Bundling in Insurance Markets: Theory and an Application to Long-term Care

Adam Solomon*

February 7, 2024

Abstract

Every insurance contract bundles risks, and explicit bundling discounts are common. I show theoretically that bundling arises whenever correlation between risk types enables insurer "cream-skimming": the willingness-to-pay for insurance against one risk must be negatively correlated with expected costs from the other risk. I analyze long-term care insurance, in which both-spouse bundles are discounted by 20-35%. I show that cream-skimming incentives are sufficient to explain the observed discounts, and rule out standard economies-of-scale. Counterfactually, banning bundling would raise welfare by 5% by correcting separate market unraveling, while mandatory family bundling would reduce welfare by 5% as it exacerbates advantageous selection.

*Department of Economics, Massachusetts Institute of Technology. Email: adamsol@mit.edu. I thank Jim Poterba, Amy Finkelstein, Jon Gruber, Nathan Hendren, Tim Layton, Frank Schilbach, Abhijit Bannerjee, Stephen Morris, Johannes Spinnewijn, Daniel Gottlieb, Cameron Ellis, Ben Handel, Evan Soltas, Glenn Ellison for helpful comments. I am grateful for support from the Bradley Foundation and the Jerry A. Hausman Dissertation Fellowship.

1 Introduction

All insurance products are bundles. Medical insurance bundles inpatient coverage with outpatient coverage, and commercial insurers will typically give a discount for home and auto insurance purchased together. At the same time, these bundles are often limited; medical insurance bundles typically do not include dental insurance, while there are generally no discounts for purchasing home and life insurance together. Moreover, regulators often mandate bundling: the Affordable Care Act requires health insurance plans to cover 10 'essential health benefits'.

Why is insurance sometimes bundled and sometimes not? And what is the proper role of the government in regulating – or restricting – such bundling? These are critical questions which are central to the insurance industry. Yet they are not addressed by either theory or empirical work on insurance. Canonical insurance models such as Akerlof (1978) or Einav et al. (2010) feature only one risk. These models draw important conclusions about the social inefficiency of unregulated insurance markets, for a single risk, that arise due to information asymmetries. With multiple risks, bundling can accentuate or mitigate typical insurance market failures. This raises critical questions of whether the private market is behaving optimally and what role the government might play in regulating or mandating bundling in the presence of private market failures.

In this paper, I address these issues both theoretically and empirically. I begin by extending the widely used Einav et al. (2010) model, based on Akerlof (1978). There are two risks and a fixed insurance contract for each. Markets clear on price. In each perfectly competitive separate insurance market, there is selection: costs are not constant and covary with willingness-to-pay (WTP). These two assumptions - selection and perfect competition - distinguish this paper from the typical setup in the Industrial Organization literature.

Two key primitives drive the results throughout. First, whether selection is adverse or advantageous. Adverse (advantageous) selection means costs are an increasing (decreasing) function of WTP. Second, whether risk types in each market are positively or negatively correlated. They are positively (negatively) correlated when those that have high WTPs for contract 1 typically have high (low) WTP for contract 2. I find that incentives to bundle and welfare consequences of regulations that mandate or forbid bundling depend critically on these primitives.

I show that profit-maximizing insurers bundle when it profitably 'cream-skims' low cost consumers from the separate markets. Bundling screens on WTP: a bundling discount attracts those with high WTP for both insurance products, whereas those with high WTP for one product but low WTP for the other continue to buy the single product. Bundling will successfully cream-skim when those with high WTP for both products are lower cost than those who remain in the single market. This occurs in exactly two situations: when risks are adversely selected and negatively correlated, or when risks are advantageously selected and positively correlated. In both cases, those that buy both products, and who would be attracted by a bundling discount, are lower cost than those who just buy one.

I show that bundling driven by market forces has an ambiguous impact on welfare. When firms cream-skim by offering a bundling discount, those in the bundled market are directly better off. By definition, those that remain in the separate markets are higher cost. This can cause the separate markets to unravel, an externality that the bundling firm does not internalize. The extent of the unraveling depends on selection and correlation: for an adversely (advantageously) selected market, the more negative (positive) the correlation, the more unraveling that occurs. As the separate markets unravel, people move from the separate market to the bundling market. They are typically low cost relative to the separate market, but high cost relative to the bundled market, which can have knock-on effects on the bundled market, dampening the bundling discount. Which of these impacts on welfare prevail - separate market unraveling versus bundled market expansion - is an empirical question. Whether government interventions that *forbid* bundling can improve welfare from the private bundling equilibrium is the inverse of this question, and must also be decided empirically.

I analyze when government interventions that *force* bundling improve welfare. Choice is restricted, which directly reduces surplus. Consumers who prefer to buy one policy now must choose between buying both or buying neither. However, forced bundling can reduce prices by dampening selection. Consider whose purchase of the first policy changes after bundling is mandated. Those who have high WTP for the first policy, but very low WTP for the second, leave the market, since they can no longer just buy policy one. Those who have have high WTP for the second and moderate WTP for the first go from buying just policy two to buying the bundle. On net, the WTP of those who buy the first policy has decreased, and similarly for the second. Under adverse selection, this reduces cost, lowers equilibrium price, and mitigates the selection problem, but must be weighed against the direct welfare costs of reduced choice. Under advantageous selection, this raises equilibrium price and unambiguously lowers welfare.

I apply this framework and insights to the market for long-term care (LTCI) insurance. Long-term care is one of the largest financial risks in elder life, but LTCI take-up is very low. Large spousal bundling discounts are offered in this market: up to 35% if *both spouses purchase* a LTCI policy, in addition to a 5-10% discount simply for *having a spouse*. These discounts benefit couples who are both in good enough health to qualify for a policy, but might increase prices in the markets for single policies. I use the framework to explain these discounts, and empirical simulations using HRS data to quantify the trade-off due to the spousal bundling, and investigate whether regulatory interventions - mandating or banning bundling - can increase welfare.

To study intra-household correlation in long-term care risk, I use the HRS. The HRS is a biennial survey of older Americans and their spouses that has run for over 30 years. The HRS collects data on actual long-term care usage, as well as risk factors for long-term care, both objective and subjective. Using eventual realizations of risk, I can make unbiased ex-ante predictions of the prospective risk at younger ages when individuals had the choice to buy insurance.

Following Finkelstein and McGarry (2006), I find that LTCI is advantageously selected, and intra-household long-term care risk is positively correlated. Both of these facts are likely driven by within-household correlation in wealth and health. Wealthier individuals are more likely to buy insurance (of any kind, per Gropper and Kuhnen (2021)), and a healthy spouse can substitute for formal care, thereby lowering expected costs. In the terminology of Finkelstein and McGarry (2006): there are two types of people who buy LTCI, those that have private information that they are high risk, and those that have private information about their strong taste for insurance, but are low risk. The latter are more likely to be married to each other, since wealth (and correlates such as risk aversion and health) is common to the household. Bundling discounts separate the wealthy and healthy couples who have high WTP from those who are buying LTCI because of private information about their high risk (a part of which may be due to not having a healthy spouse who can provide care).

Our framework predicts bundling discounts should arise in precisely this setting: conditional on one's WTP for long-term care insurance, the fact that their spouse also has high WTP - and is likely healthy and wealthy - predicts low costs for the individual (and their spouse).

I directly test how an individual's own cost is predicted by their spouses WTP for long-term care insurance. I rank individuals by their risk, and given advantageous selection, construct an ordinal ranking of WTP that increases as risk decreases. Without relying on a specific model of WTP, I compare the long-term care risk of individuals with a high WTP/low risk spouse, those with a low WTP/high risk spouse, and those with no spouse.

I find a clear relationship between spousal WTP for LTCI and own risk. Those that have a spouse are 10% cheaper than those who are single. But within those who have a spouse, those with a spouse in the highest decile of WTP have almost 60% lower risk relative to those whose spouse is in the lowest decile. This rationalizes the large discount (of 30-40%) given when both spouses buy a policy, and the much smaller discount of 5-10% given just for having a spouse. The latter is equivalent to pricing on an observable factor such as smoking status or income. The former is a true bundling discount that implicitly screens on spousal WTP, over and above having a spouse. This extra information is only revealed to the insurance company by both spouses buying a policy.

I rule out two alternate explanations for bundling discounts. First, I find that 'standard' economies of scale that might arise from selling two policies at once are not sufficient to explain the discounts offered. I show that the couples discounts offered in life insurance, annuities, and disability insurance are an order of magnitude lower than in LTC, indicating that generic cost efficiencies do not explain the LTC couples discounts. Second, I show that couples are less likely than singles to lapse their LTC policies. Since LTC insurance is front-loaded (premia are received well before claims typically occur), lapsation is profitable to the insurer. Hence, lapsation pushes against couples discounts, invalidating it as an alternative explanation of the discounts offered.

Having rationalized the bundling behavior in the private market, I evaluate potential government interventions to improve welfare. To study counterfactual market equilibria, I estimate a structural model for WTP for LTCI and long-term care costs for everyone in the HRS below the age of 65 who is eligible for a policy. To identify WTP, I use increases in premiums (sometimes more than 40% annually) over the sample period owing to insurers radically underestimating the expected claims of existing policyholders. These premium increases applied to both new and existing policyholders. This is, to my knowledge, a novel source of identification for demand in long-term care insurance.

Having estimated costs and WTP, I compare the mixed-bundling equilibrium (as is the status quo in the real LTCI market) with equilibria in which either bundling discounts are banned, or household bundling is forced such that either all or neither members of a household (whether single or a couple) buy a policy.

I find that banning bundling would raise welfare by up to 5%, while forcing family bundling would lower it by 5%. The latter fact is predicted by the theory: forced bundling in an advantageously selected market unambiguously lowers welfare, as it increases prices in the bundled market and reduces consumer choice. Banning bundling increases welfare because it decreases prices in the separate markets by 10%, almost entirely mitigating the lost bundling discount that previously accrued to couples. The price reduction in the separate markets means that almost twice as many singles buy insurance in equilibrium. Overall, couples pay marginally more for two policies, but singles and couples purchasing only one policy pay substantially less, which increases welfare.

The paper proceeds as follows. In section 2 I set up the theoretical model and answer the positive question: when do insurers bundle, and what does the equilibrium look like when they do? In section 3 I show that the positive theory is sufficient to explain the couples discounts offered in LTC insurance. In section 4 I theoretically study the normative welfare implications of bundling and government policies to force or prevent bundling. In section 5 I estimate the primitives of the LTC insurance market and simulate counterfactual equilibria to analyze the quantitative welfare impacts of government interventions. Section 6 concludes.

1.1 Literature Review

This paper contributes primarily to three literatures.

First, this paper speaks to the literature on contract design in insurance contexts. Multiple papers have documented firm incentives to manipulate contracts to attract lower cost enrollees, without exploring the welfare implications: Lavetti and Simon (2018) show that selective formulary design in Medicare Advantage plans screens out individuals with high costs in parts A and B; Shepard (2022) shows how some insurer networks exclude 'star hospitals' so as to remove the loyal users of these high cost hospitals from their risk pool; Cooper and Trivedi (2012) show how Medicare Advantage plans strategically included gym memberships in their plans to attract healthier and fitter enrollees. My contribution is to formalize the conditions under which bundling occurs, relate this to selection and correlation between the risks, and provide theory and evidence on the normative implications of bundling.

Second, this paper relates to the literature on the long-term care insurance market and its shortcomings. Despite long-term care being one of the largest late-life financial risks, insurance take-up remains remarkably low.[1] In particular, my findings are consistent with the intuition suggested by Finkelstein and McGarry (2006). They find that two types of people have high WTP for LTC insurance, those with high risk and high costs, and those with high WTP (due in part to risk aversion) and low costs. To the extent the latter types are partnered with each other, due to shared high income, for example, bundling discounts allows the insurer to separate the wealthy, prudent but low cost couples from the high cost who might have private information. My contribution is to highlight the impact that bundling discounts for spouses can have on coverage and welfare by concentrating LTCI ownership in wealthy, low cost couples and causing markets for singles (who have fewer informal carers available) to unravel.

Finally, there is a well-established Industrial Organization literature on bundling products and add-on pricing. Stigler (1963), Schmalensee (1984) and Adams and Yellen (1976) showed that pure bundling can dominate separate sales for a monopolist under negative correlation. Recent papers [2] have been able to make stronger conclusions, but typically find that bundling (pure or mixed) dominates separate sales when there is negative correlation in consumer valuation. Additional rationales for bundling include softening competition in differentiated products[3] or to deter entry[4]. Relative to that literature, my main contribution is introducing selection (e.g. non-constant marginal costs) and studying the case of perfect competition instead of the oligopoly or monopoly. I find that positive or negative correlation can induce bundling, when the market exhibits advantageous or adverse selection respectively.

The most directly related paper is Nguyen (2022), which studies familial

[1] Various explanations have been offered: Advantageous selection (Finkelstein and McGarry (2006)), crowd-out from Medicaid (Brown and Finkelstein (2008)), narrow framing (Gottlieb and Mitchell (2020)) and the expectation of formal care provided by children (Ko (2021)).

[2] For example: McAfee et al. (1989), Bakos and Brynjolfsson (1999), Chen and Riordan (2013), Haghpanah and Hartline (2021)

[3] Zhou (2017), Hurkens et al. (2019) and Zhou (2021)

[4] Nalebuff (2004)

health insurance choice in Vietnam. That paper shows, in a structural estimation, that a planner, when they are able to set optimal prices by fiat, can increase welfare by forcing all members of a family to enroll in health insurance (or not) as a unit. This reduces within-family adverse selection relative to the status quo in which a family can enroll some members but not others. My contribution is to: a) move beyond prices set by a planner and study bundling in a competitive market, b) provide a theoretical model and conceptual framework that nests the empirical example of Nguyen (2022), c) distinguish between firm and planner incentives to bundle and clarify when regulatory interventions into bundled markets can improve welfare.

2 Positive Theory: When Do Insurers Bundle?

2.1 Setup

I take a sufficient statistics approach that builds on the widely used Einav et al. (2010) (EFC) framework. A microfoundation is explored in appendix A.2. There are two risks and a fixed insurance contract for each.[5] I will refer to the risks as risk 1 and 2 and the fixed insurance contracts as insurance contract 1 and 2 respectively. Each type is labeled by their willingness-to-pay (WTP) for the contract that covers each risk $w = (w_1, w_2)$. I assume each WTP is bounded: $w_i \in [0, \bar{w}_i]$ for $i = 1, 2$ and write $\mathcal{W} = [0, \bar{w}_1] \times [0, \bar{w}_2]$ for the set of all permissible WTPs. Typically, I will think of a type as an individual and each product as covering a distinct risk, for example, home insurance and car insurance. However, one can also think of a type as a family and each policy as covering the same risk for a different individual in the family. For example, long-term care insurance for two spouses in a couple is my focus in the empirical section.

The insurance contracts may or may not be bundled. If they are bundled,

[5] In appendix A.3 I study endogenous contracts in a simplified setting, and find many insights qualitatively remain.

then each individual's WTP for the bundled contract is $w_B = w_1 + w_2$.[(6)(7)] Costs are heterogeneous and covary with WTP. The cost of insuring type $w = (w_1, w_2)$ is given by $\phi_1(w_1)$ if they buy contract 1, $\phi_2(w_2)$ if they buy contract 2, and $\phi_B(w_1, w_2) = \phi_1(w_1) + \phi_2(w_2)$ if they buy the bundled contract. I assume that the functions ϕ_1, ϕ_2, ϕ_B are continuous and differentiable. There are two canonical cases of interest:

Definition 1. *If $\phi'(w) > 0$, I say that a risk is adversely selected. If $\phi'(w) < 0$, I say that a risk is advantageously selected.*

Adverse or advantageous selection is assumed to hold globally. I refer to a joint distribution of risks (w_1, w_2) as an **economy**, and within each economy there are (up to) three insurance **markets**, one for risk 1, one for risk 2, and possibly a bundled product. For an economy X I write F_X for the joint CDF function, F_1 for the CDF of the marginal distribution of (the WTP for insurance against) risk 1, and F_2 for risk 2. I will be interested in economies X and Y that have the same marginal distributions F_1, F_2 but different joint distributions F_X, F_Y due to different correlation structures. This is explained in section 2.1.1.

The supply side of the market consists of infinitely many identical firms who compete on price. Each firm chooses a vector of prices to offer (p_1, p_2, p_B) where the bundle price will have to be lower than the sum of individual prices in order to attract any demand. In equilibrium all firms will offer the same prices at which total profits will be zero, and so I assume each firm receives a representative sample of consumers and hence makes zero profit individually.

Given prices the type-space $\mathcal{W}$ will be partitioned into the set of types who purchase only product 1, those who only purchase product 2, those that purchase both, perhaps at a bundle discount, and those that purchase neither,.

[(6)] I stress that a bundled insurance product still offers the same state-contingent payoffs (except perhaps the premium) as buying the two products separately. The bundled product pays the same indemnity should risk 1 occur as does insurance contract 1, and similarly for risk 2. This is in contrast to a different policy that pays off only when some condition involving both risks is met, perhaps the sum of losses exceeds some deductible. This is studied in, for example, Klosin and Solomon (2024).

[(7)] The separability assumption is analyzed further in A.2.

Label these groups $\mathcal{D}_1, \mathcal{D}_2$, $\mathcal{D}_B$ and $\mathcal{D}_0$ respectively. These groups depend on prices, but for clarity I suppress this from the notation.

Note that types can be in both $\mathcal{D}_1$ and $\mathcal{D}_2$ if there is no bundling discount offered $p_B > p_1 + p_2$ but they still wish to purchase both products. Only if a bundling discount is offered $p_B < p_1 + p_2$ will anyone be in $\mathcal{D}_B$ and will that market have to separately clear on price.

The average cost in market 1 at prices $p_1; p_2, p_B$ is simply the expectation of the costs of all those buying product 1.

$$AC_1(p_1; p_2, p_B) = \mathbb{E}\left[\phi_1(w_1) \mid w \in \mathcal{D}_1\right]$$

and similarly for all those buying products 2 and B.

Following Einav et al. (2010) (and many subsequent papers) I define an equilibrium as a vector of prices that ensure zero profits are made in each market:

Definition 2. *Equilibrium prices* $\mathbf{p} = (p_1, p_2, p_B)$ *solve:*

$$p_m = AC\left(p_m, p_{-m}\right) \text{ for } m = 1, 2, B.$$

Following the literature, I make a single-crossing assumption about the slope of the profit curves.

Assumption 1. *For each market* $m = 1, 2, B$ *and all* p_{-m}*, I assume that* $p_m - AC_m(p_m; p_{-m})$ *crosses zero (from above) exactly once.*

This insures a unique partial equilibrium price for each market. I assume, and empirically verify, that there is a unique general equilibrium price vector (p_1^*, p_2^*, p_M^*).

Welfare in market $m = 1, 2, B$ at prices p_1, p_2, p_B that induces demand $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_B$ is the sum of consumer surplus and insurer profit:

$$\text{Welfare}_m = \int_{\mathcal{W}} \underbrace{(w_m - p_m)\,\mathbb{1}\left(w \in \mathcal{D}_m\right) dF(w)}_{\text{Consumer Surplus}} + \int_{\mathcal{W}} \underbrace{(p - \phi_m(w_m))\,\mathbb{1}\left(w \in \mathcal{D}_m\right) dF(w)}_{\text{Producer Surplus}}. \tag{2.1}$$

Overall welfare is simply the sum of the welfare in markets 1,2 and B. In equilibrium, profits will be zero, and so welfare will equal consumer surplus. Hence, when comparing equilibria, welfare will rise if and only if consumer surplus rises.

