SUPPLEMENT TO "PRIVATE INFORMATION AND INSURANCE REJECTIONS"

BY NATHANIEL HENDREN

THIS SUPPLEMENTAL MATERIAL CONTAINS FIVE APPENDIXES. 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, presents the decomposition discussed in Section 6.3, and presents additional robustness checks referred to in Section 6.5. Appendix E provides details on the specification and estimation of the structural approach referred to throughout Section 7. Finally, Appendix F presents a selection of pages from the LTC underwriting guidelines of Genworth Financial.

APPENDIX A: THEORY

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 that satisfy 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 that 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, continu-

ously 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

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

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 f constrained maximization program*

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

PROOF: Note that the constraint set is linear and the objective function is strictly concave. The first constraint is the incentive constraint (IC) in utility space. The second constraint is the individual rationality (IR) 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 IC and IR constraints. Since the constraints are linear, the allocations $C^U(t) = tB^U + (1 - t)E^U$ lie in the con-

straint 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. *Q.E.D.*

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 $A \neq E$ that 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* $\hat{p} \in \Psi \setminus \{1\}$ *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

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

PROOF: The proof follows by constructing an allocation that 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 that contains $\hat{p}$ 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 $E[P|P \geq p]$ is locally continuous in p at $\hat{p}$ so that, without loss of generality (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

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

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

$$u_L(\varepsilon, \eta) = u(w-l) + \varepsilon + \eta,$$

$$u_{NL}(\varepsilon, \eta) = u(w) - \frac{1-\hat{p}}{\hat{p}} \varepsilon.$$

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

$$
\begin{aligned}
&\big(\hat{u}_{\mathrm{L}}(p), \hat{u}_{\mathrm{NL}}(p)\big) \\
&\quad = \begin{cases} \left(u(w-l) + \varepsilon + \eta,\, u(w) - \dfrac{\hat{p}}{1-\hat{p}}\varepsilon \right) & \text{if } p \geq \hat{p}, \\[2mm] \big(u(w-l),\, u(w) \big) & \text{if } p < \hat{p}. \end{cases}
\end{aligned}
$$

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

$$
\Pi(\varepsilon, \eta) = \int \big[w - pl - pc\big(\hat{u}_{\mathrm{L}}(p)\big) - (1-p)c\big(\hat{u}_{\mathrm{NL}}(p)\big) \big] \, dF(p).
$$

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

$$
\begin{aligned}
\frac{\partial \Pi}{\partial \varepsilon}\bigg|_{\varepsilon=0} = \int \bigg[&-pc'\big(u(w-l+\eta)\big) + \frac{\hat{p}}{1-\hat{p}}(1-p)c'\big(u(w)\big) \bigg] \\
&\times \mathbf{1}\{p \geq \hat{p}\} \, dF(p),
\end{aligned}
$$

which is strictly positive if and only if

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

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

$$
\frac{\partial \Pi}{\partial \varepsilon}\bigg|_{\varepsilon=0,\eta=0} > 0 \quad \Longleftrightarrow \quad \frac{\hat{p}}{1-\hat{p}} \frac{u'(w-l)}{u'(w)} > \frac{E[P|P \geq \hat{p}]}{1 - E[P|P \geq \hat{p}]}.
$$

Thus, by continuity, the above condition holds for sufficiently small $\eta > 0$, proving the existence of an allocation that 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.   Q.E.D.

A.1.3. *Useful Results*

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

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

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

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

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

PROOF: Since $u'(c)$ 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 $u(c)$. Q.E.D.

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

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

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

$$\tilde{c}_{NL}(p) = \begin{cases} c_{NL}(p) & \text{if } p > \hat{p}, \\ w & \text{if } p \leq \hat{p}. \end{cases}$$

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

Now, suppose $c_{NL}(\hat{p}) > w$. Then the incentive compatibility constraint requires $c_{NL}(p) > w \; \forall p \leq \hat{p}$. Construct $\tilde{A}$ as above, yielding the same contradiction. Q.E.D.

