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

Adam Solomon*

March 25, 2026

Abstract

Every insurance contract bundles risks, and explicit bundling discounts are common. I show theoretically that bundling arises in a competitive market whenever correlation between risk types enables insurer 'cream-skimming': 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 these discounts, and that model-predicted equilibrium bundling discounts closely match empirical discounts. I rule out standard economies of scale and differential contract lapsation as alternative explanations of the offered discounts. Counterfactually, banning bundling would raise welfare by 3.5% by correcting separate-market unraveling, while mandatory family bundling would reduce welfare by 5.6% by exacerbating advantageous selection. Finally, I show that imperfect competition can dull the incentive to bundle, counteract the externality from cream-skimming, and reverse policy prescriptions.

*Department of Finance, Stern School of Business, New York University. Email: adam.solomon@nyu.edu. I thank Jim Poterba, Amy Finkelstein, Jon Gruber, Nathan Hendren, Tim Layton, Frank Schilbach, Abhijit Banerjee, Stephen Morris, Johannes Spinnewijn, Daniel Gottlieb, Cameron Ellis, Ben Handel, Evan Soltas, Glenn Ellison, Casey Rothschild, Yijia Lin, Richard Phillips, Joan Schmit, and Martin Boyer 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 (ACA) requires that health insurance plans cover 10 'essential health benefits'.

Why is insurance sometimes bundled and sometimes not? What is the proper role of the government in regulating – or restricting – such bundling? These are essential questions that are central to the insurance industry. Yet, they are not addressed by either the theoretical or empirical work on insurance. Canonical insurance models such as Akerlof (1970) 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 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 (1970). 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 that costs are an increasing (decreasing) function of WTP. Second, whether the risk types in each market are positively or negatively correlated. They are positively (negatively) correlated when those with high WTPs for contract 1 typically have high (low) WTP for contract 2. I find that the incentives to bundle and the welfare consequences of regulations that mandate or forbid bundling depend directly 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 who buy both products and who would be attracted by a bundle 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 who 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 prevails - separate market unraveling versus bundled market expansion - is an empirical question. Whether government intervention that *forbids* 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 intervention that *forces* bundling improves welfare.[(1)] Choice is restricted, which reduces surplus before considering price adjustments. Consumers who prefer to buy one policy must now choose between buying both and buying neither. However, forced bundling can reduce prices by dampening selection. Consider whose purchase of the first policy changes after bundling is mandated. Those

[(1)]Examples of forced bundling include: ACA 'essential health benefits'; state-based mandates that auto insurance include Personal Injury Protection (PIP) or Uninsured Motorist (UM) coverage; requirements that workers' compensation insurance cover medical care, wage replacement and rehabilitation costs; the Federal Crop Insurance Program that only offers multi-peril insurance – farmers cannot insure against only specific perils.

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 high WTP for the second and moderate WTP for the first go from buying just the second policy 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 the equilibrium price, and mitigates the selection problem, but must be weighed against the direct welfare costs of reduced choice. Under advantageous selection, this raises the equilibrium price and unambiguously lowers welfare.

I apply this framework and insights to the long-term care insurance market (LTCI). Long-term care is one of the largest financial risks in later life, but the use of LTCI is very low. Large spousal bundling discounts are offered in this market: up to 35% if *both spouses purchase* an LTCI policy, in addition to, and distinct from, 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 this framework and empirical simulations with HRS data to explain these discounts, 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 the 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. A key distinction is between total LTC risk and total risk *net of observables on which insurers price*. My analysis considers both and yields identical qualitative conclusions.

I directly test how an individual's LTC risk correlates with their spouse's risk. I study the correlation in total risk, as well as risk net of observables on which insurers price. In both cases, I find a strong positive correlation between spousal risks. Those who have a spouse are marginally lower risk than those who are single. But within those who have a spouse, those with a spouse in the highest decile of risk have almost 60% higher risk themselves relative to those whose spouse is in the lowest decile. This rationalizes the large discount (up to 35%) given when both spouses buy a policy, and the much smaller discount 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 risk and WTP, over and above merely having a spouse. An insurer only learns this extra information when both spouses buy a policy.

Following Finkelstein and McGarry (2006), I find that LTCI is advantageously selected. My theory suggests that bundling should occur in this situation: positive correlation between risk types and advantageous selection. In the phrasing of Finkelstein and McGarry (2006), there are two types of people who buy LTCI, those who have private information that they are high risk and those who 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, are common to the household, and are more likely to buy insurance (see Gropper and Kuhnen (2021)). Bundling discounts separate the wealthy and healthy couples who both have low risk 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 informal care).

I rule out two alternative 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 of 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 premia (sometimes more than 40% annually) over the sample period owing to insurers radically underestimating the expected claims of existing policyholders. These premium increases were applied to both new and existing policyholders. This is, to my knowledge, a novel source of identification for demand in

long-term care insurance.

Given these flexibly estimated costs and WTPs, 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 no members of a household (whether single or a couple) buy a policy.

I find that banning bundling would raise welfare by 3.5%, while forcing family bundling would lower it by 5.6%. 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 14%, 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 the banned-bundling equilibrium. Overall, couples pay marginally more for two policies, but singles and couples purchasing only one policy pay substantially less, which increases welfare.

Finally, I analyze how bundling incentives and welfare implications change under imperfect competition, specifically a monopolist. Recall that in competitive markets there is an incentive to cream-skim low-cost policyholders through bundling discounts, without internalizing the cost-raising impact this has on the non-bundled markets. In contrast, a monopolist internalizes this impact, since they are the only seller in the non-bundled markets. But, in this second-best world, multiple market failures of selection and imperfect competition interact in complex ways. The monopolist can exploit the multiple insurance markets to engage in even more extreme quantity restrictions. Their incentive to raise the price of one contract is amplified (relative to a single-market monopolist) because some displaced customers will switch to another of the insurance contracts rather than exiting the market entirely, softening the usual profit penalty of a price hike. This raises the question of whether bundling regulations should differ under monopoly: banning bundling might do little to alleviate externalities that are already internalized, while forcing bundling could constrain the monopolist's ability to distort prices in individual markets.

I find that in the unrestricted mixed-bundling equilibrium, the monopolist sets very high prices in the relatively thin-tailed male and bundled markets, while setting a lower price for single-female policies to capture the priced-out consumers. This illustrates the cross-cutting tradeoffs: the monopolist's modest 5% bundling discount

reflects the now-internalized cream-skimming externality, while extreme prices in the male and bundled markets demonstrate the amplified incentive for quantity restriction. Banning bundling does very little to help single markets - prices change by less than $2 - since the monopolist already internalizes cross-market effects; it merely eliminates the small discount, lowering welfare by 2.5%. In contrast, forcing family bundling raises welfare by 7.5% by moderating the monopolist's pricing power: it prevents extreme price hikes, as couples would otherwise drop coverage, leading to more balanced (though still elevated) prices across genders. These policy prescriptions are inverted relative to the competitive case. This demonstrates that one cannot neglect the bundling-related distortion identified in this paper when conducting welfare analyses of 'standard' market failures.

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 competitive equilibria to analyze the quantitative welfare impacts of government interventions. In Section 5.2, I relax the assumption of perfect competition and study the equilibrium bundling prices and welfare implications of regulatory interventions when the market is monopolistic instead of competitive. 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 markets. Multiple papers have documented firm incentives to manipulate contracts to attract lower-cost enrollees, without exploring the welfare implications: Lavetti and Simon (2018) study selective formulary design in Medicare Advantage plans; Shepard (2022) analyzes insurer networks excluding 'star hospitals' to remove high-cost users; Cooper and Trivedi (2012) show that Medicare Advantage plans strategically include gym memberships to attract healthier enrollees. My contribution is to formalize the conditions under which bundling occurs, relate this to selection and correlation be-

tween 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 low take-up, despite being one of the largest late-life financial risks.[2] Finkelstein and McGarry (2006) find that two types of people have high willingness to pay (WTP) for LTC: those with high risk and those with high risk aversion (but low costs).[3] Murtaugh et al. (2001) studies the utility of a combined annuity and long-term care insurance product. They find that the simulated break-even premia for the bundled product are 3 to 5 percent lower than for the individual products, and attribute this to reduced selection. My contribution is to micro-found these observations in a formal model and estimate market primitives using individual-level risk data.

Third, there is a well-established Industrial Organization literature on bundling products and add-on pricing. Stigler (1963), Schmalensee (1984) and Adams and Yellen (1976), and more recently McAfee et al. (1989), Bakos and Brynjolfsson (1999), Chen and Riordan (2013), and Haghpanah and Hartline (2021), showed that bundling can dominate separate sales for a monopolist under negative correlation. Additional rationales for bundling include softening competition in differentiated products[4] or to deter entry[5]. Relative to that literature, my main contribution is introducing selection (e.g., non-constant marginal costs).

The most directly related papers are Fluet and Pannequin (1997) and Nguyen (2024). Fluet and Pannequin (1997) study a Rothschild-Stiglitz-style model with two types and two risks. They find that welfare depends on (i) whether the payoffs when one loss occurs can be conditioned on the other risk's occurrence, (ii) whether the purchase of one contract can be conditioned on the purchase of the other, and (iii) whether the two types' risk propensities are ordered in the same way. Nguyen

[2] 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)).

[3] Fang and Wu (2018) study the market for a single insurance contract under multi-dimensional types (e.g., risk and risk aversion). They find that, contrary to these empirical patterns, advantageous selection cannot emerge in equilibrium. Their key assumption is that all agents have a WTP for the insurance above their costs. This assumption is plausibly violated in the context of LTC and other health insurance settings, where there is public insurance of last resort, or informal care by relatives. See, for example, Finkelstein et al. (2019).

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

[5] Nalebuff (2004)

(2024) studies familial health insurance choice in Vietnam and 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).

