Selection through Lapsation in Life Insurance Markets*

Cameron M. Ellis† & Johannes G. Jaspersen‡

March 20, 2025

Abstract

It is commonly believed that life insurance features advantageous rather than adverse selection, primarily due to evidence that policyholders have lower mortality rates than non-policyholders. We introduce lapse rates as an alternative channel for selection in life insurance. Theoretically, we demonstrate that lapsing can induce adverse selection even among naïve consumers. Empirically, we validate this theory using administrative data and price instruments derived from cost shocks in other markets. Our results confirm that lapse and surrender rates respond significantly to price changes, revealing adverse selection through a downward-sloping cost curve. However, no comparable effect emerges in mortality rates. We find that this adverse selection reduces overall demand by 16.6%, resulting in welfare losses approaching $1.3 billion annually. Furthermore, this form of selection exacerbates inequity, disproportionately benefiting wealthy consumers.

JEL Codes: **D91** · **G22** · **G52**

Keywords: Adverse Selection · Insurance Demand · Cost Function · Insurance Markets

*Acknowledgments: We thank Daniel Bauer, Benjamin Collier, Keith Ericson, Shan Ge, Daniel Gottlieb, Elisabeth Stöckl and Justin Sydnor as well as seminar participants at the University of Iowa and the MRIC/CEAR Behavioral Insurance Workshop for helpful comments. Barbara Reißer provided excellent research assistance. Ellis received financial support from the National Association of Insurance Commissioners (NAIC) through the research fellows program.

†Tippie College of Business, University of Iowa (cameron-ellis@uiowa.edu)

‡LMU Munich School of Management, LMU Munich (jaspersen@lmu.de)

1 Introduction

Life insurance protects households from financial hardship upon the death of a primary wage earner. Despite its importance, market coverage has declined, and only 52% of US households owned a policy in 2021, down from 63% in 2011 (LIMRA, 2021). This trend is concerning, as 44% of households would face financial difficulties within 6 months of the main breadwinner's death (Crail, 2023). But why is demand so low? Adverse selection is often cited as a reason for low insurance take-up in other markets such as annuities or health (Finkelstein and Poterba, 2004; Einav et al., 2010). However, it is not conventionally seen as an explanation for lacking demand in life insurance markets because the correlation between life insurance coverage and policyholder mortality tends to be positive (Cohen and Siegelman, 2010). Yet, such tests do not consider all drivers of life insurance premiums. Whether and when consumers lapse their insurance coverage influences both their own utility and the insurer's profit (Gottlieb and Smetters, 2021). Because lapse risk is not individually underwritten, it thus presents an alternative channel for adverse selection which has not been examined.

This paper fills that gap. We develop a model in the framework of Einav et al. (2010) and demonstrate that mortality and lapsation are likely to have opposing effects on the degree of selection in the life insurance market, regardless of the level of sophistication households have about their own risk. We then use data from US life insurance companies' statutory filings to estimate the degree of selection in the individual life insurance market. In contrast to previous studies, we find a strong degree of selection with a 1% increase in premiums leading to a 0.13% increase in average future costs. We find that this selection is driven entirely by a reduction in future lapse and surrender rates with a tightly estimated null effect on future mortality. Using our causal estimates, we calculate that adverse selection decreases overall life insurance demand by 16.6%, leading to an annual welfare loss of $1.3 billion. We then estimate a structural model of life insurance demand using both data from life insurers and from the Health and Retirement Survey to analyze the equity consequences of the adverse selection. We find that it is regressive – welfare losses are concentrated on low-wealth households.

Life insurance premiums have two primary drivers: mortality risk and lapse risk. Higher mortality increases an individual's expected value of the insurance contract, because payouts are more likely. On the other hand, a higher risk to lapse decreases the expected value (Gottlieb and Smetters, 2021). Since the policyholder's mortality risk increases with age and premiums are typically flat, premiums tend to be higher than the actuarial value early in the policy and lower than the actuarial value later in the policy. If policyholders lapse before reaching the later stages of a policy they will have "overpaid" the entire time. While selection through mortality risk has been an active topic of the empirical literature[1], selection through lapsation has not yet been considered. This is problematic, because lapsation is a very common event. Gottlieb and Smetters (2021) report that between 1991 and 2010, $24 trillion of $29.7 trillion in new individual life insurance coverage issued in the United States lapsed. Similarly, analysis of our own data indicates that a policy is more than twice as likely to lapse than to cease for any other reason.[2]

Results of advantageous selection in life insurance are typically explained by a correlation of both demand and mortality risk with a third factor (as has been documented in other markets, see Fang et al., 2008; Cutler et al., 2008). Wealth, income, or financial sophistication are likely candidates as they are correlated with both mortality and life insurance demand (e.g. Chetty et al., 2016; Gropper and Kuhnen, 2025). At the same time, such factors also influence the decisions to lapse (e.g. Fang and Kung, 2021).[3] To explore the complex dynamics of selection in the life insurance market, we develop a theoretical model that incorporates both mortality risk and lapse behavior and a third factor, we call consumer type. Our model builds upon and extends the framework of Einav et al. (2010) to account for the unique features of life insurance contracts, particularly their long-term nature and the possibility of lapsation. Consumer type proxies for

[1]Results are conflicting with a tendency towards advantageous selection where life insurance holders tend to have lower mortality risk (e.g., Cawley and Philipson, 1999; McCarthy and Mitchell, 2010). Nevertheless, some studies provide evidence of adverse selection, where individuals with higher mortality risk are more likely to purchase life insurance (e.g., Giesbert et al., 2011; He, 2009). Still, some studies fail to find clear evidence of either adverse or advantageous selection (e.g., Harris and Yelowitz, 2014; Wuppermann, 2017), while others find both adverse selection and advantageous selection at different points on the health distribution (Ai and Steinorth, 2024).

[2]Of all policies ceasing in a year, ca. 68% are due to lapsation, 15% to surrender, 12% to death, and 5% to expiration.

[3]The literature identifies several reasons for lapsing, including preference shocks (Fang and Kung, 2010; Liebenberg et al., 2012; Fei et al., 2015), income shocks (Linton, 1932; Outreville, 1990; Kuo et al., 2003), policy replacement (Outreville, 1990; Carson and Forster, 2000), changes in interest rates (Schott, 1971; Pesando, 1974; Kuo et al., 2003), and non-expected utility explanations(Shefrin, 2002; Mulholland and Finke, 2014; Gottlieb and Smetters, 2021). For a comprehensive review, see Eling and Kochanski (2013).

wealth and financial sophistication and influences income, mortality risk, and lapse probability. Informed by evidence from the aforementioned empirical literature on the marginal effects of wealth on these factors this allows us to analyze how the different forces interact to determine market outcomes in a two-period setting.

Our theory model yields two key insights. First, we demonstrate that both adverse and advantageous selection can emerge in the life insurance market, depending on the relative strength of opposing effects. "High-type" individuals (with higher wealth) tend to have lower mortality risk but also lower lapse probabilities. These factors have countervailing effects on insurers' costs and consumers' willingness to pay for insurance. Consequently, the direction of selection in the market is theoretically ambiguous and depends on which effect dominates. This highlights the potentially critical role of lapse behavior in shaping market equilibrium, a factor often overlooked in previous studies focusing solely on mortality risk. Second, we show that selection can occur even if consumers are naïve in the sense that they do not know their own risk. In this case, the influence of consumer type is sufficient to create either adverse or advantageous selection, depending on whether the influence of type on costs is stronger through mortality or lapsing. This latter finding is important, because Gottlieb and Smetters (2021) show that consumers are often not aware of their own lapse behavior.