2.1.1 Correlation Order

The comparative static of interest is the correlation between risk types 1 and 2, all else equal. All else equal means, in this context, to fix the marginal distributions in the population of risk types 1 and 2 and vary only the correlation structure.

Define $\Gamma(F_1, F_2)$ to be the set of joint distribution functions with marginals F_1 and F_2. Following Shaked and Shanthikumar (2007)[(8)] and Denuit et al. (2006), the correlation order is defined as follows.

Definition 3. *Suppose $X, Y \in \Gamma(F_1, F_2)$ and have CDFs F_X, F_Y respectively. I say that X is be less correlated than Y or that X precedes Y in the correlation order, written as $X \precsim Y$ iff*

$$F_X(w_1, w_2) \leq F_Y(w_1, w_2) \quad \textit{for all } (w_1, w_2) \in \mathcal{W}. \tag{2.2}$$

This says that w_1, w_2 are more likely to both be small or both likely to be large[(9)] under the more correlated Y than under X. This comports with the intuitive notion of correlation and generalizes the familiar linear correlation coefficient. In particular, if $X \precsim Y$ then $\rho(X) \leq \rho(Y)$. In the case of jointly normal random variables, this implication goes both ways:

Example 1. *If both X and Y are jointly normally distributed with the same marginal distributions, i.e. $X \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho_X)$ and $Y \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho_Y)$, then $X \precsim Y \iff \rho(X) \leq \rho(Y)$.*

---

(8) There this ordering is called Positive Quadrant Dependance (PQD)

(9) Because X, Y have the same marginal distributions, $F_X(w_1, w_2) \leq F_Y(w_1, w_2) \iff 1 - F_X(w_1, w_2) \leq 1 - F_Y(w_1, w_2)$.

Note, the correlation order can be generalized be applied to other distributions related to X and Y. For example, the distributions of X and Y conditional on event A.[10]

2.2 Incentives to Bundle

Bundling is introduced by insurers exactly when it allows for 'cream-skimming' - when those attracted by a bundle discount are cheaper than those who buy the single products. This naturally depends on the correlation between WTP for the two different products, and how WTP relates to costs (adverse versus advantageous selection).

Label the prices in the no-bundling equilibrium as p_1^{NB}, p_2^{NB}. Figure 1 illustrates who buys product 1 and 2 at the separate market equilibria, and who would buy the bundled product if an ϵ discount were offered. In the separate market equilibria, everyone who has $w_1 \geq p_1^{NB}$ buys product 1 (groups A and B1) and everyone who has $w_2 \geq p_2^{NB}$ buys product 2 (groups A and B2). Group C buys nothing. Because the separate market is at equilibrium in which profits are zero, the average costs of those that buy product 1 is exactly equal to the price, and identically for product 2.

Suppose a firm starts selling the bundle at a small discount: $p_B = p_1^{NB} + p_2^{NB} - \epsilon$ for small $\epsilon > 0$. Those that buy the bundle are exactly those that separately bought both products prior - group I. The firm makes a profit offering a small bundling discount when the expected cost of insuring risk 1 is lower for group 1 than 2a, and for risk 2 when the expected cost of group 1 is lower than group 2b.

[10] I write $X \precsim_A Y$ for the correlation order, conditional on event A. A formal definition and discussion is in appendix A.4. The most common conditional correlation order I use is conditional on being in the set buyers for product $m = 1, 2$ and price p_m. Specifically, I write: $X \precsim_{p_m} Y$ to mean that the distribution of X, conditional on $w_m \geq p_m$ is less correlated than the distribution of Y, with the same conditioning.

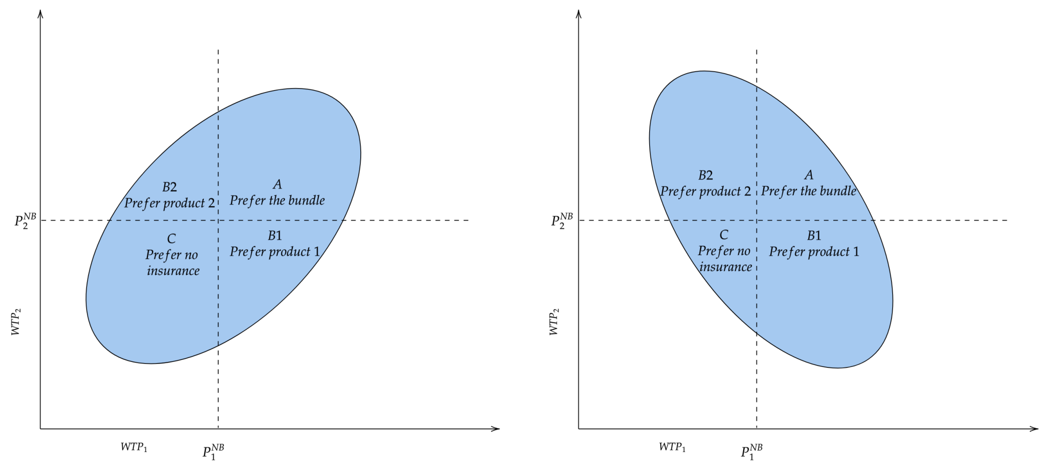

Figure 1: An illustration of which subsets of types buy in the separate markets, and which would swap to the bundled product if a small discount were offered. The left panel depicts the case of positive correlation between WTPs, the right panel depicts negative correlation.

Consider the left panel of figure 1. WTP for each separate product is positively correlated. In this case, those that buy the bundle - group 1 - have higher WTP than those who just buy the separate products - groups B1 and B2. If selection is adverse, higher WTP means higher cost, so that the expected cost of group A is higher than group B1 (for product 1) and B2 (for product 2). This makes bundling unprofitable: if selling product 1 to groups A and B1 broke even, selling it just to the higher cost group A will make a loss, and analogously for product 2. If, on the other hand, selection is advantageous, then high WTP means low cost, and so group 1 will be cheaper than group B1 or B2 for product 1 or 2 respectively, and bundling will be profitable. Hence, if WTPs are positively correlated, then bundling will occur if selection is advantageous, but not if it is adverse.

Inversely, consider the right panel in which the WTP's are negatively correlated. Here the set of people who would buy the bundle - group A - have lower WTP than the groups that would just buy the separate product. If selection is adverse, lower WTP means lower cost, making group A cheaper to sell to than groups B1 or B2 for product 1 or 2 respectively. This makes bundling profitable. If selection is advantageous, then lower WTP means higher cost, and hence group I is not profitable to sell to. Overall, the opposite conclusion applies: if WTPs are negatively correlated, bundling will occur if selection is adverse, but not if it is advantageous. This is formalized in the following proposition.

Proposition 1. *Starting from the no-bundling equilibrium, offering a bundling discount is profitable if and only if*[(11)] *there is:*

1. *Adverse selection and negative correlation*[(12)] *between risks; or*
2. *Advantageous selection and positive correlation between risks.*

This shows the conditions under which bundling will and won't occur. I now study how the introduction of bundling affects the separate markets in equilibrium.

2.2.1 Mixed Bundling Equilibrium

A mixed-bundling equilibrium is a set of prices p_1^{MB}, p_2^{MB} and p_B^{MB} in which the two separate markets as well as the bundled market break even, with a positive bundling discount offered: $p_B^{MB} < p_1^{MB} + p_2^{MB}$.

Proposition 1 shows that in markets with adverse selection bundling will occur iff there is negative correlation between the risk types, and in markets with advantageous selection iff there is positive correlation. Bundling occurs precisely when those that buy both products are cheaper than those that buy just one product. This implies that the mixed bundling equilibrium will feature higher prices in the separate product markets than the no-bundling equilibrium prices, and that those who buy the bundle will receive a discount: they will pay less than the sum of the prices in the separate product. This is formalized below:

[(11)] Assuming that the distribution is comparable in the correlation order to the independent joint distribution with the same marginals.

[(12)] Here, X having negative correlation formally means, for $X, Z \in \Gamma(F_1, F_2)$ with Z independent, $X \precsim_{p_m} Z$ for $m = 1, 2$, and inversely for positive correlation.

Proposition 2. *If the economy is adversely selected and negatively correlated or advantageously selected and positively correlated, a mixed bundling equilibrium will occur and*

- *The prices in the separate markets are increased relative to the no-bundling equilibrium prices:* $p_1^{NB} < p_1^{MB}, \quad p_2^{NB} < p_2^{MB}$;
- *Under adverse selection, the bundle price is discounted relative to the no-bundling equilibrium:* $p_B^{MB} < p_1^{NB} + p_2^{NB}$.

This coarsely speaks to the welfare trade-off from bundling. Those who buy the bundle typically do so at a discount relative to the no-bundling equilibrium (under adverse selection, at least) and are better off. This discount brings even more buyers into the market, partially solving the selection problem. But a bundling discount exerts an externality on the separate market: prices rise and demand contracts. I study the net welfare impact of bundling, and potential corrective government interventions, in section 4.

3 Empirical Tests of the Positive Theory - Spousal Bundling in LTC Insurance

In this section I test the positive prediction of the theory: bundling is used as a price discrimination device to attract low cost consumers. I study spousal discounts for long-term care insurance. First I document the substantial discounts given to spouses who both choose to buy LTC insurance policies, distinguish this from pricing on the characteristic of being married, and show that, in line with my model, bundling discounts exist here because they allow for cream-skimming of good risks from the separate markets.

LTCI Background Long-term care, either at home or at a dedicated facility, is a substantial financial and medical risk facing the elderly. Per Savell (2023), a healthy 65-year old female has a 56% chance of needing long-term care by age 89 (their life expectancy). For a 65-year-old male, it is 46% by 87. This rises further for the people who live longer than expected. If care is needed, it costs between \$55,000 (home-health care) and \$100,000 a year (nursing home), and the expected duration of required care is is between 2 and 3 years. Accounting for the timing of care and cost inflation, the expected present discounted cost of care is over \$120,000 for a healthy 65-year-old.

Despite this risk, long-term care insurance take-up is low. Administration for Community Living (2023) estimated that only 10% of 65-year-olds are insured. Different explanations for this low take-up include crowd-out by medicaid (Brown and Finkelstein (2008)), rejections (Hendren (2013)), expectations of informal care-giving by children (Ko (2021)), and behavioural frictions (Gottlieb and Mitchell (2020)). I demonstrate how bundling can also contribute to depressed demand.

3.1 LTCI Discounts for Spouses or Partners

Substantial discounts are offered to couples who both purchase LTCI: on average 25%, per Figure 6. This is computed from the universe of LTC policies in LTC Quote Plus (StrateCision (2022b)), which is software used by the insurance brokers and quotes from 9 different issuers. It shows that steep discounts for couples both buying LTC policies are the norm. Discounts range from a minimum of 19.7% to a maximum of 35%, where the sex, age, risk class and insurer have all been varied. The rest of this section shows that these bundling discounts are justified by cream-skimming, as the theory predicts.

This is in addition to smaller, separate discounts simply for being married. Discounting on marriage status is simply another form of risk classification, analogous to pricing on age. One can compare the average risk of the married and the single to understand the risk differences associated to marriage status. Pricing on observable characteristics of the insured is standard insurance practice. In contrast, offering a bundling discount when both individuals buy a policy is a screening device that can be thought of as implicitly pricing on spousal WTP.

Product	LTC
Mean Discount	25.4%
95% CI	(24.3%, 26.4%)
Range	(19.7%, 35.0%)
N	114

Figure 2: Average discount offered if both spouses buy an LTCI policy. Quotes were collected from all LTCI policies in StrateCision, for Male/Female at ages 55/65/75, risk classes Preferred/Standard/Substandard

3.2 Data and Risk Prediction

The data I use come from the Health and Retirement Study (HRS). It is a biennial survey that has been conducted since 1990, so that long-term such as as long-term care utilization or mortality have been observed for much of the sample. The HRS includes various measures of long-term care utilization, which are taken from interviews or from next of kin in the case of death. Binary measures of nursing home and home health care utilization are the primary outcomes of interest, although supplementary analysis uses measures of the duration of any nursing home stays. Moreover, the HRS collects data on whether the respondents have long-term care insurance and, if so, their premium.

The HRS also collects a wide variety of information about medical conditions, health status and any difficulties with the Activities of Daily Living (ADLAs) and the Instrumental Activities of Daily Living (IADLs). These data are important to the pricing of LTC insurance (and potential uninsurability) and they are highly predictive of subsequent LTC utilization. The partner/spouse of every HRS respondent is automatically surveyed as well, if the respondent has a spouse. This allows for risk and WTP for LTC to be matched across couples.

Data on ex-ante health status and ex-post utilization allow for a prospective prediction as to the probability of needing LTC at various time horizons. These will form the basis for my measures of respondent and spousal risk and WTP.

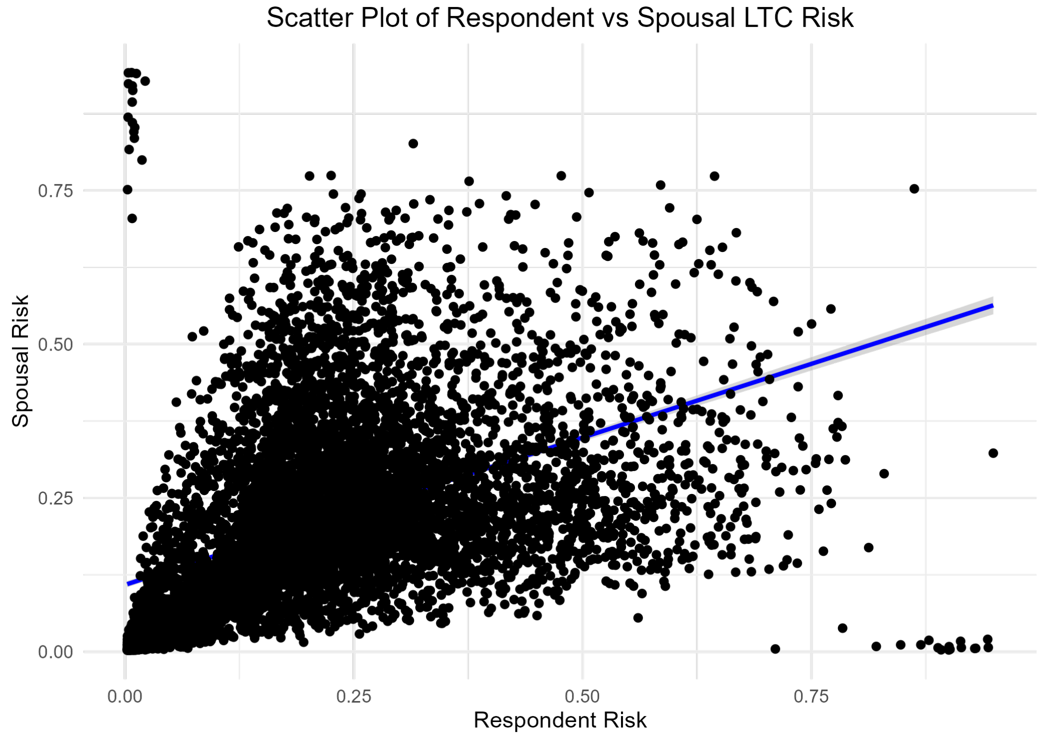

Figure 3: Scatterplot of LTC risk for the primary respondent and their spouse. A line of best fit is plotted in blue. Single respondents are excluded.

All predictions are via LASSO penalized logistic regression, and incorporate cross-validation to avoid over-fitting. Details of the variable used in the main predictions, and in the smaller predictions that mimic insurance company underwriting information, are in appendix B.6. The outcome being predicted is whether, in the next 14 years (7 HRS waves) the respondent will require home-health care or spend any time in a nursing home.

The correlation between predicted LTC risk for a respondent and their spouse can be seen in Figure 3. Simply for visual clarity, single respondents are excluded from this figure. They are included in all the analysis.

Figure 3 shows that there is a strong positive correlation between spousal risk: Low risk individuals are likely to have low risk spouses and vice versa. However, bundling screens on WTP, and the relationship between risk and WTP depends on whether selection is advantageous or adverse.

3.3 Advantageous Selection in LTC Insurance

I replicate the analysis performed by Finkelstein and McGarry (2006) on the larger HRS sample that has been collected since that analysis. I confirm, with more precision, their finding that long-term care insurance is advantageously selected. Combined with the positive correlation between spousal risk shown in Figure 3, this implies that people who have higher WTP spouses are lower risk.

To do so, I estimate probit models where the outcomes are binary indicators for using any LTC in the next 6 or 14 years, and a binary indicator for whether the respondent has LTC when surveyed. The independent variables are all variables that go into insurance company pricing, such as age, sex and major medical conditions.[13]

The results are in table 4 below.

LTC usage within:	6 Years		14 Years	
	No Controls	With Controls	No Controls	With Controls
Coefficient from Probit of LTC Usage on LTC Insurance	-0.07***	-0.04**	-0.05**	-0.03
	(0.02)	(0.02)	(0.02)	(0.02)
N	44,974	44,974	44,974	44,974

Note: ***/**/* means significant at the 1/5/10% level.

Figure 4: Testing for advantageous selection: the relationship between long-term care insurance and nursing home or home health care

The negative correlation between LTC Usage and LTC insurance is evidence of advantageous selections. This is consistent with Finkelstein and McGarry (2006). Their observation period was 5 years, and I find clearer evidence of selection over a 6 year rather than a 14 year time horizon. Nonetheless, for all the reasons in that paper, I take this as strong evidence of advantageous selection and proceed under this assumption.

Therefore, for the rest of this section, I assume that WTP decreases in risk. This allows for an **ordinal ranking** of individual's WTP without the strong

[13] For details see appendices B.6 and B.4.

assumptions needed for a cardinal measure.

3.4 High Spousal WTP Predicts Low Own Risk

An individual's risk is decreasing in their spouses (ordinal) WTP for LTC insurance. To establish this, I rank everyone according to their predicted risk. In this section, I avoid specifying a full model for WTP. I do that in section 5 and the conclusions are similar, albeit model-dependant. Here, I proceed only with an ordinal assumption consistent with advantageous selection: if person A has higher costs than person B, they have lower WTP. Hence, the WTP ranking is the reverse of the predicted risk ranking.

I bin individuals by the decile of their spouse's WTP, as well as a bin for single respondents. I compute the mean 16-year LTC risk in each bin. The results are in Figure 5. For reference, the black horizontal dashed line is the average risk for all those with a spouse.

