A Theoretical Extensions

A.1 The Planner's Problem

In addition to the competitive equilibrium studied in the prior section, the question of whether and how welfare changes under a planner's optimal allocation is of interest for multiple reasons. First, to be sure that the conclusions drawn from the previous section are not overly sensitive to the particularities of the Rothschild and Stiglitz (1976) setup. Second, a planner might want to adjust the conditions under which firms compete so as to implement the socially optimal outcome. I will show that the patterns for how misperceptions relate to welfare identically track the Rothschild and Stiglitz (1976) style formulation from section 2.1.

I continue to assume that the planner has preferences as in (2.1) and can implement any incentive compatible and budget-balanced menu of contracts. Their problem is then:

$$\max_{\boldsymbol{c}_H, \boldsymbol{c}_L} \rho V(p_H, \boldsymbol{c}_H) + (1-\rho) V(p_L, \boldsymbol{c}_L) \tag{A.1}$$

$$\text{such that} \tag{A.2}$$

$$V(q_H, \boldsymbol{c}_H) \geq V(q_H, \boldsymbol{c}_L) \tag{A.3}$$

$$V(q_L, \boldsymbol{c}_L) \geq V(q_L, \boldsymbol{c}_H) \tag{A.4}$$

$$\alpha_H \pi(p_H, \boldsymbol{c}_H) + \alpha_L \pi(p_L, \boldsymbol{c}_L) = 0 \tag{A.5}$$

$$(\boldsymbol{c}_H, \boldsymbol{c}_L) \in A = \left\{ \boldsymbol{c}_H, \boldsymbol{c}_L : MRS_H \leq \frac{1-p_H}{p_H} \wedge c_L^{NL} \geq c_L^L \right\} .. \tag{A.6}$$

The usual restriction on the contract space is that $c_{NL} \geq c_L$ for both types and is called an indemnity constraint (as in Netzer and Scheuer (2014)) such that the loss cannot make an individual better off. I wish to keep this in spirit - not allowing individuals to harm themselves through perceived over-insurance - but I do not want to shut down the wedge between subjective and objectively defined optimal insurance. So I define the set of acceptable contracts, A, such that $\boldsymbol{c}$ can only allow for small over-insurance for the high type, and I will assume that errors are small, such that the only over-insurance that can be bought is that which improves perceived utility for the high type.

This rules out equilibria where, according to the misperceived risk, a type could be made better off with a profit-neutral reduction of insurance. If no errors are made, $\boldsymbol{q} = \boldsymbol{p}$, this reduces to the standard indemnity constraint. If one omits this restriction, uniqueness of the equilibria that follow might be lost, and the new equilibria will be qualitatively similar but possibly inverted.

Equations ($A.3$) and ($A.4$) are incentive compatability constraints for the high and low risk types respectively, and equation ($A.5$) is an aggregate resource constraint which allows for cross-subsidization between types, unlike in the competitive equilibrium in which each contract individually broke even. I re-emphasize that whilst incentive constraints are evaluated according to subjectively perceived probabilities q_i, welfare and the resource constraint are evaluated according to objective probabilities p_i.

The first proposition characterizes the constraints that bind at the planners optimum.

Proposition 5. *At the planner's optimum, constraint (A.3) binds, while (A.4) is slack. Constraint (A.6) binds in that* $MRS_H = \frac{1-p_H}{p_H}$ *at the optimum.*

This demonstrates that the qualitative structure of the planners solution mirrors the competitive equilibrium. The high type will receive full insurance or infinitesimally more than full insurance if they make an upward error. The low type will receive partial insurance constrained and distorted by the high type's incentive compatibility constraint.

Initially the contract space has four free parameters. Proposition 5 shows that there is at most on degree of freedom left. It will be helpful to think in terms of the cross-subsidization from low types to high types, so I define:

$$\chi = p_H c_H^L + (1 - p_H) c_H^{NL} - (w - l p_H) \geq 0.$$

Given a level of cross-subsidization χ, this defines the profit/loss level to be earned from each type's contract. In combination with the binding part of A, this defines the high type contract. The incentive constraint of the high type then fixes the low type's contract. By this logic thinking in terms of χ as the final free-parameter is valid and parsimonious.

The next lemma relates the Pareto weight ρ to the degree of cross-subsidization χ that prevails in the optimum.

Lemma 3. *Suppose there are two planners with different Pareto weights on the high type: $\rho_1 > \rho_2$. Then at the respective optima the former weight leads to more cross-subsidization from the low types to the high types than the latter weight: $\chi^*(\rho_1) \geq \chi^*(\rho_2)$, holding strictly when $\boldsymbol{c}_H \neq \boldsymbol{c}_L$.*

When marginally more weight is placed on the high types utility, higher cross-subsidization from the low types to the high types is optimal for the planner. This result is intuitive, but is important to keep in mind for the slightly less intuitive results in the next proposition. The next propostion, the main one of this section, establishes that the qualitative features of the equilibrium in section 2.1 remain true for the planners problem.

Proposition 6. *Fix any ρ. At the planner's optimum the following hold for small enough $\xi > 0$:*

- *For upward misperceptions by the high types $q_H = p_H + \xi$, welfare increases and cross-subsidization decreases relative to no misperceptions.*
- *For downward misperceptions by the high types $q_H = p_H - \xi$, welfare decreases and cross-subsidization increases relative to no misperceptions.*
- *Misperceptions by the low types have no impact on the optimum.*

As ρ varies the entire range of constrained Pareto efficient outcomes are selected by the planner. The proposition above shows that for any particular ρ, welfare increases (decreases) but cross-subsidization decreases (increases) when there is an upward (downward) misperception is made by the high type. Intuitively, this is because the set of implementable contracts has strictly gotten bigger (smaller) due to the misperception loosening (tighetening) the high types IC constraint. The change to the IC constraint allows for welfare to strictly rise or fall. Similarly to the RS equilibrium, because the low type's IC constraint never binds at the optimum, misperceptions by them have no effect on the optimum or on welfare.

To understand the cross-subsidization result, note that without misperceptions the optimum features a net zero welfare change if one more dollar of subsidy were to go from the low types to the high types. The low types would strictly be worse off, the high types better off, but at the optimum these cancel out. Now with an upward misperception, the incentive constraint is loosened and at the previous level of cross-subsidization the low type is better off. But because the low type is better off and closer to full insurance, the marginal dollar of cross-subsidization is even less desirable to them given their relatively improved contract. The marginal dollar of cross-subsidization is, to the first order, of the same value to the high type as before the error. So the net value of the marginal dollar of cross-subsidization is strictly lower than without misperceptions, and hence less cross-subsidization prevails in equilibrium.

A.1.1 The Miyazaki-Wilson allocation as a particular constrained efficient allocation

Issues with the existence of the Rothschild-Stiglitz competitive equilibrium have lead to the use of alternative equilibrium notions. A leading alternative is that of Wilson (1977), Miyazaki (1977) and later Spence (1978). Since Miyazaki (1977) the MW allocation has been thought of as a particular solution to the planner's problem, micro-founded by the restricted deviation set in Wilson (1977).

In this context, with misperceptions, the game-theoretic logic of Wilson (1977) and the constrained efficient allocation of Miyazaki (1977), Spence (1978), Netzer and Scheuer (2014) and others are not coincident. In this section, I study the latter interpertation, which is more persuasive and interesting. In appendix A.3 I consider the game-theoretic interpretation more in line with Wilson (1977).

For now, write $(\boldsymbol{c}_H^{RS}, \boldsymbol{c}_L^{RS})$ for the allocations that prevail in the Rothschild-Stiglitz equilibrium in section 2.1. Then define an expanded program $(A.2)^*$ as $(A.2)$ subject to $(A.3)$ through $(A.6)$ with the constraint $V(q_H, \boldsymbol{c}_H) \geq V(q_H, \boldsymbol{c}_H^{RS})$ appended. That is, $(A.2)^*$ is program $(A.2)$ with the extra requirement that the high types do no worse than the competitive equilibrium. This is defined by analogy to Netzer and Scheuer (2014). We can then write:

Definition 5. *Denote the allocation $(\boldsymbol{c}_H, \boldsymbol{c}_L)$ that solves $(A.2)^*$ as the MW allocation.*

The following then follows directly from the preceding results of this section.

Corollary 2. *The MW allocation has the following properties:*

1. *The constraints bind as in Proposition 5.*
2. *(Small) upward misperceptions by the high type increase welfare, downward misperceptions decrease welfare. Misperceptions by the low type do not affect welfare.*
3. *Amongst the set of constrained efficient allocations with positive cross-subsidization, the MW allocation features the minimal level of cross-subsidization.*

A.2 Monopoly

As another robustness check that the particular equilibrium concept is not driving the qualitative results, here I consider the provision of insurance by a monopolist instead of a competitive market or planner. As standard, the monopolists problem is to, writing the endowment as $\mathbf{0} = (w, w - l)$,:

$$\max_{\boldsymbol{c}_H, \boldsymbol{c}_L} \sum_{i \in \{L,H\}} \alpha_i \pi(p_i, \boldsymbol{c}_i) \tag{A.7}$$

$$\text{such that} \tag{A.8}$$

$$V(q_H, \boldsymbol{c}_H) \geq V(q_H, \boldsymbol{c}_H) \tag{A.9}$$

$$V(q_L, \boldsymbol{c}_L) \geq V(q_L, \boldsymbol{c}_H) \tag{A.10}$$

$$V(q_H, \boldsymbol{c}_H) \geq V(q_H, \boldsymbol{0}) \tag{A.11}$$

$$V(q_L, \boldsymbol{c}_L) \geq V(q_L, \boldsymbol{0}). \tag{A.12}$$

The equilibrium configuration is a direct adaptation of the various parts of Jeleva and Villeneuve (2004) and is summarized below. Using their terminology: when I say one type receives optimal insurance means that the type is receiving their first best contract **fixing the level of profit** from that types contract. Or geometrically, it means that the type's indifference curve is tangent to the iso-profit line at their equilibrium contract.

Proposition 7 (Jeleva and Villeneuve (2004))**.** *Suppose $q_H > q_L$ and errors are small. In the unique optimum:*

1. *(A.12) and (A.9) bind.*
2. *Types are always seperated.*
3. *H types obtain optimal insurance. If $q_H > p_H$ this is over-insurance, if $q_H < p_H$ then under insurance.*
4. *L types are sub-optimal partial insurance, perhaps none.*
5. *There is a threshold $\overline{\gamma_H}$ such that low types receive no insurance if $\gamma_H > \overline{\gamma_H}$ and are served if not.*

The case where the low type is not served is not particularly interesting. In such a case low type errors clearly have no impact. The high types participation constraint binds according to their perceived probability, so a pessimistic error is bad and an optimistic error good according to their objective welfare as calculated by the planner.

The more interesting dynamics occur when both types are served. The monopolist faces a trade-off: They can extract more profit from the low types by offering them more and more insurance but this requires less (perhaps negative) profits to be earned from the high types due to the binding incentive constraint. Conversely, they can extract more profits from the high types by raising the price of optimal insurance and pushing the high types closer to their outisde option, but for the same incentive reasosn the low types contract must then feature even less insurance and profit (perhaps zero of both) so as to not tempt the high type. These forces will help with interpreting the welfare implications of errors.

First, consider a small upward misperception by the high type with no error by the low type. The low type will continue to receive utility equal to their outside option and so their welfare will not change. The high types incentive constraint has now loosened. By how much it has loosened depends on where their risk was to begin with. When they were already high risk, the loosening is greater (at the limit, consider flat indifference curves) and so per dollar of extra profit to be extracted from the low type the incentive cost to the high type is lower. This motivates the (sufficient) condition $2p_H > 1 + p_L$. The higher risk the high type is, the higher the maginal loosening in incentive cost is when they make a small error, and so the temptation to earn more profits from the low types whilst making the high types slightly better off is stronger.

On the other hand, suppose the low type makes a small error. The effect on their welfare is unambiguous. If the low types think they are riskier than they are, the monopolist will sell them insurance they don't need, still restrict them to zero utility - but according to their perceived pessimistic probability - which the planner will then determine as detrimental to their welfare as defined by the objective probability. On the other hand, if the low types are optimistic, their zero perceived utility earned will actually be an objective improvement in the planners eyes.

But the effect of a low type error on a high type is more subtle. To charge more for insurance to the low types (their IR curve has shifted left), the high types must be made better off. If the high types are very different from the low types in risk, then the insurer can charge the low types more for the same loss payout whilst not affecting their profit from the high types much at all (at the limit, again, consider high types having horizontal indifference curves). If the high types are closer in risk to the low types, then extracting profit from the low type will require a greater marginal loss from the high types. Hence, the (sufficient) condition $\frac{p_H(1-p_L)}{p_L(1-p_H)} < \frac{u'(w)}{u'(w-l)}$ says that when p_H is small and p_L large, i.e. their difference is small, the low types naivete can be exploited at lower change to high types profit.

Proposition 8. *For a small error, the effects on welfare for the high and low types are:*

For an error by the high type $\frac{\partial V_H(p_H,\boldsymbol{c}^*)}{\partial q_H} \mid_{\boldsymbol{q}=\boldsymbol{p}} > 0$ *if* $2p_H > 1 + p_L$, $\frac{\partial V_L(p_L,\boldsymbol{c}^*)}{\partial q_H} \mid_{\boldsymbol{q}=\boldsymbol{p}} = 0.$. *For an error by the low type* $\frac{\partial V_H(p_H,\boldsymbol{c}^*)}{\partial q_L} \mid_{\boldsymbol{q}=\boldsymbol{p}} > 0$ *if* $\frac{p_H(1-p_L)}{p_L(1-p_H)} < \frac{u'(w)}{u'(w-l)}$, $\frac{\partial V_L(p_L,\boldsymbol{c}^*)}{\partial q_L} \mid_{\boldsymbol{q}=\boldsymbol{p}} < 0.$.

As compared to the competitive market dynamics and the planners optimal choice problem, the key difference here is that errors by both types matter for welfare. But the main qualitative insight holds: errors that make types more disparate (which, loosely speaking, might 'increase' the amount of private information) can actually be good for welfare by weakening the incentive constraints. When there is no endogenous contract adjustment, this channel is closed down and the welfare conclusions might be changed.

A.3 Wilson-Miyazaki equilibrium, alternative interpretation

Two concerns that the have been repeatedly levelled at the Rothschild and Stiglitz (1976) equilibrium concept are that sometimes no equilibrium exists and even when the equilibrium does exist it need not be pareto optimal. To address these concerns a loosening of the equilibrium notion was developed by Wilson (1977) and subsequently further studied by Miyazaki (1977), Spence (1978) and Netzer and Scheuer (2014) amongst others. When deciding whether a given menu is an equilibrium, instead of allowing any profitable deviations, the Wilson-Miyazaiki (WM) concept only considers deviations that continue to make a non-negative profit even after all contracts rendered unprofitable by the initial deviation are withdrawn. Thus, firms anticipate best responses to their deviations in this limited manner. A formal definition is given in Wilson (1977).

There are two ways to interpret the MW equilibrium. In game-theoretic terms as described above or as a particular contrained efficient solution to the planner's problem. THe literature has largely taken the latter as canonical (see e.g. Netzer and Scheuer (2014)). For that reason that constrained efficient solution is priveleged in the main body of the paper. In this appendix I study the game-theoretic interpretation of the MW equilibrium for theoretical completeness.

