Supplementary Appendix to Private Information and Insurance Rejections

For online publication only

This Appendix contains supplementary material for Private Information and Insurance Rejections. Appendix A contains the proof of Theorem 1 and discusses Remark 1. Appendix B provides proofs for the lower bound approach discussed in Section 4.1. Appendix C discusses the data, including details for the covariate specifications and the sample selection. Appendix D discusses the empirical specification for the lower bound estimator and presents additional robustness checks referred to in Section 6.5. Section E provides details on the specification and estimation of the structural approach referred to throughout Section 7. Finally, Section F presents the a selection of pages from the LTC underwriting guidelines of Genworth Financial.

A Theory Appendix

A.1 Proof of No-Trade Theorem

I prove the no-trade theorem in several steps. First, I translate the problem to a maximization problem in utility space. Second, I prove the converse of the theorem directly by constructing an implementable allocation other than the endowment when Condition (1) does not hold. Third, I prove the no trade theorem for a finite type distribution. Fourth, I approximate arbitrary distributions satisfying Condition (1) with finite type distributions and pass to the limit, thus proving the no trade theorem for a general type distribution.

Most of the steps of the proof are straightforward. Indeed, it is arguably quite obvious that condition (1) rules out the profitability of any pooling contract. The theoretical contribution is to show that condition (1) also rules out the profitability of separating contracts. Indeed, the ability for insurance companies to offer separating contracts is an important ingredient in previous models of this environment (Spence [1978], Riley [1979], Chade and Schlee [2011]). In Lemma (A.5), I show that condition (1) implies the profitability of a menu of contracts is bounded above by the profitability of a pooling allocation.

A.1.1 Utility Space

First, translate the problem to utility space so that the incentive and individual rationality constraints are linear in utility. Let $c(u) = u^{-1}(u)$ denote the inverse of the utility function $u(c)$, which is strictly increasing, continuously differentiable, and strictly convex. I denote the endowment allocation by $E = \{(c_L(p), c_{NL}(p))\}_p = \{(w-l, l)\}_p$. Let us denote the endowment allocation in utility space by $E^U = \{u(w-l), u(w)\}_p$. To fix units, I normalize $u_{NL}(1) = u(w)$.

Given a utility allocation $A^U = \{u_L(p), u_{NL}(p)\}_{p \in \Psi}$, denote the slack in the resource constraint by

$$\Pi(A^U) = \int [w - pl - pc(u_L(p)) - (1-p)c_{NL}(p)] dF(p)$$

I begin with a useful lemma that allows us to characterize when the endowment is the only implementable allocation.

Lemma A.1 (Characterization). *The endowment is the only implementable allocation if and only if E^U is the unique solution to the following constrained maximization program, $P1$*

$$P1 \quad : \quad \max_{\{u_L(p), u_{NL}(p)\}_p} \int [w - pl - pc(u_L(p)) - (1-p)c(u_{NL}(p))] dF(p)$$
$$s.t. \quad pu_L(p) + (1-p)u_{NL}(p) \geq pu_L(\hat{p}) + (1-p)u_L(\hat{p}) \quad \forall p, \hat{p} \in \Psi$$
$$pu_L(p) + (1-p)u_{NL}(p) \geq pu(w-l) + (1-p)u(w) \quad \forall p \in \Psi$$

Proof. Note that the constraint set is linear and the objective function is strictly concave. The first constraint is the incentive constraint in utility space. The second constraint is the individual rationality constraint in utility space. The linearity of the constraints combined with strict concavity of the objective function guarantees that the solutions are unique. Suppose that the endowment is the only implementable allocation and suppose, for contradiction, that the solution to the above program is not the endowment. Then, there exists an allocation $A^U = \{u_L(p), u_{NL}(p)\}$ such that $\int [w - pl - pc(u_L(p)) - (1-p)c(u_{NL}(p))] dF(p) > 0$ which also satisfies the IC and IR constraints. Therefore, A^U is implementable, which yields a contradiction. Conversely, suppose that there exists an implementable allocation B such that $B \neq E$. Let B^U denote the associated utility allocations to the consumption allocations in B. Then, B^U satisfies the incentive and individual rationality constraints. Since the constraints are linear, the allocations $C^U(t) = tB^U + (1-t)E^U$ lie in the constraint set. By strict concavity of the objective function, $\Pi(C^U(t)) > 0$ for all $t \in (0,1)$. Since $\Pi(E^U) = 0$, E^U cannot be the solution to the constrained maximization program. $\square$

The lemma allows me to focus attention on solutions to P1, a concave maximization program with linear constraints.

A.1.2 Necessity of the No Trade Condition

I begin the proof with the converse portion of the theorem: if the no-trade condition does not hold, then there exists an implementable allocation $A \neq E$ which does not utilize all resources and provides a strict utility improvement to a positive measure of types.

Lemma A.2 (Converse). *Suppose Condition (1) does not hold so that there exists $\hat{p} \in \Psi \backslash \{1\}$ such that $\frac{\hat{p}}{1-\hat{p}} \frac{u'(w-l)}{u'(w)} > \frac{E[P|P \geq \hat{p}]}{1 - E[P|P \geq \hat{p}]}$. Then, there exists an allocation $\hat{A}^U = \{(\hat{u}_L(p), \hat{u}_{NL}(p))\}_p$*

and a positive measure of types, $\hat{\Psi} \subset \Psi$, *such that*

$$p\hat{u}_L^A(p) + (1-p)\,\hat{u}_{NL}(p) > pu(w-l) + (1-p)\,u(w) \quad \forall p \in \hat{\Psi}$$

and

$$\int \left[W - pL - pc\left(\hat{u}_L(p)\right) - (1-p)\,c\left(\hat{u}_{NL}(p)\right) \right] dF(p)$$

Proof. The proof follows by constructing an allocation which is preferable to all types $p \geq \hat{p}$ and showing that the violation of Condition (1) at $\hat{p}$ ensures its profitability. Given $\hat{p} \in \Psi$, either $P = \hat{p}$ occurs with positive probability, or any open set containing $\hat{p}$ has positive probability. In the case that $\hat{p}$ occurs with positive probability, let $\hat{\Psi} = \{\hat{p}\}$. In the latter case, note that the function $E\left[P|P \geq p\right]$ is locally continuous in p at $\hat{p}$ so that WLOG the no-trade condition does not hold for a positive mass of types. WLOG, I assume $\hat{p}$ has been chosen so that there exists a positive mass of types $\hat{\Psi}$ such that $p \in \hat{\Psi}$ implies $p \geq \hat{p}$. Then, for all $p \in \hat{\Psi}$, I have $\hat{\Psi} \subset \Psi$ such that

$$\frac{p}{1-p}\frac{u'(w-l)}{u'(w)} > \frac{E\left[P|P \geq p\right]}{1 - E\left[P|P \geq p\right]} \quad \forall p \in \hat{\Psi}$$

Now, for $\varepsilon, \eta > 0$, consider the augmented allocation to types $p \in \hat{\Psi}$:

$$
\begin{aligned}
u_L(\varepsilon, \eta) &= u(w-l) + \varepsilon + \eta \\
u_{NL}(\varepsilon, \eta) &= u(w) - \frac{1-\hat{p}}{\hat{p}}\varepsilon
\end{aligned}
$$

Note that if $\eta = 0$, ε traces out the indifference curve of individual $\hat{p}$. Construct the utility allocation $A^U(\varepsilon, \eta)$ defined by

$$
(\hat{u}_L(p), \hat{u}_{NL}(p)) = \left\{
\begin{array}{cc}
\left(u(w-l) + \varepsilon + \eta, u(w) - \frac{\hat{p}}{1-\hat{p}}\varepsilon\right) & if \quad p \geq \hat{p} \\
(u(w-l), u(w)) & if \quad p < \hat{p}
\end{array}
\right.
$$

Note that for $\varepsilon > 0$ and $\eta > 0$ the utility allocation $(\hat{u}_L(p), \hat{u}_{NL}(p))$ is strictly preferred by all types $p \geq \hat{p}$ relative to the endowment utility allocation. Therefore, A_ε^U is individually rational and incentive compatible. I now only need to verify that there exists an allocation with $\varepsilon > 0$ and $\eta > 0$ which does not exhaust resources. I have

$$\Pi(\varepsilon, \eta) = \int \left[w - pl - pc\left(\hat{u}_L(p)\right) - (1-p)\,c\left(\hat{u}_{NL}(p)\right) \right] dF(p)$$

Notice that this is continuously differentiable in ε and η. Differentiating with respect to ε and evaluating at $\varepsilon = 0$ yields

$$\frac{\partial \Pi}{\partial \varepsilon}\Big|_{\varepsilon=0} = \int \left[-pc'\left(u(w-l+\eta)\right) + \frac{\hat{p}}{1-\hat{p}}(1-p)\,c'\left(u(w)\right) \right] 1\{p \geq \hat{p}\}\,dF(p)$$

3

which is strictly positive if and only if

$$E\left[P|P \geq \hat{p}\right] c'\left(u\left(w - l + \eta\right)\right) < \frac{\hat{p}}{1 - \hat{p}}\left(1 - E\left[P|P \geq \hat{p}\right]\right) c'\left(u\left(w\right)\right)$$

Notice that this is continuous in η. So, at $\eta = 0$, I have

$$\frac{\partial \Pi}{\partial \varepsilon}\Big|_{\varepsilon=0,\eta=0} > 0 \iff \frac{\hat{p}}{1 - \hat{p}} \frac{u'\left(w - l\right)}{u'\left(w\right)} > \frac{E\left[P|P \geq \hat{p}\right]}{1 - E\left[P|P \geq \hat{p}\right]}$$

and thus by continuity, the above condition holds for sufficiently small $\eta > 0$, proving the existence of an allocation which both delivers strictly positive utility for a positive fraction of types and does not exhaust all resources.