Relative to these papers, my contribution is (a) to study, in the first 'sufficient statistics' theoretical model with multiple risks, when it is profitable for insurers to move from separate to bundled contracts, and relate this to selection and correlation; (b) to use individual-level HRS data to show that bundling discounts in long-term care insurance are rationalized by cream-skimming incentives; (c) to estimate a structural model for costs and demand which I use to evaluate the welfare impact of bundling, and regulatory interventions that force or forbid it, under competition and monopoly.

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 B.1. There are two risks and a fixed insurance contract for each.[6] I will refer to risks 1 and 2, and to the fixed insurance contracts for each risk as, respectively, contracts 1 and 2. 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, then each individual's WTP for the bundled contract is $w_B = w_1 + w_2$.[7] This strong

[6] In Appendix B.2, I study endogenous contracts in a simplified setting, and find that many insights qualitatively remain.

[7] 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,

assumption is made for theoretical tractability;[8] I relax it in the empirical model and counterfactuals in Section 5.

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 $\phi_m'(w_m) > 0$, risk m is adversely selected; if $\phi_m'(w_m) < 0$, risk m is advantageously selected.*

Adverse or advantageous selection is assumed to hold globally. I refer to a joint distribution over risks (w_1, w_2) as an **economy** X, 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, F_1 for the CDF of the marginal distribution of (the WTP for insurance against) risk 1, and F_2 for risk 2. The comparative static of interest is between economies that have the same marginal distributions but different joint distributions 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 competing 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 to attract any demand. In equilibrium, all firms will offer the same prices at which total profits will be zero. 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 who purchase both, perhaps at a bundle discount, and those who purchase neither. 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 the 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 want to purchase both products. Only if a bundle

perhaps the sum of losses exceeds some deductible. This is studied in, for example, Klosin and Solomon (2024).

[8] When utility is CARA, the assumption is without loss of generality, as I study in Appendix B.1.

discount is offered, $p_B < p_1 + p_2$, will anyone be in $\mathcal{D}_B$, and will that market have to clear separately 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:* $p_m = AC_m\left(p_m, p_{-m}\right)$ *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 ensures 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_B^*).

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

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

Overall welfare is simply the sum of welfare in markets 1, 2, and B. In equilibrium, profits will be zero, and therefore 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 to vary only the correlation structure.

We define and work within the set of all joint distributions with the same marginal distributions:

Definition 3. *Define* $\Gamma(F_1, F_2)$ *as the set of joint distribution functions with common marginals* F_1 *and* F_2*.*

This operationalizes the comparative static of interest: two joint distributions with the same marginal distributions that only differ in their correlation structures. The prices that clear the separate markets when there is no bundling will be important throughout the analysis:

Definition 4. *Assume bundling discounts are not allowed. For all distributions in* $\Gamma(F_1, F_2)$*, define the prices* $\boldsymbol{p}^\Gamma = \left(p_1^\Gamma, p_2^\Gamma\right)$ *that clear the separate markets as the solution to* $p_1^\Gamma = AC_1\left(p_1^\Gamma; p_2^\Gamma\right)$, $p_2^\Gamma = AC_2\left(p_2^\Gamma; p_1^\Gamma\right)$.

Within $\Gamma(F_1, F_2)$, these separate-market clearing prices (i.e. when there is no bundling discount) are exogenously fixed; they are not endogenous equilibrium objects.

Lemma 1. *The prices that clear the separate markets when there is no bundling,* $\boldsymbol{p}^\Gamma = (p_1^\Gamma, p_2^\Gamma)$*, (per Definition 4) are equal for all distributions in* $\Gamma(F_1, F_2)$*.*

Following Shaked and Shanthikumar (2007)[9] and Denuit et al. (2006), the correlation order between distributions in $\Gamma(F_1, F_2)$ is defined as follows.

Definition 5. *Suppose* $X, Y \in \Gamma(F_1, F_2)$ *and have CDFs* F_X, F_Y *respectively. I say that* X *is less correlated than* Y *or that* X *precedes* Y *in the correlation order, written as* $X \precsim Y$ *iff*

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

This says that w_1, w_2 are more likely to both be small or both be large[10] 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.

Note that the correlation order can be generalized to other distributions related to X and Y. For example, one can consider the distributions of X and Y conditional on event A.[11]

[9] There, this ordering is called Positive Quadrant Dependence (PQD)

[10] Because X, Y have the same marginal distributions, it is the case that $F_X(w_1, w_2) \leq F_Y(w_1, w_2) \iff 1 - F_X(w_1, w_2) \leq 1 - F_Y(w_1, w_2)$.

[11] I write $X \precsim_A Y$ for the correlation order, conditional on event A. A formal definition and dis-

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).

Recall the (exogenous, within $\Gamma(F_1, F_2)$) prices that clear the separate markets are p_1^Γ, p_2^Γ. Figure 1 illustrates who buys product 1 and 2 at (p_1^Γ, p_2^Γ), and who would buy the bundled product if an ϵ discount were offered. At the separate-market clearing prices, everyone who has $w_1 \geq p_1^\Gamma$ buys product 1 (groups A and B1) and everyone who has $w_2 \geq p_2^\Gamma$ buys product 2 (groups A and B2). Group C buys nothing. By the definition of (p_1^Γ, p_2^Γ), the average cost of those who buy product 1 is exactly equal to the price, and similarly for product 2.

Suppose that an insurer starts selling the bundle at a small discount: $p_B = p_1^\Gamma + p_2^\Gamma - \epsilon$ for small $\epsilon > 0$. Those who buy the bundle are exactly those who separately bought both products prior - group A. The firm makes a profit by offering a small bundling discount when the expected cost of insuring risk 1 (2) is lower for group A than B1 (B2).

Consider the left panel of Figure 1. WTP for each separate product is positively correlated. In this case, those that buy the bundle - group A - 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 A 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 WTPs are negatively correlated.

cussion is in Appendix B.3. The most common conditional correlation order that I use is conditional on being in the set of buyers for the product $m = 1, 2$ and the price p_m, where p_m is an exogenously fixed price. 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.

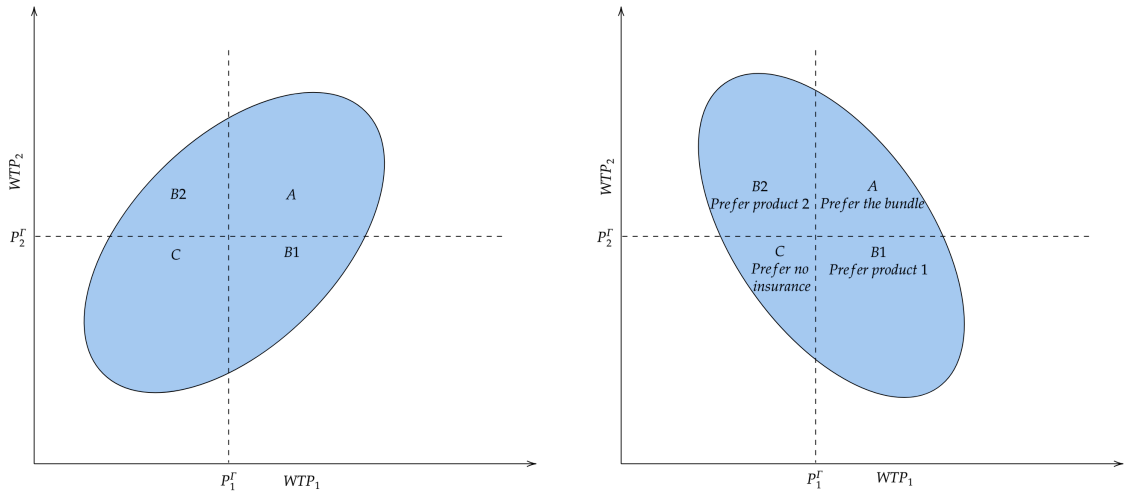


Figure 1: An illustration of which subsets of types buy in the separate markets, and which would swap to the bundled product if a small discount were offered. The left panel depicts the case of positive correlation between WTPs, the right panel depicts negative correlation.

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 A 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 intuition shows why an insurer will start offering a bundling discount: group A is cheaper to sell to than groups B1 and B2. This is true whether or not they also withdraw from or alter their price in the separate markets, since they get all of the profit from the bundled market, and only a portion of the loss in the separate markets (assuming many firms). This is formalized in the following proposition.

Proposition 1. *Starting from the separate-market clearing prices, offering a bundling discount is profitable if and only if*[12] *there is:*

1. *Adverse selection and negative correlation*[13] *between the risks, or*

[12] Assuming that both risks are adversely selected, or that both risks are advantageously selected.
[13] Here, X having negative correlation formally means, for $X, Z \in \Gamma(F_1, F_2)$ with Z independent,

2. Advantageous selection and positive correlation between the 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 who 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 separate prices. This is formalized below:

Proposition 2. *If there is adverse selection and negative correlation, or advantageous selection and positive correlation, a mixed-bundling equilibrium will occur and*

- *The prices in the separate markets will be increased relative to the separate-market clearing prices:* $p_1^\Gamma < p_1^{MB}, \quad p_2^\Gamma < p_2^{MB}$;
- *Under adverse selection, the bundle price will be discounted relative to the separate-market clearing prices:* $p_B^{MB} < p_1^\Gamma + p_2^\Gamma$.

This broadly 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. However, a bundling discount exerts an externality on the separate markets: prices rise and demand contracts. I study the net welfare impact of bundling, and potential corrective government interventions, in Section 4.

$X \precsim_{p_m^\Gamma} Z$ for $m = 1, 2$, where p_m^Γ are the separate-market clearing prices from Definition 4. Per Lemma 1, the separate market clearing prices p_m^Γ are exogenously fixed within the class of distributions $\Gamma(F_1, F_2)$ that we consider; they are not equilibrium objects. Positive correlation is defined similarly.

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 these from pricing on marital status, 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.

3.1 Connecting Sections 2 and 3

The theoretical framework in Section 2 treats willingness-to-pay (WTP) for insurance and the cost of providing insurance as primitives. Estimating these quantities from observable data requires a fully specified structural model, which I defer to Section 5. Instead, in this section, I analyze two readily estimable objects — LTC risk and insurance purchases — to show that the theory's qualitative predictions about when bundling occurs hold in this setting.