To empirically test our theoretical predictions and quantify the extent of selection in the life insurance market, we leverage a comprehensive dataset compiled from the National Association of Insurance Commissioners (NAIC) regulatory reports. This dataset spans from 1993 to 2017 and includes both national-level and state-specific data, providing a detailed view of the U.S. life insurance market across time and geography. To address potential endogeneity concerns in our empirical analysis, we follow an extensive recent tradition of instrumenting insurance prices with prices/costs by the same firm in other markets (e.g., Koijen and Yogo, 2016, 2015; Ge, 2022; Gao et al., 2023; Wenning, 2024). Specifically, we use a firm's group-market life insurance premiums in a year as an instrument for their individual-market life insurance premiums in the same year. We combine this instrument with extensive fixed effects to ensure that our identifying variation only comes from firm-year level cost shocks in the group market spilling over into individual market

pricing.

Using our instrument, we estimate the causal impact of a change in premiums on (1) average future costs, (2) contemporaneous demand, (3) future lapse rates, and (4) future death rates. Our findings provide strong support for the presence of selection in the life insurance market, while also revealing the complex nature of this selection. Notably, we find evidence of adverse selection, with a 1% increase in premiums associated with a 0.13% increase in average future costs for insurers. However, our results also highlight the crucial role of lapse behavior in driving this selection. Higher premiums are associated with lower lapse rates, suggesting that policyholders who maintain coverage at higher premium levels may have different risk characteristics or valuation of insurance compared to those who lapse. We also are able to tightly estimate a null effect of premiums on death rates, indicating that mortality risk is not the primary channel through which selection is occurring in this market. This finding underscores the importance of considering lapse behavior, rather than focusing solely on mortality risk, when analyzing selection in life insurance markets. Our results thus provide empirical validation for our theoretical model and offer new insights into the mechanisms driving selection in this crucial financial market.

We use our empirical results to estimate the efficiency losses due to lapse-based adverse selection in the US life insurance market. On average, annual new life insurance business is 16.6% smaller in equilibrium than in the efficient market outcome. This implies an annual net present value for the dead-weight loss of $1.29 billion for the companies in our estimation sample. We provide confidence intervals for this estimate based on a block-bootstrap procedure. In addition, we develop a robustness analysis in the spirit of Kang and Vasserman (2022) for the functional form assumptions used in our estimation procedure. It indicates that our estimates are conservative relative to the spectrum of possible functional forms of the demand and cost curve. The new robustness procedure can be seen as a contribution to the literature by itself and can be applied to any analysis of welfare losses in Akerlovian selection markets (such as Einav et al., 2010).

Given the documented correlation between life insurance demand and wealth and lapse behavior and wealth, our results raise concerns about the effects they have on economic equity in the population. To analyze the equity consequences of the documented selection, we also estimate

a Berry et al. (1995) (BLP) style demand model which we combine with micro-moments on purchase and lapse behavior by wealth from the Health and Retirement Survey. After estimating the model, we examine the distributional impacts of eliminating the selection through individually-underwritten lapse risk. We find that, in addition to the overall reduction in deadweight-loss from the removal of selection, there would also be a transfer from the top half of the wealth distribution to the bottom half, with the lowest ventile of wealth receiving the largest gain.

We make three primary contributions. Our first contribution is to provide a comprehensive theoretical and empirical analysis of selection in the US life insurance market, considering both mortality risk and lapse behavior. Our theoretical model demonstrates that adverse and advantageous selection can coexist within the same market, depending on the relative strength of opposing effects related to consumer type, mortality risk, and lapse probabilities. This helps explain divergent results in previous studies (e.g., Cawley and Philipson, 1999; McCarthy and Mitchell, 2010; Giesbert et al., 2011; He, 2009) and calls for a more nuanced approach to modeling selection in insurance markets.

Our second contribution is to empirically demonstrate the pivotal role of lapse behavior in influencing adverse selection in life insurance markets. A growing body of literature examines what is termed "dynamic" adverse selection, where the decision to lapse a policy is correlated with the insured's mortality risk (e.g., He, 2011; Lai et al., 2021; Ai and Steinorth, 2024). We add to this literature by showing that adverse selection in life insurance is also driven by the correlation of demand and ex ante lapse probabilities, rather that ex post lapse behavior and mortality. We also show that knowledge of lapse rates is not necessary for this selection channel to occur (Carson et al., 2020; Gottlieb and Smetters, 2021). Our results highlight the critical role of lapse behavior in shaping the risk pool and challenge prior models that focus predominantly on mortality risk as the key driver of adverse selection.

Finally, our estimate of a single company's life insurance demand elasticity is informative in its own right. In a perfectly competitive market, an exogenous price increase by one firm should lead to a strong decline in the demand for products of that firm and a corresponding increase in demand for the products of other firms. However, while we do see decreasing demand, it is

slightly inelastic. This implies significant frictions on life insurance markets. The nature of these frictions is unclear and demands further analysis. Previous evidence from other insurance markets shows that the role of distribution channels and advertising could be significant (Aizawa and Kim, 2018; Collier and Ragin, 2020), but there is also conflicting evidence for the latter (Shapiro, 2020). The magnitude of these (and potential other) frictions are as of yet not explored for life insurance markets, but our research implies that they do exist.

Our findings have significant policy implications, particularly regarding the pricing of life insurance policies. There are two potential solutions to lapse-driven selection: incorporating individual-based lapse estimates into pricing models or removing flat premiums and instead have premiums reflect contemporaneous mortality risk. Both of these changes would benefit low-income consumers, making life insurance more accessible while reducing subsidies for high-income groups who maintain their policies long-term. Moreover, these policies would reshape the secondary life insurance market, with individual lapse-based pricing creating new arbitrage opportunities for life settlements, while contemporaneous-risk premiums would reduce such dynamics, especially within stranger-originated life insurance (STOLI) schemes.

2 An Illustrative Model of Selection in Life Insurance

In the following, we are going to introduce an illustrative model of how both adverse and advantageous selection can appear in life insurance markets. For this illustration, we first need to establish the microeconomic foundations of insurance demand the producers' incurred costs. We then use these foundations to derive potential market outcomes when consumer types can be summarized by a single (scalar-valued) type, which we denote ξ. As a robustness check, we also analyze the case in which consumers are naïve in the sense that they are not aware of their own lapse probability or mortality.

2.1 Individual Level Insurance Demand and Cost

As long as they are alive, consumers derive utility from consumption according to the utility function $u_t(x)$ in two periods $t \in \{1, 2\}$.[4] In both periods, u is twice continuously differentiable with $u_t' > 0$ and $u_t'' < 0$. If alive, consumers earn income I in each period. Consumers can die after the first period with probability ρ^d. Dead consumers do non generate income, but still obtain utility from bequest motives. Thus, they can purchase a life insurance contract in the first period for a premium P to be paid in both periods. There is no saving in the model such that consumers consume all income they have left over after paying for the insurance policy. Savings would unnecessarily complicate the model without adding much information because it is a bad substitute for the contingent payment streams of life insurance.