From now on I focus on the maximization problem that defines the MW equilibrium. Although there is no planner performing this maximization, the market behaves as if the (subjectively (mis-)perceived)) welfare of the low risk type is being maximized subject to constraints. Consider the program:

$$\max_{\boldsymbol{c}_H, \boldsymbol{c}_L} V(q_L, \boldsymbol{c}_L) \tag{A.13}$$

such that (A.14)

$$V(q_H, \boldsymbol{c}_H) \geq V(q_H, \boldsymbol{c}_L) \tag{A.15}$$

$$V(q_L, \boldsymbol{c}_L) \geq V(q_L, \boldsymbol{c}_H) \tag{A.16}$$

$$V(q_H, \boldsymbol{c}_H) \geq V(q_H, \boldsymbol{c}_H^{RS}) \tag{A.17}$$

$$\alpha_H \pi(p_H, \boldsymbol{c}_H) + \alpha_L \pi(p_L, \boldsymbol{c}_L) = 0 \tag{A.18}$$

$$(\boldsymbol{c}_H, \boldsymbol{c}_L) \in A = \left\{ \boldsymbol{c}_H, \boldsymbol{c}_L : MRS_H \leq \frac{1 - p_H}{p_H} \wedge c_L^{NL} \geq c_L^L \right\}. \tag{A.19}$$

This generalizes the program in Miyazaki (1977) and Netzer and Scheuer (2014) by allowing for q_H and q_L to differ from p_H and p_L. In addition I make one more restriction on the equilibrium. The usual restriction on the contract space is that $c_{NL} \geq c_L$, for both types and is called an indemnity constraint (as in Netzer and Scheuer (2014)) such that the loss cannot make an individual better off. I wish to keep this in spirit - not allowing individuals to harm themselves through perceived over-insurance - but I do not want to shut down the wedge between subjective and objectively defined optimal insurance. So I define the set of acceptable contracts, A, such that $\boldsymbol{c}$ can only allow for small over-insurance for the high type, and somce I will assume that errors are small, such that the only over-insurance that can be bought is that which improves perceived utility for the high type. This assumption rules out non-infinitesimal over-insurance for the low types, so the standard indemnity constraint can remain for them.

This rules out equilibria where, according to the misperceived risk, a type could be made better off with a profit-neutral reduction of insurance. If no errors are made, $\boldsymbol{q} = \boldsymbol{p}$, this reduces to the standard indemnity constraint. If one omits this restriction, uniqueness of the equilibria that follow might be lost, and the new equilibria will be qualitatively similar but possibly inverted.

Proposition 9. *For small errors $\xi = \|\boldsymbol{q} - \boldsymbol{p}\|$,, the unique solution to the program (A.13) has (A.18), (A.15) and (A.19) binding and coincides with the the unique MW equilibrium also constrained by (A.19).*

There is a continuum of contract pairs that satisfy the three constraints in the above program. They differ only in their degree of cross-subsidization from low-type to high type (never the other way). It will be useful to think in terms of this cross-subsidization, so I define:

$$\chi = p_H c_H^L + (1 - p_H) c_H^{NL} - (w - l p_H) \geq 0.$$

And so program (A.13) can be written as a univariate maximization over χ where the constraints above and definition of χ implicitly define the contracts as functions of χ, e.g. $c_H^{NL} = c_H^{NL}(\chi)$ and so on. The solution can then be thought of as the optimal amount of cross-subsiziation from low types to high types to maximise the perceived utility of the former.

When $\chi = 0$ this reduces to the RS equilibrium. When $\chi > 0$ the equilibrium features a profit-making contract offered to the low types and a loss-making contract offered to the high types that makes both better off relative to $\chi = 0$. RS is the particular case where the low types are not willing to cross-subsidize the high types so that the high types will receive more in both states of the world and the low types can move closer to full insurance owing to the relaxed IC constraint. This happens, most intuitively, when there are too many high types, so that the cross-subsidization is too expensive to justify. On the other hand when there are many low types, the RS equilibrium, even if it existed, would be inefficient as the high types could all pay a small subsidy to get a relatively large relaxation of the high types IC constraint.

The prior sections study the comparative statics when $\chi = 0$. Here I focus on the case where, prior to errors being made, there is a non-zero amount of cross-subsidization: $\chi > 0$.

It will turn out that in that situation the movement of χ with q_H or q_L is welfare sufficient. In particular, welfare increases with cross-subsidization. And so if some errror increases the cross-subsidization in the equilibrium then welfare will increase, and vice versa. And so if I sign $\frac{\partial}{\partial q_H}\chi^*(q_H)$ or similarly for q_L the welfare implications will follow.

Proposition 10. *Suppose without errors $\chi^* > 0$. Then if a small error is made: welfare increases in χ, $\frac{\partial}{\partial q_H}\chi^*(q_H) < 0$ and $\frac{\partial}{\partial q_L}\chi^*(q_L) > 0$.*

To understand the intuition for this result, one should keep in mind that the χ^* in the MW equilibrium is pushed higher and higher so long as the low types are still willing to give a dollar to the high types to relax the incentive constraint. The question is how an error by either type changes the incentive for the low types to give the final dollar of cross-subsidy to the high types.

First suppose the high types make a small upward error, $q_H > p_H$. Then, holding the no-error χ^* fixed, the high types contract will move up and left on the same iso-profit line. (except with the iso-profit line not necessarily earning zero). The low types will then also receive closer to full insurance and be better off even before the cross-subsidy is adjusted. But because the low type is now closer to full insurance, the marginal utility from relaxing the cross-subsidy is now lessened, and so the final dollar of cross-subsidy that was optimal with no errors is now longer optimal with this error. There is a counterveiling force: that the marginal relaxation of the incentive constraint with a dollar of cross-subsidy is larger with the higher q_H (as the indifference curves are flatter) but this is dominated by the first effect. As such the new $\chi^* \mid_{q_H > p_H} < \chi^* \mid_{q_H = p_H}$. But welfare is evalauted according the objective probabilities, and $\chi^* \mid_{q_H = p_H}$ was optimal according to objective probabilities, and so the final foregone dollar of cross-subsidy is welfare decreasing.

Conversely, but similarly, if the low types make a downward error, $q_L < p_L$, they are no longer willing to pay for the final dollar of cross-subsidy, even though it is welfare optimal. This explains the sign of $\frac{\partial}{\partial q_L} \chi^*(q_L) > 0$.

These results reverse the logic from the RS equilibrium: errors that make a type think they are closer to the other type are welfare enhancing, errors that spread the difference between perceived types are welfare reducing. The intuition here is that **type-spreading errors are substitutes for cross-subsidies.** That is, iin the RS equilibrium where cross-subsidies are impossible, the only mechanism by which the incentive constraint can be loosened is through errors that spread the types apart. This is what was found in proposition 3. In the MS equilibrium, cross-subsidies are present without errors. Those cross-subsidies loosen the incentive constraint by as much as the low type is willing to pay. When we add an error which also loosens the incentive constraint, the low types cross-subsidize a little less. In this sense when cross-subsidies are present they are crowded out, suboptimally, when errors are introduced.

This highlights a crucial distinction: the impact of errors depends on whether or not contracts are currently cross-subsidizing each other.

A.4 Medium Errors

In the main paper, I studied the cases where errors were infinitismally small and where they were large enough to lead to changed pooling. There is a third case, where errors are large enough to change IC constraints, but not large enough to lead to increased pooling. I explore that here.

Amongst those with true risk p_H, we define $q_{H,O}$ and $q_{H,P}$ as two different perceived risk classes: The former being relative optimists, the latter relative pessimists, such that $q_{H,O} < q_{H,P}$. Similarly define $q_{L,O}$ and $q_{L,P}$ as the perceived risk of the optimists and pessimists within the low types. I write $\alpha_{H,O}, \alpha_{H,P}, \alpha_{L,O}, \alpha_{L,P}$ for the proportion of the population that falls into each of these three classes. In this section I study the case in which

$$q_{H,P} > q_{H,O} > q_{L,P} \gg q_{L,O}. \tag{A.20}$$

This leads to qualitatively similar outcomes as in the 'small' misperceptions case in section 2.1 model. The equilibrium that obtains if we assume (A.20) is illustrated in figure 7. It is formalized in the following proposition.

Proposition 11. *Suppose A.20 holds. In the unique locally-competitive equilibrium, high risk pessimistic individuals buy the contract $\boldsymbol{c}_{H,P}$ that solves:*

$$MRS(q_{H,P}, \boldsymbol{c}_{H,P}) = \frac{1-p_H}{p_H} \text{ and } \pi(p_H, \boldsymbol{c}_{H,P}) = 0,$$

high risk optimistic individuals buy the contract $\boldsymbol{c}_{H,O}$ that solves:

$$MRS(q_{H,O}, \boldsymbol{c}_{H,O}) = \frac{1-p_H}{p_H} \text{ and } \pi(p_H, \boldsymbol{c}_{H,O}) = 0,$$

and low risk pessimistic individuals receive the contract $\boldsymbol{c}_L$ *that solves*

$$V(q_{H,O}, \boldsymbol{c}_{H,O}) = V(q_{H,O}, \boldsymbol{c}_L) \text{ and } \pi(p_L, \boldsymbol{c}_L) = 0,$$

and sufficiently low risk optimistic individuals receive the contract that solves

$$MRS(q_{L,O}, \boldsymbol{c}_{L,O}) = \frac{1 - p_L}{p_L} \text{ and } \pi(p_L, \boldsymbol{c}_{L,O}) = 0.$$

The high risk types, both optimistic and pessimistic, obtain their perceived first best contract subject to zero profit being earned. Naturally, $\boldsymbol{c}_{H,P}$ features more insurance than $\boldsymbol{c}_{H,O}$ as the pessimists think the risk is more likely then the optimists and so have a greater demand for insurance.

The low risk pessimistic receive partial insurance defined by the IC constraint of the optimistic high type, and the zero-profit condition. If the low risk optimists are sufficiently optimisic, they prefer even less insurance than the partial insurance defined by $IC_{H,o}$ and receive their subjectively preferred contract on the zero profit line according to their real risk $\pi(p_L,) = 0$.

The equilibrium menu is separating in the sense that all true high types are separated from all true low types, as opposed to 'large' errors that cause pooling. The incentive constraint that binds to determine the low types contract is that of the most optimistic high type. But this is optimistic only relative to the pessimistic high types, not necessarily relative to the truth. That is, if $p_H > q_{H,O}$ (as is illustrated in figure 7) in which case the misperception will certainly create a welfare loss in the market in the same way as proposition 3. Or, if $p_H < q_{H,O} < q_{H,P}$ then if $q_{H,O} - p_H$ is small enough then the misperceptions can lead to a welfare improvement as also discussed in proposition 3.

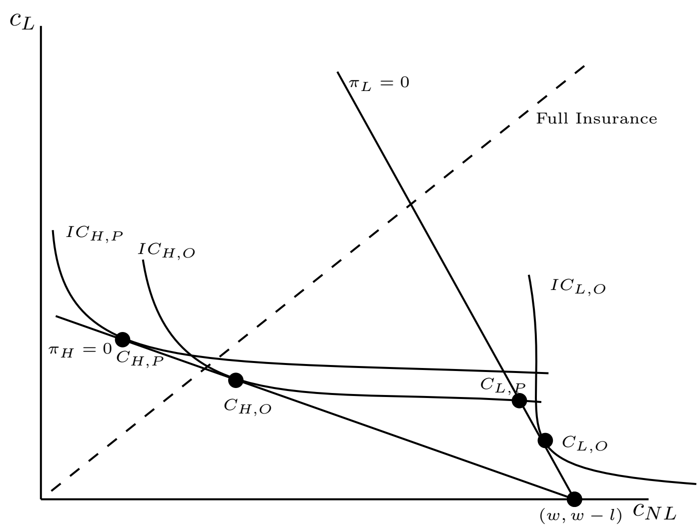


Figure 7: Separating equilibrium due to 'medium' errors

B Proofs

B.1 Proof of Proposition 1

Proof. We prove a more general claim with three types, H, M and L. The result follows by allowing M and L to have the same low risk. The claim is:

Proposition. *Suppose $\|\boldsymbol{p} - \boldsymbol{q}\|_\infty < \xi$, with $q_H > q_L$. For small enough ξ, in the unique locally-competitive equilibrium, high risk individuals buy the contract $\mathbf{c}_H$ that solves:*

$$MRS(q_H, \boldsymbol{c}'_H) = \frac{1 - p_H}{p_H} \textit{ and } \pi(p_H, \boldsymbol{c}'_H) = 0,$$

medium risk individuals receive the contract $\mathbf{c}_M$ that solves

$$V(q_H, \boldsymbol{c}'_H) = V(q_H, \boldsymbol{c}'_M) \textit{ and } \pi(p_M, \boldsymbol{c}'_M) = 0,$$

and low risk individuals receive the contract $\boldsymbol{c}_L$ that solves

$$V(q_M, \boldsymbol{c}'_M) = V(q_M, \boldsymbol{c}'_L) \text{ and } \pi(p_L, \boldsymbol{c}'_L) = 0.$$

Step 1. No pooling

First, note an intuitive but repeatedly used fact: for types i, j if $q_i > q_j$ then the equilibrium contract that i buys (and perhaps others) must offer weakly more insurance than the equilibrium contract j buys: $\boldsymbol{c}_i \succsim \boldsymbol{c}_j$. If i and j pool this is trivial. If not, then either the contracts are not comparable with respect to $\succsim$, in which case one dominates the other and an IC constraint must break, or $\boldsymbol{c}_j \succ \boldsymbol{c}_i$. In that case, I must have $V(q_j, \boldsymbol{c}_j) \geq V(q_j, \boldsymbol{c}_i)$, but also that $V(q_i, \boldsymbol{c}_j) > V(q_i, \boldsymbol{c}_i)$ since $MRS(q_i, \boldsymbol{c}_j) < MRS(q_j, \boldsymbol{c}_j)$ and $\boldsymbol{c}_j \succ \boldsymbol{c}_i$. This contradicts the i types IC constraint. Hence higher types must receive weakly more insurance that lower types in equilibrium. Denote this observation by **(*)**

The impossibility of pooling between H and M types, and between M and L types, and between all three types, in a RS equilibrium, and hence in a locally-competitive equilibrium, follows by almost identical arguments to R-S. I need only check that the deviating contract that attracts the relatively lower type in each situation does not attract the currently unpooled type, if relevant. For the latter two cases this is obvious. For the first, where I conjecture one contract that pools and one that attracts only L, $\boldsymbol{c}_{HM}, \boldsymbol{c}_L$, note simply that the the low typs IC constraint cannot bind. If it did, then the medium types IC constraint wouldn't bind (as the indifference curves are distinct), and so a deviation to offering $\boldsymbol{c}_L + (-\epsilon, \epsilon r)$ with $r \in (MRS(q_L, \boldsymbol{c}_L), \frac{1-p_L}{p_L})$ will by construction improve utility, earn positive profits, yet not attract any M or H types for small enough $\epsilon > 0$. So the low types IC constraint does not bind, so a contract marginally more attractive to the medium type will not attract the low type, and hence the pooling between high and medium types cannot be sustained. [26]

[26] Note an $r \in (MRS(q_L, \boldsymbol{c}_L), \frac{1-p_L}{p_L})$ exists when q_L receives less than full insurance. If L is has ful insurance, their IC constraint cannot bind as they are at their first-best. If they are strictly over-insured then they can move toward full insurance with a positive profit made.

Finally, H and S cannot pool in any locally competitive equilibrium. If they did, then by the observation (*) above the contracts must satisfy $\boldsymbol{c}_{HS} \succ \boldsymbol{c}_M$ or the opposite. In the former case, I contradict (*) with S types being more insured than M types. In the latter, I contradict the same (*) with M types being more insured than H types. This establishes that pooling cannot be sustained in any configuration.

Step 2. Any allocation except the solution to the maximization problem in the Proposition cannot be a local equilibrium

Step 1 shows that an equilibrium menu must feature three seperating contracts with $\boldsymbol{c}_H \succ \boldsymbol{c}_M \succ \boldsymbol{c}_L$.

