WHEN IS UTILITARIAN WELFARE HIGHER UNDER INSURANCE RISK POOLING?

By Indradeb Chatterjee†, Angus S. Macdonald‡, Pradip Tapadar†and R. Guy Thomas†

abstract

This paper considers the effect of bans on insurance risk classification on utilitarian social welfare. We consider two regimes: full risk classification, where insurers charge the actuarially fair premium for each risk, and pooling, where risk classification is banned and for institutional or regulatory reasons, insurers do not attempt to separate risk classes, but charge a common premium for all risks. For iso-elastic insurance demand, we derive sufficient conditions on higher and lower risks' demand elasticities which ensure that utilitarian social welfare is higher under pooling than under full risk classification. Using the concept of arc elasticity of demand, we extend the results to a form applicable to more general demand functions. Empirical evidence suggests that the required elasticity conditions for social welfare to be increased by a ban may be realistic for some insurance markets.

keywords

Social welfare; elasticity of demand; insurance risk classification.

contact address

†School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, CT2 7FS, UK.
‡Department of Actuarial Mathematics and Statistics, and the Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK.
Corresponding author: Pradip Tapadar, P.Tapadar@kent.ac.uk.

1. Introduction

Restrictions on insurance risk classification are common in life insurance and other personal insurance markets. Examples include the ban on gender classification in the European Union, and restrictions in many countries on insurers' use of genetic test results. Such restrictions are usually perceived by economists as having negative effects on efficiency. But because restrictions also make high risks better off and low risks worse off, they also have equity (distributional) effects. Therefore depending on distributional preferences expressed in the social welfare function, restrictions might either increase or decrease social welfare.

The social welfare function used in this paper assumes cardinal and interpersonally comparable utilities, and assigns equal weights to the utilities of all individuals. This equal-weights approach is based on the Harsanyi (1955) 'veil of ignorance' argument: that is, behind the (hypothetical) veil of ignorance, where one does not know what position in society (e.g. high risk or low risk) one occupies, the appropriate probability to assign to being any individual is $1/N$, where N is the number of individuals in society. Alternative risk classification regimes can then be compared by comparing expected utility in each regime for the (hypothetical) individual utility-maximiser behind the veil of ignorance.

We use this approach to evaluate two risk classification regimes: full risk classification, where insurers charge the actuarially fair premium for full cover for each risk, and pooling, where risk classification is banned and so insurers charge a common premium for full cover for all risks. We assume that insurers compete only on price; for institutional or regulatory reasons, they do not offer partial cover, nor menus of contracts offering different levels of cover priced at different rates. In this sense, our approach follows the tradition of Akerlof (1970) rather than Rothschild and Stiglitz (1976).

Under the pooling regime, it is intuitive that the equilibrium price – the pooled price at which insurers break even – will depend on demand elasticities of lower and higher risks. Another intuition is that pooling implies a redistribution from lower risks towards higher risks. The welfare outcome will depend on how we evaluate the trade-off between the gains and losses of the two types. This paper connects and builds on these intuitions, by establishing sufficient conditions on demand elasticities to ensure higher social welfare under pooling compared with full risk classification. The conditions encompass many plausible combinations of higher and lower risks' demand elasticities.

1.1 Literature Review

The closest precedent to the present paper is Hoy (2006), which shows that when potential losses are fixed and the fraction of high risks in the population is sufficiently small, then a ban on risk classification will increase utilitarian welfare. Polborn et al. (2006) obtain a similar result in a dynamic model of life insurance, where the quantum of insurance which an individual can purchase is not fixed, but is subject to a cap.[1] Another strand of literature (e.g. Crocker and Snow (1986), Rothschild (2011)) argues that contract-specific taxes or partial social insurance are a Pareto-superior means to implement any welfare improvements achieved by a ban. Notwithstanding this argument, bans remain of interest because for reasons of political feasibility or administrative convenience, they are invariably the preferred means in practice.

A principal departure of this paper from all those just cited is that rather than assuming all individuals have the same utility function, we assume a distribution of utility

[1] 'Dynamic model' here denotes an initial period in which the individual is uninformed about her risk level and insurance needs, then a second period where she receives information about both, and finally a third period when she is exposed to risk; she may buy insurance in either the first or second periods.

functions (not necessarily all risk-averse) across individuals who have the same probabilities of loss. This assumption leads to qualitatively different results from simpler models, through two mechanisms. First, utility functions determine individuals' insurance purchasing decisions, which determine the insurance demand curve and hence the equilibrium price of insurance when all risks are pooled. Second, utility functions determine the expected utilities which individuals assign to their outcomes *given* an insurance price. Our measure of social welfare is expected utility *given* the distributions of loss probabilities and preferences in society, but evaluated *behind* a hypothetical veil of ignorance which screens off knowledge of the decision maker's own loss probability and preferences.

This paper is also related to Hao et al. (2019) which proposes 'loss coverage', defined as expected losses compensated by insurance for the whole population, as a criterion for risk classification schemes, and points out that loss coverage has the advantage that it depends only on observables (whereas utilitarian social welfare depends on unobservable utility functions). Hao et al. (2019) shows that for iso-elastic insurance demand with elasticity the same for all risk-groups, loss coverage can be used as a proxy measure for social welfare, because it always gives the same ranking of different risk classification schemes. But for other demand specifications, the 'common ranking' property of loss coverage and social welfare may not hold. The present paper therefore focuses on direct evaluation of social welfare, and derives sufficient conditions on demand elasticities for social welfare to be higher under pooling than under full risk classification.

1.2 Outline of the Paper

The rest of this paper is organised as follows. Section 2 presents our models of insurance demand and market equilibrium. Section 3 establishes demand elasticity conditions for social welfare to be higher under pooling than under full risk classification, given iso-elastic demand the same for all risk-groups. Section 4 then considers different iso-elastic demand elasticities for different risk-groups. Section 5 uses the construct of 'arc elasticity of demand' to extend the results in Section 4 in a form applicable to more general demand functions. Section 6 summarises the results, and discusses empirical data and complementary results from other authors. Section 7 gives conclusions.

2. Model Set-up

2.1 Insurance Demand for a Single Risk-Group

Typical theories of insurance demand assume that all individuals know their own probabilities of loss and have a common utility function. Given an offered premium, individuals with the same probabilities of loss then all make the same purchasing decision. This does not correspond well to the observable reality of many insurance markets, where individuals who appear to have similar probabilities of loss often make different decisions, and substantial fractions of the population do not purchase insurance at all.[2] This section gives a theory of insurance demand which accommodates the possibility that not all individuals with the same probabilities of loss make the same decision. Key assumptions which distinguish our model from other common models are highlighted at the points where the need for each assumption arises.

First we consider demand from the perspective of a single individual. Suppose that an individual has wealth W and risks losing an amount L. The individual is offered insurance against the potential loss amount L at premium π (per unit of loss), i.e. for a payment of πL.

Assumption 1. *The individual's utility function $u(w)$, is increasing as a function of wealth, w, and differentiable, so that $u'(w) > 0$. The individual knows his own utility function.*

Note that in Assumption 1, no restriction is placed on the second derivative $u''(w)$, which may have either sign; we do not require that all individuals are risk-averse (i.e. $u''(w) < 0$).) We will show later that this departure from typical models generates the partial take-up of insurance in our demand function.

Assumption 2. *Insurance is offered in a full-cover contract which is standardised across all insurers, who compete only on price. Insurers do not offer partial cover or other contract menus.*

We justify Assumption 2 by noting that separation via contract menus is not possible in some important markets, such as life insurance, which have non-exclusive contracting. It is also often not salient to practitioners in other markets where restrictions on risk classification apply.[3]

The individual will choose to buy insurance if:

$$u(W - \pi L) > (1 - \mu)\, u(W) + \mu\, u(W - L). \tag{2.1}$$

[2]For example, in life insurance, the Life Insurance Market Research Association (LIMRA) states that 57% of US households have some individual life insurance (LIMRA (2019)). The American Council of Life Insurers states that 138m individual policies were in force in 2018 (American Council of Life Insurers (2019, p66)); the US adult population (aged 18 years and over) at 1 July 2019 as estimated by the US Census Bureau was 255m.

[3]Economists often postulate that insurers use menus of deductibles or other contract features as screening devices to separate high and low risks (e.g. Rothschild and Stiglitz (1976)). But most actuarial pricing textbooks make no reference to this concept (e.g. Gray and Pitts (2012), Friedland (2013), Parodi (2014)), and instead interpret deductibles as a device to limit moral hazard and the administrative costs of handling small claims.

Since certainty-equivalent decisions do not depend on the origin and scale of a utility function, we shall find it convenient to adopt the following standardisations: $u(W) = 1$ and $u(W - L) = 0$. Under this standardisation, the individual will purchase insurance if:

$$u(W - \pi\, L) > (1 - \mu). \tag{2.2}$$

Next we consider demand from the perspective of an insurer. The insurer observes a group of individuals comprising a *risk-group*, who all have the same probability of loss. The insurer knows the common probability of loss μ for all members of the risk-group. The individuals are, however, heterogeneous in terms of their utility functions, which the insurer cannot observe.

Assumption 3. *Utility functions are heterogeneous across individuals, and unobservable by insurers.*

Hence for any risk-group, the insurer observes μ, π and possibly each individual's W and L, but not their utility functions. So from the insurer's perspective, given a premium π, the utility of insurance of an individual chosen at random from this risk-group, $u(W - \pi\, L)$, is unobservable and we denote it by the random variable: U_I (the subscript I indicates *insurance*), which depends on W, L and π.

So the insurer can at most observe the proportion of individuals who choose to buy insurance at a given premium π. We call this a (proportional) demand function and define it as:

$$d(\pi) = \mathrm{P}\left[\, U_I > (1 - \mu) \,\right]. \tag{2.3}$$

Clearly, $0 \leq d(\pi) \leq 1$ and $d(\pi)$ is non-increasing in π (for a given value of μ) as increasing π decreases the utility of insurance for all individuals.

A related concept, the (point price) elasticity of insurance demand, is defined as:

$$\epsilon(\pi) = -\frac{\partial \log d(\pi)}{\partial \log \pi} \tag{2.4}$$

which implies that demand can also be expressed as

$$d(\pi) = \tau \exp\left[-\int_{\mu}^{\pi} \epsilon(s)\, d\log s \right] \tag{2.5}$$

where $\tau = d(\mu)$ is the *fair-premium demand* for insurance.

2.2 Insurance Market Equilibrium with n Risk-Groups

Suppose a population consists of n distinct risk-groups with probabilities of loss given by $\mu_1, \mu_2, \ldots, \mu_n$. For convenience, we assume $0 < \mu_1 < \mu_2 < \ldots < \mu_n < 1$. Let the proportion of the population belonging to risk-group i be p_i, for $i = 1, 2, \ldots, n$.

Now let the occurrence of a loss event for an individual chosen at random from the whole population be represented by the indicator random variable, X, taking the value of 1 if a loss event occurs; and 0 otherwise. Then X, conditional on risk-group i, is a Bernoulli random variable with parameter μ_i.

Suppose insurers charge premiums (per unit of loss) $\pi_1, \pi_2, \ldots, \pi_n$ for the risk-groups $i = 1, 2, \ldots, n$, respectively. For brevity, we use the notation $\underline{\pi} = (\pi_1, \pi_2, \ldots, \pi_n)$ to denote the premium regime under consideration. Define Π to be the premium which would be chargeable to an individual chosen at random from the population, if that individual purchased insurance. Then Π, conditional on risk-group i, takes the value π_i.

From insurers' perspective, the insurance purchasing decision of an individual chosen at random from the whole population can be represented by the indicator random variable Q, taking the value of 1 if insurance is purchased; and 0 otherwise. Then Q, conditional on risk-group i, is a Bernoulli random variable with parameter $d_i(\pi_i)$, where $d_i(\pi_i)$ is the demand for insurance within risk-group i at premium π_i (based on the model developed in Section 2.1).

Then for an individual chosen at random from the population, the expected premium income is $\mathrm{E}\,[\,Q\Pi L\,]$ and the expected insurance claim is $\mathrm{E}\,[\,QXL\,]$.

We assume that competition between insurers leads to zero expected profits in equilibrium. That is, the *equilibrium condition* for the insurance market under the premium regime $\underline{\pi}$, which we denote by $\rho(\underline{\pi})$, is :

$$\rho(\underline{\pi}) = \mathrm{E}\,[\,Q\Pi L\,] - \mathrm{E}\,[\,QXL\,] = 0. \tag{2.6}$$

2.3 Social Welfare

We define social welfare, $S(\underline{\pi})$, for a particular premium regime $\underline{\pi}$, as the expected utility of an individual selected at random from the entire population, i.e.:

$$S(\underline{\pi}) = \mathrm{E}\,[\,Q\,U_I + (1-Q)\,[\,(1-X)\,U_W + X\,U_{W-L}\,]\,], \tag{2.7}$$

where U_W and U_{W-L} are random variables denoting the utilities at individuals' initial wealth, W, and at their wealth after loss event, $(W-L)$, respectively. In Equation 2.7, the 'Q' term is the random utility if insurance is purchased, and the '$(1-Q)$' term is the random utility if insurance is not purchased.

In Section 2.1 we noted that certainty-equivalent decisions do not depend on the origins and scales of utility functions, and therefore the insurance decision for all individuals could be framed as one where $u(W) = 1$ and $u(W-L) = 0$, irrespective of their different individual utility functions. This was not a model requirement, but just a convenient standardisation.

However, this argument cannot be directly extended to Equation 2.7, because the utilitarian concept of social welfare does depend on the actual magnitudes of individuals' utilities at different levels of wealth. But without any standardisation, Equation 2.7 is susceptible to being dominated by a 'utility monster' who derives more utility from a given level of wealth than all other individuals combined (see Bailey (1997), Nozick (1974)). This makes it unsuitable for policy purposes. So we will continue the previous standardisation of utilities in Equation 2.7, as stated in the following assumption.

Assumption 4. *The utilities of all individuals are standardised at the 'end-points' of the range* $(W - L, W)$ *so that* $u(W) = 1$ *and* $u(W - L) = 0$*.*

Under this standardisation, Equation 2.7 simplifies to:

$$S(\underline{\pi}) = \mathrm{E}\left[\, Q\, U_I + (1 - Q)\,(1 - X)\,\right]. \tag{2.8}$$

For many insurances, insurance premiums are typically relatively small compared to an individual's wealth.[4] In this paper, we assume that the premium $\pi\, L$ is 'small' in the following technical sense.