If the insurance policy is not lapsed, consumers obtain a payment B. Utility from money in death is denoted $v_2(\kappa + x)$ with properties $v_2' > 0$ and $v_2'' < 0$ similar to utility from consumption when alive. The utility function includes a luxury good factor κ. This factor and concavity of v_2 ensure that if an individual could freely allocate x between consumption when alive and bequest when dead, there is a critical level of x below which the individual would choose to consumer 100% of the amount (see Lockwood, 2018, for a similar model set-up). The model is thus in line with the prevailing empirical evidence that bequests only appear in households above a certain level of wealth (Ameriks et al., 2011). Consumers have a chance ρ^l of lapsing the insurance policy. In this case, they do not pay a premium in the second period, but also do not obtain a payment in case of death. The expected utility of consumer i with parameters $\{I_i, \rho_i^l, \rho_i^d\}$ when buying a life insurance contract (indicated LI) for premium P is thus given by

$$EU^{LI}(I_i, \rho_i^l, \rho_i^d, P) = u_1(I_i - P) + (1 - \rho_i^l) \underbrace{\left[(1 - \rho_i^d)u_2(I_i - P) + \rho_i^d v_2(\kappa + B - P)\right]}_{\text{Expected utility without lapsing}}$$
$$+ \rho_i^l \underbrace{\left[(1 - \rho_i^d)u_2(I_i) + \rho_i^d v_2(\kappa)\right]}_{\text{Expected utility when lapsed}}. \quad (1)$$

[4]Note that by time-indexing the utility function, we subsume all additive preference models, including the commonly used additive exponential discounting model.

How do the characteristics of the consumer impact demand when price is fixed, that is, when there is asymmetric information about the characteristics in the market? For this, we first need to define the maximum premium consumer i is willing to pay for life insurance, $\hat{P}_i$. Following the assumption that the outside option is not to purchase insurance and thus to obtain expected utility $EU^0(I_i, \rho_i^d) = u_1(I_i) + (1-\rho_i^d)u_2(I_i) + \rho_i^d v_2(\kappa)$, we can define this maximum acceptable premium through

$$EU^{LI}(I_i, \rho_i^l, \rho_i^d, \hat{P}_i) - EU^0(I_i, \rho_i^d) = 0. \tag{2}$$

We can then see the influence of the relevant characteristics on $\hat{P}_i$ from the implicit function theorem. Specifically, we see that[5]

$$\frac{\partial \hat{P}_i}{\partial I_i} = \frac{u_1'(I_i - \hat{P}_i) - u_1'(I_i) + (1-\rho_i^d)(1-\rho_i^l)\left[u_2'(I_i - \hat{P}_i) - u_2'(I_i)\right]}{u_1'(I_i) + (1-\rho_i^l)\left[(1-\rho_i^d)u_2'(I_i - \hat{P}_i) + \rho_i^d v_2'(\kappa + B - \hat{P}_i)\right]} > 0, \tag{3}$$

$$\frac{\partial \hat{P}_i}{\partial \rho_i^d} = \frac{(1-\rho_i^l)\left[u_2(I_i) - u_2(I_i - \hat{P}_i) + v_2(\kappa + B - \hat{P}_i) - v_2(\kappa)\right]}{u_1'(I_i) + (1-\rho_i^l)\left[(1-\rho_i^d)u_2'(I_i - \hat{P}_i) + \rho_i^d v_2'(\kappa + B - \hat{P}_i)\right]} > 0, \tag{4}$$

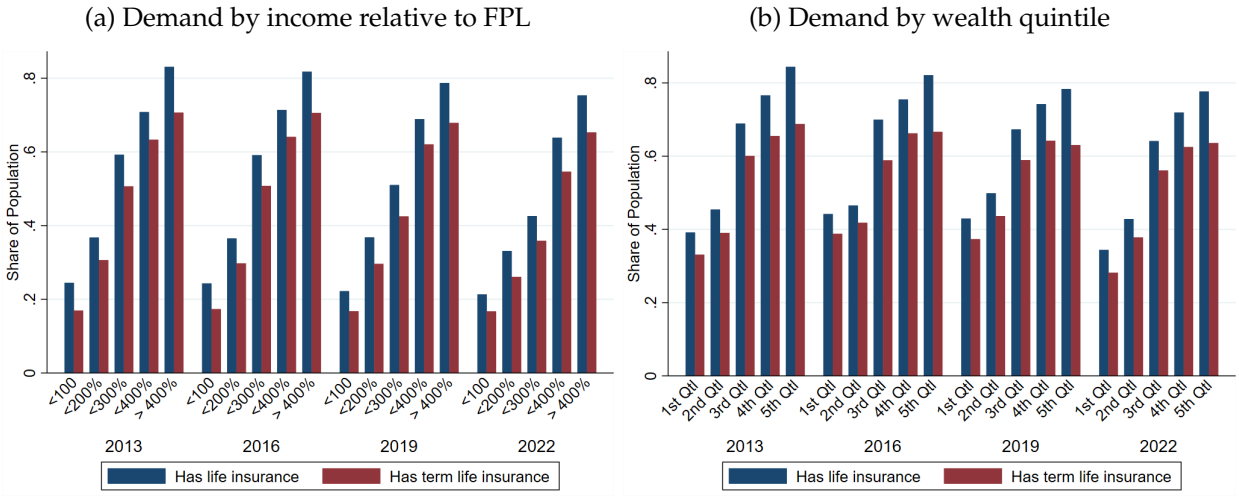
$$\frac{\partial \hat{P}_i}{\partial \rho_i^l} = \frac{(1-\rho_i^d)\left[u_2(I_i) - u_2(I_i - \hat{P}_i)\right] + \rho_i^d\left[v_2(\kappa) - v_2(\kappa + B - \hat{P}_i)\right]}{u_1'(I_i) + (1-\rho_i^l)\left[(1-\rho_i^d)u_2'(I_i - \hat{P}_i) + \rho_i^d v_2'(\kappa + B - \hat{P}_i)\right]} < 0. \tag{5}$$

The positive influence of income I_i stems from the concavity of the utility function. As the income of the consumer increases, the marginal cost of the premium decreases, while the benefit stays constant. This effect is in line with empirical evidence, as can be seen in Panel (a) of Figure 1 that shows the effect of income on (term-) life insurance demand using data from the Survey of Consumer Finances. We can clearly see a strong relationship between income and insurance demand, as well as the declining trend of life insurance ownership over time that was mentioned in the introduction. This mirrors results from the literature that wealthier and higher income

[5]The inequality in (3) follows from concavity of u. The inequality in (5) follows from rearranging Equation (2), realizing that $u_1(I_i - \hat{P}_i) - u_1(I_i) < 0$ and dividing both sides by $-(1-\rho_i^l)$. The inequality in (4) follows from $B > \hat{P}_i$ which has to hold by first order stochastic dominance.

individuals purchase more life insurance (Truett and Truett, 1990; Bernheim, 1991; Chen et al., 2001; Gropper and Kuhnen, 2025).