We first verify that the incentive constraints cannot bind upwards. Assume that $\boldsymbol{c}$ underinsures.[27] For a contradiction suppose $V(q_M, \boldsymbol{c}_M) = V(q_M, \boldsymbol{c}_H)$. This implies $V(q_H, \boldsymbol{c}_M) < V(q_H, \boldsymbol{c}_H)$. Hence, for small enough $\epsilon > 0$, there is a new contract $\boldsymbol{c}'_M = \boldsymbol{c}_M + (-\epsilon, \epsilon r)$ with $r \in \left(MRS(q_M, \boldsymbol{c}_M), \min\left\{\frac{1-q_M}{q_M}, MRS(q_L, \boldsymbol{c}_L)\right\}\right)$ that strictly improves on the utility of the medium types, yet makes positive profits as it doesn't attract any other types. Hence the IC constraint preventing the medium type from reporting themselves as high doesn't bind in any equilibrium.

Similarly, I cannot have $V(q_L, \boldsymbol{c}_L) = V(q_L, \boldsymbol{c}_M)$. Assume that $\boldsymbol{c}_L$ features underinsurance.[28]. If I did have $V(q_L, \boldsymbol{c}_L) = V(q_L, \boldsymbol{c}_M)$, then I must have $V(q_M, \boldsymbol{c}_L) < V(q_M, \boldsymbol{c}_M)$. This also implies that $V(q_H, \boldsymbol{c}_L) < V(q_H, \boldsymbol{c}_M)$ as the MRS curves are always shallower for H types. Hence with $r \in \left(MRS(q_L, \boldsymbol{c}_L), \frac{1-q_L}{p_L}\right)$, then $\boldsymbol{c}_L + (-\epsilon, \epsilon r)$ will make low types strictly better off, earning strictly positive profits by just attracting low types so long as $\epsilon > 0$ is small enough.

Then, suppose the H types received any contract except the specific contract whilst

[27] If not, then since $MRS\left(q_M, \boldsymbol{c}_M\right) > \frac{1-p_M}{p_M}$ for overinsured contracts, the M types indifference curve is always above $(1-p_M)(W - c_M^{NL}) + p_M(c_M^L - (W-l)) = 0$ for any contract that features more insurance than $\boldsymbol{c}_M$, and hence cannot intersect with $(1-p_H)(W - c_H^{NL}) + p_H(c_H^L - (W-l)) = 0$ at such contract, and hence the IC constraint for the medium type cannot bind.

[28] If not, then the logic of the footnote above applies

maintaining zero profits. Then I must have $MRS\left(q_H, \boldsymbol{c}_H\right) \neq \frac{1-p_H}{p_H}$. In either case there is a contract slightly north-west or south-east of the current contract (depending on whether the inequality is $>$ or $<$, that offers positive profits and strictly higher utility to the H types without attracting any lower types when $\epsilon > 0$ is small since their IC constraints do not bind upwards.

Then, suppose that $V(q_H, \boldsymbol{c}_H) > V(q_H, \boldsymbol{c}_M)$. Then the deviation from I considered M's IC constraint (pretending to be H) is a profitable, utility improving deviation here as well (assuming underinsurance at $\boldsymbol{c}_M$)[(29)] . Similarly, suppose that $V(q_M, \boldsymbol{c}_M) > V(q_M, \boldsymbol{c}_L)$. Then the deviation from when I considered Ls IC constraint suffices here as well. Note that that deviation relies on L being underinsured, which must be the case since they can always be offered less insurance without affecting the IC constraints.

This shows that the given menu is the only possible equilibrium configuration.

Step 3. Existence

It remains to verify that the proposed menu is indeed a locally competitive equilibrium.

Consider first $\boldsymbol{c}_H$. At $\boldsymbol{c}_H$ I have $MRS(q_H, \boldsymbol{c}_H) = \frac{1-p_H}{p_H}$. Further, since $U'' < 0$ these curves are tangent.This directly means that any contract attracting only H types cannot both make a non-negative profit and also strictly improve upon H's utility. As above, the IC constraints of the lower types do not bind upward, and so no local deviation will attract anyone but the H types. It follows there are no local deviations near $\boldsymbol{c}_H$.

Consider next $\boldsymbol{c}_M$. Recall that $MRS(q_M, \boldsymbol{c}_M) < \frac{1-p_M}{p_M}$ for small enough ξ. Consider a deviation in the direction $(-1, r)$ from $\boldsymbol{c}_M$ with $r \in \left[MRS\left(q_M, \boldsymbol{c}_M\right), \frac{1-p_M}{p_M}\right]$. Since $MRS\left(q_M, \boldsymbol{c}_M\right) > MRS\left(q_M, \boldsymbol{c}_M\right)$ this contract attracts high and medium types, but for small ϵ this will be loss making since $p_{HM} > p_M$ and so the pooled contract will make profits arbitrarily close to $(1-p_{HM})(W-c_M^{NL})+p_{HM}(c_M^L-(W-l)) < (1-p_M)(W-c_M^{NL})+p_M(c_M^L-(W-l)) = 0$. Similarly, any contract deviations with $r \in [MRS\left(q_M, \boldsymbol{c}_M\right), MRS\left(q_M, \boldsymbol{c}_M\right))$

(29) If not, and $\boldsymbol{c}_M$ is overinsured, then and since Ls IC constraint does nto bind, M can be moved toward full insurance slightly at a profit and utility gain since $MRS\left(q_M, \boldsymbol{c}_M\right) > \frac{1-p_M}{p_M}$ for overinsured $\boldsymbol{c}_M$

will attract only high types and also earn negative profits. With $r > MRS(q_M, \boldsymbol{c}_M)$ no one will demand the contract.

Then if I consider deviations of the form $(1, r)$ from $\boldsymbol{c}_M$, if $r > -\frac{1-p_M}{p_M}$ then the contract will make negative profit if it attracts medium types, and hence also if it attracts medium and high or just high types. If $r < -\frac{1-p_M}{p_M}$ in particular $r < -MRS(q_M, \boldsymbol{c}_M) < MRS(q_H, \boldsymbol{c}_H)$ for small ξ and so no medium or high types will swap to it, and for small enough deviations neither will low types. This shows that there are no local deviations from $\boldsymbol{c}_M$. Similar working rules out local deviations from $\boldsymbol{c}_L$. [30]

From this I can conclude that the proposed menu is a locally competitive equilibrium , and from step 2 that there are no others. This concludes the proof.

■

B.2 Proof of Proposition 2

Proof. **Step 1. High-type pessimists are separated.** Recall the fact from the proof of proposition 1: if $q_i > q_j$ then in equilibrium the former type receives weakly more insurance than the latter $\boldsymbol{c}_i \succsim \boldsymbol{c}_j$. Since $q_{H,P} > q_L > q_{H,O}$ immediately we have that either the high-type pessimists separated, are pooled with the low types, or all three are pooled together.

In the case all three are pooled together at a zero-profit making contract, a cream-skimming deviation exists in which a contract with slightly less insurance is offered that attracts only the low types and the high-type optimists. Since the pooling contract made zero profit, and the deviating contract excludes the high-type pessimists, it makes a positive profit. Hence all three types cannot be pooled together.

Now consider the case in which the high type pessimists and the low types are pooled together, while the high-type optimists are separated. Either no incentive constraint binds, or the low type has to be indifferent between their pooled contract with the high type pessimists and the contract of the high type optimists. To see this, suppose that the pessimistic high

[30] This all assumes that $\boldsymbol{c}_M$ features underinsurance, which is true for small enough ξ.

types were indifferent between the two contracts, then the low types must strictly prefer the optimistic high types' contract (that offers less insurance), a contradiction. Suppose the high type optimists were indifferent between their contract and the pooled contract offered to the high type pessimists and the low types. Then a pooled contract that offered slightly more insurance would be strictly preferred by the high type pessimists and the low types, and hence could make a profit. This establishes that the incentive constraint that binds, if any, is that of the low types. It follows that a regular cream-skimming contract that splits apart the pooled types and just attracts the low-type pessimists is a profitable deviation.

This establishes that the high-type pessimists are separated.

Step 2. High-type pessimists receive their (subjectively) preferred separating contract. The only way this won't occur is if one of the other two types IC constraints prevent the high-type pessimists from receiving full or over-insurance.

Suppose the other two types are pooled together. The low types IC constraint must then upwards. But then the pooling contract could move toward full insurance. The only way that deviation is not possible is if the pooling contract is already at full insurance. But then the high-type pessimist could themselves receive (arbitrarily) close to their own actuarially own full insurance, which is their preference.

Step 3. The low types and high-type optimists are pooled Given step 2, the high-type pessimists receive their subjectively preferred separating contract, which is infinitesimally more than full insurance. THis means for the high-type optimists to be separated from the low types, they must be receiving some contract on the same zero profit line as the high-type pessimists, which the high-type pessimists are therefore not indifferent to. Next, the low type's IC constraint must bind (if not and the high-type pessimist was indifferent between their contract and the low types, the latter of which offered more insurance, then so would the low types (who think they are high risk) prefer it, a contradiction). Hence, there a deviation near the high-type optimists separating contract that attracts the high-type optimists and the low types (not the high-type pessimists, who are receiving their first-best on the same zero profit line) and therefore makes a profit since it attracts the low types.

Step 4. The high-type pessimists IC constraint binds. If no IC constraint binds, then by the assumption that $MRS(q_L, c_{HL}) < \frac{1-p_{HL}}{p_{HL}}$, the low types would prefer to move toward full insurance. Can the low types IC constraint bind? No, since if they are indifferent between their contract and the high-type pessimists, deviation near the latter than also attracts the low types will make a profit.

Step 5. The proposed allocation is a local equilibrium. The prior steps imply that only the proposed allocation can possibly be a local equilibrium. It remains to check there are no profitable local deviations. Since the high-type pessimists are receiving their first-best and their IC constraint binds, no profitable deviation can attract them or any of the other types.

Consider deviations near the pooling contract. Deviations to more insurance attract the low types and the high types, which are not profitable since $\alpha_{H,P} \geq \alpha_{H,O}$. Deviations to less insurance attract only the high-type optimists, who are more costly than the pool, and so these deviations are not profitable.

This completes the proof.

■

B.3 Proof of Proposition 3

Proof. First I suppose that the error made is a small upward error by the high types. Proposition 3 shows that the high types contract can be found as the solution to the constrained maximization problem:

$$\hat{V}_H(\boldsymbol{q}) \equiv \max_{\boldsymbol{c}} V(q_H, \boldsymbol{c}) \quad \text{s.t } \pi_H(p_h, \boldsymbol{c}) = 0.$$

By the envelope theorem, for the multiplier $\lambda > 0$ I have

$$\frac{\partial \hat{V}}{\partial q_H} = \frac{\partial V}{\partial q_H} + \lambda \frac{\partial g}{\partial q_H} = \frac{\partial V}{\partial q_H} = u(c_H^L) - u(c_H^{NL}).$$

In particular, at $q_H = p_H$ I have $c_H^L = c_H^{NL}$ and hence :

$$\frac{\partial \hat{V}}{\partial \xi}|_{q_H = p_H} \frac{\partial \hat{V}}{\partial \xi}|_{q_H = p_H} = 0.$$

For low types, I can write their equilibrium contract, for a small error, as the solution to the (degenerate) maximization program:

$$\hat{V}_L(\boldsymbol{q}) \equiv \max_{\boldsymbol{c}} V(q_L, \boldsymbol{c}) \quad \text{s.t } \pi_L(p_L, \boldsymbol{c}) = 0$$
$$\text{and } h_L \equiv V(q_H, \boldsymbol{c}_H) - V(q_H, \boldsymbol{c}) = 0.$$

This is degenerate in the sense that the constraints exactly pin down the solution, and hence the maximization is superfluous. Nevertheless, with multiplier $\lambda > 0$ on the IC constraint I immediately get that:

$$\frac{\partial \hat{V}_L}{\partial q_H} = u(c_H^L) - u(c_H^{NL}) - (u(c_L^L) - u(c_L^{NL})).$$

At $q_H = p_H$ I have $c_H^L = c_H^{NL}$ and $c_L^L < c_L^{NL}$ and so

$$\frac{\partial \hat{V}_L}{\partial q_H}|_{q_H = p_H} = \frac{\partial \hat{V}_L}{\partial \xi}|_{q_H = p_H} > 0.$$

It follows that for a small change ξ, I have

$$\frac{\partial W}{\partial \xi} > 0 \text{ when } \xi > 0, \text{ and } \frac{\partial W}{\partial \xi} < 0 \text{ when } \xi < 0,$$

which completes the proof.

■

B.4 Proof of Lemma 1

Proof. First note that the high type optimists derive the same experienced utility from their old contract C_H as their new pooling contract C'_{HL} since the latter is defined by the incentive constraint of the (unbiased) high types.

Second, the unbiased high type is indifferent between C_H, C'_{HL} and C_L by construction. Hence, in terms of experienced utility, the low type strictly prefers C_L to either of the former, as their risk is lower and their indifference curves steeper. It follows that they are strictly worse off in the pooling contract following the large misperception.

Finally, the unbiased low types move from C_L to C'_L. The former is defined by the indifference curve of the unbiased high types through C'_{HL}, the latter by the indifference curve of the downward-biased high type through C'_{HL}. Hence the latter is down and right of the former. Along $\pi_L = 0$ experienced utility is maximized at full insurance and monotonically increases down and right. Hence the unbiased low type is worse off as well.

In sum, experienced utility declines.

■

B.5 Proof of Lemma 2

Proof. The high type optimists, as they go from a small to a large error, move form a contract that the unbiased high types strictly disliked relative to c_H to a contract the unbiased high types are indifferent between. Hence, in experienced utility the biased high types are better off.

The welfare of the low types is ambiguous. When the small error is sufficiently small, then it is approaches no error, and so the negative conclusion from lemma 1 holds. But as $\alpha_{L,P}$ gets large compared to $\alpha_{H,O}$ the iso-profits $\pi_{HL} = 0$ and $\pi_L = 0$ converge. As such since the pooled contract moves up and left on $\pi_{HL} = 0$ compared to the $IC_{H,O}$ intersection with $\pi_{HL} = 0$ so is the pooled contract up and left on $\pi_L = 0$ relative to c_L as $\alpha_{L,P}$ gets large. Hence the experienced utility of both the biased and unbiased low types will increase after pooling occurs.

■

B.6 Proof of Propositions 4

First I require some preliminary lemmas.

Lemma 4. *In a pooling allocation stable to cream skimming, pools consist only of contiguous types in terms of q. Formally, for types i, j, k with $q_i > q_j > q_k$, if i and k are pooled together then so is j*

Proof. From the proof of proposition 1, types with higher perceived risk must receive contracts with weakly more insurance. This is regardless of which binding IC constraints, if any, define the allcocation. Hence, writing C_i for is contract, I must have $C_i \succsim C_j \succsim C_k$, where $\succsim$ is the partial order defined in the proof of proposition 1 (verbally: $C \succsim$ C' if C offers more consumption in the loss state of the world and less in the no-loss state of the world than C'). Since i and k are pooled I have $C_i = C_k$ and so I must have $C_j = C_i = C_k$.

■

Lemma 5. *Algorithm 5d produces a pooling allocation stable to cream-skimming.*

Proof. First note algorithm 5d, by construction, constructs pools only of contiguous types. Now, I prove the claim by induction.

Clearly any pool of size two produced by the algorithm on its first iteration is stable to cream-skimming. Suppose the pool consists of types with contiguous (in q) types q_1, q_2 and as usual the indices label that $q_1 > q_2$. The only cream-skimming deviation possible attracts q_2. But if the only way a pool of size two could be produced by the algorithm the $p_1 < p_2$. As such, the cream-skimming deviation attracts p_2, the higher risk. Since the pool allocation, by definition broke even with p_1 and p_2 buying, it loses money on just p_2.