This requires a translation of the model primitives into empirically estimable objects:

Risk vs. Cost

- Risk (estimated here) serves as a direct proxy for cost (the primitive in the prior section). Individuals with higher risk are assumed to have higher costs.

Selection

- *Prior section:* Adverse (advantageous) selection meant high-WTP individuals had high (low) costs.
- *This section:* Adverse (advantageous) selection means high-risk individuals are more (less) likely to purchase insurance.

Correlation

- *Prior section:* Defined in terms of WTP for each contract.

- *This section:* Defined in terms of risk: spouses' LTC risks are positively (negatively) correlated if a high-risk individual typically has a high- (low-)risk spouse.

Restating the second half of Proposition 1 in terms of these empirical objects: bundling occurs in equilibrium when (i) there is *advantageous selection* (low-risk individuals are more likely to purchase LTC insurance) and (ii) spouses' risks are positively correlated. These are the predictions I test in this section.

3.2 LTC Insurance 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 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 behavioral frictions (Gottlieb and Mitchell (2020)). I demonstrate how bundling can also contribute to depressed demand.

3.3 LTCI Discounts for Spouses or Partners

Substantial discounts are offered to couples who both purchase LTCI: on average 25%, per Figure 4. This is computed from the universe of LTC policies in LTC Quote Plus (StrateCision (2022b)), which is software used by insurance brokers that provides quotes from nine different issuers. It shows that steep discounts for couples who both purchase 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.

Product	LTC
Mean Couples' 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.

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 unmarried to understand the risk differences associated with 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.

3.4 Data and Risk Prediction

The data I use come from the Health and Retirement Study (HRS (2018)). The HRS is a biennial survey that has been conducted since 1990, so that long-run outcomes such as long-term care utilization or death 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 (ADLs) 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. I restrict to 64-66 -year olds, to avoid repeated sampling of the same individuals. All predictions are via LASSO penalized logistic regression, and incorporate cross-validation to avoid overfitting. 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.

A key distinction is between risk observable to the insurer and already priced on, and risk unobservable to the insurer. Bundling discounts implicitly screen on the latter. To capture this, I make two predictions. First, I predict risk using the full set of relevant variables in the HRS, which are substantially richer than those available to insurers. Second, I make a coarser prediction of risk using only variables (health, age, sex, etc.) that are observable to the insurer. The difference between these two predictions is a proxy for the unobservable risk that the insurers use bundling discounts to screen on. Details of the variables used in the main predictions, and in the smaller predictions that mimic insurance company underwriting information, are in Appendix C.7.

In the analysis that follows, I exhibit the results both for the total risk (from my rich predictive model) and for the risk unobservable to the insurer (the residual between my rich predictive model and the coarser model using only insurer-observed variables). The results are qualitatively identical.

3.5 Spousal Risks (Total and Residualized on Priced Observables) are Positively Correlated

An individual's risk is increasing in their spouse's (ordinal) risk for LTC insurance. This is true for total risk (observable to the econometrician) and for the component of risk unobservable to insurers. I bin individuals by the decile of their spouse's risk, as well as a bin for single respondents. I compute the mean 14-year LTC risk in each bin. The results are in Figure 3. For reference, the black horizontal dashed line is the average risk for all those with a spouse.

First, single individuals are higher risk than those who have a spouse, after conditioning on the other information insurers price on. In the bottom panel, compare

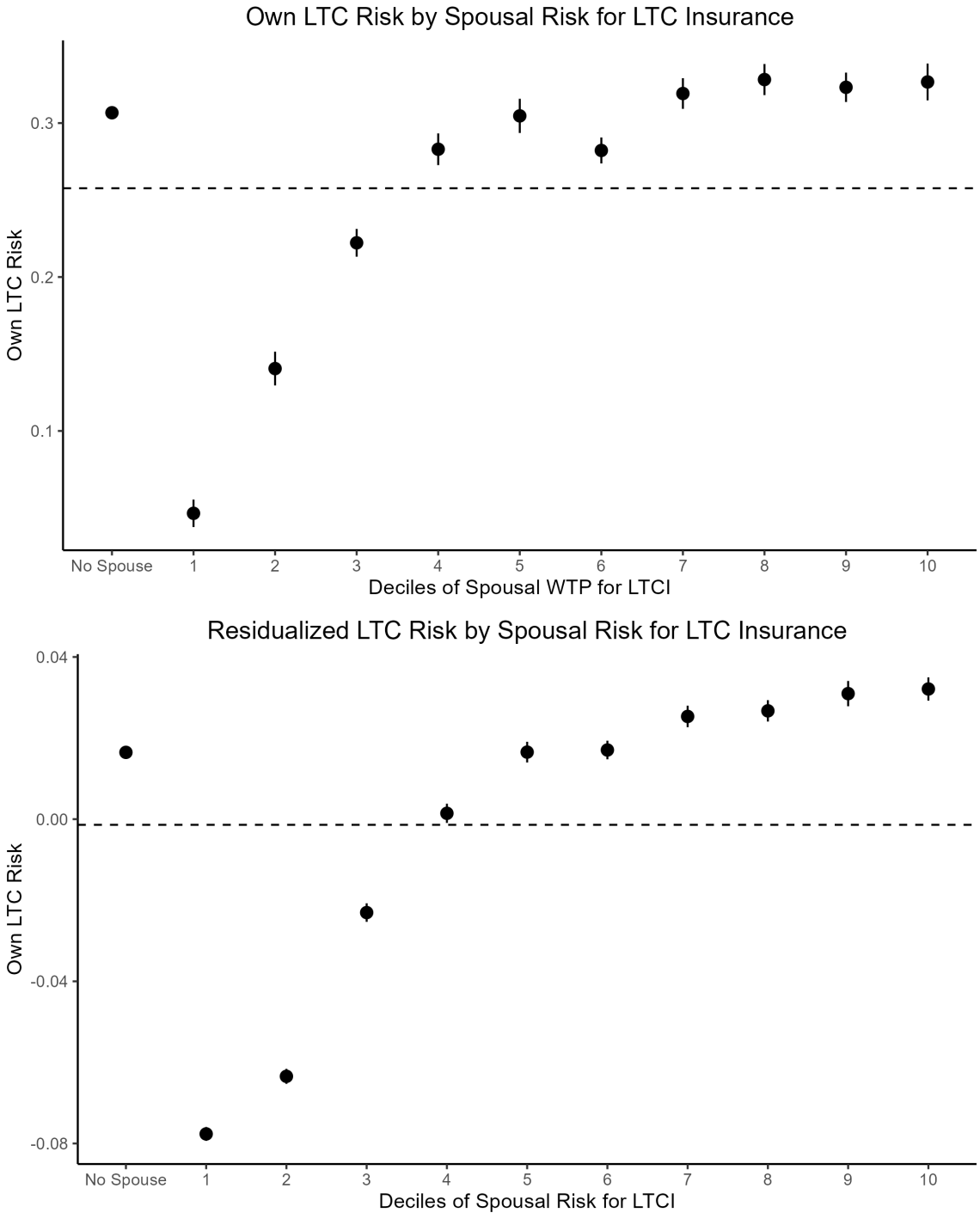


Figure 3: Binscatter of LTC risk by deciles of spousal risk. The horizontal dashed line is the average risk amongst all those who have a spouse. The top panel uses total own and spousal risk. The bottom panel uses own and spousal risk residualized on the variables insurers observe.

the average risk in the leftmost 'No Spouse' bucket with the average risk of all those with a spouse, shown by the horizontal dashed line. Although hard to numerically interpret, the residual risk is higher for the average non-married respondent than the average married respondent. This explains why simply being married attracts a small but positive discount.

Second and most importantly, *conditional on having a spouse*, a respondent's LTC risk increases in their spouse's risk. In the top panel, those with high spousal risk (the rightmost bins) are over 66% higher risk than those with the lowest spousal risk (the bin second from the left). The bottom panel exhibits a similar pattern: having a high-risk spouse strongly predicts high own risk, after residualizing on insurer-observable factors.

This shows that a bundling discount plays a screening role over and above simply pricing on being married. Within those who are married, knowing that an individual has a high-risk spouse predicts high risk. This is true even net of observables on which insurers price.

Given the strong positive correlation observed, Proposition 1 predicts that bundling will occur when there is advantageous selection. When low risk is correlated with high WTP, those who buy the bundle - couples who both have high WTP - will be lower cost than those who do not. I now test for advantageous selection.

Correlations in Propensities versus Outcomes. The correlations studied here are correlations in ex-ante propensities within a household. This is related to, but distinct from, correlations in ex-post outcomes. In Appendix C.3, I show that outcomes are positively correlated, conditional on propensities. That is, conditional on ex-ante risk, if a person is unlucky and requires LTC, the chance of their spouse requiring LTC increases. Intuitively, this is because informal home care can no longer be provided to the second spouse if the first spouse is themselves receiving LTC. This correlation in outcomes is a *mechanism* for the positive correlation in propensities observed here: in addition to correlated wealth and health, high-risk individuals are more likely to need LTC, causally increasing the chance of their spouse needing LTC. Nevertheless, correlation in ex-ante propensities, which incorporates any effect of ex-post correlation in outcomes, is the sufficient statistic.

3.6 LTC Insurance is Advantageously Selected

I replicate the analysis performed by Finkelstein and McGarry (2006) on the larger and longer-duration HRS sample that has been collected since then. I confirm their finding that long-term care insurance is advantageously selected over a longer time horizon. 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, for all 65-year-olds, I estimate probit models in which the outcomes are binary indicators for using any LTC in the next 14 years, and a binary indicator for whether the respondent has LTCI when surveyed. The control variables are those used in insurance company pricing, such as age, sex, and major medical conditions.[(14)]

The results are in Table 1 below.

LTC usage within:	14 Years		
	(1) No Controls	(2) Controls for Insurance Company Prediction	(3) (2) Restricted to Non-Rejects
Coefficient from Probit of LTC Usage on LTC Insurance	-0.105^{*} (0.063)	-0.105 (0.064)	-0.168^{**} (0.085)
N	4038	4038	2397

Note: ***/**/* denotes significance at the 1/5/10% level.