Assumption 5. *All individuals' utility functions are such that for small premium amounts* πL *(compared to initial wealth* W*), the second and higher-order terms in the Taylor series of expansion of* $u(W - \pi L)$ *can be ignored as negligible.*

It is important to highlight here that, we are <u>not</u> suggesting that the curvatures of individuals' utility functions are unimportant. Indeed, the specific values of individuals' $u(W - \pi\, L)$ depend on the evolution of the utility curves over the range $(W - L, W - \pi\, L)$, taking into account all the characteristics of the underlying utility functions, including slopes and curvatures. Assumption 5 only requires that for small premium amounts $\pi\, L$, the utility function $u(w)$ over the short interval $(W - \pi\, L, W)$ can be approximated by a straight line.

To illustrate the effect of Assumption 5, Figure 1 shows standardised utility functions over the range $(W - L, W)$ for four hypothetical individuals with different risk preferences. The straight diagonal line from $u(W - L)$ to $u(W)$ through point C represents a risk-neutral individual. The concave curves through points A and B each represent risk-averse individuals and the convex curve through point D represents a risk-loving individual.[5] The role of utility functions' slopes and curvatures, over the range $(W - L, W - \pi\, L)$ to portray individual risk preferences, is evident in the four distinctive curves and also in the relative

[4] There are some notable exceptions, such as health insurance in some jurisdictions, and our analysis will not apply in these cases.

[5] Although 'risk-loving' or 'risk-seeking' are the usual stylised descriptions, it might be more appropriate to characterise this phenomenon as 'risk-neglecting'.

differences in the values of $u(W - \pi L)$. Assumption 5 says that, for small πL, each individual's utility curve over the short interval $(W - \pi L, W)$ can be approximated by a straight line .

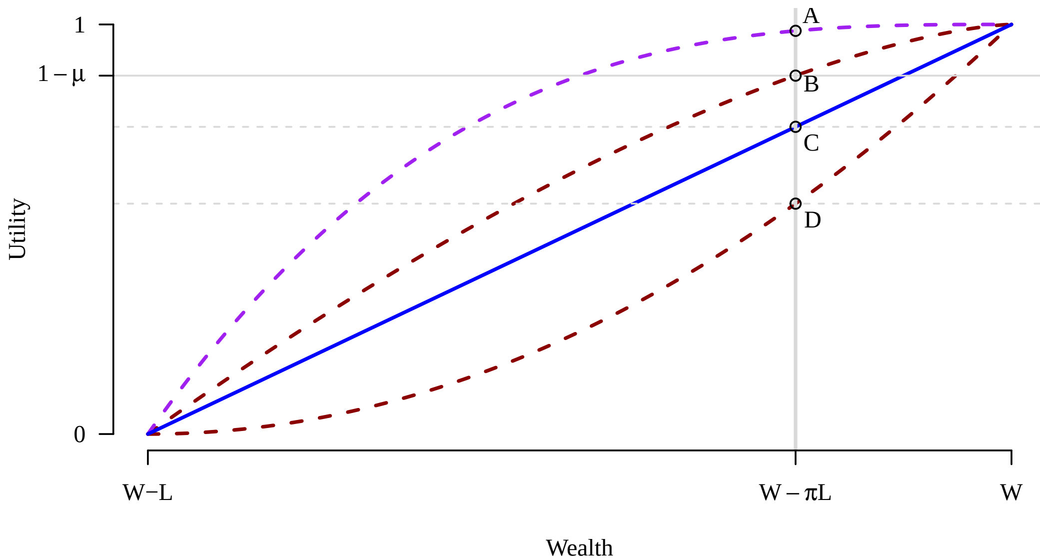

Figure 1: Intuition for $\gamma = L\,u'(W)$ as an index of risk preferences .

From Equation 2.2, an individual's decision rule for purchasing insurance is:

$$u(W - \pi L) > (1 - \mu). \tag{2.9}$$

Using Assumption 5, the left-hand side of Equation 2.9 can be evaluated as:

$$u(W - \pi L) \approx u(W) - \pi L\,u'(W) = 1 - \pi L\,u'(W), \quad \text{as } u(W) = 1. \tag{2.10}$$

Using the approximation in Equation 2.10, the individual's decision rule in Equation 2.9 becomes:

$$L\,u'(W) < \frac{\mu}{\pi}. \tag{2.11}$$

Now, if we define $\gamma = L\,u'(W)$, then for a given individual, the decision rule can be written as:

$$\gamma < \frac{\mu}{\pi}. \tag{2.12}$$

The quantity $\gamma = L\,u'(W)$ can be interpreted as a *risk preferences index*, in the sense illustrated in Figure 1. The straight diagonal line, representing a risk-neutral individual, has a slope of $1/L$, giving the index $\gamma = L\,u'(W) = 1$. The concave curves through points A and B representing risk-averse individuals have lower slopes $u'(W)$ than for the straight diagonal line, and hence the index $\gamma = Lu'(W) < 1$ for risk-averse individuals. For the convex curve through point D, representing a risk-loving individual, an analogous geometric intuition confirms $\gamma = L\,u'(W) > 1$. Provided that Assumption 5 holds, the index $\gamma = L\,u'(W)$ is then sufficient to characterise an individual's risk preferences at wealth $(W - \pi\,L)$.

As an example, consider the special case of power utility function $u(w) = w^{\gamma}$, with $W = L = 1$. The parameter γ fully characterises an individual's risk preferences. For this particular example, Assumption 5 implies that for small premium π:

$$u(1-\pi) = (1-\pi)^{\gamma} \approx 1 - \pi\,\gamma, \quad \text{as } u(1) = 1 \text{ and } u'(1) = \gamma. \tag{2.13}$$

And for this specific power utility example, the decision rule then becomes:

$$u(1-\pi) > (1-\mu) \Leftrightarrow (1 - \pi\,\gamma) > (1-\mu) \Leftrightarrow \gamma < \frac{\mu}{\pi}, \tag{2.14}$$

reproducing the same general decision rule as obtained in Equation 2.12.

Note that in accordance with the decision rule in Equation 2.9, insurance is purchased if $u(W - \pi\,L) > (1-\mu)$: so in this illustration, A purchases, B is indifferent, and C and D do not purchase. The variation across individuals in utility functions drives the partial take-up of insurance (i.e. $d(\pi) < 1$) in our model.

Since insurers cannot observe individuals' utility functions (Assumption 3), γ is not observable and appears to be sampled randomly from some underlying random variable Γ with distribution function $F_{\Gamma}(\gamma)$. Following on from Equation 2.12, the (proportional) insurance demand function in Equation 2.3 can be expressed as:

$$d(\pi) = \mathrm{P}\left[\,U_I > (1-\mu)\,\right] = \mathrm{P}\left[\,\Gamma < \frac{\mu}{\pi}\,\right]. \tag{2.15}$$

By applying Taylor series approximation as in Equation 2.10, the expression for social welfare in Equation 2.8 can now be approximated by:

$$S(\underline{\pi}) \approx \mathrm{E}\left[\,Q\,(1 - \Pi\,\Gamma) + (1-Q)\,(1-X)\,\right], \tag{2.16}$$

$$= \mathrm{E}\left[\,Q\,(X - \Pi\,\Gamma)\,\right] + K, \tag{2.17}$$

where $K = E[1-X]$ does not depend on the premium regime under consideration.

The development to this point accommodates the possibility that potential loss amounts L vary across individuals. But to obviate the need to model this variation in this paper, we make our next assumption:

Assumption 6. *For all individuals, the potential loss amount L is the same constant.*

Under this assumption, the equilibrium condition $\rho(\underline{\pi}) = 0$ from Equation 2.6 simplifies to:

$$\mathrm{E}\left[\, Q\Pi \,\right] - \mathrm{E}\left[\, QX \,\right] = 0. \tag{2.18}$$

To progress to a parameterised version of Equation 2.18, we need to assume that there is no moral hazard. Technically:

Assumption 7. *Conditional on a given risk-group, Q and X are independent.*

Given this assumption, conditioning over the different risk-groups and then taking conditional expectation, the equilibrium condition in Equation 2.18 yields:

$$\mathrm{E}\left[\, Q\,\Pi - Q\,X \,\right] = 0$$

$$\Leftrightarrow \sum_{i=1}^{n} \mathrm{P}[\text{Risk-group } i] \left[\mathrm{E}\left[\, Q\,\Pi \mid \text{Risk-group } i\right] - \mathrm{E}\left[\, Q\,X \;\mid \text{Risk-group } i\,\right]\right] = 0 \tag{2.19}$$

$$\Leftrightarrow \sum_{i=1}^{n} p_i \left[\pi_i \mathrm{E}\left[\, Q \mid \text{Risk-group } i\right] - \mathrm{E}\left[\, Q \mid \text{Risk-group } i\,\right] \mathrm{E}\left[\, X \;\mid \text{Risk-group } i\,\right]\right] = 0 \tag{2.20}$$

(as $\Pi = \pi_i$ for risk-group i; and Q and X are independent given a risk-group),

$$\Leftrightarrow \sum_{i=1}^{n} p_i\, d_i(\pi_i)\; (\pi_i - \mu_i) = 0, \tag{2.21}$$

as given a risk-group i, Q and X are Bernoulli random variables with parameters $d_i(\pi_i)$ and μ_i respectively. Equation 2.21 is intuitively appealing as it can be interpreted as the demand-weighted average profits generated by different risk-groups.

By inspection, $\underline{\pi} = \underline{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)$ is a solution to Equation 2.21, and we will refer to this as the *full risk classification* regime.

At the other end of the spectrum is the *pooling* regime where risk-classification is banned and all risk-groups are charged the same premium $\pi_i = \pi_0$ for $i = 1, 2, \ldots, n$. Since the insurance demand in our model is a continuous function of premium, there exists at least one premium π_0 where $\mu_1 \leq \pi_0 \leq \mu_n$ and $\rho(\pi_0) = 0$.[6]

In between these two extremes of *full risk classification* and *pooling*, there can be many possible equilibria (subject to suitable regulation). For example, some variables

[6]For notational convenience, we specify only one argument for multivariate functions if all arguments are equal, e.g. we write $\rho(\pi)$ for $\rho(\pi, \pi, \ldots, \pi)$.

(e.g. gender) may be banned, but not others (e.g. smoking status). However, in this paper we focus our attention only on the two extreme premium regimes.

The final assumption in our model set-up is not a strict requirement, but is made for presentational convenience:

Assumption 8. *No risk-group is fully insured under any risk classification regimes.*

It is possible to envisage a situation where an entire risk-group is fully insured, if the premium charged is sufficiently small. This special case can also be analysed using the same model framework. However for ease of exposition, we present our findings based on Assumption 8 in the main text of this paper, while the special case of possible full take-up of insurance for certain risk-groups is given in Appendix F.

3. Iso-elastic Insurance Demand

In this section, we apply the framework created in Section 2 to develop the simple case of iso-elastic insurance demand and derive the analytical form for social welfare, $S(\underline{\pi})$, so that we can compare the pooling regime against the full risk classification regime.

One tractable form of insurance demand function, for risk-group i, is the iso-elastic demand:

$$d_i(\pi_i) = \tau_i \left(\frac{\mu_i}{\pi_i}\right)^{\lambda_i}, \quad \text{(subject to a cap of 1)}, \tag{3.1}$$

which produces a constant demand elasticity:

$$\epsilon(\pi_i) = -\frac{\partial \log(d_i(\pi_i))}{\partial \log \pi_i} = \lambda_i. \tag{3.2}$$

The parameter τ_i can be interpreted as the *fair-premium demand*, that is the demand when an actuarially fair premium is charged.

The above iso-elastic insurance demand can be constructed within our model set-up as follows. Consider an individual from risk-group i, with initial wealth W, who risks losing an amount L. Suppose her risk preferences are driven by a power utility function:

$$u(w) = \left[\frac{w-(W-L)}{L}\right]^{\gamma}, \tag{3.3}$$

so that $u(W) = 1$ and $u(W-L) = 0$. This particular form of utility function leads to:

$$u'(w) = \frac{\gamma}{L}\left[\frac{w-(W-L)}{L}\right]^{\gamma-1}, \text{ and so consequently:} \tag{3.4}$$

$$L\,u'(W) = \gamma. \tag{3.5}$$

So under the framework of power utility functions, the *risk preferences index*, $L\,u'(W)$, defined in Section 2.3, can be interpreted as the underlying parameter, γ, of the power utility function.

As outlined in Section 2.3, γ is sampled randomly from some underlying random variable Γ_i with distribution function $F_{\Gamma_i}(\gamma)$, and the demand for insurance for risk-group i at a given premium π_i is then:

$$d_i(\pi_i) = \mathrm{P}\left[\Gamma_i < \frac{\mu_i}{\pi_i}\right]. \tag{3.6}$$

The demand for insurance for risk-group i takes the form of iso-elastic demand given in Equation 3.1 if Γ_i has the following distribution:

$$F_{\Gamma_i}(\gamma) = \mathrm{P}\left[\Gamma_i \le \gamma\right] = \begin{cases} 0 & \text{if } \gamma < 0 \\ \tau_i\,\gamma^{\lambda_i} & \text{if } 0 \le \gamma \le (1/\tau_i)^{1/\lambda_i} \\ 1 & \text{if } \gamma > (1/\tau_i)^{1/\lambda_i}, \end{cases} \tag{3.7}$$

where τ_i and λ_i are positive parameters. λ_i controls the shape of the distribution function and τ_i controls the range over which Γ_i takes its values.[7]

Using the specific form of iso-elastic demand, the analytical form of social welfare given in Equation 2.17 for a particular premium regime $\underline{\pi}$, is provided in Lemma 1 (proof in Appendix A).

Lemma 1. *Suppose there are n risk-groups with risks $\mu_1 < \mu_2 < \cdots < \mu_n$ with iso-elastic demand elasticities $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively, then for a given premium regime $\underline{\pi}$, the expression for social welfare is given by:*

$$S(\underline{\pi}) = \sum_{i=1}^{n} p_i\,\tau_i\,\frac{1}{(\lambda_i+1)}\left(\frac{\mu_i}{\pi_i}\right)^{\lambda_i+1}\pi_i + K, \tag{3.8}$$

where the premium regime $\underline{\pi}$ satisfies the equilibrium condition:

$$\sum_{i=1}^{n} p_i\tau_i\left(\frac{\mu_i}{\pi_i}\right)^{\lambda_i}(\pi_i - \mu_i) = 0, \tag{3.9}$$

and the constant K does not depend on the premium regime under consideration.

[7]This is a generalised version of the Kumaraswamy distribution, which in its standard form takes values only over [0,1] (Kumaraswamy (1980)). Note that $\tau_i = \lambda_i = 1$ leads to a uniform distribution.