This shows that Condition (1) is necessary for the endowment to be the only implementable allocation. $\qquad \square$

A.1.3   Useful Results

Before showing that Condition (1) is sufficient for no trade, it is useful to have a couple of results characterizing solutions to P1.

Lemma A.3. *Suppose Condition (1) holds. Then for all $c_L, c_{NL} \in [w - l, l]$,*

$$\frac{p}{1 - p} \frac{u'\left(c_L\right)}{u'\left(c_{NL}\right)} \leq \frac{E\left[P|P \geq p\right]}{1 - E\left[P|P \geq p\right]} \quad \forall p \in \Psi \backslash \{1\}$$

and if $c_L, c_{NL} \in (w - l, l)$,

$$\frac{p}{1 - p} \frac{u'\left(c_L\right)}{u'\left(c_{NL}\right)} < \frac{E\left[P|P \geq p\right]}{1 - E\left[P|P \geq p\right]} \quad \forall p \in \Psi \backslash \{0, 1\}$$

Proof. Since $u'\left(c\right)$ is decreasing in c, I have $\frac{u'(c_L)}{u'(c_{NL})} \leq \frac{u'(w-l)}{u'(w)}$. Therefore, the result follows immediately from Condition (1). The strict inequality follows from strict concavity of $u\left(c\right)$. $\quad \square$

Lemma A.4. *In any solution to P1, $c_L\left(p\right) \geq w - l$ and $c_{NL}\left(p\right) \leq w$.*

Proof. Suppose $A = \{c_L\left(p\right), c_{NL}\left(p\right)\}_p$ is a solution to $P1$. First, suppose that $c_L\left(\hat{p}\right) < w - l$. For this contract to be individually rational, I must have $c_{NL}\left(\hat{p}\right) > w$. Incentive compatibility requires $c_L\left(p\right) \leq c_L\left(p\right) < w - l \ \forall p < \hat{p}$ and $c_{NL}\left(p\right) \geq c_{NL}\left(\hat{p}\right) > w \ \forall p < \hat{p}$. Consider the new allocation $\tilde{A} = \{\tilde{c}_L\left(p\right), \tilde{c}_{NL}\left(p\right)\}$ defined by

$$\tilde{c}_L\left(p\right) = \begin{cases} c_L\left(p\right) & if \ p > \hat{p} \\ w - l & if \ p \leq \hat{p} \end{cases}$$

4

$$\tilde{c}_{NL}\left(p\right)=\begin{cases}c_{NL}\left(p\right) & if\ p>\hat{p}\\ w & if\ p\le\hat{p}\end{cases}$$

Then $\tilde{A}$ is implementable (IC holds because of single crossing of the utility function). It only remains to show that $\Pi\left(A\right)<\Pi\left(\tilde{A}\right)$. But this follows trivially. Notice that the IR constraint and concavity of the utility function requires that points $\left(c_L\left(p\right),c_{NL}\left(p\right)\right)$ lie above the zero profit line $p\left(w-l-c_L\right)+\left(1-p\right)\left(w-c_{NL}\right)$. Thus, each point $\left(c_L\left(p\right),c_{NL}\left(p\right)\right)$ must earn negative profits at each $p\le\hat{p}$.

Now, suppose $c_{NL}\left(\hat{p}\right)>w$. Then, the incentive compatibility constraint requires $c_{NL}\left(p\right)>w\ \forall p\le\hat{p}$. Construct $\tilde{A}$ as above, yielding the same contradiction. $\qquad\square$

I now prove the theorem in two steps. First, I prove the result for a finite type distribution. I then pass to the limit to cover the case of arbitrary distributions.

A.1.4 Sufficiency of the No Trade Condition for Finite Types

To begin, suppose that $\Psi=\{p_1,...p_N\}$. I first show that Condition (1) implies that the solution to $P1$ is a pooling allocation which provides the same allocation to all types.

Lemma A.5. *Suppose $\Psi=\{p_1,...,p_N\}$ and that condition (1) holds (note that this requires $p_N=1$). Then, the solution to $P1$ is a full pooling allocation: there exists $\bar{u}_L,\bar{u}_{NL}$ such that $\left(u_L\left(p\right),u_{NL}\left(p\right)\right)=\left(\bar{u}_L,\bar{u}_{NL}\right)$ for all $p\in\Psi\backslash\{0,1\}$, $u_L\left(1\right)=\bar{u}_L$, $u_{NL}\left(0\right)=\bar{u}_{NL}$.*

Proof. Let $A^U=\{u_L^*\left(p\right),u_{NL}^*\left(p\right)\}_p$ denote the solution to P and suppose for contradiction that the solution to P is not a full pooling allocation. Let $\hat{p}=\min\{p|u_L^*\left(p\right)=u_L^*\left(1\right)\}$, let $\hat{p}_-=\max\{p|u_L^*\left(p\right)\ne u_L^*\left(1\right)\}$. The assumption that Ψ is finite implies that $\hat{p}>\hat{p}_-$. Let us define the pooling sets $J=\{p|u_L^*\left(p\right)=u_L^*\left(1\right)\}$ and $K=\{p|u_L^*\left(p\right)=u_L^*\left(\hat{p}_-\right)\}$. I will show that a profitable deviation exists which pools groups J and K into the same allocation. First, notice that if $\hat{p}=1$, then clearly it is optimal to provide group J with the same amount of consumption in the event of a loss as group K, since otherwise the IC constraint of the type $\hat{p}=1$ type would be slack. So, I need only consider the case $\hat{p}<1$.

Notice that if the IR constraint of any member of group J binds (i.e. if the IR constraint for $\hat{p}$ binds), then their IC constraint implies that the only possible allocation for the lower risk types $p<\hat{p}$ is the endowment. This standard result follows from single crossing of the utility function. Therefore, I have two cases. Either all types $\tilde{p}\in\Psi\backslash J$ receive their endowment, $\left(c_L,c_{NL}\right)=\left(w-l,w\right)$, or the IR constraint cannot bind for any member of J. I consider these two cases in turn.

Suppose $u_L^*\left(p\right)=u\left(w-l\right)$ and $u_{NL}^*\left(p\right)=u\left(w\right)$ for all types $\tilde{p}\in\Psi\backslash J$. Clearly, I must then have that the IR constraint must bind for type $\hat{p}$, since otherwise profitability could be improved by lowering the utility provided to types $\tilde{p}\in\Psi\backslash J$. I now show that the profitability

of the allocation violates the no-trade condition. The profitability of A^U is

$$\Pi\left(A^U\right) = \int_{p \in J} \left[w - pl - pc\left(u_L^*\left(\hat{p}\right)\right) - (1-p)\,c\left(u_{NL}^*\left(\hat{p}\right)\right)\right] dF\left(p\right)$$

Now, I construct the utility allocation A_t^U by

$$\left(u_L^t\left(p\right), u_{NL}^t\left(p\right)\right) = \begin{cases} \left(u\left(w-l\right) + t, u\left(w\right) - \frac{\hat{p}}{1-\hat{p}}t\right) & if \quad p \in J \\ \left(u\left(w-l\right), u\left(w\right)\right) & if \quad p \notin J \end{cases}$$

Since the IR constraint binds for type $\hat{p}$, I know that there exists $\hat{t}$ such that $A_{\hat{t}}^U = A^U$. By Lemma A.4, $\hat{t} > 0$ and A_t^U satisfies IC and IR for any $t \in \left[0, \hat{t} + \eta\right]$ for some $\eta > 0$. Since profits are maximized at $t = \hat{t}$ and since the objective function is strictly concave, it must be the case that

$$\frac{d\Pi\left(A_t^U\right)}{dt}\Big|_{t=\hat{t}} = 0$$

where

$$\frac{d\Pi\left(A_t^U\right)}{dt}\Big|_{t=\hat{t}} = \int_{p \in J} \left[pc'\left(u_L^*\left(p\right)\right) - (1-p)\,c'\left(u_{NL}^*\left(p\right)\right)\frac{\hat{p}}{1-\hat{p}}\right] dF\left(p\right)$$

Re-arranging and combining these two equations, I have

$$\frac{\hat{p}}{1-\hat{p}}\frac{u'\left(c\left(u_L^*\left(\hat{p}\right)\right)\right)}{u'\left(c\left(u_{NL}^*\left(\hat{p}\right)\right)\right)} = \frac{E\left[P|P \geq \hat{p}\right]}{1 - E\left[P|P \geq \hat{p}\right]}$$

which, by strict concavity of u, implies

$$\frac{\hat{p}}{1-\hat{p}}\frac{u'\left(w-l\right)}{u'\left(w\right)} > \frac{E\left[P|P \geq \hat{p}\right]}{1 - E\left[P|P \geq \hat{p}\right]}$$

which contradicts Condition (1).