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

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

PROOF: Let $A^U = \{u_L^*(p), u_{NL}^*(p)\}_p$ denote the solution to P and suppose, for contradiction, that the solution to P is not a full-pooling allocation. Let $\hat{p} = \min\{p | u_L^*(p) = u_L^*(1)\}$ and let $\hat{p}_- = \max\{p | u_L^*(p) \neq u_L^*(1)\}$. The assumption that Ψ is finite implies that $\hat{p} > \hat{p}_-$. Let us define the pooling sets $J = \{p | u_L^*(p) = u_L^*(1)\}$ and $K = \{p | u_L^*(p) = u_L^*(\hat{p}_-)\}$. I will show that a profitable deviation exists that 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 $\hat{p} = 1$ 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 its 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 $\tilde{p} \in \Psi \setminus J$ receive their endowment, $(c_L, c_{NL}) = (w - l, w)$, or the IR constraint cannot bind for any member of J. I consider these two cases in turn.

Suppose $u_L^*(p) = u(w - l)$ and $u_{NL}^*(p) = u(w)$ for all types $\tilde{p} \in \Psi \setminus J$. 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 $\tilde{p} \in \Psi \setminus J$. I now show that the profitability of the allocation violates the no-trade condition. The profitability of A^U is

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

Now, I construct the utility allocation A_t^U by

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

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 $t \in [0, \hat{t} + \eta]$ 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

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

where

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

Rearranging and combining these two equations, I have

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

which, by strict concavity of u, implies

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

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

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

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

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

$$\times \left[\frac{E[P|P \geq \hat{p}]}{(1 - E[P|P \geq \hat{p}])} - \frac{u'(c(u_{\mathrm{L}}^*(\hat{p})))}{u'(c(u_{\mathrm{NL}}^*(\hat{p})))} \frac{\hat{p}}{1 - \hat{p}} \right]$$
$$= 0,$$

which implies

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

which, by strict concavity of u, implies

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

which contradicts condition (1). Therefore, if condition (1) holds, the only possible solution to P1 is a full-pooling allocation.   *Q.E.D.*

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}^* and u_{NL}^* denote the full-pooling allocations A^U. Recall that $p_1 = \min \Psi$ 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

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

$$u_{\mathrm{NL}}^* + (1 - t)\big(u(w) - u_{\mathrm{NL}}^*\big) \Big)$$

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 that runs through the endowment. Differentiating with respect to t and evaluating at $t = 0$ yields

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

$$- \big(1 - E[P|P \geq p_1]\big)c'\big(u(w)\big)\frac{p_1}{1 - p_1},$$

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

$$\frac{d\Pi(C_t^U)}{dt}\bigg|_{t=0}$$

$$= -E[P|P \ge p_1]\frac{1}{u'(w-l)} + \left(1 - E[P|P \ge p_1]\right)\frac{1}{u'(w)}\frac{p_1}{1-p_1}$$

$$= \frac{1 - E[P|P \ge p_1]}{u'(W-L)}\left(-\frac{E[P|P \ge p_1]}{1 - E[P|P \ge p_1]} + \frac{u'(w-l)}{u'(w)}\frac{p_1}{1-p_1}\right) < 0,$$

which yields a contradiction to condition (1) at $p = p_1$.   Q.E.D.

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 $F(p)$ 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 satisfies the no-trade condition.

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

$$E[P_N|P_N \ge p] \ge E[P|P \ge p] \quad \forall p, \forall N.$$

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 $\lim_{\varepsilon \to +0} F(\delta_i) - \lim_{\varepsilon \to -0} F(\delta_i)$ is decreasing in i (so that δ_1 is the point of largest discontinuity). Then the distribution F is continuous on $\Psi \setminus D$. 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 $j = 0, \ldots, 2^N$ and $\{\delta_i | i \le N\}$. I write $\omega_N = \{p_j^N\}_{j=1}^{2^N + \min\{N, |D|\}+1}$. Now define $\hat{F}^N : \omega_N \to [0, 1]$ by

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

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 $G(x)$ on $[0, 1]$, define the random variable X_λ to be the random variable with c.d.f. $\lambda G(x)$ 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 $E[X_\lambda | X_\lambda \geq x]$ is continuously decreasing in λ and $E[X_0 | X_0 \geq x] = 1 \ \forall x$.