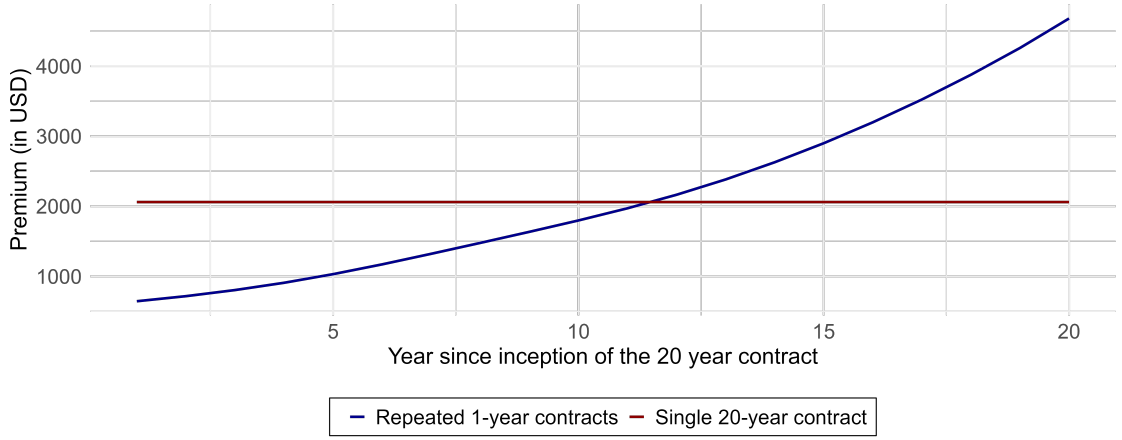
Figure 1: Life Insurance Demand by Income and Wealth



Note: Panel (a) shows the amount of households with 18-64 year-old household heads owning any type of life insurance coverage (blue) or term life (red) in the US. Demand is differentiated by year and income relative to the federal poverty line of the corresponding year. Panel (b) shows the same data, but differentiated by wealth quintile with wealth defined as the sum of all assets minus the sum of all debt. Data is taken from four waves of the Survey of Consumer Finances, see Appendix B for details on the variable construction.

The influences of ρ_i^d and ρ_i^l are intuitive. The maximum willingness to pay increases in ρ_i^d because with increased likelihood to die, the bequest case and thus the payout from the policy becomes more likely and consumers are willing to pay a higher premium for it. If consumers have a higher chance of lapsing and this chance is not reflected in the premium, then they are more likely to pay a premium without having a potential payout from the policy and the maximum acceptable premium declines. To illustrate the potential magnitude of this effect in a real life insurance contract, Figure 2 shows the annual premiums for a 20-year term life insurance policy and those of repeated 1-year policies. We can see that in the beginning of the 20-year contract, policyholders pay more than three times the risk-adequate annual premium. This equalizes with later premiums, beginning in contract year 12. However, any policyholder who cancels early will have paid more than the aggregate 1-year premiums.

Figure 2: Annual payments in 20-year and annual term life contracts



Note: Figure displays the premiums for a 20 year term life insurance contract of a 40-year old German male with a death benefit of $500,000 and the premiums of repeated annual contracts renewed each year. Mortality data is taken from the 2008 DAV German Endowment mortality table.

Models of multiple periods are necessary when considering life insurance demand, especially when analyzing lapse behavior (e.g., Gottlieb and Smetters, 2021). However, using a two-period model requires clarification of the price that is used to determine insurance demand in the EFC equilibrium. It is not equal to the premium P in Equation (1), because the consumers pay this premium multiple times. To clear the disambiguation, we will determine the market equilibrium using Π to denote the price. Consumers always pay the premium in the first period and pay it in the second period, if they do not lapse such that $\Pi_i = P_i(2 - \rho_i^l)$. The corresponding maximum price consumer i is willing to pay is given by $\hat{\Pi}_i = \hat{P}_i(2 - \rho_i^l)$.

We now aim to collapse the three consumer characteristics into a single, scalar-valued type ξ_i. We define ξ as a factor which has a positive influence on income and, at the same time, reflects the other empirical regularities which are relevant for our analysis. ξ can be seen as a combination of wealth, financial literacy and sophistication. This interpretation allows us to find clear indications from the empirical literature, on how ξ influences ρ^d and ρ^l. It is a long-standing result in public health that good financial standing and longevity are correlated positively, such that we can assume $\frac{\partial \rho^d}{\partial \xi} < 0$ (see, e.g., Chetty et al., 2016; Kinge et al., 2019). Analyses of life insurance lapse behavior also show financial literacy and wealth or related concepts such as unexpected liquidity

needs as important determinants (Fier et al., 2013; Nolte and Schneider, 2017; Fang and Kung, 2021; Gottlieb and Smetters, 2021). We can thus assume that $\frac{\partial \rho^l}{\partial \xi} < 0$.

Based on the above considerations and the assumptions that I, ρ^d, and ρ^l are only driven by type, we can set $EU^{LI}(I_i, \rho_i^l, \rho_i^d, P) = EU^{LI}(I(\xi_i), \rho^l(\xi_i), \rho^d(\xi_i), P) = EU^{LI}(\xi_i, P)$ and similarly $EU^0(I_i, \rho_i^d) = EU^0(\xi_i)$ and then determine the influence of type on the maximum willingness to pay for insurance as

$$\frac{\partial \hat{\Pi}}{\partial \xi} = \left(\frac{\partial \hat{P}}{\partial I}\frac{\partial I}{\partial \xi} + \frac{\partial \hat{P}}{\partial \rho^l}\frac{\partial \rho^l}{\partial \xi} + \frac{\partial \hat{P}}{\partial \rho^d}\frac{\partial \rho^d}{\partial \xi} \right)(1-\rho^l) - \frac{\partial \rho^l}{\partial \xi}\hat{P}. \tag{6}$$

Here, the first term in the brackets is the influence of (wealth-) type on willingness to pay, which is positive given Equation (3). This is in congruence with the empirical evidence given in panel (b) of Figure 1 which shows how life insurance demand increases by wealth quintile. The second term describes the influence of type through lapse probabilities and is positive given $\frac{\partial \rho^l}{\partial \xi} < 0$ and Equation (5). The third term describes the influence of type through mortality and is negative given the above $\frac{\partial \rho^d}{\partial \xi} < 0$ and Equation (4). The last term in Equation (6) is positive due to $\frac{\partial \rho^l}{\partial \xi} < 0$ and stochastic dominance ensuring that $\hat{P} > 0$. $\frac{\partial \hat{\Pi}}{\partial \xi}$ can thus be positive or negative, depending on which of the above effects dominates.

Costs of the individual i and the influence of type on them are somewhat simpler to derive. In the above model, and in fact as a necessary assumption in EFC equilibrium models, the insurance company pays the fixed indemnity B in case the consumer dies and has not lapsed the policy. The costs are thus $c(\xi_i) = \rho^d(\xi_i)(1-\rho^l(\xi_i))B$. The influence of type on costs is given by

$$\frac{\partial c}{\partial \xi} = (1-\rho^l)B\frac{\partial \rho^d}{\partial \xi} - \rho^d B\frac{\partial \rho^l}{\partial \xi}. \tag{7}$$

Using the assumptions from above we can see that the first term is negative, while the second is positive. Like the effect of type on the willingness to pay, the effect of type on cost is also ambiguous and can be both positive and negative.