Lemma 1 provides the basis for comparing any two premium regimes. Specifically, in this paper we will focus on comparing pooling regimes against full risk classification regimes.

Under pooling, it is sometimes notationally convenient to express the equilibrium condition and social welfare in terms of the *risk-premium ratios*: $v_i = \mu_i/\pi_0$. A risk-premium ratio of $v_i < 1$ indicates that the i-th risk-group pay more than than their fair actuarial premium, and conversely for $v_i > 1$. Using this notation, the pooling equilibrium in Equation 3.9 becomes:

$$\sum_{i=1}^{n} \alpha_i v_i^{\lambda_i+1} = \sum_{i=1}^{n} \alpha_i v_i^{\lambda_i}, \tag{3.10}$$

$$\text{or, equivalently:} \quad \sum_{i:\, v_i>1} \alpha_i \left[v_i^{\lambda_i+1} - v_i^{\lambda_i}\right] = \sum_{i:\, v_i\leq 1} \alpha_i \left[v_i^{\lambda_i} - v_i^{\lambda_i+1}\right], \tag{3.11}$$

where $\alpha_i = \frac{p_i \tau_i}{\sum_{j=1}^{n} p_j \tau_j}$ and the social welfare condition Equation 3.8 can be expressed as:

$$S(\pi_0) \gtreqless S(\underline{\mu}) \Leftrightarrow \sum_{i=1}^{n} \frac{\alpha_i\, v_i^{\lambda_i+1}}{\lambda_i+1} \gtreqless \sum_{i=1}^{n} \frac{\alpha_i\, v_i}{\lambda_i+1},^{8} \tag{3.12}$$

$$\Leftrightarrow \sum_{i:\, v_i>1} \frac{\alpha_i}{\lambda_i+1} \left[v_i^{\lambda_i+1} - v_i\right] \gtreqless \sum_{i:\, v_i\leq 1} \frac{\alpha_i}{\lambda_i+1} \left[v_i - v_i^{\lambda_i+1}\right]. \tag{3.13}$$

Equation 3.11 says that under pooling equilibrium losses from the high risk-groups are exactly offset by the profits from the low risk-groups. And Equation 3.13 can be interpreted as the comparison between the (aggregate) utility gains by the high risk-groups (from pooling as compared against full risk classification) against the (aggregate) utility losses of the low risk-groups.

We can now derive the conditions for which social welfare under pooling is higher than that under full risk classification. In the first instance, we make the simplest assumption that all risk-groups have the same positive constant demand elasticity λ. Under this assumption, we obtain the following result (proof in Appendix B) :

Theorem 1. *Suppose there are n risk-groups with risks $\mu_1 < \mu_2 < \cdots < \mu_n$ with the same positive constant demand elasticity λ for all risk-groups. Then:*

$$\lambda \lesseqgtr 1 \Rightarrow S(\pi_0) \gtreqless S(\underline{\mu}). \tag{3.14}$$

[8] We use the notation $\gtreqless$ in the following sense: $A \gtreqless B \Rightarrow C \gtreqless D$ is shorthand for $A > B \Rightarrow C > D$ and $A = B \Rightarrow C = D$ and $A < B \Rightarrow C < D$. A similar interpretation applies for the notation $\lesseqgtr$.

Figure 2: Curve showing the ratio of $(S(\pi_0) - K)$ to $\big(S(\underline{\mu}) - K\big)$, for different values of constant demand elasticity λ, where $(\mu_1, \mu_2) = (0.01, 0.04)$ and $(\alpha_1, \alpha_2) = (0.8, 0.2)$.

Figure 2 provides a graphical representation of Theorem 1 showing the ratio of $(S(\pi_0) - K)$ to $\big(S(\underline{\mu}) - K\big)$ as a function of constant demand elasticity λ for two risk-groups with risks $(\mu_1, \mu_2) = (0.01, 0.04)$ and $(\alpha_1, \alpha_2) = (0.8, 0.2)$. Recall from Equation 3.8, in the expression for $S(\underline{\pi})$, K is a constant which does not depend on the premium regime $\underline{\pi}$. So the ratio of $(S(\pi_0) - K)$ to $\big(S(\underline{\mu}) - K\big)$ focuses solely on the effect of changes in premium regimes.

It can be clearly seen that $\lambda = 1 \Rightarrow S(\pi_0) = S(\underline{\mu})$, while $\lambda < 1 \Rightarrow S(\pi_0) > S(\underline{\mu})$ and vice versa, as postulated in Theorem 1.

4. Different Iso-elastic Demand Elasticities for Different Risk-Groups

Theorem 1 assumes the same constant iso-elastic demand elasticity for all individuals. However, it is entirely possible that different risk-groups have different sensitivities to price changes. In particular, for higher risk consumers, insurance premiums may represent a larger part of their total budget constraint, and so the effect of a small percentage change in price on their insurance demand might be larger. In this section, for ease of exposition, we first consider two risk-groups with iso-elastic demand, but with different demand elasticities. We then generalise our result to more than two risk-groups.

Typical insurance underwriting processes often classify a majority of insurance risks as *standard* (or low risks in the terminology of this paper), with the remaining risks rated higher based on their individual characteristics. The empirical evidence (cited in Table 1) suggests that the more numerous low risk-group's demand elasticity may often be less than 1. But, as noted above, the high risk-group's demand elasticity is likely to be higher than that the low risk-group, and may often exceed 1. This motivates Theorem 2 (proof in Appendix C).

Theorem 2.1 states a sufficient condition on λ_1 and λ_2 for social welfare to be higher under pooling than under full risk classification for any population structures and underlying risks. Theorem 2.2 then extends it for some of the ranges of λ_2 not covered in Theorem 2.1, but this involves introduction of additional conditions.

Theorem 2. *Suppose there are two risk-groups with risks $\mu_1 < \mu_2$ with positive constant demand elasticities λ_1 and λ_2 respectively.*

2.1. *For any population structure:*

$$\lambda_1 \leq 1 \text{ and } \lambda_1 \leq \lambda_2 \leq \frac{1}{\lambda_1} \Rightarrow S(\pi_0) \geq S(\underline{\mu}). \tag{4.1}$$

2.2. *For any population structure there exists a threshold premium π^* such that:*

$$\lambda_1 \leq 1 \text{ and } \lambda_2 > \frac{1}{\lambda_1} \text{ and } \pi_0 \geq \pi^* \Rightarrow S(\pi_0) \geq S(\underline{\mu}). \tag{4.2}$$

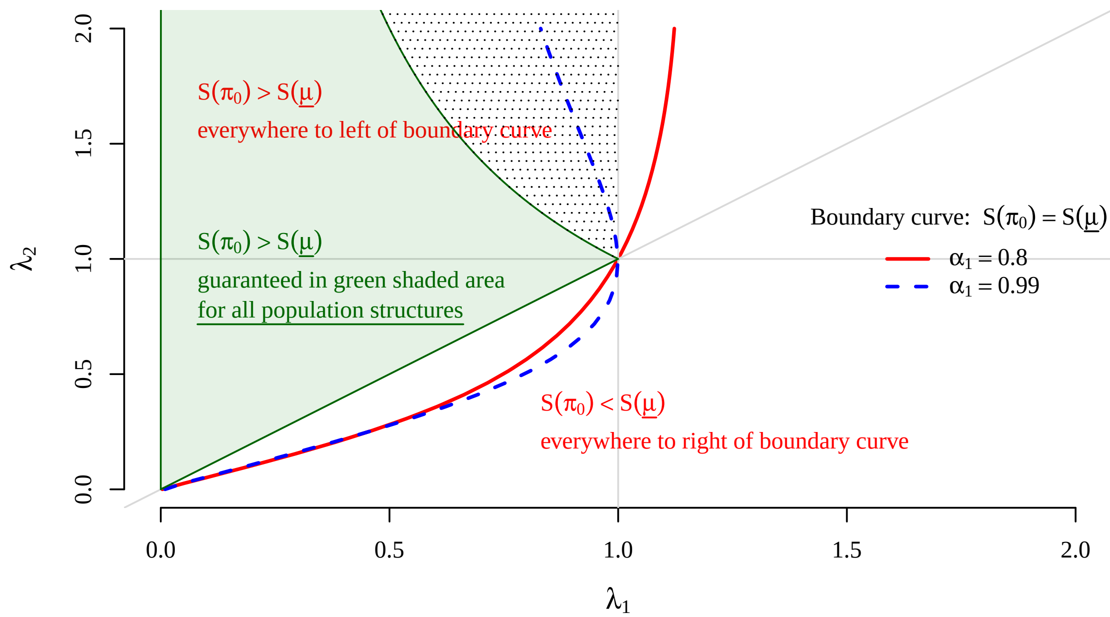

Figure 3: Curve demarcating the regions where social welfare under pooling is greater than under full risk-differentiation when $(\mu_1, \mu_2) = (0.01, 0.04)$ and for different values of (α_1, α_2).

Theorem 2.1 is illustrated in Figure 3, where $(\mu_1, \mu_2) = (0.01, 0.04)$ and the two axes represent the demand elasticities λ_1 and λ_2. The two curves emanating from the origin show the boundary at which $S(\pi_0) = S(\underline{\mu})$ for two possible population structures. The bold red curve demarcates the boundary for a moderate population structure with $(\alpha_1, \alpha_2) = (0.8, 0.2)$; while the dashed blue curve is the boundary for an extreme population structure with very few high risks, $(\alpha_1, \alpha_2) = (0.99, 0.01)$. Social welfare under pooling is higher than under full risk classification on the left of the boundary curves, and lower on the right. The sufficient conditions in Theorem 2.1 specify that in the green shaded region where $\lambda_1 \leq 1$ and $\lambda_1 \leq \lambda_2 \leq 1/\lambda_1$, social welfare under pooling is *always* higher than that under full risk classification, *irrespective* of the population structure (and also the risks μ_1 and μ_2).

To understand the green shaded area of Figure 3, first note that moving from full risk classification to pooling always leads to (i) a beneficial *increase* in both the *number* of high risks insured, and the *per capita* utility of insured high risks and (ii) a detrimental *decrease* in both the *number* of low risks insured, and the *per capita* utility of each insured low risk. An initial intuition is that pooling will tend to "work well" when lower risks' elasticity is low compared with higher risks' elasticity.

As we move leftwards in the graph with λ_2 fixed, λ_1 eventually becomes *sufficiently low* compared with λ_2, so that pooling "works well" and effect (i) dominates. As we move upwards in the graph with λ_1 fixed (where $\lambda_1 \leq 1$), λ_2 eventually becomes *sufficiently high* compared with λ_1, so that pooling again "works well" and effect (i) dominates. However, if the high risk-group is small and has high demand elasticities, it may not have the required capacity to absorb all the aggregate utility losses of the low risk-group. This is illustrated by the curvature of the dashed blue line for $\alpha_1 = 0.99$ (a very small fraction of high risks) back towards the vertical axis for $\lambda_2 > 1$ (high elasticities of the high risks). This explains the upper bound on λ_2, i.e. $\lambda_2 \leq 1/\lambda_1$, in Theorem 2.1. Where the conditions of Theorem 2.1 are satisfied (i.e. inside the green shaded region), effect (i) is guaranteed to dominate for *any* population structures and risks.

Note that the conditions in Theorem 2.1 are sufficient, but *not* necessary. This non-necessity is illustrated by the white and dotted regions adjacent to the green shaded region, but to the left of the red boundary curve, where $S(\pi_0) > S(\underline{\mu})$ for the population structure $\alpha_1 = 0.8$ even though the conditions of Theorem 2.1 are not satisfied. Where the conditions of Theorem 2.1 are not satisfied, social welfare may still be higher under pooling than under full risk classification, but this might require additional conditions. For the region $\lambda_1 \leq 1$ and $\lambda_2 > 1/\lambda_1$ (dotted in Figure 3), Theorem 2.2 identifies the additional condition in the form of the equilibrium premium π_0 needing to exceed a threshold premium π^* for social welfare under pooling to be higher.

An implication of Theorem 2.2 is that the high risk group needs to be of a large enough size to pull the equilibrium premium above the threshold. This can be interpreted as the need for the high risk-group to be of a reasonably large size to absorb the impact of aggregate utility losses for the low risk-group. Figure 3 shows that for the extreme population structure with very few high risks, $\alpha_1 = 0.99$, the dashed blue boundary curves back into the dotted region indicating that the condition $\pi_0 \geq \pi^*$ may not always be satisfied. In contrast, for a moderate population structure with $\alpha_1 = 0.8$, the bold red boundary curves back into the dotted region only at much higher values of λ_2 (not shown in the figure).

Theorem 2 can be generalised for more than two risk-groups with iso-elastic demand for all risk-groups. While generalising our results to more than two risk-groups, under pooling it will be convenient to classify the different risk-groups into two broad categories:

- 'lower' risk-groups, for whom pooled premium is higher than fair premium, i.e. $\mu_i \leq \pi_0$;
- 'higher' risk-groups, for whom pooled premium is lower than fair premium, i.e. $\mu_i > \pi_0$.

For these two broad categories, we define the following:

- $\lambda_{lo}^{min} = \min\{\lambda_i : \mu_i \le \pi_0\}$, i.e. minimum demand elasticity for lower risk-groups;
- $\lambda_{lo}^{max} = \max\{\lambda_i : \mu_i \le \pi_0\}$, i.e. maximum demand elasticity for lower risk-groups;
- $\lambda_{hi}^{min} = \min\{\lambda_i : \mu_i > \pi_0\}$, i.e. minimum demand elasticity for higher risk-groups;
- $\lambda_{hi}^{max} = \max\{\lambda_i : \mu_i > \pi_0\}$, i.e. maximum demand elasticity for higher risk-groups.

For the case of two risk-groups, we simply have: $\lambda_{lo}^{min} = \lambda_{lo}^{max} = \lambda_1$ and $\lambda_{hi}^{min} = \lambda_{hi}^{max} = \lambda_2$.

Using these notations, we present our general result (for proof see Appendix C):

Theorem 3. *Suppose there are n risk-groups with risks $\mu_1 < \mu_2 < \cdots < \mu_n$ with iso-elastic demand elasticities $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively.*

3.1. *For any underlying population structures:*

$$\lambda_{lo}^{max} \le 1 \text{ and } \lambda_{hi}^{max} \le 1 \text{ and } \lambda_{lo}^{max} \le \lambda_{hi}^{min} \Rightarrow S(\pi_0) \ge S(\underline{\mu}), \tag{4.3}$$

3.2. *For any underlying population structures:*

$$\lambda_{lo}^{max} \le 1 \text{ and } \lambda_{hi}^{min} \ge 1 \text{ and } \lambda_{hi}^{max} \le \frac{1}{\lambda_{lo}^{max}} \Rightarrow S(\pi_0) \ge S(\underline{\mu}), \tag{4.4}$$

3.3. *There exists a threshold premium π^* such that:*

$$\lambda_{lo}^{max} \le 1 \text{ and } \lambda_{hi}^{min} > \frac{1}{\lambda_{lo}^{min}} \text{ and } \pi_0 \ge \pi^* \Rightarrow S(\pi_0) \ge S(\underline{\mu}), \tag{4.5}$$

It is easy to see that Theorem 2.1 can be obtained as a special case of Theorems 3.1 and 3.2; while Theorem 2.2 is a special case of Theorem 3.3.