Now, suppose that the IR constraint does not bind for any member of J. Then, clearly the IC constraint for type $\hat{p}$ must bind, otherwise profit could be increased by lowering the utility provided to members of J. So, construct the utility allocation B_ε^U to be

$$\left(u_L^\varepsilon\left(p\right), u_{NL}^\varepsilon\left(p\right)\right) = \begin{cases} \left(u_L^*\left(\hat{p}\right) - \varepsilon, u_{NL}^*\left(\hat{p}\right) + \frac{\hat{p}}{1-\hat{p}}\varepsilon\right) & if \quad p \geq \hat{p} \\ \left(u_L^*\left(p\right), u_{NL}^*\left(p\right)\right) & if \quad p < \hat{p} \end{cases}$$

so that B_ε^U consists of allocations equivalent to A^U except for $p \in J$. By construction, B_ε^U, is IR for any ε. Moreover, because of single crossing and because types are separated (finite types), B_ε^U continues to be IC and IR for $\varepsilon \in (-\eta, \eta)$ for some $\eta > 0$ sufficiently small. Therefore, I

6

must have $\frac{d\Pi(B_\varepsilon^U)}{d\varepsilon}|_{\varepsilon=0} = 0$, which implies

$$
\begin{aligned}
\frac{d\Pi\left(B_\varepsilon^U\right)}{d\varepsilon}|_{\varepsilon=0} &= \int_{p\in J}\left[pc'\left(u_L^*\left(\hat{p}\right)\right) - (1-p)c'\left(u_{NL}^*\left(\hat{p}\right)\right)\frac{\hat{p}}{1-\hat{p}}\right]dF\left(p\right) \\
&= \Pr\{p \in J\}\left[E\left[P|P \geq \hat{p}\right]\frac{1}{u'\left(c\left(u_L^*\left(\hat{p}\right)\right)\right)} - (1-E\left[P|P \geq \hat{p}\right])\frac{1}{u'\left(c\left(u_{NL}^*\left(\hat{p}\right)\right)\right)}\frac{\hat{p}}{1-\hat{p}}\right] \\
&= \Pr\{p \in J\}\frac{(1-E\left[P|P \geq \hat{p}\right])}{u'\left(c\left(u_L^*\left(\hat{p}\right)\right)\right)}\left[\frac{E\left[P|P \geq \hat{p}\right]}{(1-E\left[P|P \geq \hat{p}\right])} - \frac{u'\left(c\left(u_L^*\left(\hat{p}\right)\right)\right)}{u'\left(c\left(u_{NL}^*\left(\hat{p}\right)\right)\right)}\frac{\hat{p}}{1-\hat{p}}\right] \\
&= 0
\end{aligned}
$$

which implies
$$
\frac{\hat{p}}{1-\hat{p}}\frac{u'\left(c\left(u_L^*\left(\hat{p}\right)\right)\right)}{u'\left(c\left(u_{NL}^*\left(\hat{p}\right)\right)\right)} = \frac{E\left[P|P \geq \hat{p}\right]}{1-E\left[P|P \geq \hat{p}\right]}
$$

which, by strict concavity of u, implies
$$
\frac{\hat{p}}{1-\hat{p}}\frac{u'\left(w-l\right)}{u'\left(w\right)} > \frac{E\left[P|P \geq \hat{p}\right]}{1-E\left[P|P \geq \hat{p}\right]}
$$

which contradicts Condition (1). Therefore, if Condition (1) holds, the only possible solution to $P1$ is a full pooling allocation. □

All that remains to show is that a full pooling allocation cannot be a solution to $P1$.

Lemma A.6. *Suppose Condition* (1) *holds. Then, the only possible full-pooling solution to $P1$ is E^U.*

Proof. Suppose for contradiction that $A^U \neq E^U$ is a full-pooling solution to $P1$. Let u_L^*, u_{NL}^* denote the full pooling allocations A^U. Recall $p_1 = \min \Psi$ is the lowest risk type. Note that the IR constraint for the $p_1 = \min \Psi$ type must bind in any solution to $P1$. Otherwise, profits could be increased by providing all types with less consumption, without any consequences on the incentive constraints of types $p > p_1$. Consider the allocations C_t^U defined by

$$
\left(u_L^t, u_{NL}^t\right) = \left(u_L^* + (1-t)\left(u\left(w-l\right) - u_L^*\right), u_{NL}^* + (1-t)\left(u\left(w\right) - u_{NL}^*\right)\right)
$$

so that when $t = 1$ these allocations correspond to A^U and $t = 0$ corresponds to the endowment. Because the IR constraint of the p_1 type must hold, I know that these allocations must follow the iso-utility curve of the p_1 type which runs through the endowment. Differentiating with respect to t and evaluating at $t = 0$ yields

$$
\frac{d\Pi\left(C_t^U\right)}{dt}|_{t=0} = E\left[P|P \geq p_1\right]c'\left(u\left(w-l\right)\right) - (1-E\left[P|P \geq p_1\right])c'\left(u\left(w\right)\right)\frac{p_1}{1-p_1}
$$

where $\frac{p_1}{1-p_1}$ comes from the fact that I can parameterize the iso-utility curve of the p_1 type by

7

$u_L - \tau, u_{NL} + \frac{p_1}{1-p_1}\tau$. But re-arranging the equation, I have

$$\frac{d\Pi\left(C_t^U\right)}{dt}\Big|_{t=0} = -E\left[P|P \geq p_1\right]\frac{1}{u'\left(w-l\right)} + \left(1 - E\left[P|P \geq p_1\right]\right)\frac{1}{u'\left(w\right)}\frac{p_1}{1-p_1}$$
$$= \frac{1 - E\left[P|P \geq p_1\right]}{u'\left(W-L\right)}\left(-\frac{E\left[P|P \geq p_1\right]}{1 - E\left[P|P \geq p_1\right]} + \frac{u'\left(w-l\right)}{u'\left(w\right)}\frac{p_1}{1-p_1}\right) < 0$$

which yields a contradiction of Condition (1) at $p = p_1$. $\qquad\square$

Therefore, I have shown that if Ψ is finite, then if Condition (1) holds, the only possible allocation is the endowment. It only remains to show that this property holds when Ψ is not finite.

A.1.5   Extension to Arbitrary Type Distribution

If $F\left(p\right)$ is continuous or mixed and satisfies the no-trade condition, I first show that F can be approximated by a sequence F_n of finite support distributions on $[0, 1]$, each of which satisfy the no-trade condition.

Lemma A.7. *Let P be any random variable on $[0, 1]$ with c.d.f. $F\left(p\right)$. Then, there exists a sequence of random variables, P_N, with c.d.f. $F^N\left(p\right)$, such that $F^N \to F$ uniformly and*

$$E\left[P_N|P_N \geq p\right] \geq E\left[P|P \geq p\right] \quad \forall p, \forall N$$

Proof. Since F is increasing, it has at most a countable number of discontinuities on $[0, 1]$. Let $D = \{\delta_i\}$ denote the set of discontinuities and WLOG order these points so that $\lim_{\varepsilon \to +0} F\left(\delta_i\right) - \lim_{\varepsilon \to -0} F\left(\delta_i\right)$ is decreasing in i (so that δ_1 is the point of largest discontinuity). Then, the distribution F is continuous on $\Psi \backslash D$. For any N, let ω_N denote a partition of $[0, 1]$ given by $2^N + \min\{N, |D|\} + 1$ points equal to $\frac{j}{2^N}$ for $j = 0, ..., 2^N$ and $\{\delta_i | i \leq N\}$. I write $\omega_N = \left\{p_j^N\right\}_{j=1}^{2^N + \min\{N, |D|\} + 1}$. Now, define $\hat{F}^N : \omega_N \to [0, 1]$ by

$$\hat{F}^N\left(p\right) = F\left(\max\left\{p_j^N | p_j^N \leq p\right\}\right)$$

so that $\hat{F}^N$ converges to F uniformly as $N \to \infty$.

Unfortunately, I cannot be assured that $\hat{F}^N$ satisfies the no-trade condition at each N. But, I can perform a simple modification to $\hat{F}^N$ to arrive at a distribution that does satisfy the no-trade condition for all N and still converges to F. To do so, consider the following modification to any random variable. For any $\lambda \in [0, 1]$ and for any random variable X distributed $G\left(x\right)$ on $[0, 1]$ define the random variable X_λ to be the random variable with c.d.f. $\lambda G\left(x\right)$ and $\Pr\{X_\lambda = 1\} = 1 - \lambda$. In other words, with probability λ the variable is distributed according to X and with probability $1 - \lambda$ the variable takes on a value of 1 with certainty. Notice that $E\left[X_\lambda | X_\lambda \geq x\right]$ is continuously decreasing in λ and $E\left[X_0 | X_0 \geq x\right] = 1 \; \forall x$.

8

Now, given $\hat{F}^N$ with associated random variable $\hat{P}^N$, I define P_λ^N to be the random variable with c.d.f. $\lambda \hat{F}^N(p)$. I now define a sequence $\{\lambda_N\}_N$ by

$$\lambda_N = \max \left\{ \lambda | E\left[P_\lambda^N | P_\lambda^N \geq p \right] \geq E\left[P | P \geq p \right] \quad \forall p \right\}$$

Note that for each N fixed, the set $\left\{ \lambda | E\left[P_\lambda^N | P_\lambda^N \geq p \right] \geq E\left[P | P \geq p \right] \quad \forall p \right\}$ is a compact subset of $[0, 1]$, so that the maximum exists. Given λ_N, I define the new approximating distribution, $F^N(p)$, by

$$F^N(p) = \lambda_N F^N(p)$$

which satisfies the no-trade condition for all N. The only thing that remains to show is that $\lambda_N \to 1$ as $N \to \infty$.