For induction, suppose all pools produced in the first k iterations of the algorithm are stable to cream skimming. Now, consider the $k + 1$th iteration. If no new pooling occurs I am done. Else, suppose two pools from the kth iteration are pooled together, and the new pool has size n_k. Visualize this as the contiguous n_k types in the new pool, where the 'join' is where the two pools at the kth iteration are now joined at the $k+1$th.

$$q_1 \quad q_2 \quad \underbrace{\quad | \quad}_{\text{join}} q_3 \quad \cdots \quad q_{n_k-2} \quad q_{n_k-1} \quad q_{n_k}$$

Now, to show that the new pool produced is stable to cream skimming, consider any possible cream skimming deviation. By definition, if a cream-skimming deviation attracts q it also attracts all $q' < q$. So I label as 'split' the proposed cream skimming deviation.

$$q_1 \quad q_2 \quad \underbrace{\quad | \quad}_{\text{join}} q_3 \quad \cdots \quad q_{n_k-2} \quad \underbrace{\quad | \quad}_{\text{split}} q_{n_k-1} \quad q_{n_k}$$

Since the algorithm joined $q_1 \quad q_2$ to $q_3 \quad q_4 \quad \ldots \quad q_{n_k}$ it must be that the average risk of the former pool, which I write as p_{12}, is lower than the average risk of the latter, written as $p_{34\ldots n_k}$.

By induction, the pool $q_3 \quad q_4 \quad \ldots \quad q_{n_k}$ was stable. In particular, the cream skimming at the new 'split' can't have been profitable, and it must have been the case that the average risk to the left of the split, $p_{34\ldots(n_k-2)}$ is lower than to the right, $p_{(n_k-1)n_k}$.

Putting these together, I have that $p_{12} < p_{34\ldots n_k}$, and that $p_{34\ldots(n_k-2)} < p_{(n_k-1)n_k}$ from which it follows that $p_{12} < p_{(n_k-1)n_k}$. If the split occurred, the resulting group skimmed would have higher risk than those not skimmed $p_{(n_k-1)n_k} > p_{12\ldots(n_k-2)}$. Since I assumed the larger pool's contract made zero profits, this split must be unprofitable.

It follows that any pooling formed at iteration $k+1$ are stable to cream skimming. By induction, the final pooling outputted also is.

∎

Now I am able to prove the main proposition.

Proof. Suppose, for a contradiction, that there is an alternate pooling, not generated by algorithm 5d, that is stable to cream skimming. By lemma 4, the pooling must still be contiguous. I visualize the pooling by the algorithm through square brackets $[\cdot]$ and through the alternate with regular $(\cdot)$. There are three options.

First, if the pools in the alternate are strict subsets of the algorithim's pools, for example:

$$\Big[(q_1 \quad q_2 \quad q_3)\,(q_4 \quad q_5) \Big]$$

Suppose the average risk in the left pool p_{123} is lower than the right, in which case the alternate is an earlier iteration of the algorithm, since the algorithm applied to $(q_1 \quad q_2 \quad q_3)\,(q_4 \quad q_5)$ would pool them into $[q_1 \quad q_2 \quad q_3 \quad q_4 \quad q_5]$. Suppose the average risk in the left pool is higher than the right. Then cream-skimming the right pool would be profitable, and the algorithm's pool $[q_1 \quad q_2 \quad q_3 \quad q_4 \quad q_5]$ can't have been stable. In either case, I have a contradiction.

Second, if the pools in the alternate are strict supersets of the algorithm's pools, for example:

$$\Big([q_1 \quad q_2 \quad q_3]\,[q_4 \quad q_5] \Big)$$

In this case the alternate cannot be stable, since a split between q_3 and q_4 must be profitable, else the algorithm would have combined it's two pools.

Third, if the pools from the alternate are distinct but overlap with those from the algorithm. For example:

$$q_1 \Big(q_2 \Big] \Big[q_3 \quad q_4 \Big) \quad q_5$$

Consider, in the alternate pooling, a split between q_2 and q_3. Consider the average risk of those cream skimmed: p_{34}.

Without loss of generality, I assume the algorithms pool is a weak superset of those cream-skimmed (q_3 and q_4) (since if not, just run this logic on the algorithms pool to the right). So the averge risk of those cream-skimmed is weakly lower than cost of the algorithms pool that contains them (since if not, one could cream-skim the algorithms pool between q_4

and q_5).

Similarly, the average risk of the post-cream skimming alternate pool containing q_2 must be weakly higher than the algorithms pool containing q_2 If the algorithms pool is a weak superset, this is immediate, else the algorithms pool would have been cream-skimmed at the boundary of the alternate (ie left of q_1). If the algorithms pool is a subset, then it is lower risk than all to the left (else it would have been pooled) and I conclude similarly.

So I have that the post-split alternate pool containing q_3 has weakly lower average risk than the algorithms pool containing q_3. Similarly the post-split alternate pool containing q_2 has higher average risk than the algorithm's pool containing q_2. But the algorithm's pool containing q_2 has higher average risk than the algorithm's pool containing q_3, else it would have been pooled. It follows that that proposed cream-skimming split in the alternate pooling is profitable. Hence the alternate pooling cannot have been stable. This shows that only the algorithm can produce a pooling stable to cream skimming.

■

B.7 Proof of Propositions 5 and 6

Proof. First, I begin with IC_H. Suppose IC_H doesn't bind. If the optimum is pooling then this is an immediate contradiction. So the optimum must seperate the types. Then IC_L must bind else one type can be made better off holding cross-subsidization fixed (since both cannot be at their tangency point, by the assumption of small errors, as then both individuals tangency points are arbitrarily close to full insurance and one will be strictly preferred to the other by both types).

If $\boldsymbol{c}_L \succ \boldsymbol{c}_H$ I have an immediate contradiction as the indifference curve of the high type is strictly shallower than the low type, so if $V_L(\boldsymbol{c}_L, q_L) = V_L(\boldsymbol{c}_L, q_L)$ then it must be that $V_H(\boldsymbol{c}_L, q_H) > V_L(\boldsymbol{c}_H, q_H)$, contracting IC_H. On the other hand, if $\boldsymbol{c}_H \succ \boldsymbol{c}_L$ then, since $(\boldsymbol{c}_H, \boldsymbol{c}_L) \in A$ and the contracts are seperate, the low type must receive less than full insurance, $c_L^{NL} > c_L^L$. But, fixing a level of cross-subsidization, their optimal contract from the planner's perspective is full insurance. They can then be made better off by moving along the profit line $\pi_L = -\chi$ toward full insurance. That IC_H will continue to be slack for small enough movements, giving a contradiction to welfare optimality. Hence IC_H must bind. It immediately follows that unless there is pooling, IC_L does not bind.

Then suppose there is a small error by the high type only. Whereas the no-error optimum for the high type featured $c_H^{NL} = c_H^L$, now slight over-insurance $c_H^{NL} + \epsilon = c_H^L$ for some small $\epsilon > 0$ is feasible, according to the constraints A. Holding constant the level of cross-subsidazation, the proof of Proposition 3 shows that this slight over-insurance for the high types and a correspondingly better contract for the low types owing to the relaxed IC_H strictly improves welfare. This proves the first welfare property (with respect to the high type misperceptions) and shows that part of requirement (A.6) binds, in that $MRS_H = \frac{1-p_H}{p_H}$.

Now if the low type makes a small error, that IC_H binds and IC_L doesn't immediately demonstrates the welfare property with respect to misperceptions by the low type.

It remains to establish the cross-subsidization results. First, the comparative static results for $c_H^{NL}(\chi), c_H^L(\chi), c_L^{NL}(\chi), c_L^L(\chi)$ due to the implicit function theorem in the proof of proposition 10 hold here. So I calculate

$\frac{\partial}{\partial \chi} Welfare$,, substitute in comparative statics (B.56) through (B.63), and then differientiate with respect to q_H to get the cross partial, and then evaluate at $\boldsymbol{q} = \boldsymbol{p}$ to study small errors. The resulting expression is

$$\frac{(p_L-1)p_L(\rho-1)\left(u'(c_L^L)-u'(c_L^{NL})\right)\left((a_H-1)u'(c)\left((p_L-1)u'(c_L^L)-p_Lu'(c_L^{NL})\right)+a_Hu'(c_L^L)u'(c_L^{NL})\right)}{(a_H-1)\left(p_H(p_L-1)u'(c_L^L)-(p+H-1)p_Lu'(c_L^{NL})\right)^2}.$$

The denominator is negative as $a_H < 1$. The numerator is positive and hence the cross partial is negative. That means the return to a little more cross-subsidization decreases in q_H. The claims follow. ■

B.8 Proof of Lemma 3

Proof. Similarly to the final part of the proof above, I substitute into the welfare function the comparative static results for $c_H^{NL}(\chi), c_H^L(\chi), c_L^{NL}(\chi), c_L^L(\chi)$ from (B.56) through (B.63), and then evaluate $\frac{\partial^2}{\partial\rho\partial\chi} Welfare \mid_{\boldsymbol{q}=\boldsymbol{p}}$.

This yields

$$\frac{(p_H - p_L)\left((a_H - 1)u'(c)\left((p_L - 1)u'(c_L^L) - p_L u'(c_L^{NL})\right) + a_H u'(c_L^L)u'(c_L^{NL})\right)}{(a_H - 1)\left(p_H(p_L - 1)u'(c_L^L) - (p_H - 1)p_L u'(c_L^{NL})\right)}.$$

The denominator is easily seen to be positive, and the numerator as well, making the whole expression positive. This shows that the return to cross-subsidaztion increases in ρ. The results follow.

■

B.9 Proof of Proposition 8

First I calculate the change in welfare with respect to a small high type error. The process for a low type is similar and not explicated at length.

Proof. The three binding constraints are respectively the tangency constraint for the high type, the incentive constraint for the high type, and the participation constraint for the low type:

$$\frac{(1 - q_H)u'(c_H^{NL})}{q_H u'(c_H^L(c_H^{NL}))} = \frac{1 - p_H}{p_H} \tag{B.1}$$

$$\text{IC} = q_H u(c_H^L(c_H^{NL})) + (1 - q_H)u(c_H^{NL}) = q_H u(c_L^L(c_H^{NL})) + (1 - q_H)u(c_L^{NL}(c_H^{NL})) \tag{B.2}$$

$$\text{IR} = q_L u(c_L^L(c_H^{NL})) + (1 - q_L)u(c_L^{NL}(c_H^{NL})) = q_L u(w - l) + (1 - q_L)u(w). \tag{B.3}$$

Given these, I can implicitly express each of c_H^L, c_L^L, c_L^{NL} in terms of c_H^{NL} in which case the monopolists maximization problem is solely a function of c_H^{NL}. As such, from the implicit function theorem I have

$$\frac{\partial c_H^{NL}}{\partial qH} = -\frac{\partial^2 \Pi / \partial qH \partial c_H^{NL}}{\partial^2 \Pi / \partial^2 c_H^L}.$$

Recalling that

$$\Pi(c_H^{NL}) = \alpha_H \left[(1-p_H)\left(W - c_H^{NL}\right) - p_H(c_H^L(c_H^{NL}) - (W-l))\right] + \alpha_H \left[(1-p_L)\left(W - c_L^{NL}(c_H^{NL})\right) - p_L(c_L^L(c_H^{NL}$$

differentiating I have

$$\frac{\partial^2 \Pi(c_H^{NL})}{\partial qH \partial c_H^{NL}} = \alpha_H \left[p_H \frac{\partial^2}{\partial qH \partial c_H^{NL}}(c_H^L(c_H^{NL}))\right]$$
$$+\alpha_H \left[(1-p_L)\left(\frac{\partial^2}{\partial qH \partial c_H^{NL}} c_L^{NL}(c_H^{NL})\right) - p_L(\frac{\partial^2}{\partial qH \partial c_H^{NL}} c_L^L(c_H^{NL})))\right]$$

and

$$\frac{\partial^2 \Pi(c_H^{NL})}{\partial (c_H^{NL})^2} = \alpha_H \left[p_H \frac{\partial^2}{\partial (c_H^{NL})^2}(c_H^L(c_H^{NL}))\right] + \alpha_H \left[(1-p_L)\left(\frac{\partial^2}{\partial (c_H^{NL})^2} c_L^{NL}(c_H^{NL})\right) - p_L(\frac{\partial^2}{\partial (c_H^{NL})^2} c_L^L(c_H^{NL})))\right]$$

To calculate $\frac{\partial c_H^L(c_H^{NL})}{\partial c_H^{NL}}, \frac{\partial c_L^L(c_H^{NL})}{\partial c_H^{NL}}, \frac{\partial c_L^{NL}(c_H^{NL})}{\partial c_H^{NL}}$ differientiate each of the binding constraints with respect to c_H^{NL} to get respectively:

$$0 = \frac{(q_H - 1)\left(u''(c_H^{NL})u'(c_H^L(c_H^{NL})) - (c_H^L)'(c_H^{NL})u'(c_H^{NL})u''(c_H^L(c_H^{NL}))\right)}{q_H u'(c_H^L(c_H^{NL}))} \tag{B.4}$$

$$q_H(c_H^L)'(c_H^{NL})u'(c_H^L(c_H^{NL})) + (1-q_H)u'(c_H^{NL}) = q_H(c_L^L)'(c_H^{NL})u'(c_L^L(c_H^{NL})) + (1-q_H)(c_L^{NL})'(c_H^{NL})u'(c_L^{NL}(c_H^{NL})) \tag{B.5}$$

$$0 = q_L(c_L^L)'(c_H^{NL})u'(c_L^L(c_H^{NL})) + (1-q_L)(c_L^{NL})'(c_H^{NL})u'(c_L^{NL}(c_H^{NL})) \tag{B.6}$$

Solving this system yields

$$\frac{\partial c_L^L(c_H^{NL})}{\partial c_H^{NL}} = -\frac{(q_L-1)\left((q_H-1)u'(c_H^{NL}) - \frac{q_H u''(c_H^{NL})u'(c_H^L(c_H^{NL}))^2}{u'(c_H^{NL})u''(c_H^L(c_H^{NL}))}\right)}{(q_L-q_H)u'(c_L^L(c_H^{NL}))} \tag{B.7}$$

$$\frac{\partial c_L^{NL}(c_H^{NL})}{\partial c_H^{NL}} = -\frac{q_L\left((q_H-1)u'(c_H^{NL}) - \frac{q_H u''(c_H^{NL})u'(c_H^L(c_H^{NL}))^2}{u'(c_H^{NL})u''(c_H^L(c_H^{NL}))}\right)}{(q_L-q_H)u'(c_L^{NL}(c_H^{NL}))} \tag{B.8}$$

$$\frac{\partial c_H^L(c_H^{NL})}{\partial c_H^{NL}} = \frac{u''(c_H^{NL})u'(c_H^L(c_H^{NL}))}{u'(c_H^{NL})u''(c_H^L(c_H^{NL}))} \tag{B.9}$$

To get the cross partials, differientiate with respect to q_H:

$$\frac{\partial^2}{\partial qH\partial c_H^{NL}}c_L^L(c_H^{NL})) = \frac{(q_L-1)\left(q_L u''(c_H^{NL})u'(c_H^L(c_H^{NL}))^2 - (q_L-1)u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))\right)}{(q_H-q_L)^2u'(c_H^{NL})u''(c_H^L(c_H^{NL}))u'(c_L^L(c_H^{NL}))} \tag{B.10}$$

$$\frac{\partial^2}{\partial qH\partial c_H^{NL}}c_L^{NL}(c_H^{NL})) = \frac{q_L\left(q_L u''(c_H^{NL})u'(c_H^L(c_H^{NL}))^2 - (q_L-1)u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))\right)}{(q_H-q_L)^2u'(c_H^{NL})u''(c_H^L(c_H^{NL}))u'(c_L^{NL}(c_H^{NL}))} \tag{B.11}$$

$$\frac{\partial^2}{\partial qH\partial c_H^{NL}}c_H^L(c_H^{NL})) = 0. \tag{B.12}$$