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

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

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

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

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

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

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

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

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

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

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

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

and $F^{N_k}(p)$ converges uniformly to $F(p)$.   Q.E.D.

Now, let us return to problem P1 for an arbitrary distribution $F(p)$ that 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 that satisfy the no-trade condition and converge to F. Let $\omega_N = \{p_j^N\}$ 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 preferred bundle from the set $\{u_L(p_j^N), u_{NL}(p_j^N)\}_j$. Since $\hat{A}$ is incentive compatible, clearly I will have $(\hat{u}_L^N(p_j^N), \hat{u}_{NL}^N(p_j^N)) = (\hat{u}_L(p_j^N), \hat{u}_{NL}(p_j^N))$. By single crossing, for $p \neq p_j^N$, agents with $p \in (p_{j-1}^N, p_j^N)$ 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

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

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

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

which yields a contradiction that $\hat{A}$ was the optimal solution (which required $\Pi(\hat{A}|F) > 0$) 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 fact that the no-trade condition holds for values $p \leq F^{-1}(1 - \alpha)$ 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.

APPENDIX B: PROPERTIES OF THE LOWER BOUND ESTIMATOR

This appendix 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,

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

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

$$
Q_P(\alpha) = \inf_q\big\{q|\Pr\{P \le q\} \ge \alpha\big\},
$$

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

$$
Q_{P_Z}(a) = \inf_q\big\{q|\Pr\{P_Z \le q\} \ge \alpha\big\}.
$$

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

$$
\begin{aligned}
E\big[m(P)\big] &= \int_0^1 \Big[E_{\tilde{\alpha}}\big[Q_P(\tilde{\alpha}) - Q_P(\alpha)|\tilde{\alpha} \ge \alpha\big]\Big] d\alpha \\
&= \int_0^1 \frac{1}{1-\alpha}\bigg[\int_{\tilde{a} \ge \alpha}\big[Q_P(\tilde{\alpha}) - Q_P(\alpha)\, d\tilde{\alpha}\big]\bigg] d\alpha \\
&= \int_0^1 \int_{\tilde{a} \ge \alpha} \frac{Q_P(\alpha)}{1-\alpha}\, d\tilde{\alpha}\, d\alpha - E[P] \\
&= \int_0^1 Q_P(\tilde{\alpha}) \int_0^{\tilde{\alpha}} \frac{1}{1-\alpha}\, d\alpha\, d\tilde{\alpha} - E[P] \\
&= \int_0^1 \big[Q_P(\alpha) - E[P]\big]\log\Big(\frac{1}{1-\alpha}\Big)\, d\alpha.
\end{aligned}
$$

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

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

So

$$E\big[m(P)\big] - E\big[m_Z(P_Z)\big] = \int_0^1 \big[Q_P(\alpha) - Q_{P_Z}(\alpha)\big] \log\left(\frac{1}{1-\alpha}\right) d\alpha$$

$$= \int_0^1 \big[Q_P(\alpha) - Q_{P_Z}(\alpha)\big],$$

$$\int_0^\alpha \frac{1}{1-\tilde{\alpha}}\, d\tilde{\alpha}\, d\alpha = \int_0^1 \int_0^\alpha \big[Q_P(\alpha) - Q_{P_Z}(\alpha)\big] \frac{1}{1-\tilde{\alpha}}\, d\tilde{\alpha}\, d\alpha$$

$$= \int_0^1 \int_{\tilde{\alpha}}^1 \big[Q_P(\alpha) - Q_{P_Z}(\alpha)\big] \frac{1}{1-\tilde{\alpha}}\, d\alpha\, d\tilde{\alpha}$$