By definition of λ_N, for each N there exists $\tilde{p}_N$ such that

$$E\left[P_{\lambda_N}^N | P_{\lambda_N}^N \geq \tilde{p}_N \right] = E\left[P | P \geq \tilde{p}_N \right]$$

Moreover, because λ_N is bounded, it has a convergent subsequence, $\lambda_{N_k} \to \lambda^*$. Therefore,

$$E\left[P_{\lambda^*}^{N_k} | P_{\lambda^*}^{N_k} \geq q \right] \to E\left[P_{\lambda^*} | P_{\lambda^*} \geq q \right]$$

uniformly (over q) as $k \to 0$, where P_{λ^*} is the random variable with c.d.f. $\lambda^* F(p)$. Moreover,

$$E\left[P_{\lambda_{N_k}}^{N_k} | P_{\lambda_{N_k}}^{N_k} \geq q \right] \to E\left[P_{\lambda^*} | P_{\lambda^*} \geq q \right]$$

uniformly (over q) as $k \to 0$. Therefore,

$$E\left[P_{\lambda^*}^{N_k} | P_{\lambda^*}^{N_k} \geq \tilde{p}_N \right] \to E\left[P | P \geq \tilde{p}_N \right]$$

so that I must have $\lambda^* = 1$.

Therefore, the distribution P^{N_k} with c.d.f. $F^{N_k}(p) = \lambda_{N_k} F^{N_k}(p)$ for $k \geq 1$ has the property

$$E\left[P^{N_k} | P_\lambda^{N_k} \geq p \right] \geq E\left[P | P \geq p \right] \quad \forall p$$

and $F^{N_k}(p)$ converges uniformly to $F(p)$. $\qquad\square$

Now, returning to problem $P1$ for an arbitrary distribution $F(p)$ which satisfies the no-trade condition. Let $\Pi(A|F)$ denote the value of the objective function for allocation A under distribution F. Suppose for contradiction that an allocation $\hat{A} = (\hat{u}_L(p), \hat{u}_{NL}(p)) \neq (w - l, w)$ is the solution to $P1$ under distribution F, so that $\Pi(A|F) > 0$. Let $F^N(p)$ be a sequence of finite approximating distributions which satisfy the no-trade condition and converge to F. Let $\omega_N = \left\{ p_j^N \right\}$ denote the support of each approximating distribution. For any N, define the augmented allocation $\hat{A}_N = (\hat{u}_L^N(p), \hat{u}_{NL}^N(p))$ by choosing $(\hat{u}_L(p), \hat{u}_{NL}(p))$ to be the most

9

preferred bundle from the set $\left\{ u_L \left(p_j^N \right), u_{NL} \left(p_j^N \right) \right\}_j$. Since $\hat{A}$ is incentive compatible, clearly I will have $\left(\hat{u}_L^N \left(p_j^N \right), \hat{u}_{NL}^N \left(p_j^N \right) \right) = \left(\hat{u}_L \left(p_j^N \right), \hat{u}_{NL} \left(p_j^N \right) \right)$. By single crossing, for $p \neq p_j^N$ agents with $p \in \left(p_{j-1}^N, p_j^N \right)$ will prefer either allocation for type p_{j-1}^N or p_j^N.

Clearly, $\hat{A}_N$ converges uniformly to $\hat{A}$. Since $\hat{A}_N$ satisfies IC and IR by construction, the no-trade condition implies that the allocation $\hat{A}_N$ cannot be as profitable as the endowment, so that

$$\Pi \left(\hat{A}_N | F_N \right) \leq \Pi \left(E | F_N \right) = 0 \ \ \forall N$$

By the Lebesgue dominated convergence theorem ($\Pi \left(\hat{A}_N | F_N \right)$ is also bounded below by $-(W + L)$),

$$\Pi \left(\hat{A} | F \right) \leq 0$$

Which yields a contradiction that $\hat{A}$ was the optimal solution (which required $\Pi \left(\hat{A} | F \right) > 0$) and concludes the proof.

A.2 Remark 1

A proof of Remark 1 follows in the same manner as the proof of the no trade condition. It is straightforward to see how the no trade condition holding for values $p \leq F^{-1} \left(1 - \alpha \right)$ rules out the tradability of pooling contracts that attract a fraction α of the population. To see how it also rules out separating contracts, one can repeat the analysis of Lemma A.5 noting that the measure of the sets J and K must be at least α.

B   Properties of the Lower Bound Estimator

This section formally derives the properties of the nonparametric lower bound approach presented in Section 4.1 and provides a proof of Proposition 2.

First, note that P is a mean-preserving spread of P_Z:

$$
\begin{aligned}
E\left[P|X,Z\right] &= E\left[\Pr\left\{L|X,P\right\}|X,Z\right] \\
&= E\left[\Pr\left\{L|X,Z,P\right\}|X,Z\right] \\
&= \Pr\left\{L|X,Z\right\} \\
&= P_Z
\end{aligned}
$$

where the first equality follows from Assumption 1, the second equality follows from Assumption 2, the third equality follows from the law of iterated expectations (averaging over realizations of P given X and Z), and the fourth equality is the definition of P_Z.

Let $Q_P\left(\alpha\right)$ to be the α-quantile of P,

$$
Q_P\left(\alpha\right) = \inf_q \left\{q|\Pr\left\{P \leq q\right\} \geq \alpha\right\}
$$

and $Q_\alpha\left(P_Z\right)$ to be the α-quantile of the analogue,

$$
Q_{P_Z}\left(a\right) = \inf_q \left\{q|\Pr\left\{P_Z \leq q\right\} \geq \alpha\right\}
$$

Note that $E\left[m\left(P\right)\right]$ can be represented as a weighted average of these quantiles:

$$
\begin{aligned}
E\left[m\left(P\right)\right] &= \int_0^1 \left[E_{\tilde{\alpha}}\left[Q_P\left(\tilde{\alpha}\right) - Q_P\left(\alpha\right)|\tilde{\alpha} \geq \alpha\right]\right]d\alpha \\
&= \int_0^1 \frac{1}{1-\alpha}\left[\int_{\tilde{a} \geq \alpha}\left[Q_P\left(\tilde{\alpha}\right) - Q_P\left(\alpha\right)d\tilde{\alpha}\right]\right]d\alpha \\
&= \int_0^1 \int_{\tilde{a} \geq \alpha} \frac{Q_P\left(\alpha\right)}{1-\alpha}d\tilde{\alpha}d\alpha - E\left[P\right] \\
&= \int_0^1 Q_P\left(\tilde{\alpha}\right)\int_0^{\tilde{\alpha}} \frac{1}{1-\alpha}d\alpha d\tilde{\alpha} - E\left[P\right] \\
&= \int_0^1 \left[Q_P\left(\alpha\right) - E\left[P\right]\right]\log\left(\frac{1}{1-\alpha}\right)d\alpha
\end{aligned}
$$

Now it is straightforward to prove Proposition 2.

Proof of Proposition 2   The fact that P is a mean-preserving spread of P_Z implies that

$$
\int_x^1 Q_{P_Z}\left(\alpha\right)d\alpha \leq \int_x^1 Q_P\left(\alpha\right)d\alpha \quad \forall x \in [0,1]
$$

So,

$$
\begin{aligned}
E\left[m\left(P\right)\right]-E\left[m_Z\left(P_Z\right)\right] &= \int_0^1 \left[Q_P\left(\alpha\right)-Q_{P_Z}\left(\alpha\right)\right] log\left(\frac{1}{1-\alpha}\right)d\alpha \\
&= \int_0^1 \left[Q_P\left(\alpha\right)-Q_{P_Z}\left(\alpha\right)\right]\int_0^\alpha \frac{1}{1-\tilde{\alpha}}d\tilde{\alpha}d\alpha \\
&= \int_0^1 \int_0^\alpha \left[Q_P\left(\alpha\right)-Q_{P_Z}\left(\alpha\right)\right]\frac{1}{1-\tilde{\alpha}}d\tilde{\alpha}d\alpha \\
&= \int_0^1 \int_{\tilde{\alpha}}^1 \left[Q_P\left(\alpha\right)-Q_{P_Z}\left(\alpha\right)\right]\frac{1}{1-\tilde{\alpha}}d\alpha d\tilde{\alpha} \\
&= \int_0^1 \left(\int_{\tilde{\alpha}}^1 \left[Q_P\left(\alpha\right)-Q_{P_Z}\left(\alpha\right)\right]d\alpha\right)\frac{1}{1-\tilde{\alpha}}d\tilde{\alpha} \\
&\geq 0
\end{aligned}
$$

where the last inequality follows from the fact that $\int_{\tilde{\alpha}}^1 \left[Q_P\left(\alpha\right)-Q_{P_Z}\left(\alpha\right)\right]d\alpha \geq 0$ for all $\tilde{\alpha}$ because P is a mean-preserving spread of P_Z.