2.2 Determining the Market Equilibrium

Following Einav et al. (2010, henceforth EFC), the equilibrium of the insurance market is defined by the intersection of the demand curve and the average cost curve in the quantity/ price plane. Assuming types to be distributed according to $g(\xi)$, we can define the quantity demanded at price Π as

$$D(\Pi) = \int \mathbf{1}(\hat{\Pi}(\xi) \geq \Pi) g(\xi) d\xi. \tag{8}$$

Note that the demand curve is always falling in price because there is only a single insurance contract and consumers are assumed to purchase one of these contracts at most. Marginal costs at price Π are then given as $MC(\Pi) = E[c(\xi)|\hat{\Pi}(\xi) = \Pi]$, which allows us to determine average costs as

$$AC(\Pi) = \frac{1}{D(\Pi)} \int c(\xi) \mathbf{1}(\hat{\Pi}(\xi) \geq \Pi) g(\xi) d\xi. \tag{9}$$

From Equation (9) it is apparent that average costs and marginal costs have the same slope in the quantity/ price plane. To determine the type of the market equilibrium, it is relevant whether the cost curves are increasing or decreasing the price. An analytical solution cannot be obtained for the general case, because without further assumption $\frac{\partial MC}{\partial \Pi}$ may not be well-behaved or might not exist at all. We thus make the simplifying assumption that both c and Π are monotonous in type ξ. In terms of the micro-economic foundations, this assumption implies that one of the opposing effects in Equations (6) and (7) dominates the others. While this assumption may not be realistic for the entire support of $g(\xi)$, we would argue that it is realistic for the relevant area of the support where the equilibrium solution lies. Using this assumption, the following proposition results.

Proposition 1. *If $c(\xi)$ and $\Pi(\xi)$ are monotonous in ξ, the resulting equilibrium is one of adverse selection if $sgn\left(\frac{\partial c}{\partial \xi}\right) = sgn\left(\frac{\partial \Pi}{\partial \xi}\right)$ and one of advantageous selection if $sgn\left(\frac{\partial c}{\partial \xi}\right) \neq sgn\left(\frac{\partial \Pi}{\partial \xi}\right)$, where sgn denotes the sign of an expression.*

Proof. Let Ξ denote the support of $g(\xi)$. Because $\hat{\Pi}(\xi) : \Xi \mapsto \mathbb{R}$ is monotone, $\hat{\Pi}^{-1}(\Pi) : \mathbb{R} \mapsto \Xi$ is well-defined. We can thus write

$$MC(\Pi) = E[c(\xi)|\hat{\Pi}(\xi) = \Pi] = E[c(\xi)|\xi = \hat{\Pi}^{-1}(\Pi)] = c(\hat{\Pi}^{-1}(\Pi)). \tag{10}$$

The last step follows, because monotonicity in Π ensures that there is a single value of ξ such that $\hat{\Pi}(\xi) = \Pi$. From the above it follows that $\frac{\partial MC}{\partial \Pi} = \frac{\partial c}{\partial \xi}\frac{\partial \hat{\Pi}^{-1}}{\partial \Pi}$ and thus $\frac{\partial MC}{\partial \Pi} > 0$ if and only if $sgn\left(\frac{\partial c}{\partial \xi}\right) = sgn\left(\frac{\partial \hat{\Pi}^{-1}}{\partial \Pi}\right)$. The proposition follows, because monotonicity implies that $sgn\left(\frac{\partial \hat{\Pi}^{-1}}{\partial \Pi}\right) = sgn\left(\frac{\partial \hat{\Pi}}{\partial \xi}\right)$. □

Proposition 1 in combination with Equations (6) and (7) makes it clear that both adverse and advantageous selection can be potential outcomes for the market of life insurance, depending on which of the effects dominates.

As an illustration, it is informative to consider the cases in which type only affects one of the two probabilities. We being with the situation that is most commonly considered in the literature, which is type only affecting mortality. In this case, $\frac{\partial c}{\partial \xi} < 0$. The effect of type on the maximum willingness to pay is, however, unclear. If the effect of type through the death probability dominates, $\hat{\Pi}$ will increase in the type and adverse selection will appear. Empirical evidence for this case is given in He (2009). If, however, the direct effect of type dominates the effect through the death probability, then advantageous selection results. This situation is, for example, documented by Cawley and Philipson (1999) and Hedengren and Stratmann (2016).

The second illustrative case is the situation in which type only affects the lapse probability. In this case, both derivatives are clearly signed. Type decreases the lapse probabilities and thus increases costs, siuch that $\frac{\partial c}{\partial \xi} > 0$. Similarly, type increases the willingness to pay both directly and though the lapse probability, such that $\frac{\partial \hat{\Pi}}{\partial \xi} > 0$. As a consequence, the resulting market equilibrium will be one of adverse selection. Evidence consistent with this outcome is documented in Carson et al. (2020).

2.3 The Case of Consumer Naïveté

The derivation of $\hat{\Pi}$ in Equations (1) through (6) make the implicit assumption that consumers not only know their respective death and lapse probabilities, but that they also take these probabilities into account when deciding whether or not to purchase a life insurance policy. Whether this is a valid assumption in insurance markets is unclear (Cawley and Philipson, 1999). Gottlieb and Smetters (2021) make the case that consumers are not acting in this manner, at least when their lapse probability is concerned. They conclude from a survey of life insurance policyholders that the respondent do not consider their own beliefs about future income shocks when forming a subjective lapse probability. A similar argument can be made about mortality risk. Subjective mortality estimates of survey respondents have substantial differences to their objective counterparts (Bissonnette et al., 2017).

It is thus fathomable that consumers do vary in their insured costs (through both ρ^d and ρ^l), but that they are not aware of this variation and do not take it into account in their insurance choices. Naïve consumers make the occurrence of adverse selection less likely. After all, adverse selection assumes that consumers know their underlying risk and act accordingly. We model consumer naïveté by assuming that they determine their maximum willingness to pay without regard to their idiosyncratic values of ρ_i^d and ρ_i^l. Rather consumers simply assume themselves to be equal to the average in both values. Formally, if average values are denoted by index μ, consumers with type ξ_i have a naively determined maximum acceptable premium, $\hat{P}_i^n$, defined implicitly by

$$EU^{LI}(I(\xi_i), \rho_\mu^l, \rho_\mu^d, \hat{P}_i^n) - EU^0(I(\xi_i), \rho_\mu^d) = 0. \tag{11}$$

In consequence, $\frac{\partial \hat{P}_i^n}{\partial \rho_i^l} = \frac{\partial \hat{P}_i^n}{\partial \rho_i^d} = 0$ and thus the derivative of the maximum willingness to pay towards type greatly simplifies to

$$\frac{\partial \hat{\Pi}^n}{\partial \xi} = (1 - \rho_\mu^l)\frac{\partial \hat{P}^n}{\partial I}\frac{\partial I}{\partial \xi} > 0. \tag{12}$$

Because the effect of type on the willingness to pay is always positive, the type of market

outcome is determined by the influence of type on the cost function. In case type affects costs more strongly through the death probability, we observe $\frac{\partial c}{\partial \xi} < 0$ and advantageous selection. In case the effect through the lapse probability is stronger, $\frac{\partial c}{\partial \xi} < 0$ and adverse selection will be the result. Thus, even with naïve consumers, both market outcomes are possible.

The naïve consumer model acts as a robustness check of our result that both adverse and advantageous selection can appear in the live insurance market. Whether the model is realistic is unclear. There are also a variety of results which oppose the view of naïve consumers when it comes to lapse and death probabilities. Carson et al. (2020) show that life insurance policyholders in their sample select themselves into contracts which are more suited to their own lapse behavior, thus indicating sophistication about their lapse probabilities. Hurd and McGarry (2002) find that subjective mortality estimates do predict actual mortality and various studies have found that people make economic choices based on their own mortality (e.g. Jayachandran and Lleras-Muney, 2009).

3 Data

Data Source: Our primary data source comes from the annual statutory reports that life insurers file with the National Association of Insurance Commissioners (NAIC). These reports offer comprehensive information on premiums, face amounts, death and surrender benefits, and other financial details at both the state and national levels. We focus on the U.S. individual life insurance market because our framework and empirical strategy target long-term contractual features and related lapse behavior.