5. General Demand Functions

So far, we have only considered constant demand elasticities, either for all individuals in the population, or for all individuals belonging to a particular risk-group. Iso-elastic demand functions are easy to understand and are also analytically convenient. However, they may also be criticised as being unrealistic. In this section, we use the concept of *arc elasticity of demand* to extend the results in Section 4 to a form applicable to more general demand functions.

The formulation for iso-elastic demand arose from the particular choice of distribution function in Equation 3.7 for the random variable Γ_i (denoting the risk preferences index) for risk-group i. However, the framework developed in Section 2 is general and can be applied to any distribution for the risk preferences index. In this section, we will just assume that Γ_i is a positive continuous random variable[9] with a distribution function:

$$F_{\Gamma_i}(\gamma) = \mathrm{P}\left[\,\Gamma_i \leq \gamma\,\right]. \tag{5.1}$$

Under this general framework, social welfare for a given premium regime $\underline{\pi}$ is given by Lemma 2 (for proof see Appendix D).

Lemma 2. *Suppose there are n risk-groups with risks $\mu_1 < \mu_2 < \cdots < \mu_n$ and any general demand functions. Then for a given premium regime $\underline{\pi}$, for which no risk-group is fully insured, the expression for social welfare is given by:*

$$S(\underline{\pi}) = \sum_{i=1}^{n} p_i\, G_i\left(\frac{\mu_i}{\pi_i}\right) \pi_i + K, \quad \textit{where} \quad G_i(g) = \int_0^g P\left[\,\Gamma_i < \gamma\,\right] d\gamma, \tag{5.2}$$

where the premium regime $\underline{\pi}$ satisfies the equilibrium condition:

$$\sum_{i=1}^{n} p_i\, d_i(\pi_i)\;(\pi_i - \mu_i) = 0, \tag{5.3}$$

and the constant K does not depend on the premium regime under consideration.

Comparing social welfare under pooling to that under full risk classification gives:

$$S(\pi_0) - S(\underline{\mu}) = \sum_{i=1}^{n} p_i\, G_i\left(\frac{\mu_i}{\pi_0}\right) \pi_0 - \sum_{i=1}^{n} p_i\, G_i\left(\frac{\mu_i}{\mu_i}\right) \mu_i. \tag{5.4}$$

where the equilibrium premium π_0 satisfies:

$$\sum_{i=1}^{n} p_i\, d_i(\pi_0)\;(\pi_0 - \mu_i) = 0. \tag{5.5}$$

Using the notations involving risk-premium ratios, $v_i = \mu_i/\pi_0$, we get:

$$S(\pi_0) \gtreqless S(\underline{\mu}) \Leftrightarrow \sum_{i=1}^{n} p_i\,\left[\,G_i\left(v_i\right) - v_i\, G_i\left(1\right)\,\right] \gtreqless 0. \tag{5.6}$$

[9]The derivations in this section can also be suitably adapted for any positive discrete random variable.

To make analytical progress with the general relationship in Equation 5.6, we need to establish a connection between general demand elasticity functions, $\epsilon_i(\cdot)$, and general distribution functions for the risk preferences index, $F_{\Gamma_i}(\cdot)$. The link arises from Equations 2.5 and 2.15, reproduced below with appropriate adaptation for risk-group i:

$$d_i(\pi) = \tau_i \exp\left[-\int_{\mu_i}^{\pi} \epsilon_i(s)\, d\log s\right], \tag{5.7}$$

$$d_i(\pi) = \mathrm{P}\left[\Gamma_i < \frac{\mu_i}{\pi}\right] = \mathrm{P}\left[\Gamma_i \leq v\right], \quad \text{where } v = \frac{\mu_i}{\pi}. \tag{5.8}$$

Note the distinction between v_i (earlier in the paper) and v for risk-group i: v_i is the risk-premium ratio at the equilibrium premium π_0, whereas v is the risk-premium ratio as a function of premium π.

We now need the concept of *arc elasticity of demand* (Vazquez (1995)), defined as:

$$\lambda_i(v) = \frac{\int_{\mu_i}^{\pi} \epsilon_i(s)\, d\log s}{\int_{\mu_i}^{\pi} d\log s}, \text{ for } i = 1, 2, \ldots, n. \tag{5.9}$$

which can be interpreted as the weighted average of (point) elasticity for risk-group i, $\epsilon_i(s)$, over the arc of the demand curve from premium μ_i to premium π, where the weights are the log premiums.

Using the concept of arc elasticity of demand, Equation 5.8 can be written as:

$$d_i(\pi) = \mathrm{P}\left[\Gamma_i \leq v\right] = \tau_i \exp\left[-\lambda_i(v) \int_{\mu_i}^{\pi} d\log s\right] = \tau_i \left(\frac{\mu_i}{\pi}\right)^{\lambda_i(v)} = \tau_i\, v^{\lambda_i(v)}. \tag{5.10}$$

and the equilibrium condition in Equation 5.5 as:

$$\sum_{i=1}^{n} p_i\, \tau_i\, v_i^{\lambda_i(v_i)+1} = \sum_{i=1}^{n} p_i\, \tau_i\, v_i^{\lambda_i(v_i)}, \text{ as } d_i(\pi_0) = \tau_i\, v_i^{\lambda_i(v_i)}. \tag{5.11}$$

Now consider a *hypothetical* population with the same probabilities of loss, i.e. $\mu_1 < \mu_2 < \cdots < \mu_n$, as in the actual population. But suppose that in the hypothetical population, demand for insurance is iso-elastic with constant elasticity parameters set at values $\lambda_1(v_1), \lambda_2(v_2), \ldots, \lambda_n(v_n)$ respectively. Then all the results obtained in Section 4 are applicable for the hypothetical population with iso-elastic demand. This creates an avenue for extending the results for iso-elastic demand to general demand functions.

Specifically, if the relevant conditions of iso-elastic demand functions given in Theorem 3 of Section 4 apply for the hypothetical population, we know that pooling increases social welfare as compared to full risk classification. In that case, Equation 3.13 implies that for the hypothetical population:

$$\sum_{i=1}^{n} p_i \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i^{\lambda_i(v_i)+1} - v_i\right] \geq 0. \tag{5.12}$$

However, insurance demand of the *actual* population is not necessarily iso-elastic. But, interestingly, by construction, the equilibrium condition in Equation 5.11 is the same for both the hypothetical population and the actual population, i.e. the pooled equilibrium premium, π_0, will be the same under both set-ups.

Now for the higher risk-groups, i.e. for those risk-groups for which $\mu_i > \pi_0$, it is shown in Lemma 4 in Appendix E that if the demand elasticity, $\epsilon_i(\pi)$, is either increasing or iso-elastic as a function of premium π, then:

$$G_i(v_i) - v_i G_i(1) \geq \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i^{\lambda_i(v_i)+1} - v_i\right]. \tag{5.13}$$

In other words: for the higher risk-groups, under the assumption of increasing or iso-elastic demand elasticities, the increase in social welfare in the actual population when we move to pooling is *higher* than that in the hypothetical population.

Conversely, for the lower risk-groups, i.e. for those risk-groups for which $\mu_i \leq \pi_0$, it is shown in Lemma 5 in Appendix E that if the demand elasticity, $\epsilon_i(\pi)$, is either decreasing or iso-elastic as a function of premium π, then:

$$v_i G_i(1) - G_i(v_i) \leq \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i - v_i^{\lambda_i(v_i)+1}\right]. \tag{5.14}$$

In other words: for the lower risk-groups, under the assumption of decreasing or iso-elastic demand elasticities, the fall in social welfare in the actual population when we move to pooling is *lower* than that in the hypothetical population.

Putting Equations 5.13 and 5.14 together, we get the following expression for the increase in social welfare in the actual population when we move to pooling:

$$\sum_{i=1}^{n} p_i \left[G_i\left(v_i\right) - v_i\, G_i\left(1\right)\right] \tag{5.15}$$

$$= \sum_{\mu_i > \pi_0} p_i \left[G_i\left(v_i\right) - v_i\, G_i\left(1\right)\right] - \sum_{\mu_i \leq \pi_0} p_i \left[v_i\, G_i\left(1\right) - G_i\left(v_i\right)\right], \tag{5.16}$$

$$\geq \sum_{\mu_i > \pi_0} p_i \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i^{\lambda_i(v_i)+1} - v_i\right] - \sum_{\mu_i \leq \pi_0} p_i \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i - v_i^{\lambda_i(v_i)+1}\right], \tag{5.17}$$

$$= \sum_{i=1}^{n} p_i \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i^{\lambda_i(v_i)+1} - v_i\right]. \tag{5.18}$$

This implies that if the actual population is such that the hypothetical population satisfies the relevant conditions of iso-elastic demand functions given in Theorem 3.1 of Section 4, then pooling gives higher social welfare than full risk classification in the actual population. The following theorem outlines the required conditions in the actual population.

Theorem 4. *Suppose there are n risk-groups with risks $\mu_1 < \mu_2 < \cdots \mu_n$. If the insurance demand elasticities have the following properties over their respective ranges from μ_i to the pooled premium π_0:*

(i) for each lower risk-group, demand elasticity is either decreasing or iso-elastic as a function of premium;

(ii) for each higher risk-group, demand elasticity is either increasing or iso-elastic as a function of premium;

(iii) risk-groups with higher risks have higher arc elasticities of demand; and

(iv) demand elasticities do not exceed 1

then pooling increases social welfare as compared against full risk classification.

Theorem 4 thus partly relaxes the iso-elasticity condition on higher risk-groups in Theorem 3.1. Specifically, condition (ii) allows higher risk-groups to have either iso-elastic *or increasing* demand elasticities (as a function of premium), *provided that* they also have higher *arc elasticities* than all lower risk-groups (condition (iii)) and their demand elasticities do not exceed 1 (condition (iv)).

Technically, Theorem 4 also partly relaxes the iso-elasticity condition on lower risk-groups. Specifically, condition (i) allows lower risk-groups to have either iso-elastic *or decreasing* demand elasticities (as a function of premium). However, as discussed previously, demand elasticities are more likely to be increasing as a function of premium. So, for all practical purposes, condition (i) amounts to a restriction to iso-elastic demand functions.

We emphasise that the conditions presented in Theorem 4 are sufficient, but *not* necessary. In fact, experimentation using simple functions reveals that pooling can sometimes increase social welfare even where lower risk-groups have increasing demand elasticity (as a function of premium), as long as the marginal increase in their demand elasticities does not exceed a certain threshold which depends on the high risk-groups' demand elasticities. However, we do not include these results here as they are not generic and apply to specific analytic forms of demand elasticity functions.

6. Discussion

6.1 Summary and Empirical Comparisons

The results obtained in this paper give sufficient conditions for social welfare to be higher under pooling than under full risk classification. They can be summarised as follows.

(a) Theorem 1 for iso-elastic demand (common elasticity for all risk-groups) requires only that the common demand elasticity is less than 1.
(b) Theorems 2 and 3 for iso-elastic demand (different elasticities for different risk-groups) requires that all higher risk-groups' demand elasticities are higher than all lower risk-groups' demand elasticities, *and* all demand elasticities are less than 1. They also provide sufficient conditions when higher risk-groups' demand elasticities exceed 1, as long as all lower risk-groups' demand elasticities are less than 1.
(c) Theorem 4 then uses the concept of arc elasticity of demand to extend the results in a form applicable to more general demand functions.

The conditions above are stringent because they are sufficient for *any* population structures and relative risks. But the conditions are *not* necessary, and where they are not fully satisfied, social welfare under pooling may still be higher than under risk-differentiated premiums for some combinations of population structures and demand elasticities.

Given that the conditions all relate to demand elasticities, an obvious question is: what elasticities do we typically observe? Table 1 shows some relevant empirical estimates. It can be seen that most estimates are of magnitude significantly less than 1. Whilst the various contexts in which these estimates were made may not correspond closely to the set-up in this paper, the figures are at least suggestive of the possibility that insurance demand elasticities may often be less than 1. This tends to affirm the possibility that social welfare in some insurance markets could be higher under pooling than under full risk classification.

Table 1: Estimates of demand elasticity for various insurance markets.

Market & country	Demand elasticities [a]	Authors
Term life insurance, USA	0.66	Viswanathan et al. (2006)
Yearly renewable term life, USA	0.4 to 0.5	Pauly et al. (2003)
Whole life insurance, USA	0.71 to 0.92	Babbel (1985)
Health insurance, USA	0 to 0.2	Chernew et al. (1997), Blumberg et al. (2001), Buchmueller and Ohri (2006)
Health insurance, Australia	0.35 to 0.50	Butler (1999)
Farm crop insurance, USA	0.32 to 0.73	Goodwin (1993)

[a] Estimates in empirical papers are generally given as negative values, but we have presented the absolute values here for consistency with the definition of demand elasticity used in this paper.

6.2 Comparison with Loss Coverage

The results for social welfare can be compared with the analogous results for loss coverage in Hao et al. (2018). That paper shows that on the loss coverage criterion, pooling is sure to be beneficial in the both the green and dotted regions in Figure 3. The additional dotted area where pooling is sure to increase loss coverage (but increases social welfare only subject to further conditions) arises because the loss coverage criterion focuses on compensation of losses for the population as a whole, and places no weight on the premium cross-subsidies implied by pooling; on the other hand, social welfare takes account of the premium cross-subsidies. For moderate dispersion of elasticities (and hence utility functions), taking account of premium cross-subsidies typically does not change the ranking of pooling versus full risk classification. But with large dispersion of elasticities (and hence utility functions) – in particular, $\lambda_2 \gg \lambda_1$, that is where high risks have much higher demand elasticities than low risks – then pooling may be beneficial in terms of loss coverage, but not in terms of social welfare. However, $\lambda_2 \gg \lambda_1$ is probably an unrealistic parameterisation; for more realistic parameters (e.g. all elasticities not much more than 1), loss coverage and social welfare usually give the same ranking of pooling versus full risk classification.[10]

6.3 Comparison with other authors

The results can also be compared with those of Hoy (2006), who finds that utilitarian welfare is increased by pooling, provided only that the fraction of high risks is sufficiently small. Hoy (2006) assumes a utility function which is state-dependent (i.e. lower utility in the loss state) and uniformly risk-averse for the whole population; this leads all individuals to buy insurance under either pooling or full risk classification, albeit the pooling contract provides only partial insurance.[11] When pooling is mandated, there is (i) a loss in efficiency because the pooling contract offers only partial insurance (ii) a redistribution from low risks (previously better off, because they paid lower premiums) to high risks. Behind the veil of ignorance, effect (i) reduces welfare, but effect (ii) increases welfare. For a sufficiently small high-risk fraction, effect (ii) dominates (i.e. for a risk-averse utility function, expected utility behind the veil of ignorance is always increased by a sufficiently small redistribution towards the previously worse off).