Hence I have

$$\frac{\partial^2\Pi(c_H^{NL})}{\partial q_H\partial c_H^{NL}} = \frac{\alpha_L\left(q_L u''(c_H^{NL})u'(c_H^L(c_H^{NL}))^2 - (q_L-1)u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))\right)}{(q_H-q_L)^2u'(c_H^{NL})u''(c_H^L(c_H^{NL}))u'(c_L^L(c_H^{NL}))u'(c_L^{NL}(c_H^{NL}))}$$
$$\times\left((p_L-1)q_L u'(c_L^L(c_H^{NL})) - p_L(q_L-1)u'(c_L^{NL}(c_H^{NL}))\right)$$

On the other hand, differentiating (B.7) with respect to c_H^{NL} again, and using the values from $(B.7)$ yields

$$\frac{\partial^2}{\partial(c_H^{NL})^2}c_L^L(c_H^{NL})) = \Big[(q_L-1)((1-q_L)u''(c_H^L(c_H^{NL}))u''(c_L^L(c_H^{NL}))(q_H u''(c_H^{NL})u'(c_H^L(c_H^{NL}))^2 \tag{B.13}$$

$$-(q_H-1)u'(c_H^{NL})^2u''(c_H^L(c_H^{NL})))^2 \tag{B.14}$$

$$-(q_H-q_L)u'(c_L^L(c_H^{NL}))^2(-q_H u^{(3)}(c_H^L(c_H^{NL}))u''(c_H^{NL})^2u'(c_H^L(c_H^{NL}))^3 \tag{B.15}$$

$$+q_H u'(c_H^L(c_H^{NL}))^2u''(c_H^L(c_H^{NL}))^2(u^{(3)}(c_H^{NL})u'(c_H^{NL})+u''(c_H^{NL})^2) \tag{B.16}$$

$$-(q_H-1)u'(c_H^{NL})^2u''(c_H^{NL})u''(c_H^L(c_H^{NL}))^3))\Big] \tag{B.17}$$

$$/\left[(q_H-q_L)^2u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3u'(c_L^L(c_H^{NL}))^3\right] \tag{B.18}$$

$$\frac{\partial^2}{\partial(c_H^{NL})^2}c_L^{NL}(c_H^{NL})) = \Big[-q_L((q_H-q_L)u'(c_L^{NL}(c_H^{NL}))^2(-q_H u^{(3)}(c_H^L(c_H^{NL}))u''(c_H^{NL})^2u'(c_H^L(c_H^{NL}))^3 \tag{B.19}$$

$$+q_H u'(c_H^L(c_H^{NL}))^2u''(c_H^L(c_H^{NL}))^2(u^{(3)}(c_H^{NL})u'(c_H^{NL})+u''(c_H^{NL})^2) \tag{B.20}$$

$$-(q_H-1)u'(c_H^{NL})^2u''(c_H^{NL})u''(c_H^L(c_H^{NL}))^3) \tag{B.21}$$

$$+q_L u''(c_H^L(c_H^{NL}))u''(c_L^{NL}(c_H^{NL}))(q_H u''(c_H^{NL})u'(c_H^L(c_H^{NL}))^2-(q_H-1)u'(c_H^{NL})^2u''(c_H^L(c_H^{NL})) \tag{B.22}$$

$$/\left[(q_H-q_L)^2u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3u'(c_L^{NL}(c_H^{NL}))^3\right] \tag{B.23}$$

$$\frac{\partial^2}{\partial(c_H^{NL})^2}c_H^L(c_H^{NL})) = \frac{u'(c_H^L(c_H^{NL}))(u^{(3)}(c_H^{NL})u'(c_H^{NL})u''(c_H^L(c_H^{NL}))^2-u^{(3)}(c_H^L(c_H^{NL}))u''(c_H^{NL})^2u'(c_H^L(c_H^{NL})))}{u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3}. \tag{B.24}$$

And so the second order total derivative of profit is

$$\frac{\partial^2\Pi(c_H^{NL})}{\partial(c_H^{NL})^2} = \frac{\gamma \mathrm{H}(p_H-1)u'(c_H^L(c_H^{NL}))(u^{(3)}(c_H^{NL})u'(c_H^{NL})u''(c_H^L(c_H^{NL}))^2 - u^{(3)}(c_H^L(c_H^{NL}))u''(c_H^{NL})^2u'(c_H^L(c_H^{NL})))}{u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3} \tag{B.25}$$

$$- \frac{(1-\alpha_H)p_L(q_L-1)(1-q_L)u''(c_H^L(c_H^{NL}))u''(c_L^L(c_H^{NL}))(q_Hu''(c_H^{NL})u'(c_H^L(c_H^{NL}))^2}{(q_H-q_L)^2u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3u'(c_L^L(c_H^{NL}))^3} \tag{B.26}$$

$$- \frac{-(q_H-1)u'(c_H^{NL})^2u''(c_H^L(c_H^{NL})))^2}{(q_H-q_L)^2u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3u'(c_L^L(c_H^{NL}))^3} \tag{B.27}$$

$$- \frac{(q_H-q_L)u'(c_L^L(c_H^{NL}))^2(-q_Hu^{(3)}(c_H^L(c_H^{NL}))u''(c_H^{NL})^2u'(c_H^L(c_H^{NL}))^3}{u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3} \tag{B.28}$$

$$- \frac{q_Hu'(c_H^L(c_H^{NL}))^2u''(c_H^L(c_H^{NL}))^2(u^{(3)}(c_H^{NL})u'(c_H^{NL})}{u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3} \tag{B.29}$$

$$+ \frac{u''(c_H^{NL})^2) - (q_H-1)u'(c_H^{NL})^2u''(c_H^{NL})u''(c_H^L(c_H^{NL}))^3))}{u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3} \tag{B.30}$$

$$+ \frac{(1-\alpha_H)(1-p_L)q_Lq_Lu''(c_H^L(c_H^{NL}))u''(c_L^{NL}(c_H^{NL}))(q_Hu''(c_H^{NL})u'(c_H^L(c_H^{NL}))^2}{(q_H-q_L)^2u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3u'(c_L^{NL}(c_H^{NL}))^3} \tag{B.31}$$

$$+ \frac{(q_H-1)u'(c_H^{NL})^2u''(c_H^L(c_H^{NL})))^2}{(q_H-q_L)^2u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3u'(c_L^{NL}(c_H^{NL}))^3} \tag{B.32}$$

$$+ \frac{(q_H-q_L)u'(c_L^{NL}(c_H^{NL}))^2 - q_Hu^{(3)}(c_H^L(c_H^{NL}))u''(c_H^{NL})^2u'(c_H^L(c_H^{NL}))^3}{u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3} \tag{B.33}$$

$$+ \frac{q_Hu'(c_H^L(c_H^{NL}))^2u''(c_H^L(c_H^{NL}))^2(u^{(3)}(c_H^{NL})u'(c_H^{NL})}{u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3} \tag{B.34}$$

$$+ \frac{u''(c_H^{NL})^2) - (q_H-1)u'(c_H^{NL})^2u''(c_H^{NL})u''(c_H^L(c_H^{NL}))^3)}{u'(c_H^{NL})^2u''(c_H^L(c_H^{NL}))^3}. \tag{B.35}$$

Now, knowing how a change in q_H affects the optimized value of c_H^{NL}, I calculate how this will affect the optimized values of c_H^L, c_L^L, c_L^{NL} . The optimized values solve the same constraints of course, now written as dependant on q_H. Differentiating them with respect to q_H and substituting the known expression for $\frac{\partial^2\Pi(c_H^{NL})}{\partial(c_H^{NL})^2}$ yields a system that implicitly defines $\frac{\partial(c_L^L)^*}{\partial q_H}, \frac{\partial(c_L^{NL})^*}{\partial q_H}, \frac{\partial(c_H^L)^*}{\partial q_H}$.

Solving these gives, evaluated now at $\boldsymbol{p} = qvec$:

$$\frac{\partial (c_L^L)^*}{\partial q_H} = \frac{(p_L - 1)}{(p_H - p_L)u'(c_L^L)} \tag{B.36}$$

$$\times\left(\frac{u'(c)^2 u'(c_L^L)^2 u'(c_L^{NL})^2 \left(u'(c_L^{NL}) - u'(c_L^L)\right)}{u'(c_L^{NL})^2 \left((p_H - p_L)u''(c)u'(c_L^L)^2 \left(u'(c_L^L) - u'(c_L^{NL})\right) - (p_L - 1)u'(c)^2 u''(c_L^L)u'(c_L^{NL})\right) + p_L u'(c)^2 u'(c_L^L)^3 u''(c_L^{NL})}\right. \tag{B.37}$$

$$\left. - \frac{u'(c)^2}{(p_H - 1)u''(c)} + u(c_L^L) - u(c_L^{NL})\right) \tag{B.38}$$

$$\frac{\partial (c_L^{NL})^*}{\partial q_H} = \frac{p_L}{(p_H - p_L)u'(c_L^{NL})}\left(u'(c)^2 u'(c_L^L)^2 u'(c_L^{NL})^2 \left(u'(c_L^{NL}) - u'(c_L^L)\right)\right) \tag{B.39}$$

$$/\left(u'(c_L^{NL})^2 \left((p_H - p_L)u''(c)u'(c_L^L)^2 \left(u'(c_L^L) - u'(c_L^{NL})\right) - (p_L - 1)u'(c)^2 u''(c_L^L)u'(c_L^{NL})\right)\right. \tag{B.40}$$

$$\left. + p_L u'(c)^2 u'(c_L^L)^3 u''(c_L^{NL}) - \frac{u'(c)^2}{(p_H - 1)u''(c)} + u(c_L^L) - u(c_L^{NL})\right) \tag{B.41}$$

$$\frac{\partial (c_H^L)^*}{\partial q_H} = \left((p_H^2 - p_L)u'(c)u''(c)u'(c_L^L)^2 u'(c_L^{NL})^2 \left(u'(c_L^L) - u'(c_L^{NL})\right) + u'(c)^3 \left[p_L u'(c_L^L)^3 u''(c_L^{NL})\right.\right. \tag{B.42}$$

$$\left.\left. - (p_L - 1)u''(c_L^L)u'(c_L^{NL})^3\right]\right) \tag{B.43}$$

$$/\left((p_H - 1)p_H u''(c)(u'(c_L^{NL})^2((p_H - p_L)u''(c)u'(c_L^L)^2 \left(u'(c_L^L) - u'(c_L^{NL})\right)\right. \tag{B.44}$$

$$\left. - (p_L - 1)u'(c)^2 u''(c_L^L)u'(c_L^{NL})) + p_L u'(c)^2 u'(c_L^L)^3 u''(c_L^{NL}))\right). \tag{B.45}$$

Now I can evaluate the change in welfare at the new optimum versus the old optimum: Directly from the constraints I have

$$\frac{\partial W_L^*}{\partial q_H} \mid_{\boldsymbol{q}=\boldsymbol{p}} = 0.$$

For the high type:

$$\frac{\partial W_H^*}{\partial q_H}\,|_{\boldsymbol{q}=pvec} = \alpha_H p_H (c_H^L)'(q_H)u'(c_H^L) + \alpha_H(1-p_H)(c_H^{NL})'(q_H)u'(c_H^{NL}) \tag{B.46}$$

$$= \Bigg(\alpha_H(2p_H - p_L - 1)u'(c)^2u''(c)u'(c_L^L)^2u'(c_L^{NL})^2\left(u'(c_L^L) - u'(c_L^{NL})\right) \tag{B.47}$$

$$+ \alpha_H u'(c)^4\left(p_L u'(c_L^L)^3u''(c_L^{NL}) - (p_L-1)u''(c_L^L)u'(c_L^{NL})^3\right)\Bigg) \tag{B.48}$$

$$/\Bigg((p_H-1)u''(c)(u'(c_L^{NL})^2\Big[(p_H-p_L)u''(c)u'(c_L^L)^2\left(u'(c_L^L)-u'(c_L^{NL})\right) \tag{B.49}$$

$$-(p_L-1)u'(c)^2u''(c_L^L)u'(c_L^{NL}) + p_L u'(c)^2u'(c_L^L)^3u''(c_L^{NL})\Big]\Bigg) \tag{B.50}$$

The denominator can be seen to be negative, since $u'(c_L^L) - u'(c_L^{NL}) > 0$ and utility is concave. As such, the welfare change will be positive when, the numerator is negative. A sufficient condition for this is $2pH > 1 + pL > 0$, which proves the claim.

The calculations for a change in the low types error are completely analogous and omitted for brevity.

■

B.10 Proof of Proposition 9

Proof. First I show that each of the constraints must hold else the allocation cannot be an MW equilibrium.

First the budget constraint (A.18). If negative profit is earnt the contract will be withdrawn. If positive profit is made and the equilibrium is pooling, then consumption in each state of the world can be increased, and everyone will swap. If the equilibrium is seperating, then increase consumption for both contracts to make everyone better off whilst keeping the IC constraint binding. The old contracts will be withdrawn, the new contracts will be demanded by the appropriate types. Hence (A.18) must hold in the MW equilibrium.

Next, suppose that (A.19) does not hold. Write the level of profit earned from each type as

$\pi_H = \pi(p_H, \boldsymbol{c}_H), \pi_L = \pi(p_L, \boldsymbol{c}_L)$. First, suppose IC_H binds and the equilibrium is separating. This means IC_L does not bind. By assumption (A.19), this means that $\boldsymbol{c}_H$ is down and left of the tangency point of iso-profit line for π_H with the appropriate H indifference curve, Then suppose a firm deviated to offer the contract $\boldsymbol{c}' = \boldsymbol{c} - (\epsilon, \frac{\epsilon(1-p_H)}{p_H}$. The high types are better off, as the slope of their utility function is monotone along an iso-profit line with a maximum at the point of tangency, whilst the low types still buy their old contract, and total profit has not changed by construction. So no contracts are withdrawn, and this is a valid deviation. So (A.19) must hold. Suppose the equilibrium is pooling, then the deviation to each types optimal contract subject to the pooling profit iso-profit line still breaks even and makes the high type strictly better off, and the low type weakly better off. All types change, so no contracts are withdrawn, so this is a valid deviation. Hence (A.19) must hold here by contradiction.

Suppose IC_L binds and the equilibrium is seperating. As $\boldsymbol{c}_H$ is down and right of the optimal insurance point on the π_H isoprofit line, then $MRS(q_H, \boldsymbol{c}_H) < \frac{1-p_H}{p_H}$ and also $MRS(q_H, \boldsymbol{c}_H) < MRS(q_L, \boldsymbol{c}_H)$ and so there exists a slope $\psi \in ([MRS(q_H, \boldsymbol{c}_H), \frac{1-p_H}{p_H}) \cap (MRS(q_H, \boldsymbol{c}_H) < MRS(q_L, \boldsymbol{c}_H))$ such that the contract $\boldsymbol{c} + \epsilon(-1, \xi)$ is preferred by H types to $\boldsymbol{c}_H$, not preferred by low types to $\boldsymbol{c}_L$ and costs less. This deviation shows the initial menu was not an MW equilibrium and hence (A.19) must hold.

Next, suppose that (A.17) does not hold, i.e. $MRS(q_H, \boldsymbol{c}_H) < V(q_H, \boldsymbol{c}_H^{RS})$. If a firm then offers a contract arbitrarily close (but earning positive profit) to $\boldsymbol{c}_H^{RS}$ all the high types will swap to it, and even if the low types prefer (perhaps after other contracts are withdrawn) it then profit will be even higher. So this is a valid deviation and hence (A.17) must hold.

Finally, suppose that (A.15) does not hold. Then as in RS the low types can be made better off with positive profit without inducing the high types to swap. So it must be that (A.15).

