac{N-1}{N}$, *and* $\boldsymbol{\theta}_N^B\equiv\arg\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{N}\frac{1}{3}\sum_{i=1}^N w_{Ni}\sum_{j\in\{L,M,H\}} d(\boldsymbol{\theta}'\mathbf{m}_{ij},\mathcal{I}_{ij})$, *with* $(w_{N1},\dots,w_{NN})\sim\mathrm{Multinomial}(N;1/N,\dots,1/N)$, *is the classical Efron bootstrap estima-*

tor. Then

$$\sup_{t\in\mathbb{R}}\left|F_B(t)-F_N(t)\right|=o_p(1)\quad \text{as } N\to\infty.$$

PROOF. The result follows immediately by observing that the assumptions of Theorem 2.4 in Bose and Chatterjee (2003) are verified in the proofs of our Theorems 1 and 2. □

F.2 *Lower order polynomials*

In this section, we show that the minimum distance Ω is robust to specifying a lower order polynomial. Figure S4 compares the minimum distance Ω that we present in the paper, which is based on a fifth-degree polynomial, with the minimum distance Ω that would result if we instead specify a second- or first-degree polynomial. As the figure shows, all three specifications yield very similar functions.

G. RANK CORRELATION OF CHOICES

In this section, we show that the results presented in Table 5 are very similar under quadraticity and linearity. Table S5 is an extension of Table 5. Column (c) breaks out the rank correlations for the rationalizable households that satisfy and violate quadraticity. Column (d) breaks out the rank correlations for the rationalizable households that satisfy and violate linearity.

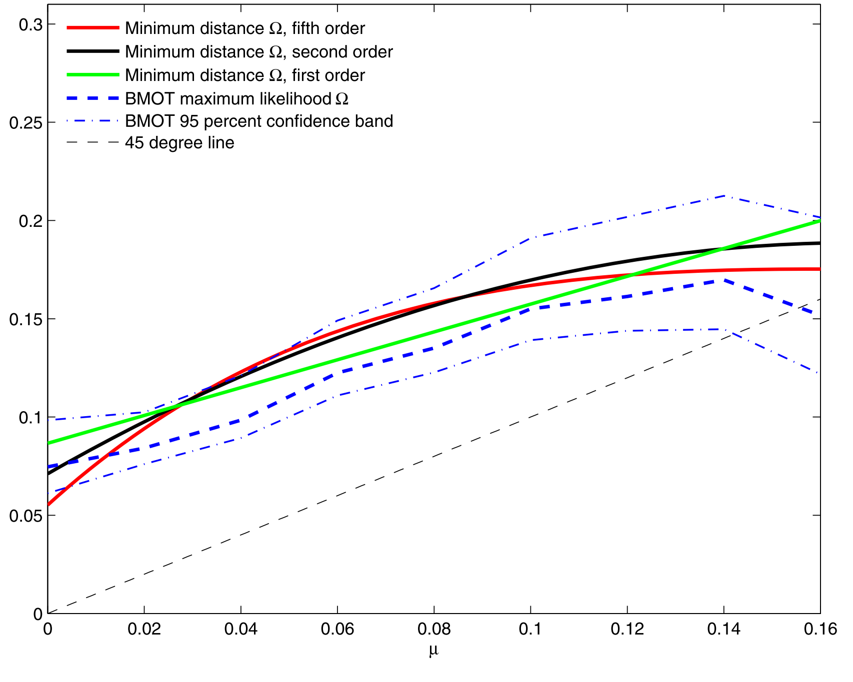

FIGURE S4. Minimum distance Ω.

Table S5. Rank correlation of deductible choices. Rationalizable subsample (3629 households).

