Heterogeneity, Demand for Insurance and Adverse Selection

Johannes Spinnewijn*
London School of Economics

January 25, 2012
COMMENTS VERY WELCOME.

Abstract

Recent empirical work finds that surprisingly little variation in the demand for insurance is explained by heterogeneity in risks. I distinguish between heterogeneity in risk preferences and risk perceptions underlying the unexplained variation. Heterogeneous risk perceptions induce a systematic difference between the revealed and actual value of insurance as a function of the insurance price. Using a sufficient statistics approach that accounts for this alternative source of heterogeneity, I find that the welfare conclusions regarding adversely selected markets are substantially different. The source of heterogeneity is also essential for the evaluation of different interventions intended to correct inefficiencies due to adverse selection like insurance subsidies and mandates, risk-adjusted pricing and information policies.

1 Introduction

Adverse selection due to heterogeneity in risks has been considered a prime reason for governments to intervene in insurance markets. The classic argument is that the presence of higher risk types increases insurance premia and drives lower risk types out of the market (Akerlof 1970). However, empirical work has found surprisingly little evidence supporting the importance of adverse selection in insurance markets. An individual's risk type plays only a minor role in explaining his or her demand for insurance, which raises the important question what type of heterogeneity is actually driving the variation in insurance demand. Recent work attributes the unexplained variation to heterogeneity in preferences (Cohen and Einav 2007, Einav, Finkelstein

*Department of Economics, STICERD R515, LSE, Houghton Street, London WC2A 2AE, United Kingdom (email: j.spinnewijn@lse.ac.uk, web: http://personal.lse.ac.uk/spinnewi/).
I thank Pedro Bordalo, Gharad Bryan, Arthur Campbell, Raj Chetty, Erik Eyster, Philipp Kircher, Henrik Kleven, Botond Koszegi, David Laibson, Sendhil Mullainathan, Gerard Padró i Miquel, Matthew Rabin, Frans Spinnewyn and seminar participants at LSE, Lausanne, Zurich and the CESifo meetings for valuable discussions and comments. I would also like to thank Shantayne Chan for excellent research assistance.

and Cullen 2010a, Einav, Finkelstein and Schrimpf 2010b) and finds that the estimated welfare cost of inefficient pricing due to adverse selection is very small. The main reason is that the value of insurance for the uninsured is estimated to be small. Heterogeneity in preferences thus reduces the scope for policy interventions in insurance markets.

An alternative explanation why risks do not explain the demand for insurance is the discrepancy between perceived and actual risks. A large literature documents the biases and heuristics in the formation of risk perceptions and the inherently subjective nature of people's formed risk perceptions.[1] Risk perceptions are likely to be only a noisy measure of one's actual risk, which drives a wedge between the *actual* value of insurance and the value of insurance as *revealed* by an individual's demand.[2] The tenuous relation between insurance choices and the value of insurance is more generally confirmed by recent empirical evidence on inertia (Handel 2011) and insurance choice inconsistencies (Abaluck and Gruber, 2011) and on the substantial role played by cognitive ability for insurance choices (Fang, Keane and Silverman 2008). To the extent that we care about the actual value rather than the revealed value of insurance, the *non-welfarist* heterogeneity driving the demand for insurance may change earlier welfare and policy conclusions.

This paper presents a framework with different dimensions of heterogeneity underlying the variation in the demand for insurance and the resulting adverse selection. I use this framework to extend the sufficient statistics approach by Einav et al. (2010a) and to analyze the importance of different sources of heterogeneity for welfare and policy analysis regarding adverse selection. The analysis leads to two key insights. First, instead of increasing the uncertainty of welfare conclusions, non-welfarist heterogeneity has an unambiguous impact on the welfare cost of adverse selection due to a simple selection effect. Second, the source of heterogeneity underlying the demand for insurance determines the relative effectiveness of the standard policy interventions used to tackle adverse selection. Calibrations based on the empirical analysis in Einav et al. (2010a) suggest that accounting for non-welfarist heterogeneity substantially changes both welfare and policy conclusions.

I consider a simple model where individuals are heterogeneous in risks, preferences and perceptions (or any other non-welfarist noise), and decide whether or not to buy insurance. Even when perceptions are accurate on average, the insured tend to overestimate, while the uninsured tend to underestimate the value of insurance and this at any price. That is, as overly pessimistic beliefs encourage individuals to buy insurance, individuals buying insurance are more likely to be too pessimistic and vice versa.[3] The welfare implication is that the demand curve overstates the surplus for the insured

[1] See Tversky and Kahneman (1974) and Slovic (2000) for the seminal contributions to this literature.

[2] Neighbors in a coastal area have very different perceptions about the risk of a natural disaster damaging their property, even though they face the same actual risk (Peacock et al. 2005).

[3] The selection effect is structurally similar to the mechanisms underlying for example the winner's curse, regression towards to the mean, and choice-driven optimism (Van Den Steen 2004), conditioning an expected value on a particular choice or outcome.

individuals and understates the potential surplus for the uninsured individuals. When taking the demand curve at face value, the evaluation of policy interventions, targeting the insured and uninsured respectively, will be unambiguously biased in opposite directions. For instance, the welfare gain of a universal mandate is unambiguously higher than the demand for insurance would suggest. In order to extend the results from infra-marginal individuals to marginal individuals and thus to evaluate more targeted policy interventions, I then derive conditions under which the *value curve*, depicting the actual value of insurance for the marginally insured at a given price, is a counter-clockwise rotation of the demand curve.[4] With normal heterogeneity, the value curve becomes flatter the lower the correlation between the perceived and actual risk and the larger the variance in perceived risks relative to the variance in actual risks.

I use this systematic relation between the value and demand curve to extend the sufficient statistics approach by Einav et al. (2010a) for non-welfarist heterogeneity. When the demand reveals the actual value, the demand and cost curves are sufficient statistics for welfare analysis. In the presence of non-welfarist heterogeneity, the one additional statistic that is required captures the extent to which heterogeneous choices - left unexplained by heterogeneity in risks - are explained by heterogeneous risk perceptions (or other noise) rather than by heterogeneous preferences. Einav et al. (2010a) illustrate their sufficient statistics approach using data on employer-provided health insurance. I build on their empirical analysis and find that the actual cost of adverse selection would be thirty percent higher when ten percent of the unexplained variation is driven by variation in perceptions and four times as high when this share increases to fifty percent. While disentangling the underlying heterogeneity is challenging, I briefly consider different empirical approaches and find that back-of-the-envelope calculations using existing empirical evidence make a fifty percent share plausible. The cost of adverse selection in this setting may thus be substantially larger than previously estimated.

I also use the framework to analyze and calibrate the impact of non-welfarist heterogeneity on standard government interventions in insurance markets. First, the heterogeneity introduces a disconnect between price and quantity policies. Price policies aiming at increasing insurance coverage are constrained by individuals' perceived valuations, while the welfare effect of an increase in insurance coverage solely depends on their actual valuations. The calibrations suggest that a universal mandate becomes welfare improving when accounting for non-welfarist heterogeneity, while a subsidy inducing the efficient price would increase the welfare loss. Second, policies that inform individuals about their risk have an ambiguous effect on welfare. While information makes individuals better off at a given price, it also changes the selection of individuals buying

[4] Johnson and Myatt (2006) analyze rotations of the demand curve when marketing and advertizing changes the distribution of the value of insurance. Here, the value curve is also a rotation of the demand curve, but the underlying distribution of perceived values is explicitly correlated with the distribution of actual values underlying the original demand curve.

insurance and thus the equilibrium price.[5] The framework with multi-dimensional heterogeneity allows to disentangle these two effects. Improving individuals' information about the expected risk decreases welfare, since those who previously underestimated their cost become more likely to buy insurance and vice versa, regardless of the net value of insurance for these individuals. Improving individuals' information about the variance of the risk increases welfare, since those who previously underestimated the insurance value become more likely to buy insurance and vice versa, regardless of their cost. Finally, adjusting the pricing of insurance contracts for the buyer's particular risk type is often argued to reduce adverse selection and thus decrease inefficiency. I find, however, that the adjustment for risks may in fact increase the inefficiency it is expected to correct when individuals do not perceive these risks accurately.

1.1 Related Literature

Starting with the work by Chiappori and Salanié (1997, 2000), several papers have tested for the presence of adverse selection in different insurance markets, using the testable implication that the correlation between insurance coverage and risk is positive. The mixed evidence, with some insurance markets being advantageously rather than adversely selected (Cohen and Siegelman 2010), inspired a new series of studies which estimate the heterogeneity in risk preferences jointly with the heterogeneity in risk types (Cohen and Einav, 2007; Einav et al. 2010a, 2010b). The estimated heterogeneity allows to move beyond testing for adverse selection and actually analyze the welfare cost of inefficient pricing. This cost is generally found to be small (see Einav, Finkelstein and Levin 2010c).

While attributing heterogeneity in insurance choices - unexplained by heterogeneity in risks - to heterogeneity in preferences is a natural first step and in line with the revealed preference paradigm, several empirical papers have recently made the case that insurance behavior cannot be adequately explained with standard preferences and risk perceptions. Abaluck and Gruber (2011) identify important inconsistencies in the insurance choices of the elderly and document examples of insurance plans that offer better risk protection at a lower cost which are available, but not chosen. Fang et al. (2008) find that heterogeneity in cognitive ability is important (relative to risk aversion) in explaining the choice of elderly to buy Medigap. Cutler and Zeckhauser (2004) argue that distorted risk perceptions are one of the main reasons why some insurance markets do not work efficiently. A number of related empirical papers emphasize the importance of deviations from expected utility maximization in explaining insurance coverage and other choices under risk. For example, Barseghyan, Molinari, O'Donoghue and Teitelbaum (2011) find that a structural model with nonlinear prob-

[5] Condon, Kling and Mullainathan (2011) also discuss the potential welfare loss when people are better informed about their risks. Handel (2010) provides an empirical welfare analysis of a similar trade-off for a nudging policy when people's decisions are subject to switching costs or inertia.

ability weighting explains the data better than a model with standard risk aversion looking at deductible choices in auto and house insurance. Other examples are Bruhin et al. (2010), Snowberg and Wolfers (2010) and Sydnor (2010). Notice that these papers restrict individuals with the same actual risk to perceive the same risk. Most recently, the stability of an individual's risk preference across insurance domains has been tested as well. Barseghyan, Prince and Teitelbaum (2011) reject the hypothesis of stable risk preferences across domains using a structural model. Einav, Finkelstein, Pascu and Cullen (2011) cannot reject the presence of a domain-general component, but also find that an individual's domain-specific risk plays a minor role in explaining insurance choices.

Accounting for non-welfarist heterogeneity when analyzing welfare and policy interventions in insurance markets seems the natural next step in light of the evidence above. The analysis in the paper relates to two recent trends in public economics; the first is the inclusion of non-standard decision makers in welfare analysis, the second is the expression of optimal policies in terms of sufficient statistics.[6] In a similar spirit, Chetty, Kroft and Looney (2009) extend the sufficient statistics approach to tax policy for tax salience and Spinnewijn (2010a) extends the sufficient statistics approach to unemployment policy for biased perceptions of employment prospects. Sandroni and Squintani (2007, 2010) and Spinnewijn (2010b) focus on differences in risk and risk perceptions, analyzing the screening contracts offered in the equilibrium of Rotschild-Stiglitz type models and revisit whether an insurance mandate is Pareto-improving in the respective settings.

The remainder of the paper is as follows. Section 2 introduces a simple model of insurance demand and characterizes the difference between actual and perceived insurance values along the demand curve. Section 3 introduces heterogeneity in risk types and preferences to analyze and calibrate the cost of adverse selection depending on the role of non-welfarist heterogeneity, building on Einav et al. (2010a). Section 4 analyzes the effectiveness of different government interventions depending on the importance of non-welfarist heterogeneity. Section 5 discusses the empirical implementation and the robustness of the welfare and policy analysis. Section 6 concludes.

2 Demand and Welfare

This section introduces a simple model of insurance demand and analyzes the difference between the value of insurance, as revealed by an individual's demand for insurance, and the true value of insurance. I find this difference to be systematic as a function of the insurance price. The analysis deviates from the revealed preference paradigm and assumes that the variation in insurance decisions may be driven by heterogeneity in *non-welfarist* variables, unrelated to the true value of insurance. These non-welfarist

[6] See Congdon et al. 2011 and Chetty 2010 for recent discussions.

variables relate to the notion of ancillary conditions, as introduced by Bernheim and Rangel (2009), which are features of the choice environment that may affect behavior, but not relevant to a social planner's choice. While the distinction between welfarist and non-welfarist variables may be subject to discussion, I will refer to the policy maker who ignores non-welfarist heterogeneity as *naive.* Notice also that the analysis does not only apply to heterogeneity in 'behavioral' variables like misperceptions, inattention, cognitive inability or inertia, but also to heterogeneity in 'economic' variables, like liquidity constraints or adjustment costs, which also restrict people's ability to buy insurance regardless of the value of insurance for those individuals. The analysis is general, but I will mostly interpret the source of the non-welfarist heterogeneity as coming from differences between perceived and actual risks.