In contrast, we allow for a distribution of utility functions in the population, leading to an equilibrium where not all individuals purchase insurance at an actuarially fair price. In our model, if we pool a very small high-risk population with high elasticity with a large low-risk population with low elasticity, many of the high risks who now choose to

[10] In the equivalent of the present paper's Figure 3 in Hao et al. (2018), the boundary curve between "pooling better" and "full risk classification better" on the loss coverage criterion follows a very similar arc inside the unit square to the red boundary curve in Figure 3.

[11] The partial-cover pooling contract is that predicted by the anticipatory (E2) equilibrium concept in Wilson (1977).

participate at the (cheap) pooled price have low risk aversion, so their gain in utility from participating is relatively small. On the other hand, the low-elasticity lower risks' loss in utility (from either leaving the market or paying the (expensive) pooled price) is relatively large. Therefore overall, pooling might not be advantageous, even with a very small high-risk fraction. In Figure 3, this is represented by the curvature of the dashed boundary for $\alpha_1 = 0.99$ (i.e. very few high risks) back towards the vertical axis for $\lambda_2 \gg \lambda_1$.

But this feature in our model probably has little practical significance, because $\lambda_2 \gg \lambda_1$ is not a realistic parameterisation. For more typical parameter values (e.g. $\lambda_1 < \lambda_2 < 1$), the relative position of the two curves in Figure 3 suggests that reducing the size of the high risk-group makes pooling slightly more likely to be beneficial (in the sense that pooling gives higher social welfare for a slightly wider range of (λ_1, λ_2) parameter values). This is more in accordance with (albeit not the same as) Hoy's result.

7. Conclusions

This paper has evaluated the welfare effects of bans on risk classification, in circumstances where institutional or regulatory factors lead insurers to pool all risks at a common price. Such bans have both efficiency and equity effects. Depending on the distribution of utility functions in the population, utilitarian social welfare can increase or decrease.

The distribution of utility functions in the population influences social welfare through two mechanisms. First, utility functions determine individuals' insurance purchasing decisions, which determine the insurance demand curve and hence the equilibrium price of insurance when all risks are pooled. Second, utility functions determine the utilities which individuals assign to their outcomes *given* an equilibrium pooled price.

Because the distribution of utility functions and the insurance demand function are mutually implicative, the distribution of utility functions across the population is completely characterised by demand elasticities. Hence in this paper, demand elasticity functions have been used to specify both demand and (implicitly) the distribution of utility functions in the population.

This paper has stated sufficient conditions on demand elasticities of higher and lower risks which ensure that social welfare will be higher under pooling than under fully risk-differentiated premiums. The conditions were stated first for iso-elastic demand with a single elasticity parameter; then for iso-elastic demand with different elasticity parameters for different risk-groups; and then generalised in a form applicable to other demand functions using the concept of arc elasticity. The conditions for higher social welfare under pooling encompass many plausible combinations of higher and lower risks' demand elasticities, particularly in scenarios where all demand elasticities are less than 1.

Acknowledgement

Indradeb Chatterjee thanks Radfall Charitable Trust for research funding. Some preliminary investigations on the subject matter of this paper were carried out by Dr MingJie Hao in her PhD thesis.

References

G.A. Akerlof. The market for lemons: quality uncertainty and the market mechanism. The Quarterly Journal of Economics, 84:488–500, 1970.

American Council of Life Insurers. 2019 life insurers factbook, October 2019. http://www.acli.org (accessed 27 January 2019).

D. Babbel. The price elasticity of demand for whole life insurance. Journal of Finance, 40(1): 225–239, 1985.

J.W. Bailey. Utilitarianism, institutions and justice. Oxford University Press, 1997.

L. Blumberg, L. Nichold, and J. Banthin. Worker decisions to purchase health insurance. International Journal of Health Care Finance and Economics, 1:305–325, 2001.

T.C. Buchmueller and S. Ohri. Health insurance take-up by the near-elderly. Health Services Research, 41:2054–2073, 2006.

J.R. Butler. Estimating elasticities of demand for private health insurance in australia. Working Paper No. 43. National Centre for Epidemiology and Population Health ,Australian National University, Canberra, 1999.

M. Chernew, K. Frick, and C. McLaughlin. The demand for health insurance coverage by low-income workers: Can reduced premiums achieve full coverage? Health Services Research, 32: 453–470, 1997.

K. Crocker and A. Snow. The efficiency effects of categorical discrimination in the insurance industry. Journal of Political Economy, 94:321–344, 1986.

Z. Cvetkovski. Inequalities: Theorems, Techniques and Selected Problems. Springer, 2012.

J. Friedland. Fundamentals of general insurance actuarial analysis. Society of Actuaries, 2013.

B.K. Goodwin. An empirical analysis of the demand for multiple peril crop insurance. American Journal of Agricultural Economics, 75(2):424–434, 1993.

R.J Gray and S. Pitts. Risk modelling in general insurance. Cambridge University Press, 2012.

M. Hao, A.S. Macdonald, P. Tapadar, and R.G. Thomas. Insurance loss coverage and demand elasticities. Insurance: Mathematics and Economics, 79:15–25, 2018. https://doi.org/10.1016/j.insmatheco.2017.12.002.

M. Hao, A.S. Macdonald, P. Tapadar, and R.G. Thomas. Insurance loss coverage and social welfare. Scandinavian Actuarial Journal, 2019:113–128, 2019. https://doi.org/10.1080/03461238.2018.1513865.

G.H. Hardy, J. E. Littlewood, and G. Pólya. Inequalities. Cambridge University Press, 1988.

J.C. Harsanyi. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. Journal of Political Economy, 63:309–321, 1955.

M. Hoy. Risk classification and social welfare. Geneva Papers on Risk and Insurance, 31:245–269, 2006.

P. Kumaraswamy. A generalized probability density function for double-bounded random processes. Journal of Hydrology, 46:79–88, 1980.

LIMRA. Facts about life 2019, September 2019. http://www.limra.com (accessed 27 Janaury 2020).

R. Nozick. Anarchy, state and utopia. Basic Books, N.Y., 1974.

P. Parodi. Pricing in general insurance. Chapman and Hall, 2014.

M.V. Pauly, K.H. Withers, K.S. Viswanathan, J. Lemaire, J.C. Hershey, K. Armstrong, and D.A. Asch. Price elasticity of demand for term life insurance and adverse selection. NBER Working Paper (9925), 2003.

M.K. Polborn, M. Hoy, and A. Sadanand. Advantageous effects of regulatory adverse selection in the life insurance market. The Economic Journal, 116:327–354, 2006.

C. Rothschild. The efficiency of categorical discrimination in insurance markets. Journal of Risk and Insurance, 78:267–285, 2011.

M. Rothschild and J. Stiglitz. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. Quarterly Journal of Economics, 90(4):630–649, 1976.

A. Vazquez. A note on arc elasticity of demand. Estudios Economicos, 10(2):221–228, 1995.

K.S. Viswanathan, J. Lemaire, K. K. Withers, K. Armstrong, A. Baumritter, J. Hershey, M. Pauly, and D.A. Asch. Adverse selection in term life insurance purchasing due to the brca 1/2 genetic test and elastic demand. Journal of Risk and Insurance, 74:65–86, 2006.

C. Wilson. A model for insurance with incomplete information. Journal of Economic Theory, 16:167–207, 1977.

Appendices

A. Expressions for Social Welfare Under Iso-elastic Demand

Lemma 1. *Suppose there are n risk-groups with risks $\mu_1 < \mu_2 < \cdots < \mu_n$ with iso-elastic demand elasticities $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively, then for a given premium regime $\underline{\pi}$, the expression for social welfare is given by:*

$$S(\underline{\pi}) = \sum_{i=1}^{n} p_i \, \tau_i \, \frac{1}{(\lambda_i + 1)} \left(\frac{\mu_i}{\pi_i}\right)^{\lambda_i + 1} \pi_i + K, \tag{3.8}$$

where the premium regime $\underline{\pi}$ satisfies the equilibrium condition:

$$\sum_{i=1}^{n} p_i \tau_i \left(\frac{\mu_i}{\pi_i}\right)^{\lambda_i} (\pi_i - \mu_i) = 0, \tag{3.9}$$

and the constant K does not depend on the premium regime under consideration.

Proof. The equilibrium condition follows directly by inserting the specific expression for iso-elastic insurance demand in Equation 2.21.

Now recall that, given a risk-group i, insurance is purchased when $\Gamma_i < \mu_i/\pi_i$ (a subscript i in Γ_i is used to denote the random variable specific to risk-group i). Hence:

$$[\,Q \mid \text{Risk-group } i\,] = I\left[\Gamma_i < \frac{\mu_i}{\pi_i}\right] \Rightarrow \mathrm{E}\,[\,Q \mid \text{Risk-group } i\,] = \mathrm{P}\left[\Gamma_i < \frac{\mu_i}{\pi_i}\right] = d_i(\pi_i), \tag{A.1}$$

where $I(\cdot)$ is the indicator function.

Using the expression for social welfare as given in Equation 2.17 we have:

$$S(\underline{\pi}) = \mathrm{E}\,[\,Q\,(X - \Pi\,\Gamma)\,] + K = \mathrm{E}\,[\,Q\,X\,] - \mathrm{E}\,[\,Q\,\Pi\,\Gamma\,] + K. \tag{A.2}$$

Evaluating each of these terms separately:

$$\mathrm{E}\,[\,Q\,X\,] = \sum_{i=1}^{n} \mathrm{P}[\text{Risk-group } i]\,\mathrm{E}\,[\,Q\,X \mid \text{Risk-group } i\,] \tag{A.3}$$

$$= \sum_{i=1}^{n} p_i\,\mathrm{E}\,[\,Q \mid \text{Risk-group } i\,]\;\mathrm{E}\,[\,X \mid \text{Risk-group } i\,], \quad \text{using Assumption 7,} \tag{A.4}$$

$$= \sum_{i=1}^{n} p_i\, d_i(\pi_i)\,\mu_i, \tag{A.5}$$

$$= \sum_{i=1}^{n} p_i\, \tau_i \left(\frac{\mu_i}{\pi_i}\right)^{\lambda_i} \mu_i, \tag{A.6}$$

$$= \sum_{i=1}^{n} p_i\, \tau_i \left(\frac{\mu_i}{\pi_i}\right)^{\lambda_i+1} \pi_i, \tag{A.7}$$

and:

$$\mathrm{E}\,[\,Q\,\Pi\,\Gamma\,] = \sum_{i=1}^{n} \mathrm{P}[\text{Risk-group } i]\, \mathrm{E}\,[\,Q\,\Pi\,\Gamma\ \mid \text{Risk-group } i\,] \tag{A.8}$$

$$= \sum_{i=1}^{n} p_i\, \mathrm{E}\left[I \left[\Gamma_i < \frac{\mu_i}{\pi_i} \right] \Gamma_i \right] \pi_i, \tag{A.9}$$

$$= \sum_{i=1}^{n} p_i \left[\int_0^{\frac{\mu_i}{\pi_i}} \gamma \tau_i \lambda_i \gamma^{\lambda_i - 1} d\gamma \right] \pi_i, \quad \text{using the distribution of } \Gamma_i \text{ in Equation 3.7,} \tag{A.10}$$

$$= \sum_{i=1}^{n} p_i\, \tau_i \frac{\lambda_i}{(\lambda_i+1)} \left(\frac{\mu_i}{\pi_i}\right)^{\lambda_i+1} \pi_i. \tag{A.11}$$

Putting these together, we have:

$$S(\underline{\pi}) = \sum_{i=1}^{n} p_i\, \tau_i \frac{1}{(\lambda_i+1)} \left(\frac{\mu_i}{\pi_i}\right)^{\lambda_i+1} \pi_i + K, \tag{A.12}$$

where $K = E[1 - X]$ does not depend on the premium regime under consideration. □

B. Same Iso-elastic Demand Elasticity and Social Welfare

Theorem 1. *Suppose there are n risk-groups with risks $\mu_1 < \mu_2 < \cdots < \mu_n$ with the same positive constant demand elasticity λ for all risk-groups. Then:*

$$\lambda \lesseqqgtr 1 \Rightarrow S(\pi_0) \gtreqqless S(\underline{\mu}). \tag{3.14}$$

Proof. Using the construction involving risk-premium ratios, $v_i = \mu_i/\pi_0$, we observe that, under the assumption of the same constant demand elasticity, λ, for all risk-groups, the equilibrium condition in Equation 3.10 simply becomes:

$$\sum_{i=1}^{n} \alpha_i v_i^{\lambda+1} = \sum_{i=1}^{n} \alpha_i v_i^{\lambda}. \tag{B.1}$$

And the condition comparing social welfare under pooling against that under the full risk classification regime in Equation 3.12 simplifies to:

$$S(\pi_0) \gtreqqless S(\underline{\mu}) \Leftrightarrow \sum_{i=1}^{n} \frac{\alpha_i\, v_i^{\lambda+1}}{\lambda+1} \gtreqqless \sum_{i=1}^{n} \frac{\alpha_i\, v_i}{\lambda+1} \Leftrightarrow \sum_{i=1}^{n} \alpha_i\, v_i^{\lambda+1} \gtreqqless \sum_{i=1}^{n} \alpha_i\, v_i. \tag{B.2}$$

We will consider the three cases $\lambda = 1$, $0 < \lambda < 1$ and $\lambda > 1$ separately:

Case: $\lambda = 1$: Due to the equilibrium condition in Equation B.1, for $\lambda = 1$:

$$\sum_{i=1}^{n} \alpha_i v_i^{\lambda+1} = \sum_{i=1}^{n} \alpha_i v_i^{\lambda} = \sum_{i=1}^{n} \alpha_i v_i \Rightarrow S(\pi_0) = S(\underline{\mu}). \tag{B.3}$$

Case: $0 < \lambda < 1$: (Weighted) Hölder's inequality (Hardy et al. (1988); Cvetkovski (2012)) states:

(Weighted) Hölder's inequality. *Let $a_1, a_2, \ldots, a_n; b_1, b_2 \ldots, b_n; m_1, m_2, \ldots, m_n$ be three sequences of positive reals numbers and If $p, q > 1$ be such that $1/p + 1/q = 1$, Then:*

$$\left(\sum_{i=1}^{n} m_i\, a_i^p\right)^{1/p} \left(\sum_{i=1}^{n} m_i\, b_i^q\right)^{1/q} \geq \sum_{i=1}^{n} m_i\, a_i\, b_i. \tag{B.4}$$

Equality occurs if and only if $\frac{a_1^p}{b_1^q} = \frac{a_2^p}{b_2^q} = \cdots = \frac{a_n^p}{b_n^q}$.