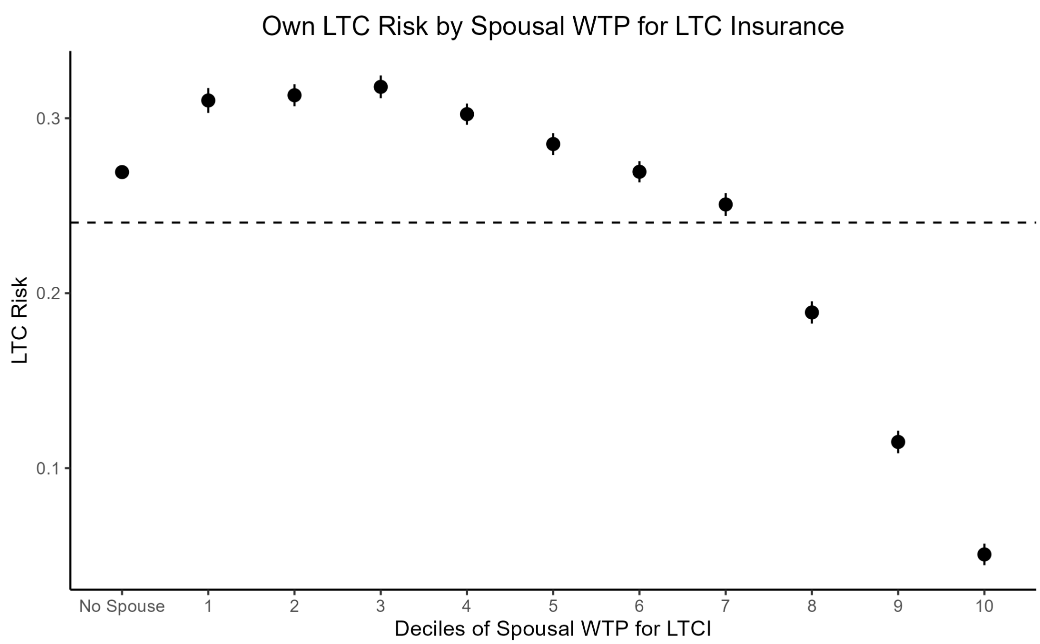

Figure 5: Binscatter of LTC risk by deciles of spousal (ordinal) WTP. The horizontal dashed line is the average risk amongst all those who have a spouse.

First, those who have a spouse are on average 10% lower risk than those without a spouse. Compare the average risk in the leftmost 'No Spouse' bucket - approximately 24%, with the average risk of all those with a spouse shown by the horizontal line - approximately 22%. This explains why simply being married attracts a small, but positive discount.

Second, *conditional on having a spouse*, a respondent's LTC risk decreases in their spouse's WTP. Those with high spousal WTP (the rightmost bins) have over 66% lower chance of using LTC within 14 years than those with the lowest spousal WTP (the bin second from the left).

This shows that a bundling discount plays a screening role over and above simply pricing on being married. Within those that are married, knowing that an individual has a high WTP spouse predicts low costs. This screening on WTP cannot be implemented by pricing on further observables, even observables that predict a high WTP, low cost couple. In appendix B.3 I show that, when only information available to the insurer is used, the difference in risk between the first and tenth decile of spousal WTP is one quarter the size of Figure 5. Hence, it is unobservable information, only implicitly revealed through WTP, that drives Figure 5.

Moreover, in appendix B.4 I show that on the intensive margin - the number of nights in a nursing home, conditional on entry - a similar pattern prevails: singles stay longer than the married, but within the married those with a high WTP spouse have shorter stays than those with a low WTP spouse.

This is one step away from a direct test of Proposition 1. Figure 5 conditions only on spousal WTP being low or high, without regard to own WTP. Proposition 1 only considers the predictive power of spousal WTP amongst those who themselves have high WTP. To precisely test Proposition 1, I define the percentage cost saving that a small bundling discount would attract, conditional on the current market 'price' being q.[14]

[14] Because, in this section, I do not specify a full cardinal model for WTP, q is not exactly a price. It is actually a risk level, and all those with lower risk (hence higher WTP, due to advantageous selection) buy the product. Nevertheless, for brevity I will continue to refer to q, since it does define who buys and who doesn't within the ordinal measure of WTP.

$$\Delta\% = E\left[\phi(w_1) \mid w_1 > q \wedge w_2 > q\right] - E\left[\phi(w_1) \mid w_1 > q \wedge w_2 \leq q\right] < 0. \quad (3.1)$$

Defined this way, $\Delta\%$ compares the average costs, amongst people who buy at price q, of those whose spouse who also buys at price q against those whose spouse doesn't buy at price q. Singles are included as having a spouse with zero WTP: they do not buy the bundle at any price. In the terminology of Figure 1, $\Delta\%$ compares the average cost of those in group A to those in group B1 (risk 1) or group B2 (risk 2). I randomize which spouse is counted as risk 1 versus risk 2, thereby ensuring symmetry. This symmetry means that whatever price I conjecture as the equilibrium price in the market for spouse 1 will also be the equilibrium price in market 2.

I compute $\Delta\%$ for different possible prices q, so as to avoid taking an exact stand on market primitives at this stage. The results are in Table 1. The different possible prices q I use are the deciles of the ordinal WTP distribution.

$LTCI$ Market ”Price”	$\Delta\%$
10th Pcile of WTP Distribution	-0.07^{***}
	(0.01)
20th Pcile of WTP Distribution	-0.11^{***}
	(0.01)
30th Pcile of WTP Distribution	-0.15^{***}
	(0.01)
40th Pcile of WTP Distribution	-0.2^{***}
	(0.02)
50th Pcile of WTP Distribution	-0.25^{***}
	(0.02)
60th Pcile of WTP Distribution	-0.29^{***}
	(0.02)
70th Pcile of WTP Distribution	-0.29^{***}
	(0.02)
80th Pcile of WTP Distribution	-0.28^{***}
	(0.02)
90th Pcile of WTP Distribution	-0.18^{***}
	(0.03)

*** = significant at the 1 % level.

Table 1: The difference in risk between the bundle buyers (region B1/B2, per Figure 1) and the non-bundle buyers (region A) at 9 different market prices. The prices tested are the deciles of the ordinal WTP distribution.

Table 1 demonstrates a strong rationale for bundling discounts. The cost saving amongst those who would buy a bundled product are between 7% and 29%, depending on the cutoff. This saving is present at all the different prices tested. I show in appendix section B.4 that there are also cost savings on the intensive margin: conditional on using a nursing home, nights spent there are approximately 40% lower for partnered versus unpartnered. Together, these facts justify the sharp discounts offered to partnered LTC insurance buyers, consistent with the cost-saving ’cream-skimming’ motive predicted by Proposition 1.

3.5 Ruling Out Alternate Explanations

I analyze standard (non-selection based) economies of scale, and differential lapsation, as competing explanations for the bundling discounts offered to couples in LTC insurance.

When a couple buys an insurance policy, there might be efficiencies in variable costs relative to single policies. This might explain some of the couples discount offered.

I offer evidence against this hypothesis. First, I study couples discounts in hybrid LTC/life insurance, pure life insurance, annuities and private disability insurance. For each of these, I compute the average couples discount offered, if any. Details of the data sources and assumptions are in appendix B.1. The results are reported in Figure 6.

Product	LTC	Hybrid LTC/LI	Life Insurance	Disability	Annuities
Mean Discount	25.4%	4.20%	1.01%	0%	-3.50%
95% CI	(24.3%, 26.4%)	(3.57%,4.84%)	(0.004%, 1.7%)	-	(-5.05%, -1.96%)
Range	(19.7%,35.0%)	(0.09%, 10.7%)	(0%, 33.3%)	-	(-11.3% 5.5%)
N	114	60	257	12	567

Figure 6: For LTC/Hybrid/LI/Disability, these are direct couples discounts. For annuities, it is the discount/cost of arbitraging a 50% joint and survivor annuity with two single life annuities. There were no couples discounts in any disability products surveyed.

We see that life insurance, disability and annuities attract, respectively, a very small couples discount, no discount, or an anti-discount.(15) A hybrid LTC/life policy attracts a 4% couples discount, which is between the large LTC couples discount and the tiny life insurance discount. Since the other products considered do not attract substantial bundling discounts, I rule out generic economies of scale as a justification for couples discounts in LTC.

(15)There is no explicit couples discount for annuities. I compute this anti-discount by comparing the cost of a joint life annuity to a payoff-identical pair of single life annities. Details are in appendix B.1.

This is consistent with intuition. The vast majority of the cost of servicing an insurance policy come from acquisition costs - primarily commissions payed to brokers and underwriting costs (see, for example, Tables 8 and 15 in Society of Actuaries (2010)). Commissions are proportional to premium, and underwriting costs are still paid for both members of the couple, ruling either out as potential economies of scale. Operational costs such as policy issue expenses that might plausibly be economized on are less than 6% of total overhead.

The largest expense other expenses are policy-owner services that primarily consist of lapsation and surrender costs. If lapsation patterns amongst couples were more favourable to the insurers than singles, this might explain some of discount offered. In appendix B.2 I show that couples lapse their policies at a lower rate than singles. LTC insurance policies are front-loaded: the premia are paid early and the claims many years later. This makes lapsation profitable for insurers. Hence, the fact that couples lapse at lower rates than singles pushes *against* discounts as it makes couples more expensive to insure.

4 Normative Theory: Welfare and Counterfactual Policies

In this section I theoretically study the welfare impacts of the introduction of bundling discounts. Further, I analyze when government interventions such as mandating or forbidding bundling can improve welfare. There are four types of market outcomes I will study throughout:

1. The **mixed-bundling** equilibrium is when both separate products are offered, and the bundle is offered at a discount to the sum of the separate prices. I label outcomes in this type of equilibrium with the superscript MB, e.g. p_B^{MB} means the price of the bundle in the mixed-bundling equilibrium.

2. The **no-bundling** equilibrium, when no bundle discount is offered and only the separate markets exist. This might occur because there is no in-

centive for a firm to bundle, or because the government outlaws bundling. I label the equilibrium objects in this case with a superscript NB, e.g. p_1^{NB} means the price of product 1 in the no-bundling equilibrium.

3. The **pure-bundling** equilibrium , where only the bundled product can be bought, will be denoted with a PB superscript, e.g. P_B^{PB}.

4. The **family-bundling** equilibrium, where I think of each type as a family, forces families that have two members to buy both policies or neither, but allows singles to buy only one policy. The family-bundling equilibrium forbids bundling discounts, and equilibrium prices are denoted p_1^{FB}, p_2^{FB}.

4.1 Mixed Bundling vs No Bundling

The characterization of the mixed bundling equilibrium in Proposition 2 showed that once bundling is introduced, prices in the separate markets will rise, and the bundle price will be lower than the sum of pre-bundling separate prices. This allows us to qualitatively analyze who is better off and who is worse off relative to the equilibrium without bundling. Consider figure 7, which illustrates the equilibrium prices under mixed bundling and no bundling, and which groups of people buy different combinations of products under each regime.

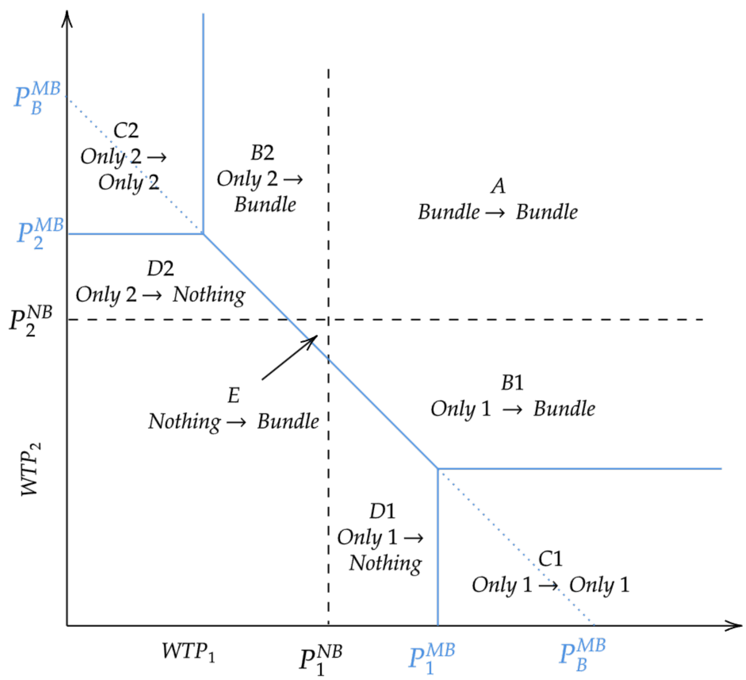

Figure 7: Equilibrium prices and buying patterns in the mixed-bundling and no-bundling equilibrium. The changes from no-bundling to mixed-bundling equilibrium insurance purchases for different groups are shown. $X \to Y$ means that the group chose X under no-bundling, and Y under mixed-bundling.

The separate equilibrium prices without bundling are p_1^{NB} and p_2^{NB}. Once bundling is introduced, the prices in the separate markets rise to p_1^B and p_2^B, and the bundle price is p_{Bundle}. Under the the no-bundling equilibrium groups A, B1, C1 and D1 buy product 1, while A, B2, C2 and D2 buy product 2. Under the mixed-bundling equilibrium, groups A, B1, B2 and E (if there is a discount relative to no bundling) buy the bundle, group C1 buys only product 1, and C2 only product 2.

Under adverse selection, groups A and E are better off. Group A buys both products in both the no-bundling and mixed-bundling equilibrium, but does so at a lower price in the latter. Under advantageous selection, as I discuss below, the bundle price might end up being higher than the sum of the no-bundling equilibrium separate market prices. In that case there is no group F, and group A can be smaller and worse off.

Under either adverse or advantageous selection, groups C1 and C2 are worse off. They still only buy one product, but at a higher price. Groups D1 and D2 used to buy something with positive surplus but now buy nothing.

Finally, recall that the second part of proposition 2 only holds for adversely selected markets. In advantageously selected markets, the bundling price can end up being higher than the sum of the separate no-bundling equilibrium prices. This is because, if a firm offers an ϵ bundling discount, and once prices in the separate markets adjust upwards, instead of selling product 1 to groups A, B1, C1 and D1, they now sell product 1 (and 2) to groups A, B1 and B2. In other words, in terms of product 1, the insurer has replaced groups C1 and D1 with group B2. Those in group B2 have lower WTP for product 1 than groups C1 and D1. Under adverse selection, this makes them cheaper, but under advantageous selection this makes them more expensive. Hence, the equilibrium increase in separate market prices can fully undo the cost-savings that motivated bundling, leaving everyone worse off. As I show in in section 5, this is the case in the long-term care market.

Overall, I find that the welfare consequences are ambiguous, but intuitively more likely to be positive under adverse than advantageous selection. This is because the marginal buyers attracted to a bundle are lower cost than the infra-marginal bundle buyer under adverse selection, but higher cost under advantageous selection. Given the ambiguous welfare consequences of bundling introduced by the private market, there is scope for welfare enhancing government interventions. I consider three interventions: mandating pure bundling, forbidding bundling and mandating family bundling.

4.2 Government Interventions: Pure Bundling

Once the government mandates bundling, only bundled risk type matters. How the distribution of bundled risk types relate to the separate risk types depends on the correlation structure and the type of selection.

Under positive correlation, the demand and cost curves in the bundled market will look similar to the separate markets. At the limit, with perfect positive correlation, the bundled market equilibrium will be identical to the separate market equilibria, in the sense that the price of the bundle will be the sum of the pre-mandate separate market equilibrium prices, and the same set of people will buy the bundled product as bought both product prior.

As correlation gets less positive or even negative, the bundled market looks more and more different from the separate markets. With less correlation, it is less likely that individuals have high WTP for both contracts, or low WTP for both contracts. It becomes more likely that they have high WTP for one and low for the other. This homogenizes and flattens the selection: there are few people with very high or very low WTP for the bundle, and many more in the middle.

A reduction of selection looks slightly different in an adversely versus advantageously selected market. In a bundled market with adverse selection, costs are highest for the first purchasers of insurance. As selection is lessened due to lower correlation, the first to buy insurance look more like the last to buy. That is, their costs decrease toward the average. Toward the limit of perfect negative correlation, average cost is almost flat with respect to WTP. The opposite is true in advantageously selected markets. In that setting, those that are the first to buy have the lowest costs. As correlation increases, the costs of the first types to buy move toward the average. That is, the average cost function increases point-wise.

This is formalized in the proposition below:

Proposition 3. *If the separate markets are adversely (advantageously) selected then in the bundled market:*

- *The average cost curve is everywhere lower (higher) under less correlated separate distributions.*
- *The equilibrium bundle price decreases (increases) as correlation decreases.*

Welfare. Moving from price changes to welfare is more difficult, since the demand curve, hence surplus and the equilibrium quantity, in the bundled market also changes. In appendix A.1, I make further assumptions on the demand curve and derive results about quantity changes.

I focus on a blunt but important fact: *forced bundling is always a bad idea for advantageously selected markets.* Suppose the pre-mandate prices are p_1, p_2, p_B (where perhaps there is no bundling). What happens if the government forces bundling and the bundle price is set to be the same as the pre-mandate price: $p_B^{PB} = p_B^{MB}$? In the terminology of 7, when bundling is forced: the upper portion of group C1 moves from buying just product 1 to the bundle, and the upper portion of group C2 moves from buying product 2 to the bundle. On net, the group of people who end up with product 1 has decreased in WTP :the lower portion of C1 has been replaced with the upper portion of C2. Similarly for product 2. Under advantageous selection this means average cost has increased. Since pre-forced bundling profits were zero in all markets, selling to a higher cost set of buyers at a price lower than the sum of pre-forced bundling profits separate prices cannot be profitable. Hence, the bundling price under forced bundling will have to rise. Consumers are choosing from a smaller set of products at higher prices, and so their consumer surplus falls. Since profits are zero in equilibrium, this means welfare also falls. This is formalized below:

Proposition 4. *If the government forces the products to be bundled:*

- *Under adverse selection, the effects on welfare are ambiguous.*
- *Under advantageous selection, under approximate symmetry*[(16)] *in the risks, the price of the bundle rises and welfare unambiguously falls.*

4.3 Government Interventions: Family Bundling

Family bundling synthesizes the potential benefits of forced bundling while not disadvantaging singles. Bundling discounts are disallowed, families with two

[(16)] Formally, that the density of types in B2 is equal to D1, defined at the mixed-bundling equilibrium prices.

members must either both buy policies or neither can. Singles face the same choice, and I code them as having zero WTP for the other contract.

The logic is similar to pure bundling. Family bundling constrains the choice set of couples. They must either buy an extra policy for a family member they preferred not to insure or forego coverage altogether. This initially hurts these families. Whether overall welfare can rise depends on the extent prices fall due to the blunting of selection. When prices rise after forced bundling, welfare unambiguously falls: everyone is worse off. As in the case, family bundling typically leads to higher prices under advantageous selection, but lower prices under adverse selection.

This speaks to the empirical work in Nguyen (2022). That paper found that, empirically, family bundling for adversely selected health risk increased welfare. The theory suggests that this is driven by adverse selection, and not guaranteed. Forced bundling with advantageous selection, such as in the LTCI setting I study in section 5, I find to be bad for welfare.

4.4 Government Interventions: Banning Bundling

An alternative intervention the government might undertake is forbidding bundling discounts, i.e. enforcing separate markets. For example, while household discounts are allowed in Medigap insurance, they are not allowed in Part D or Medicare Advantage (Part C) plans.

If there is bundling to forbid, then per Proposition 1, the market is either adversely selected and negatively correlated or advantageously selected and positively correlated. The welfare implications are ambiguous and the inverse of those discussed in section 2.2.1. Per Figure 7, forcing the market back to separate equilibria with no bundling will decrease the prices in the separate market, benefiting groups C1/C2 and D1/D2. Groups A and E (when the latter exists) are worse off under adverse selection (where Proposition 2 guarantees the bundling price will be discounted relative to sum of the separate market price), but might be better off under advantageous selection.