2.1 Simple Model

Individuals decide whether or not to buy insurance against a risk. I assume that only one contract is provided and all individuals can buy this contract at a variable price p. Individuals may differ in several dimensions and these different characteristics are captured by a vector ζ. Examples of characteristics are individuals' risk preferences, risk types, perceptions of their risk types, cognitive ability, wealth and liquidity constraints,... I distinguish between the *true* value of insurance $v(\zeta)$ and the *perceived* value of insurance $\hat{v}(\zeta)$ for an individual with characteristics ζ. The true value refers to the actual value of the insurance contract for a given individual and is relevant for evaluating welfare and policy interventions. The perceived value, however, refers to the value as perceived by this individual and determines his or her demand for insurance.

I assume that the perceived value equals the sum of the true value and a noise term ε,

$$\hat{v}(\zeta) = v(\zeta) + \varepsilon(\zeta) \text{ with } E_{\zeta}(\varepsilon) = 0,$$

and that the distributions $F_{\hat{v}}, F_v$ and F_{ε} are continuous. The noise reduces the correlation between the perceived and true value in the population. When capturing misperceptions of the risk, the noise term would be positive when an individual overestimates the risk she is facing and negative when the individual underestimates that risk. I assume that on average the noise cancels out across the entire population so that the average true and perceived value coincide. However, since the demand for insurance depends only on the perceived value, the true and perceived value may differ substantially conditional on the insurance decision.

An individual with characteristics ζ will buy an insurance contract if her perceived value exceeds the price, $\hat{v}(\zeta) \geq p$. The demand for insurance at price p equals $D(p) = 1 - F_{\hat{v}}(p)$. As well known, the demand curve reflects the marginal willingness to pay of marginal buyers at different prices. That is, the price reveals the perceived value for the marginal buyers at that price, $p = E_{\zeta}(\hat{v}|\hat{v} = p)$. However, to evaluate welfare, we would like to know the (average) true value for the marginal buyers, which I denote by $MV(p) \equiv E_{\zeta}(v|\hat{v}=p)$.[7] The central question is thus to what extent the true value co-varies with the perceived value. A central statistic capturing this co-movement is the ratio of the covariance between the true and perceived value to the variance of the perceived value, $cov(v,\hat{v})/var(\hat{v})$.

Graphically, one can construct the *value curve*, depicting the expected true value for the marginal buyers for any level of insurance coverage q, and compare this to the demand curve, depicting the perceived value $D^{-1}(q)$ for that level of insurance coverage, as shown in Figure 1. The mistake made by an *naive* policy maker who incorrectly assumes that the demand curve reveals the true value of insurance depends on the wedge between the two curves. For policies that target marginal individuals, the difference in levels between the two curves is relevant. For policies that target infra-marginal individuals, either the insured or the uninsured, the difference in the areas below the two curves is relevant. I analyze the systematic nature of these differences along the demand curve.

2.2 Infra-marginally Insured and Uninsured

I start by analyzing the average insurance value for the infra-marginal individuals. This is given by $E_{\zeta}(v|\hat{v} \geq p)$ for the insured and by $E_{\zeta}(v|\hat{v} < p)$ for the uninsured. For the insured, the average value of insurance determines the actual surplus generated in the insurance market and thus the value of any policy affecting all insured individuals, like banning an insurance product. For the uninsured, the average value determines the value of any policy affecting all uninsured individuals, like a universal mandate to buy insurance.

Random Noise I first consider the case where the noise determining the perceived value is independent of the true value. The implied co-movement of the actual and perceived value only depends on the relative variances of the true value and the noise term,

$$\frac{cov(v,\hat{v})}{var(\hat{v})} = \frac{var(v)}{var(v)+var(\varepsilon)}.$$

Not surprisingly, an increase in the perceived value is less indicative of an increase in the actual value if noise is more important. Moreover, since the noise term determines the perceived value of insurance, the expected noise realization will be different among those who buy and do not buy insurance.

Proposition 1 *If the true value v and the noise term ε are independent, the demand curve overestimates the value of the insured and underestimates the value of the unin-*

[7] Individuals with the same perceived value may have very different actual values. I take the unweighted average of the insurance value to evaluate welfare. This utilitarian approach implies that in the absence of noise, total welfare is captured by the consumer surplus.

sured,

$$E_\zeta\left(\varepsilon|\hat{v}\geq p\right)\geq 0\geq E_\zeta\left(\varepsilon|\hat{v}<p\right).$$

The Proposition relies on a simple selection effect; characteristics that affect the decision to buy insurance will be differently represented among the insured and the uninsured. Even if these characteristics average out over the entire population, they do not conditional on the decision to buy insurance. For example, since optimistic beliefs discourage individuals from buying insurance and pessimistic beliefs encourage individuals to buy insurance, those who buy insurance are more likely to be too pessimistic, while those who do not buy insurance are more likely to be too optimistic, even when beliefs are accurate on average. This simple argument has important policy consequences. The selection effect unambiguously signs the mistake naive policy makers make by using the demand curve to evaluate welfare consequences of policy interventions targeting either all the insured or uninsured. They overestimate the surplus generated in the insurance market and underestimate the potential value of insurance for the uninsured. As a consequence, universal insurance mandates, often central in the insurance policy debate, are always underappreciated.

Normal Heterogeneity Random noise decreases the correlation between the perceived and true value of insurance and increases the dispersion in the perceived value relative to the dispersion in the actual value. Both a reduction in the correlation and an increase in the relative dispersion decrease the extent to which the true value covaries with the perceived value. For tractability, I only illustrate this here for normal distributions, allowing for correlation between the true value and the noise term, but I extend this insight for more general distributions in Appendix. Denote the mean value of insurance by $\mu_{\hat{v}}=\mu_v$ and the covariance matrix for the joint distribution by

$$\Sigma=\begin{pmatrix}\sigma_v^2 & \rho_{v,\hat{v}}\sigma_v\sigma_{\hat{v}}\\ \rho_{v,\hat{v}}\sigma_v\sigma_{\hat{v}} & \sigma_{\hat{v}}^2\end{pmatrix}.$$

Proposition 2 *If the true and perceived value are normally distributed, the demand curve overestimates the value of the insured and underestimates the potential value for the uninsured if and only if* $\rho_{v,\hat{v}}\times\dfrac{\sigma_v}{\sigma_{\hat{v}}}\leq 1$.

The proposition shows that naive policy makers underestimate the value of insurance for the uninsured when either the correlation between the perceived and actual value of insurance or the dispersion in the perceived values is relatively high. When the wedge between the actual and perceived value is driven by risk perceptions, the correlation is less than perfect for individuals confounding their risk types and thus misperceiving who is facing a higher risk. When individuals also exaggerate the differences in their risk types, this increases the dispersion in perceived values relative to the dispersion in true values and thus increases the wedge even further. However, when individuals underappreciate the differences in their actual risks, the wedge is reduced and could even be reversed. Expressing this in terms of noise, the condition states that the result derived under independence is robust as long as the correlation between the noise term and the true value $\rho_{v,\varepsilon}$ is bounded below by the negative ratio of the standard deviations, $-\frac{\sigma_\varepsilon}{\sigma_v}$.

2.3 Marginally Insured

I continue by analyzing the average value of insurance for the marginal buyers, who are indifferent about buying insurance at a price p. From the selection argument before, we expect that, *on average*, people with high perceived value are more likely to overestimate the value of insurance than people with low perceived value. However, to have that higher perceived values *always* signal stronger overestimation of the true values, we require more structure corresponding to the monotone likelihood ratio property (Milgrom 1981).

Proposition 3 *If $f(\hat{v}|\varepsilon)$ satisfies the monotone likelihood ratio property, $\frac{f(\hat{v}_H|\tilde{\varepsilon})}{f(\hat{v}_H|\varepsilon)} \geq \frac{f(\hat{v}_L|\tilde{\varepsilon})}{f(\hat{v}_L|\varepsilon)}$ for any $\hat{v}_H \geq \hat{v}_L$, $\tilde{\varepsilon} \geq \varepsilon$, the difference between the true and perceived value of insurance is increasing in the price,*

$$\frac{\partial}{\partial p} E_\zeta\left(\varepsilon|\hat{v}=p\right) \geq 0.$$

The immediate policy implication is that a naive policy maker underestimates the value of an increase in insurance coverage more, the thicker the market is. Moreover, since the population averages of the perceived and actual value are assumed to be equal, the demand curve and thus the naive policy maker overestimate the true value of additional insurance if and only if the market is sufficiently thin.

Graphically, the Proposition implies that the value curve is a counter-clockwise rotation of the demand curve, as shown in Figure 1. The value curve lies below the demand curve when prices are high and above the demand curve when prices are low, and the difference between the two curves is monotone in the price. If both the perceived and true values are symmetrically distributed, the intersection of the demand and value curve will be exactly where the price equals the median value, which coincides with the average value. The counter-clockwise rotation naturally implies that the area to the left of any q is larger below the demand curve than below the value curve, while to the right of any q it is smaller, which implies Proposition 1. When the true and perceived value are normally distributed, the condition for the value curve to be a counter-clockwise rotation of the demand curve is $\rho_{v,\hat{v}} \times \dfrac{\sigma_v}{\sigma_{\hat{v}}} \leq 1$, exactly the same as in Proposition 2.

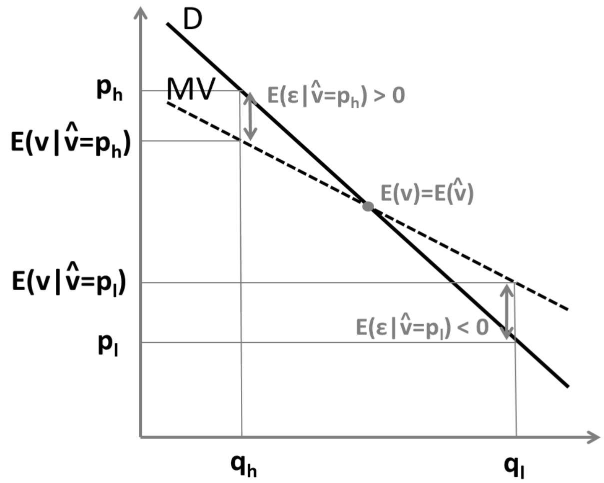

Figure 1: The Demand Curve and the Value Curve.

3 Adverse Selection

I now introduce risk heterogeneity into the analysis and consider the supply of insurance contracts. Particular to insurance markets is that an individual's risk type influences not only her demand for insurance, but also the cost to the insurer of providing insurance to that individual. I decompose a type's valuation of insurance in a risk component and a preference component with only the former determining the cost of insuring that type. Following the approach by Einav et al. (2010a), I derive a sufficient statistics formula to evaluate the welfare cost of inefficient pricing due to adverse selection. This formula shows the mistake made by a *naive* policy maker when determining the efficient price and estimating the cost of adverse selection, by ignoring the non-welfarist heterogeneity underlying the heterogeneous choices.

3.1 Heterogeneity in the Simple Model

The true value of insurance $v(\zeta)$ for an individual with characteristics ζ depends on a risk term, denoted by $\pi(\zeta)$, and a preference term, denoted by $r(\zeta)$,

$$v(\zeta) \equiv \pi(\zeta) + r(\zeta).$$

The risk term not only determines the true value of insurance, but also the expected cost for the insurance company of providing insurance $c(\zeta) \equiv \pi(\zeta)$. Like before, the perceived value equals the true value plus a noise term. The model thus captures heterogeneity in three different dimensions: risk types, risk preferences and risk perceptions.