	(a)	(b)		(c)		(d)	
	All Rationalizable Households (100 Percent)	Rationalizable Households That Satisfy Monotonicity (84.8 Percent)	Rationalizable Households That Violate Monotonicity (15.2 Percent)	Rationalizable Households That Satisfy Quadraticity (82.0 Percent)	Rationalizable Households That Violate Quadraticity (18.0 Percent)	Rationalizable Households That Satisfy Linearity (80.4 Percent)	Rationalizable Households That Violate Linearity (19.6 Percent)
Auto collision and auto comprehensive	0.490*	0.553*	0.335*	0.562*	0.359*	0.563*	0.334*
Auto collision and home	0.290*	0.363*	−0.019	0.409*	−0.068	0.390*	−0.062
Auto comprehensive and home	0.285*	0.352*	0.029	0.358*	−0.056	0.349*	−0.023

Note: Each cell reports a pairwise Spearman rank correlation coefficient. *Significant at the 1 percent level.

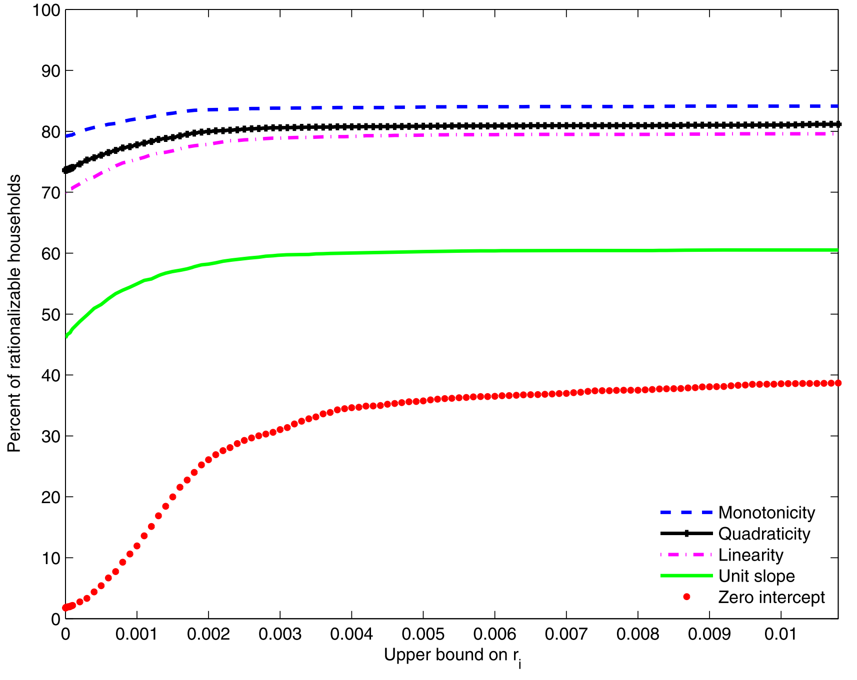

FIGURE S5. Full pricing menu in home.

H. PRICING MENU IN HOME

In this section, we show that including the \$2500 and \$5000 deductible options in the home menu would not materially change our results. Figure S5 displays the percentage of rationalizable households that satisfy each shape restriction as we increase the upper bound on r_i from 0 to 0.0108, after restoring the \$2500 and \$5000 deductible options to the home menu. The patterns displayed in Figure S5 are nearly identical to the patterns displayed in Figure 2.

I. ASYMMETRIC INFORMATION

In this section, we address the concern that the asymmetric information twins—moral hazard (unobserved action) and adverse selection (unobserved type)—may be biasing our claim rate estimates and hence our results.

I.1 *Moral hazard*

Throughout our analysis, we assume that deductible choice does not influence claim risk. That is, we assume there is no deductible-related moral hazard. There are two types of moral hazard that might operate in our setting. First, a household's deductible choice might influence its incentives to take care (ex ante moral hazard). Second, a household's deductible choice might influence its incentives to file a claim after experiencing a loss (ex post moral hazard), especially if its premium is experience rated or if the loss results in a "nil" claim (i.e., a claim that does not exceed its deductible). For either type of moral hazard, the incentive to alter behavior (take more care or file fewer claims) is stronger for households with larger deductibles. Hence, we investigate whether moral hazard is a significant issue in our data by examining whether our claim rate estimates change if we exclude households with high deductibles.