The overall welfare impact of this intervention is specific to the empirical setting. However, all else equal, I expect it to be more likely to increase welfare under advantageous rather than adverse selection, as the bundling discount (that is being forbidden) is typically higher in the latter. Indeed, in section 5, I find this: in the LTCI market, which is advantageously selected, disallowing bundling raises welfare.

5 Equilibrium and Welfare Consequences of LTCI Bundling

I return to the market for long-term care insurance and analyze the impacts of bundling discounts on equilibrium and welfare. Specifically, I compare the status quo equilibrium that prevails in the private market - where both single products and bundling discounts exist - to two counterfactual policies: banning bundling and family bundling (all family members must buy, or none can). I find that family bundling lowers welfare while banning bundling increases it.

5.1 Market Simulation Details

I study the LTCI market for 60-65 year old couples consisting of one man and one woman, or singles. Of course, same-sex spouses/partners are eligible for the same discounts, but for brevity I focus on heterosexual partnerships. To simulate equilibria, I require a measure of WTP and costs for each person.

Estimating Costs

In section 4, I predicted LTC risk over the next 14 years for each individual and their spouse. However, welfare simulations require an estimate of expected dollar costs. There are two additional steps needed. First, we need to extrapolate the estimate of LTC risk in the next 14 years to LTC risk over the whole lifetime. Since our respondents are at most 65, much of the long-term care risk is expected to occur in more than 14 years. Second, we need to translate the estimate of LTC risk (a probability) into an expected cost (in dollars) that accounts for possible LTC cost inflation, and is discounted appropriately.

Costs are extrapolated from the estimates of risk described in section 3.2. The predictions only cover LTC usage in the next 14 years. To account for risk after 14 years, I use computations of LTC risk at each 5-year age bucket from Savell (2023) (which are conditional on survival) and estimates of population survival from Social Security Administration (2023). Using these, I compute the lifetime probability of long-term care for an average 65 year old. I then scale the personalized predicted risks from section 3.2 so that the mean risk in the distribution of HRS predictions matches the average lifetime risk computed from Savell (2023) and Social Security Administration (2023). This procedure produces an estimate of lifetime LTC risk for each individual in the sample.

To convert lifetime risk into dollar costs, I take the cost of a year long-term care ($69,508, a blended average of nursing home and home health care) from Savell (2023), multiply by the gender-averaged years of care required conditional on requiring care (2.26 years), and take the present discounted value, accounting for expected LTC cost inflation and time to care being needed, using the cost deflators in that report.

I present the primary results with expected costs defined as above. However, to make sure the findings are robust to the numerous choices made to define expected costs, in section 5.3 I run the same analyses with a wide range of (distributions of) expected cost, and the sign of the welfare conclusions obtained are identical.

Estimating WTP

I would like to regress demand for LTCI on premia and recover an estimated WTP. There are two issues: first, I only know the LTC premia for those that hold LTCI; second, the classical endogeneity issues between premia and WTP are problematic in this setting in which equilibrium price changes based on who buys insurance.

First, to impute the premium that those who do not hold LTCI would have to pay, I rely on the structure of LTCI pricing. LTC is only priced on sex, age, marital status and risk (coarsely). In the sample, approximately 13% of individuals hold LTC insurance. I observe these individual's premiums. Hence, given the observable information on variables used in pricing, I can impute the premia that would be offered to those without LTCI.

Second, to overcome the endogeneity problem, I use the sharp increases in premia throughout the sample period (1996 to 2006) owing to incorrect insurer assumptions regarding lapsation and longevity. Premium increases were sharp, sometimes over 40% *each year* (Duhigg (2007)), and applied to existing and new policies. By regulation, premium increases have to be approved by state insurance commissioners. Rate increases are only allowed when the insurer can prove that costs are higher than expected, which is justified by historical claims experiences (Genworth (2023)). The identifying assumption is that these changes over time shocked premia without affecting unobservable demand factors such as the outside option of self-insuring. The main threat to this is if individuals are learning from rate hikes that their expected cost of care is likely to be higher should they self-insure, possibly affecting demand for formal insurance. Individuals, if they update at all, should do so all at once, whereas changes in premia over time are only allowed as the historical loss experience actualizes, even if it is clear that future claims are likely to be far higher. For this reason, I do not pursue this issue further.

I model the purchase of LTCI as a logistic function of price p and covariates X that include age, sex and risk (which is itself a summary prediction using many risk factors). I instrument p with the calendar year in which the individual is surveyed as a proxy for premium growth over time for the reasons explained above.

Given the demand model, I impute each individuals WTP to rationalize the rate of actual LTCI purchase in the sample. In particular, for each individual I find the price p_i^* at which the predicted probability from the demand model equals the rate of LTCI purchase in the HRS data - a matching of moments. I set individual WTP to be equal to this pivotal price: $WTP_i = p_i^*$.

Since demand varies with individual characteristics, so will WTP. For example, demand strongly increases in income, hence so does WTP. The premium and hence the WTP are expressed as recurring monthly amounts. To match costs, I convert this to a expected present discounted value based on expected years of longevity at 65 and an interest rate of 5%. The average WTP in these terms is approximately \$25,000, over 30% lower than the average expected cost of \$35,000.

5.2 Equilibrium, Welfare and Counterfactuals

I define the markets under study as the M(ale) LTCI market, the F(emale) market and the B(undled) market, unless the regulator counterfactually outlaws the latter. Each type w is a family in the HRS that consists of a single individual, or a male and female couple. WTP for LTCI for the male is w_M, for the female w_F and for the bundle is $w_M + w_F$. In the case the family consists only of a single man or women, I set w_F or w_M to be zero respectively. As in section 2, costs are $\phi_M(w_M), \phi_F(w_F)$ and $\phi_M(w_M) + \phi_F(w_F)$ for each of the three products. Given prices, the set families who choose to buy product $m = M, F, B$ is denoted $\mathcal{D}_m$ respectively, as defined in section 2.

I compute equilibrium prices that solve $p_m = AC\left(p_m, p_{-m}\right)$ where $m = M, F$ and B when the latter is not banned. To remove lumpiness, smooth out the market and finite sample issues, I duplicate the data and add small random noise to costs and WTP. Since, by construction, equilibrium prices ensure zero profits, welfare is the sum of consumer surplus in the three markets:

$$\text{Welfare} = \sum_{m=M,F,B} \int_{\mathcal{W}} \left(w_M - p_M\right) \mathbb{1}\left(w \in \mathcal{D}_m\right) dF(w). \tag{5.1}$$

5.2.1 Equilibrium Simulations: Mixed Bundling, Banned Bundling, Forced Bundling

I simulate the equilibrium under three regimes. First, where firms are free offer products separately or together with a bundling discount. This mimics the current market conditions. Second, a counterfactual where no bundling discount is allowed, so that only the separate markets exist. Third, a counterfactual where family bundling is forced: either all members of the family buy (for singles, this is just themselves) or none do.

The equilibrium prices, expressed as premia per month, welfare changes relative to the mixed-bundling status quo, and proportions of: singles who buy insurance, couples who buy insurance for one member, and couples who buy insurance for both members, are in Table 2. A scatter plot that illustrates who buys which products in each of the regimes is in Figure 8.

	Mixed Bundling (status quo)	Family Bundling	Bundling Banned
Equilibrium Prices:			
p_M	\$373.80	\$366.65	\$338.05
p_F	\$473.76	\$380.73	\$427.85
p_B	\$743.47	-	-
Proportion Insured:			
Singles	2.2%	5.4%	4.2%
Couples With One Insured	7.3%	-	18.8%
Couples With Both Insured	19.4%	22.8%	12.5%
Welfare Change (Relative to Mixed Bundling)	-	-5.1%	+4.7%

Table 2: Results of Equilibrium Simulations

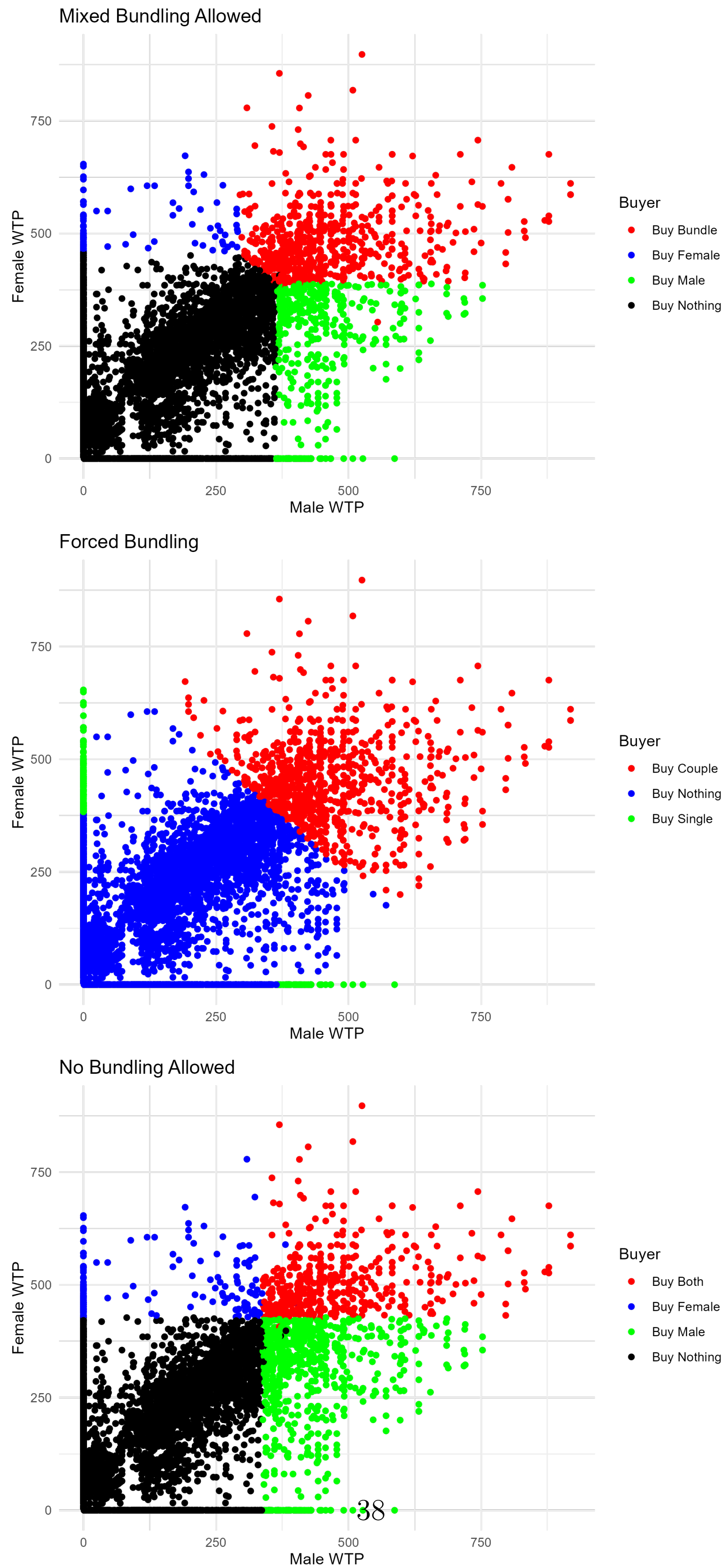

Figure 8: Illustrations of demand for male LTCI, female LTCI and both/bundle. The top, middle and bottom panels respectively illustrate the regimes of mixed bundling (no regulatory intervention), forced bundling, and

The mixed bundling regime that has no restrictions on the products that can be offered closely mimics the empirically observed market. In the simulated equilibrium, 11.1% of individuals buy LTCI, compared to 13% in the HRS data. The equilibrium prices are close to actual premia in the market: $\approx$ \$250 to \$400 for a 65 year old male (depending on product and underwriting class) and $\approx$ \$350 to \$550 for a 65-year-old female. The simulated bundling discount offered is 19% of the male premium and 15% of the female premium.

Forced Family Bundling

The first counterfactual of interest is forced family bundling. This means that couples have to either both buy insurance, otherwise neither can buy. Singles choose either to buy just the gender-appropriate policy, or not.

Forced bundling leads to reduced prices and increased quantity insured. Comparing the second to the first column of Table 2: the equilibrium price for men falls from \$373.80 under mixed bundling to \$366.65 under forced bundling, but for women falls from \$473.76 to \$380.73. However, notice that the sum of the prices under family bundling is \$747.47, four dollars more than the bundle price under the status quo mixed bundling. This is consistent with Proposition 4. These price changes help singles, 5.4% of whom now buy insurance instead of 2.2%. Family bundling lowers total demand for couples (22.8% buy under family bundling compared to 26.7% = 7.3% + 19.4% who bought something under mixed bundling). This is due to a) couples who previously bought both now choose to buy nothing as the price has increased, and b) many couples who previously bought one policy now choose no insurance at all.

Forced bundling has multiple effects on welfare. First, the price for a couple to both be insured has increased, since the high cost singles are now mixed in with the cheap couples who previously bought the bundle. Second, couples no longer have the option of insuring just one member. This contraction of the choice set lowers their surplus. Third, singles are better off, especially women. Ignoring the marginals (whose surplus is near zero), only 2.2% of singles are infra-marginal to the price reduction under family bundling, whereas 19.4% of couples are impacted by the price rise for the bundle, and 7.3% are mechanically made worse by not allowing half a couple to be insured. This is why the benefits to the singles are overpowered by the costs to the couples, and on net welfare declines.

Banned Bundling

The second counterfactual I study is when the regulator bans bundling discounts from being offered. Unlike forced family bundling, couples can still buy just no policy, one policy, or both, but the latter cannot be discounted.

When bundling is banned the separate market prices fall. This is true generally, as proposition 2 shows. Bundling occurs exactly because those that buy the bundle are cheaper than those that do not. Naturally, when bundling is banned and the former are mixed into the latter's separate policy, the cost falls.

When bundling is banned, welfare increases. Singles and couples who bought only one policy are much better off: prices have dropped by about \$40. Couples who who previously bought the bundle at a discount are slightly worse off. This is because the sum of the post-ban separate prices is $\$338.05 + \$427.85 = \$765.90$, \$22 more than the bundled price under mixed bundling. However, the large price reduction for those buying one policy leads to demand expansions: 4.2% of singles buy a policy (instead of 2.2%) and 31.3% of couples buy at least one policy (compared to 26.7%). Overall, the demand effect and the reduction in single prices outweigh the removal of the bundling discount, and welfare increases.

This highlights why observing a large bundling discount can be misleading in terms of welfare. The bundling discount is calculated relative to separate market prices that are distorted upwards by the existence of bundling. The bundling discount is nowhere near as large when compared to the prices that would prevail in the separate markets if bundling were banned.

5.3 Robustness

Our estimates of costs and WTP for each individual in the HRS involved multiple data choices. To check that the welfare conclusions obtained above are not sensitive to these choices, I re-run the analysis above under different costs assumptions in a flexible and parsimonious way.

Our baseline assumption for the cost of LTC care, conditional on it being needed, was approximately \$157,000 (\$67,000 per year $\times$ 2.26 expected years of care). This was then discounted and multiplied by each individual's predicted probability of care to estimate their expected costs. Hence, to check for robustness, I vary the expected cost of care, should it be need (i.e. I vary the \$157,000). Given the linearity of expected cost, this proxies for changing the discounting assumptions as well. Moreover, since there is a scale invariance between WTP and costs (if I multiplied them both by 2, the equilibrium diagrams would look the same, just with axes relabelled), checking for robustness to costs is structurally similar to checking for robustness to WTP.

I vary the expected cost of care 10% either way from baseline. Personalized expected costs are then computed as before, by multiplying each individual's probability of needing care, as estimated in section 3.2, with the varied expected cost of care. I compute the three equilibria (mixed bundling, forced family bundling and no bundling allowed) as before, and compare welfare. The welfare results are in Figure 9 below. For each cost assumption I express welfare relative to the mixed bundling status quo.

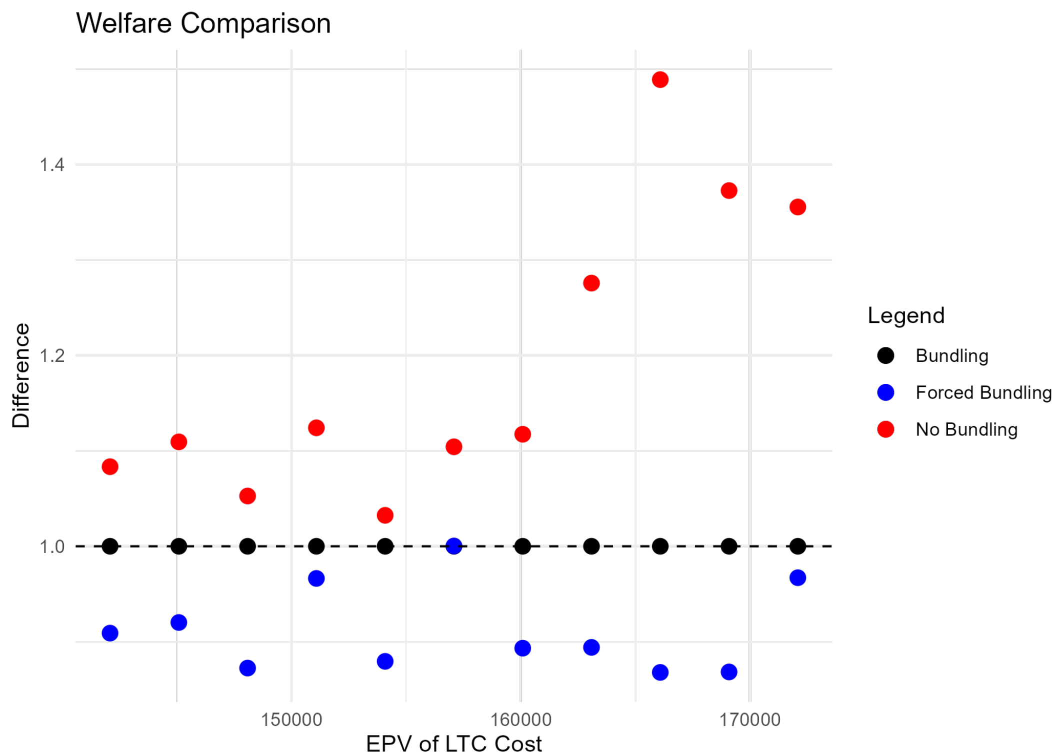

Figure 9: Welfare in the three regimes (mixed bundling, forced family bundling and no bundling allowed) under different cost assumptions. For each cost assumption, welfare is relative to the mixed bundling status quo.

The welfare results are robust to different cost assumptions. Banning bundling increases welfare, forcing family bundling decreases it. Note that as costs increase, the markets get smaller and less numerically stable, and so the quantitative magnitudes should not be taken too literally. Nevertheless, the qualitative insights are clear. The private market equilibrium that features mixed bundling is not optimal, and government interventions that ban bundling would improve welfare in the market.

6 Conclusion

This paper offers a theoretical foundation and empirical evidence for the pervasive practice of bundling insurance products. The theory predicts that bundling is a cream-skimming device: firms will bundle if and only if those that buy both products are cheaper than those who buy one. For the planner, bundling is useful when it homogenizes the risk pool, which occurs when the risks are more negatively (or less positively) correlated. Empirically, I study LTCI. I offer evidence consistent with firm behavior that the model predicts: discounts are offered to couples who both buy LTCI because they are lower risk than those who would only buy for one member. Using an estimated model for costs and WTP for LTCI, I study counterfactual regulations. I find that a ban on spousal bundling discounts would raise welfare, and compulsory spousal purchase would lower it.