C   Data Appendix

C.1   Covariate Specification

The variables used in the pricing and full controls specifications for each market are presented in Table A1. These specifications, along with the baseline age and gender specification, cover a wide range of variables that insurance companies could potentially use to price insurance and allow for an assessment of how the potential frictions imposed by private information would vary with the observable characteristics insurance companies use to price insurance.

LTC   In LTC, the pricing specification primarily follows Finkelstein and McGarry [2006] to control for variables insurers use to price insurance, along with the interaction of a rich set of health conditions to capture how insurance companies would price contracts to whom they currently reject. I include age and age squared, both interacted with gender; indicators for ADL restrictions, an indicator for performance in the lowest quartile on a word recall test, and indicators for numerous health conditions: presence of an ADL or IADL, psychological condition, diabetes, lung disease, arthritis, heart disease, cancer, stroke, and high blood pressure.[73] For the extended controls specification, I add full interactions for age and gender, along with interactions of 5 year age bins with measures of health conditions, indicators for the number of living relatives (up to 3), census region, and income deciles.

Disability   For disability, I construct the pricing specification using underwriting guidelines and also rely on feedback from interviews with a couple of disability insurance underwriters at major US insurers. In general, there are three main categories of variables used in pricing: demographics, health, and job information. The pricing specification includes age, age squared, and gender interactions; indicators for self employment, obesity (BMI > 40), the presence of a psychological condition, back condition, diabetes, lung disease, arthritis, a heart condition, cancer, stroke, and high blood pressure. I also include a linear term in BMI to capture differential pricing based on weight. Finally, I include wage deciles to capture differential pricing by wage.

The extended controls specification includes full interactions of age and gender, full interactions of wage deciles, a part time working status indicator, job tenure quartiles, and a self employment indicator. I also include interactions between 5 year age bins and the following health variables: arthritis, diabetes, lung disease, cancer, heart condition, psychological condition, back condition, and BMI quartiles. I also include full interactions between 5 year age bins and BMI quartiles. I also include full interactions of several job characteristic variables: an indicator that the job requires stooping, the job requires lifting, and the job requires physical activity. Finally, I include interactions between 5 year age bins and census region (1-5).

---

[73]Note that for the no reject sample many of these health conditions will in practice drop out of the estimation because, for example, there are no people with ADLs in the no reject sample.

In general, my conversations with underwriters suggest I have a decent approximation to the way in which insurers currently price insurance. However, as discussed in the main text, I do not observe the results of medical tests and attending physician statements, which sometimes feed into the underwriting process. Underwriters suggest the primary role of such tests is to verify application information, not for independent use in pricing; but there may be some additional factors not included in my regressions that disability insurers could use to price insurance.[74]

Life   For life, the pricing specification primarily follows He [2009] who tests for adverse selection in life insurance. The preferred specification includes age, age squared, and gender interactions, smoking status, indicators for the death of a parent before age 60, BMI, income decile, and indicators for a psychological condition, diabetes, lung disease, arthritis, heart disease, cancer, stroke, and high blood pressure. I also include a set of indicators for the years between the survey date and the AGE corresponding to the loss.[75]

The extended controls specification adds full interactions of age and gender; full interactions between age and the AGE in the subjective probability question; interactions between 5 year age bins and smoking status, income decile, census region, and various health conditions (heart condition, stroke, non-basal cell cancer, lung disease, diabetes and high blood pressure); BMI; and an indicator for death of a parent before age 60.

In general, I approximate the variables insurers use to price insurance fairly well. As with disability insurance, life insurers often require medical tests and attending physician statements from applicants; and, as with disability insurance, my conversations with underwriters suggest that the primary role of such tests is to verify application information and ensure that there is no presence of a rejection condition. But, I cannot rule out that such information could be used by insurance companies to price insurance.

Although I can well approximate the variables insurers use currently to price insurance, the data does have one key limitation in constructing the variables insurers would use to price insurance to the rejectees. A common rejection condition is the presence of cancer. If insurers were to offer insurance to people with cancer, they would potentially price differentially based on the organ in which the cancer occurs. Unfortunately, the HRS does not report the organ in which the cancer occurs in all years. Fortunately, the 2nd wave (1993/1994) of the survey does provide the organ in which a cancer occurs; therefore, to assess whether pricing differentially based on the organ of the cancer would reduce the amount of (or potentially remove all) private information, I construct a sample from 1993/1994 and include a full set of indicators for the cancer organs (54 indicators). These results are discussed in Section D.2.2 and the main conclusions of the

---

[74]Even if one believes insurers would use more information to price policies to the rejectees, it should be clear that my approach will still be able to simulate the extent to which private information would afflict an insurance market if insurers priced using the set of observables I use from the HRS. With additional data future work could explore different specifications and perhaps even make prescriptive recommendations to underwriters about relevant variables for reducing informational asymmetries.

[75]I also include this in my age & gender and extended control specifications for life.

lower bound analysis in life insurance continue to hold.

C.2 Sample Selection

For all three settings, I begin with years 1993-2008 (waves 2-9) of the HRS survey (subjective probability elicitations are not asked in wave 1).

LTC For LTC, I exclude individuals I cannot follow for a subsequent five years to construct the loss indicator variable; years 2004-2008 are used but only for construction of the loss indicator. Also, I exclude individuals who currently reside in a nursing home or receiving specialized home care. Finally, I exclude individuals with missing observations (either the subjective probabilities, or observable covariates). For consistency, I exclude any case missing any of the extended control variables (results are similar for the price controls and age/gender controls not excluding these additional missing cases).

The primary sample consists of 9,051 observations from 4,418 individuals for the no reject sample, 10,108 observations from 3,215 individuals for the reject sample, and 10,690 observations from 5,190 individuals for the uncertain sample. In each sample, I include multiple observations for a given individual (which are spaced roughly two years apart) to increase power. All standard errors are clustered at the household level.

In addition to the primary sample, I construct a sample that excludes those who own insurance to assess robustness of my results to moral hazard. For this, I drop the 13% of the sample that owns insurance, along with an additional 5% of the sample currently enrolled in Medicaid.

Disability For disability, I begin with the set of individuals between the ages of 40 and 60 who are currently working and report no presence of work-limiting disabilities. Although individuals are To construct the corresponding loss realization, I limit the sample to individuals who I can observe for a subsequent 10 years (years 2000-2008 are used solely for the construction of the loss indicator). The final sample consists of 936 observations from 491 individuals for the no reject classification, 2,216 observations from 1,280 individuals for the reject classification, and 5,361 observations from 1,280 individuals for the uncertain classification.[76] Note that the size of the no reject sample is quite small. This is primarily due to the restriction that income must be above $70,000. As discussed in Section _, the individual disability insurance market primarily exists for individuals with sufficient incomes. Thus, many of these individuals enter the uncertain classification.

Life For the life sample, I restrict to individuals I can follow through the age corresponding to the subjective probability elicitation 10-15 years in the future, so that years 2000-2008 are used

---

[76]Ideally, I would also test the robustness of my results using a sample of those who do not own disability insurance, but unfortunately the HRS does not ask about disability insurance ownership.

solely for the construction of the loss indicator. Since the earliest age used in the elicitation is 75, my sample consists of individuals aged 61 and older. The final sample consists of 2,689 observations from 1,720 individuals for the no reject classification, 2,362 observations from 1,371 individuals for the reject classification, and 6,800 observations from 4,270 individuals for the uncertain classification. Similar to LTC, I include those who own life insurance in the primary sample (64% of the sample) but present results excluding this group for robustness.

D   Lower Bound Appendix

D.1   Lower Bound Specification

Here, I discuss the construction of the lower bound estimates. I begin with a detailed discussion of the specification for the pricing controls specification and then discuss the modifications for the age/gender and extended controls specifications.

Aside from differences in the variables X, Z, and L, the specifications do not vary across the 9 samples (LTC, Life, Disability + Reject, No reject, Uncertain classifications). For the pricing controls specification, I model $\Pr\{L|X,Z\}$ as a probit,

$$\Pr\{L|X,Z\} = \Phi\left(X\beta + \Gamma\left(age, Z\right)\right)$$

where X contains all of the price control variables. The function $\Gamma\left(age, Z\right)$ captures the way in which the subjective probabilities affect the probability of a loss. In principle, one could allow this effect to vary with all observables, X; in practice, this would generate far too many interaction terms to estimate. Therefore, I allow Z to interact with age but not other variables. Note that this does not restrict how the distribution of $\Pr\{L|X,Z\}$ varies with X and Z; it only limits the number of interaction coefficients. The distribution of $\Pr[L|X,Z]$ can and does vary because of variation in Z conditional on X. Indeed, the results are quite similar if one adopts a simple specification of $\Pr\{L|X,Z\} = \Phi\left(X\beta + \gamma Z\right)$.