Specifically, we rerun our claim rate regressions using a restricted sample of the full data set in which we drop all household–coverage–year records with deductibles of $1000 or larger. We then use the new estimates to generate new claim rates for all households in the core sample (including those with deductibles of $1000 or larger). Comparing the new claim rates with the benchmark claim rates, we find that they are essentially indistinguishable: in each coverage, pairwise correlations exceed 0.995 and linear regressions yield intercepts less than 0.001 and coefficients of determination (R^2) greater than 0.99. Moreover, the estimates of the variance of unobserved heterogeneity in claim rates are nearly identical.[3]

The foregoing analysis suggests that moral hazard is not a significant issue in our data. This is perhaps not surprising, for two reasons. First, the empirical evidence on moral hazard in auto insurance markets is mixed. (We are not aware of any empirical evidence on moral hazard in home insurance markets.) Most studies that use "positive correlation" tests of asymmetric information in auto insurance do not find evidence of a correlation between coverage and risk (e.g., Chiappori and Salanié (2000); for a recent review of the literature, see Cohen and Siegelman (2010)).[4] Second, there are theoretical reasons to discount the force of moral hazard in our setting. In particular, because deductibles are small relative to the overall level of coverage, ex ante moral hazard strikes us as implausible in our setting.[5] As for ex post moral hazard, households have countervailing incentives to file claims no matter the size of the loss: under the terms of the company's policies, if a household fails to report a claimable event (especially an event that is a matter of public record; e.g., collision events typically entail police reports), it risks denial of all forms of coverage (notably liability coverage) for such event and also cancellation (or nonrenewal) of its policy.

[3]The revised estimates are 0.22, 0.56, and 0.44 in auto collision, auto comprehensive, and home, respectively, whereas the corresponding benchmark estimates are 0.22, 0.57, and 0.45.

[4]Beginning with Abbring, Chiappori, Heckman, and Pinquet (2003) and Abbring, Chiappori, and Pinquet (2003), a second strand of literature tests for moral hazard in longitudinal auto insurance data using various dynamic approaches. Abbring, Chiappori, and Pinquet (2003) find no evidence of moral hazard in French data. A handful of subsequent studies present some evidence of moral hazard using data from Canada and Europe. The only study of which we are aware that uses U.S. data is Israel (2004), which reports a small moral hazard effect for drivers in Illinois. Each of these studies, however, identifies a moral hazard effect with respect to either liability coverage or a composite coverage that confounds liability coverage with other coverages. None of them identifies a separate moral hazard attributable to the choice of deductible in the auto coverages we study.

[5]We note that Cohen and Einav (2007) reach the same conclusion. Furthermore, we note that the principal justification for deductibles is the insurer's administrative costs (Arrow (1963)).