$$= \int_0^1 \left(\int_{\tilde{\alpha}}^1 \big[Q_P(\alpha) - Q_{P_Z}(\alpha)\big] d\alpha \right) \frac{1}{1-\tilde{\alpha}}\, d\tilde{\alpha}$$

$$\geq 0,$$

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

APPENDIX C: DATA

C.1. *Covariate Specification*

The variables used in the pricing and full controls specifications for each market are presented in Table A-I. 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 those individuals 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, that is, the presence of an ADL or instrumental ADL (IADL), psychological condition, diabetes, lung disease, arthritis, heart disease, cancer, stroke, and high blood pressure.[1] For the extended controls specification, I add full interactions for age and gender, along

---

[1]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.

TABLE A-I

COVARIATE SPECIFICATIONS

Price Controls	Extended Controls

Long-Term Care

Price Controls	Extended Controls
Age, ageˆ2, gender	Full interactions of
Gender ∗ age	Age
Gender ∗ ageˆ2	Gender
Word recall performance[a]	Word recall performance[a]
Indicators for	Indicators for
ADL/IADL restriction	ADL/IADL restriction
Psychological condition	Psychological condition
Diabetes	Diabetes
Lung disease	Lung disease
Arthritis	Arthritis
Heart disease	Heart disease
Cancer	Cancer
Stroke	Stroke
High blood pressure	High blood pressure
	Interactions between 5 yr age bins and the presence of
	Number of health conditions (high blood preassure, diabetes, heart condition, lung disease, arthritis, stroke, obesity, psychological condition)
	Number of ADL/IADL restrictions
	Number of living relatives (≤ 3)
	Past home care usage
	Census region (1–5)
	Income decile

Disability

Price Controls	Extended Controls
Age, ageˆ2, gender	Full interactions of
Gender ∗ age	Age
Gender ∗ ageˆ2	Gender
Indicators for	Full interactions of
Self employed	Wage decile
Obese	Part time indicator
Psychological condition	Job tenure quartile
Back condition	Self-employment indicator
Diabetes	
Lung disease	Interactions between 5 yr age bins and the presence of
Arthritis	
Heart condition	Arthritis
Cancer	Diabettes
Stroke	Lung disease
High blood pressure	Cancer

TABLE A-I—*Continued*

Price Controls	Extended Controls
	Heart condition
BMI	Psychological condition
	Back condition
Wage decile	BMI quartile
	Full interactions of
	BMI quartile
	5 year age bins
	Full interactions of
	Job requires stooping
	Job requires lifting
	Job requires phys activity
	Full interactions of
	5 year age bins
	Census region (1–5)

<div align="center">Life</div>

Price Controls	Extended Controls
Age, age^2, gender	Full interactions of
Gender ∗ age	Age
Gender ∗ age^2	Gender
Smoker status	
	Full interactions of
Indicator for years to question[b]	Age
	AGE in subj prob question
Indicator for death of parent	
before age 60	Interactions of 5 yr age bins with
	Smoker status
BMI	Income decile
	Heart condition
Indicators for	Stroke
Psychological condition	Cancer
Diabetes	Lung disease
Lung disease	Diabetes
Arthritis	High blood pressure
Heart disease	Census region
Cancer	
Stroke	BMI
High blood pressure	
	Indicator for death of parent
Income decile	before age 60

[a]Indicator for lowest quartile performance on word recall test.

[b]Full indicator variables for number of years to AGE reported in subjective probability question.

with interactions of 5-year age bins with measures of health conditions, indicators for the number of living relatives (up to three), 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 U.S. 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 (body mass index (BMI) > 40), the presence of a psychological condition, back condition, diabetes, lung disease, arthritis, 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).

In general, my conversations with underwriters suggest that 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 that 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.[2]

Life

For life, the pricing specification primarily follows He (2009), who tested for adverse selection in life insurance. The preferred specification includes age,

---

[2]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.