The setup is kept as simple as possible to keep the analysis insightful, clear and tractable. Notice that this exact setup arises when individuals have CARA preferences and face a normally distributed risk x. In this particular case, the actual value of full insurance equals the sum of the expected risk, $\pi(\zeta) = E(x|\zeta)$, and the risk premium, $r(\zeta) = \frac{\eta(\zeta) Var(x|\zeta)}{2}$, where $\eta(\zeta)$ is the individual's parameter of absolute risk aversion. This suggests that in the decomposition above the preference term should be interpreted as the net value of insurance, i.e., the valuation that is not related to the cost of insurance. The nature of the results would not change if the value and cost function do not depend in an identical way on the individual's risk type $\pi(\zeta)$, neither if the value were not additive in the risk and preference type. Notice that the additivity is not restrictive without restrictions on the distribution of the heterogeneity in the different dimensions.[8]

3.2 Cost of Adverse Selection

The expected cost of an insurance contract depends on the types who decide to buy the contract. The average and marginal cost of providing a contract at price p equal respectively

$$AC(p) = E_\zeta(\pi|\hat{v} \geq p) \text{ and } MC(p) = E_\zeta(\pi|\hat{v} = p).$$

Adverse selection results when the marginal cost is an increasing function of the price. That is, the willingness to buy insurance is lower for lower risk types and they thus decide not to buy insurance at lower prices. The average cost function is increasing as well and lies above the marginal cost function. On the contrary, in advantageously selected markets, individuals with higher risk are less likely to buy insurance and the average cost function will be below rather than above the increasing marginal cost function. The less an individual's risk affects her insurance choice, the less the marginal cost will depend on the price.

In a competitive equilibrium, following Einav et al. (2010a), the competitive price p^c equals the average cost of providing insurance given that competitive price,

$$AC(p^c) = p^c.$$

However, it is efficient for an individual to buy insurance as long as her valuation exceeds the cost of insurance. Hence, at the constrained efficient price p^*, the marginal cost of insurance equals the marginal actual value of insurance,[9]

$$MC(p^*) = MV(p^*) \;\; (= E_\zeta(r + \pi|\hat{v} = p^*)).$$

When the market is adversely selected and the marginal cost is thus below the av-

[8] The assumption of CARA preferences or additivity of the risk premium in the contract valuation is standard in the recent empirical insurance literature (see Einav et al. 2010c).

[9] In the unconstrained efficient allocation, an individual buys insurance if and only if $r \geq 0$. Since individuals with the same perceived value cannot be separated, the constrained efficient allocation has individuals with perceived value $\hat{v}$ buying insurance if and only if $E_\zeta(r|\hat{v}) \geq 0$.

erage cost ($MC(p) < AC(p)$), the competitive price is inefficiently high under the assumption that the demand curve reflects the value of insurance. When the market is sufficiently thick and the demand curve thus underestimates the value of insurance ($p < MV(p)$), the inefficiency is further increased.

The total cost of adverse selection depends on the difference between the value and cost for the pool of inefficiently uninsured individuals with a perceived value between p^* and p^c,

$$\Gamma = \int_{p^*}^{p^c} [MV(p) - MC(p)]\, dD(p).$$

Graphically, the cost equals the area between the value curve and the marginal cost curve from the competitive to the efficient level of insurance coverage, as shown in Figure 2. The level in the competitive equilibrium q^c is given by the intersection of the demand and average cost curve, while the efficient coverage level q^* is given by the intersection of the value curve and the marginal cost curve. When the perceived and actual values coincide, the demand and cost curves are sufficient to determine the cost of adverse selection, as shown by Einav et al. (2010a). However, when the perceived and actual values do not coincide, the demand and cost curves are no longer sufficient. A *naive* policy maker mistakenly beliefs that the efficient price p^n is given by

$$MC(p^n) = p^n$$

and evaluates the inefficiency comparing the wedge between the price and the associated marginal cost. The policy maker thus misestimates this welfare cost as he (1) misidentifies the pool of individuals who should be insured and (2) misestimates the welfare loss for the adversely uninsured,

$$\Gamma = \Gamma^n + \underbrace{\int_{p^*}^{p^n} [MV(p) - MC(p)]\, dD(p)}_{(1)} + \underbrace{\int_{p^n}^{p^c} [MV(p) - p]\, dD(p)}_{(2)},$$

where $\Gamma^n = \int_{p^n}^{p^c} [p - MC(p)]\, dD(p)$ denotes the welfare cost as estimated by a naive policy maker. The difference between Γ and Γ^n depends on the thickness of the market ($MV(p) \gtrless p$) and whether the market is adversely selected or not ($p^c \gtrless p^n$). Figure 2 illustrates the difference between the actual and naively estimated inefficiency cost for a thick and adversely selected market. The inefficiency is higher than a naive policy maker thinks, both because the extent of underinsurance is worse and the welfare loss of underinsurance at a given price is larger than expected.

3.3 Sufficient Statistics Formula

In order to derive a closed-form expression for the cost of adverse selection, I assume normal heterogeneity in all three dimensions. I put no restrictions on the covariance and

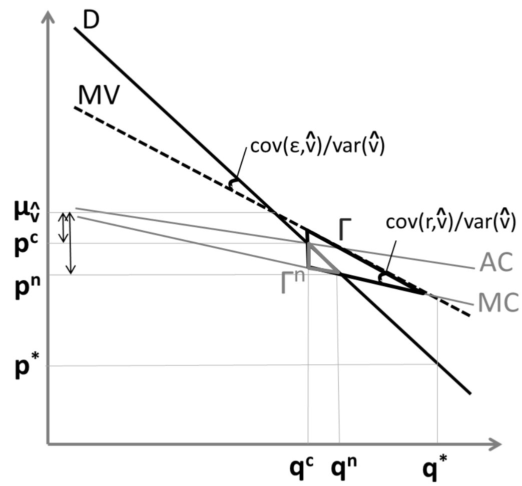

Figure 2: Adverse Selection: the naively estimated cost Γ^n vs. the actual cost Γ.

use notation as before. Under normality, the expected value of any variable $z \in \{\pi, r, \varepsilon\}$, conditional on the perceived value, equals

$$E_\zeta (z|\hat{v} = p) = \frac{cov(z, \hat{v})}{var(\hat{v})} [p - \mu_{\hat{v}}] + \mu_z.$$

The covariance between a term and the perceived value itself indicates how much this term moves with the price. The variation in demand can thus be attributed to the different sources of heterogeneity depending on the relative covariance of each component with the perceived value. If all terms are independent, the covariance of each term with the perceived value is equal to the variance of that term.

The misestimation by a naive policy maker crucially depends on the covariance ratio $cov(\varepsilon, \hat{v}) / cov(r, \hat{v})$, capturing the extent to which the variation in demand is explained by noise rather than by preferences. First, the covariance ratio determines the wedge between the true and perceived marginal surplus of insurance and thus the misestimation of the welfare loss of being uninsured,[10]

$$\frac{E_\zeta (\varepsilon|\hat{v} = p)}{E_\zeta (r|\hat{v} = p) - \mu_r} = \frac{cov(\varepsilon, \hat{v})}{cov(r, \hat{v})}.$$

Second, the covariance ratio determines the difference between the price that is perceived to be efficient and the price that is actually efficient and thus the misidentification

[10] The sign of the misestimation depends not only on whether the expected noise is increasing or decreasing in the perceived value ($cov(\varepsilon, \hat{v}) \gtrless 0$), but also on whether the market is thick or thin $(p \gtrless \mu_{\hat{v}})$.

of the pool of inefficiently uninsured,[11]

$$p^n - p^* = \frac{cov\,(\varepsilon, \hat{v})}{cov\,(r + \varepsilon, \hat{v})} \mu_r.$$

By linearizing the demand curve through (p^n, q^n) and (p^c, q^c), we obtain the following approximate result.

Proposition 4 *With normal heterogeneity, the bias in welfare cost estimation equals*

$$\frac{\Gamma}{\Gamma^n} \cong \frac{[1 + \frac{cov\,(\varepsilon, \hat{v})}{cov\,(r, \hat{v})} \mathcal{P}]^2}{1 + \frac{cov\,(\varepsilon, \hat{v})}{cov\,(r, \hat{v})}} \text{ where } \mathcal{P} \equiv \frac{\mu_{\hat{v}} - p^n}{p^c - p^n}.$$

The impact of the covariance ratio $cov\,(\varepsilon, \hat{v})\,/cov\,(r, \hat{v})$ on the bias in the welfare cost estimation depends on the price ratio $\mathcal{P} = \frac{\mu_{\hat{v}} - p^n}{p^c - p^n}$. Since these prices are known from the intersections of the demand curve with the average and marginal cost curves, as shown in Figure 2, the covariance ratio is the only additional sufficient statistic required when accounting for non-welfarist heterogeneity in the welfare analysis of an adversely selected market. If the price ratio $\mathcal{P}$ is larger than one, the policy maker unambiguously underestimates the efficiency cost of selection. This is the case if the market is thick and adversely selected, $\mu_{\hat{v}} \geq p^c \geq p^n$, since all adversely uninsured are underestimating the value of insurance on average.[12] This is the case that arises in the empirical application.

When half of the market is covered and the competitive price thus equals the average valuation in the market $(p^c = \mu_{\hat{v}})$, the misestimation is approximately linear in the covariance ratio,

$$\frac{\Gamma}{\Gamma^n} \cong 1 + \frac{cov\,(\varepsilon, \hat{v})}{cov\,(r, \hat{v})}.$$

For a thicker market, the bias is larger and increases at faster rate. If the market is sufficiently thin $(p^c > p^n \geq \mu_v)$, the price ratio $\mathcal{P}$ is smaller than one and the policy maker underestimates the inefficiency cost of selection. The thinner the market, the more of the adversely uninsured are overestimating rather than underestimating the value of insurance, but this reverses if the role of noise is sufficiently important such that p^* decreases below μ_v.[13]

[11] Notice that the sign of the price differential is reversed for a thin market, implied by $\mu_r < 0$, where r could be negative, for example due to administrative costs borne by the insured.

[12] Notice that the price ratio $\mathcal{P}$ is also larger than one if the market is thin and advantageously selected, $\hat{\mu}_v \leq p^n \leq p^n$.

[13] Not surprisingly, if the market is thin, but adversely selected, the efficient price may be above the competitive price such that it becomes welfare improving to decrease rather than increase the level of market coverage.

3.4 Calibration

In order to assess the potential importance of the bias, I build on the empirical analysis of employer-provided health insurance by Einav, Finkelstein and Cullen (2010a), henceforth EFC, illustrating their sufficient statistics approach. Based on the health insurance choices and medical insurance claims of the employees of Alcoa, a multinational producer of aluminium, EFC estimate the demand for (additional) insurance coverage and the associated cost of providing (additional) insurance.[14,15] They find that the marginal cost is increasing in the price, but the increase is small. The increase indicates the existence of adverse selection, but the small magnitude of the increase suggests that relatively little heterogeneity in insurance choices is explained by heterogeneity in risks. EFC assume that the residual heterogeneity in insurance choices is due to heterogeneity in (welfarist) preferences and estimate a very small welfare cost of adverse selection, equal to $9.55 per employee per year, with a 95% confidence interval ranging from $1 to $40 per employee. Relative to the average price of $463.5 - the maximum amount of money at stake - this suggest a welfare cost of only 2.2 percent. Relative to the estimated surplus at efficient pricing, this suggests a welfare cost of only 3 percent.

I relax the assumption that the demand curve reveals the actual value of insurance and use the estimates in EFC to illustrate how welfare conclusions are affected when non-welfarist noise explains people's insurance choices. Since the market is thick and adversely selected, the bias in the estimation of the welfare cost increases as a function of $cov(\varepsilon, \hat{v}) / cov(r + \varepsilon, \hat{v})$, as shown in Table 1. I apply the formula derived in Proposition 4, which was derived for a linear approximation of the demand curve under normal heterogeneity. Since EFC estimate a linear system, the formula is exact when the value curve is a rotation of the demand curve like in the case with normal heterogeneity.[16,17] Using the earlier interpretation, I find that if 1 percent of the residual variation is explained by noise, the actual cost of adverse selection is 3 percent higher than estimated when using the demand function. If this share increases to 10 percent, the actual cost of adverse selection is already 31 percent higher. If half of the residual variation is explained by noise, the actual cost of adverse selection is more than 4 times higher

[14] The price variation is argued to be exogenous, as business unit managers set the prices for a menu of different health insurance options, offered to all employees within their business unit.

[15] In particular, they consider a sample of 3,779 salaried employees, who chose one of the two modal health insurance choices, where one option is more expensive, but provides more coverage.

[16] I thus assume that the value curve has slope $\frac{cov(\pi+r,\hat{v})}{var(\hat{v})} p'(q)$ and crosses the demand curve at $q = 0.5$.