Finally, we note that even if our claim rates are roughly correct, the possibility of nil claims could bias our results, as they violate our assumption that every claim exceeds the highest available deductible (which underlies how we define the deductible lotteries). To investigate this potential, we make the extreme counterfactual assumption that claimable events invariably result in losses between \$500 and \$1000—specifically \$750—and we recalculate (i) the distribution of the minimum plausible r_i for each shape restriction on $\Omega_i(\cdot)$ and (ii) the percentage of rationalizable households that satisfy each shape restriction on $\Omega_i(\cdot)$ as we increase the upper bound on r_i from 0 to 0.0108.[6] For each shape restriction, the distribution of the minimum plausible r_i shifts to the right, such that for each nondegenerate shape restriction the median increases from 0 to between 0.0005 and 0.0014, and for zero intercept the median increases from 0.0015 to 0.0029. Consequently, for values of r_i below 0.0014, for each shape restriction a lesser percentage of rationalizable households satisfy the restriction, though it still is the case that a greater percentage satisfy monotonicity, quadraticity, linearity, and unit slope than zero intercept. Importantly, however, for values of r_i above 0.0015, for each shape restriction the percentage of rationalizable households that satisfy the restriction increases rapidly and more or less returns to its benchmark level and trajectory. The intuition behind these findings is straightforward. Under the assumption that claimable events invariably result in losses of \$750, the lottery associated with a \$1000 deductible becomes $L_{1000} \equiv (-p_{1000}, 1-\mu; -p_{1000} - 750, \mu)$. This increases the lower bound on the Ω-interval for households choosing a deductible less than \$1000, and for many households the lower bound ends up exceeding the upper bound. The only way to restore $\mathrm{LB}_{ij} \le \mathrm{UB}_{ij}$ is then to increase r_i. Once that happens, the need for probability distortions remains more or less the same.

I.2 *Adverse selection*

I.2.1 *Heterogeneity unobserved by the econometrician*

In terms of adverse selection, the standard concern is that there may be heterogeneity in claim risk that is observed by the households but unobserved by the econometrician. That is, a household may have better information about its claim risk than does the econometrician. To assess the potential effect on our results of heterogeneity that may be unobserved by us, we utilize the distributions of $\exp(\varepsilon_{ij})$ that we estimated in the claim rate regressions in Section 2.2 to simulate the distribution of the percentage of rationalizable households that satisfy each restriction on $\Omega_i(\cdot)$. More specifically, for every rationalizable household i and every coverage j, we construct $\widetilde{\lambda}_{ij} = \exp(\mathbf{X}'_{ij}\widehat{\boldsymbol{\beta}}_j)\exp(\varepsilon_{ij})$, where $\exp(\varepsilon_{ij})$ is drawn from the gamma distribution estimated in the claim rate regression for coverage j, conditional on household i's ex post claims experience in coverage j. Next, we let $\widetilde{\mu}_{ij} \equiv 1 - \exp(-\widetilde{\lambda}_{ij})$ and we use $\widetilde{\mu}_{ij}$ in constructing the rationalizable households' Ω-intervals. We then recalculate the percentage of rationalizable households that satisfy each shape restriction on $\Omega_i(\cdot)$. We repeat this procedure 200 times and record the 5th, 25th, 50th, 75th, and 95th percentiles of each percentage.

[6] We emphasize that this is an extreme counterfactual assumption, as it surely is the case that most, if not all, claimable events result in losses that exceed \$1000.

Table S6. Unobserved heterogeneity in risk. Rationalizable subsample (3629 households).

Shape Restriction	(a) Percent of Households Satisfying Restriction With $\mu =$ Est. Claim Prob.	(b) Simulated Distribution					(c) Percent of Households Satisfying Restriction With $\mu =$ Avg. Claim Prob.
		5th Pctl.	25th Pctl.	50th Pctl.	75th Pctl.	95th Pctl.	
Monotonicity	84.8	84.0	84.4	84.6	84.7	85.1	86.2
Quadraticity	82.0	81.8	82.0	82.2	82.4	82.7	80.0
Linearity	80.4	80.5	80.7	80.9	81.2	81.4	78.6
Unit slope	61.6	55.9	56.4	56.9	57.3	57.7	62.1
Zero intercept	39.6	31.3	31.9	32.3	32.7	33.2	42.9

Table S6, column (b) reports the results. For each shape restriction, the 5th–95th interpercentile range is narrow, 1–2 percentage points. For monotonicity and quadraticity, the percentage we report in Table 1, column (a) (which is reproduced in Table S6, column (a) for the reader's convenience) lies between the 5th and 95th percentiles of the simulated distribution. It is unlikely, therefore, that unobserved heterogeneity is biasing our results and conclusions regarding monotone or quadratic probability distortions. For linearity, the percentage we report in Table 1 lies just below the 5th percentile of the simulated distribution. This suggests that our results may understate somewhat the extent to which the data are consistent with linear probability distortions. Conversely, for unit slope and zero intercept, the percentage we report in Table 1 exceeds the 95th percentile of the simulated distribution. This suggests that our results may overstate the extent to which the data are consistent with the unit slope distortions model and the objective expected utility model.