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.[3]

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 do 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 is present. Unfortunately, the HRS does not report the organ in which the cancer occurs in all years. Fortunately, the second 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 Appendix D.3.2 and the main conclusions of the 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 were not asked in wave 1).

LTC

For LTC, I exclude individuals I cannot follow for a subsequent 5 years to construct the loss indicator variable; years 2004–2008 are used but only for

---

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

construction of the loss indicator. Also, I exclude individuals who currently reside in a nursing home or receive specialized home care. Finally, I exclude individuals who have missing observations (either the subjective probabilities or observable covariates). For consistency, I exclude any case that is missing any of the extended control variables (results are similar for the price controls and age/gender controls that do not exclude these additional missing cases).

The primary sample consists of 9051 observations from 4418 individuals for the no reject sample, 10,108 observations from 3215 individuals for the reject sample, and 10,690 observations from 5190 individuals for the uncertain sample. In each sample, I include multiple observations for a given individual (which are spaced roughly 2 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 who own insurance, along with an additional 5% of the sample who currently are 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 used 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, 2216 observations from 1280 individuals for the reject classification, and 5361 observations from 1280 individuals for the uncertain classification.[4] 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 5.1, the individual disability insurance market primarily exists for individuals who have 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 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 2689 observations from 1720 individuals for the no reject classification, 2362 observations from

---

[4]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.

1371 individuals for the reject classification, and 6800 observations from 4270 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 that exclude this group for robustness.

APPENDIX D: LOWER BOUND

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, and then discuss 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 nine samples (LTC, life, disability + reject, no reject, uncertain classifications). For the pricing controls specification, I model $\Pr\{L|X, Z\}$ as a probit,

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

where X contains all of the price control variables. The function $\Gamma(\text{age}, Z)$ 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 $\Pr\{L|X, Z\} = \Phi(X\beta + \gamma Z)$.

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

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

For the basis in Z, I use second-order Chebyshev polynomials for the normalized variables, $\tilde{Z} = 2(Z - 50\%)$, plus separate indicators for focal point responses at $Z = 0$, 50, and 100:

$$g_1(Z) = \tilde{Z},$$
$$g_2(Z) = (2\tilde{Z}^2 - 1),$$
$$g_3(Z) = 1\{Z = 0\%\},$$
$$g_4(Z) = 1\{Z = 50\%\},$$
$$g_5(Z) = 1\{Z = 100\%\}.$$

For the basis in age, I use a linear specification, $f_1(\text{age}) = \text{age}$ (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 $\Pr\{L|X, Z\}$. Given these predicted values, I plot the distribution of $\Pr\{L|X, Z\} - \Pr\{L|X\}$ aggregated within each setting and rejection classification. To do so, I also need an estimate of $\Pr\{L|X\}$. For this, I use the same specification as above, except I exclude $\Gamma(\text{age}, Z)$, so that

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

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

Now, for each observation, I have an estimate of $\Pr\{L|X, Z\}$ and $\Pr\{L|X\}$. Therefore, I can plot the predicted empirical distribution of $\Pr\{L|X, Z\} - \Pr\{L|X\}$ 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 $E[m_Z(P_Z)|X]$ 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 $\Pr\{L|X, Z\} - \Pr\{L|X\}$ does not vary conditional on age. Thus, I can aggregate across the residual distribution to construct, for each age, the average difference between one's own probability and the probability of worse risks. I construct the residual $r_i = \Pr\{L|X, Z\} - \Pr\{L|X\}$ for each case in the data. Then, within each age, I compute the average residual, $\Pr\{L|X, Z\} - \Pr\{L|X\}$, of those who have higher residuals within a given age (i.e., for an observation with $r_i = x$, I construct $\hat{r}_i = E[r_i|r_i \geq x, \text{age}]$). Note that this is where I use the assumption that the distribution of $\Pr\{L|X, Z\} - \Pr\{L|X\}$ does not vary conditional on age. I then construct the average of $\hat{r}_i$ in the sample, which equals $E[m_Z(P_Z)|X \in \Theta]$ 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, $L = \beta X + \Gamma(\text{age}, Z) + \epsilon$, 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. *Statistical Drivers of Lower Bound Results*