Table 1: 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 selection. This is consistent with Finkelstein and McGarry (2006). Their observation period was 5 years, yet I find clear evidence of advantageous selection over a 14-year time horizon. In all specifications, the coefficient is negative, and in the most realistic model (3), which conditions on controls that insurers price on and removes those with a rejectable health condition, the coefficient is most negative and significant at the 5% level.

I have demonstrated two facts: couples have positively correlated LTC risk, and LTC insurance purchases are advantageously selected. Together, these justify the sharp discounts offered to couples buying LTC insurance, consistent with the 'cream-

[(14)] For details, see Appendices C.6 and C.7.

skimming' motive predicted by Proposition 1. In Section 5, I demonstrate this to be true when WTP and costs are flexibly estimated instead of proxied for via risk.

3.7 Ruling Out Alternative Explanations

In the remainder of this section, I analyze standard (i.e. non-selection-based) explanations for the bundling discounts offered to couples in LTC insurance. Specifically, I show that economies of scale and differential lapsation could not plausibly generate these discounts.

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' discounts 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 C.1. The results are reported in Figure 4.

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 4: 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

[(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 annuities. Details are in Appendix C.1.

substantial bundling discounts, I rule out generic economies of scale as a justification for couples' discounts in LTC.

This is consistent with intuition. The vast majority of the cost of servicing an insurance policy comes from acquisition costs - primarily commissions paid 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 other expenses are policy-owner services that primarily consist of lapsation and surrender costs. If lapsation patterns amongst couples were more favorable to the insurers than singles, this might explain some of the discount offered. In Appendix C.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 restrictions I will study:

1. The **mixed-bundling** market is when both separate products are offered, and the bundle can be offered at a discount. This is the default status quo outcome when there is no government intervention.

2. The **no-bundling** market, is when the government bans bundling, and only the separate markets exist.

3. The **pure-bundling** market, where the government disallows the sale of separate products, and only the bundled product can be bought.

4. The **family-bundling** market, where I think of each type as a family, is when the government compels families that have two members to buy both policies or neither, but allows singles to buy only one policy or none. Family bundling forbids bundling discounts.

4.1 Government Interventions: Pure Bundling

Once the government mandates bundling, only bundled risk type matters. How the distribution of bundled risk types relates 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 equilibria prices, and the same set of people will buy the bundled product as bought both products prior.

As correlation gets less positive or even negative, the bundled market looks increasingly 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 fewer people with very high or very low WTP for the bundle, and many more in the middle.

A reduction in 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 who 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 pointwise.

This is formalized in the following proposition:

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, and therefore surplus and the equilibrium quantity, in the bundled market also change.

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? When bundling is forced, some consumers who were buying just product 1 (those with sufficiently high combined WTP but low WTP for product 2) move from buying just product 1 to the bundle, and similarly some who were buying just product 2 (those with sufficiently high combined WTP but low WTP for product 1) move to the bundle. On net, the group of people who end up with product 1 has decreased in WTP: the consumers with higher WTP for product 1 who previously bought only product 1 but now buy nothing (because their combined WTP falls below p_B) have been replaced with consumers who have lower WTP for product 1 but previously bought only product 2 and now buy the bundle. 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 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 in the following:

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

- *Under adverse selection, the effect on welfare is ambiguous.*

- *Under advantageous selection and approximate symmetry*[16] *in the risks, the price of the bundle rises and welfare unambiguously falls.*

[16] Formally, that the density of types in B2 equals the density in D1, defined at the mixed-bundling equilibrium prices.

4.2 Government Interventions: Family Bundling

Family bundling synthesizes the potential benefits of forced bundling, while not disadvantaging singles. Bundling discounts are disallowed; families with two members must either both buy policies or neither can buy. 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 prefer not to insure or forgo coverage altogether. This initially hurts these families. Whether overall welfare can rise depends on the extent to which prices fall due to the blunting of selection. When prices rise after family bundling, welfare unambiguously falls: everyone is worse off. As in the case of pure bundling, family bundling typically leads to higher prices under advantageous selection, but lower prices under adverse selection.

This speaks to the empirical work in Nguyen (2024). 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 is 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.

4.3 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 bundling is banned, the prices in the separate markets fall while the price of buying both products rises (per Proposition 2). As a result, some groups are better off, while many are worse off. Which of these dominates will be specific to the empirical setting.

Consumers who buy both products in both situations are worse off since they do not receive a discount. Similarly, consumers who would buy the bundle with the discount but, when bundling is banned, buy nothing are worse off.

Those who purchase exactly one product in both regimes are better off when bundling is banned as the low-cost would-be bundlers lower prices in the single markets. Additionally, those who go from purchasing nothing to purchasing one product

when bundling is banned go from zero to positive surplus. The welfare change of those who go from buying the bundle to buying a single product when bundling is banned is ambiguous.

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. To do so, I specify and estimate a rich structural model of WTP and costs for insurance that relaxes many of the restrictive assumptions made in Section 2. In particular, I do not assume that WTP for the bundled insurance is additive: I allow for insurance purchases within the household to be substitutes or complements.

After estimating this flexible model, I consider counterfactual regulations and equilibria. 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 buy). 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.

5.1.1 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, I need to extrapolate the estimate of LTC risk in the next 14 years to LTC risk over the whole lifetime. Since the respondents are at most 65, much of the long-term care risk is expected to occur

in more than 14 years. Second, I 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 3.4. 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 3.4 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 of long-term care ($66,327, a blended average of nursing home and home health care) from Savell (2023), multiply by the gender-averaged years of care required conditional on needing care (2.26 years), and take the present discounted value, accounting for expected LTC cost inflation and time until care is 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 C.4 I run the same analyses with a wide range of (distributions of) expected costs, and the signs of the welfare conclusions obtained are identical.

5.1.2 Modeling and Estimating WTP

Two-step procedure. I estimate household WTP in two steps. First, individual WTPs are estimated using a standard procedure described in Appendix C.8. Second, I estimate the intrafamily measure of substitutability/complementarity from a multinomial choice model described below.

Choice set. A married household chooses from $\mathcal{J} = \{\text{none}, m, f, b\}$, corresponding to buying *no* LTC insurance, insurance only for the male (m), only for the female (f), or a *bundle* covering both spouses (b).

Indirect utilities. For spouse $k \in \{m, f\}$, the latent indirect utility from a standalone contract is

$$U_k \;=\; \beta_1 p_k + \beta_2 a_k + \beta_3\,\texttt{married}_k + \boldsymbol{\beta}^{\text{inc}}\cdot\mathbf{inc}_k + \boldsymbol{\beta}^{\text{risk}}\cdot\mathbf{risk}_k + \mathbf{1}\{k=m\}\beta_{\text{sex}}, \quad (5.1)$$

where p_k is the quoted premium[17], a_k is age, and $\mathbf{inc}_k$ and $\mathbf{risk}_k$ are vectors of income-group and risk dummies.

The bundle delivers

$$U_b \;=\; U_m + U_f \;+\; \beta_{\text{bundle}}\big(p_m + p_f\big)\delta \;+\; \gamma, \quad (5.2)$$

with $\delta \in (0,1)$ the price discount on the second policy (assumed to be 35%) and γ encodes the intra-household substitutability or complementarity of two insurance policies.

Utility for 'none' is normalized to 0. Assuming i.i.d. type-I extreme-value errors, the purchase probabilities are

$$P_j \;=\; \frac{\exp\big(U_j\big)}{1 + \sum_{\ell \neq \text{none}} \exp\big(U_\ell\big)}, \qquad j \in \{m, f, b\}. \quad (5.3)$$

Identification and estimation. To overcome the endogeneity problem, I use the sharp and unexpected increases in premia throughout the sample period (1996 to 2006) owing to incorrect insurer assumptions regarding lapsation and longevity.[18]

[17] Note, 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 based on sex, age, marital status, and risk (coarsely). In the sample, approximately 13% of the individuals have LTC insurance. I observe these individuals' premia. Hence, given the observable information on variables used in pricing, I can impute the premia that would be offered to those without LTCI.

[18] Premium increases were sharp, sometimes over 40% *each year* (Duhigg (2007)), and were applied to existing and new policies. By regulation, premium increases must be approved by state insurance commissioners. Rate increases are only allowed when the insurer can show that costs are higher than expected, which is justified by historical claims experience (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 materializes, even if it is clear that future claims are likely to be much higher. For this reason, I do not pursue this issue further.

The structural parameters $\boldsymbol{\theta} = \left(\beta_1, \beta_2, \beta_3, \boldsymbol{\beta}^{\text{inc}}, \boldsymbol{\beta}^{\text{risk}}, \beta_{\text{sex}}, \beta_{\text{bundle}}, \gamma\right)$ are estimated by maximum likelihood: $\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \sum_i \sum_{j \in \mathcal{J}} \mathbf{1}\{j_i = j\} \log P_{ij}(\boldsymbol{\theta})$, where j_i denotes the observed choice of the household i.

Recovering intra-household substitutability/complementarity. Within each household, I recover WTPs for the (as relevant) male, female and bundled policies, $(\widehat{w}_m, \widehat{w}_f, \widehat{w}_b)$ so as to minimize the distance between the choice probabilities predicted by (5.3) and the observed choice frequencies in the data.

After recovering WTPs, I define the dollarized measure of intrafamily substitutability/complementarity as $\overline{\gamma} = \widehat{w}_b - (\widehat{w}_m + \widehat{w}_f)$.

WTP Estimates Individual WTPs (as estimated in Appendix C.8) average \$101 for singles and \$170 for members of a couple. This is substantially lower than the average expected present discounted cost of \$276 per month.

Within-family demand is substitutable: $\overline{\gamma}$ is estimated to be -\$57 on average. This comports with economic intuition: all else equal, once a family is insured against one shock, their WTP for insurance against another shock decreases. Hence, the average household WTP for a couple is approximately \$295, well below the sum of their WTPs for the individual products.