Now I show that the maximization has a unique solution. After which I will show that the MW equilibrium is exactly that unique solution. The three binding constraints mean that this is reducible to a one-dimensional problem. In particular, I can index the problem by the cross-subsidy χ. The constraint (A.17) implies $\chi \geq 0$. Moreover, an upper bound on feasible χ is the maximum that which induces pooling: $\chi_1 = (p_H - P_L)l$ since for any higher χ the incentive constraint for the high type cannot possibly bind, or that which hits the indemnity constraint for the low type (A.19), χ_2. Write $\overline{\chi} = \max\{\chi_1, \chi_2\}$. Moreover, a choice of $\chi \in [0, \overline{\chi}]$ uniquely defines $\boldsymbol{c}_H$ and $\boldsymbol{c}_L$. To see this: once a χ is chosen, the iso-profit for each type is pinned down. The high types iso-profit, together with (A.19), uniquely pins down $\boldsymbol{c}_H$ as the MRS is monotonic along an iso-profit line. Then (A.15) and the low types isoprofit line intersect at two places. One of them is up and left of $\boldsymbol{c}_H$ and violates (A.19). Hence only the intersection that offers the low type less insurance remains. So $\boldsymbol{c}_L$ is pinned down.

Given all this, the problem becomes:

$$\max_{\chi \in [0, \overline{\chi}]} V(q_L, \boldsymbol{c}_L(\chi)) \tag{B.51}$$

so that all the constraints hold with $\boldsymbol{c}_H, \boldsymbol{c}_L$ implicit functions of χ. The existence of a solution follows by Weierstrass' theorem. Uniqueness follows directly from the proof of uniqueness in Netzer and Scheuer (2014), the fact that errors are small and so the solutions to this problem are arbitrarily close to the problem without errors, and continuity of the second derivative. In particular, they showed that the second derivative is globally concave, which then holds here.

This shows that any MW equilibrium must satisfy the given constraints, and that a unique solution to the optimization problem exists. It remains to show that only the amongst the allocations that satisfy the constraints, only that which maximizes the objective is an MW equilibrium. Suppose not. Then suppose a firm does offer the menu that satisfies the constraints and maximizes the objective. By definition the low types will prefer to switch to the new contract. That contract makes weakly positive profit, by constraint (A.17). The high types will either remain in their old contract or swap to the newly offered contract for them. In the former case the deviating firm will make zero profits, in the latter they will make positive profits. No withdrawal of contracts will make the low types switch as their subjective utility is being maximized. This shows their is a MW deviation, and so the supposition was false. Hence the objective must be maximized. This concludes the proof.

■

B.11 Proof of Proposition 10

Proof. THe argument above establishes that to sign $\frac{\partial}{\partial q_H}\chi^*(q_H)$ when without errors it is the case that $\chi^* > 0$ the comparative statics with respect to q_H can be computed in the following way. (The result for q_L is entirely similar and omitted for brevity.

The following binding constraints (and a definition) implicitly define $c_H^{NL} = c_H^{NL}(\chi)$ and similarly for c_H^L, c_L^L, c_L^{NL}. :

$$\frac{u'((\chi))}{u'(c_H^L(\chi))} = \frac{(1-p_H)q_H}{p_H(1-q_H)} \tag{B.52}$$

$$q_H u(c_H^L(\chi)) + (1-q_H)u(c_H^{NL}(\chi)) = q_H u(c_L^L(\chi)) + (1-q_H)u(c_L^{NL}(\chi)) \tag{B.53}$$

$$w - l(p_H + p_L) = (1-\alpha_H)p_L c_L^L(\chi) + (1-\alpha_H)(1-p_L)c_L^{NL}(\chi) + \alpha_H(l(w-p_H)+\chi) \tag{B.54}$$

$$\chi = p_H c_H^L(\chi) + (1-p_H)c_H^{NL}(\chi) - (w - lp_H). \tag{B.55}$$

Differentiating each with respect to χ and solving yields

$$(c_H^{NL})'(\chi) = \frac{u''(c_H^L(\chi))u'(c_H^{NL}(\chi))}{p_H u'(c_H^L(\chi))u''(c_H^{NL}(\chi)) - (p_H-1)u''(c_H^L(\chi))u'(c_H^{NL}(\chi))} \tag{B.56}$$

$$(c_H^L)'(\chi) = \frac{u'(c_H^L(\chi))u''(c_H^{NL}(\chi))}{p_H u'(c_H^L(\chi))u''(c_H^{NL}(\chi)) - (p_H-1)u''(c_H^L(\chi))u'(c_H^{NL}(\chi))} \tag{B.57}$$

$$(c_L^L)'(\chi) = \Big[u'(c_H^L(\chi))u''(c_H^{NL}(\chi))\left(q_H(\alpha_H(-p_L)+\alpha_H+p_L-1)u'(c_H^L(\chi)) + \alpha_H p_H(q_H-1)u'(c_L^{NL}(\chi))\right) \tag{B.58}$$

$$+ \alpha_H(p_H(-q_H)+p_H+q_H-1)u''(c_H^L(\chi))u'(c_H^{NL}(\chi))u'(c_L^{NL}(\chi)) \tag{B.59}$$

$$+ (\alpha_H-1)(p_L-1)(q_H-1)u''(c_H^L(\chi))u'(c_H^{NL}(\chi))^2\Big] \tag{B.60}$$

$$/\Big[(\alpha_H-1)\left((p_H-1)u''(c_H^L(\chi))u'(c_H^{NL}(\chi)) - p_H u'(c_H^L(\chi))u''(c_H^{NL}(\chi))\right) \tag{B.61}$$

$$\times \left((p_L-1)q_H u'(c_L^L(\chi)) - p_L(q_H-1)u'(c_L^{NL}(\chi))\right)\Big] \tag{B.62}$$

$$(c_L^{NL})'(\chi) = \Big[u''(c_H^L(\chi))u'(c_H^{NL}(\chi))\left((\alpha_H-1)p_L(q_H-1)u'(c_H^{NL}(\chi)) - \alpha_H(p_H-1)q_H u'(c_L^L(\chi))\right) \tag{B.63}$$

$$+ q_H u'(c_H^L(\chi))u''(c_H^{NL}(\chi))\left((p_L-\alpha_H p_L)u'(c_H^L(\chi)) + \alpha_H p_H u'(c_L^L(\chi))\right)\Big] \tag{B.64}$$

$$= /\Big[(\alpha_H-1)\left((p_H-1)u''(c_H^L(\chi))u'(c_H^{NL}(\chi)) - p_H u'(c_H^L(\chi))u''(c_H^{NL}(\chi))\right) \tag{B.65}$$

$$\times \left((p_L-1)q_H u'(c_L^L(\chi)) - p_L(q_H-1)u'(c_L^{NL}(\chi))\right)\Big]. \tag{B.66}$$

Being overly explicit,the objective can be written:

$$V_L(q_L,\chi) = q_L u(c_L^L(\chi(q_H),q_H)) + (1-q_L)u(c_L^{NL}(\chi(q_H),q_H))$$

By the implicit function theorem, the sign of $\chi'(q_H)$ is the same as the sign of $\partial^2 V_L(q_L,\chi)/\partial\chi\partial q_L$. So computing:

$$\frac{\partial^2 V_L(q_L,\chi)}{\partial\chi\partial q_H} = q_L \frac{\partial c_L^L \chi}{\partial q_H} u'(c_L^L(\chi(q_H), q_H)) + (1-q_L)\frac{\partial c_L^{NL}(\chi(q_H), q_H)}{\partial q_H} u'(c_L^{NL}(\chi(q_H), q_H)) \tag{B.67}$$

$$+ q_L (c_L^L)'(q_H)(c_L^L)'(\chi)u''(c_L^L(\chi(q_H), q_H)) + (1-q_L)(c_L^{NL})'(q_H)(c_L^{NL})'(\chi)u''(c_L^{NL}(\chi(q_H), q_H)) \tag{B.68}$$

where $(c_L^{NL})'(q_H)$ is shorthand for the partial derivative with respect to q_H when χ is fixed (i.e. only the explicit dependence is taken into account, not the implicit dependence through $\chi(q_H)$. Such quantities can be found by considering each consumption quantity to be a function of q_H, not χ in (B.52) - (B.55), differentiating them all and solving. The solution to this is, suppressing dependence on q_H for brevity:

$$(c_H^{NL})'(q_H) = \frac{(p_H-1)u'(c_H^L(q_H))^2}{(q_H-1)^2\left((p_H-1)u''(c_H^L(q_H))u'(c_H^{NL}) - p_H u'(c_H^L)u''(c_H^{NL})\right)}$$

$$(c_H^L)'(q_H) = \frac{(p_H-1)^2 u'(c_H^L)^2}{p_H(q_H-1)^2\left((p_H-1)u''(c_H^L)u'(c_H^{NL}) - p_H u'(c_H^L)u''(c_H^{NL})\right)}$$

$$(c_L^L)'(q_H) = -\Big[(p_L-1)(p_H^2(q_H-1)^2 u'(c_H^L)u''(c_H^{NL})(u(c_H^L) - u(c_H^{NL}) - u(c_L^L) + u(c_L^{NL}))$$
$$- (p_H-1)p_H(q_H-1)^2 u''(c_H^L)u'(c_H^{NL})(u(c_H^L) - u(c_H^{NL}) - u(c_L^L) + u(c_L^{NL}))$$
$$+ (p_H-1)p_H(q_H-1)u'(c_H^L)^2 u'(c_H^{NL}) - (p_H-1)^2 q_H u'(c_H^L)^3)\Big]$$
$$/\Big[p_H(q_H-1)^2((p_L-1)q_H u'(c_L^L) - p_L(q_H-1)u'(c_L^{NL}))\left((p_H-1)u''(c_H^L)u'(c_H^{NL}) - p_H u'(c_H^L)u''(c_H^{N\ldots}\right.$$

$$(c_L^{NL})'(q_H) = \Big[p_L(-p_H^2(q_H-1)^2 u'(c_H^L)u''(c_H^{NL})(u(c_H^L) - u(c_H^{NL}) - u(c_L^L) + u(c_L^{NL}))$$
$$+ (p_H-1)p_H(q_H-1)^2 u''(c_H^L)u'(c_H^{NL})(u(c_H^L) - u(c_H^{NL}) - u(c_L^L) + u(c_L^{NL}))$$
$$+ p_H(p_H(-q_H) + p_H + q_H - 1)u'(c_H^L)^2 u'(c_H^{NL}) + (p_H-1)^2 q_H u'(c_H^L)^3)\Big]$$
$$/\Big[p_H(q_H-1)^2((p_L-1)q_H u'(c_L^L) - p_L(q_H-1)u'(c_L^{NL}))((p_H-1)u''(c_H^L)u'(c_H^{NL}) - p_H u'(c_H^L)u''(c_H^{NI\ldots}$$

The appropriate second order partials with respect to χ can be found by differientiating

(B.56)-(B.63) with respect to χ once more. Substituting these all in, evaluating at $\boldsymbol{q} = \boldsymbol{p}$ and hence at $c_H^{NL} = c_H^L = c$ yields

$$
\begin{aligned}
&\frac{\partial}{\partial q_H}\chi'(q_H) \mid_{\boldsymbol{q}=\boldsymbol{p}} = (p_L - 1)p_L((u(c_L^L) - u(c_L^{NL})) \\
&\times \left[\frac{(-(\alpha_H - 1)u'(c)((p_L - 1)u''(c_L^L) - p_L u''(c_L^{NL})) + \alpha_H(p_H - 1)u''(c_L^L)u'(c_L^{NL}) - \alpha_H p_H u'(c_L^L)u''(c_L^{NL}))}{(\alpha_H - 1)(p_H(p_L - 1)u'(c_L^L) - (p_H - 1)p_L u'(c_L^{NL}))^2} \right. \\
&+ \left. \frac{-(u'(c_L^L) - u'(c_L^{NL}))((\alpha_H - 1)u'(c)((p_L - 1)u'(c_L^L) - p_L u'(c_L^{NL})) + \alpha_H u'(c_L^L)u'(c_L^{NL})))}{(\alpha_H - 1)(p_H(p_L - 1)u'(c_L^L) - (p_H - 1)p_L u'(c_L^{NL}))^2} \right] \\
&< 0.
\end{aligned}
$$

The denominator is negative and the numerator is positive as $c_L^L < c_L^{NL}$, hence $u(c_L^L) < u(c_L^{NL})$ and $u'(c_L^L) > u'(c_L^{NL})$. This shows $\frac{\partial}{\partial q_H}\chi'(q_H) \mid_{\boldsymbol{q}=\boldsymbol{p}} < 0$. Similar working shows that for a small error by the low type:

$$
\frac{\partial}{\partial q_L}\chi'(q_L) \mid_{\boldsymbol{q}=\boldsymbol{p}} = \frac{p_L\left((\alpha_H - 1)u'(c)\left((p_L - 1)u'(c_L^L) - p_L u'(c_L^{NL})\right) + \alpha_H u'(c_L^L)u'(c_L^{NL})\right)}{(\alpha_H - 1)\left(p_H(p_L - 1)u'(c_L^L) - (p_H - 1)p_L u'(c_L^{NL})\right)} > 0.
$$

once one notices that $p_H > p_L, 1 - p_L > 1 - p_H$ and $u'(c_L^L) > u'(c_L^{NL})$ and hence the denominator is positive.

Lastly it remains to show that welfare increases in χ. Recall welfare is given by

$$
\alpha_H(p_H u(c_H^L(\chi)) + (1 - p_H)u(c_H^{NL}(\chi))) + (1 - \alpha_H)(p_L u(c_L^L(\chi)) + (1 - p_L)u(c_L^{NL}(\chi))).
$$

Differentiating with respect to χ and substituting in $W = (B.58)$ and $(B.63)$, and then evaluating at $\boldsymbol{p} = \boldsymbol{q}$, yields

$$\frac{\partial W}{\partial \chi} = \frac{u'(c)\left((p_L - 1)(\alpha_H p_H - \alpha_H p_L + p_L)u'(c_L^L) + p_L(-\alpha_H p_H + (\alpha_H - 1)p_L + 1)u'(c_L^{NL})\right)}{p_H(p_L - 1)u'(c_L^L) - (p_H - 1)p_L u'(c_L^{NL})} \tag{B.69}$$

$$+ \frac{\alpha_H(p_H - p_L)u'(c_L^L)u'(c_L^{NL})}{p_H(p_L - 1)u'(c_L^L) - (p_H - 1)p_L u'(c_L^{NL})}. \tag{B.70}$$

Note that the first order condition here reads:

$$p_L \frac{\partial c_L^L}{\partial \chi} + (1 - p_L)\frac{\partial c_L^{NL}}{\partial \chi} = 0,$$

and substituting this in and simplifying yields such that the change in welfare is just given by the effect on the high types utility:

$$\frac{\partial W}{\partial \chi} = \alpha_H u'(c) > 0.$$

■

C Empirical Appendix

C.1 Implementing the Beresteanu and Molinari (2008) and Manski and Molinari (2010) Algorithm

As an additional check that the patterns described above are robust to noise in individual's elicitations due to rounding, I implement the procedure from Manski and Molinari (2010) based on the theory in Beresteanu and Molinari (2008).