Figure 2(a)–(c) shows that the elicitations are more predictive of L given X for the rejectees than for nonrejectees in each setting. To help provide greater statistical clarity on the drivers of this result and to relate to more standard regression data analysis used in previous literature, I use a linear functional form to decompose the variance of the predictive distribution into two components: (i) the "slope" of L and Z given X, $\frac{\text{cov}(L,Z|X)}{\text{var}(Z|X)}$, and (ii) the dispersion of the elicitations given X, $E[\text{var}(Z|X)]$.

To do so, I adopt a linear specification

$$\Pr\{L|X, Z\} = \gamma Z + \beta X,$$

so that we have the decomposition[5]

$$\text{var}\big(\Pr\{L|X, Z\} - \Pr\{L|X\}\big) = \gamma^2 E\big[\text{var}(Z|X)\big],$$

so that greater dispersion in $\Pr\{L|X, Z\} - \Pr\{L|X\}$ can be driven by two components: $\gamma = \frac{\text{cov}(L,Z|X)}{\text{var}(Z|X)}$ (assumed to be constant across values of X by the additive specification) and $E[\text{var}(Z|X)]$.

Table A-II presents estimates of γ and $E[\text{var}(Z|X)]$ for rejectees and nonrejectees in each market setting.[6] The first row shows that in all three market settings, the coefficient γ is larger for the rejectees relative to the nonrejectees. This indicates that differences in elicitations lead to larger differences in experienced loss probabilities.

The third row presents the estimated mean residual squared error of a regression of Z on X. This provides an estimate of $E[\text{var}(Z|X)]$. In all three settings, we find larger estimates for rejectees relative to nonrejectees: rejectees have greater dispersion in their subjective probability elicitations.

In short, the greater dispersions in predicted probabilities, $\Pr\{L|X, Z\} - \Pr\{L|X\}$, presented in Figure 2(a)–(c) are driven by a combination of a larger "slope" to the relationship between Z and L, conditional on X, and greater dispersion in Z, conditional on X.

D.3. *Lower Bound Robustness Checks*

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

---

[5]The advantage of the linear specification is that it allows this simple decomposition. All of the results in the paper can be qualitatively reproduced with this specification; however, I use the more flexible (nonlinear) specification for the main results because L is a binary variable and hence the linear specification can be inappropriate.

[6]I estimate $E[\text{var}(Z|X)]$ by adopting the linear specification

$$Z = \Omega X + \varepsilon.$$

TABLE A-II

DECOMPOSITION OF DISPERSION IN $\Pr\{L|X, Z\}$

	LTC		Disability		Life	
Classification	No Reject	Reject	No Reject	Reject	No Reject	Reject
Regression of L on Z, cond'l on X						
Gamma (Y)	−0.004	0.132	−0.033	0.206	0.062	0.159
s.e.[a]	0.013	0.018	0.045	0.041	0.031	0.032
RMSE of regression of Z on X						
$E[\mathrm{var}(Z|X)]$	0.194	0.250	0.239	0.257	0.294	0.320

[a]Standard errors clustered at household level.

D.3.1. *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 A-1 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 nonrejectees, conditional on age.