Broadly this paper speaks to a literature about the optimal scope of insurance products: which risks should be insured together and which separately. The framework developed here could be applied different markets to clarify the limits and robustness of the empirical findings. Alternate bundling combinations (in terms of different risk products or different people) can be considered. For example, which services are included and excluded from the design of Medicare Part C contracts, or contracts that combine long-term care insurance with annuities or life insurance can be evaluated in this framework.

Moreover, the use of credit scores in the pricing of various insurance products is a form of bundling of two contracts. For example, mortgages and auto insurance are both selection markets with private information, and a good credit score provides a discount in both. Current or counterfactual regulations that mediate this can be analyzed through the framework proposed here. Likely those with good credit who want auto insurance are better off under the current mixed-bundling regime, but this might be outweighed by the damage done to those who are safe drivers but less safe borrowers, or vice versa.

Bibliography

Adams, W. J. and J. L. Yellen (1976). Commodity Bundling and the Burden of Monopoly. *The Quarterly Journal of Economics 90*(3), 475–498.

Administration for Community Living (2023). How much care will you need? Accessed: 2023-11-07.

Akerlof, G. A. (1978). The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning. *The American Economic Review 68*(1), 8–19.

Bakos, Y. and E. Brynjolfsson (1999, December). Bundling Information Goods: Pricing, Profits, and Efficiency. *Management Science 45*(12), 1613–1630.

Brown, J. R. and A. Finkelstein (2008, June). The interaction of public and private insurance: Medicaid and the long-term care insurance market. *American Economic Review 98*(3), 1083–1102.

Chen, Y. and M. H. Riordan (2013). Profitability of Product Bundling*. *International Economic Review 54*(1), 35–57.

Cooper, A. L. and A. N. Trivedi (2012). Fitness memberships and favorable selection in medicare advantage plans. *New England Journal of Medicine 366*(2), 150–157. PMID: 22236225.

Denuit, M., J. Dhaene, M. Goovaerts, and R. Kaas (2006, May). *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. John Wiley and Sons.

Duhigg, C. (2007, 03). Aged, frail and denied care by their insurers. *The New York Times*. Accessed: 2023-11-14.

Einav, L., A. Finkelstein, and M. R. Cullen (2010, August). Estimating Welfare in Insurance Markets Using Variation in Prices*. *The Quarterly Journal of Economics 125*(3), 877–921.

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

Genworth (2023). Long-term care insurance premiums: Faqs. Accessed: 2023-11-14.

Gottlieb, D. and O. S. Mitchell (2020). Narrow framing and long-term care insurance. *Journal of Risk and Insurance 87*(4), 861–893.

Gropper, M. J. and C. M. Kuhnen (2021). Wealth and insurance choices: Evidence from us households. Technical report, National Bureau of Economic Research.

Haghpanah, N. and J. Hartline (2021, May). When Is Pure Bundling Optimal? *The Review of Economic Studies 88*(3), 1127–1156.

Hendren, N. (2013). Private Information and Insurance Rejections. *Econometrica 81*(5), 1713–1762.

Hurkens, S., D.-S. Jeon, and D. Menicucci (2019, August). Dominance and Competitive Bundling. *American Economic Journal: Microeconomics 11*(3), 1–33.

ImmediateAnnuities.com. Annuity shopper buyer's guide.

Klosin, S. and A. Solomon (2024, January). Optimal insurance scope: Theory and evidence from us crop insurance. Manuscript in preparation.

Ko, A. (2021, 11). An Equilibrium Analysis of the Long-Term Care Insurance Market. *The Review of Economic Studies 89*(4), 1993–2025.

Kotz, S. (2000). *Continuous multivariate distributions. Volume 1, Models and applications* (2nd ed. ed.). Wiley series in probability and statistics. New York: John Wiley and Sons, Inc.

Lavetti, K. and K. Simon (2018, August). Strategic Formulary Design in Medicare Part D Plans. *American Economic Journal: Economic Policy 10*(3), 154–192.

McAfee, R. P., J. McMillan, and M. D. Whinston (1989, May). Multiproduct Monopoly, Commodity Bundling, and Correlation of Values*. *The Quarterly Journal of Economics 104*(2), 371–383.

Mutual of Omaha Insurance Company (2023). *Long-Term Care Insurance Underwriting Guide.* For Producer Use Only. Not for use with the general public.

Nalebuff, B. (2004, February). Bundling as an Entry Barrier*. *The Quarterly Journal of Economics 119*(1), 159–187.

Nguyen, A. (2022, May). Household Bundling to Reduce Adverse Selection: Application to Social Health Insurance. Technical Report ID 3173424, Rochester, NY.

Savell, M. (2023, March). Long-term care and financial planning. `https://hvsfinancial.com/white-papers/ltc-and-financial-planning/`. Accessed: 2023-11-14.

Schmalensee, R. (1984). Gaussian Demand and Commodity Bundling. *The Journal of Business 57*(1), S211–S230.

Shaked, M. and J. G. Shanthikumar (2007, April). *Stochastic Orders*. Springer Science & Business Media. Google-Books-ID: rPiToBK2rwwC.

Shepard, M. (2022, February). Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange. *American Economic Review 112*(2), 578–615.

SOA (2016). 2009-2016 individual life insurance mortality experience report. Technical report, Society of Actuaries.

Social Security Administration (2023). 2020 period life table. Social Security Administration Actuarial Life Table. Accessed: 2023-11-14.

Society of Actuaries (2010). 2010 individual life and annuity expense study. `https://www.soa.org/research-reports/2010/research-2010-ind-life-annuity-exp-study/`. Accessed: 2024-01-10.

Stigler, G. J. (1963). United states v. loew's inc.: A note on block-booking. *The Supreme Court Review 1963*, 152–157.

StrateCision (2022a, June). Combocompare. https://www.ltca.com/Combo.htm. https://www.ltca.com/Combo.htm.

StrateCision (2022b, June). Ltc quote plus. https://www.ltca.com/qp.htm. https://www.ltca.com/qp.htm.

Zhou, J. (2017). Competitive Bundling. *Econometrica 85*(1), 145–172.

Zhou, J. (2021, June). Mixed bundling in oligopoly markets. *Journal of Economic Theory 194*, 105257.

For Online Publication Only

A Theoretical Appendix

A.1 Further Analysis of the Forced Bundled Market

To obtain sharper results in this appendix make assumptions about how the marginal distributions relate to the joint distribution. For ease of exposition I will focus on joint normality, but alternative assumptions suffice as well. I can now consider the impact on the (forced) bundled market of more or less correlation.

First, the WTP curve for the bundled product, i.e. demand. Consider the two extreme case: perfect positive correlation. Under perfect positive between the two risks, the high risk in market one are exactly the high risk in market two. The WTP for the bundled product is the (vertical) sum of the WTP for the two separate policies. This is illustrated as the black WTP curves in Figure 10.

As the risks become less positively or negatively correlated, the WTP curve rotates. The less positively correlated, the less likely is an individual with high WTP for risk 1 to also be high WTP for risk 2, and inversely for low WTP. This means tgere are fewer people with very high or very low WTP for the bundle, and many in the middle. These rotated curves are the red WTP' curves in Figure 10.

Since each individual's cost is a function of their WTP for each risk, the effect on the marginal cost curves is analagous. Under perfect positive correlation, the marginal cost curves in the bundled market are the sum of the marginal cost curves in the separate market. These are the black MC curves in Figure 10. As correlation decreases, the marginal cost curves rotate. These are the red MC' curves in Figure 10. Under adverse selection, in the same way as the WTP curve, since lower WTP means lower MC. Under advantageous selection, lower WTP means higher cost, and so the rotation is inverted.

Perhaps surprisingly, the average cost curve does not rotate, instead as correlation weakens or becomes more negative, the AC curve monotonically decreases (increases) when selection is adverse (advantageous). Under adverse selection, as correlation weakens, the marginal cost curve falls for low q and rises for high q. Clearly this means the average cost curve must fall for low q. But that is sufficient to counteract the subsequent rise in marginal costs and allow the average cost curve to always fall. Inversely for advantageous selection.

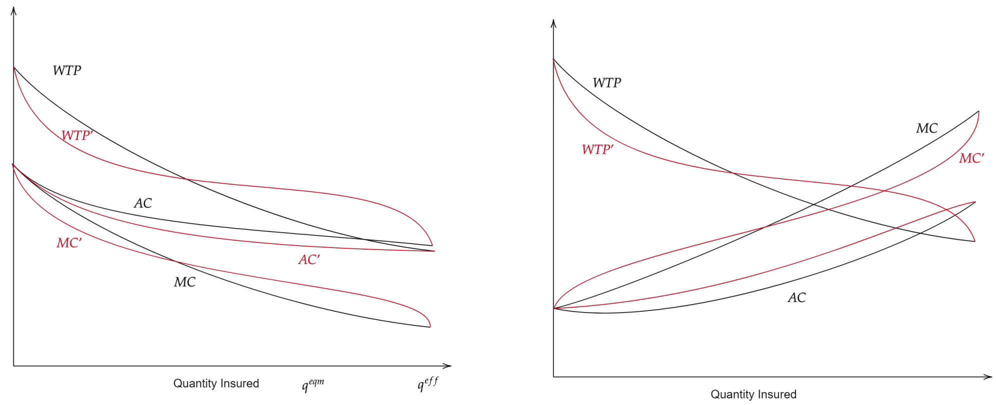

Figure 10: Equilibrium in the bundled market under adverse selection (left panel) and advantageous selection (right panel). The curves in black are under perfect positive correlation, in which the bundled market equilibrium is identical to the sum of the no-bundling separate equilibria. The curves in red are under less than perfect positive correlation.

The market failiure due to selection is due to the correlation between WTP and costs. More precisely, even if WTP differed between individuals, if they all had the same cost there would be no market failiure. When correlation gets weaker or more negative, average cost curves fall (rise) for adverse (advantageous) selection. In both cases, this is an example of costs becoming more homogeneous, in the sense that average cost is moving closer to marginal cost. This, conceptually speaking, weakens the underlying cost heterogeneity that is leading to a market failure. At the limit, under perfect negative correlation, costs are constant and there is no market failure.

What happens to equilibrium quantity as correlation changes is harder to say in general. Equilibrium quantity is determined by the intersection of average cost and WTP. Correlation monotonically moves average cost, but rotates WTP. Hence, we can sign quantity changes only when are sure that we are in the part of the WTP curve moving in opposite direction to the average cost curve. That means, under adverse selection, if the market is 'large' (in that equilibrium q with perfect positive correlation is beyond the rotation point) then we can be sure weaker correlation leads to a greater quantity insured. Conversely, under advantageous selection, if the market is 'small', weaker correlation leads to lesser quantity insured. The preceding discussion is formalized in the following proposition.

A.2 Microfoundation

In this section I present a microfoundation for the individual WTP and cost functions that were taken as given in the main body of the paper.

Each agent begins with wealth w. There are two risks that might occur, causing a loss of l_1 and l_2 dollars respectively. Each agent i privately knows their probability of each risk occurring. The risks occur independently.[17] Label these p_1^i, p_2^i. Consumption utility is evaluated according to utility function $u(\cdot)$ which is twice continuous differentiable. The expected utility that an individual would derive without any insurance is

$$V^i = (1-p_1^i)(1-p_2^i)u(w)+p_1^i(1-p_2^i)u(w-l_1)+(1-p_1^i)p_2^i u(w-l_2)+p_1^i p_2^i u(w-l_1-l_2).$$

Suppose two fixed insurance contracts are offered to this agent. An insurance contract for risk 1 specifies a premium $\rho_i$to be paid in all states of the world and an indemnity ι_1 to be paid from the insurer to the insuree in case risk 1 occurs. I summarize the random variable that the insurance contract offers by the consumption vector in the no loss and loss state of the world: $I_1 = (-p_1, -p_1+\iota_1)$, in place of the the raw loss random variable $L_1 = (0, -l_1)$. Similarly for risk 2.

Individual i's WTP for insurance contract 1 generally depends on whether they are insured against risk 2 or not. That is, their WTP for insurance contract 1 when they are insured against risk 2, $WTP_{1,B2}$ will differ from their WTP when they are not insured against risk 2 $WTP_{1,NB2}$.

These will respectively solve:

$$\mathbb{E}\left[u\left(w - WTP_{1,B2} - I_2\right)\right] = \mathbb{E}\left[u\left(w - L_1 - I_2\right)\right]$$
$$\mathbb{E}\left[u\left(w - WTP_{1,NB2} - L_2\right)\right] = \mathbb{E}\left[u\left(w - L_1 - L_2\right)\right].$$

Below I present a sufficient condition for $WTP_{1,B2} = WTP_{1,NB2}$ and hence for the seperability assumption in section 2 that there is just one WTP_1 and WTP_2 regardless of whether the other contract is purchased.

Definition 4. *The utility function $u(\cdot)$ is of CARA form when the coefficient of absolute risk aversion, $-u''(c)/u'(c) = \alpha$ is constant.*

[17] The distinction drawn throughout this paper is between correlation in *probabilities* (which I focus on) and correlation in *outcomes* which in this paper I assume away.

Proposition 5. *Suppose $u(\cdot)$ is of CARA form. Then $WTP_{1,B2} = WTP_{1,NB2}$.*

Proof. First note that for CARA u and independent risks X_1 and X_2 it is the case that $\mathbb{E}(u(X_1 + X_2)) = \mathbb{E}(u(X_1)) \cdot \mathbb{E}(u(X_2))$. Hence, we have that

$$\begin{aligned} &\mathbb{E}\left[u\left(w - WTP_{1,NB2} - L_2\right)\right] = \mathbb{E}\left[u\left(w - L_1 - L_2\right)\right] \\ \iff &\mathbb{E}\left[u\left(w - WTP_{1,NB2}\right)\right]\mathbb{E}\left[u\left(-L_2\right)\right] = \mathbb{E}\left[u\left(w - L_1\right)\right]\mathbb{E}\left[u\left(-L_2\right)\right] \\ \iff &\mathbb{E}\left[u\left(w - WTP_{1,NB2}\right)\right] = \mathbb{E}\left[u\left(w - L_1\right)\right] \end{aligned}$$

Similarly,

$$\begin{aligned} &\mathbb{E}\left[u\left(w - WTP_{1,B2} - I_2\right)\right] = \mathbb{E}\left[u\left(w - L_1 - I_2\right)\right] \\ \iff &\mathbb{E}\left[u\left(w - WTP_{1,B2}\right)\right]\mathbb{E}\left[u\left(-I_2\right)\right] = \mathbb{E}\left[u\left(w - L_1\right)\right]\mathbb{E}\left[u\left(-I_2\right)\right] \\ \iff &\mathbb{E}\left[u\left(w - WTP_{1,B2}\right)\right] = \mathbb{E}\left[u\left(w - L_1\right)\right] \end{aligned}$$

Together this implies that $WTP_{1,B2} = WTP_{1,NB2}$ as required.

■

An alternate set of assumptions that are sufficient for the WTP for insurance ocntract oen to not depend on whether contract 2 is bought or not is that both risks are small in the following sense.

A.3 Endogenous contracts

Until this point I have assumed that the insurance contracts under consideration are fixed, in particular that they do not change as the correlation structure changes, beyond being bundled or not. In this section I show that the qualitative conclusions of the prior sections hold in a simplified model where the contracts offered endogenously respond to the correlation structure. The model closely follows Crocker and Snow (2011). Here, however, I focus on the (locally) competitive equilibrium which will not allow cross-subsidization between types.

There is a unit mass of individuals who begin with wealth w. There are two states of the world, L(oss) and NL and two types, $t = H, L$ who have different probabilities of experiencing a loss: $p^H > p^L$. Conditional on a loss occuring, there are two possible (exhaustive and exclusive) versions of this loss, which I refer to as *perils.* For simplicity assume both perils lead to a loss of the same size l. Write θ^H and θ^L for the likelihood that peril 1 occurs, conditional on a loss, for the H and L types respectively. Peril 2 occurs with the complementary probability. Without loss of generality assume $\theta^H \geq \theta^L$.

An insurance contract specifies a premium to be paid α to be paid in all states of the world in exchange for indemnities ι_1, ι_2 to be paid in case a loss occurs due to peril 1 or 2 respectively. On net, an insurance contract promises consumption of $c_0 = W - \alpha$ in the no-loss state, and consumption of $c_1 = w - \alpha - l + \iota_1$ if the loss occurs due to peril 1, with c_2 defined similarly.

If an individual of type $t = H, L$ purchases an insurance contract that promises consumption vector $\boldsymbol{c} = (c_0, c_1, c_2)$ their expected utility is given by

$$V^i(\boldsymbol{c}) = (1 - p^t)u(c_0) + p^t\theta^t u(c_1) + p^t(1 - \theta^t)u(c_2).$$

The flow utility function $u(\cdot)$ is assumed twice continuously differentiable and weakly concave. When evaluating welfare, I do so according to, where α is the pareto weight:

$$W = \alpha V^H\left(\boldsymbol{c}^H\right) + (1 - \alpha)V^L\left(\boldsymbol{c}^L\right).$$

As in section 2 I am interested in two questions: when will bundling occur and when will it be welfare improving. To that end I define:

Definition 5. *If a contract promises $c_1 = c_2$ then say that contract is unbundled. If $c_1 \neq c_2$ then that contract is bundled.*

A bundled contract means there is some cross-subsidization across perils so that the contract only breaks even when both perils are insured. Conversely an unbundled contract doesn't differientiate between perils and hence breaks even no matter which of the perils occurs. In this sense it could be unbundled to be peril specific while remaining weakly profitable. This conceptually mirrors the choice to be made by an insurer in section 2.

Definition 6. *Define the degree of correlation between types as $\rho = \theta_L/\theta_H$.*

When $\rho = 1$ types are 'perfectly correlated' in the sense that conditional on a loss, it is equally likely to be from the same peril for each type. As ρ decreases to zero the peril causing the loss is more likely to be different for each type. Nevertheless, as ρ changes the expected loss doesn't change, nor does the overall probability of some loss occurring, simply which of the two perils.

The first proposition speaks to when policies will be bundled in equilibrium. This is essentially Theorem 1 of Crocker and Snow (2011).

Proposition 6. *The equilibrium will feature only unbundled contracts when $\rho = 1$. Whereas for $\rho < 1$ the low types will receive a bundled contract.*

When $\rho = 1$, that is, when there is 'perfect correlation' between the two types, there is no benefit to bundling the perils and cross-subsidizing between them. On the other hand, when $\rho < 1$, and one peril is more likely than the other to affect the low type, then a bundled contract can be more attractive to the low types. Specifically, a contract that slightly overpays against the peril comparatively more likely to afflict the low types, and underpays against the peril comparatively more likely to afflict the high types will impose a second order utility cost on the low types (as marginal utility will be slightly unequal in each of the peril states) but a first order gain as the IC constraint that is holding the low types to partial insurance will be loosened.