[17] I have also evaluated the exact welfare cost when the demand components are normally distributed. The approach to calibrate the covariance matrix from the estimates in EFC is the same as introduced in the next subsections. The demand, value and cost curves are calculated using this matrix. Table App1 in the web appendix shows that the welfare results are very similar for this system of normal demand, cost and value curves. Moreover, the final column shows the estimated bias based on the linear approximation in Proposition 4, suggesting that this approximation works very well when $cov(\varepsilon, \hat{v}) / cov(\varepsilon + r, \hat{v})$ is small.

than estimated based on the demand function. I find even a fifty percent share to be plausible based on back-of-the-envelope calculations using empirical evidence discussed in Section 5. This would imply that rather than \$9.55 per employee per year, the cost of adverse selection would be \$38.4 per employee per year, corresponding to 25 percent of the surplus generated in this market at the efficient price.[18]

4 Policy Interventions

If the cost of inefficient pricing due to adverse selection is small, the net gain from correcting interventions will be small or even negative. The analysis in the previous section showed that the gain may be substantially higher in the presence of non-welfarist heterogeneity. The analysis in this section shows that the nature of the heterogeneity driving the demand for insurance is also crucial for the evaluation of different policy interventions. I discuss insurance subsidies and mandates, information policies and risk-adjustments of insurance premia. For these policy results, I restrict the analysis to a thick ($p^c \leq \mu_v$) and adversely selected ($p^n \leq p^c$) market and assume that the value curve is a counter-clockwise rotation of the demand curve, as discussed in the previous section. I also continue to assume normal heterogeneity in all dimensions.

4.1 Price vs. Quantity

The two most standard interventions in insurance markets are price subsidies and insurance mandates. The question whether health insurance should be subsidized or mandated has been central in the policy debate in the US and other countries during the last decennia. While in some circumstances price and quantity policies are equivalent (Weitzman 1974), this is no longer the case when a wedge is driven between the perceived and actual values. A price subsidy gives financial incentives to individuals to buy insurance. Individuals still decide, based on their perceived value of insurance, whether or not to buy the contract. While the actual value of insurance determines the welfare impact of such a price policy, the perceived value determines how big the price incentives need to be. Hence, inducing an individual to buy insurance through a price policy will be more costly the more the individual underestimates the value of insurance. In contrast, a mandate forces an individual to buy insurance, regardless of her perceived value of the contract. Hence, the cost of implementation does not depend on the perceived values.[19]

To compare the effectiveness of the two types of policies in the presence and absence of non-welfarist heterogeneity, I consider an *efficient-price subsidy* and a *universal mandate*. An efficient-price subsidy reduces the price paid by the insured to the efficient

[18] Notice that the actual efficient allocation is bounded by complete market coverage. The calculations take this into account.

[19] Notice that people's resistance to a mandate or the benefit of implementing a non-universal mandate will depend on the perceived values as well.

price p^*. By inducing the pool of inefficiently uninsured individuals to buy insurance, the welfare gain from such subsidy equals Γ. The cost from such a subsidy is often approximated by $\Phi^S = \lambda q^* [p^c - p^*]$, where λ is the cost of public funds. A counter-clockwise rotation of the value curve due to non-welfarist heterogeneity unambiguously increases Γ, but also increases the cost of implementing the subsidy Φ^S by reducing the efficient price p^*. The change in the net welfare gain $\Gamma - \Phi^S$ is thus ambiguous. A universal mandate realizes the welfare gain Γ as well, but also entails the welfare cost Φ^M from forcing individuals to buy insurance for whom the expected surplus is negative. A counter-clockwise rotation of the value curve does not only increase the gain Γ, but also decreases this cost. Hence, the presence of non-welfarist heterogeneity unambiguously increases the net welfare gain $\Gamma - \Phi^M$.[20]

Policy Result 1 *The presence of non-welfarist heterogeneity underlying the demand curve makes a universal mandate more desirable relative to an efficient-price policy.*

A naive policy maker underestimates the welfare gain Γ, but also overestimates the welfare cost Φ^M and thus will underestimate the value of a universal mandate. When intending to induce the efficient price, a naive policy maker would implement a subsidy equal to $p^c - p^n$ that is too small, thereby realizing a smaller gain than Γ, but this at the lower cost $\Phi^N = \lambda q^n [p^c - p^n]$.

Calibration EFC evaluate the welfare gains and losses from an efficient-price subsidy and a universal mandate, using their estimated demand and cost curves. Table 2 shows how the implied estimates would change when the relative importance of noise underlying the estimated demand curve increases. Setting the cost of public funds λ equal to 0.3, EFC find that the welfare cost of the efficient price subsidy Φ^S equals \$45 per employee per year, almost five times as large as the welfare gain Γ. The net loss from the efficient-price subsidy is even higher when some non-welfarist noise is driving the variation in demand. Despite the increased social gain, the willingness to pay for insurance of the employees for whom insurance is marginally efficient drops substantially. A larger subsidy is required to induce these employees to buy insurance. In contrast, the net gain from a universal mandate unambiguously increases as the role of noise increases in importance, in line with Proposition 2. Columns (1) and (2) show that the increase in the net gain from a universal mandate dominates the change in the welfare gain from a price subsidy, in line with Policy Result 1. The calibration also illustrates that the source of heterogeneity may change the net welfare impact of a policy intervention. Without noise underlying the insurance, the estimates of EFC

[20] Notice that the results could be easily restated by considering an increase in the dispersion of perceived values causing a clockwise rotation of the demand curve, but keeping the value curve fixed. In case of independence, this is simply implied by an increase in the variance of the noise term. This would keep the level of efficient coverage fixed, but reduce the efficient price that induces that level of coverage even if full market coverage is efficient. In contrast with the other case, this would also break the equivalence between a universal mandate and a subsidy inducing universal coverage.

imply that a universal mandate decreases welfare by $20 per employee. However, if more than 17 percent of the unexplained heterogeneity is due to noise, the conclusion is reversed and a universal mandate becomes welfare increasing.

4.2 Information Policies

When individuals take misguided choices, a natural government intervention is to provide them with information to reduce or overcome their mistakes. In this context, information does not unambiguously increase welfare as it may affect the pool of insured and thus the competitive price. While an individual is always as well off when her perceived valuation is closer to her actual valuation for a given price, the higher price when more costly individuals buy insurance reduces coverage in equilibrium. This implies a potential trade-off for information policies. To disentangle the two effects, I consider the impact of two information policies on the equilibrium surplus $S^c = E_\zeta\left(r|\hat{v} \geq p^c\right) \Pr_\zeta\left(\hat{v} \geq p^c\right)$. A first policy increases the correlation between the actual risk π and the *perceived* risk $\hat{\pi} \equiv \pi + \varepsilon$. A second policy increases the correlation between the actual net-value r and the *perceived* net-value $\hat{r} \equiv r + \varepsilon$. In both cases, everything else remains unchanged.[21] As the variance in the perceived values is unaffected by the information policies, the same number of individuals buy insurance at a given price, but the selection of individuals buying insurance will depend on the policy.

The first policy will induce individuals with high risk π rather than individuals with high perceived risk $\hat{\pi}$ to buy insurance. The average expected cost of the individuals buying insurance at a given price level increases, which increases the equilibrium price as the demand function is unaffected. However, the expected net-value of the individuals buying insurance at a given price is still the same. The same surplus is generated for those buying insurance, but less individuals buy insurance. Hence, the equilibrium surplus welfare is unambiguously lower.

Policy Result 2 *An information policy that increases the correlation between the actual and perceived risk, ceteris paribus, unambiguously reduces the equilibrium surplus.*

The second information policy has an opposite effect. While the same number of individuals buy insurance, a higher welfare surplus is generated for those buying insurance. The information policy induces people with a high net-value r to buy insurance, but the competitive price remains unchanged as the expected cost of the individuals buying insurance is not affected. Hence, welfare unambiguously increases.

[21] An alternative interpretation is that the information policy reduces the variance in the noise term, where the noise term is independent of the one term x, but negatively related to the other term y (i.e., $\rho_{\varepsilon,x} = 0$ and $\rho_{\varepsilon,y} = -\frac{1}{2}\frac{\sigma_\varepsilon}{\sigma_y}$ for $x = r, \pi$, $y = \pi, r$). In this interpretation, ε could be interpreted as a misperception of y where the dispersion of the perceived and true y is the same.

Policy Result 3 *An information policy that increases the correlation between the actual and perceived net-value, ceteris paribus, unambiguously increases the equilibrium surplus.*

Better information induces people to make better decisions, but increases the scope for adverse selection. This potential trade-off can be relaxed by providing the right type of information. Information regarding the cost-related value of information will be detrimental, as it only affects the market price, while information regarding the net-value of insurance will be beneficial, as it only affects the selection of the individuals buying insurance. For instance, when individuals with CARA-preferences face a normally distributed risk, the net-value of insurance equals the risk premium, which depends on both the risk aversion and the variance of the risk. Hence, while providing information about one's expected risk may be a bad idea, providing information about the variance of the risk one is facing may be a very good idea. With other factors driving a wedge between the perceived and actual value, the trade-off is similar, but it may be more difficult to separate the two sides. For instance, when switching costs prevent some individuals from buying a new insurance contract, as considered by Handel (2010), a policy that eliminates the switching costs will always be welfare-improving, unless the individuals facing switching costs have particularly high risk types.

Calibration I use the empirical analysis in EFC to evaluate the impact of information policies and shed light on the potential trade-off. While in the previous calibrations the estimated demand and cost functions were taken as given, information policies will change the selection of employees buying insurance contracts and thus the demand and cost functions. I assume that all conditional expectations $E_\zeta(x|\hat{v})$ are linear with slopes equal to $\frac{cov(x,\hat{v})}{var(\hat{v})}p'(q)$, like with normal heterogeneity.[22]

I consider the welfare consequence of information policies reducing the importance of noise by decreasing $var(\varepsilon)$ under three different scenarios. The first scenario assumes that the three demand components r, π and ε are independent. The second scenario assumes that the underlying dispersion in perceived and actual risks is the same (i.e., $var(\pi + \varepsilon) = var(\pi)$) such that the reduction in noise increases only the correlation between the perceived and actual risks, like in Proposition 2. The final scenario assumes that the underlying dispersion in perceived and actual net-value is the same (i.e., $var(r + \varepsilon) = var(r)$) such that a reduction in noise increases only the correlation between the perceived and actual net-value, like in Proposition 3. For each scenario, I

[22] The slope of the normal demand curve equals $\sigma_{\hat{v}}\left[\Phi^{-1}\right]'(1-q)$. The estimated linear slope equals $-1/0.0007$. We need to approximate $\left[\Phi^{-1}\right]'(1-q)$ to back out the implied $\sigma_{\hat{v}}$ from our linear demand curve. Looking at the estimated demand for $q = 0.5$ and $q = 0.7$, in between which all observations in EFC are, we find

$$\frac{\Phi^{-1}(1-0.5) - \Phi^{-1}(1-0.7)}{0.5-0.7} = -2.5.$$

Hence, I set $\sigma_{\hat{v}} = 571.43 = \frac{1}{2.5 \times 0.0007}$.

assume that an initial value for $cov\left(\varepsilon,\hat{v}\right)/cov\left(\varepsilon+r,\hat{v}\right)$ of 0.25, capturing the relative importance of noise underlying the demand function. This seems quite conservative given the back-of-the-envelope calculations in Section 5. Table 3 shows the change in equilibrium welfare $S^c = E_\zeta\left(r|\hat{v}\geq p^c\right)\Pr\left(\hat{v}\geq p^c\right)$ when reducing σ_ε^2 and the cost of inefficient pricing Γ in the new equilibrium.

In the first scenario, assuming independence, a reduction in noise increases the net-value of the insured, but also their expected cost and thus the equilibrium price. While in theory the net impact is ambiguous, column (1a) in table 3 shows that in this case the first effect dominates. When the information policy eliminates all the underlying noise, the surplus in the new equilibrium is 4 percent higher. The cost of inefficient pricing has decreased from \$18.6 to \$14.2. The second and third scenario disentangle the importance of the two opposing effects. Under the second scenario, the information policy induces the more costly types to buy insurance and thus worsens the adverse selection. Welfare is lower in the new equilibrium and the cost of inefficient pricing has increased. With all noise eliminated, welfare decreases by 3 percent, while the cost of inefficient pricing has increased to \$26.5. Finally, under the third scenario, the information policy improves the selection of individuals, without affecting the equilibrium price and thus welfare increases. With all noise eliminated, welfare increases by 12 percent, while the cost of inefficient pricing is halved. The conclusions regarding the desirability of information policies thus crucially depend on the nature of the noise underlying the demand for insurance.[23]

4.3 Risk-Adjusted Pricing

The uniform price paid by all the insured regardless of their risks causes the competitive price to be inefficient. Adjusting the price to reflect an individual's risk could reduce adverse selection, but also introduces inequality between higher and lower risk types, with the higher risk types facing higher prices for the same insurance contract. The policy maker thus faces a difficult trade-off between efficiency and equity. While the equity argument has inspired more regulation of risk-adjustment pricing in recent times (e.g., the ban on gender discrimination in insurance pricing by the European Court of Justice), some recent work emphasizes the efficiency argument, showing that risk-adjusted pricing may substantially increase the net surplus generated in equilibrium (Bundorf, Levin and Mahoney 2011). However, the efficiency gain from adjusting premia to an individuals risk types crucially depends on these individuals perceiving their risk types accurately. If not, the risk-adjustment may decrease rather than increase the net surplus generated in equilibrium.