D.3.2. *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 1993/1994 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 A-III, I report results from a specification restricted to years 1993/1994 that 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 among the rejectees continues to hold ($p = 0.0204$). Moreover, the estimate of $E[m_Z(P_Z)|X \in \Theta^{\mathrm{Reject}}]$ 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.
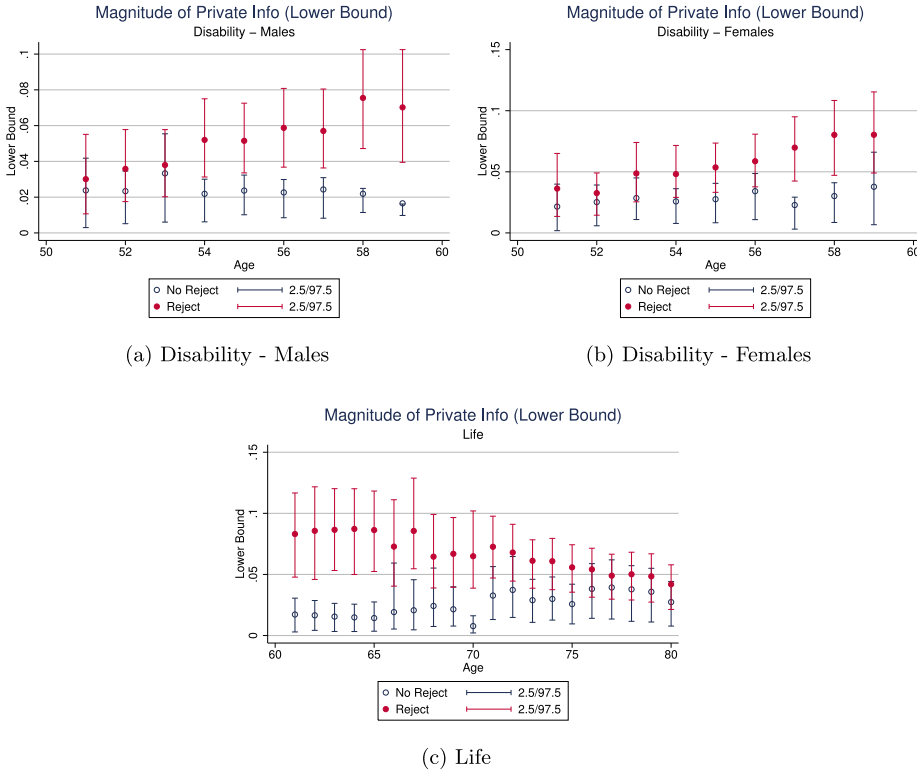
(a) Disability - Males

(b) Disability - Females

(c) Life

FIGURE A-1.—Magnitude of private information by age.

APPENDIX E: STRUCTURAL ESTIMATION

E.1. *Specification Details*

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

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

where $\text{Beta}(\mu, \psi)$ is the p.d.f. of the beta distribution with mean μ and shape parameter ψ. Note that this parameterization of the beta distribution is slightly nonstandard; 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, $\psi_1 = \psi_{\text{central}}$, and one for all other beta distributions, $\psi_i = \psi_{\text{noncentral}}$ ($i = 2, 3$).

TABLE A-III

CANCER ORGAN CONTROLS (LIFE SETTING)

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

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

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

[c]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.

***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

This helps reduce the nonconvexity of the likelihood function.[7] 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 nontrivial 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 small 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 who have 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 $\psi_i \leq 200$.

---

[7]For example, nonconvexity 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.

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 readjust 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 such a case, I adjust the weights w_2 and w_3 to 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 noncentrality of the beta distributions (a_1 and a_2), and the two shape parameters, ψ_{central} and $\psi_{\text{noncentral}}$. Finally, for the LTC no reject and disability no reject samples, I estimate a seventh parameter, which is given by the weight on the point mass, w_{ptmass}.

Estimation

In each of the six samples, estimation is done in two steps.[8] First, I estimate $\Pr\{L|X\}$ using the probit specification described in Section D.1. Second, I estimate the six beta mixture parameters, $\{w_1, w_2, a_1, a_2, \psi_{\text{central}}, \psi_{\text{noncentral}}\}$, along with the four elicitation error parameters $\{\sigma, \kappa, \lambda, \alpha\}$, using maximum likelihood. As is standard with mixture estimation, the likelihood is nonconvex 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 $\sigma \geq \min\{\text{var}(Z^{\text{nf}}) - \text{cov}(Z^{\text{nf}}, L), \sqrt{\frac{3}{8}}\}$. 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 $\sigma \geq \min\{\text{var}(Z^{\text{nf}}) - \text{cov}(Z^{\text{nf}}, L), \sqrt{\frac{3}{8}}\}$, these pathological outcomes are removed. Reassuringly, the constraint does not locally bind in any of my samples (i.e., I find estimates of σ between 0.3

---

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

and 0.45, whereas values of $\text{var}(Z^{\text{nf}}) - \text{cov}(Z^{\text{nf}}, L)$ fall consistently around 0.2 in each setting).