This logic makes the next proposition clear:

Proposition 7. *Welfare is minimized at $\rho = 1$ and monotonically increases as ρ falls to zero.*

As ρ decreases companies can offer a contract to the low types that redistribute from H's more likely peril to L's. Low types will only buy the new contract if their welfare improves, and H types are getting full insurance throughout. Hence as soon as the equilibrium contracts change due to a ρ change, welfare must increase, analogously to the logic explained by Crocker and Snow (2011). As ρ declines, the 'comparative advantage' enjoyed by L at one peril vs the other increases, and therefore does the screening benefits possible by over-paying in that peril and underpaying in the other.

Overall, the qualitative message from this section coheres with the prior sections. In a setting with adverse selection, when the less correlation there is between types in what type of loss will occur, fixing the overall probability of *some* loss occurring, the more easily can the types be screened apart and the incentive constraint holding the low types to partial insurance be relaxed.

In section 2, bundling was introduced when the types with WTP for one contract had relatively low WTP for the other. Similarly here, the degree to which the WTPs for the different types of losses differ encourages bundling to the benefit of the firms and the planner.

A.4 Conditional Correlation Order

Recall that the original (unconditional) correlation order relates two joint distributions with the same marginals. We write $X \precsim Y \iff F_X(w_1, w_2) \leq$

$F_Y(w_1, w_2)$ for all (w_1, w_2). In particular this is equivalent to:

$$X \precsim Y \iff F_X(w_1, w_2) \leq F_Y(w_1, w_2) \tag{A.1}$$

$$\iff P_X(W_1 \leq w_1)P_X(W_2 \leq w_2 \mid W_1 \leq w_1) \leq P_Y(W_1 \leq w_1)P_Y(W_2 \leq w_2 \mid W_1 \leq w_1) \tag{A.2}$$

$$\iff \frac{P_X(W_1 \leq w_1)P_X(W_2 \leq w_2 \mid W_1 \leq w_1)}{P_X(W_1 \leq w_1)} \leq \frac{P_Y(W_1 \leq w_1)P_Y(W_2 \leq w_2 \mid W_1 \leq w_1)}{P_Y(W_1 \leq w_1)} \tag{A.3}$$

$$\iff P_X(W_2 \leq w_2 \mid W_1 \leq w_1) \leq P_Y(W_2 \leq w_2 \mid W_1 \leq w_1) \tag{A.4}$$

and similarly

$$X \precsim Y \iff P_X(W_1 \leq w_1 \mid W_2 \leq w_2) \leq P_Y(W_1 \leq w_1 \mid W_2 \leq w_2).$$

This is a transparent version of the correlation order. X is less correlated than Y when knowing one variable is small makes it more likely under Y than X that the other is small. Similarly, if one variable was large then the other would be more likely to be large under Y than X.

This definition generalizes easily to conditional distributions.

Definition 7. *Suppose $X, Y \in \Gamma(F_1, F_2)$ and have CDFs F_X, F_Y respectively. I say that X is be less correlated than Y or that X precedes Y in the correlation order, conditional on A, written as $X \precsim_A Y$ when*

$$\begin{aligned} X \precsim_A Y \iff & P(X_1 \leq w_1 \mid X_2 \leq w_2, A) \leq P(Y_1 \leq w_1 \mid Y_2 \leq w_2, A), \quad \textit{for all } (w_1, w_2) \in \mathcal{W} \\ \wedge \quad & P(X_2 \leq w_2 \mid X_1 \leq w_1, A) \leq P(Y_2 \leq w_2 \mid Y_1 \leq w_1, A), \quad \textit{for all } (w_1, w_2) \in \mathcal{W} \end{aligned}$$

For example, we might wish to condition on $A = \{X_1 \geq \underline{w}\}$. There isn't always an exact relationship between the correlation of the non-truncated distributions X, Y and that of the truncated distributions. But typically the correlation order of the underlying distributions is retained under truncation. For example, (e.g. see Kotz (2000)):

Example 2. *If both X and Y are jointly normally distributed with identical marginals, $X \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho_X)$ and $Y \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho_Y)$ and we write $X_{\underline{w}}, Y_{\underline{w}}$ for the (singly) truncated distributions at w then we have*

$$\rho_X \leq \rho_Y \iff \rho_{X_{\underline{w}}} \leq \rho_{Y_{\underline{w}}}.$$

The assumptions in the main theory section were typically about truncated distributions. These are attractive as they correspond to the exactly the individuals that would show up in a dataset of the insured. But in the case when one has data on populations, perhaps because the insurance market doesn't exist, or because counter-factual insurance policies are being considered, assumptions on the global distribution of types might be preferable. In this section I present some results that are close equivalents of results in the main paper except with different assumptions used. Moreover, results that extend the theory in the main paper that require alternative or additional assumptions are studied.

An alternatively sufficient condition for many of the results that follow is that the change in total cost when correlation increases has the same sign as the change in average cost, conditional on being in some set S. Note that, fixing prices, if the joint distribution changes there are two effects: 1) those that buy products $1, 2$ or B have higher or lower costs and 2) the mass of people who buy a product changes. The second effect is of sole importance in case of no adverse or advantageous selection, such as the IO literature. I focus on the first effect. The assumption is essentially an assumption that the second effect is small. I refer to the first effect as the cost effect and the second as the quantity effect.

Note that the average cost of all those who buy a product $i = 1, 2, B$ can be written as

$$AC = \frac{\mathbb{E}\left[\mathbb{1}\left(w \in W_i\right) \times \phi_i(w)\right]}{\mathbb{E}\left[\mathbb{1}\left(w \in W_i\right)\right]}.$$

When the distribution changes the change in the value of the numerator is effect (1) and the denominator is effect (2). The following assumption therefore assumes (2) is small and will be an alternatively sufficient condition for the propositions of the main theory section to go through.

Assumption 2. *Given two distributions $X, Y \in \Gamma(F_1, F_2)$, we say that the quantity effect is small if*

$$\mathbb{E}_X\left[\mathbb{1}\left(w \in W_i\right)\right] \approx \mathbb{E}_Y\left[\mathbb{1}\left(w \in W_i\right)\right].$$

Given these alternate definitions we can restate the main results of section 2 in terms of correlation structures of the entire distributions with the additional assumptions. In particular, all assumptions about $Y \succsim^{\overline{\boldsymbol{p}}} X$ or $Y \succsim^{\backslash p_B} X$ are changed to simpler assumptions on $Y \succsim X$ and the additional assumption 2.

First the results about adverse selection, in which $\phi' > 0$

Proposition 8. *(Analogue of Proposition 1). Suppose $\phi' > 0$. Suppose $X, Y, Z^{\perp} \in \Gamma(F_1, F_2)$ with $X \precsim Y$. Denote the profit earned per person on* *a bundle contract offered at price $p_B = \bar{p}_1 + \bar{p} - \epsilon$ by π^ϵ. We have the following:*

1. $\pi^\epsilon_{Z^\perp} = 0$

2. *Suppose Y, X, Z satisfy assumption 2. If $Y \succsim X \succsim Z^\perp$ then $\pi^\epsilon_Y \leq \pi^\epsilon_X \leq \pi^\epsilon_{Z^\perp} = 0$, and conversely if $Z^\perp \succsim X \succsim Y$ then $\pi^\epsilon_Y \geq \pi^\epsilon_X \geq \pi^\epsilon_{Z^\perp} = 0$*

Proposition 9. *(Analogue of Proposition 3). Suppose $\phi' > 0$. Suppose $X, Y \in \Gamma(F_1, F_2)$ and satisfy assumption 2 with $Y \succsim X$. The following comparative statics hold:*

- *The average cost curve is everywhere lower under less correlated distributions: $AC_X(p) \leq AC_Y(p)$ for all p.*

- *The equilibrium bundle price increases in correlation: $p^X_B \leq p^Y_B$.*

- *Assuming joint normality, for 'large' markets, i.e. if $q^Y_B \geq \underline{q}$ for some $\underline{q}$ then the equilibrium quantity insured increases under the less correlated X: $q^X_B \geq q^Y_B$.*

And the analogues for the case of advantageous selection, $\phi' < 0$.

Proposition 10. *. Suppose $\phi' < 0$. Suppose $X, Y, Z^\perp \in \Gamma(F_1, F_2)$ with $X \precsim Y$. Denote the profit earned per person on a bundle contract offered at price $p_B = \bar{p}_1 + \bar{p} - \epsilon$ by π^ϵ. We have the following:*

1. $\pi^\epsilon_{Z^\perp} = 0$

2. *Suppose Y, X, Z satisfy assumption 2. If $Y \succsim X \succsim Z^\perp$ then $\pi^\epsilon_Y \geq \pi^\epsilon_X \geq \pi^\epsilon_{Z^\perp} = 0$, and conversely if $Z^\perp \succsim X \succsim Y$ then $\pi^\epsilon_Y \leq \pi^\epsilon_X \leq \pi^\epsilon_{Z^\perp} = 0$*

Proposition 11. *Suppose $\phi' < 0$. Suppose $X, Y \in \Gamma(F_1, F_2)$ and satisfy assumption 2 with $Y \succsim X$. The following comparative statics hold:*

- *The average cost curve is everywhere higher under less correlated distributions: $AC_X(p) \geq AC_Y(p)$ for all p.*

- *The equilibrium bundle price decreases in correlation: $p^X_B \leq p^Y_B$.*

- *Assuming joint normality, for 'small' markets, i.e. if $q^Y_B \leq \underline{q}$ for some $\underline{q}$ then the equilibrium quantity insured decreases under the less correlated X: $q^X_B \leq q^Y_B$.*

B Empirical Appendix

B.1 Ruling out 'standard' cost efficiencies

In this section I provide more detail on the other insurance products considered in section 3.5:life insurance, Long-term care insurance, hybrid long-term care and life insurance, and life annuities. In all cases only individual (i.e. non-group) products are included.

First I measure the average discount offered in each product when spouses apply and are accepted together. The data for life insurance comes from Compulife , software which collects historical quotes for over 100 life insurance issuers. The data for hybrid life-long-term care policies come from Combo-Compare (StrateCision (2022a)) which quotes from 6 different hybrid products. Annuity prices come from the Annuity Shopper (ImmediateAnnuities.com (ImmediateAnnuities.com)). Disability (non-group) discounts come directly from underwriting guides available publicly.

The general idea is straightforward for long-term care, life and hybrid products. I simply compare the price of two identical policies when bought as a couple versus when bought individually. For example, two 10 year term life insurance policies from Ameritas Life Insurance Corp, each with a face value of $1 million for a 40 year old man and a 40 year old women in the highest rating category cost $370 and $330 per month when purchased separately, but $660 when purchased together, a discount of $40 or 6%. In the case of disability insurance, I was unable to obtain direct prices, but did study the underwriting guides for 12 different products. None offered couples discounts.

The only exception are annuity prices. To measure a discount (or anti-discount) offered to couples I construct a payoff equivalent version of a joint life annuity out of two single life annuities, and compare the prices. In particular, for each company at each time period, I take the payout offered to a joint and survivor annuity bought at a cost of $100,000 that has payments that reduce to 50% after one spouse dies. For example, for a joint annuity for a couple consisting of a 65 year old man and 60 year old woman, Nationwide promised payouts of $428 per month reducing to $214 after one spouse days. This could alternatively be constructed by spending $44,958 on a 65 year old male single life annuity from Nationwide (the quoted payout rate is $476 per month per $100,000) and $53,768 on a 60 year old female single life annuity (the quoted rate is $398 per month per $100,000). This will generate $428 in income per month while both annuitants are alive, and $214 after one dies. But this synthetic version of the joint annuity costs only $98,727. Hence the joint product is sold at an *anti-discount* of $1,273, or 1.275%. That final number,

Table 3: Single vs Married Voluntary Lapse Rates

	Without Controls	With Controls
Single	0.038***	0.021***
	(0.003)	(0.003)
Num.Obs.	682 104	682 104
R2	0.183	0.445

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

-1.275%, is what I record for each company x year x joint policy combination.

The resulting average and range of discounts for each product are what is reported in table 6 in the main body of the paper.

B.2 Ruling out differential lapsation as the driver of discounts

The data come from an SOA experience study. The dependent variable is lapsation of the contract due to any factor other than mortality. The results are similar even if mortality is included.

The equation estimated is

$$lapse_i = \alpha_0 + \alpha_1 LivesExposed_i + \beta Married_i + \gamma Controls_i + \epsilon_i. \tag{B.1}$$

The controls included are as exhaustive as possible. Specifically: Observation year, gender, policy year, attained age (grouped), issued age (group), premium payment frequency, rate increase flag, inflation coverage (grouped), elimination period length (grouped). The results are in 3 below.

We see strong evidence that voluntary lapses are lower for married couples. Voluntary lapsation is profitable for the insurer, since premia are front-loaded relative to claims. This demonstrates that it is *not* a favorable lapsation pattern that leads to a spousal discount. This gives more weight to the evidence presented in the main body of the paper that the discount is in fact due to favorable cost correlations.

B.3 The Predictive Content of WTP Relative to Information Collected by Insurers

Figure 11 below is an analogue of Figure 5, but where the predictions of LTC risk use only the smaller set of information available to insurers.[18] For example, subjective questions in the HRS, such as perceptions of physical and mental health, which could never be elicited in a verifiable, and hence incentive compatible, way by insurers, go into the prediction of Figure 5 but not 11.

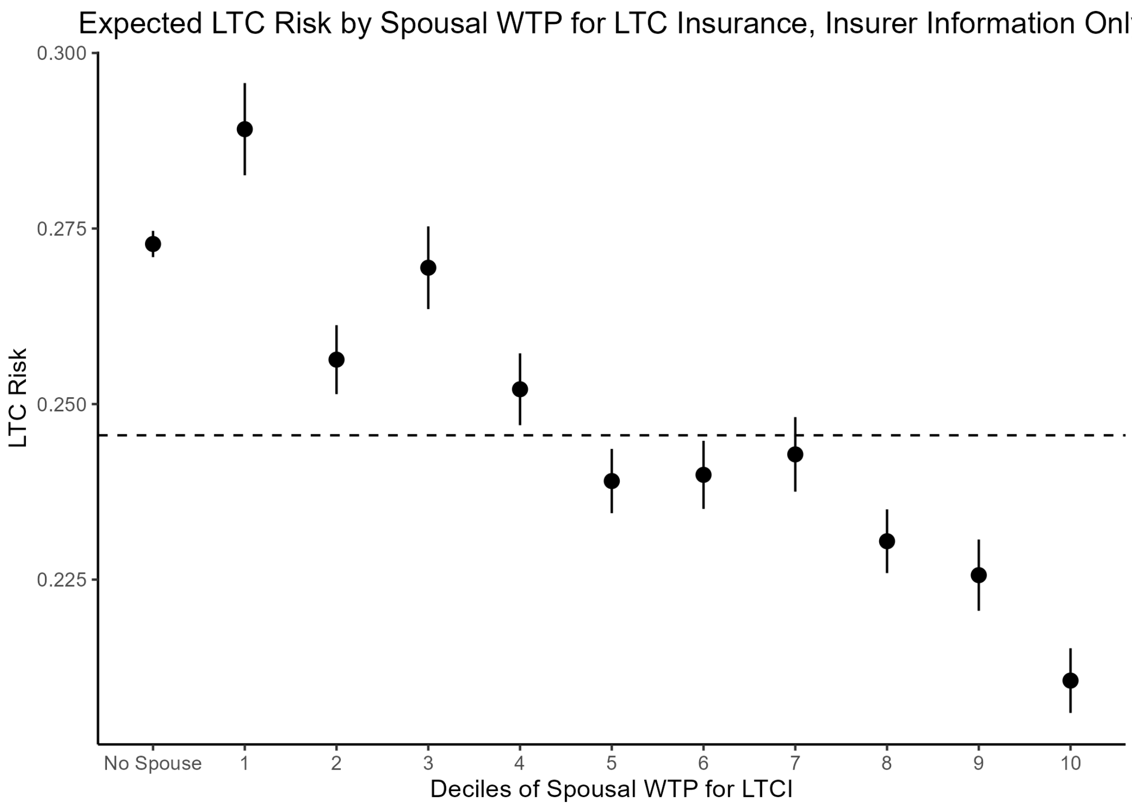

Figure 11: Binscatter of expected LTC risk that is predictable by insurers, by deciles of spousal WTP. The horizontal dashed line is the average risk amongst all those who have a spouse.

The slope of observable risk with respect to spousal WTP (Figure 11)

[18] These are: age, Activities of Daily Living Assistance (ADLA), current and past memory status, Instrumental Activities of Daily Living Assistance (IADLA), high blood pressure, diabetes, cancer, lung disease, heart disease, stroke history, arthritis, Body Mass Index (BMI), sex, nursing home use, and home health care use. See Mutual of Omaha Insurance Company (2023)

is much shallower than that of total (observable and unobservable) risk with respect to spousal WTP (Figure 5). Whereas total risk falls from 0.3 to 0.05 as spousal WTP goes from the first to tenth decile, observable risk falls only from 0.28 to 0.22. This is why, even if the insurer tried to predict which individuals had high WTP spouses using information observable to them, they could not that close to the information revealed by letting individuals self-select using their private information.

B.4 Intensive Margin of LTC Utilization

Figure 12 is an analogue of figure 5 where the outcome is the expected length of stay in a nursing home, conditional on entry, instead of the probability of entry. A similar pattern prevails, although less pronounced. Being married predicts lower intensive margin usage, and within the married having a spouse with higher WTP predicts even lower intensive margin usage. This further compounds the extensive margin effect.

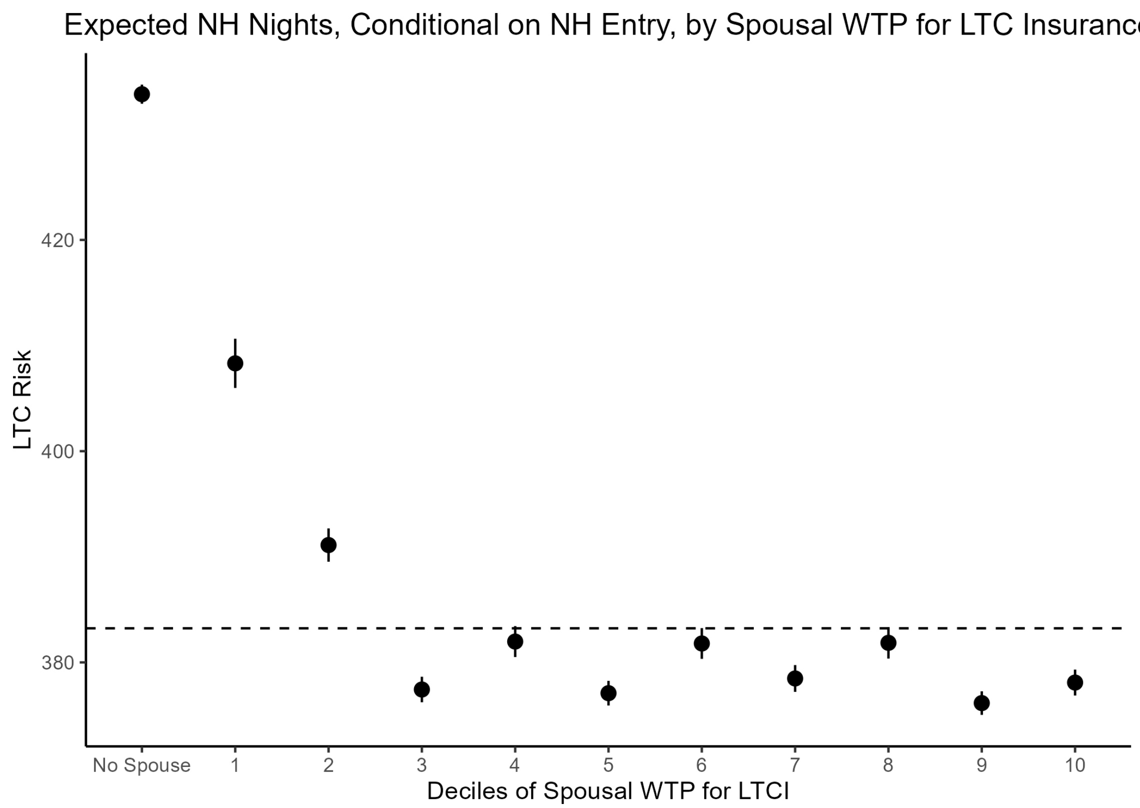

Figure 12: Binscatter of expected number of nights in a nursing home, conditional on entry, by deciles of spousal WTP. The horizontal dashed line is the average risk amongst all those who have a spouse.