I choose a flexible functional form for $\Gamma\left(age, Z\right)$ that uses full interactions of basis functions in age and Z:

$$\Gamma\left(age, Z\right) = \sum_{i,j} \alpha_{ij} f_i\left(age\right) g_j\left(Z\right)$$

For the basis in Z, I use second-order Chebyshev polynomials for the normalized variables,

$\tilde{Z} = 2 \left(Z - 50\% \right)$, plus separate indicators for focal point responses at $Z = 0$, 50, and 100:

$$
\begin{aligned}
g_1 \left(Z \right) &= \tilde{Z} \\
g_2 \left(Z \right) &= \left(2\tilde{Z}^2 - 1 \right) \\
g_3 \left(Z \right) &= 1 \left\{ Z = 0\% \right\} \\
g_4 \left(Z \right) &= 1 \left\{ Z = 50\% \right\} \\
g_5 \left(Z \right) &= 1 \left\{ Z = 100\% \right\}
\end{aligned}
$$

For the basis in age, I use a linear specification, $f_1 \left(age \right) = age$ (note that any constant terms are absorbed into $X\beta$).

I estimate β and $\{\alpha_{ij}\}$ using MLE (the standard probit command in stata) and construct the predicted values for $\Pr \left\{ L | X, Z \right\}$. Given these predicted values, I plot the distribution of $\Pr \left\{ L | X, Z \right\} - \Pr \left\{ L | X \right\}$ aggregated within each setting and rejection classification. To do so, I also need an estimate of $\Pr \left\{ L | X \right\}$. For this, I use the same specification as above, except I exclude $\Gamma \left(age, Z \right)$, so that

$$
\Pr \left\{ L | X \right\} = \Phi \left(X\tilde{\beta} \right)
$$

I again estimate $\tilde{\beta}$ using MLE and construct the predicted values of $\Pr \left\{ L | X \right\}$.

Now, for each observation, I have an estimate of $\Pr \left\{ L | X, Z \right\}$ and $\Pr \left\{ L | X \right\}$. Therefore, I can plot the predicted empirical distribution of $\Pr \left\{ L | X, Z \right\} - \Pr \left\{ L | X \right\}$ in each sample. For ease of viewing, I estimate a kernel density, using the optimal bandwidth selection (the default option in Stata), and plot the density in Figure 2.

I then construct an estimate of the average magnitude of private information implied by Z. With infinite data, I could construct an estimate of $E \left[m_Z \left(P_Z \right) | X \right]$ for each value of X; in practice, I need to aggregate across values of X within a sample to gain statistical power. To do this aggregation, I rely on the assumption that the distribution of $\Pr \left\{ L | X, Z \right\} - \Pr \left\{ L | X \right\}$ does not vary conditional on age. Thus, I can aggregate across the residual distribution to construct, for each age, the average difference between ones own probability and the probability of worse risks. I construct the residual, $r_i = \Pr \left\{ L | X, Z \right\} - \Pr \left\{ L | X \right\}$ for each case in the data. Then, within each age, I compute the average residual, $\Pr \left\{ L | X, Z \right\} - \Pr \left\{ L | X \right\}$ of those with higher residuals within a given age (i.e. for an observation with $r_i = x$, I construct $\hat{r}_i = E \left[r_i | r_i \geq x, age \right]$). Note that this is where I use the assumption that the distribution of $\Pr \left\{ L | X, Z \right\} - \Pr \left\{ L | X \right\}$ does not vary conditional on age. I then construct the average of $\hat{r}_i$ in the sample, which equals $E \left[m_Z \left(P_Z \right) | X \in \Theta \right]$ for the given sample Θ.

For the age/gender controls specification, I use the same specification as for the price controls, but replace X with the saturated set of age/gender variables. However, for the extended controls specification, the number of covariates is too large for a probit specification. Aside from the computational difficulties of maximizing the probit likelihood, it is widely known that the probit

yields inconsistent estimates of Γ in this setting when the dimensionality of X increases (this is analogous to the problem of doing a probit with fixed effects). I therefore adopt a linear specification, $L = \beta X + \Gamma\left(age, Z\right) + \epsilon$, to ease estimation with the very high dimensionality of X. Under the null hypothesis that the linear model is true, this approach continues to deliver consistent estimates of Γ even as the dimensionality of X increases. For Γ, I use the same basis function approximation as used above (of course it now has a different interpretation).

D.2 Lower Bound Robustness Checks

This section presents several robustness checks of the lower bound analysis.

D.2.1 Age Analysis

First, I present estimates of the average magnitude of private information implied by Z separately by age for the disability and life settings. Figure 6 presents the results, along with bootstrapped standard errors. I also split the results separately for males and females in disability to ensure that the results are not driven by age-based sample selection in the HRS (the HRS samples near retirement individuals and includes their spouses regardless of age).

As one can see, the results suggest generally that there is more private information for the rejectees relative to non-rejectees, conditional on age.

D.2.2 Organ Controls for Life Specification

The specifications for life insurance did not include controls for the affected organ of cancer sufferers. As a result, the main results identify the impact of private information assuming that the insurer would not differentially price insurance as a function of the organ afflicted by cancer. It seems likely that insurers, if they sold insurance to cancer patients, would price differentially based on the afflicted organ. Fortunately, organ information is provided in the 1993/4 wave of the survey (it is not provided in other waves). Therefore, I can assess the robustness of my finding of private information using a sample restricted to this wave alone.

In the second column of Table A2, I report results from a specification restricted to years 1993/1994 which includes a full set of 54 indicators for the affected organ added to the extended controls specification. The finding of statistically significant amounts of private information amongst the rejectees continues to hold ($p = 0.0204$). Moreover, the estimate of $E\left[m_Z\left(P_Z\right) | X \in \Theta^{Reject}\right]$ remains similar to the preferred (pricing) specification (0.0308 versus 0.0338 for the primary specification). While insurers could potentially price differentially based on the afflicted organ, doing so would not eliminate or significantly reduce the amount of private information held by the potential applicant pool.
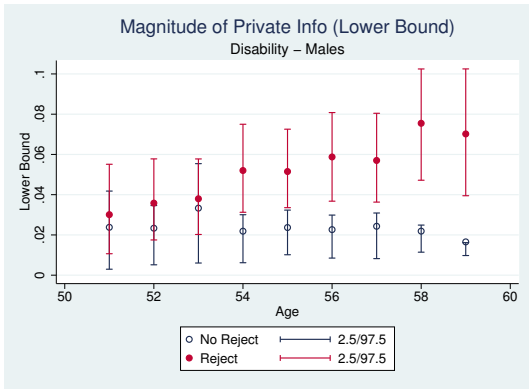
(a) Disability - Males
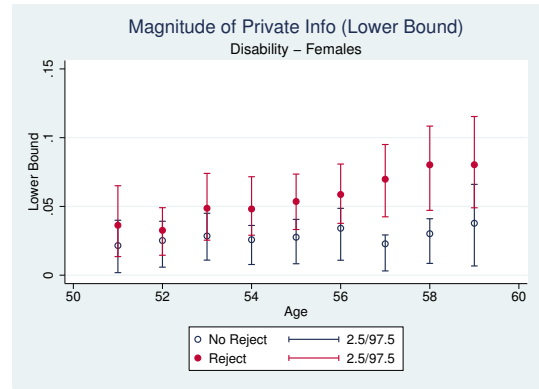
(b) Disability - Females
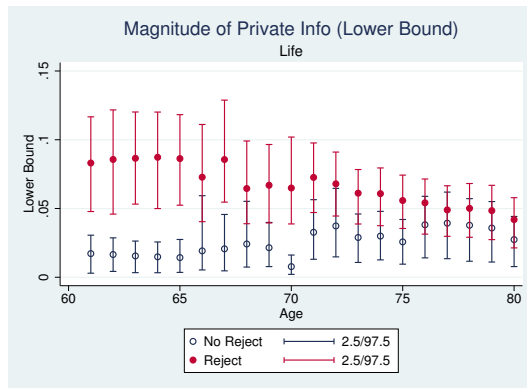
(c) Life

Figure 6: Magnitude of private information by age

E Structural Estimation

E.1 Specification Details

I approximate the distribution, $f(p|X)$, using mixtures of beta distributions,

$$f(p|X) = \sum_i w_i Beta(a_i + \Pr\{L|X\}, \psi_i)$$

where $Beta(\mu, \psi)$ is the p.d.f. of the beta distribution with mean μ and shape parameter ψ. Note this parameterization of the beta distribution is slightly non-standard; the Beta distribution is traditionally defined with parameters α and β such that the mean is $\mu = \frac{\alpha}{\alpha+\beta}$ and the shape parameter $\psi = \alpha + \beta$.

In the main specification, I use three beta distributions, $i = 1, 2, 3$. Also, I make a couple of simplifying restrictions to ease estimation. First, I only estimate two values of the shape parameter; one for the most central beta, $\psi_1 = \psi_{central}$, and one for all other beta distributions, $\psi_i = \psi_{noncentral}$ ($i = 2, 3$). This helps reduce the non-convexity of the likelihood function.[77] Second, I constrain the shape parameters, ψ_i, such that $\psi_i \leq 200$. This restriction prevents ψ_i from reaching large values that introduce non-trivial approximation errors in the numerical integration of the likelihood over values of p (these numerical errors arise when $f_P(p|X)$ exhibits extreme curvature). Changing the levels of this constraint does not affect the results in the LTC Reject, Disability Reject, Life No Reject, and Life Reject samples. However, the LTC No Reject and Disability No Reject initial estimates did lie on the boundary, $\psi_i = 200$ for the most central beta. Intuitively, these samples have little amounts of private information so that the model attempts to construct a very highly concentrated distribution, $f_P(p|X)$. To relax this constraint, I therefore include an additional point mass at the mean, $\Pr\{L|X\}$, that helps capture the mass of people with no private information (note that inserting a point mass at the mean is equivalent to inserting a beta distribution with $a_i = 0$ and $\psi_i = \infty$). This computational shortcut improves the estimation time and helps remove the bias induced by the restriction $\psi_i \leq 200$.