E.2. *Robustness*

Table A-IV 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 quantiles 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 nonrejectees. In general, the results are quite similar to the values reported in Tables V and VI, which considered a characteristic that corresponds to the mean loss, $\Pr\{L|X\} = \Pr\{L\}$.

E.3. *Estimation Results Details*

Measurement Error Parameters

Table A-V 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 the moderate bias of magnitudes less than 10% in all samples except the LTC rejectees, for whom I estimate a substantial 28.6 percentage point 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 A-VI presents the estimated parameters for $f_P(p|X)$, along with the bootstrapped standard errors.

APPENDIX F: Selected Pages From Genworth Financial Underwriting Guidelines

The following four pages contain a selection from Genworth Financial's LTC underwriting guideline that is provided to insurance agents for use in screening applicants. Although marked "[n]ot 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 four of the 152 pages

TABLE A-IV

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%[a]	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%[a]	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 Reject)	**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%[b]	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

[a]5/95% CI computed using bootstrap block re-sampling at the household level ($N = 250$ repetitions); 5% level extended to include 1.00 if p-value of F-test for presence of private information is less than 0.05; Bootstrap CI is bias corrected using the nonaccelerated procedure in Efron (1982).

[b]5/95% CI computed using bootstrap block re-sampling at the household level ($N = 1000$ repetitions); 5% level extended to include 1.00 if p-value of F-test for presence of private information for the rejectees is less than 0.05; Bootstrap CI is the union of the percentile-t bootstrap and bias corrected (nonaccelerated) percentile intervals from Efron and Gong (1983).

TABLE A-V

ELICITATION ERROR PARAMETERS[a]

	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)

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

of the guidelines. The conditions documented below are not exhaustive for the list of conditions that lead to rejection: they constitute the set of conditions that 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.

TABLE A-VI

BETA MIXTURE PARAMETERS[a]

	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)
Noncentrality 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)
Noncentrality 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)			

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

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.

For agent use only.
Not for use with
consumers or to be
distributed to the public.

85440 06/10/09

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)
   Camprai (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

REFERENCES

CHADE, H., AND E. SCHLEE (2011): "Optimal Insurance With Adverse Selection," *Theoretical Economics*, 7 (3), 571–607. [1]

EFRON, B. (1982): *The Jackknife, the Bootstrap, and Other Resampling Plans*. Society of Industrial and Applied Mathematics CBMS-NSF Monographs, Vol. 38. Philadelphia: SIAM. [27]

EFRON, B., AND G. GONG (1983): "A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation," *The American Statistician*, 37 (1), 36–48. [27]

FINKELSTEIN, A., AND K. MCGARRY (2006): "Multiple Dimensions of Private Information: Evidence From the Long-Term Care Insurance Market," *American Economic Review*, 96 (4), 938–958. [13]

HE, D. (2009): "The Life Insurance Market: Asymmetric Information Revisited," *Journal of Public Economics*, 93 (9–10), 1090–1097. [16]

RILEY, J. G. (1979): "Informational Equilibrium," *Econometrica*, 47 (2), 331–359. [1]

SPENCE, A. M. (1978): "Product Differentiation and Performance in Insurance Markets," *Journal of Public Economics*, 10 (3), 427–447. [1]

Dept. of Economics, Harvard University, Cambridge, MA 02138, U.S.A. and NBER; nhendren@fas.harvard.edu.