B.5 Further Evidence from long-term care Policyholders

A possible drawback of the HRS data is that I can measure only underlying risk, not actual insurance purchases. In this section I provide evidence from a large experience study of long-term care insurance policyholders consistent with the HRS evidence. A downside of policyholder data is that this may be a sample already induced to buy by couples discounts. Given my data I cannot rule this out. Nevertheless, the HRS and policyholder evidence are complementary and together more convincing.

The data come from the Society of Actuaries (SOA) SOA (2016) who periodically run 'experience studies'. This experience study collected data from long-term care insurance policyholders from 18 different insurers from

2000-2016. The outcomes of interest were claim rates for different types of long-term care benefits. The data contain approximately 60 million contract years of exposure and 600,000 claims. The SOA amalgamated contract level data such that each observation contains multiple years of exposure for the individuals with the same covaraites. This is controlled for in my analysis, but explains how 60 million contract years of exposure becomes an N of just over 1 million.

I run 8 different regressions. All combinations of four different outcome variables (total claims, nursing home claims (NH), assisted living facility claims (ALF), home health care (HHC) and two different specifications, with and without controls, are included in the table. Note that I don't observe price, although I do observe the above factors that enter into pricing.

The equation estimated is

$$outcome_i = \alpha_0 + \alpha_1 LivesExposed_i + \beta Single_i + \gamma Controls_i + \epsilon_i. \tag{B.2}$$

The outcome is either all claims, NH claims, ALF claims or HHC claims. The controls, when included, contain the following exhaustive set of variables that are priced upon: , Coverage Type , Issue Age, Current Age, Issue Year , Premium Class, Underwriting Type, Coverage Type, Inflation Rider, Rate Increase Flag, Restoration of Benefits Flag, NH Daily Benefit, ALF Daily Benefit, HHC Daily Benefit, NH Benefit Period, ALF Benefit Period , HHC Benefit Period, NH Elimination Period, ALF Elimination Period, HHC Elimination Period.

Table 4 below illustrates the coefficient on being single (as opposed to married) without controls (column 1), with controls (column 2) and column 3 is the overall claim rate per contract-year exposed to allow for interpretation of the effect of being single.

We see that those who are single have substantially higher claim rates than those who are married. Even with controls,and relative to the overall claim rate, the probability of making a claim is approximately 69%, 60%, 66% and 100% higher for all claims, NH claims, ALF claims and HHC claims respectively.

This is evidence that being married strongly predicts the probability of utilizing long-term care insurance. This is consistent with the substantial discounts offered to married couples.

	Coefficient on Single (β)		
Outcome	Without Controls	With Controls	Overall Claim Rate
All claims	0.17*** (0.004)	0.09*** (0.004)	0.13*** (0.002)
NH claims	0.06*** (0.001)	0.03*** (0.001)	0.05*** (0.001)
ALF claims	0.04*** (0.001)	0.02*** (0.001)	0.03*** (0.001)
HHC claims	0.07*** (0.003)	0.04*** (0.003)	0.04*** (0.001)
N	1,507,275	1,507,275	1,507,275

Table 4: Results from estimating equations (B.2) with different outcomes, with and without controls. The rightmost column is the sample average claims per contract-year. *** = significant at the 1 % level.

B.6 Details of LTC Risk Prediction

The following variables are included in the prediction of LTC risk.

1. The individual's age.
2. Marital status.
3. Mother's survival status.
4. Father's survival status.
5. Mother's age.
6. Father's age.
7. Self-assessed health status.
8. History of hospitalization.
9. Residency in a nursing home.
10. Number of doctor visits.
11. Receipt of home care.

12. Out-of-pocket medical expenses.
13. Assistance with personal care activities.
14. Assistance with household tasks.
15. Assistance with mobility.
16. Assistance with activities requiring muscle strength.
17. Assistance with major physical activities.
18. Assistance with delicate or precise activities.
19. Feelings of depression.
20. Past happiness.
21. Alcohol consumption frequency.
22. Smoking history.
23. Current smoking status.
24. High blood pressure.
25. Diabetes.
26. History of cancer.
27. Lung disease.
28. Heart disease.
29. Stroke history.
30. Arthritis.
31. Current memory status.
32. Past memory status.
33. Government health insurance.
34. Number of private health insurance plans.
35. Life insurance coverage.

36. Number of children.
37. Retirement status.
38. Past high blood pressure.
39. Past diabetes.
40. Past cancer.
41. Past lung disease.
42. Past heart disease.
43. Past stroke.
44. Past arthritis.
45. Body Mass Index.
46. Gender.
47. Asset prediction less than $10,000.
48. Asset prediction less than $100,000.
49. Income group classification.
50. Participation in a training program.
51. Occupational Health Facility usage in the past six years.
52. Data collection wave.
53. Interaction of sex with: diabetes, cancer, lung disease, heart disease, stroke, arthritis, ADLA, IADLA, BMI and being married.

C Proofs of Main Results

C.1 Proof of Proposition 1

Proof. Suppose that the equilibrium prices in markets 1 and when there is no bundling are given by $\bar{p}_1$ and $\bar{p}_2$ respectively. Then a bundle is introduced at price $p_B = \bar{p}_1 + \bar{p}_2 - \epsilon$. For small enough *epsilon* > 0 individuals will buy the bundle at price p_B are those who separately bought products 1 and 2 before. Hence the set that purchase the bundle, which I denote $\mathcal{W}_B$ is given by $\mathcal{B} = \{w \in \mathcal{W} : w_1 \geq \bar{p}_1 \wedge w_2 \geq \bar{p}_2\}$. As before, the sets fo types that only buy 1 or 2 after the bundle is offered are denoted by $\mathcal{W}_1$ and $\mathcal{W}_2$.

To the first point, note simply that by the definition of independence, and that the market for product 1 was initially in equilibrium with $E[\phi(w_1) \mid \mathcal{W}_1] = \bar{p}_1$, we have

$$\begin{aligned}
\mathbb{E}[\phi_1(w_1) \mid w_1 \geq \bar{p}_1 \wedge w_2 \geq \bar{p}_2] &= \mathbb{E}[\phi_1(w_1) \mid w_1 \geq \bar{p}_1 \wedge w_2 \leq \bar{p}_2] \\
&= \mathbb{E}[\phi_1(w_1) \mid w_1 \geq \bar{p}_1] \\
&= \bar{p}_1
\end{aligned}$$

Similarly

$$\mathbb{E}[\phi_2(w_2) \mid w_1 \geq \bar{p}_1 \wedge w_2 \geq \bar{p}_2] = \bar{p}_2.$$

Hence the total cost of selling the bundle is $\bar{p}_1 + \bar{p}_2$, and so this breaks even for small ϵ, i.e. $\pi^{\epsilon}_{Z^{\perp}} \to 0$ as $\epsilon \to 0$.

To prove the rest, suppose we have two distributions with $X, Y \in \Gamma(F_1, F_2)$ with Y more correlated than X in the conditional sense ($Y \succsim_{p_1} X, Y \succsim_{p_2} X$).

We then have, first for adverse selection:

$$\begin{aligned}
\mathbb{E}_X[\phi_1(w_1) \mid w_1 \geq \bar{p}_1 \wedge w_2 \geq \bar{p}_2] &= \int_{\bar{p}_1}^{\overline{w}_1} \phi(w_1) P_X(W_1 = w_1 \mid W_2 \geq \bar{p}_2 \wedge W_1 \geq \bar{p}_1)\, dw_1 \\
&= \left[\phi(w_1)\left(P_X(W_1 \leq w_1 \mid W_2 \geq \bar{p}_2 \wedge W_1 \geq \bar{p}_1) - 1\right)\right]_{w_1=\bar{p}_1}^{w_1=\overline{w}_1} \\
&+ \int_{\bar{p}_1}^{\overline{w}_1} \phi'(w_1)\left(1 - P_X(W_1 \leq w_1 \mid W_2 \geq \bar{p}_2 \wedge W_1 \geq \bar{p}_1)\right) dw_1 \qquad \text{(inte} \\
&= \phi(\bar{p}_1) + \int_{\bar{p}_1}^{\overline{w}_1} \phi'(w_1) P_X(W_1 > w_1 \mid W_2 \geq \bar{p}_2 \wedge W_1 \geq \bar{p}_1)\, dw_1 \\
&\leq \phi(\bar{p}_1) + \int_{\bar{p}_1}^{\overline{w}_1} \phi'(w_1) P_Y(W_1 > w_1 \mid W_2 \geq \bar{p}_2 \wedge W_1 \geq \bar{p}_1)\, dw_1 \\
&= \mathbb{E}_Y[\phi_1(w_1) \mid w_1 \geq \bar{p}_1 \wedge w_2 \geq \bar{p}_2].
\end{aligned}$$

where the inequality follows from $\phi_1'(w_1) > 0$ (adverse selection) and equation (C.3). Under advantageous selection the inequality is reversed. This

Similarly we conclude for risk 2 that

$$\mathbb{E}_X[\phi_2(w_2) \mid w_1 \geq \bar{p}_1 \wedge w_2 \leq \bar{p}_2] \leq \mathbb{E}_Y[\phi_2(w_2) \mid w_1 \geq \bar{p}_1 \wedge w_2 \leq \bar{p}_2].$$

Now if Y were independent, making X negatively correlated, this would show that bundling is profitable under adverse selection but not under advantageous. Conversely, if X were independent, making Y positively correlated, we see that bundling is profitable under adverse selection but not under advantageous selection. This immediately shows that the equilibrium must feature bundling in the adverse & negatively correlated or advantageous & positively correlated case, since there is a profitable deviation away from the separate market equilibrium.

It also shows that an infinitesimal bundling discount is not profitable under the adverse & positively correlated or advantageous & negatively correlated case. By our assumption of the monotonicity of the profit function, at any price below $p_B = p_1^{NB} + p_2^{NB} - \epsilon$ an even greater loss is made, and at any price above zero profits are made, since no one buys the bundle. This concludes the proof.

■

C.2 Proof of Proposition 2

Proof. First, we prove the first part of the proposition. Fix the prices of product 2 and the bundle price $p_2,^{MB}, p_B^{MB}$ to be the mixed bundling equilibrium prices, and suppose the separate price in market 1 is the same as the no-bundling equilibrium market 1 price: $p_1^{MB} = p_1^{NB}$. The set of people who buy the separate product 1 are those with $w_1 > p_1^{MB}$ and $w_1 - p_1^{MB} > w_1 + w_2 - p_B \iff w_2 < p_B - p_1^{MB}$.

Suppose we have adverse selection. The case of advantageous selection is symmetrical. Re-running the proof of proposition 1, we have that, for X less correlated than Y in the same sense as in that proposition:

$$\mathbb{E}_X\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \leq p_B - p_1^{MB}\right] \leq \mathbb{E}_Y\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \leq p_B - p_1^{MB}\right]$$

and

$$\mathbb{E}_X\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \geq p_B - p_1^{MB}\right] \geq \mathbb{E}_Y\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \geq p_B - p_1^{MB}\right].$$

If Y is independent, we have that we have that

$$\mathbb{E}_Y\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \geq p_B - p_1^{MB}\right] = \mathbb{E}_Y\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB}\right] \quad \text{(C.1)}$$

$$= \mathbb{E}_Y\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \leq p_B - p_1^{MB}\right] \quad \text{(C.2)}$$

from which it follows that

$$\mathbb{E}_X\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \geq p_B - p_1^{MB}\right] \geq \mathbb{E}_X\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \leq p_B - p_1^{MB}\right].$$

By the definition of the separate market clearing price, we have that

$$\mathbb{E}_X\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB}\right] = p_1^{MB}.$$

Hence, by the law of total expectation, we have that

$$\mathbb{E}_X\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \leq p_B - p_1^{MB}\right] < p_1^{MB}.$$

That is, when X is negatively correlated, in a mixed bundling situation, if the separate price for product 1 was equal to the no-bundling equilibrium price, product 1 would make a loss. By assumption 1, it follows that the equilibrium price must be higher. This argument works identically for product 2, and symmetrically for advantageous selection with positive correlation.

Next, to prove the second part of the proposition: having established that the separate market prices must be higher, we show that the bundling price must be lower than the sum of the *no-bundling* equilibrium prices. To see this, suppose that the bundling price was set equal to the sum of the no-bundling equilibrium prices: $p_B = p_1^{NB} + p_2^{NB}$.

Consider who buys product 1 thorough the bundle relative to those who bought product 1 in the no-bundling equilibrium. In the notation of figure 7, except with $p_B = p_1^{NB} + p_2^{NB}$ define three groups. Group α consists of 1 and 2a from figure 7, group β consists of 3a,4a and 5a, while group γ consists of 2b. Under no-bundling equilibrium, groups α and β buy product 1, whereas under mixed bundling, groups α and γ buy product 1 in the bundle.

In terms of relative costs of these groups, note that α is cheaper than β (this is why there was bundling in the first place, and the proof above can be adapted to show this), and α is more expensive than γ *under adverse selection*, because the WTP of β are higher than γ. Hence the average cost of risk 1 is ordered: group $\gamma <$ group $\alpha <$ group β under the bundle is the weighted average of α and β. This means the average cost of those that buy under the bundle - groups α and γ is lower than the average cost of those who buy in the no-bundling equilibrium = groups α and β, no matter the relative weights of these groups. But groups α and β broke even according to p_1 and so α and γ are profitable to sell to. Symmetrically for risk 2. Hence selling the bundle at price $p_B = p_1^{NB} + p_2^{NB}$ gives a positive profit. We conclude that $p_B < p_1^{NB} + p_2^{NB}$ as required.

■

C.3 Proof of Proposition 3

Proof. Consider two correlation structures X and Y with $Y \succsim X$. **First, for adverse selection.**

This assumption means, in words, that knowing W_2 is large (larger than $p_B - w$ in particular) says more about W_1 being large under Y than under X. Formally,

$$P_X\left(W_1 > w_1 \mid W_2 \geq p_B - w_1\right) \leq P_Y\left(W_1 > w_1 \mid W_2 \geq p_B - w_1\right) \tag{C.3}$$

$$\tag{C.4}$$

For brevity I write this set as

$$\mathcal{W}_{\setminus p_B} = \{w \in \mathcal{W} : W_B \geq p_B\} \tag{C.5}$$

We then have

$$\begin{aligned}
\mathbb{E}_X\left[\phi_1(w_1) \mid \mathcal{W}_{\setminus p_B}\right] &= \int_0^{\overline{w}_1} \phi(w_1) P_X\left(W_1 = w_1 \mid W_2 \geq p_B - w_1\right) dw_1 \\
&= \left[\phi(w_1)\left(P_X\left(W_1 \leq w_1 \mid W_2 \geq p_B - w_1\right) - 1\right)\right]_{w_1=\bar{p}_1}^{w_1=\overline{w}_1} \\
&+ \int_{\bar{p}_1}^{\overline{w}_1} \phi'(w_1)\left(1 - P_X\left(W_1 \leq w_1 \mid W_2 \geq p_B - w_1\right)\right) dw_1 \quad \text{(integrating by parts)} \\
&= \phi(\bar{p}_1) + \int_{\bar{p}_1}^{\overline{w}_1} \phi'(w_1) P_X\left(W_1 > w_1 \mid W_2 \geq p_B - w_1\right) dw_1 \\
&\leq \phi(\bar{p}_1) + \int_{\bar{p}_1}^{\overline{w}_1} \phi'(w_1) P_Y\left(W_1 > w_1 \mid W_2 \geq p_B - w_1\right) dw_1 \\
&= \mathbb{E}_Y\left[\phi_1(w_1) \mid W_2 \geq p_B - w_1\right].
\end{aligned}$$

Similarly we conclude for $\phi_2(w_2)$ and hence for $\phi_B(w_B)$.

This establishes that for a fixed price, the average cost of those who buy increases under Y relative to X when Y is more correlated than X in this sense.

This immediately implies that the equilibrium price is lower under X than under Y, since

$$p_B^Y = \mathbb{E}_Y \left[\phi_1(w_1) \mid \mathcal{W}_{\backslash p_B^Y}\right] \geq \mathbb{E}_X \left[\phi_1(w_1) \mid \mathcal{W}_{\backslash p_B^Y}\right]$$

and hence if firms sold the bundle to a population with true distribution X at price p_B^Y they would make a profit, and by assumption 1 the true equilbirium price under X must be lower.

Finally, for the third part, by proposition 13, the demand curve for the bundle under X is above that for Y for all $q \geq \bar{q}$ where $\bar{q}$ is defined in the proof of proposition 13. That means, for any quantity insured beyond $\bar{q}$, the gap between demand and AC is higher under X than Y. In particular, at the equilibrium quantity q_B^Y the gap is zero under Y and positive under X, hence $q_B^X > q_B^Y$ as required.

Under advantageous selection, the proof works almost identically. That $\phi_2'(w_2) < 0$ reverses the first inequality and shows that costs are lower under more correlated Y than less correlated X. The second and third parts of the proof are analogous.