Every year, each U.S. licensed life insurer submits detailed financial statements and exhibits to the National Association of Insurance Commissioners (NAIC). We use two exhibits from these filings: the "State Page", which provides state-by-state data on total premiums, the quantity (number of policies and total face amount) of in-force, newly issued, and newly ceased policies as well as data on surrender benefits and death benefits paid to policyholder. We also use the "Exhibit of Life Insurance", which offers similar data on a nationwide basis, with additional details on the

premiums earned for newly issued policies and a detailed breakdown of why policies ceased.

Variable Construction: For our analysis, we will be estimating both a demand function and a cost function. For the demand function, we define quantities as the face value of policies newly issued each year, which we observe at the firm-state-year level. Prices are more complicated. While we observe total premiums, total face value of policies in force, and the face value of newly written policies at the state level for every firm, we only observe the premiums associated with the newly issued policies at the national level. However, prices for individual life insurance are largely set nationally (Wenning, 2024). Thus, using prices on the national firm-year basis is not a large limitation for our analysis.

Estimating costs is more challenging. The NAIC reports have detailed, firm-state-year level data on cashflow, but there is no way to connect payouts in one period to when that policy was initially purchased. As a workaround, we calculate a (discounted) "future average costs per ceased" variable for each firm-state-year. For firm i in state s in year t, we calculate

$$C_{ist} = \frac{1}{\sum_{\tau=t+1}^{T} \beta^{\tau}} \sum_{\tau=t+1}^{T} \beta^{\tau} \frac{D_{is\tau} + S_{is\tau}}{N_{is\tau}} \tag{13}$$

Where $D_{js\tau}$ and $S_{js\tau}$ are the death and surrender benefits incurred in year τ and $N_{js\tau}$ is the total face value of policies that "cease" in year τ. We calculate future death, lapse, and surrender rates in a similar manner.

Sample Selection: We begin from the sample of all insurer-state-years with positive issued individual (or ordinary) life insurance. We then sequentially apply filters to ensure data quality and analytical suitability, requiring: (1) valid positive values for key variables (policy costs, premiums (group and individual), and in-force amounts), (2) plausible rate values between 0 and 1 for lapse, surrender, and mortality rates, and (3) a minimum 5-year observation window to assess long-term outcomes. The impact of each filter on our sample is shown in Table 1. The final sample retains around 25% of original observations, representing 77,652 firm-state-years and $9.2 trillion of issued face value. The largest sample restriction is the requirement that the insurer sells both individual and group insurance (which is required for our instrument).

Table 1: Sample Selection Steps

Sample Selection Step	Observations	% of Full	Total Issued ($B)	% of Total
Full Sample	316,819	100.0	37,065.34	100.0
Valid Lapse Rate	286,080	90.3	34,941.97	94.3
Valid Death Rate	285,684	90.2	34,935.99	94.3
Valid Surrender Rate	283,567	89.5	34,891.2	94.1
Valid In-Force Amounts	280,073	88.4	34,824.21	94.0
Valid Ind. Premiums	275,204	86.9	34,786.84	93.9
Valid Grp. Premiums	98,519	31.1	12,877.18	34.7
Valid Future Costs	97,852	30.9	12,871.16	34.7
>5 Years Observed	77,652	24.5	9,296.34	25.1

Notes: This table shows the sequential impact of data filters on the sample size. The full sample includes all firm-state-year observations from the NAIC data that issue individual life insurance. Subsequent rows show observations remaining after applying each filter criterion. "Valid" means non-missing, positive, and finite.

Final Sample Descriptive Statistics: Table 2 summarizes key descriptive statistics. The average insurance company issues life insurance coverage with a face value of $101 million per firm-state-year. This coverage has a mean future costs when ceased of $25 cents (with a median of $10 cents) per $1 of face value. This is contrasted with a mean individual premium rate of about $2–3 cents per $1 of face value (with a median of 1.14 cents).

The premium income and cost parameters differ by about an order of magnitude. To have profitable life insurance companies, coverage thus needs to cease at a rate of 10% or less, on average. This is is supported by the data. On average, 5% of policies lapse annually, 2% terminate due to surrender and 1.5% due to death. Figure 8 shows these parameters and the expiration rate over time. We can see that the order of the different reasons to cease is the same over time. While the death rate is approximately constant over the whole time period, the surrender rate decreases over time, ending at about half of its 1993 value in 2017. The lapse rate does not have a discernible time trend, but shows temporary spikes during the economic downturns of the dot-com bubble and the financial crisis.

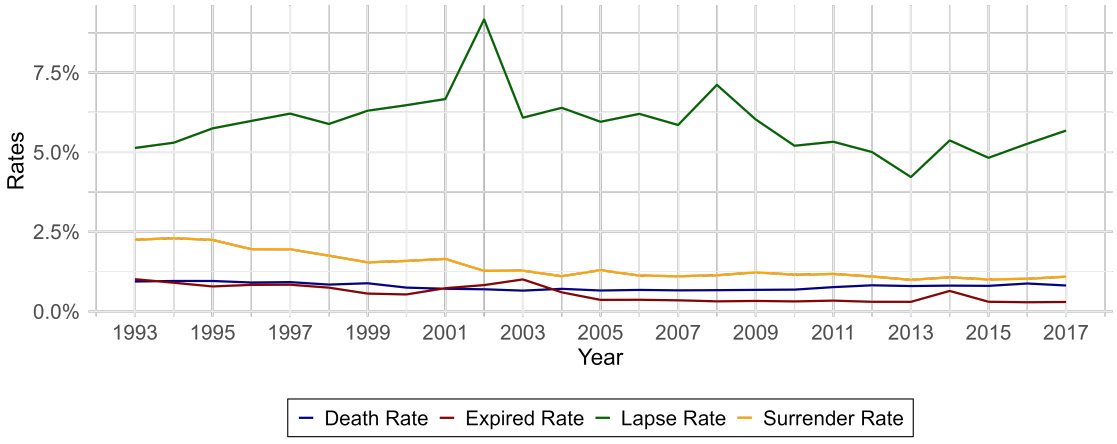
Table 2: Summary Statistics

Variable	Mean	SD	P25	Median	P75
Life Insurance Issued (000,000s)	101.41	292.36	0.8	8.33	55.15
Avg. Future Cost when Ceased	0.25	0.47	0.04	0.1	0.22
Avg. Ind. Life Premium	0.02	0.03	0.01	0.01	0.02
Avg. Grp. Life Premium	0.03	0.06	0.00	0.01	0.02
Future Lapse Rate (%)	4.69	3.87	2.02	3.69	6.03
Future Surrender Rate (%)	2.23	1.23	1.39	1.98	2.89
Future Death Rate (%)	1.52	1.37	0.56	1.1	1.9
In-force Prior Year (000,000s)	758.47	2,239.95	12.23	77.94	407.2
Years Observed	14.59	5.44	10.00	14.00	19.00

Unique:	
Firms:	287
Firm-Years:	2,446
Firm-States:	9,868
Firm-State-Years:	77,652

Note: This table presents the summary statistics for subsample used in the analysis. All variables (except years observed) have been Winsorized at the 1% and 99% level. Future Cost, and Premiums are all per $ of Face Value. Amount Issued, and Amount In-Force Prior are all in units of $ of Face Value. Future Costs are the average costs per dollar of face value incurred for policy surrender and policyholder death over the next 10 years weighted by policies in force and discounted using a 3% discount rate.