Consider the adjustment $\beta\left(\pi\right)$ to the insurance premium for an individual with

[23] Table App3 in the web appendix shows that the results are again very similar when the demand components are normally distributed.

risk π, with $\beta(\pi)$ weakly increasing in π and equal to 0 if $\pi = \mu_\pi$.[24] Perfect risk-adjusted pricing is obtained when $\beta(\pi) = \pi - \mu_\pi$. In general, the risk-adjustment can be only based on observable dimensions of the risk (e.g., $\pi = \pi_{unobs.} + \pi_{obs.}$, $\beta(\pi) = \pi_{obs.} - \mu_{\pi_{obs.}}$). With risk-adjusted pricing, an individual buys insurance if and only if

$$\hat{v}(\zeta) \geq p + \beta(\pi(\zeta)) \Leftrightarrow \hat{v}^\beta(\zeta) \geq p,$$

where $\hat{v}^\beta(\zeta)$ denotes the perceived value of insurance net of the risk-adjustment. The cost for the insurer, net of the risk-adjustment, now equals

$$\begin{aligned} AC^\beta(p) &= E_\zeta\left(\pi - \beta(\pi)\,|\hat{v}^\beta \geq p\right). \\ MC^\beta(p) &= E_\zeta\left(\pi - \beta(\pi)\,|\hat{v}^\beta = p\right). \end{aligned}$$

Given these adjusted expressions, we can apply the equilibrium and welfare analysis from before.

The cost of adverse selection depends on the wedge between the competitive and the efficient price and the selection of individuals buying insurance at the competitive price. Adjusting the insurance price for risks affects both. First, pricing the risk (or part of the risk) mechanically reduces the difference between the average and marginal net-cost of providing insurance, conditional on the demand for insurance. That is, the difference between the unpriced risk among the insured and the unpriced risk for the marginal individual is reduced. With perfect risk-adjusted pricing, the average net-cost becomes independent of the price p and coincides with the marginal net-cost curve. The resulting competitive price p^c equals the average risk $AC(p^c)$, which is (constrained) efficient if the demand and value curve coincide. Second, pricing the risk also makes high risk types less likely to buy insurance and low risk types more likely to buy insurance. This selection change lowers the average cost curve and thus the competitive price even further, potentially leading to advantageous selection. Moreover, the change also affects the surplus generated at a given price p,

$$E_\zeta\left(r|\hat{v}^\beta \geq p\right).$$

Intuitively, the insurance surplus will be higher the more risk preferences rather than any other variable drive the demand for insurance. Since the risk type π does not affect the net value of insurance, reducing the role that one's risk plays in the decision to buy insurance increases the equilibrium surplus. The main issue is that when perceived risks are different from true risks, adjusting the prices for the true risks does not reduce the impact of risk on insurance decisions as expected and, in fact, may even increase its impact. For example, when the preference term is independently distributed, risk-

[24] As the average risk is reflected in the competitive price, we can analyze risk-adjustments depending on an individual's risk relative to the average risk without loss of generality.

adjusted pricing increases the surplus at a given equilibrium price only if

$$var\left(\pi+\varepsilon\right)\geq var\left(\pi+\varepsilon-\beta\left(\pi\right)\right).$$

With perfect risk-adjusted pricing, this condition simplifies to $\rho_{\varepsilon,\pi}\geq-\frac{1}{2}\frac{\sigma_{\pi}}{\sigma_{\varepsilon}}$. Hence, if the correlation between the risk and noise term is below this lower bound, the introduction of risk-adjusted pricing reduces the surplus. The following Proposition considers two extreme cases to illustrate the change in surplus in the competitive equilibrium, $S^{c}=E_{\zeta}\left(r|\hat{v}^{\beta}\geq p\right)\Pr_{\zeta}\left(\hat{v}^{\beta}\geq p\right)$.

Policy Result 4 *With accurate risk perceptions, $\hat{\pi}=\pi$, perfect risk-adjusting pricing unambiguously increases the equilibrium surplus. With no heterogeneity in risk perceptions, $\hat{\pi}=E\left(\pi\right)$, perfect risk-adjusted pricing unambiguously decreases the equilibrium surplus.*

When deciding whether or not to buy insurance, an individual does not internalize the cost she is imposing on the insurer. Perfect risk-adjusted pricing corrects this type of externality and induces an efficient decision if risk perceptions are accurate. If not, an individual does not accurately internalize the value of buying insurance for herself either. With no heterogeneity in perceived risks, $\hat{\pi}=E\left(\pi\right)$, this 'internality' exactly offsets the externality such that the introduction of risk-adjusted pricing creates the inefficiency that it is supposed to eliminate. The two considered cases are extreme, but make the policy implications very clear. The impact of risk-adjusted policies very much depend on how the priced risks are perceived. By ignoring the heterogeneity in risk perceptions, a naive policy maker is likely to overestimate the efficiency gain realized by risk-adjusted pricing.

Calibration I build on the empirical analysis in EFC to shed more light on the welfare impact of risk-adjusted pricing. I consider a linear risk-adjustment of the insurance premium $\beta\left(\pi\right)=\beta\times\left[\pi-\mu_{\pi}\right]$, where $\beta=1$ implies perfect risk-adjusted pricing. Like for the information policies, I simulate the new equilibrium price and calculate the equilibrium surplus S^{c} and the cost of inefficient pricing Γ, driven by the wedge between p^{c} and p^{*}.

The first two columns (0a) and (0b) in Table 4 show the positive welfare impact of risk-adjusted pricing in the absence of non-welfarist heterogeneity like in EFC. Equilibrium welfare increases by up to 11 percent when the risk-adjustment is perfect, $\beta=1$. Notice that the new equilibrium is first-best, since the new equilibrium price is (constrained) efficient, $p^{c}=p^{*}$, and individuals buy insurance if and only if $r\geq0$. The elimination of the inefficient wedge between the equilibrium and efficient price accounts for about one third of the welfare increase. These estimates are very similar to the estimates in Bundorf et al. (2011), analyzing the choice between HMO plans and PPO plans offered by 11 employers in the United States between 2004 and 2005. Bundorf et al. (2011) allow for private information about risks next to the observed risk scores, but assume accurate risk perceptions. They find a potential welfare increase of 2-11 percent from pricing the observable risk, where about one fourth is due to eliminating the wedge between the equilibrium and efficient price.

The remaining columns of Table 4 show how different the welfare conclusions are when the actual and perceived risks do not coincide. Like for the information policies, I assume an initial value for $cov(\varepsilon, \hat{v}) / cov(\varepsilon + r, \hat{v})$ of 0.25, capturing the relative importance of noise, and consider three different scenarios showing the importance of the negative correlation between this noise and the actual risk. The first two scenario's are the same as for the information policies. In the first scenario, all components are independent and the welfare impact of risk-adjusted pricing is hardly affected (columns (1a) and (1b)). In the second scenario, the noise term risk is negatively correlated with the actual risks such that the dispersion in perceived and actual risks is the same, i.e., $var(\hat{\pi}) = var(\pi)$. Risk-adjusted pricing still increases welfare, but the increase is reduced to 7 percent for $\beta = 1$ (column 2a). The new third scenario increases the magnitude of this negative correlation resulting in a variance in perceived risk that is half of the variance in actual risks, i.e., $var(\hat{\pi}) = 0.5 var(\pi)$. This scenario illustrates that noisy risk perceptions may not only reduce but even reverse the positive welfare effect of risk-adjusted pricing. With little dispersion in the perceived relative to the actual risks, risks hardly affect the insurance choice. However, risk-adjusted pricing changes this, reducing the prices for the low risk types and thus inducing them to buy insurance, regardless of the net-value of insurance for these types. The opposite is true for the high risk types. The market thus becomes more advantageously selected. When the risk-adjustment is less than perfect, the advantageous selection initially offsets the adverse selection and thus increases welfare, as shown in column (3a). However, with perfect risk-adjustment, the inefficiently low price of insurance increases the cost Γ and welfare is reduced by 3 percent.[25]

5 Discussion

In this final section, I briefly discuss some existing empirical evidence and potential empirical approaches to shed more light on the role of non-welfarist heterogeneity for insurance demand. Related to this, I also discuss the robustness of the welfare and policy results.

Empirical Implementation The analysis shows that the welfare cost of adverse selection and the effectiveness of potential government interventions depend crucially on the sources of heterogeneity driving differences in the demand for insurance. Disentangling the different sources is essential, but challenging as well. The ideal data

[25] Table App4 in the web appendix shows the equilibrium welfare and cost of adverse selection when the demand components are normally distributed. The results are very similar.

contains next to information on the insurance choices and claim rates, also price variation, like in Einav et al. (2010a), and, in addition, provides some way to disaggregate the revealed value of insurance into actual value and noise.

In response to the empirical evidence that risk plays a minor role in explaining insurance choices, some recent studies started analyzing the role of heterogeneity in risk preferences. While some papers have used the heterogeneity in insurance choices, left unexplained by heterogeneity in risks, to estimate heterogeneity in risk preferences (e.g., Cohen and Einav, 2007), only few papers use explicit measures of risk preferences to explain insurance choices (e.g., Cutler, Finkelstein and McGarry 2008). As discussed before, the empirical evidence across individuals (see Cohen and Siegelman, 2010) and across domains (Barseghyan et al., 2011, and Einav et al., 2011b) can attribute only a minor part of the variation in insurance demand to heterogeneous preferences. While these findings seem suggestive, they are not sufficient to conclude that choices hardly reveal the actual value of insurance and that $cov\left(r,\hat{v}\right)/var\left(\hat{v}\right)$ would be small.

Empirical work is needed to provide direct evidence of the role of non-welfarist heterogeneity and the potential importance of $cov\left(\varepsilon,\hat{v}\right)/var\left(\hat{v}\right)$. One natural approach is to identify a variable which does not affect the actual value of insurance, but can be shown to affect the insurance decision.[26] An application of this approach is the evidence in Fang et al. (2008) that cognitive ability is a strong predictor of Medigap insurance coverage, while cognitive ability is unlikely to be related to the actual value of Medigap insurance. Similarly, wealth, income and education are also often found to explain insurance choices. While these variables may be related to the true value of insurance, empirical evidence suggests that they are also strongly related to the mere quality of decisions under uncertainty (Choi et al. 2011).

An alternative application is the empirical analysis of the difference between actual and perceived risks. Clearly, estimating an individual's (ex ante) risk type is very difficult. However, it is sufficient to know the extent to which an increase in the perceived risk corresponds to an increase in the actual risk. For this, one only requires to observe an (ex post) risk realization x_i, drawn from F_{π_i}, and the perceived risk $\hat{\pi}_i$ for each individual, since regressing x_i on $\hat{\pi}_i$ provides an estimate of $cov\left(\pi,\hat{\pi}\right)/var\left(\hat{\pi}\right)$.[27] This estimate can be used to recover the importance of the noise in risk perceptions underlying the demand for insurance, when combined with the estimated relation between the insurance demand and the actual risk types, $cov\left(\pi,\hat{v}\right)/var\left(\hat{v}\right)$. This relation, however, corresponds to the slope of the marginal cost curve relative to the demand curve, like

[26] In a similar spirit, one could try to identify a group of individuals for which the importance of non-welfarist noise is expected to be smaller to uncover the value function associated with the observed demand.