In addition to these constraints, Assumptions 1 and 2 also yield the constraint $\Pr\{L|X\} = E[P|X]$, which requires $\sum_i w_i a_i = 0$. Imposing this constraint further reduces the number of estimated parameters. I also censor the mean of each beta distribution, $a_i + \Pr\{L|X\}$ to lie in $[0.001, 0.999]$. I accomplish this by censoring the value of a_i given to observations with values of X such that $a_i + \Pr\{L|X\}$ is greater than 0.999 or less than 0.001. I then re-adjust the other values of a_i and w_i for this observation, to ensure the constraint $\sum_i w_i a_i = 0$ continues to hold. If the parameter values and values of X are such that $a_2 + \Pr\{L|X\} < 0.001$, I then define $a_2 = 0.001 - \Pr\{L|X\}$ and then adjust a_3 such that $a_2 w_2 + a_3 w_3 = 0$. In some instances, it may be the case that $a_2 = 0.001$ and $a_3 = 0.999$; in this case, I adjust the weights w_2 and w_3 to

---

[77]For example, non-convexity arises because a dispersed distribution can be accomplished either with one beta distribution with a high value of ψ or with two beta distributions with lower values of the shape parameters but differing values of a_i.

ensure that $\sum w_i a_i = 0$ (note that weight w_1 is unaffected because $a_1 = 0$).

Given this specification with three beta distributions and the above mentioned restrictions, there are six parameters to estimate: two parameters capture the relative weights on the three betas, two parameters capture the non-centrality of the beta distributions (a_1 and a_2), and the two shape parameters, $\psi_{central}$, and $\psi_{noncentral}$. Finally, for the LTC No Reject and Disability No Reject samples I estimate a seventh parameter given by the weight on the point-mass, w_{ptmass}.

Estimation   In each of the six samples, estimation is done in two steps.[78]   First, I estimate $\Pr\{L|X\}$ using the probit specification described in Section E.1. Second, I estimate the six beta mixture parameters, $\{w_1, w_2, a_1, a_2, \psi_{central}, \psi_{noncentral}\}$, along with the four elicitation error parameters, $\{\sigma, \kappa, \lambda, \alpha\}$ using maximum likelihood. As is standard with mixture estimation, the likelihood is non-convex and can have local minima. I therefore start the maximization algorithm from 100+ random starting points in the range of feasible parameter values.

In addition, I impose a lower bound on σ in the estimation process. It is straightforward to verify that, under the null hypothesis, I have $\sigma \geq min\left\{var\left(Z^{nf}\right) - cov\left(Z^{nf}, L\right), \sqrt{\frac{3}{8}}\right\}$. In reality, the distribution of Z is concentrated on integer values between 0 and 100%, and in particular, multiples of 5% and 10%. In some specifications, the unconstrained maximum likelihood procedure would yield estimates of $\sigma \approx 0$ and distributions of P that attempt to match the integer patterns of Z. In other words, the model attempts to match the dearth of Z values between 5.01% and 9.99%, and the higher frequency at $Z = 10\%$. By imposing the constraint $\sigma \geq min\left\{var\left(Z^{nf}\right) - cov\left(Z^{nf}, L\right), \sqrt{\frac{3}{8}}\right\}$, these pathological outcomes are removed. Re-assuringly, the constraint does not locally bind in any of my samples (i.e. I find estimates of σ between 0.3 and 0.45, whereas values of $var\left(Z^{nf}\right) - cov\left(Z^{nf}, L\right)$ fall consistently around 0.2 in each setting).

E.2   Robustness

Table A3 presents the minimum pooled price ratio evaluated at other points along the distribution of $\Pr\{L|X\}$ in each sample. The table presents the estimates at the 20th, 50th, and 80th quantile of the $\Pr\{L|X\}$ distribution. The first set of rows presents the results for the Reject samples. The first row presents the point estimates, followed by the 5/95% confidence intervals, and finally by the value of $\Pr\{L|X\}$ corresponding to the given quantile. The second set of rows repeats these figures for the non-rejectees. In general, the results are quite similar to the values reported in Tables 5 and 6, which considered a characteristic corresponding to the mean loss, $\Pr\{L|X\} = \Pr\{L\}$.

---

[78]The bootstrapping procedure for standard errors will repeat the entire estimation process (i.e. both steps 1 and 2) for each bootstrap iteration.

E.3 Estimation Results Details

Measurement error parameters Table A4 presents the estimated measurement error parameters. In general, I estimate values of σ between 0.29 and 0.46, indicating that elicitations are quite noisy measures of true beliefs. Roughly 30-42% of respondents are focal point respondents, and the focal point window estimate ranges from 0 to 0.173. The estimate of $\kappa = 0$ indicates that focal point respondents choose to report an elicitation of 50%, regardless of their true beliefs. Finally, I estimate moderate bias of magnitudes less than 10% in all samples except the LTC Rejectees, for whom I estimate a substantial 28.6pp downward bias. Although many factors could be driving this result, it is consistent with the hypothesis that many individuals do not want to admit to a surveyor that they are going to have to go to a nursing home.

Beta Mixture Parameters Table A5 presents the estimated parameters for $f_P\left(p|X\right)$, along with the bootstrapped standard errors.

F Selected Pages from Genworth Financial Underwriting Guidelines

The following 4 pages contain a selection from Genworth Financial's LTC underwriting guideline which is provided to insurance agents for use in screening applicants. Although marked "Not for use with consumers or to be distributed to the public", these guidelines are commonly left in the public domain on the websites of insurance brokers. The printed version here was found in public circulation at http://www.nyltcb.com/brokers/pdfs/Genworth_Underwriting_Guide.pdf on November 4, 2011. I present 4 pages of the 152 pages of the guidelines. The conditions documented below are not exhaustive for the list of conditions which lead to rejection - they constitute the set of conditions which solely lead to rejection (independent of other health conditions); combinations of other conditions may also lead to rejections and the details for these are provided in the remaining pages not shown here.

Genworth® Financial

LONG TERM CARE INSURANCE UNDERWRITING GUIDE

PROVIDED BY THE GENWORTH UNDERWRITING DEPARTMENT

Long Term Care Insurance Underwritten
by Genworth Life Insurance Company,
and in New York
by Genworth Life Insurance Company of New York
Administrative Offices: Richmond, VA.

INTRODUCTION

Underwriting is the process by which an applicant's current health, medical history and lifestyle are evaluated to determine a risk profile. The underwriter's decision to accept or decline an applicant is determined by matching the profile to guidelines, which outline the limits of acceptable risk to the company.

We underwrite applicants in the age range 18-79. We do not modify the coverage applied for, nor do we apply extra premiums. We make every attempt to issue the desired coverage at the corresponding published premium.

The information in this manual reflects over 30 years of experience...the longest in the Long Term Care insurance industry. While not all-inclusive, enough information is presented to help you in most situations you will encounter. A hotline number is included should you have questions or run into an unusual circumstance.

An appeal process is also outlined in the event you disagree with our underwriting evaluation. We are always willing to have a second look, especially when additional information not included in the original application file is made available.

We value our relationship with you and look forward to providing high quality service and underwriting for you and your clients.

UNINSURABLE CONDITIONS

Acquired Immune Deficiency Syndrome (AIDS)
ADL limitation, present
AIDS Related Complex (ARC)
Alzheimer's Disease
Amputation due to disease, e.g., diabetes or atherosclerosis
Amyotrophic Lateral Sclerosis (ALS) , Lou Gehrig's Disease
Ascites present
Ataxia, Cerebellar
Autonomic Insufficiency (Shy-Drager Syndrome)
Autonomic Neuropathy (excluding impotence)
Behçet's Disease
Binswanger's Disease
Bladder incontinence requiring assistance
Blindness due to disease or with ADL/IADL limitations
Bowel incontinence requiring assistance
Buerger's Disease (thromboangiitis obliterans)
Cerebral Vascular Accident (CVA)
Chorea
Chronic Memory Loss
Cognitive Testing, failed
Cystic Fibrosis
Dementia
Diabetes treated with insulin
Dialysis, Kidney (Renal)
Ehlers-Danlos Syndrome
Forgetfulness (frequent or persistent)
Gangrene due to diabetes or peripheral vascular disease
Hemiplegia
Hoyer Lift
Huntington's or other forms of Chorea
Immune Deficiency Syndrome
Korsakoff's Psychosis
Leukemia-except for Chronic Lymphocytic Leukemia (CLL) and Hairy Cell Leukemia (HCL)
Marfan's Syndrome
Medications
       Antabuse (disulfiram)
       Aricept (donepezil HCl)
       Campral (acamprosate calcium)
       Cognex (tacrine)
       Depade (naltrexone)
       Exelon (rivastigmine)
       Hydergine (ergoloid mesylate)
       Namenda (memantine)
       Razadyne (galantamine hydrobromide)
       Reminyl (galantamine hydrobromide)
       ReVia (naltrexone)
       Vivitrol (naltrexone)
Memory Loss, chronic
Mesothelioma
Multiple Sclerosis (MS)