Setting $1/p = \lambda$, $1/q = 1 - \lambda$, $a_i = v_i^{\lambda^2}$, $b_i = v_i^{1-\lambda^2}$ and $m_i = \alpha_i$; and noting that the ratios, $a_i^p / b_i^q = 1/v_i$, are not constant (unless all $v_i = 1$), (weighted) Hölder's inequality gives:

$$\left[\sum_{i=1}^{n} \alpha_i \left(v_i^{\lambda^2}\right)^{\frac{1}{\lambda}}\right]^{\lambda} \left[\sum_{i=1}^{n} \alpha_i \left(v_i^{1-\lambda^2}\right)^{\frac{1}{1-\lambda}}\right]^{1-\lambda} > \sum_{i=1}^{n} \alpha_i \, v_i^{\lambda^2} \, v_i^{1-\lambda^2}, \tag{B.5}$$

$$\Rightarrow \left[\sum_{i=1}^{n} \alpha_i \, v_i^{\lambda}\right]^{\lambda} \left[\sum_{i=1}^{n} \alpha_i \, v_i^{1+\lambda}\right]^{1-\lambda} > \sum_{i=1}^{n} \alpha_i \, v_i, \tag{B.6}$$

$$\Rightarrow \sum_{i=1}^{n} \alpha_i \, v_i^{1+\lambda} > \sum_{i=1}^{n} \alpha_i \, v_i, \quad \text{by the equilibrium condition in Equation B.1,} \tag{B.7}$$

$$\Rightarrow S(\pi_0) > S(\underline{\mu}), \quad \text{by the social welfare condition in Equation B.2.} \tag{B.8}$$

Case: $\lambda > 1$: Young's inequality (Hardy et al. (1988); Cvetkovski (2012)) states that:

> **Young's inequality.** *For $a, b > 0$ and $p, q > 1$ such that $1/p + 1/q = 1$:*
>
> $$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{B.9}$$
>
> *Equality occurs if and only if $a^p = b^q$.*

Setting $p = \lambda$, $q = \frac{\lambda}{\lambda-1}$, $a = v_i^{\frac{1}{\lambda}}$, $b = v_i^{\lambda - \frac{1}{\lambda}}$ and noting that $a^p \neq b^q$ unless $v_i = 1$, Young's inequality gives:

$$v_i^{\frac{1}{\lambda}} v_i^{\lambda - \frac{1}{\lambda}} < \frac{1}{\lambda} v_i^{\frac{1}{\lambda}\lambda} + \frac{\lambda - 1}{\lambda} v_i^{(\lambda - \frac{1}{\lambda})\frac{\lambda}{\lambda-1}}, \tag{B.10}$$

$$\Rightarrow v_i^{\lambda} < \frac{1}{\lambda} v_i + \frac{\lambda - 1}{\lambda} v_i^{\lambda+1}, \tag{B.11}$$

$$\Rightarrow \sum_{i=1}^{n} \alpha_i \, v_i^{\lambda} < \frac{1}{\lambda} \sum_{i=1}^{n} \alpha_i \, v_i + \frac{\lambda - 1}{\lambda} \sum_{i=1}^{n} \alpha_i \, v_i^{\lambda+1}, \tag{B.12}$$

$$\Rightarrow \sum_{i=1}^{n} \alpha_i \, v_i^{\lambda+1} < \sum_{i=1}^{n} \alpha_i \, v_i, \quad \text{by the equilibrium condition in Equation B.1,} \tag{B.13}$$

$$\Rightarrow S(\pi_0) < S(\underline{\mu}), \quad \text{by the social welfare condition in Equation B.2.} \tag{B.14}$$

□

C. Different Iso-elastic Demand Elasticities and Social Welfare

In this section, we prove Theorem 3. As discussed in Section 4, Theorem 2 is a special case of Theorem 3.

Theorem 3. *Suppose there are n risk-groups with risks $\mu_1 < \mu_2 < \cdots < \mu_n$ with iso-elastic demand elasticities $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively.*

3.1. *For any underlying population structures:*

$$\lambda_{lo}^{max} \leq 1 \text{ and } \lambda_{hi}^{max} \leq 1 \text{ and } \lambda_{lo}^{max} \leq \lambda_{hi}^{min} \Rightarrow S(\pi_0) \geq S(\underline{\mu}), \tag{4.3}$$

3.2. *For any underlying population structures:*

$$\lambda_{lo}^{max} \leq 1 \text{ and } \lambda_{hi}^{min} \geq 1 \text{ and } \lambda_{hi}^{max} \leq \frac{1}{\lambda_{lo}^{max}} \Rightarrow S(\pi_0) \geq S(\underline{\mu}), \tag{4.4}$$

3.3. *There exists a threshold premium π^* such that:*

$$\lambda_{lo}^{max} \leq 1 \text{ and } \lambda_{hi}^{min} > \frac{1}{\lambda_{lo}^{min}} \text{ and } \pi_0 \geq \pi^* \Rightarrow S(\pi_0) \geq S(\underline{\mu}), \tag{4.5}$$

Proof. (of Theorem 3.1) The proof is presented in the following steps:

Step 1: If $a > 0$ and $0 < b \leq 1$, then since Arithmetic Mean $\geq$ Geometric Mean:

$$(1-b)a^{b+1} + ba^b \geq a^{(b+1)(1-b)} \times a^{b^2} = a \Rightarrow \left(\frac{a^{b+1}-a}{b}\right) \geq \left(a^{b+1} - a^b\right). \tag{C.1}$$

Step 2: As $v_i > 0$ and $0 < \lambda_i \leq 1$ for all risk-groups, using Step 1, we get:

$$\sum_{i=1}^{n} \alpha_i \frac{v_i^{\lambda_i+1} - v_i}{\lambda_i} \geq \sum_{i=1}^{n} \alpha_i \left(v_i^{\lambda_i+1} - v_i^{\lambda_i}\right) = \sum_{i=1}^{n} \alpha_i v_i^{\lambda_i+1} - \sum_{i=1}^{n} \alpha_i v_i^{\lambda_i} = 0, \tag{C.2}$$

by equilibrium condition in Equation 3.10.

Step 3: Using Step 2, and separating out the terms involving $v_i > 1$ from $v_i \leq 1$ we get:

$$\sum_{i:\, v_i>1} \alpha_i \frac{v_i^{\lambda_i+1} - v_i}{\lambda_i} \geq \sum_{i:\, v_i \leq 1} \alpha_i \frac{v_i - v_i^{\lambda_i+1}}{\lambda_i} \geq 0. \tag{C.3}$$

Step 4: As $0 < x \leq y \Rightarrow \frac{x}{x+1} \leq \frac{y}{y+1}$, if $0 < v_j \leq 1 \leq v_k$, for some j and k, then

$$\lambda_j \leq \lambda_{lo}^{max} \leq \lambda_{hi}^{min} \leq \lambda_k \Rightarrow \frac{\lambda_j}{\lambda_j+1} \leq \frac{\lambda_{lo}^{max}}{\lambda_{lo}^{max}+1} \leq \frac{\lambda_{hi}^{min}}{\lambda_{hi}^{min}+1} \leq \frac{\lambda_k}{\lambda_k+1}. \tag{C.4}$$

Step 5: Using Steps 3 and 4, we get:

$$\sum_{i:\, v_i>1} \frac{\alpha_i}{\lambda_i+1}\left[v_i^{\lambda_i+1} - v_i\right] = \sum_{i:\, v_i>1} \alpha_i \frac{\lambda_i}{\lambda_i+1} \frac{v_i^{\lambda_i+1} - v_i}{\lambda_i}, \tag{C.5}$$

$$\geq \frac{\lambda_{hi}^{min}}{\lambda_{hi}^{min}+1} \sum_{i:\, v_i>1} \alpha_i \frac{v_i^{\lambda_i+1} - v_i}{\lambda_i}, \tag{C.6}$$

$$\geq \frac{\lambda_{lo}^{max}}{\lambda_{lo}^{miax}+1} \sum_{i:\, v_i \leq 1} \alpha_i \frac{v_i - v_i^{\lambda_i+1}}{\lambda_i}, \tag{C.7}$$

$$\geq \sum_{i:\, v_i \leq 1} \alpha_i \frac{\lambda_i}{\lambda_i+1} \frac{v_i - v_i^{\lambda_i+1}}{\lambda_i}, \tag{C.8}$$

$$= \sum_{i:\, v_i \leq 1} \frac{\alpha_i}{\lambda_i+1}\left[v_i - v_i^{\lambda_i+1}\right] \tag{C.9}$$

Hence by Equation 3.13, $S(\pi_0) \geq S(\underline{\mu})$. □

Proof. (of Theorem 3.2) The proof is presented in the following steps:

Step 1: Let $0 < a \leq 1$, $b \geq a$ such that $a\, b \leq 1$ and function $g(v)$ be defined as:

$$g(v) = (b-a)v^a + (a+1)v^{a-1} - (b+1), \text{ for } v > 0. \tag{C.10}$$

If $a = 1$, then $b = 1$ (as $b \geq a$ and $ab \leq 1$), in which case: $g(v) = 0$ for $v > 0$.

If $0 < a < 1$ i.e. $(a-1) < 0$, $\lim_{v \to 0+} g(v) = +\infty$, $g(1) = 0$ and:

$$g'(v) = (b-a)\, a\, v^{a-2} \left[v - \frac{1-a^2}{ab-a^2} \right] < 0, \text{ for } 0 < v < 1 \text{ as } ab \leq 1. \tag{C.11}$$

So $g(v)$ is a non-negative decreasing function over $0 < v \leq 1$. Hence $g(v) \geq 0$ for $0 < v \leq 1$.

Step 2: For $v_i \leq 1$, set $a = \lambda_i$ and $b = \lambda_{hi}^{max} \Rightarrow ab = \lambda_i\, \lambda_{hi}^{max} \leq \lambda_{lo}^{max}\, \lambda_{hi}^{max} \leq 1$. By Step 1:

$$(\lambda_{hi}^{max} - \lambda_i)v_i^{\lambda_i} + (\lambda_i + 1)v_i^{\lambda_i - 1} - (\lambda_{hi}^{max} + 1) \geq 0. \tag{C.12}$$

Rearranging and multiplying by $\alpha_i\, v_i$ on both sides, we get:

$$\frac{\alpha_i}{\lambda_{hi}^{max} + 1} \left[v_i^{\lambda_i} - v_i^{\lambda_i + 1} \right] \geq \frac{\alpha_i}{\lambda_i + 1} \left[v_i - v_i^{\lambda_i + 1} \right]. \tag{C.13}$$

As this holds for all $v_i \leq 1$, summing over all such risk-groups leads to:

$$\frac{1}{\lambda_{hi}^{max} + 1} \sum_{i:\, v_i \leq 1} \alpha_i \left[v_i^{\lambda_i} - v_i^{\lambda_i + 1} \right] \geq \sum_{i:\, v_i \leq 1} \frac{\alpha_i}{\lambda_i + 1} \left[v_i - v_i^{\lambda_i + 1} \right]. \tag{C.14}$$

Step 3: For all risk-groups with $v_i > 1$, $\lambda_i \geq 1$ (since $\lambda_{hi}^{min} \geq 1$). So:

$$\sum_{i:\, v_i > 1} \frac{\alpha_i}{\lambda_i + 1} \left[v_i^{\lambda_i + 1} - v_i \right] \geq \frac{1}{\lambda_{hi}^{max} + 1} \sum_{i:\, v_i > 1} \alpha_i \left[v_i^{\lambda_i + 1} - v_i \right], \text{ as } \lambda_{hi}^{max} \geq \lambda_i \tag{C.15}$$

$$\geq \frac{1}{\lambda_{hi}^{max} + 1} \sum_{i:\, v_i > 1} \alpha_i \left[v_i^{\lambda_i + 1} - v_i^{\lambda_i} \right], \text{ as } v_i > 1 \text{ and } \lambda_i \geq 1 \tag{C.16}$$

$$= \frac{1}{\lambda_{hi}^{max} + 1} \sum_{i:\, v_i \leq 1} \alpha_i \left[v_i^{\lambda_i} - v_i^{\lambda_i + 1} \right], \text{ by Equation 3.11.} \tag{C.17}$$

Step 4: Combining Steps 2 and 3, we get:

$$\sum_{i:\, v_i > 1} \frac{\alpha_i}{\lambda_i + 1} \left[v_i^{\lambda_i + 1} - v_i \right] \geq \frac{1}{\lambda_{hi}^{max} + 1} \sum_{i:\, v_i \leq 1} \alpha_i \left[v_i^{\lambda_i} - v_i^{\lambda_i + 1} \right] \geq \sum_{i:\, v_i \leq 1} \frac{\alpha_i}{\lambda_i + 1} \left[v_i - v_i^{\lambda_i + 1} \right], \tag{C.18}$$

Hence by Equation 3.13, $S(\pi_0) \geq S(\underline{\mu})$. □

Proof. (of Theorem 3.3) The proof is presented in the following steps:

Step 1: Let $0 < a \leq 1$, $b > a$ such that $a\,b > 1$ and function $h(v)$ be defined as:

$$h(v) = (b-a)v^b - (b+1)v^{b-1} + (a+1), \text{ for } v > 0. \quad \text{(C.19)}$$

$\lim_{v\to 0+} h(v) = a+1 > 1$, $\lim_{v\to +\infty} h(v) = +\infty$, $h(1) = 0$ and:

$$h'(v) = (b-a)\,b\,v^{b-2}\left[v - \frac{b^2-1}{b^2-ab}\right] \Rightarrow h'(v_m) = 0 \Rightarrow v_m = \frac{b^2-1}{b^2-ab} > 1. \quad \text{(C.20)}$$

$h''(v_m) > 0 \Rightarrow v_m$ is minimum. So there exists a $v^* > 1$ such that, $h(v) \leq 0$ for $1 < v \leq v^*$.