■

C.4 Proof of Proposition 4

Proof. For ease, we label various groups according to the to the following figure

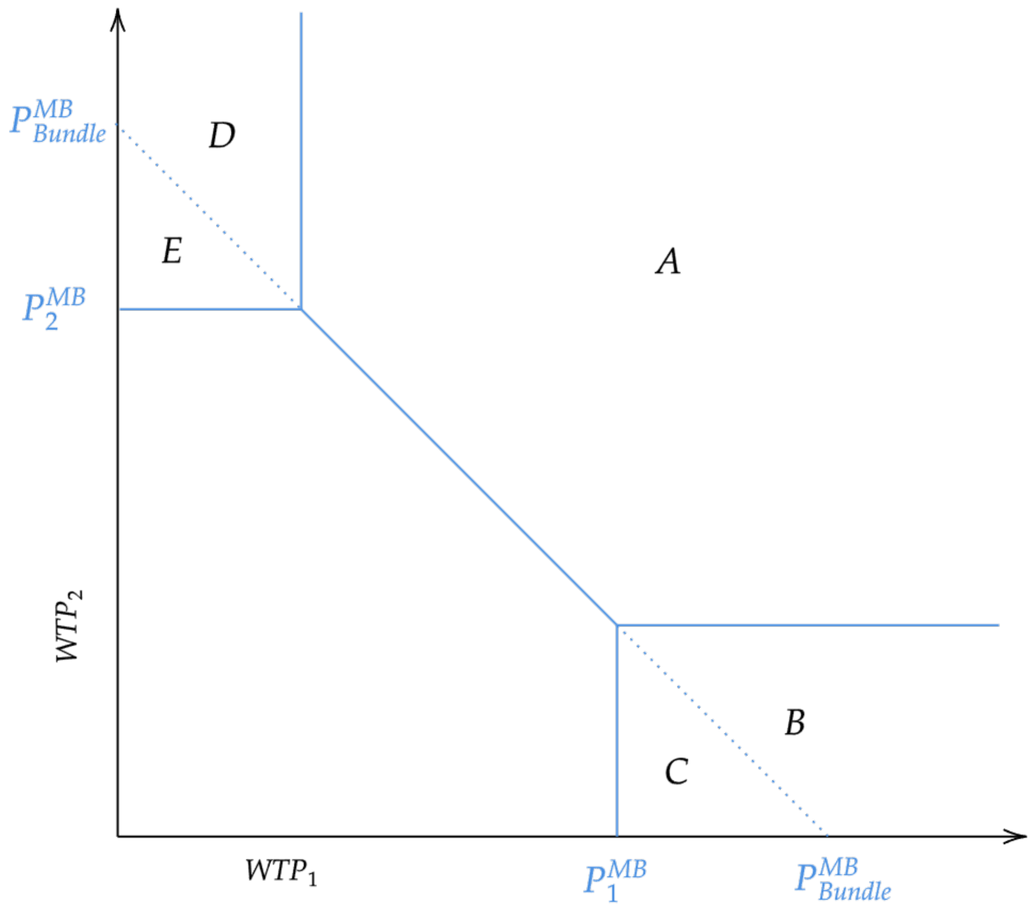

Figure 13: Different groupings under mixed and forced bundlings.

Write α_A for the mass of individuals in group A, AC_A for their average cost and so on. α_{BC} means for combination of groups B and C, etc. Under the mixed bundling regime, zero profit conditions ensure that

$$0 = \alpha_A p_B - \alpha_A AC^1_A - \alpha_A AC^2_A \tag{C.6}$$
$$0 = \alpha_{BC} p_1 - \alpha_{BC} AC^1_{BC} \tag{C.7}$$
$$0 = \alpha_{DE} p_2 - \alpha_{DE} AC^2_{DE}. \tag{C.8}$$

Now, if the government forces bundling, the bundle will be sold to groups A, B and D. Note in particular that since those in group C have higher WTP for product 1 than group D, under advantageous selection, they have lower cost.

By assumption, they have approximately the same mass. It follows that the average cost of selling product 1 to groups B and C is higher than to groups B and D: $AC_{BC} = AC_{BD}$ as well as the masses being equal $\alpha_{BC} = \alpha_{BD}$. Analagously for selling product 2 to DE vs BD.

Next, we compute the profit of selling the bundle, under forced bundling, at the same bundle price from mixed-bundling:

$$\pi = \alpha_A p_B - \alpha_A AC_A^1 - \alpha_A AC_A^2 + \alpha_{BD} p_B - \alpha_{BD} AC_{BD}^1 - \alpha_{BD} AC_{BD}^2 \tag{C.9}$$

$$= \alpha_{BD} p_B - \alpha_{BD} AC_{BD}^1 - \alpha_{BD} AC_{BD}^2 \qquad \text{by (C.6)} \tag{C.10}$$

$$< \alpha_{BD}(p_1 + p_2) - \alpha_{BD} AC_{BD}^1 - \alpha_{BD} AC_{BD}^2 \qquad \text{(since } p_B < p_1 + p_2\text{)} \tag{C.11}$$

$$< \alpha_{BC} p_1 + \alpha_{DE} p_2 - \alpha_{BC} AC_{BC}^1 - \alpha_{DE} AC_{DE}^2 \qquad \text{by the arguments above} \tag{C.12}$$

$$= 0 \qquad \text{by (C.7) \& (C.8).} \tag{C.13}$$

Hence, selling the bundle, under forced bundling, at the same price as the bundle sold for under mixed bundling leads to a loss. Then, by assumption 1, the equilibrium price is even higher. This establishes the proposition.

■

D Proofs of Supplementary Results

D.1 Preliminary Results

Definition 8. *We say that X precedes Y in the stop-loss order, written $X \precsim_{SL} Y$, iff*

$$\mathbb{E}_X\left[\max\{X - d, 0\}\right] \leq \mathbb{E}_Y\left[\max\{Y - d, 0\}\right] \quad \textit{for all } d \in \mathbb{R}.$$

The quantity $E(\max\{X - d, 0\})$ is known as the stop-loss premium. For example, it is the expected loss an insurer faces when they insure risk X with a deductible d.

The main proposition to be used comes from Denuit et al. (2006) or a variant from Cambanis et al. (1976).

Proposition 12. *(Denuit et al. (2006)) Suppose $X, Y \in \Gamma(F_1, F_2)$. If $X \precsim Y$ then $\psi(X) \precsim_{SL} \psi(Y)$ for any non-decreasing super-modular function ψ.*

Alternatively, per Cambanis et al. (1976), if ψ is supermodular and right continous, the stop-loss ordering holds, and if ψ is submodular and right continuous the stop-loss ordering is reversed.

This says that when X is less correlated than Y, the stop-loss of any supermodular function of the margins is higher under Y than X for any stop-loss premium. In particular, $w_1 + w_2 = w_B$ is super-modular, and by assumption so is $\phi_B(w_1 + w_2)$ and so the expectation of the excess of these functions relative to a deductible d is always higher under Y than X.

Next, I show that for any joint distribution satisfying joint normality or of FGM form, the WTP for the bundle W_B 'rotates' as the correlation changes.

Proposition 13. *Suppose X, Y are jointly normal or of FGM form (defined below). Then if $X \succsim Y$ the demand curve for the bundle under X, relative to the demand curve under Y satisfies:*

- $WTP_X(q) \leq WTP_Y(q)$ *for* $q \in [0, \underline{q}]$ *and then*
- $WTP_X(q) \geq WTP_Y(q)$ *for* $q \in [\underline{q}, 1]$.

Proof. **Joint Normality**

In the case of joint normality with identical margins, the CDF of the convolution $W_1 + W_2$ is normally distributed with mean $\mu_X + \mu_Y$ and standard deviation $\sigma_{W_B} = \sqrt{\sigma^2_{W_1} + \sigma^2_{W_2} + \rho\sigma_{W_1}\sigma_{W_2}}$. That the CDF's of W_B (and hence the WTP for the bundle) are rotated as described is a standard fact about two normal distributions with the same mean and differing variances.

FGM form

Writing $u(w_1) = F_1(w_1), v = F_2(w_2)$ as the uniformly distributed CDFs of marginals or a joint distribution of FGM form,

$$F_X(w_1, w_2) = u \cdot v \cdot [1 + \rho\,(1 - u)\,(1 - v)]\,.$$

Differentiating we get that the density is:

$$f_X(w_1, w_2) = 1 + \rho(1 - 2w_1)(1 - 2w_2).$$

Hence the CDF of the convolution $W_B = W_1 + W_2$ can be written as

$$\begin{aligned} Prob(W_B < s) &= \int_0^s \int_0^{s - w_2} f_X(w_1, w_2)\, dw_1\, dw_2 \\ &= \frac{1}{6} s^2 (\rho(s-3)(s-1) + 3) \end{aligned}$$

We are interested in how this CDF changes with the correlation parameter ρ. Hence differentiating we have

$$\frac{\partial Prob(W_B < s)}{\partial \rho} = \frac{1}{6}(s-3)(s-1)s^2.$$

The derivative is weakly positive on $[0, 1]$, and becomes weakly negative on $[1, 2]$. Hence for all $s \geq 1$ the CDF of the sum gets lower under more correlation, while for $s \leq 1$ the CDF of the sum gets higher under more correlation. This is what we needed to show.

■

D.2 Proof of Proposition 7

Before the proof, I repeat the key equilibrium characterization result from Crocker and Snow (2011).

Proposition 14. *(Crocker and Snow (2011)) In any constrained efficient allocation (in particular in the equilibrium with no cross-subsidization between types):*

- *The high types receive full insurance*
- *The low types receive partial insurance, limited by the high type's IC constraint*
- *The low types IC constraint is slack.*

Now we can prove proposition 7.

Proof. From proposition 14 the high type will get full insurance ($\boldsymbol{c}^*$)) and the low type will get the contract that solves:

$$\max V^L(\boldsymbol{c}^L) \tag{D.1}$$

$$\text{subject to} \tag{D.2}$$

$$IC_H : V^H(\boldsymbol{c}^*) \geq V^L(\boldsymbol{c}^L) \tag{D.3}$$

$$\text{Zero profit: } \pi_L = 0 \tag{D.4}$$

Substituting in the zero profit constraint gives the Lagrangian:

$$\mathcal{L} = V^L + \mu IC_H = u(c^*)(\mu - \mu p^H) - (\mu - \mu p^H + p^L - 1)u\left(\frac{p^L(c^1\theta^L + c^2(-\theta^L) + c^2 + l - 2w) + w}{1 - p^L}\right)$$
$$+ u(c^1)(\theta^L p^L - \theta^H \mu p^H) + \theta^H \mu p^H u(c^*) + \theta^H \mu p^H u(c^2) - \mu p^H u(c^2) - \theta^L p^L u(c^2)$$
$$+ p^L u(c^2) - \theta^H \mu p^H u(c^*) + \mu p^H u(c^*)$$

The change in welfare due to a decrease in θ^L is, since the high types get full insurance, proportional simply to the change in the maximized $V^L(\boldsymbol{c}^L)$. Hence, by the envelope theorem, after substituting in the zero profit constraint, the derivative of V^L and hence welfare with respect to θ^L is $\frac{\partial \mathcal{L}}{\partial \theta^L}$.

To compute this, first we calculate the value of μ, the multiplier, at the optimum. The first order condition with respect to c_1 is:

$$\frac{\theta^L p^L(\mu - \mu p^H + p^L - 1)u'\left(\frac{p^L(c_1\theta^L + c_2(-\theta^L) + c_2 + l - 2w) + w}{1 - p^L}\right)}{p^L - 1} + u'(c_1)(\theta^L p^L - \theta^H \mu p^H) = 0$$

Solving this we get:

$$\mu = \frac{\theta^L(p^L - 1)p^L\left(u'\left(\frac{p^L(c_1\theta^L + c_2(-\theta^L) + c_2 + l - 2w) + w}{1 - p^L}\right) + u'(c_1)\right)}{\theta^L(p^H - 1)p^L u'\left(\frac{p^L(c_1\theta^L + c_2(-\theta^L) + c_2 + l - 2w) + w}{1 - p^L}\right) + \theta^H p^H(p^L - 1)u'(c_1)}$$

at the optimum.

Hence $\frac{\partial \mathcal{L}}{\partial \theta^L}$ after substituting in for μ is given by

$$\frac{\partial \mathcal{L}}{\partial \theta^L} = p^L\left(\frac{(c_1 - c_2)u'(c_1)(\theta^H p^H(p^L - 1) - \theta^L p^H p^L + \theta^L p^L)u'\left(\frac{p^L(c_1\theta^L + c_2(-\theta^L) + c_2 + l - 2w) + w}{1 - p^L}\right)}{\theta^L(p^H - 1)p^L u'\left(\frac{p^L(c_1\theta^L + c_2(-\theta^L) + c_2 + l - 2w) + w}{1 - p^L}\right) + \theta^H p^H(p^L - 1)u'(c_1)} + u(c_1) - \right. \tag{D.5}$$

Since $c_1 \leq c_0 = \frac{p^L(c_1\theta^L + c_2(-\theta^L) + c_2 + l - 2w) + w}{1 - p^L}$ the denominator of the fraction is negative. The numerator of the fraction is positive after noting that:

$$(\theta^H p^H(p^L - 1) - \theta^L p^H p^L + \theta^L p^L = -p_H\theta_H(1 - p_L) + p_L\theta_L(1 - p_H) < 0.$$

And since $\theta_L \leq \theta_H \iff \rho \leq 1$ this means (see Crocker and Snow (2011)

Theorem 1) that $c_1 \leq c_2$ hence we have $\frac{\partial \mathcal{L}}{\partial \theta^L} < 0$ and so welfare increases as θ_L falls $\iff$ ρ goes towards zero.

■

D.3 Proof of Proposition 8

Proof. First, I show that, for a fixed price p, the total cost of those that buy the bundle $E(\phi_B(w_B) \mid w_B > p_B)$ is higher under Y than X when $X \precsim Y$. Note that $w_B > p_B$ iff $\phi(w_B) > \phi(p_B) \equiv \underline{\phi}$. Since $\phi_B(w_1+w_2)$ is supermodular, by proposition 12, we have that

$$\phi_B(W_X) \precsim_{SL} \phi_B(W_Y).$$

In particular, at a stop-loss of $\underline{w}_B$ we have that

$$\mathbb{E}_X\left[\max\{\phi(w_B) - \underline{\phi}, 0\}\right] \leq \mathbb{E}_Y\left[\max\{\phi(w_B) - \underline{\phi}, 0\}\right].$$

This is equivalent to:

$$\begin{aligned} & \mathbb{E}_X\left[\mathbb{1}\left[w_B > p_B\right] \times \left(\phi(w_B) - \underline{\phi}\right)\right] \leq \mathbb{E}_Y\left[\mathbb{1}\left[w_B > p_B\right] \times \left(\phi(w_B) - \underline{\phi}\right)\right] \\ \iff & \mathbb{E}_X\left[\mathbb{1}\left[w_B > p_B\right]\phi(w_B) - \mathbb{1}\left[w_B > p_B\right]\underline{\phi}\right] \leq \mathbb{E}_Y\left[\mathbb{1}\left[w_B > p_B\right]\phi(w_B) - \mathbb{1}\left[w_B > p_B\right]\underline{\phi}\right] \\ \iff & \mathbb{E}_X\left[\mathbb{1}\left[w_B > p_B\right]\phi(w_B)\right] - \underline{\phi}\mathbb{E}_X\left[\mathbb{1}\left[w_B > p_B\right]\right] \leq \mathbb{E}_Y\left[\mathbb{1}\left[w_B > p_B\right]\phi(w_B)\right] - \underline{\phi}\mathbb{E}_X\left[\mathbb{1}\left[w_B > p_B\right]\right] \\ \iff & \mathbb{E}_X\left[\mathbb{1}\left[w_B > p_B\right]\phi(w_B)\right] \leq \mathbb{E}_Y\left[\mathbb{1}\left[w_B > p_B\right]\phi(w_B)\right] \end{aligned}$$

where the final line follows from assumption 2.

But since average cost can be written as

$$AC_B = \mathbb{E}_X\left[\phi(w_B) \mid w_B > p_B\right] = \frac{\mathbb{E}\left[\mathbb{1}\left(w \in W_B\right) \times \phi_B(w)\right]}{\mathbb{E}\left[\mathbb{1}\left(w \in W_B\right)\right]}$$

it follows again by assumption 2 that we have

$$\mathbb{E}_X\left[\phi(w_B) \mid w_B > p_B\right] \leq \mathbb{E}_Y\left[\phi(w_B) \mid w_B > p_B\right]$$

This shows that at every price, the average cost decreases. In particular, at the equilibrium price under Y, P_B^Y, if before we had $p_B^Y = \mathbb{E}_Y\left[\phi(w_B) - \underline{\phi} \mid w_B > p_B^Y\right]$, under X we have $p_B^Y > \mathbb{E}_X\left[\phi(w_B) - \underline{\phi} \mid w_B > p_B^Y\right]$ implying, by the assumption of single crossing, that $p_B^X < p_B^Y$ the second part of the proposition states.

Finally, for the third part, by proposition 13, the demand curve for the bundle under X is above that for Y for all $q \geq \bar{q}$ where $\bar{q}$ is defined in the proof of proposition 13. That means, for any quantity insured beyond $\bar{q}$, the gap between demand and AC is higher under X than Y. In particular, at the equilibrium quantity q_B^Y the gap is zero under Y and positive under X, hence $q_B^X > q_B^Y$ as required.

■

D.4 Proof of Proposition 9

Proof. The total cost of those that buy the bundle at a small discount is

$$\mathbb{E}\left[\mathbb{1}\left[[w_1 \geq p_1 \wedge w_2 \geq p_2\right] \times (\phi_1(w_1) + \phi_2(w_2)\right].$$

Both $\mathbb{1}\left[[w_1 \geq p_1 \wedge w_2 \geq p_2\right]$ and $\phi_1(w_1) + \phi_2(w_2)$ are supermodular and increasing functions and hence so is their product. Then by proposition 12 with a stop-loss of zero we have

$$\mathbb{E}_X\left[\mathbb{1}\left[[w_1 \geq p_1 \wedge w_2 \geq p_2\right] \times (\phi_1(w_1) + \phi_2(w_2)\right] \leq \mathbb{E}_Y\left[\mathbb{1}\left[[w_1 \geq p_1 \wedge w_2 \geq p_2\right] \times (\phi_1(w_1) + \phi_2(w_2)\right].$$

Then by assumption 2 this implies that the average costs of those that buy the bundle are ordered in the same way:

$$\frac{\mathbb{E}_X\left[\mathbb{1}\left[[w_1 \geq p_1 \wedge w_2 \geq p_2\right] \times (\phi_1(w_1) + \phi_2(w_2)\right]}{\mathbb{E}_X\left[\mathbb{1}\left[[w_1 \geq p_1 \wedge w_2 \geq p_2\right]\right]} \leq \frac{\mathbb{E}_Y\left[\mathbb{1}\left[[w_1 \geq p_1 \wedge w_2 \geq p_2\right] \times (\phi_1(w_1) + \phi_2(w_2)\right]}{\mathbb{E}_Y\left[\mathbb{1}\left[[w_1 \geq p_1 \wedge w_2 \geq p_2\right]\right]}.$$

Since the bundled price is the same under each distribution, it follows that the profits from offering the bundled product are ordered $\pi_X^\epsilon \geq \pi_Y^\epsilon$ as required.

■

D.5 Proof of Propositions 10 and 11

Proof. Similarly to the proof of proposition 8 and 9 except we apply the logic to the increasing $-\phi_B(w_B)$ instead of $\phi_B(w_B)$ with a stop loss of $-\underline{\phi}$ not $\underline{\phi}$. In particular, the bundle is bought if $w_B \geq p_B \iff \phi(w_B) \leq \underline{\phi} \iff -\phi(w_B) \geq -\underline{\phi}$. We then conclude that

$$\mathbb{E}_X\left[-\phi(w_B) \mid w_B > p_B\right] \leq \mathbb{E}_Y\left[-\phi(w_B) \mid w_B > p_B\right].$$

Multiplying by minus one and reversing the inequality shows the first part. The rest are analagous to the proof of proposition 8. ■

References for Appendices

Cambanis, S., G. Simons, and W. Stout (1976, December). Inequalities for Ek(X, Y) when the marginals are fixed. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 36*(4), 285–294.

Crocker, K. J. and A. Snow (2011). Multidimensional Screening in Insurance Markets with Adverse Selection. *The Journal of Risk and Insurance 78*(2), 287–307.

Denuit, M., J. Dhaene, M. Goovaerts, and R. Kaas (2006, May). *Actuarial Theory for Dependent Risks: Measures, Orders and Models.* John Wiley and Sons.