Figure 3: Policy Cease Rates Over Time



Note: Figure displays the timeline of the average lapse, surrender, death, and expiry rate for all policies in the NAIC data.

4 Estimation Strategy and Results

4.1 Methods

In the reduced form, we start by estimating a demand curve and an average cost curve. Assuming (log) linearity, we define the these two curves as follows:

$$ln(Q_{ist}) = \alpha_{ist} + \beta ln(P_{ist}) \tag{14}$$

$$ln(AC_{ist}) = \gamma_{ist} + \delta ln(P_{ist}) \tag{15}$$

where Q_{ist} is the demand for life insurance at firm i in state s and year t, P_{ist} is the premium, AC_{ist} is the average cost curve, where "costs" are defined as the present value of future cash outflows of the associated policy. β is the price-elasticity of demand and δ is the price elasticity of the average cost curve and represents the combination of mortality and lapse-induced selection. The intercept terms α and γ are specific to the firm-state-year observation to reflect the presence of control variables and fixed effects in our estimation. Note that the assumption of log-linearity translates into an assumption of constant price elasticities in both the demand and the cost function.

Suppressing indexes for legibility, we can then express the inverse demand function, that is the function mapping demand to prices, and average cost as a function of demand as

$$P(Q) = exp\left(-\frac{\alpha}{\beta}\right) \times Q^{\frac{1}{\beta}} \tag{16}$$

$$AC(Q) = exp\left(\gamma - \frac{\alpha\delta}{\beta}\right) \times Q^{\frac{\delta}{\beta}} \tag{17}$$

In their initial model, Einav et al. (2010) assume a zero profit equilibrium which would imply that $P = AC(P)$. This assumption, however, is not realistic in the market for life insurance where we know markups to exist. We thus adopt the model of Hackmann et al. (2015) which allows for profits of the firm. Further, our average costs are discounted total outflows of cash, while our premiums are measured annually. To adjust for both profits and the difference in time frame, we introduce a multiplicative firm-state-year specific factor ζ on average costs. The equilibrium

average cost curve of firm i in state s at time t then has the intercept $\tilde{\gamma} = \gamma - ln(\zeta)$. Here, and below, the tilde indicates the adjusted version of the cost curve. To obtain its value, we set the market equilibrium for a given firm in a given state-year as the price and quantity observed in the data. Because this point defines both demand and adjusted average cost, we can use it to get the value of α through Equation (14). For this, we use the fact that we know β and δ, which are not firm-state-year specific, from our estimations. Leveraging Equation (17), we obtain $\tilde{\gamma}$ through[6]

$$\tilde{\gamma} = ln\left(\frac{P}{Q^{\frac{\delta}{\beta}}}\right) + \frac{\alpha\delta}{\beta}. \tag{18}$$

This allows us to express the adjusted average cost function as

$$\widetilde{AC}(q) = exp\left(\tilde{\gamma} - \frac{\alpha\delta}{\beta}\right) \times Q^{\frac{\delta}{\beta}} \tag{19}$$

To obtain welfare estimates, we further require the marginal cost function. Again, we use an adjusted version. We know from Einav et al. (2010) that

$$\widetilde{MC}(P) = \left(\frac{\partial Q(P)}{\partial P}\right)^{-1} \frac{\partial \widetilde{AC}(P) Q(P)}{\partial P} \tag{20}$$

$$\widetilde{MC}(P) = \frac{\delta + \beta}{\beta} \times \widetilde{AC}(P) \tag{21}$$

We substitute Equation (19) and obtain the adjusted marginal cost curve as

$$\widetilde{MC}(Q) = \frac{\delta + \beta}{\beta} \times exp\left(\tilde{\gamma} - \frac{\alpha\delta}{\beta}\right) \times Q^{\frac{\delta}{\beta}}. \tag{22}$$

We have set the equilibrium demand quantity Q_{eq} such that it implies $\widetilde{AC}(Q_{\text{eq}}) = P(Q_{\text{eq}})$ mechanically. The efficient level of demand, Q_{eff}, is given by $\widetilde{MC}(Q_{\text{eff}}) = P(Q_{\text{eff}})$, which renders

$$Q_{\text{eff}} = \left(\frac{\beta}{\delta + \beta}\right)^{\frac{\beta}{\delta - 1}} \times Q_{\text{eq}} \tag{23}$$

[6] An alternative way to approach this is that in equilibrium it holds that $\tilde{AC}_{\text{eq}} = P(Q_{\text{eq}})$. One can thus calculate a factor $\iota = \frac{AC_{\text{eq}}}{P(Q_{\text{eq}})}$ which implies that $\tilde{AC} = exp(\gamma)P^{\delta}/\iota$, such that $\tilde{\gamma}$ can be found from $\tilde{\gamma} = \gamma - ln(\iota)$.

The dead-weight welfare loss is then given by

$$DWL = \int_{Q_{\text{eq}}}^{Q_{\text{eff}}} P(Q)dQ - \int_{Q_{eqm}}^{Q_{\text{eff}}} MC(Q)dQ \tag{24}$$

$$DWL = \left[exp\left(-\frac{\alpha}{\beta}\right) \frac{\beta}{\beta+1} Q^{\frac{\beta+1}{\beta}} - exp\left(\tilde{\gamma} - \frac{\alpha\delta}{\beta}\right) Q^{\frac{\delta+\beta}{\beta}} \right]_{Q_{\text{eq}}}^{Q_{\text{eff}}} \tag{25}$$

Now that we have defined the dead-weight loss (DWL) as a function of parameters, we next directly estimate β and δ, which we can then combine along with observed values of Q_{eq} and P_{eq} for each market to back out α and $\tilde{\gamma}$ and calculate the firm-state-year specific value of DWL.

The central challenge to these estimations are unobserved product characteristics that jointly influence price and cost/demand. To account for these unobserved factors, we use a "Hausman"-style instrument (Hausman, 1996). The instrument we use are group-insurance premiums for plans offered by the same firm in the same year. Because our analysis focuses on the individual market, the goal of this instrument is to capture firm-level cost shocks that are external to the individual market.[7] These cost shocks could either be at the general firm level, such as wages or taxes, or they could be at the group-insurance level, such as unexpectedly high costs at a few large clients, and spill over. This instrument follows an extensive recent tradition of instrumenting insurance prices with prices/costs by the same firm in other markets (e.g., Koijen and Yogo, 2016, 2015; Ge, 2022; Gao et al., 2023; Wenning, 2024).

We exploit this instrument using two-stage least squares (2SLS). The first stage regresses the endogenous variable, individual life premiums, on the instrument, group life premiums. Formally, we estimate

$$\ln(P_{it}) = \pi \ln(P_{it}^{\text{group}}) + \gamma \mathbf{X}_{it} + \phi_{is} + \tau_{st} + u_{it} \tag{26}$$

where $\ln(P_{it})$ is the log of individual life premiums for firm i in year t, $\ln(P_{it}^{\text{group}})$ is the log of group life premiums, which serves as the instrument, for the same firm-year and $\mathbf{X}_{it}$ includes firm-year-

[7] Any local mortality shocks that impact both group and individual insurance will be picked up by the state-year fixed effects.

specific control variables.[8] Firm-state fixed effects (ϕ_{is}) and state-year fixed effects (τ_{st}) are also included to isolate identifying variation at the firm-year level instead of biasing the estimates due to effects of the general firm-level quality (Nevo, 2001). The error term u_{it} then captures the portion of premiums that is unexplained by the instrument and the included controls / fixed effects. Note that we estimate this first stage at the firm-state-year level because that is the level of our second stage. This explains why the fixed effects are defined state-wise even though all variation of the other variables in Equation (26) is on the firm-year level.