I.2.2 *Heterogeneity unobserved by the households* The reverse concern is that the econometrician may have better information about the households' claim risk than do the households themselves. To assess the potential effect on our results of heterogeneity that may be unobserved by the households, we recompute the percentage of rationalizable households that satisfy each shape restriction on $\Omega_i(\cdot)$ under the extreme assumption that, in each line of coverage, every household's claim probability corresponds to the sample mean reported in Table S3. Table S6, column (c) reports the results. The percentages increase under monotonicity, unit slope, and zero intercept, and decrease under quadraticity and linearity. Thus, if there is any bias, it does not operate in a consistent direction. Moreover, the differences are small, 2.0 percentage points or less, except in the case of zero intercept, where the difference is somewhat larger at 3.3 percentage points. Hence, if there is any bias, it likely is not material to our results and conclusions regarding probability distortions; at most, our results may understate somewhat the extent to which the data are consistent with the objective expected utility model. Of course, the potential bias here runs in the opposite direction of the potential bias from heterogeneity that is unobserved by the econometrician.

I.3 *Alternative claim probabilities*

In this subsection, we explore further the sensitivity of our results to our claim risk estimates. We consider three alternative cases: (i) claim probabilities that are derived from fitted claim rates that do not condition on ex post claims experience, (ii) claim probabilities that are half as large as our estimates, and (iii) claim probabilities that are twice as large as our estimates.

Figures S6, S7, and S8 display, for cases (i), (ii), and (iii), respectively, the percentage of rationalizable households that satisfy each shape restriction as we increase the upper bound on r_i from 0 to 0.0108. In case (i), the patterns are very similar to the patterns displayed in Figure 2. In cases (ii) and (iii), the patterns for monotonicity, quadraticity, and linearity are very similar to those displayed in Figure 2. However, the patterns for unit slope and zero intercept are somewhat different. Generally speaking, compared to the base case (Figure 2), the percentage of rationalizable households that satisfy unit slope and zero intercept are a bit higher in case (ii) and quite a bit lower in case (iii). This suggests that if we have grossly overestimated (resp. grossly underestimated) the households' claim probabilities, then our results may understate somewhat (resp. overstate