5.1.3 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 is w_F and for the bundle it is $w_M + w_F + \overline{\gamma}$. In the case that the family consists only of a single man or woman, 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 of families who choose to buy product $m = M, F, B$ is denoted $\mathcal{D}_m$, as defined in Section 2.

I compute equilibrium prices that solve $p_m = AC_m\left(p_m, p_{-m}\right)$ where $m = M, F, B$ when the latter is allowed. To remove lumpiness, smooth out the market and address finite-sample issues, I duplicate the data and add small random noise to costs and WTPs. Since, by construction, equilibrium prices ensure zero profits, welfare is the sum of consumer surplus in the three markets:

$$\text{Welfare} = \sum_{m=M,F,B} \int_{\mathcal{W}} (w_m - p_m) \mathbb{1}(w \in \mathcal{D}_m) \, dF(w). \tag{5.4}$$

5.1.4 Equilibrium Simulations: Mixed Bundling, Banned Bundling, Family Bundling

I simulate the equilibrium under three regimes. First, where firms are free to offer products separately, or together with a bundling discount. This mimics the current free-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 restriction is moot), 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 5.

	Mixed Bundling (status quo)	Family Bundling	Banned Bundling
Equilibrium Prices:			
p_M	\$252.06	\$263.88	\$259.06
p_F	\$358.53	\$273.01	\$269.68
p_B	\$516.76	-	-
Bundle Discount	18.2%	-	-
Proportion Insured:			
Singles	10.8%	18.4%	19.3%
Couples With One Insured	8.1%	-	19.5%
Couples With Both Insured	49.1%	51.9%	43.1%
Welfare Change (Relative to Mixed Bundling)	-	-5.6%	+3.5%

Table 2: Results of Equilibrium Simulations

The mixed bundling regime that has no restrictions on the products that can be offered closely mimics the empirically observed market. The equilibrium prices are close to actual premia in the market: $\approx$ \$250 to \$400 for a 65 year old male (depending

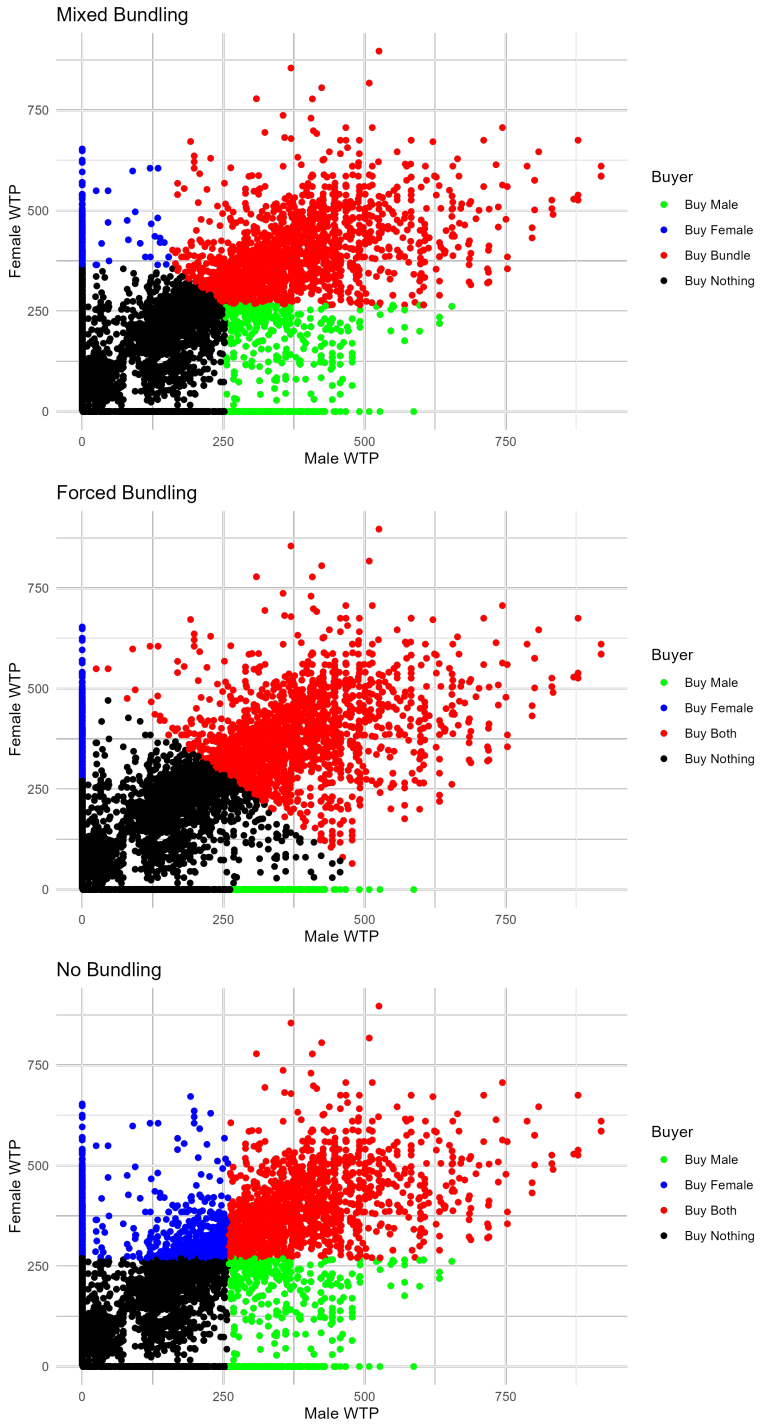


Figure 5: Illustrations of demand for male LTCI, female LTCI and both/bundle. The top, middle and bottom panels respectively illustrate the regimes of mixed bundling (no regulatory intervention), forced family bundling, and banned bundling. The singles are coded as having zero WTP for the other product, and appear on the axes.

on product and underwriting class) and ≈ $350 to $550 for a 65-year-old female. The simulated bundling discount offered is 18% of the sum of male and female individual premia, closely matching observed bundling discounts (see Table 4).

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

Forced family bundling leads to changes in prices and quantity covered. Comparing the second column to the first column of Table 2: the equilibrium price for men increases from $252.06 under mixed bundling to $263.88 under family bundling, while for women the price decreases substantially from $358.53 to $273.01. The sum of the prices under family bundling is $536.89, higher than the bundle price of $516.76 under the status quo of mixed bundling. This is consistent with Proposition 4.

These price changes significantly help singles, 18.4% of whom now buy insurance instead of 10.8%. Family bundling reduces the total demand for couples (only 52% of couples buy insurance under family bundling, compared to more than 57% who buy at least one policy under mixed bundling). This is due to a) couples who had previously bought both now choose to buy nothing as the price has increased, and b) some couples who had previously bought one policy now choose no insurance at all.

Forced family bundling decreases welfare by more than 5% relative to mixed bundling. This is because the welfare losses from restricting choice and increasing the cost of both members of a couple insuring outweigh the price decline in the female market.

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

When bundling is banned, the prices in the separate markets change significantly. The price for men rises slightly from $252.06 to $259.06, while the price for women drops from $358.53 to $269.68, a reduction of 25%. This is consistent with Proposition 2, which shows that bundling occurs exactly because those who buy the bundle are cheaper than those who do not. When bundling is banned and the former are mixed into the latter's separate policy, the cost falls.

When bundling is banned, welfare increases by 3.5% relative to mixed bundling. Even though couples now pay $12 more to buy two policies ($259.06 + $269.68 = $528.74 > $516.76), this is outweighed by the substantial price decreases in the female

market.

The general conclusion, which in Appendix C.4 I show to be qualitatively robust to alternative data choices regarding costs, is that forced family bundling typically reduces welfare by limiting choice, whereas banning bundling increases welfare by undoing the effects of cream-skimming in the single markets.

5.2 Monopolist Incentives to Bundle

The primary mechanism by which bundling occurs in a competitive market is 'cream-skimming'. Insurers bundle to steal the low-cost buyers and leave their competitors with the high-cost buyers. This can lead to welfare losses, since bundling firms do not internalize the impact that their bundling discount has on the individual markets. This relies on the fact that the insurer doing the 'cream-skimming' in the bundled market is not the same insurer hurt in the single markets.

In this section, I study the extreme case, monopoly, and show that these bundling incentives are entirely absent. Monopolists internalize the impacts of 'cream-skimming' via bundling, because there are no competitors from whom to steal the low-cost buyers. This could lead to a welfare gain over the mixed bundling equilibrium in the competitive market. However, monopolists distort markets in standard ways: by reducing supply so as to raise price. I study how regulatory interventions - mandating or banning bundling - should change as the market structure moves from competitive to imperfectly competitive.

I simulate the monopolist's profit-maximizing prices under the mixed bundling, banned bundling, and forced family bundling regimes. The results, under the baseline cost assumptions, are illustrated in Figure 6.

The monopolist almost completely shuts down the male single-policy market: they set a price of $\approx$ \$587 instead of the competitive equilibrium price of $\approx$ \$250. In contrast, the monopolist sets a low price for a single-female policy: $\approx$ \$252 instead of $\approx$ \$358. The female market (given other prices) is dense in the middle of the WTP distribution, and so the monopolist sets a low price because the marginal quantity effects outweigh the inframarginal price effects. This allows them to set a very high price in the bundled market ($\approx$ \$795 versus $\approx$ \$516) and extract surplus from the relatively thick-tailed bundled market. This demonstrates that a version of the standard distortion from monopoly holds in this bundling setting.