Muscular Dystrophy (MD)
Myelofibrosis
Organ Transplants, except kidney transplants
Organic Brain Syndrome (OBS)
Oxygen use except if used for headaches or sleep apnea
Paralysis/Paraplegia
Parkinson's Disease
Pneumocystis Pneumonia
Polyarteritis Nodosa
Postero-Lateral Sclerosis
Quad Cane use
Quadriplegia
Senility
Spinal Cord Injury with ADL/IADL limitations
Stroke (CVA)
Surgery scheduled or anticipated (except cataract surgery under local anesthesia)
Takayasu's Arteritis
Thalassemia Major
Total Parenteral Nutrition (TPN) for regular or supplementary feeding or
     administration of medication
Waldenstrom's Macroglobulinemia
Walker use
Wegener's Granulomatosis
Wernicke-Korsakoff Syndrome
Wheelchair use
Wilson's Disease

Table A1: Covariate Specifications

Long-Term Care		Disability		Life	
Price Controls	Extended Controls	Price Controls	Extended Controls	Price Controls	Extended Controls
Age, Age^2, Gender Gender*age Gender*age^2	Full interactions of Age Gender	Age, Age^2, Gender Gender*age Gender*age^2	Full interactions of Age Gender	Age, Age^2, Gender Gender*age Gender*age^2 Smoker Status	Full interactions of Age Gender
Word Recall Performance[1]	Word Recall Performance[1]	Indicators for Self Employed Obese Psych condition Back condition Diabetes Lung Disease Arthritis Heart Condition Cancer Stroke High Blood Pressure	Full interactions of wage decile part time indicator job tenure quartile self-employment indicator	Indicator for years to question[2]	Full Interactions of age AGE in subj prob question
Indicators for ADL/IADL Restriction Psych Condition Diabetes Lung Disease Arthritis Heart Disease Cancer Stroke High blood pressure	Indicators for ADL/IADL Restriction Psychological Condition Diabetes Lung Disease Arthritis Heart Disease Cancer Stroke High blood pressure		Interactions between 5 yr age bins and the presence of: Arthritis Diabetes Lung disease Cancer Heart condition Psychological condition Back condition BMI Quartile	Indicator for death of parent before age 60	Interactions of 5 yr age bins with: Smoker Status Income Decile Heart condition Stroke Cancer Lung disease Diabetes High blood pressure Census Region
	Interactions between 5 yr age bins and the presence of:	BMI		BMI	BMI
	Number of Health Conditions (High bp, diabetes, heart condition, lung disease, arthritis, stroke, obesity, psych condition) Number of ADL / IADL Restrictions Number of living relatives (<=3) Past home care usage Census region (1-5) Income Decile	Wage Decile	Full interactions of BMI quartile 5 year age bins	Indicators for Psychological Condition Diabetes Lung Disease Arthritis Heart Disease Cancer Stroke High blood pressure	Indicator for death of parent before age 60
			Full interactions of Job requires stooping Job requires lifting Job requires phys activity	Income decile	
			Full Interactions of 5 year age bins Census region (1-5)		

[1]Indicator for lowest quartile performance on word recall test
[2]Full indicator variables for number of years to AGE reported in subjective probability question

Table A2: Cancer Organ Controls (Life Setting)

	Preferred Specification	Organ + Extended Controls (1993/1994 Only)
Reject	0.0587***	0.0526***
s.e.[1]	(0.0083)	(0.0098)
p-value[2]	0.000	0.002
No Reject	0.0249	0.0218
s.e.[1]	(0.007)	(0.007)
p-value[2]	0.1187	0.3592
Difference: Δ_z	0.0338***	0.0308**
s.e.[1]	(0.0107)	(0.0121)
p-value[3]	0.0000	0.0260
Uncertain	0.0294***	0.0342***
s.e.[1]	(0.0054)	(0.0063)
p-value[2]	0.0001	0.0003

[1]Bootstrapped standard errors computed using block re-sampling at the household level (results shown for N=1000 repetitions)

[2]p-value for the Wald test which restricts coefficients on subjective probabilities equal to zero

[3]p-value is the maximum of the p-value for the rejection group having no private information (Wald test) and the p-value for the hypothesis that the difference is less than or equal to zero, where the latter is computed using bootstrap

*** p<0.01, ** p<0.05, * p<0.10

Table A3: Minimum Pooled Price Ratio: Robustness to Alternative Pr{L|X} Locations

Quantile of Index, Pr{L|X}	LTC				Disability				Life			
	Mean	20%	50%	80%	Mean	20%	50%	80%	Mean	20%	50%	80%
Reject	**1.827**	**2.090**	**1.849**	**1.776**	**1.661**	**1.687**	**1.659**	**1.741**	**1.428**	**1.416**	**1.436**	**1.609**
5%[1]	1.657	1.901	1.684	1.562	1.524	1.550	1.522	1.550	1.076	0.987	0.987	0.987
95%	2.047	2.280	2.280	2.280	1.824	1.825	1.825	1.879	1.780	1.846	1.846	2.054
Pr{L|reject}	0.225	0.124	0.207	0.314	0.441	0.293	0.430	0.578	0.572	0.351	0.589	0.791
No Reject	**1.163**	**1.168**	**1.160**	**1.171**	**1.069**	**1.064**	**1.068**	**1.072**	**1.350**	**1.621**	**1.390**	**1.336**
5%[1]	1.000	1.000	1.000	1.000	0.932	0.926	0.926	0.926	1.000	1.000	1.000	1.000
95%	1.361	1.665	1.665	1.665	1.840	1.967	1.967	1.967	1.702	2.050	2.050	2.050
Pr{L|No Reject}	0.052	0.021	0.041	0.076	0.115	0.069	0.105	0.147	0.273	0.073	0.194	0.458
Difference (Reject - No Reje	**0.664**	**0.922**	**0.689**	**0.605**	**0.592**	**0.623**	**0.591**	**0.669**	**0.077**	**-0.204**	**0.045**	**0.272**
5%[2]	0.428	0.583	0.444	0.341	0.177	0.069	0.069	0.069	-0.329	-5.050	-5.050	-5.050
95%	0.901	1.261	1.261	1.261	1.008	1.178	1.178	1.178	0.535	4.641	4.641	4.641

[1]5/95% CI computed using bootstrap block re-sampling at the household level (N=250 Reps); 5% level extended to include 1 if p-value of F-test for presence of private information is less than .05; Bootstrap CI is bias corrected using the non-accelerated procedure in Efron (1982).

[2]5/95% CI computed using bootstrap block re-sampling at the household level (N=1000 Reps); 5% level extended to include 1.00 if p-value of F-test for presence of private information for the rejectees is less than .05; Bootstrap CI is the union of the percentile-t bootstrap and bias corrected (non-accelerated) percentile invervals from Efron and Gong (1983).

Table A4: Elicitation Error Parameters

	LTC		Disability		Life	
	No Reject	Reject	No Reject	Reject	No Reject	Reject
Standard Deviation (σ)	0.293	0.443	0.298	0.311	0.422	0.462
s.e.	(0.015)	(0.009)	(0.025)	(0.016)	(0.014)	(0.013)
Fraction Focal Respondents (Λ)	0.364	0.348	0.292	0.417	0.375	0.383
s.e.	(0.046)	(0.01)	(0.032)	(0.018)	(0.014)	(0.013)
Focal Window (κ)	0.173	0.001	0.000	0.000	0.001	0.000
s.e.	(0.058)	(0.015)	(0.073)	(0.053)	(0.014)	(0.003)
Bias (α)	-0.078	-0.286	0.086	-0.099	0.034	0.014
s.e.	(0.025)	(0.01)	(0.041)	(0.017)	(0.014)	(0.016)

Note: Bootstrapped standard errors computed using block re-sampling at the household level (results shown for N=1,000 repetitions)

Table A5: Beta Mixture Parameters

	LTC		Disability		Life	
	No Reject	Reject	No Reject	Reject	No Reject	Reject
Weight on Beta 1	0.005	0.848	0.001	0.890	0.000	0.714
s.e.	(0.066)	(0.247)	(0.134)	(0.421)	(0.377)	(0.388)
Weight on Beta 2	0.142	0.065	0.059	0.007	0.094	0.285
s.e.	(0.057)	(0.246)	(0.093)	(0.182)	(0.341)	(0.281)
Non-centrality of Beta 1	0.500	0.065	0.059	0.007	0.094	0.285
s.e.	(0.057)	(0.246)	(0.093)	(0.182)	(0.341)	(0.281)
Non-centrality of Beta 2	-0.527	0.021	-0.054	0.030	0.410	0.000
s.e.	(0.036)	(0.058)	(0.116)	(0.111)	(0.062)	(0.154)
Shape parameter for Beta 1	11.185	31.488	190.752	11.261	16.005	10.941
s.e.	(38.499)	(81.148)	(57.545)	(14.604)	(19.697)	(14.208)
Shape parameter for Beta 2 and Beta 3	27.940	36.318	66.992	18.272	17.977	10.714
s.e.	(36.391)	(46.674)	(46.846)	(11.719)	(18.026)	(27.965)
Weight on point mass at mean	0.817		0.833			
s.e.	(0.116)		(0.263)			

Note: Bootstrapped standard errors computed using block re-sampling at the household level (results shown for N=1,000 repetitions).