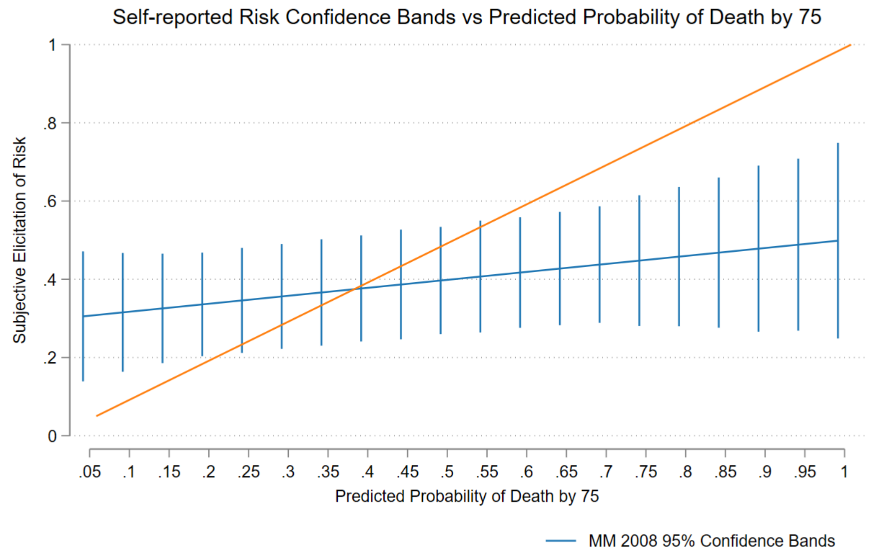
The procedure has multiple parts. First, given individual's point elicitation their degree of rounding is inferred. Next, using their degree of rounding an interval is constructed in which the econometrician is confident the true, de-rounded, belief lies. Then, using the asymptotic properties established in Beresteanu and Molinari (2008), valid inference can be performed on the relationship between (point-identified) predicted risk and (interval-identified) subjective beliefs about the risk.

Adapting the procedure of Manski and Molinari (2010) and section 3.4 I categorize individual's into rounding categories and assign them intervals into which their true de-rounded beliefs must fall as follows:

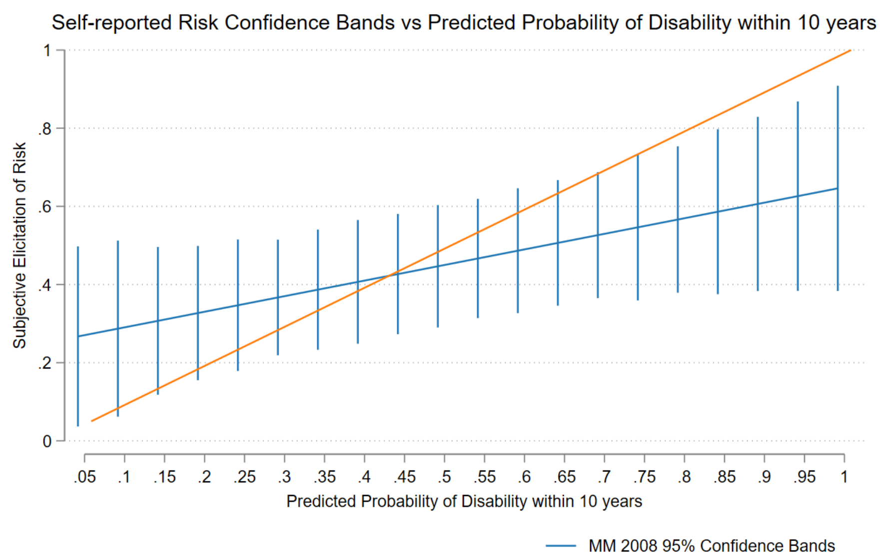
- If $q_i \in \{0, 50\%, 100\%\}$, then $\left[\underline{q}_i, \overline{q}_i\right] = [\max\{0, q_i - 0.25\}, \min\{1, q_i = 0.25\}]$
- If $q_i \in \{25\%, 75\%\}$, then $\left[\underline{q}_i, \overline{q}_i\right] = [\max\{0, q_i - 0.125\}, \min\{1, q_i = 0.125\}]$
- If $q_i \in \{10\%, 20\%, \ldots, 90\%\}$, then $\left[\underline{q}_i, \overline{q}_i\right] = [\max\{0, q_i - 0.05\}, \min\{1, q_i = 0.05\}]$
- If $q_i \in \{5\%, 15\%, \ldots, \%95\}$, then $\left[\underline{q}_i, \overline{q}_i\right] = [\max\{0, q_i - 0.025\}, \min\{1, q_i = 0.025\}]$
- For everyone else, I assume no rounding.

Then to study the relationship between de-rounded subjective risk and predicted objective risk the procedure is akin to interval regression. The regressor is point valued and the regressand is interval valued. 95% confidence intervals are constructed that are asymptotically valid per Beresteanu and Molinari (2008).

The results are shown graphically in figure 9.

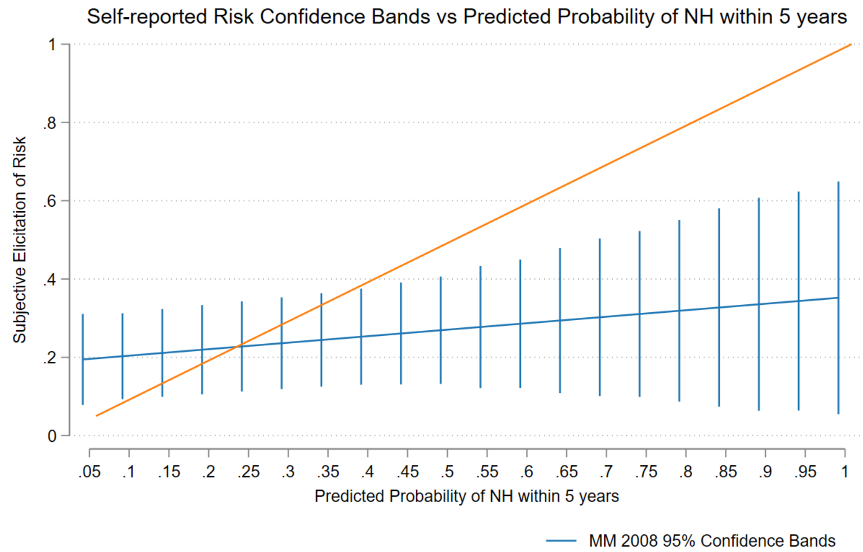


(a) **Mortality Risk**

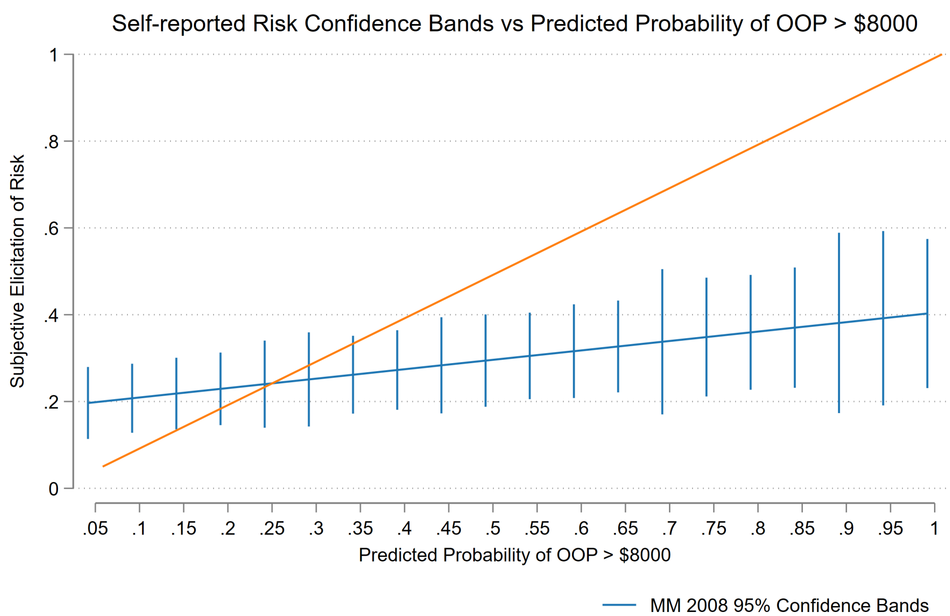


(b) **Disability Risk**

We see that the general patterns described above are robust to this accounting of rounding error. In the case of mortality and long-term care risk the ’S’ shape holds true. Disability risk is still entirely explicable by rounding error and hence I cannot reject the hypothesis of no true misperceptions. Catastrophic medical risk still exhibits, over the empirically observed range of objective risk, either no or upward misperception, suggesting a welfare increase that will be explored in section 5.



(c) **Long-term Care Risk**



(d) **Catastrophic Medical Risk**

Figure 9: Predicted probability of each risk versus confidence bands for subjective elicitation, per the algorithm of Beresteanu and Molinari (2008) and Manski and Molinari (2010)
Note: Only those in the training set are included. Individuals are sorted by their predicted probability, then divided into 20 evenly sized bins. In each bin the average self-report and average risk realization rate, with standard errors, are graphed.

C.2 Mortality Risk Evidence from Wave 1

A concern with the evidence presented above is that the rounding categories into which I have assigned individuals are incorrect. I have attempted to be conservative and err in the direction that would be adversarial to my results (by assigning individuals to coarser categories whenever in doubt). Nevertheless this concern might persist. To circumvent this concern, I offer evidence below that comes only from the responses in the first wave of the HRS.

In the first wave of the HRS, the same question was asked of survival to age 75, but the responses were constrained to be an integer from 0-10. These responses are then aligned with later HRS waves by treating the integer response of x as signifying a response of x in 10, or $10x$ percent[(31)]. So a response of 4 (out of 10) in wave one is equivalent to a response of 40% in any subsequent wave. Interpreted this way, the unit of rounding is clear. I assume that an individual considered all possibilities and rounded to the nearest 10%.

The same prediction and inference exercise is performed , except now restricted to the individuals for whom a prediction and subjective elicitation can be obtained in wave 1. This restricts the sample size considerably and there is more noise in the prediction and inference. The results, again with the two different specifications (raw and derounded elicitations) are presented in table 4.

[(31)]Study (2018)

	Mortality Risk	
Quantity	Raw q_i	Derounded $\tilde{q}_i$
$E(q_i \mid p_i = 0)$	0.26*** (0.01)	0.23*** (0.01)
β	0.33*** (0.03)	0.37*** (0.03)
$E(q_i \mid p_i = 1)$	0.59*** (0.02)	0.60*** (0.02)
Observations (Individuals)	2,690	1,467

Table 4: Results from estimating equations (4.1) and (4.2) restricted to wave 1. Notes: *** means significant at the 1% level against 0 (first row) or 1 (second and third rows). Only those in the training set are included.

The patterns are the same as table 2. In all specifications I observe pessimism at the bottom of the risk distribution, optimism at the top and attenuation as we move from the bottom to the top. I continue reject the hypothesis of constant misperception (in particular, of zero misperception).

C.3 Evidence from the Panel Data

The HRS is a panel data set that has been collected for almost 30 years. Up to this point I have pooled all observations without meaningfully using the panel structure, except when clustering standard errors. In this section I offer within-person evidence using the panel data to reinforce the conclusions so far. The questions such evidence can answer include: Is the within-person time variation in subjective perceptions concordant with the within-person time variation in predicted objective risk?

Introducing a time subscript, I model the relationship between objective risk and subjective perception as:

$$q_{it} = \alpha + \beta p_{it} + u_i + \epsilon_{it}. \tag{C.1}$$

I estimate (C.1) by time-demeaning within person, such that u_i drops out and β is identified using only within person variation. In particular, any time-invariant optimism or pessimism or unobservable factor drops out. It is very unlikely there is a true time invariant unobservable. Suppose an individual had some disease that was unobserved to the HRS. If at age 60 say this disease decreased the probability of survival by 10%, it is unlikely at age 62 the decreased probability of survival is still 10%, because the individual is now 2 years closer to 75. So the primary purpose of the fixed effect is to remove a hypothetical baseline optimism or pessimism, such that an individual always inflates or deflates their perception by a constant amount.

The first specification in table 5 is the fixed effect regression of (C.1) with raw elicitation q_i used. The second specification is again (C.1) with de-rounded $\tilde{q}_i$ used.

	Mortality Risk		Disability Risk		LTC Risk		Health Risk	
Quantity	Raw q_i	Derounded $\tilde{q}_i$	Raw q_i	Derounded $\tilde{q}_i$	Raw q_i	Derounded $\tilde{q}_i$	Raw q_i	Derounded $\tilde{q}_i$
β	0.10*** (0.03)	0.41*** (0.02)	0.62*** (0.03)	0.75*** (0.02)	0.10*** (0.01)	0.34*** (0.01)	0.24*** (0.05)	0.44*** (0.04)
Observations (Ind $\times$ wave)	5,252	5,252	7,184	7,184	44,773	44,773	8.904	8.904
Individuals	2,241	2,241	3,751	3,751	12,677	12,677	5,658	5,658

Table 5: Fixed Effect estimation of equation (C.1) by taking differences over 1 wave (2 years).
Note: *** means significant at the 1% level against 0 (first row) or 1 (second and third rows).

As before similar patterns obtain. Disability risk is quite well understood by individuals, with almost three quarters (after de-rounding) of within-person variation in objective risk updated into subjective elicitations. This speaks to the relative predictability of disability risk, as opposed to the rarer but more severe shocks that dramatically move health and mortality risk.

For the other three risks - mortality, long-term care and catastrophic health - table 5 shows that individuals do not appropriately update their risk perceptions. Even accounting for rounding bias, a 1% change in objective risk is reflected in an, at most, 0.41% change in subjective elicitation.

In combination with the main results in table 2, I conclude that the patterns of misperception are not simply due to baseline bias that is anchored to. Rather the attenuation pattern we see in the three risks except for catastrophic medical risk are due to within person attenuation. Whether individuals' risk changes favorably or unfavorably the full magnitude of this change isn't factored in, resulting in subjective perceptions that are bent toward the middle. Even though catastrophic medical risk exhibits pessimism throughout the risk distribution, the same within-person mechanism is likely the cause. There is baseline pessimism and since most movements are in the direction of lower out-of-pocket risk (either due to treatment or the upgrading of insurance), an underreaction to the lowered risk increases the pessimism.

C.4 Cognitive Foundations

I have shown that there is a strong covariance between risk type and risk misperception. The general pattern is that when there are high risk types, they under-perceive their risk, and when there are low-risk types they over-perceive this risk. This typically generates an 'S' shaped pattern. But with an important exception: when the risk in question is almost always low, such as catastrophic medical risk, then only the pessimism remains.

This empirical pattern is broadly consistent with the experimental literature on how individuals perceive probabilities. Beginning with Tversky and Kahneman (1992) and summarized in Enke and Graeber (2019), in many contexts people have held (or acted as if they held) subjective probabilities that are 'S' shaped. For low probabilities, the subjective probability is above the objective, while for high probabilities the reverse is true. This is true in the evaluation of lotteries, choice under ambiguity, Bayesian updating and more.

The main cognitive foundation consistent with this pattern is confusion causing movement toward the middle as a default. While the findings in this paper are independent of the precise cause of the misperception, it is worth noting some possibilities. Fischhoff and

Bruine De Bruin (1999) show that individuals who have no idea about the answer often respond with 50% when faced with a probability scale. If these confused individuals are spread throughout the risk distribution, this might mechanically generate this pattern.[32] A continuous version of this dynamic is posited by Enke and Graeber (2019) in which individuals choose something between their true belief and some cognitive default based on their degree of cognitive uncertainty. The more cognitively uncertain choose closer to the default, which might too be 50% on this scale.

Alternative cognitive causes might include insufficient updating away from a mean based on personal characteristics. This is consistent with the panel evidence shown in section C.3. Even as time passes, and these risks usually change due to aging or as medical conditions arise or resolve or as mechanically the number of years left until the age of 75 fall, individuals on average fail to update their subjective perceptions sufficiently.

This paper provides evidence for the same pattern in a new setting. In individual's subjective perceptions of their own mortality and their long-term care risk, the probabilities too are 'S' shaped: They are compressed toward the middle. Novelly, this is true in the present context of probabilities that are fundamentally personal or private information, not simply in a carefully controlled lab setting or in the context of a future piece of public information such as a macroeconomic statistic. It is reassuring though that the literature has broadly concluded that this pattern actually reflects distorted views of probabilities, not just noise induced by surveys.