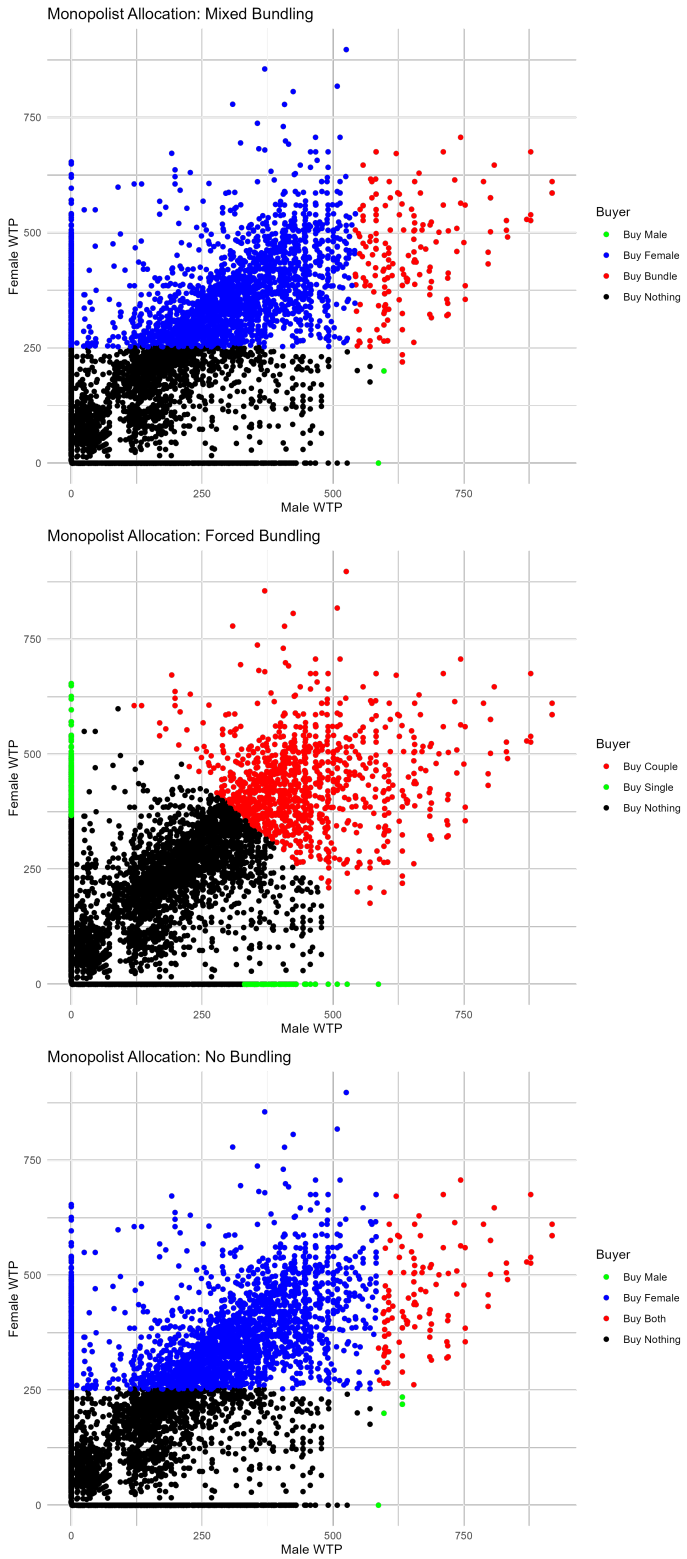


Figure 6: Illustrations of monopolist-optimal prices and corresponding demand under different bundling regimes. The top, middle and bottom panels respectively illustrate the regimes of mixed bundling (no regulatory intervention), forced family bundling, and banned bundling. The singles are coded as having zero WTP for the other product and therefore appear on the axes.

Banning bundling lowers welfare by 2.5%. This is in contrast to perfect competition, in which banning bundling raised welfare. When bundling is allowed, the monopolist offers a bundling discount of only 5%. Because a monopolist already internalizes the cream-skimming externality, banning bundling does little to help the single markets – prices change by less than $2; it simply removes the small bundling discount, reducing welfare.

On the other hand, also in direct contrast to the competitive case, forced family bundling raises welfare relative to mixed bundling by 7.5%. Under family bundling, the monopolist's incentive to set a very high male price is lessened, as this would cause many couples to buy nothing instead of both. Instead, the monopolist sets moderately high prices in both markets ($325 for men, $367 for women). These are still substantially higher than competitive family-bundling prices, but because an entire market is no longer shut out, welfare rises relative to mixed bundling.

This has general implications for imperfect competition in selection markets with multiple products. Imperfect competition and the uninternalized impact of cream-skimming are both market failures that reduce welfare. As often seen with multiple market failures, correcting one need not increase welfare if it worsens the other. As the market structure moves from competitive to monopolistic, the distortion from imperfect competition increases, while the distortion from cream-skimming decreases. Which effect dominates is an empirical question. Quantifying these welfare tradeoffs with richer models of imperfect competition is an important avenue for future research.

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 who buy both products are cheaper than those who buy one. Empirically, I study LTCI. I offer evidence consistent with the 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, while compulsory spousal purchase would lower it. In simulations, I

show that imperfect competition reverses these policy prescriptions.

The theoretical model and computational tools developed here can be applied to a wide variety of selection and insurance markets. For example, this framework could analyze: the effect of pervasive multi-line discounts for home and auto insurance; the combination of health, dental, and vision plans by some employers; bundled long-term care and life insurance products; credit cards that include travel insurance, purchase protection, and rental car coverage. The insight from this paper – that bundling can act as a screening device and impose an externality on the unbundled markets – can be combined with richer models of product design and imperfect competition to deliver welfare evaluations and policy prescriptions tailored to those settings.

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. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics 84*(3), 488–500.

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.

Fang, H. and Z. Wu (2018). Multidimensional private information, market structure, and insurance markets. *The RAND Journal of Economics 49*(3), 751–787.

Finkelstein, A., N. Hendren, and M. Shepard (2019, April). Subsidizing health insurance for low-income adults: Evidence from massachusetts. *American Economic Review 109*(4), 1530–67.

Finkelstein, A. and K. McGarry (2006). Multiple dimensions of private informa-

tion: evidence from the long-term care insurance market. *American Economic Review 96*(4), 938–958.

Fluet, C. and F. Pannequin (1997). Complete versus incomplete insurance contracts under adverse selection with multiple risks. *Geneva Risk Insur Rev 22*, 81–101.

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.

HRS (2018). Health retirement study produced and distributed by the university of michigan with funding from the national institute on aging (grant number nia u01ag009740).

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.

Murtaugh, C. M., B. C. Spillman, and M. J. Warshawsky (2001). In sickness and in health: An annuity approach to financing long-term care and retirement income. *The Journal of Risk and Insurance 68*(2), 225–253.

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. (2024, May). Household bundling of 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.

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.

A Proofs of Main Results

A.1 Proof of Lemma 1

Proof. The separate-market clearing prices solve $p_1 = AC_1(p_1; p_2)$ and $p_2 = AC_2(p_2; p_1)$. Since bundling is not allowed, the demand and therefore cost functions depend only on the marginal distributions of each risk. By definition, all joint distributions $X, Y \in \Gamma(F_1, F_2)$ have the same marginal distributions. Hence, the prices that clear the separate markets are identical. ■

A.2 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 of 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 breaking even 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 ($Y \succsim_{\bar{p}_1} X, Y \succsim_{\bar{p}_2} X$) as defined in Definition 9.

We then have, first for adverse selection:

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

where the inequality follows from $\phi_1'(w_1) > 0$ (adverse selection) and Definition 9, which is the definition of the (truncated) correlation order in this proposition. Under advantageous selection the inequality is reversed.

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. ■

A.3 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 with the reverse inequality when we condition on $w_2 \geq p_B - p_1^{MB}$ instead of $w_2 \leq p_B - p_1^{MB}$.

If Y is independent, we have that $\mathbb{E}_Y\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \geq p_B - p_1^{MB}\right] = \mathbb{E}_Y\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB} \wedge w_2 \leq p_B - p_1^{MB}\right]$

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 definition, we have that $\mathbb{E}_X\left[\phi_1(w_1) \mid w_1 \geq p_1^{MB}\right] = p_1^{MB}$. Hence, 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. I define three groups. Group α consists of those that buy product one (or both products) under both mixed bundling and banned bundling. Group β consists of those those that buy product one under banned bundling but nothing under mixed bundling. Group γ consists of those that buy only product 2 under mixed bundling but both under banned bundling. In sum, under tje 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. ■

A.4 Proof of Proposition 3

Proof. Consider X and Y with $Y \gtrsim X$. **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,

$$P_X\left(W_1 > w_1 \mid W_2 \geq p_B - w_1\right) \leq P_Y\left(W_1 > w_1 \mid W_2 \geq p_B - w_1\right) \tag{A.1}$$

For brevity I write this set as $\mathcal{W}_{\backslash p_B} = \{w \in \mathcal{W} : W_B \geq p_B\}$. Hence,

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

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

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

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, the demand curve for the bundle under X is above that for Y for all $q \geq \bar{q}$. 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. ■

A.5 Proof of Proposition 4

Proof. Consider the following groups. Group A: consumers who buy the bundle under mixed bundling. Group B: buy product 1 under mixed, bundle under forced. Group

C: buy product 1 under mixed, nothing under forced. Group D: buy product 2 under mixed, bundle under forced. Group E: buy product 2 under mixed, nothing under forced.

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

$$0 = \alpha_A p_B - \alpha_A AC_A^1 - \alpha_A AC_A^2 \tag{A.2}$$

$$0 = \alpha_{BC} p_1 - \alpha_{BC} AC_{BC}^1 \tag{A.3}$$

$$0 = \alpha_{DE} p_2 - \alpha_{DE} AC_{DE}^2. \tag{A.4}$$

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:

$$\begin{aligned}
\pi &= \alpha_A p_B - \alpha_A AC_A^1 - \alpha_A AC_A^2 + \alpha_{BD} p_B - \alpha_{BD} AC_{BD}^1 - \alpha_{BD} AC_{BD}^2 \\
&= \alpha_{BD} p_B - \alpha_{BD} AC_{BD}^1 - \alpha_{BD} AC_{BD}^2 && \text{by (A.2)} \\
&< \alpha_{BD}(p_1 + p_2) - \alpha_{BD} AC_{BD}^1 - \alpha_{BD} AC_{BD}^2 && (\text{since } p_B < p_1 + p_2) \\
&< \alpha_{BC} p_1 + \alpha_{DE} p_2 - \alpha_{BC} AC_{BC}^1 - \alpha_{DE} AC_{DE}^2 && \text{by the arguments above} \\
&= 0 && \text{by (A.3) \& (A.4).}
\end{aligned}$$

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.