Step 2: For all $v_i > 1$, there exists a v_i^* such that for $1 < v_i \leq v_i^*$,

$$\sum_{i:\,v_i>1} \frac{\alpha_i}{\lambda_i+1}\left[v_i^{\lambda_i+1} - v_i\right] \geq \frac{1}{\lambda_{lo}^{min}+1} \sum_{i:\,v_i>1} \alpha_i \left[v_i^{\lambda_i+1} - v_i^{\lambda_i}\right]. \quad \text{(C.21)}$$

To prove this, set $a = \lambda_{lo}^{min}$ and $b = \lambda_i$, so $ab = \lambda_i\,\lambda_{lo}^{min} \geq \lambda_{hi}^{min}\,\lambda_{lo}^{min} > 1$. So, by Step 1:

$$(\lambda_i - \lambda_{lo}^{min})v_i^{\lambda_i} - (\lambda_i+1)v_i^{\lambda_i-1} + (\lambda_{lo}^{min}+1) \leq 0. \quad \text{(C.22)}$$

Rearranging and multiplying by $\alpha_i\,v_i$ on both sides, we get:

$$\frac{\alpha_i}{\lambda_i+1}\left[v_i^{\lambda_i+1} - v_i\right] \geq \frac{\alpha_i}{\lambda_{lo}^{min}+1}\left[v_i^{\lambda_i+1} - v_i^{\lambda_i}\right]. \quad \text{(C.23)}$$

As this holds for all $v_i > 1$, summing over all such risk-groups leads to Equation C.21.

Step 3: Based on all risk-groups for which $v_i \leq 1$:

$$\sum_{i:\,v_i\leq 1} \frac{\alpha_i}{\lambda_i+1}\left[v_i - v_i^{\lambda_i+1}\right] \leq \frac{1}{\lambda_{lo}^{min}+1} \sum_{i:\,v_i\leq 1} \alpha_i \left[v_i - v_i^{\lambda_i+1}\right], \text{ as } \lambda_{lo}^{min} \leq \lambda_i \quad \text{(C.24)}$$

$$\leq \frac{1}{\lambda_{lo}^{min}+1} \sum_{i:\,v_i\leq 1} \alpha_i \left[v_i^{\lambda_i} - v_i^{\lambda_i+1}\right], \text{ as } v_i \leq 1 \text{ and } \lambda_i \leq 1 \quad \text{(C.25)}$$

$$= \frac{1}{\lambda_{lo}^{min}+1} \sum_{i:\,v_i>1} \alpha_i \left[v_i^{\lambda_i+1} - v_i^{\lambda_i}\right], \text{ by Equation 3.11.} \quad \text{(C.26)}$$

Step 4: Combining Steps 2 and 3, we get

$$\sum_{i:\,v_i>1} \frac{\alpha_i}{\lambda_i+1}\left[v_i^{\lambda_i+1} - v_i\right] \geq \frac{1}{\lambda_{lo}^{min}+1} \sum_{i:\,v_i>1} \alpha_i \left[v_i^{\lambda_i+1} - v_i^{\lambda_i}\right] \geq \sum_{i:\,v_i\leq 1} \frac{\alpha_i}{\lambda_i+1}\left[v_i - v_i^{\lambda_i+1}\right], \quad \text{(C.27)}$$

for $1 < v_i \leq v_i^*$ for all $v_i > 1$.

As $v_i = \mu_i/\pi_0$, $v_i \leq v_i^* \Rightarrow \pi_0 \geq \mu_i/v_i^*$ for all risk-groups for which $v_i > 1$. So if we define $\pi^* = \max_{i:\,v_i>1}(\mu_i/v_i^*)$, then $\pi_0 \geq \pi^* \Rightarrow S(\pi_0) \geq S(\underline{\mu})$ by Equation 3.13. □

D. Expression for Social Welfare Under General Insurance Demand

Lemma 2. *Suppose there are n risk-groups with risks $\mu_1 < \mu_2 < \cdots < \mu_n$ and any general demand functions. Then for a given premium regime $\underline{\pi}$, for which no risk-group is fully insured, the expression for social welfare is given by:*

$$S(\underline{\pi}) = \sum_{i=1}^{n} p_i \, G_i \left(\frac{\mu_i}{\pi_i} \right) \pi_i + K, \quad \text{where} \quad G_i(g) = \int_0^g P\,[\,\Gamma_i < \gamma\,]\; d\gamma, \tag{5.2}$$

where the premium regime $\underline{\pi}$ satisfies the equilibrium condition:

$$\sum_{i=1}^{n} p_i \, d_i(\pi_i) \; (\pi_i - \mu_i) = 0, \tag{5.3}$$

and the constant K does not depend on the premium regime under consideration.

Proof. Recall that, given a risk-group i, insurance is purchased when $\Gamma_i < \mu_i/\pi_i$. Hence:

$$[\,Q \mid \text{Risk-group } i\,] = I\left[\Gamma_i < \frac{\mu_i}{\pi_i}\right] \Rightarrow \mathrm{E}\,[\,Q \mid \text{Risk-group } i\,] = \mathrm{P}\left[\Gamma_i < \frac{\mu_i}{\pi_i}\right] = d_i(\pi_i). \tag{D.1}$$

Using the expression for social welfare as given in Equation 2.17 we have:

$$\begin{aligned}
S(\underline{\pi}) &= \mathrm{E}\,[\,Q\,(X - \Pi\,\Gamma)\,] + K, && \text{(D.2)}\\
&= \mathrm{E}\,[\,Q\,X\,] - \mathrm{E}\,[\,Q\,\Pi\,\Gamma\,] + K, && \text{(D.3)}\\
&= \mathrm{E}\,[\,Q\,\Pi\,] - \mathrm{E}\,[\,Q\,\Pi\,\Gamma\,] + K, \text{ as under equilibrium: } E\,[\,Q\,X\,] = E\,[\,Q\,\Pi\,] && \text{(D.4)}\\
&= \mathrm{E}\,[\,(1-\Gamma)\,Q\,\Pi\,] + K, && \text{(D.5)}\\
&= \sum_{i=1}^{n} p_i \, \mathrm{E}\left[(1-\Gamma_i)\, I\left[\Gamma_i \le \frac{\mu_i}{\pi_i}\right]\right] \pi_i + K. && \text{(D.6)}
\end{aligned}$$

Now using Lemma 3:

$$\begin{aligned}
S(\underline{\pi}) &= \sum_{i=1}^{n} p_i \left[\left(1 - \frac{\mu_i}{\pi_i}\right) \mathrm{P}\left[\Gamma_i \le \frac{\mu_i}{\pi_i}\right] + \int_0^{\frac{\mu_i}{\pi_i}} \mathrm{P}\,[\,\Gamma_i \le \gamma\,]\; d\gamma \right] \pi_i + K, && \text{(D.7)}\\
&= \sum_{i=1}^{n} p_i \left(1 - \frac{\mu_i}{\pi_i}\right) \mathrm{P}\left[\Gamma_i \le \frac{\mu_i}{\pi_i}\right] \pi_i + \sum_{i=1}^{n} p_i \int_0^{\frac{\mu_i}{\pi_i}} \mathrm{P}\,[\,\Gamma_i \le \gamma\,]\; d\gamma\, \pi_i + K, && \text{(D.8)}\\
&= \sum_{i=1}^{n} p_i \, d_i(\pi_i)\; (\pi_i - \mu_i) + \sum_{i=1}^{n} p_i \, G_i \left(\frac{\mu_i}{\pi_i}\right) \pi_i + K, \quad \text{as } \mathrm{P}\left[\Gamma_i \le \frac{\mu_i}{\pi_i}\right] = d_i(\pi_i), && \text{(D.9)}\\
&= \sum_{i=1}^{n} p_i \, G_i \left(\frac{\mu_i}{\pi_i}\right) \pi_i + K, \quad \text{as in equilibrium: } \sum_{i=1}^{n} p_i \, d_i(\pi_i)\; (\pi_i - \mu_i) = 0. && \text{(D.10)}
\end{aligned}$$

as required. ∎

Lemma 3. *For a positive continuous random variable, X:*

(i) $E[X] = \int_0^\infty P[X > y]\, dy$;

(ii) $E[X\, I[X \le c]] = c\, P[X \le c] - \int_0^c P[X \le y]\, dy$;

(iii) $E[(1 - X)\, I[X \le c]] = (1 - c)\, P[X \le c] + \int_0^c P[X \le y]\, dy$.

Proof. Assuming the density function of X is given by $p(x)$

(i)

$$\begin{aligned} \mathrm{E}[X] &= \int_0^\infty x\, p(x)\, dx = \int_0^\infty \left[\int_0^x dy\right] p(x)\, dx = \int_0^\infty \left[\int_y^\infty p(x)\, dx\right] dy \\ &= \int_0^\infty \mathrm{P}[X > y]\, dy. \end{aligned} \tag{D.11}$$

(ii)

$$\mathrm{E}[X\, I[X \le c]] = \int_0^c x\, p(x)\, dx, \tag{D.12}$$

$$= \int_0^c \left[\int_0^x dy\right] p(x)\, dx, \tag{D.13}$$

$$= \int_0^c \left[\int_y^c p(x)\, dx\right] dy, \quad \text{by interchanging integrals,} \tag{D.14}$$

$$= \int_0^c \mathrm{P}[y < X \le c]\, dy, \tag{D.15}$$

$$= \int_0^c [\mathrm{P}[X \le c] - \mathrm{P}[X \le y]]\, dy, \tag{D.16}$$

$$= c\,\mathrm{P}[X \le c] - \int_0^c \mathrm{P}[X \le y]\, dy. \tag{D.17}$$

(iii)

$$\mathrm{E}[(1 - X)\, I[X \le c]] = \mathrm{E}[I[X \le c]] - \mathrm{E}[X\, I[X \le c]], \tag{D.18}$$

$$= \mathrm{P}[X \le c] - \left[c\,\mathrm{P}[X \le c] - \int_0^c \mathrm{P}[X \le y]\, dy\right], \tag{D.19}$$

$$= (1 - c)\,\mathrm{P}[X \le c] + \int_0^c \mathrm{P}[X \le y]\, dy \tag{D.20}$$

□

E. Derivations for General Demand Elasticities

First note that if demand elasticity is an increasing function of premium π, then it is a decreasing function of $v = \mu_i/\pi$; and hence a weighted average such as arc elasticity $\lambda_i(v)$ is also decreasing function of v. The inverse statements (i.e. with increasing replaced by decreasing and vice versa) also hold.

Lemma 4. *If for a risk-group i, $\mu_i > \pi_0$ (i.e. $v_i > 1$) and the demand elasticity, $\epsilon_i(\pi)$, is an increasing function of premium π, then:*

$$G_i(v_i) - v_i G_i(1) \geq \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i^{\lambda_i(v_i)+1} - v_i\right]. \tag{E.1}$$

Proof. Firstly:

$$\begin{aligned}
G_i(v_i) - G_i(1) &= \int_1^{v_i} P\left[\Gamma_i \leq v\right]\, dv, && \text{(E.2)}\\
&= \int_1^{v_i} \tau_i\, v^{\lambda_i(v)} dv, \text{ by Equation 5.10,} && \text{(E.3)}\\
&\geq \int_1^{v_i} \tau_i\, v^{\lambda_i(v_i)} dv, \text{ as } \lambda_i(v) \text{ is a decreasing function,} && \text{(E.4)}\\
&= \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i^{\lambda_i(v_i)+1} - 1\right]. && \text{(E.5)}
\end{aligned}$$

And,

$$\begin{aligned}
(v_i - 1)G_i(1) &= (v_i - 1)\int_0^1 P\left[\Gamma_i \leq v\right]\, dv, && \text{(E.6)}\\
&= (v_i - 1)\int_0^1 \tau_i\, v^{\lambda_i(v)} dv, \text{ by Equation 5.10,} && \text{(E.7)}\\
&\leq (v_i - 1)\int_0^1 \tau_i v^{\lambda_i(v_i)} dv, \text{ as } v < 1 \Rightarrow v^{\lambda_i(v)} \leq v^{\lambda_i(v_i)}, && \text{(E.8)}\\
&= \frac{(v_i - 1)\tau_i}{\lambda_i(v_i)+1}. && \text{(E.9)}
\end{aligned}$$

Hence:

$$G_i(v_i) - v_i G_i(1) = \left[G_i(v_i) - G_i(1)\right] - \left[(v_i - 1)G_i(1)\right] \geq \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i^{\lambda_i(v_i)+1} - v_i\right], \tag{E.10}$$

as required. □

Lemma 5. *If for a risk-group i, $\mu_i \leq \pi_0$ (i.e. $v_i \leq 1$) and the demand elasticity, $\epsilon_i(\pi)$, is a decreasing function of premium π, then:*

$$v_i G_i(1) - G_i(v_i) \leq \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i - v_i^{\lambda_i(v_i)+1}\right]. \tag{E.11}$$

Proof. Firstly:

$$\begin{aligned}
v_i\left[G_i(1) - G_i(v_i)\right] &= v_i \int_{v_i}^{1} P\left[\Gamma_i \leq v\right] dv, && \text{(E.12)}\\
&= v_i \int_{v_i}^{1} \tau_i\, v^{\lambda_i(v)} dv, \text{ by Equation 5.10,} && \text{(E.13)}\\
&\leq v_i \int_{v_i}^{1} \tau_i\, v^{\lambda_i(v_i)} dv, \text{ as } v < 1 \Rightarrow v^{\lambda_i(v)} \leq v^{\lambda_i(v_i)}, && \text{(E.14)}\\
&= \frac{v_i \tau_i}{\lambda_i(v_i)+1}\left[1 - v_i^{\lambda_i(v_i)+1}\right]. && \text{(E.15)}
\end{aligned}$$

And

$$\begin{aligned}
(1-v_i)G_i(v_i) &= (1-v_i)\int_0^{v_i} P\left[\Gamma_i \leq v\right] dv, && \text{(E.16)}\\
&= (1-v_i)\int_0^{v_i} \tau_i\, v^{\lambda_i(v)} dv, \text{ by Equation 5.10,} && \text{(E.17)}\\
&\geq (1-v_i)\int_0^{v_i} \tau_i\, v^{\lambda_i(v_i)} dv, \text{ as } \lambda_i(v) \text{ is an increasing function,} && \text{(E.18)}\\
&= \frac{(1-v_i)\tau_i}{\lambda_i(v_i)+1}\left[v_i^{\lambda_i(v_i)+1}\right]. && \text{(E.19)}
\end{aligned}$$

Hence, as required:

$$v_i G_i(1) - G_i(v_i) = v_i\left[G_i(1) - G_i(v_i)\right] - (1-v_i)G_i(v_i) \leq \frac{\tau_i}{\lambda_i(v_i)+1}\left[v_i - v_i^{\lambda_i(v_i)+1}\right]. \tag{E.20}$$

□

F. Social welfare when higher risks are fully insured under pooling

In the main text of the paper, we have explicitly assumed that no risk-groups are fully insured under any premium regime. However, for sufficiently small pooled equilibrium premium, it is possible that all individuals purchase insurance, in some higher risk-groups.