C.5 Additional Figures

C.5.1 AUROC curves

Figure 10 exhibits the operator-receiver curve for the mortality predictions LASSO logit and classical logit make. Also included are two blunt predictions of mortality risk: 1) US

[32] The HRS asks a follow-up question to those who respond with 50%, querying whether they truly think it is equally likely or are just not sure. Excluding those who respond they are not sure does not change the empirical findings above.

population life tables that are differentiated only by age and sex, 2) Individuals subjective perception of their mortality risk. Clearly the logit and LASSO logit routines make substantial improvements over the subjective perceptions or the life tables. This rules out individuals fully understanding and correctly quantifying their own mortality risk. Secondarily, the LASSO logit routine does not do much better than classical logit. For computational ease and to minimize concerns about arbitrary variable selection by the econometrician, I stick with the former.

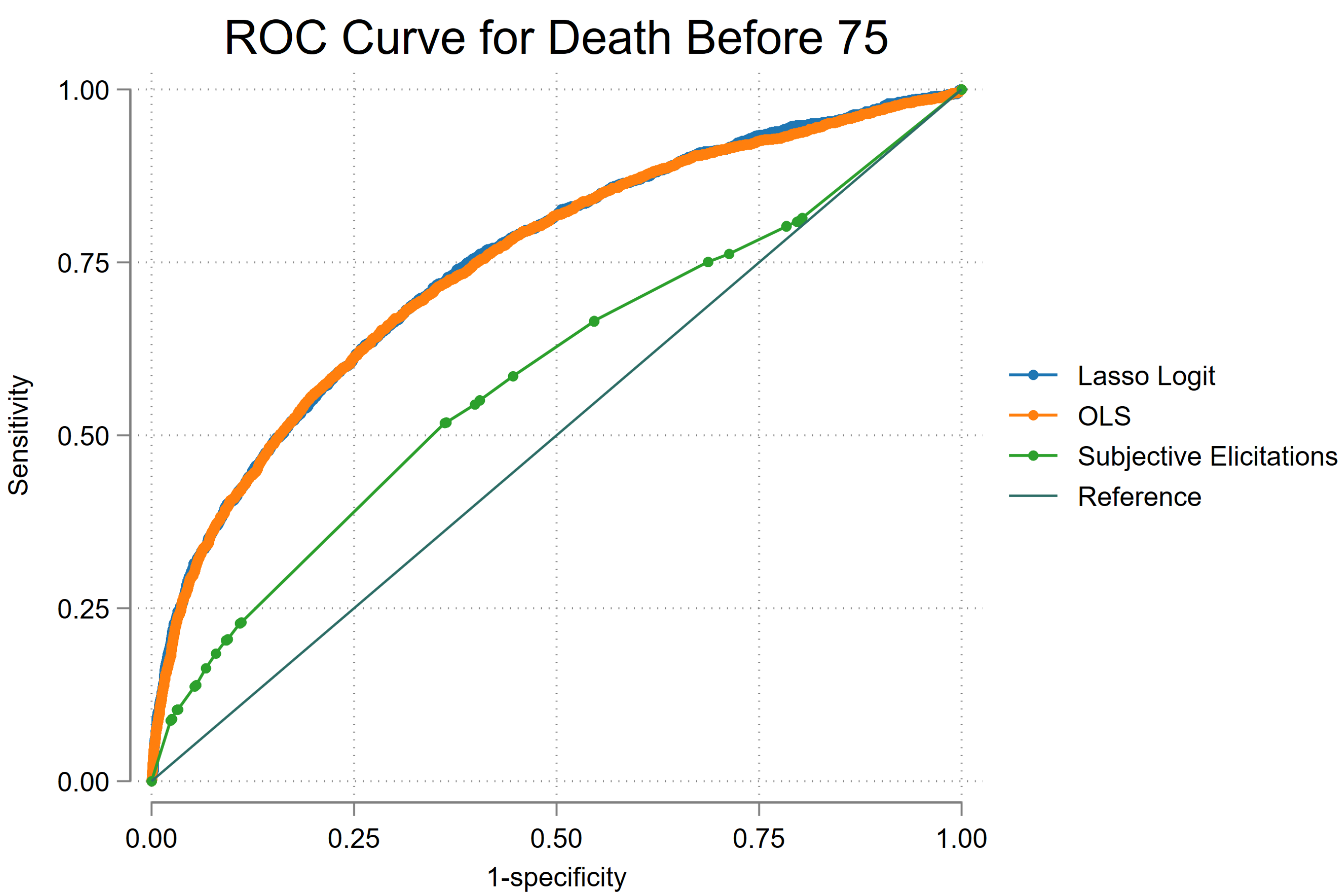


Figure 10: Receiver-operator curves for lasso logit, long logit, subjective elicitations (self-reports) and life tables predictions of mortality by age 75

C.5.2 Misperceptions among older versus younger annuitants

The simulations in section 5.1 show that older annuitants should be offered fewer contracts in equilibrium than younger annuitants. That is directly tested in actual annuity offerings in section 5.2. Here we demonstrate that the reason for this might be that older annuitants are making larger errors in risk perception than younger, and so for a given ω, there is more mixing of types, more pooling and hence fewer contracts offered.

The first key fact I establish is that older people have less information about their long-run longevity than younger people. Specifically, that 65-year-olds have less information about their longevity to 85 than 55-year-olds have about their longevity to 75. This is intuitive: middle age (55-75) mortality is more likely due to forseeable or partially foreseeable events such a chronic heart, lung or kidney diseases, cancer or diabetes. On the other hand, there are a suite of unpredictable medical problems (e.g. infection, flu, falls) that are more likely and more consequential for the elderly (75-85) than the middle aged. The difficulty of predicting longevity at old ages is likely due to the prominence of these unforeseeable events relative to middle age.

In addition to the "death by 75" question that has been used for all my mortality risk analysis thus far, the HRS also asks individuals what their subjective perception of the probability of living to 85 is. Using this question, I compare how accurate 55-year-olds perception of their 20-year mortality risk is relative to 65-year-olds'.

The 55-year-olds' beliefs about whether they will live to 75 are more predictive than the 65-year olds' beliefs about whether they will live to 85, reflecting the greater difficulty of predicting mortality outcomes at older ages. This is consistent with the intuitive story that unpredictable health shocks are more frequent and important at older ages, relative to the types of chronic, long-term diseases that cause most of middle aged mortality.

To quantify the degree of predictiveness, figure 11 below shows receiver operating characteristic curves (ROCs) for these perceptions of 20-year longevity. ROCs are a commonly used illustration and measure of predictive accuracy. On the horizontal axis is the false-positive rate (i.e. the type-I error rate). On the vertical axis is the true positive rate. A prediction that was no better than random guessing would have an ROC curve on the 45-degree line. That is, a randomly guessed prediction would have as many false positives as true positives, for any given rate of false positives.[(33)] An ROC curve above the 45-degree line reflects the subjective elicitations having strong predictive power for *actual longevity* (my LASSO prediction from section 3 is not used) - for a given rate of false positives, individual elicitations produce a high rate of true positive predictions.

55-year-olds' 20-year longevity predictions are, overall, more accurate than 65-year-olds'. Figure 11 shows that the ROC lies further above the 45 degree line for 55-year-olds than for 65 year-olds. Moreover, the area between the ROC (AUROC) and the 45 degree line is greater for 55-year-olds than for 65-year-old. The AUROC for 20-year longevity at age 55 are 0.59 for the raw elicitations and 0.68 for the de-rounded elicitations. Analagously for the 20-year mortality of 65-year-oldss the AUROCs are 0.48 and 0.55. This indicates that the elicitations of 65-year-olds are substantially less accurate than 55-year-olds.[(34)] Indeed, the subjective elicitations at age 65 are no better than random.

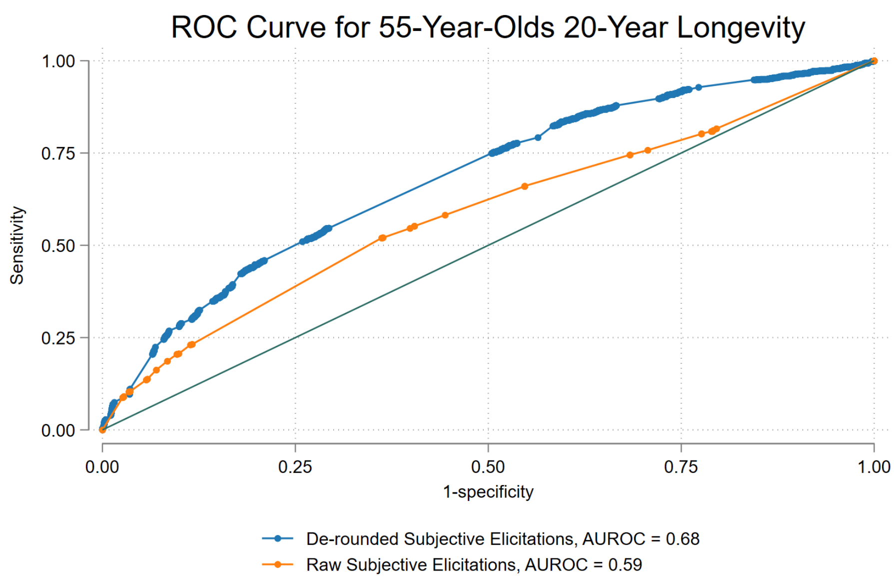
Additionally, 55-year-olds' longevity predictions are informative at either extreme of the risk distribution, whereas 65-year-olds' have no predictive power. This is illustrated by the ROCs for 65-year-olds being coincident with the 45-degree line at very high and very low

[(33)] To understand this, suppose we flipped a balanced coin and predicted 1's and 0's from heads and tails. Among those with a 0 predicted, 1/2 will be true ones, hence the false positive rate is 1/2. Among those with a 1 predicted, 1/2 will actually be true zeroes, hence a true-positive rate of 1/2. This leads to the point (1/2,1/2) being on the ROC for random guessing. By using a coin weighted toward heads or tails, this draws out the entire 45-degree line).

To the extent a prediction is better than random guessing, then that prediction combined with a cutoff rule (e.g. for all those with continuous prediction $\geq x$, assign 1, for all others assign 0) will lead to a ROC well above the 45-degree line - more true-positives for a given rate of false-positives. In my case, I take individual beliefs, pick a cutoff, assign 1's and 0's according to the cutoff, and compute the false positive and true positive rate accordingly. By varying the cutoff the entire ROC curve is traced out. These are plotted in Figure 11.

[(34)] An AUROC can be interpreted as: if we randomly select one person for whom the risk actually realized, and one for whom it didn't, what is the probability the former was assigned a higher predicted risk than the latter. Hence, in my case, the probability that someone who actually survives to 75 has a higher subjective elicitation at 55 than someone who died by 75 is about 20% higher the probability that someone who actually survives to to 85 had a higher subjective belief than someone that died by 85.

perceptions [35]



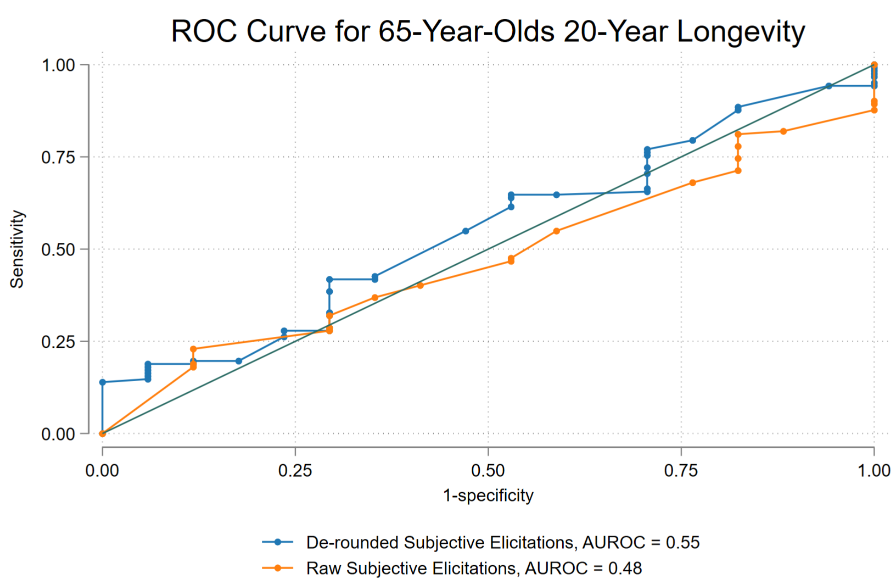


Figure 11: Receiver operating characteristic curves for 20-year longevity risk for 55- vs 65-year-olds.

[35] Note, very high or very low means relative to the distribution of beliefs, not in absolute magnitude. In particular, whether there are many or few beliefs very close to zero or one does not affect the ROC, which in a sense is drawn in quantile space.

C.6 Full set of mortality prediction covariates

The set of variables that appear in the long logit specification, and that were the starting point for the variable selection in the lasso logit specification, are as follows:

1. The respondents subjective elicitation.
2. Sex and age(nearest year)
3. Dummies for diabetes, cancer, lung disease, heart disease, stroke, arthritis, and the first time difference for all of these.
4. Dummies for whether the respondent has ever had any of these diseases.
5. The full set of interactions of 2. and 3.
6. BMI, dummies for being married, seperated, divorced, never marraid. The length of the current marriage.
7. A dummy for whether the respodnents mother and father are alive, and their current age (or age of death).
8. The subjective self health assessment and the first time difference.
9. The number of overnight hospital visits, nursing home visits, doctor visits and episodes of home care since the previous interview. Out of pocket medical expenditures and the first time difference.
10. Indices for activities of daily living (and a time difference), mobility, large muscle, gross motor skills, fine motor skills, and instrumental activities of daily living.
11. Indicators for whether depression was experienced, whether the respondent was happy, the number of days each week in which alchohol was drunk and the first time difference, whether the respondent now smokes or ever smoked, an indicator for high blood pressure.

12. Indicators for past or current memory problems, whether the repondent has public or private health insurance (if the latter how many plans), has life insurance.

13. The number of children, an indicator for whether the respondent is retired, current income and the first time difference in income.

References for Appendix

Beresteanu, A. and F. Molinari (2008). Asymptotic Properties for a Class of Partially Identified Models. *Econometrica 76*(4), 763–814.

Enke, B. and T. Graeber (2019, November). Cognitive Uncertainty. Working Paper 26518.

Fischhoff, B. and W. Bruine De Bruin (1999). Fifty–Fifty=50%? *Journal of Behavioral Decision Making 12*(2), 149–163.

Jeleva, M. and B. Villeneuve (2004, May). Insurance contracts with imprecise probabilities and adverse selection. *Economic Theory 23*(4), 777–794.

Manski, C. and F. Molinari (2010, April). Rounding Probabilistic Expectations in Surveys. *Journal of business & economic statistics : a publication of the American Statistical Association 28*, 219–231.

Miyazaki, H. (1977). The rat race and internal labor markets. *The Bell Journal of Economics 8*(2), 394–418.

Netzer, N. and F. Scheuer (2014). A game theoretic foundation of competitive equilibria with adverse selection. *International Economic Review 55*(2), 399 – 422.

Rothschild, M. and J. Stiglitz (1976, November). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics 90*(4), 629–649.

Spence, M. (1978, December). Product differentiation and performance in insurance markets. *Journal of Public Economics 10*(3), 427–447.

Study, H. . R. (2018). public use dataset. produced and distributed by the university of michigan with funding from the national institute on aging (grant number nia u01ag009740).

Tversky, A. and D. Kahneman (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty 5*(4), 297–323.

Wilson, C. (1977, December). A model of insurance markets with incomplete information. *Journal of Economic Theory 16*(2), 167–207.