[27] This estimate is consistent if $x_i = \pi_i + \nu_i$ with $E\left(\nu_i|\pi_i\right) = 0$, like for a binary risk with probability π_i.

estimated in EFC using price variation. Decomposing the covariances, we find

$$\begin{aligned}\frac{cov\left(\pi,\hat{v}\right)}{var\left(\hat{v}\right)}\Big/\frac{cov\left(\pi,\hat{\pi}\right)}{var\left(\hat{\pi}\right)} &= \left[\frac{cov\left(\hat{\pi},\hat{v}\right)}{var\left(\hat{v}\right)}-\frac{cov\left(\hat{\pi},r\right)}{var\left(\hat{v}\right)}\right]\times\left[\frac{cov\left(\pi,\hat{\pi}\right)}{cov\left(\pi,\hat{\pi}\right)}+\frac{cov\left(\pi,r\right)}{cov\left(\pi,\hat{\pi}\right)}\right]\\ &\cong \frac{cov\left(\hat{\pi},\hat{v}\right)}{var\left(\hat{v}\right)}.\end{aligned}$$

The approximation depends on the covariance between preferences and perceived or actual risks being small. After subtracting $\frac{cov(\pi,\hat{v})}{var(\hat{v})}$ from both sides, we find

$$\frac{cov\left(\pi,\hat{v}\right)}{var\left(\hat{v}\right)}\left[1\Big/\frac{cov\left(\pi,\hat{\pi}\right)}{var\left(\hat{\pi}\right)}-1\right]\cong\frac{cov\left(\varepsilon,\hat{v}\right)}{var\left(\hat{v}\right)}.$$

Since the EFC analysis implies an estimate for $cov\left(\pi,\hat{v}\right)/var\left(\hat{v}\right)$ of about 1/3, the approximation suggests that if $cov\left(\pi,\hat{\pi}\right)/var\left(\hat{\pi}\right)$ is smaller than 1/2, $cov\left(\varepsilon,\hat{v}\right)/var\left(\hat{v}\right)$ is greater than 1/3. Since by definition, $cov\left(\pi,\hat{v}\right)+cov\left(\varepsilon,\hat{v}\right)+cov\left(r,\hat{v}\right)=var\left(\hat{v}\right)$, $cov\left(r,\hat{v}\right)/var\left(\hat{v}\right)$ is thus smaller than 1/3. Hence, this implies that the heterogeneity in risk perceptions explains more than 50 percent of the variation in demand that is left unexplained by the heterogeneity in actual risks. This back-of-the-envelope calculation thus suggests that our sufficient statistic $cov\left(\varepsilon,\hat{v}\right)/cov\left(\varepsilon+r,\hat{v}\right)$, used in Table 1 and 2, would exceed 0.5 when $cov\left(\pi,\hat{\pi}\right)/var\left(\hat{\pi}\right)$ is indeed smaller than 1/2.

Previous empirical work analyzed the relation between actual and perceived risks, using surveyed risk perceptions, reviewed by Hurd (2009). While surveyed risk perceptions are predictive and often explain more variation in risk realizations than any other set of covariates, the estimated relation is very small.[28] For example, Finkelstein and McGarry (2006) find estimates smaller than 0.10 when estimating a linear probability model of nursing home use in the five years between 1995 and 2000 on the 1995 self-reported beliefs of this probability. When taking the self-reported beliefs at face value, this would imply that $cov\left(\pi,\hat{\pi}\right)/var\left(\hat{\pi}\right)=0.10$. Clearly, the self-reported probability does not measure the demand-driving perceived probability $\hat{\pi}$ without error and measurement error attenuates the regression estimate of $cov\left(\pi,\hat{\pi}\right)/var\left(\hat{\pi}\right)$ towards 0.[29] However, the empirical evidence is generally suggestive that an average increase in the perceived risk may well be associated with an average increase in the actual risk that is less than half.[30]

[28] For an empirical analysis of the relation between subjective life expectations and survival probabilities, see for instance (Hamermesh 1985, Hurd and McGarry 1995 and 2002).

[29] Notice that Finkelstein and McGarry (2006) find a positive relationship between the self-reported probability and insurance coverage, but no significant relationship between the actual risk and insurance coverage.

[30] Kircher and Spinnewijn (2011) suggest an alternative approach using price variation to disentangle perceived risks from risk preferences. Another alternative to measuring perceived risks is to provide information about risks in a controlled experiment and analyze the effect on the demand for insurance and the associated costs.

Robustness Most of the welfare and policy conclusions were derived under the assumption that non-welfarist heterogeneity causes the value curve to be a counter-clock wise rotation of the demand curve. The empirical discussion regarding the relation between the actual and perceived risks suggests already that the assumption regarding the direction of the rotation is appropriate, but even on theoretical grounds I would argue that this is the case. The assumption that heterogeneity in risk perceptions perfectly reflects heterogeneity in risk types seems particularly strong. John C. Harsanyi (1968) observed that "by the very nature of subjective probabilities, even if two individuals have exactly the same information and are at exactly the same high level of intelligence, they may very well assign different subjective probabilities to the very same events." While rationality may restrict individuals to be Bayesian, it puts no restrictions on priors themselves, which are primitives of the model (Van Den Steen 2004). Hence, one's prior $\hat{\pi}$ is likely to be different from one's actual risk type π. Learning about one's risk type will increase the correlation, but learning is likely to be incomplete and any correlation lower than one induces $cov\left(\pi,\hat{\pi}\right)/var\left(\hat{\pi}\right)$ and thus $cov\left(v,\hat{v}\right)/var\left(\hat{v}\right)$ to be smaller than one, suggesting that the discussed bias in evaluating infra-marginal policies. An alternative model of non-welfarist heterogeneity that only affects the correlation assumes that for some individuals the perceived value (or risk) is a random draw from the distribution of the true values (or risks), while for all other individuals the perceived value equals the true value. In this model, the value curve will be a counter-clockwise rotation of the demand curve that is flatter the higher the proportion α of non-standard decision makers, with $cov\left(v,\hat{v}\right)/var\left(\hat{v}\right) = 1-\alpha$.

As discussed in Section 2, the wedge is also affected by the relative dispersion of the perceived and actual risk types. The wedge is further increased when the perceived expected risks are more dispersed than the actual expected risks, while it is reduced and potentially reversed if the perceived risks are less dispersed than the actual risks. To illustrate the importance of the dispersion, a potential model of the relation between perceived and actual risk types has $\hat{\pi} = \mu_\pi + \lambda\left(\pi - \mu_\pi\right)$, also considered in Einav et al. (2010b).[31] This model allows the relative dispersion in the perceived risk types to be larger ($\lambda > 1$) or smaller ($\lambda < 1$), such that the sign of the wedge would only depend on the sign of $\lambda - 1$, since $cov\left(\pi,\hat{\pi}\right)/var\left(\hat{\pi}\right) = \frac{1}{\lambda}$. While it seems plausible that the dispersion in actual risks is larger than in perceived risks (e.g., when everyone perceives a particular risk to be the same), this particular model assumes perfect correlation between the perceived and actual risk types. With imperfect correlation, the dispersion in perceived risk should be sufficiently smaller ($\lambda << 1$) to reverse the wedge. Maybe this explains why empirically, the estimates of $cov\left(\pi,\hat{\pi}\right)/var\left(\hat{\pi}\right)$ tend to be much smaller than one. In the web appendix, I extend the insights that both a decreased correlation and increase in the relative dispersion induce a wedge in the same direction

[31] One could interpret this model as coming from the updating of a uniform prior, equal to the population distribution, using individual signals. The updating is non-Bayesian if indeed $\alpha \neq 1$.

for general, discrete distributions of risk types.

The analysis assumes that the population averages of the actual value of insurance and the value as revealed by the demand are the same. Regarding risk perceptions, various studies suggest that people are too optimistic or too pessimistic on average, depending on the context, the size of the probability, the own control, etc. (see Tversky and Kahneman 1974, Slovic 2000, Weinstein 1980, 1982 and 1984). The presence of an average bias in beliefs clearly induces a wedge between the actual and perceived value of insurance, as analyzed in Spinnewijn (2010). However, the average bias does not affect the nature of the insights regarding the impact of heterogeneity, changing the wedge between the perceived and actual value along the demand curve. Still, the sign of the wedge is relevant for evaluating policy interventions. For instance, heterogeneous risk perceptions induce the uninsured to be more optimistic than the average individual. However, if the average individual is too optimistic, the underappreciation of the insurance value for the uninsured will be even larger and vice versa. The welfare analysis can be easily extended for such average biases. For instance, in Proposition 4, only the price ratio $\mathcal{P} = \frac{\mu_{\hat{v}} - p^n}{p^c - p^n}$ should be adjusted to $\mathcal{P}^x \equiv \frac{x - p^n}{p^c - p^n}$, where x is determined by the intersection of the demand curve and the value curve, solving $F_{\hat{v}}(x) = F_v(x)$. Notice that $x \geq \mu_{\hat{v}}$ if and only if $\mu_v \geq \mu_{\hat{v}}$. Hence, in a thick and adversely selected market, the wedge Γ/Γ^n further increases if there is a pessimistic bias next to heterogeneity in perceptions. Similarly, changes in the symmetry of the distribution of the actual or perceived values would require the use of $\mathcal{P}^x$ rather than $\mathcal{P}$. Graphically, heterogeneity in perceptions induces a rotation of the value curve relative to the demand curve around $(p, q) = (\mu_{\hat{v}}, 0.5)$, while an average optimistic or pessimistic bias introduces a shift and thus changes the intersection of the demand and the value curve. Similarly, if an individual does not buy or switch insurance contracts because of liquidity constraints or inertia, the value curve will rotation of the demand curve around $(p, q) = (\hat{v}_{\max}, 1)$. The demand curve would underestimate the actual value of insurance, but heterogeneity in liquidity constraints or inertia causes the bias to be particularly large for those not buying insurance relative to those buying insurance.

6 Conclusion

What drives the heterogeneity in the demand for insurance? This difficult question has been central in a recent, but already prominent empirical literature. While a number of recent empirical studies suggest that what drives the selection into insurance contracts is often unrelated to the actual value of these contracts, the studies analyzing the importance of adverse selection in insurance markets, have mostly evaluated potential government interventions under the assumption that individuals' choices reveal the actual value of insurance. This paper provides a simple framework to analyze the consequences of heterogeneity in the differences between the actual and revealed value of insurance. The analysis presents a simple selection argument that shows that even without an average bias in the valuation, the welfare conclusions will be systematically biased. Not only the welfare cost of adverse selection, but also the relative welfare gains from standard policy intervention in insurance markets depend on the source of the heterogeneity underlying the demand for insurance. A calibration of the model illustrates that for plausible differences between the actual and perceived value of insurance, the policy conclusions are substantially different.

References

Abaluck, J., and J. Gruber, 2011. Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program. American Economic Review. forthcoming

Akerlof, G., 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. The Quarterly Journal of Economics, 84(3), 488-500.

Arrow, K. J., 1963. Uncertainty and the Welfare Economics of Medical Care. The American Economic Review, 53(5), 941–973.

Barseghyan, L., Molinari, F., O'Donoghue, T., and J. Teitelbaum, 2011. The Nature of Risk Preferences: Evidence from Insurance Choices. mimeo

Barseghyan, L., Prince, J., and J. Teitelbaum, 2011. Are Risk Preferences Stable Across Contexts? Evidence from Insurance Data. American Economic Review, 101(2), 591-631.

Bernheim, D., and A. Rangel, 2009. Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics. Quarterly Journal of Economics, 124(1), 51-104.

Bruhin, A., Fehr-Duda, H. and T. Epper, 2010. Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. Econometrica, 78(4), 1375-1412.

Bundorf, K., Levin, J. and N. Mahoney, 2011. Pricing and Welfare in Health Plan Choice. American Economic Review. forthcoming

Chetty, R., 2009. Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. Annual Review of Economics, 1, 451-488.

Chetty, R., Kroft, L., and A. Looney 2009. Salience and Taxation: Theory and Evidence. American Economic Review, 99(4), 1145-1177.

Chiappori, P., and B. Salanié, 1997. Empirical Contract Theory: The Case of Insurance Data. European Economic Review, 41(3), 943-950.

Chiappori, P., and B. Salanié, 2000. Testing for asymmetric information in insurance markets. Journal of Political Economy, 108(1), 56-78.

Chiappori, P.-A., Jullien, B., Salanié, B., Salanié, F., 2006. Asymmetric Information in Insurance: General Testable Implications. Rand Journal of Economics 37, 783-798.

Choi, S., Kariv, S., Muller, W. and D. Silverman, 2011. Who is (More) Rational? mimeo

Cohen, A., Einav, L., 2007. Estimating Risk Preferences from Deductible Choice. American Economic Review 97(3), 745-788.

Cohen, A., Siegelman, P., 2010. Testing for Adverse Selection in Insurance Markets. Journal of Risk and Insurance 77(1), 39-84.

Congdon, W., Kling, J. and S. Mullanaithan 2011. Policy and Choice: Public Finance Through the Lens of Behavioral Economics. Brookings Institute.

Cutler, D., Finkelstein, A., and K. McGarry, 2008. Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance. American Economic Review Papers and Proceedings.