If there are more than two risk-groups, the analysis of implications of full insurance would require consideration of many possible combinations. For ease of exposition, while analysing the case of full take-up of insurance, we will only consider two risk-groups, where the high risk-group is fully insured under pooling. We assume that fair-premium demand $\tau_i < 1$ for all risk-groups, which is consistent with most empirical evidence. (The special case of $\tau_i = 1$ can also be analysed using the same techniques.)

Assuming $\tau_i < 1$, social welfare under full risk classification follows from Lemma 1:

$$S(\underline{\mu}) = p_1\,\tau_1 \frac{1}{(\lambda_1+1)}\mu_1 + p_2\,\tau_2 \frac{1}{(\lambda_2+1)}\mu_2 + K. \tag{F.1}$$

For pooling we obtain the following lower bound for social welfare:

Lemma 6. *Suppose there are two risk-groups with risks $\mu_1 < \mu_2$ with positive constant demand elasticities λ_1 and λ_2 respectively. If the high risk-group is fully insured under pooling, then social welfare under pooled premium $S(\pi_0)$ satisfies:*

$$S(\pi_0) \geq p_1\,\tau_1 \frac{1}{(\lambda_1+1)} \left(\frac{\mu_1}{\pi_0}\right)^{\lambda_1+1} \pi_0 + p_2 \frac{1}{(\lambda_2+1)}\mu_2 + K, \tag{F.2}$$

where the pooled premium π_0 satisfies the equilibrium condition:

$$p_1\,\tau_1 \left(\frac{\mu_1}{\pi_1}\right)^{\lambda_1} (\pi_0 - \mu_1) + p_2\,(\pi_0 - \mu_2) = 0, \tag{F.3}$$

and the constant K does not depend on the premium regime under consideration.

Proof. The equilibrium condition follows from Equation 2.21, by inserting the specific expression for iso-elastic insurance demand for low risk-group and noting that proportional demand for high risk-group is 1 under pooling.

Using the general expression for social welfare given in Equation 2.17, we have:

$$S(\pi_0) = \mathrm{E}\,[\,Q\,X - \,Q\,\Pi\,\Gamma\,] + K, \tag{F.4}$$

$$= \sum_{i=1}^{2} \mathrm{E}\,[\,Q\,X - \,Q\,\Pi\,\Gamma \mid \text{Risk-group } i\,]\,p_i + K. \tag{F.5}$$

As not all low risks will purchase insurance, the same steps in Lemma 1 will give:

$$\mathrm{E}\,[\,Q\,X - \,Q\,\Pi\,\Gamma \mid \text{Risk-group } 1\,] = p_1\,\tau_1 \frac{1}{(\lambda_1+1)} \left(\frac{\mu_1}{\pi_0}\right)^{\lambda_1+1} \pi_0. \tag{F.6}$$

But all high risks buy insurance under pooling, i.e. $[Q \mid \text{Risk-group } 2] = 1$. So:

$$\begin{aligned}
\mathrm{E}\left[\, Q\,X - Q\,\Pi\,\Gamma \mid \text{Risk-group } 2\,\right] &= \mathrm{E}\left[\, X \mid \text{Risk-group } 2\,\right] - \mathrm{E}\left[\, \Pi\,\Gamma \mid \text{Risk-group } 2\,\right], && \text{(F.7)}\\
&= \mu_2 - \mathrm{E}\left[\, \Gamma \mid \text{Risk-group } 2\,\right]\, \pi_0, && \text{(F.8)}\\
&= \mu_2 - \int_0^{\left(\frac{1}{\tau_2}\right)^{\frac{1}{\lambda_2}}} \gamma\tau_2\lambda_2\gamma^{\lambda_2-1} d\gamma\, \pi_0, && \text{(F.9)}\\
&= \mu_2 - \frac{\lambda_2}{(\lambda_2+1)} \left(\frac{1}{\tau_2}\right)^{\frac{1}{\lambda_2}} \pi_0, && \text{(F.10)}\\
&\geq \frac{1}{(\lambda_2+1)}\, \mu_2, \quad \text{since } \tau_2 \left(\frac{\mu_2}{\pi_0}\right)^{\lambda_2} \geq 1 \Rightarrow \left(\frac{1}{\tau_2}\right)^{\frac{1}{\lambda_2}} \pi_0 \leq \mu_2. && \text{(F.11)}
\end{aligned}$$

Using Equations F.6 and F.11 in Equation F.5 gives the required relationship in Equation F.2. □

Equation F.2 of Lemma 6 implies that, when high risks are fully insured under pooling (but partially insured under full risk classification), social welfare under pooling exceeds that under full risk classification, i.e. $S(\pi_0) \geq S(\underline{\mu})$ if:

$$\begin{aligned}
&p_1\,\tau_1 \frac{1}{(\lambda_1+1)} \left(\frac{\mu_1}{\pi_0}\right)^{\lambda_1+1} \pi_0 + p_2 \frac{1}{(\lambda_2+1)}\, \mu_2 \geq p_1\,\tau_1 \frac{1}{(\lambda_1+1)} \mu_1 + p_2\,\tau_2 \frac{1}{(\lambda_2+1)} \mu_2, && \text{(F.12)}\\
\Leftrightarrow\ &p_1\,\tau_1 \frac{1}{(\lambda_1+1)} {v_1}^{\lambda_1+1} + p_2 \frac{1}{(\lambda_2+1)} v_2 \geq p_1\,\tau_1 \frac{1}{(\lambda_1+1)}\, v_1 + p_2\,\tau_2 \frac{1}{(\lambda_2+1)}\, v_2, && \text{(F.13)}
\end{aligned}$$

using the notations involving risk-premium ratios: v_1 and v_2. And Equation F.3 becomes:

$$p_1\,\tau_1\,{v_1}^{\lambda_1}\,(1 - v_1) + p_2\,(1 - v_2) = 0 \qquad \text{(F.14)}$$

We can then state the sufficient condition on λ_1 and λ_2, for social welfare to be higher under pooling than under full risk classification for any population structures and underlying risks, when high risks are fully insured under pooling.

Theorem 5. *Suppose there are two risk-groups with risks $\mu_1 < \mu_2$ with positive constant demand elasticities λ_1 and λ_2 respectively. If high risks are fully insured under pooling while low risks are not, and neither risk-group is fully insured under full risk classification, then:*

$$\lambda_1 \leq 1 \text{ and } \lambda_2 \leq \left(1 + \frac{1}{\lambda_1}\right)(1 - \tau_2) - 1 \Rightarrow S(\pi_0) \geq S(\underline{\mu}). \qquad \text{(F.15)}$$

Proof. The proof is presented in the following steps:

Step 1: The equilibrium condition in Equation F.14 leads to:

$$p_2\, v_2 = p_1\, \tau_1\, \left({v_1}^{\lambda_1} - {v_1}^{\lambda_1+1}\right) + p_2. \tag{F.16}$$

Step 2: Using Equation F.16 in the social welfare condition in Equation F.13 gives:

$$S(\pi_0) \geq S(\underline{\mu}) \tag{F.17}$$

$$\text{if} \quad p_1\, \tau_1\, \frac{1}{(\lambda_1+1)} {v_1}^{\lambda_1+1} + p_2\, \frac{1}{(\lambda_2+1)} v_2 \geq p_1\, \tau_1 \frac{1}{(\lambda_1+1)}\, v_1 + p_2\, \tau_2 \frac{1}{(\lambda_2+1)}\, v_2, \tag{F.18}$$

$$\begin{aligned} \text{i.e. if} \quad & p_1\, \tau_1\, \frac{1}{(\lambda_1+1)} {v_1}^{\lambda_1+1} + p_1\, \tau_1\, \frac{1}{(\lambda_2+1)} \left({v_1}^{\lambda_1} - {v_1}^{\lambda_1+1}\right) + p_2\, \frac{1}{(\lambda_2+1)} \\ & \geq p_1\, \tau_1 \frac{1}{(\lambda_1+1)}\, v_1 + p_1\, \tau_1 \frac{\tau_2}{(\lambda_2+1)}\, \left({v_1}^{\lambda_1} - {v_1}^{\lambda_1+1}\right) + p_2\, \frac{\tau_2}{(\lambda_2+1)}, \end{aligned} \tag{F.19}$$

$$\begin{aligned} \text{i.e. if} \quad & p_1\, \tau_1\, \frac{1}{(\lambda_1+1)} {v_1}^{\lambda_1+1} + p_1\, \tau_1\, \frac{1}{(\lambda_2+1)} \left({v_1}^{\lambda_1} - {v_1}^{\lambda_1+1}\right) \\ & \geq p_1\, \tau_1 \frac{1}{(\lambda_1+1)}\, v_1 + p_1\, \tau_1 \frac{\tau_2}{(\lambda_2+1)}\, \left({v_1}^{\lambda_1} - {v_1}^{\lambda_1+1}\right), \quad \text{as } \tau_2 < 1, \end{aligned} \tag{F.20}$$

$$\text{i.e. if} \quad \frac{(1-\tau_2)}{(\lambda_2+1)} \geq \frac{1}{(\lambda_1+1)}\, \frac{\left(v_1 - {v_1}^{\lambda_1+1}\right)}{\left({v_1}^{\lambda_1} - {v_1}^{\lambda_1+1}\right)}. \tag{F.21}$$

Step 3: As $0 < \lambda_1 \leq 1$ and $0 < v_1 < 1$, using Arithmetic Mean $\geq$ Geometric Mean:

$$(1-\lambda_1)\, v_1^{\lambda_1+1} + \lambda_1\, v_1^{\lambda_1} \geq v_1 \Rightarrow \frac{\lambda_1}{(\lambda_1+1)} \geq \frac{1}{(\lambda_1+1)}\, \frac{\left(v_1 - {v_1}^{\lambda_1+1}\right)}{\left({v_1}^{\lambda_1} - {v_1}^{\lambda_1+1}\right)}. \tag{F.22}$$

Step 4: Finally:

$$\lambda_2 \leq \left(1 + \frac{1}{\lambda_1}\right)\, (1-\tau_2) - 1 \Rightarrow \frac{(1-\tau_2)}{(\lambda_2+1)} \geq \frac{\lambda_1}{(\lambda_1+1)}, \tag{F.23}$$

$$\Rightarrow \frac{(1-\tau_2)}{(\lambda_2+1)} \geq \frac{1}{(\lambda_1+1)}\, \frac{\left(v_1 - {v_1}^{\lambda_1+1}\right)}{\left({v_1}^{\lambda_1} - {v_1}^{\lambda_1+1}\right)}, \quad \text{by Step 3}, \tag{F.24}$$

$$\Rightarrow S(\pi_0) \geq S(\underline{\mu}), \quad \text{by Step 2}. \tag{F.25}$$

□

Figure 4 provides a graphical representation of Theorem 5, where the fair-premium demand is 50% for both low and high risk-groups. Social welfare is guaranteed to be higher under pooling for all population structures and risks in the shaded region to the left of the bold green curve.

For specific population structures and risk parameters, the region where social welfare is higher under pooling is a much larger area than the shaded region in Figure 4. For example,

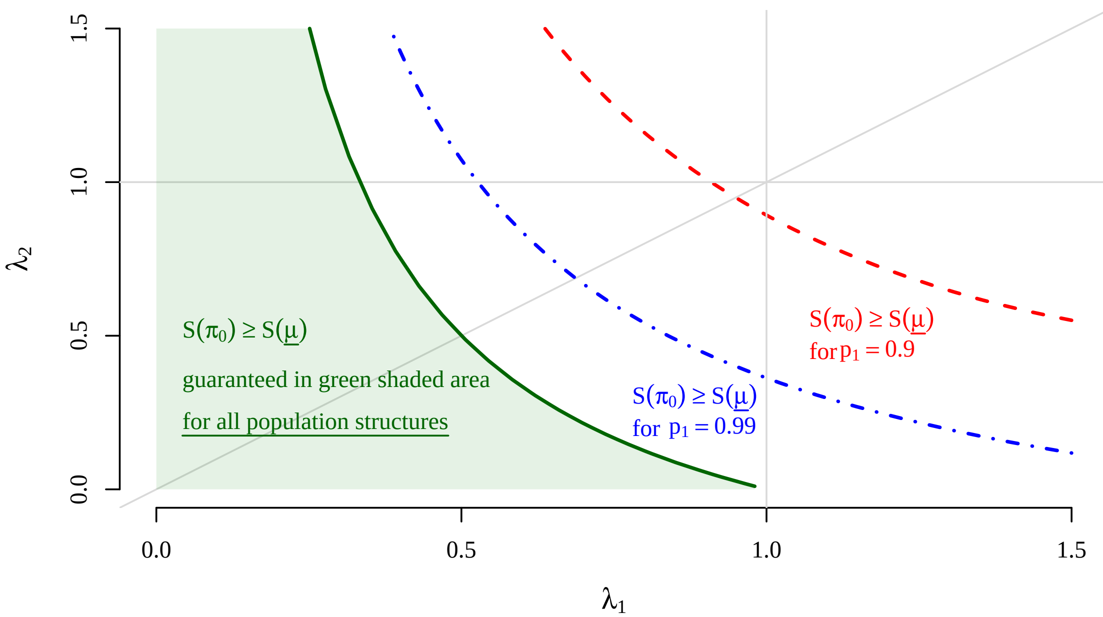

Figure 4: Curve demarcating the regions where social welfare under pooling is greater than under full risk-differentiation where $(\mu_1, \mu_2) = (0.01, 0.04)$, fair-premium demand is 50% for both risk-groups and high risks are fully insured under pooling.

social welfare is guaranteed to be higher under pooling in the region to the left of the blue dot-dashed line for $p_1 = 0.99$ and $(\mu_1, \mu_2) = (0.01, 0.04)$. Similarly, the region to the left of the red dashed line represents the region where social welfare is guaranteed to be higher under pooling for $p_1 = 0.9$ and $(\mu_1, \mu_2) = (0.01, 0.04)$. The region where social welfare is guaranteed to be higher under pooling increases with the size of the higher risk-group, because larger high risk-group's gain in welfare from pooling has greater capacity to offset the lower risk-group's loss in welfare from pooling.