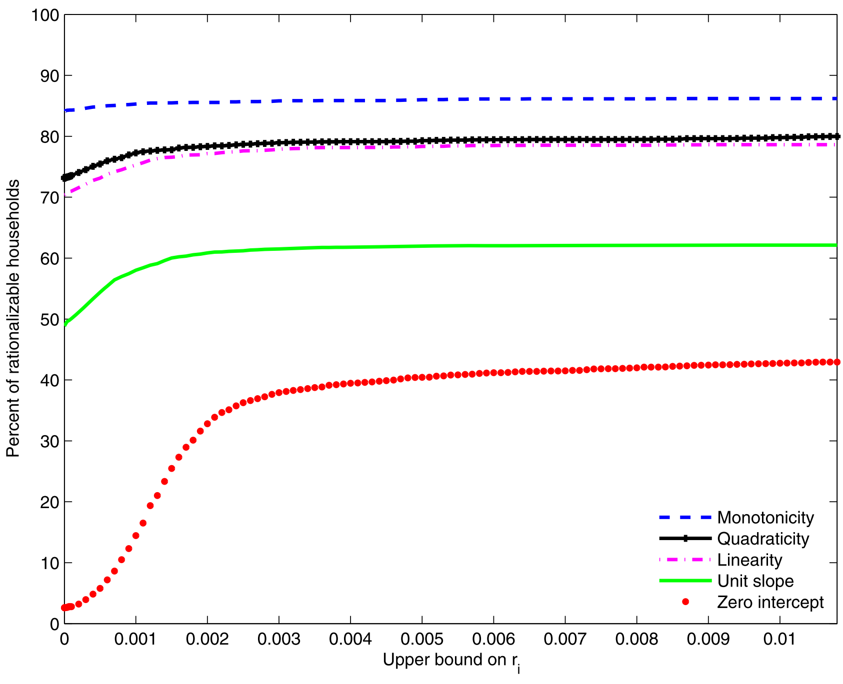

FIGURE S6. Unconditional claim probabilities.

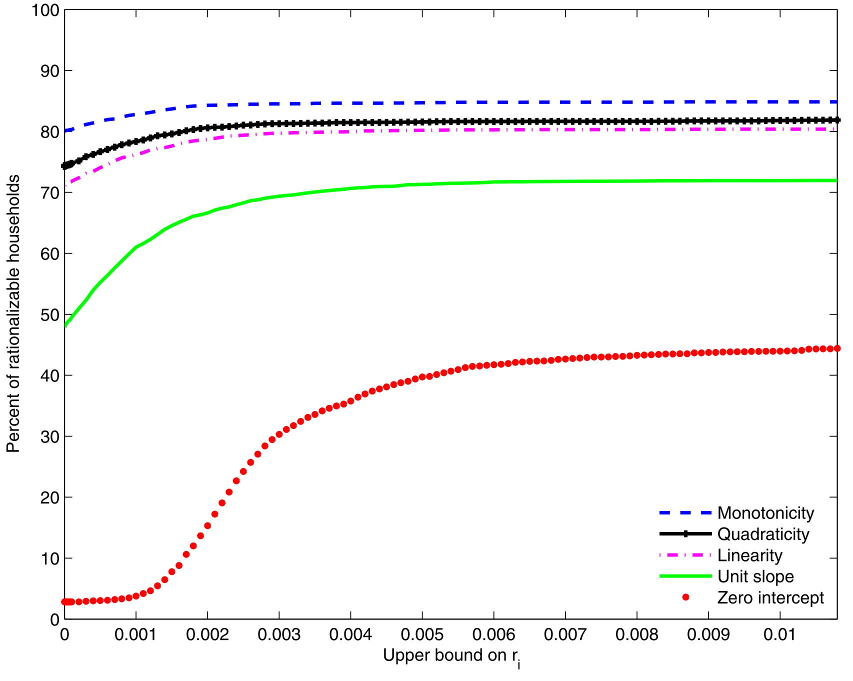

FIGURE S7. $\frac{1}{2}\times$ claim probabilities.