Cutler, D., and R., Zeckhauser, 2004. Extending the Theory to Meet the Practice of Insurance. In R. Litan and R. Herring, eds., Brookings-Wharton Papers on Financial Service. Washington, D.C.: Brookings Institution.

Einav, L., Finkelstein, A., Cullen, M., 2010a. Estimating Welfare in Insurance Markets using Variation in Prices. Quarterly Journal of Economics 125(3), 877-921.

Einav, L., Finkelstein, A., Schrimpf, P., 2010b. Optimal Mandates and The Welfare Cost of Asymmetric Information: Evidence from the U.K. Annuity Market. Econometrica, 78(3), 1031-1092.

Einav, L., Finkelstein, A., Levin, J., 2010c. Beyond Testing: Empirical Models of Insurance Markets. Annual Review of Economics 2, 311-336.

Einav, L., Finkelstein, A., Cullen, M., and I. Pascu 2011. How general are risk preferences? Choices under uncertainty in different domains. American Economic Review. forthcoming

Fang, H., Keane, M. and D. Silverman, 2008. Sources of Advantageous Selection: Evidence from the Medigap Insurance Market. Journal of Political Economy, 116(2), 303-350.

Finkelstein, A. and K. McGarry, 2006. Multiple dimensions of private information: Evidence from the long-term care insurance market. American Economic Review, 96(4), 938–958.

Hamermesh, D., 1985. Expectations, Life Expectancy, and Economic Behavior. Quarterly Journal of Economics, 100(2), 389-408.

Hurd, M. and K. McGarry, 1995. Evaluations of the Subjective Probability of Survival in the Health and Retirement Survey. Journal of Human Resources, 30, 268-292.

Hurd, M. and K. McGarry, 2002. The Predictive Validity of Subjective Probabilities of Survival. Economic Journal, 112, 966-985.

Hurd, M., 2009. Subjective Probabilities in Household Surveys. Annual Review of Economics, 1, 543-562.

Peacock, W., Brody, S., and W. Highfield, 2005. Hurricane Risk Perceptions among Florida's Single Family Homeowners. Landscape and Urban Planning, 73(2-3), 120-135.

Rothschild, M., Stiglitz, J., 1976. Equilibrium in Competitive Insurance markets: An Essay on the Economics of Imperfect Information. Quarterly Journal of Economics 90, 630-649.

Sandroni, A., Squintani, F., 2007. Overconfidence, insurance and Paternalism. American Economic Review 97(5), 1994-2004.

Sandroni, A., Squintani, F., 2010. Overconfidence and Adverse Selection: The Case of Insurance. working paper

Slovic, P., 2000. The Perception of Risk. London: Earthscan.

Snowberg, E., and J. Wolfers, 2010. Explaining the Favorite–Long Shot Bias: Is it Risk-Love or Misperceptions? Journal of Political Economy, 118(4), 723-746.

Spinnewijn, J., 2010a. Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs. LSE mimeo.

Spinnewijn, J., 2010b. Insurance and Perceptions: How to Screen Optimists and Pessimists. LSE mimeo.

Sydnor, J., 2010. (Over)insuring Modest Risks. American Economic Journal: Applied Economics, 2(4), 177–99.

Tversky, A., Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Biases. Science, 185, 1124-1131.

Weinstein, N., 1980. Unrealistic Optimism about Future Life Events. Journal of Personality and Social Psychology 39(5), 806-820.

Weinstein, N., 1982. Unrealistic Optimism about Susceptibility to Health Problems. Journal of Behavioral Medicine 5(4), 441-460.

Weinstein, N., 1984. Why It Won't Happen to Me: Perceptions of Risk Factors and Susceptibility. Health Psychology 3 (5), 431-457.

Weitzman, M. 1974. Prices vs. Quantities. Review of Economic Studies 41(4), 477-491.

Tables

Table 1: The cost of adverse selection as a function of the noise ratio.

Noise Ratio	Cost of Adverse Selection		
$\frac{cov(\varepsilon, \hat{v})}{cov(\varepsilon + r, \hat{v})}$	Γ (1)	Γ / S^* (2)	Γ / Γ^n (3)
0	9.5	.04	1
.01	9.8	.04	1.03
.10	12.4	.06	1.31
.25	18.6	.10	1.95
.50	38.4	.25	4.03
1	96.6	.62	10.1

Column (1) shows the actual cost of inefficient pricing due to adverse selection Γ expressed in \$ / indiv. Column (2) expresses this actual cost relative to the surplus S^* when the price is (constrained) efficient $p = p^*$. Column (3) expresses this actual cost relative to the estimated cost when ignoring non-welfarist noise, Γ^n. The first row corresponds to the welfare estimates in Einav, Finkelstein and Cullen (2010a), assuming the absence of non-welfarist heterogeneity. The covariance ratio $cov(\varepsilon, \hat{v})/cov(\varepsilon + r, \hat{v})$ captures the importance of non-welfarist noise relative to welfarist noise in explaining insurance choices, conditional on risk.

Table 2: The welfare gain of subsidies and mandates

Noise Ratio	Government Interventions	
$\frac{cov(\varepsilon, \hat{v})}{cov(\varepsilon + r, \hat{v})}$	Price Subsidy $\Gamma - \Phi^S$ (1)	Universal Mandate $\Gamma - \Phi^M$ (2)
0	−35.4	−19.8
.01	−35.7	−18.6
.10	−37.2	−8.1
.25	−41.1	9.3
.50	−125.7	38.4
1	−67.2	96.6

Column (1) shows the net welfare gain from the efficient-price subsidy closing the gap between the equilibrium price p^c and the efficient price p^*, with $\Phi^S = \lambda q^* [p^c - p^*]$. Column (2) shows the net welfare gain from a universal mandate obliging all individuals to buy insurance, where Φ^M denotes the welfare loss from mandating individuals with expected valuation below the expected marginal cost to buy insurance. The first row corresponds to the welfare estimates in Einav, Finkelstein and Cullen (2010a), assuming the absence of non-welfarist heterogeneity. The covariance ratio $cov(\varepsilon, \hat{v})/cov(\varepsilon + r, \hat{v})$ captures the importance of non-welfarist noise relative to welfarist noise in explaining insurance choices, conditional on risk.

Table 3: The Welfare Impact of Information Policies.

Noise Reduction	**Scenario 1**		**Scenario 2**		**Scenario 3**	
	Independence		$var\,(\pi+\varepsilon) = var\,(\pi)$		$var\,(r+\varepsilon) = var\,(r)$	
$\Delta\sigma_\varepsilon^2/\sigma_\varepsilon^2$	$\Delta S^c/S^c$	Γ	$\Delta S^c/S^c$	Γ	$\Delta S^c/S^c$	Γ
	$(1a)$	$(1b)$	$(2a)$	$(2b)$	$(3a)$	$(3b)$
0	0	18.6	0	18.6	0	18.6
.10	.00	18.1	$-.00$	19.2	.01	17.4
.25	.01	17.5	$-.01$	20.3	.03	15.7
.50	.02	16.4	$-.01$	22.2	.06	13.3
1	.04	14.2	$-.03$	26.5	.12	9.5

Columns (1a),(2a) and (3a) show the change in equilibrium welfare $S^c = E_\zeta(r|\hat{v} \geq p^c)\,\Pr(\hat{v} \geq p^c)$ when reducing the variance in noise under the three respective scenario's (relative to the case with no noise reduction). Columns (1b),(2b) and (3b) show the welfare cost in the new equilibrium due to the inefficient pricing Γ. Scenario 1 assumes independence between r, π and ε. Scenario 2 assumes that the variance in perceived risks equals the variance in actual risks. Scenario 3 assumes that the variance in perceived net-values equals the variance in actual net-values. The three scenario's start from an initial value for $cov(\varepsilon, \hat{v})/cov(\varepsilon + r, \hat{v})$ equal to .25. Notice that equilibrium welfare equals S^c= \$243 given this initial value. The demand, value and cost curves are linear with the slopes determined like with normal heterogeneity.

Table 4: The Welfare Impact of Risk-Adjusted Pricing.

Risk Adj.	**No Noise**		**Scenario 1**		**Scenario 2**		**Scenario 3**	
			Independ.		$var\,(\pi{+}\varepsilon) = var\,(\pi)$		$var\,(\pi{+}\varepsilon) = \frac{1}{2}var\,(\pi)$	
β	$\Delta S^c/S^c$	Γ	$\Delta S^c/S^c$	Γ	$\Delta S^c/S^c$	Γ	$\Delta S^c/S^c$	Γ
	$(0a)$	$(0b)$	$(1a)$	$(1b)$	$(2a)$	$(2b)$	$(3a)$	$(3b)$
0	0	9.5	0	18.6	0	18.6	0	18.6
.10	.02	6.8	.02	15.0	.02	14.5	.03	13.2
.25	.05	3.6	.05	10.5	.05	9.9	.05	8.5
.50	.08	.8	.09	5.4	.08	5.7	.06	6.8
.75	.10	.1	.11	2.9	.08	4.7	.02	10.6
1	.11	0	.11	2.2	.07	6.4	$-.03$	19.8

Columns (0a),(1a),(2a) and (3a) show the change in equilibrium welfare $S^c = E_\zeta(r|\hat{v} \geq p^c)\,\Pr(\hat{v} \geq p^c)$ for positive linear shares of the risk-premium adjustment $\beta(\pi) = \beta[\pi - \mu_\pi]$ (relative to the case with no risk-adjustment, $\beta = 0$). Columns (0b),(1b),(2b) and (3b) show the welfare cost in the new equilibrium due to the inefficient pricing Γ. Scenario 1 assumes independence between r, π and ε. Scenario 2 assumes that the variance in perceived risks equals the variance in actual risks. Scenario 3 assumes that the variance in perceived net-values equals the variance in actual net-values. The three scenario's start from an initial value for $cov(\varepsilon, \hat{v})/cov(\varepsilon + r, \hat{v})$ equal to .25. Notice that equilibrium welfare equals S^c= \$243 given this initial value, while it equals S^c= \$272 without noise.

Appendix: Proofs

Proof of Proposition 1

I assume that the random variables are draws from continuous distributions. Denote the density functions of $\hat{v}, v$ and ε by $f(\hat{v}), h(v)$ and $g(\varepsilon)$ respectively. Since by Bayes' law $g(\varepsilon|\hat{v}) = \frac{f(\hat{v}|\varepsilon)g(\varepsilon)}{f(\hat{v})}$, we can rewrite

$$\begin{aligned} g(\varepsilon|\hat{v} \geq p) &= \frac{\int_p g(\varepsilon|\hat{v})\, d\hat{v}}{\int_p f(\hat{v})\, d\hat{v}} \\ &= \frac{\int_p f(\hat{v}|\varepsilon)\, g(\varepsilon)\, d\hat{v}}{\int_p f(\hat{v})\, d\hat{v}} = \frac{\Pr(\hat{v} \geq p|\varepsilon)}{\Pr(\hat{v} \geq p)} g(\varepsilon), \end{aligned}$$

with $\int \frac{\Pr(\hat{v} \geq p|\varepsilon)}{\Pr(\hat{v} \geq p)} g(\varepsilon) = 1$. Moreover, since v and ε are independent, we have that $\Pr(\hat{v} \geq p|\varepsilon) = \int_{p-\varepsilon} h(v)\, dv$ is increasing in ε. Hence, the conditional distribution of $\varepsilon|\hat{v} \geq p$ first-order stochastically dominates the unconditional distribution of ε and thus

$$E(\varepsilon|\hat{v} \geq p) = \int \varepsilon g(\varepsilon) \frac{\Pr(\hat{v} \geq p|\varepsilon)}{\Pr(\hat{v} \geq p)} d\varepsilon \geq \int \varepsilon g(\varepsilon)\, d\varepsilon = E(\varepsilon) = 0.$$

Similarly, we find

$$E(\varepsilon|\hat{v} \leq p) = \int \varepsilon g(\varepsilon) \frac{\Pr(\hat{v} \leq p|\varepsilon)}{\Pr(\hat{v} \leq p)} d\varepsilon \leq \int \varepsilon g(\varepsilon)\, d\varepsilon = E(\varepsilon) = 0.\square$$