■

For Online Publication Only

B Theoretical Appendix

B.1 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.[19] 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:

[19] 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.

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

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

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 contract one to not depend on whether contract 2 is bought or not is that both risks are small in the following sense.

B.2 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 7. *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 differentiate 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 8. *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.

B.3 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:

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

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

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

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

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 9. *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*

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

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 1. *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*

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

C Empirical Appendix

C.1 Ruling out 'standard' cost efficiencies

In this section, I provide more detail on the other insurance products considered in Section 3.7: 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. Life insurance data come from Compulife, software that collects historical quotes from more than 100 life insurance issuers. Data for hybrid long-term care policies come from ComboCompare (StrateCision (2022a)) which quotes six 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 dies. 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, -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 4 in the main body of the paper.

C.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

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

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.

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

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.

C.3 Correlation in Outcomes versus Propensities

The main paper studies correlation in ex-ante propensities between couples. This is related to, but distinct from, correlation in outcomes. Intuitively the difference is between correlation in marginal probabilities versus the full joint distribution. Both spouses can have high marginal probabilities (which sums over outcomes where the other spouse does and doesn't require LTC) while the probability of one spouse requiring care conditional on the other requiring care can be low or high.

Consider Table 4:

Spouse 1 → Spouse 2 ↓	No LTC	LTC	
No LTC	p_0	p_1	$1-p_2^M$
LTC	p_2	p_B	p_2^M
	$1-p_1^M$	p_1^M	

Table 4: The full joint distribution of outcomes.

The probability of no spouse needing LTC is p_0, of spouse 1 (2) needing care but the other spouse not p_1 (p_2) and the probability of both spouses requiring care p_B. This leads to marginal probabilities of spouse 1 requiring care equal to $p_1^M = p_1 + p_B$ and similarly for spouse 2.

The correlations discussed in the main paper are between marginal probabilities p_1^M and p_2^M. For the purposes of costs, only these matter. In section 3, I found strong positive correlation between p_1^M and p_2^M.

Correlations in outcomes, on the other hand, are about the joint probabilities p_0, p_1, p_2, p_B. These are inputs into the marginal probabilities, and hence correlation in outcomes is a possible mechanism for correlation in propensities. Positively correlated outcomes would occur when p_0 and p_B are relatively high, and p_1 and p_2 are relatively low. Note, this doesn't imply anything directly about the marginal probabilities.

Positive correlation in propensities (marginal probabilities) means that p_1^M and p_2^M are high together or low together, but this can be beecause p_1 and p_2 are high together or low together, or because p_B/p_0 is high or low for both.

To empirically study this, I estimate a series of linear probability models of the form:

$$\text{Own Usage of LTC} = \beta_1 \text{Spousal Usage of LTC} + \beta_2 \text{Own Ex-ante Predicted Risk} + \epsilon. \tag{C.2}$$

I include different predictions of own risk: either the full prediction on the rich model, or the restricted prediction using only insurer observed variables, or both. The results are in Table 5.

	(1) Base Model	(2) Controlling for Own Risk	(3) Controlling for Insurer Prediction of Own Risk	(4) (1) and (2)
Spousal Usage of LTC	0.113***	0.049**	0.115***	0.050**
	(0.017)	(0.015)	(0.016)	(0.015)
Predicted Own Risk		1.036***		1.018***
		(0.025)		(0.037)
Predicted Own Risk (Insurer Variables Only)			1.062***	0.034
			(0.037)	(0.051)
N	5148	5148	5148	5148
R2	0.004	0.296	0.180	0.296

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

Table 5: Testing for correlation in outcomes. The estimating equation is (C.2).

Looking at the estimate for β_1 (the first row), I find a strong positive correlation in outcomes. Conditional on a person's ex-ante risk, if a person's spouse requires LTC, the probability of that person requiring LTC themselves increases by 5% to 12%, depending on the specification.

This is a mechanism for the correlation in ex-ante risk observed in the primary analysis. The marginal probabilities remain the sufficient statistics for the analysis performed there.

C.4 Welfare Robustness

My baseline assumption for the cost of LTC care, conditional on it being needed, was approximately \$150,000 (\$66,237 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 needed (i.e. I vary the \$150,000). Given the linearity of expected cost, this proxies for changing the discounting assumptions. Moreover, since there is a scale invariance between WTP and costs, checking for robustness to costs is structurally similar to checking for robustness to WTP.

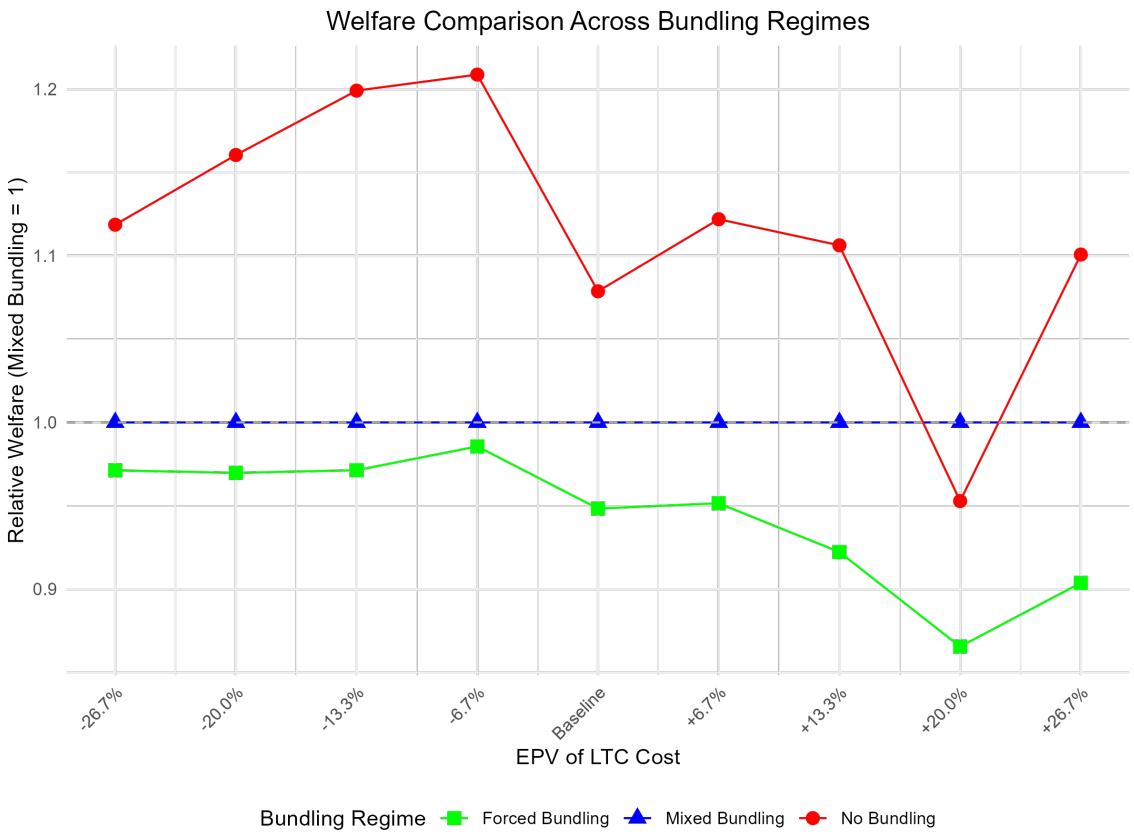


Figure 7: Welfare in the three regimes (mixed bundling, forced family bundling and no bundling allowed) under different cost assumptions. For each cost assumption, welfare is expressed relative to the mixed bundling status quo.

I vary the expected cost of care by approximately 25% up and down from the baseline. Personalized expected costs are then computed as before, by multiplying each individual's probability of needing care, as estimated in Section 3.4, 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 7 below. For each cost assumption, I express welfare relative to the mixed bundling status quo.

The welfare results are largely robust to different cost assumptions. Banning bundling almost always increases welfare, whereas forcing family bundling always decreases it. Hence, the robust qualitative takeaway is that the status quo private market equilibrium that features mixed bundling is not optimal, and government interventions that ban bundling would likely improve welfare in the market.

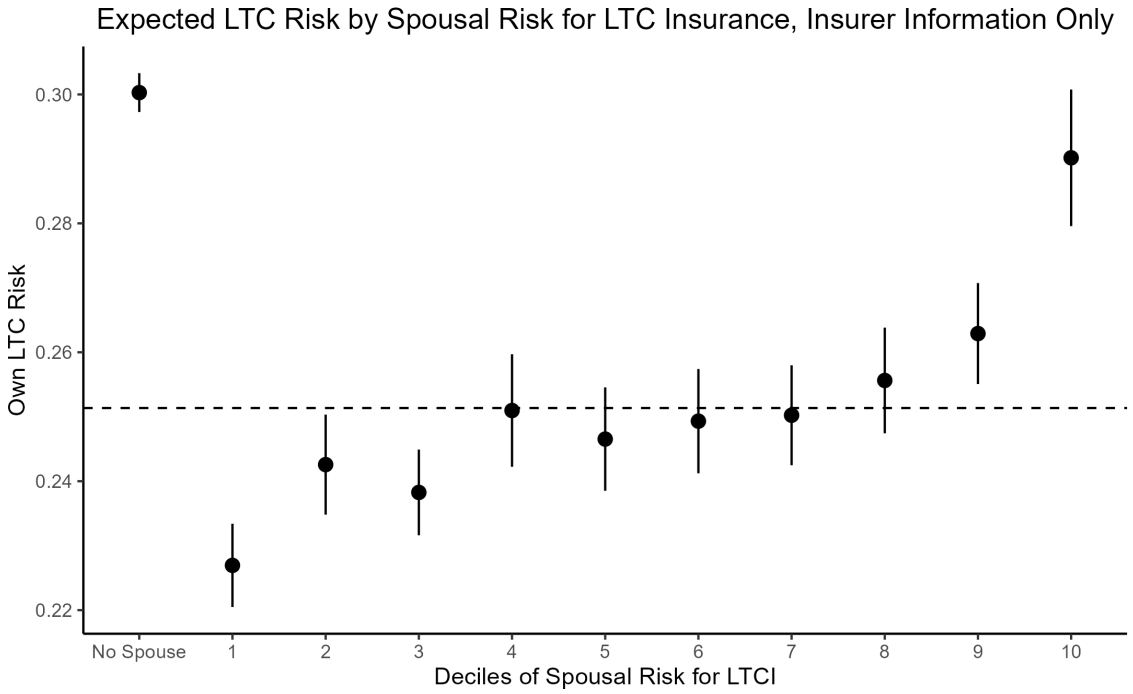


Figure 8: Binscatter of expected LTC risk that is predictable by insurers, by deciles of spousal WTP. The horizontal dashed line is the average risk amongst all those who have a spouse.