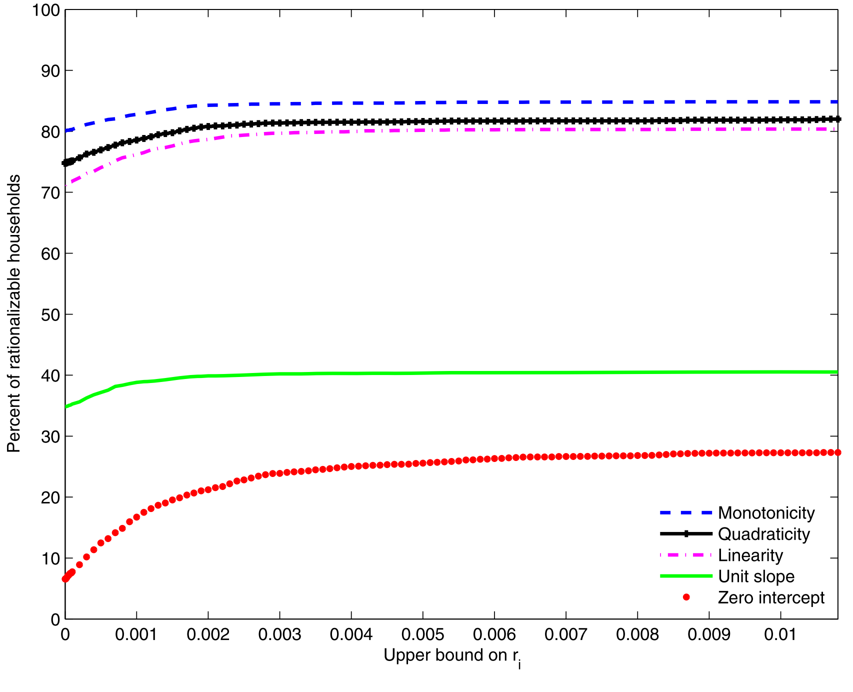

FIGURE S8. $2\times$ claim probabilities.

quite a bit) the extent to which the data are consistent with the unit slope distortions models and the objective expected utility model.

References

Abbring, J. H., P.-A. Chiappori, J. Heckman, and J. Pinquet (2003), "Adverse selection and moral hazard in insurance: Can dynamic data help to distinguish?" *Journal of the European Economic Association*, 1 (2–3), 512–521. [16]

Abbring, J. H., P.-A. Chiappori, and J. Pinquet (2003), "Moral hazard and dynamic insurance data." *Journal of the European Economic Association*, 1 (4), 767–820. [16]

Arrow, K. J. (1963), "Uncertainty and the welfare economics of medical care." *American Economic Review*, 53 (5), 941–969. [16]

Barseghyan, L., F. Molinari, T. O'Donoghue, and J. C. Teitelbaum (2013), "The nature of risk preferences: Evidence from insurance choices." *American Economic Review*, 103 (6), 2499–2529. [1]

Barseghyan, L., J. Prince, and J. C. Teitelbaum (2011), "Are risk preferences stable across contexts? Evidence from insurance data." *American Economic Review*, 101 (2), 591–631. [1]

Beatty, T. K. M. and I. A. Crawford (2011), "How demanding is the revealed preference approach to demand?" *American Economic Review*, 101 (6), 2782–2795. [7]

Bose, A. and S. Chatterjee (2003), "Generalized bootstrap for estimators of minimizers of convex functions." *Journal of Statistical Planning and Inference*, 117 (2), 225–239. [13]

Chiappori, P.-A. and B. Salanié (2000), "Testing for asymmetric information in insurance markets." *Journal of Political Economy*, 108 (1), 56–78. [16]

Cohen, A. and L. Einav (2007), "Estimating risk preferences from deductible choice." *American Economic Review*, 97 (3), 745–788. [1, 16]

Cohen, A. and P. Siegelman (2010), "Testing for adverse selection in insurance markets." *Journal of Risk and Insurance*, 77 (1), 39–84. [16]

Dean, M. and D. Martin (forthcoming), "Measuring rationality with the minimum cost of revealed preference violations." *Review of Economics and Statistics*. [7]

Israel, M. (2004), "Do we drive more safely when accidents are more expensive? Identifying moral hazard from experience rating schemes." CSIO Working Paper 0043, Northwestern University. [16]

Newey, W. K. and D. McFadden (1994), "Large sample estimation and hypothesis testing." In *Handbook of Econometrics*, Vol. 4 (R. F. Engle and D. L. McFadden, eds.), 2111–2245, Chapter 36, Elsevier, Amsterdam. [9, 10]

Pollard, D. (1991), "Asymptotics for least absolute deviation regression estimators." *Econometric Theory*, 7 (2), 186–199. [11]

van der Vaart, A. W. and J. A. Wellner (1996), *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer, New York. [11, 12]

Co-editor Rosa L. Matzkin handled this manuscript.

Submitted November, 2014. Final version accepted July, 2015.