Proof of Proposition 2

By normality, we have

$$\begin{aligned} E(\hat{v}|\hat{v} \geq p) - E(v|\hat{v} \geq p) &= \mu_{\hat{v}} - \mu_v + \sigma_{\hat{v}} \frac{\phi\left(\frac{p-\mu_{\hat{v}}}{\sigma_{\hat{v}}}\right)}{1 - \Phi\left(\frac{p-\mu_{\hat{v}}}{\sigma_{\hat{v}}}\right)} - \sigma_v \rho \frac{\phi\left(\frac{p-\mu_{\hat{v}}}{\sigma_{\hat{v}}}\right)}{1 - \Phi\left(\frac{p-\mu_{\hat{v}}}{\sigma_{\hat{v}}}\right)} \\ &= [\sigma_{\hat{v}} - \sigma_v \rho] \frac{\phi\left(\frac{p-\mu_{\hat{v}}}{\sigma_{\hat{v}}}\right)}{1 - \Phi\left(\frac{p-\mu_{\hat{v}}}{\sigma_{\hat{v}}}\right)}. \end{aligned}$$

Hence, $E(\hat{v}|\hat{v} \geq p) \geq E(v|\hat{v} \geq p)$ iff $\sigma_{\hat{v}} \geq \sigma_v \rho$.$\square$

Proof of Proposition 3

This is an immediate application of Proposition 1 in Milgrom (1981). That is,

$$\int \varepsilon g(\varepsilon|\hat{v}_H)\, d\varepsilon \geq \int \varepsilon g(\varepsilon|\hat{v}_L)\, d\varepsilon \text{ for any } \hat{v}_H \geq v_L$$

iff

$$\frac{f(\hat{v}_H|\tilde{\varepsilon})}{f(\hat{v}_H|\varepsilon)} \geq \frac{f(\hat{v}_L|\tilde{\varepsilon})}{f(\hat{v}_L|\varepsilon)} \text{ for any } \tilde{\varepsilon} \geq \varepsilon.$$

Hence, the expected value of noise, conditional on the perceived value, is increasing in the perceived value.□

Proof of Proposition 4

The perceived cost of adverse selection equals

$$\begin{aligned}
\Gamma^n &= \int_{p^n}^{p^c} [p - MC(p)]\, dD(p) \\
&= \int_{p^n}^{p^c} \left[p - \frac{cov(\pi,\hat{v})}{var(\hat{v})}[p-\mu_v] - \mu_\pi\right] dD(p) \\
&= \int_{p^n}^{p^c} \left(1 - \frac{cov(\pi,\hat{v})}{var(\hat{v})}\right)[p - p^n]\, dD(p),
\end{aligned}$$

where $p = MC(p)$ evaluated at $p = p^n$. Hence, the perceived cost of adverse selection is equal to the area between two proportional functions, relative to p^n. Now linearizing the demand function, (i.e., assuming that the density at each price level is the same and equal to $\bar{f}$), this is approximately equal to

$$\begin{aligned}
\Gamma^n &\cong \left(1 - \frac{cov(\pi,\hat{v})}{var(\hat{v})}\right)[p^c - p^n]^2 \frac{\bar{f}}{2} \\
&= \frac{cov(r+\varepsilon,\hat{v})}{var(\hat{v})}[p^c - p^n]^2 \frac{\bar{f}}{2}.
\end{aligned}$$

A similar argument allows to approximate the actual cost of adverse selection,

$$\begin{aligned}
\Gamma &= \int_{p^*}^{p^c} [MV(p) - MC(p)]\, dD(p) \\
&= \int_{p^*}^{p^c} \left[\frac{cov(\pi+r,\hat{v})}{var(\hat{v})}[p-\mu_v] + \mu_v - \frac{cov(\pi,\hat{v})}{var(\hat{v})}[p-\mu_v] - \mu_\pi\right] dD(p) \\
&= \int_{p^*}^{p^c} \frac{cov(r,\hat{v})}{var(\hat{v})}[p - p^*]\, dD(p) \\
&\cong \frac{cov(r,\hat{v})}{var(\hat{v})}[p^c - p^*]^2 \frac{\bar{f}}{2}.
\end{aligned}$$

Hence, the ratio equals

$$\begin{aligned}
\frac{\Gamma}{\Gamma^n} &\cong \frac{cov(r,\hat{v})}{cov(r+\varepsilon,\hat{v})}\frac{[p^c - p^*]^2}{[p^c - p^n]^2} \\
&= \frac{cov(r,\hat{v})}{cov(r+\varepsilon,\hat{v})}\left[1 + \frac{p^n - p^*}{p^c - p^n}\right]^2.
\end{aligned}$$

Now we still want to substitute for the unobservable p^*. By normality, we find that

$$\begin{aligned} p - MC(p) &= \frac{cov(r+\varepsilon, \hat{v})}{var(\hat{v})}[p - p^n], \\ MV(p) - MC(p) &= \frac{cov(r, \hat{v})}{var(\hat{v})}[p - p^*], \end{aligned}$$

since respectively $p^n = MC(p^n)$ and $MV(p^*) = MC(p^*)$. Moreover, notice that at $p = \mu_v$,

$$p - MC(p) = MV(p) - MC(p).$$

Hence,

$$\frac{cov(r+\varepsilon, \hat{v})}{var(\hat{v})}[\mu_v - p^n] = \frac{cov(r, \hat{v})}{var(\hat{v})}[\mu_v - p^*].$$

Rearranging, we find

$$[p^n - p^*] = \frac{cov(\varepsilon, \hat{v})}{cov(r, \hat{v})}[\mu_v - p^n].$$

Substituting this in the previous expression, we find

$$\begin{aligned} \frac{\Gamma}{\Gamma^n} &\cong \frac{cov(r, \hat{v})}{cov(r+\varepsilon, \hat{v})}\left[1 + \frac{cov(\varepsilon, \hat{v})}{cov(r, \hat{v})}\frac{\mu_v - p^n}{p^c - p^n}\right]^2 \\ &= \frac{\left[1 + \frac{cov(\varepsilon, \hat{v})}{cov(r, \hat{v})}\frac{\mu_v - p^n}{p^c - p^n}\right]^2}{1 + \frac{cov(\varepsilon, \hat{v})}{cov(r, \hat{v})}}. \square \end{aligned}$$

Proof of Policy Result 1

We consider the impact of a counter-clockwise rotation of the value curve, keeping the demand and cost functions unchanged (i.e., an increase in $\frac{cov(\varepsilon,\hat{v})}{cov(r,\hat{v})}$, keeping $cov(\pi, \hat{v})$ and $var(\hat{v})$ fixed). The counter-clockwise rotation increases $MV(p)$ for all $p \leq \mu_v$. It also increases the efficient price p^*, solving $MV(p) = MC(p)$ and thus

$$\frac{cov(\pi + r, \hat{v})}{var(\hat{v})}[p^* - \mu_{\hat{v}}] + \mu_\pi + \mu_r = \frac{cov(\pi, \hat{v})}{var(\hat{v})}[p^* - \mu_{\hat{v}}] + \mu_\pi.$$

Hence,

$$\begin{aligned} p^* &= \mu_{\hat{v}} - \mu_r \frac{var(\hat{v})}{cov(r, \hat{v})} \\ &= \mu_r\left(\frac{cov(r, \hat{v}) - var(\hat{v})}{cov(r, \hat{v})}\right) + \mu_\pi \\ &= \mu_\pi - \mu_r\left(\frac{cov(\varepsilon, \hat{v})}{cov(r, \hat{v})} + \frac{cov(\pi, \hat{v})}{cov(r, \hat{v})}\right). \end{aligned}$$

In an adversely selected and thick market, $\mu_r > 0$ and $cov(\pi, \hat{v})$. Hence, an increase in $cov(\varepsilon, \hat{v})$ or decrease in $cov(r, \hat{v})$ decreases the efficient price p^*. The competitive price p^c, however, remains the same.

Hence, the cost of the efficient-price subsidy $\Phi^S = \lambda q^* [p^c - p^*]$ thus increases, since p^* decreases and $q^* = D(p^*)$ increases. The cost of the universal mandate $\Phi^M = \int_{-\infty}^{p^*} [MC(p) - MV(p)]\, dD(p)$ decreases, since $MC(p) \geq MV(p)$ for $p \leq p^*$. Since the gain from both policies is the same Γ, this proves the proposition.□

Proof of Policy Result 2

The correlation $\rho_{\varepsilon,\pi} = -\frac{1}{2}\frac{\sigma_\varepsilon}{\sigma_\pi}$ implies $cov(\pi,\varepsilon) = -\frac{1}{2}var(\varepsilon)$, while $\rho_{\varepsilon,r} = 0$ implies that $cov(r,\varepsilon) = 0$ and thus

$$\begin{aligned} var(\hat{v}) &= var(v) + var(\varepsilon) + 2cov(v,\varepsilon) \\ &= var(v). \end{aligned}$$

The demand function $D(p) = 1 - F_{\hat{v}}(p)$ is thus unaffected by σ_ε. Moreover, $\rho_{\varepsilon,r} = 0$ implies that $\frac{cov(r,\hat{v})}{var(\hat{v})} = \frac{cov(r,v)}{var(v)}$, such that the expected net-value at a price, $E(r|\hat{v} = p) \geq 0$, is unaffected by σ_ε as well. Finally, since $\frac{cov(\pi,\hat{v})}{\sqrt{var(\hat{v})}} = \frac{var(\pi) - \frac{1}{2}var(\varepsilon)}{\sqrt{var(\hat{v})}}$, the average cost,

$$AC(p) = \mu_\pi + \frac{cov(\pi,\hat{v})}{\sqrt{var(\hat{v})}} \frac{\phi\left(\frac{p-\mu_{\hat{v}}}{\sqrt{var(\hat{v})}}\right)}{1-\Phi\left(\frac{p-\mu_{\hat{v}}}{\sqrt{var(\hat{v})}}\right)},$$

increases when σ_ε decreases for any p. Hence, the competitive price $p^c = AC(p^c)$ increases. The welfare surplus, $\int_{p^c}^{\infty} E(r|\hat{v} = p)\, dF(p)$, decreases unambiguously.□

Proof of Policy Result 3

The correlation $\rho_{\varepsilon,r} = -\frac{1}{2}\frac{\sigma_\varepsilon}{\sigma_r}$ implies $cov(r,\varepsilon) = -\frac{1}{2}var(\varepsilon)$, while $\rho_{\varepsilon,\pi} = 0$ implies that $cov(\pi,\varepsilon) = 0$ and thus $var(\hat{v}) = var(v)$. The demand function $D(p) = 1 - F_{\hat{v}}(p)$ is thus unaffected by σ_ε. Moreover, $\rho_{\varepsilon,\pi} = 0$ implies that $\frac{cov(\pi,\hat{v})}{\sqrt{var(\hat{v})}} = \frac{cov(\pi,v)}{\sqrt{var(v)}}$, such that the average cost $AC(p)$ is unaffected by σ_ε as well. Hence, the competitive price p^c remains the same. Finally, since $\frac{cov(r,\hat{v})}{\sqrt{var(\hat{v})}} = \frac{var(r) - \frac{1}{2}var(\varepsilon)}{\sqrt{var(\hat{v})}}$, the expected net-value at a price p,

$$E(r|\hat{v} \geq p) = \mu_r + \frac{cov(r,\hat{v})}{\sqrt{var(\hat{v})}} \frac{\phi\left(\frac{p-\mu_{\hat{v}}}{\sqrt{var(\hat{v})}}\right)}{1-\Phi\left(\frac{p-\mu_{\hat{v}}}{\sqrt{var(\hat{v})}}\right)},$$

is increasing in σ_ε. Hence, the welfare surplus, $\int_{p^c}^{\infty} E(r|\hat{v} = p)\, dF(p) = \Pr(\hat{v} \geq p^c)\, E(r|\hat{v} \geq p)$, decreases unambiguously.□

Proof of Policy Result 4

Consider first the case with accurate risk perceptions, $\hat{\pi} = \pi$. With perfect risk-adjusted pricing, $\beta(\pi) = \pi - \mu_\pi$, the average cost $E(\pi - \beta(\pi)\,|\pi + r \geq p + \beta(\pi)) = \mu_\pi$, independent of the price. Hence, $p^c = \mu_\pi$. An individual thus buys insurance if and only if

$$\pi + r \geq p^c + \beta(\pi) \Leftrightarrow r \geq 0.$$

This is the first-best. Hence, perfect risk-adjusted pricing improves welfare in an adversely selected market. Consider now the case with no heterogeneity in risk perceptions, $\hat{\pi} = E(\pi)$. Without risk-adjusted pricing, the average cost $E(\pi | \mu_\pi + r \geq p) = \mu_\pi$, independent of the price. Hence, $p^c = \mu_\pi$. An individual thus buys insurance if and only if

$$\mu_\pi + r \geq p^c \Leftrightarrow r \geq 0.$$

This is the first-best. However, with perfect risk-adjusted pricing, the competitive price still equals $p^c = \mu_\pi$. However, an individual buys insurance if and only $r \geq \pi$, which is inefficient.$\square$