C.5 The Predictive Content of WTP Relative to Information Collected by Insurers

Figure 8 below is an analogue of Figure 3, but where the predictions of LTC risk use only the smaller set of information available to insurers.[(20)] 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 3 but not 8.

The slope of observable risk with respect to spousal WTP (Figure 8) is much shallower than that of total (observable and unobservable) risk with respect to spousal WTP (Figure 3). 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,

[(20)] 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)

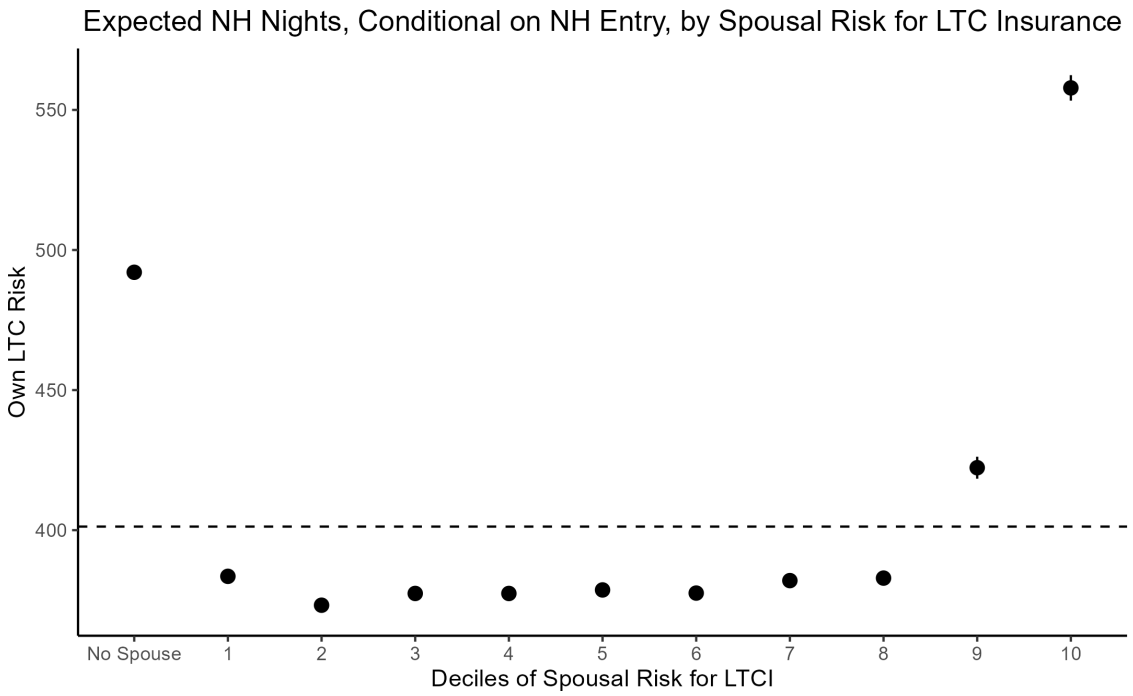


Figure 9: Binscatter of expected number of nights in a nursing home, conditional on entry, by deciles of spousal WTP. The horizontal dashed line is the average risk amongst all those who have a spouse.

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.

C.6 Intensive Margin of LTC Utilization

Figure 9 is an analogue of figure 3 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.

C.7 Details of LTC Risk Prediction

The following variables are included in the prediction of LTC risk: the individual's age, marital status, mother's survival status, father's survival status, mother's age, father's age, self-assessed health status, history of hospitalization, residency in a nursing home, number of doctor visits, receipt of home care, out-of-pocket medical expenses, assistance with personal care activities, assistance with household tasks, assistance with mobility, assistance with activities requiring muscle strength, assistance with major physical activities, assistance with delicate or precise activities, feelings of depression, past happiness, alcohol consumption frequency, smoking history, current smoking status, high blood pressure, diabetes, history of cancer, lung disease, heart disease, stroke history, arthritis, current memory status, past memory status, government health insurance, number of private health insurance plans, life insurance coverage, number of children, retirement status, past high blood pressure, past diabetes, past cancer, past lung disease, past heart disease, past stroke, past arthritis, body mass index, gender, asset prediction less than $10,000, asset prediction less than $100,000, income group classification, participation in a training program, occupational health facility usage in the past six years, data collection wave, and interaction of sex with diabetes, cancer, lung disease, heart disease, stroke, arthritis, ADLA, IADLA, BMI, and being married.

C.8 Estimation Details

A model for individual demand and WTP for LTCI is estimated in parallel to the model for couples described in Section 5. All individuals are included here: whether single or married.

Choice set. An individual chooses from the two-element set $\mathcal{J}^{\text{single}} = \{\text{none}, s\}$, where *none* denotes self–insurance and s denotes purchasing a stand–alone long-term-care insurance (LTCI) policy.[(21)]

Indirect utility. Let p be the quoted premium for an LTCI contract and let the sub-script $k \in \{\text{m}, \text{f}\}$ index the individual's sex. The latent indirect utility from

[(21)] Throughout, "singles" are HRS respondents who report being widowed, divorced, separated, or never married.

purchasing the policy is

$$U_s \;=\; \beta_1 p + \beta_2 a + \beta_3\,\texttt{married} + \boldsymbol{\beta}^{\text{inc}}\cdot\mathbf{inc} + \boldsymbol{\beta}^{\text{risk}}\cdot\mathbf{risk} + \mathbf{1}\{k=m\}\beta_{\text{sex}}, \qquad \text{(C.3)}$$

where a is age, **inc** and **risk** are the same income-group and risk dummies used in (5.1), and $\texttt{married} = 0$ for every observation in the singles sample. Premiums for non-purchasers are imputed with a pricing model that regresses observed premia on age, sex, (lagged) wave dummies, and coarse risk groups,[(22)] so that every individual faces a counterfactual quote p.

Utility from *none* is normalised to 0. Assuming i.i.d. type-I extreme-value errors, the purchase probability is

$$P_s \;=\; \frac{\exp(U_s)}{1+\exp(U_s)}, \qquad P_{\text{none}} = 1 - P_s. \qquad \text{(C.4)}$$

Identification and estimation. As in the married-household model, identification of the price coefficient β_1 relies on the sharp, regulator-approved premium hikes that occurred between 1996 and 2006. Formally, I treat wave dummies as excluded instruments in a two-stage procedure:

1. Regress p on wave and demographic covariates to obtain the fitted premium $\hat{p}$;
2. Estimate (C.4) by maximum likelihood after replacing p with $\hat{p}$.

The parameter vector is $\boldsymbol{\theta}^{\text{single}} \;=\; \big(\beta_1, \beta_2, \beta_3, \boldsymbol{\beta}^{\text{inc}}, \boldsymbol{\beta}^{\text{risk}}, \beta_{\text{sex}}\big)$ and the log-likelihood is $\widehat{\boldsymbol{\theta}}^{\text{single}} \;=\; \arg\max_{\boldsymbol{\theta}} \sum_i \Big[\mathbf{1}\{j_i = s\}\log P_{s,i}(\boldsymbol{\theta}) + \mathbf{1}\{j_i = \text{none}\}\log\big(1 - P_{s,i}(\boldsymbol{\theta})\big)\Big]$.

Recovering individual WTP. With $\widehat{\boldsymbol{\theta}}^{\text{single}}$ in hand, I recover each individual's willingness-to-pay w_s by finding the premium that makes the model to match the observed aggregate coverage rate $\bar{P}_{\text{data}}$:

$$w_s \;=\; \arg\min \Big| P_s(p; \widehat{\boldsymbol{\theta}}^{\text{single}}) - \bar{P}_{\text{data}} \Big|. \qquad \text{(C.5)}$$

Operationally, I solve (C.5) with a one-dimensional root-finding algorithm on the support $p \in [0, \bar{p} \times 5]$, where $\bar{p}$ is the maximum observed quote. The procedure yields an

[(22)] Because LTCI prices are based only on sex, age, marital status, and coarse risk categories, a simple linear model fits the premium schedule well. Approximately 11% of singles hold LTCI in the estimation sample, providing sufficient variation to estimate the pricing function credibly.

average WTP of roughly $155, which—after converting to a present-discounted value using remaining life expectancy at age 65 and a 5% real interest rate—corresponds to an average willingness to pay of about $26,000. Consistent with the evidence for married households, I find mild advantageous selection: individuals with lower estimated nursing-home risk have, on average, higher w_s.

Reconciling the singles and couples models. The procedure described above generates a WTP for long-term care insurance for all individuals, whether married or not. But, per Section 5, I want married couples demand to allow for complementarity or substitutability within the household. Hence, we merge in the estimated $\overline{\gamma}$ as estimated by the procedure in Section 5. So individual WTPs (for single *and* married people) are estimated per this section, and the complementarity/substitutability term $\overline{\gamma}$ is estimated per Section 5.

References for Appendices

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.