The predicted values from this first stage, $\widehat{\ln(P_{it})}$, are then used in the second stage to estimate the causal relationship between premiums and the outcome variables of interest. Formally, the second-stage regressions are

$$\ln(Q_{ist}) = \beta\widehat{\ln(P_{it})} + \theta_q \mathbf{X}_{it} + \phi_{q,is} + \tau_{q,st} + \varepsilon_{q,ist} \tag{27}$$

$$\ln(C_{ist}) = \delta\widehat{\ln(P_{it})} + \theta_c \mathbf{X}_{it} + \phi_{c,is} + \tau_{c,st} + \varepsilon_{c,ist} \tag{28}$$

where $\ln(Q_{ist})$ and $\ln(C_{ist})$ represent (logged) quantity and the (logged) present value of future costs, respectively, for firm i in state s in year t, and $\widehat{\ln(P_{it})}$ is the predicted log of individual life premiums from the first stage. Note that the use of the log-linear specification allows us to use the discounted future costs in Equation (28) rather than the time-frame and profit adjusted cost and still obtain the same estimate for δ. As in the first stage, the model also includes firm-year controls $\mathbf{X}_{it}$, firm-state fixed effects (ϕ_{is}), and state-year fixed effects (τ_{st}). The error terms ε_{ist} capture idiosyncratic shocks and are uncorrelated with the error term from the first stage. Standard errors are clustered at the firm-year level, which is the unit of random variation in our instrument (Abadie et al., 2023). Our price instrument also allows to separately test the two potential selection channels through the same set up. We estimate each channel through

$$\ln(L_{ist}) = \lambda\widehat{\ln(P_{it})} + \theta_l \mathbf{X}_{it} + \phi_{l,is} + \tau_{l,st} + \varepsilon_{l,ist} \tag{29}$$

$$\ln(D_{ist}) = \omega\widehat{\ln(P_{it})} + \theta_d \mathbf{X}_{it} + \phi_{d,is} + \tau_{d,st} + \varepsilon_{d,ist} \tag{30}$$

[8] Firms set life insurance premiums largely at the national level rather than separately for each state that they operate in. See Wenning (2024) for more discussion.

Where $\ln(L_{ist})$ and $\ln(D_{ist})$ the (logged) average future lapse and death rates. λ and ω then represent how random shocks in firms' prices influence the two primary drivers of firms' costs. If either of the coefficients are zero, then that is evidence that consumers do not select along that dimension.

4.2 Demand and Cost Estimation Results

Table 3 present the results of our two-stage least squares (2SLS) analysis. Column (1) shows the first-stage regression results, while columns (2) and (3) display the second-stage results examining the impact of premiums on future costs and demand in the individual life insurance market.

Table 3: Estimation Results

	log(Avg. Ind. Premium)	log(Current Demand)	log(Future Avg. Costs)
log(Avg. Group Premium)	0.047***	–	–
	(0.002)	–	–
log(Avg. Ind. Premium) (fitted)	–	-0.904***	0.125***
	–	(0.093)	(0.017)
Firm Size Control	Yes	Yes	Yes
Firm-State FE	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes
Years Observed FE	Yes	Yes	Yes
Observations	77,652	77,652	77,652
KP First-stage F	–	668.9	668.9

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: The instrumental variable for all models is *ln(Average Group Premium)*, which is the premium for the group market. In the second stage, the main dependent and independent variables are on the firm-state-year level. The estimation equations are given in Equation (26) for the first stage and Equations (27) and (28) for the second stage. Current demand is given by the face value of new policies sold. Future average costs are discounted monetary outflows in future periods as described in Equation (13). The size control is the face vale of in-force life insurance prior to the current year. Data is taken from annual reports submitted to the National Association of Insurance Commissioners. Standard errors, clustered by Firm-Year (the unit of randomization for our instrument, Abadie et al., 2023), are in parentheses.

The first-stage results demonstrate the strength of our instrument, group life insurance premiums. A 1% increase in group premiums is associated with a 0.05% increase in individual premiums, indicating a small, but significant, spillover of costs between markets. The strength of this relationship is evidenced by the high Kleibergen-Paap first-stage F-statistics (>650), supporting the relevance condition of our instrumental variable.

Column (2) of Table 3 shows the cost function estimation and our estimates of δ. We find that a 1% increase in premiums is associated with approximately a 0.125% increase in future costs for insurers. The positive coefficient suggests that as premiums rise, the insurer's risk pool changes in a way that increases expected future payouts, indicating adverse selection. Column (3) estimates the demand model and reveals a price elasticity β of -0.904, indicating that a 1% increase in premiums leads to about a 0.9% decrease in the quantity of life insurance demanded.

Table 4 isolates the two potential channels for selection. In column (1), we find that higher premiums are associated with lower lapse rates. A 1% increase in premiums corresponds to approximately 0.145% decrease in future lapse rates and a 0.034% decrease in future surrender rates. This result suggests that lapse and surrender behavior is a channel through which selection operates in this market. As detailed in Section 2.3 it does not necessarily indicate that people incorporate knowledge of their future lapse rates into their current purchasing decisions. The negative relationship could also indicate that policyholders who maintain coverage at higher premium levels may have different risk characteristics or valuation of insurance compared to those who lapse. In contrast, we find no statistically significant relationship between premiums and death rates. This tightly estimated null result is not a placebo test, but rather indicates that mortality risk is not a primary channel through which selection is occurring in this market. The absence of a relationship between premiums and death rates, coupled with the significant relationship between premiums and lapse and surrender rates, suggests that the observed adverse selection is primarily driven by lapse and surrender behavior rather than mortality risk.

Table 4: Estimation Results

	log(Future Lapse Rates)	log(Future Death Rates)	log(Future Surrender Rates)
log(Avg. Ind. Premium) (fitted)	-0.145***	0.003	-0.034***
	(0.014)	(0.009)	(0.008)
Firm Size Control	Yes	Yes	Yes
Firm-State FE	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes
Years Observed FE	Yes	Yes	Yes
Observations	77,652	77,652	77,652
KP First-stage F	668.9	668.9	668.9

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: The instrumental variable for all models is *log(Average Group Premium)*, which is the premium for the group market. The main dependent and independent variables are on the firm-year level and are the (discounted) observed future rates at which existing policies either lapse, surrender, or end in death. The size control is the face vale of in-force life insurance prior to the current year. Data is taken from annual reports submitted to the National Association of Insurance Commissioners. Standard errors, clustered by Firm-Year (the unit of randomization for our instrument, Abadie et al., 2023), are in parentheses.

4.3 Welfare Results

We now use our estimated coefficients from Table 3 to calculate the welfare consequences of adverse selection in the US life insurance market according to the method laid out in Section 4.1. The resulting parameters provide us with demand, average cost, and marginal cost curves for each firm-state-year observation.[9] One such observation is displayed in panel (a) of Figure 4. It shows the dead-weight loss as a consequence of the marginal cost curve lying below the average cost curve which is a consequence of average costs being increasing in price. We can see in this picture that adverse selection leads to less policies sold in equilibrium than would be the efficient outcome in a world without information asymmetries. Due to the structure of our estimation, the relative decrease in the number of policies due to adverse selection is the same in every market at about 16.6%. This number is constant across all observations due to our assumption of constant elasticities in the demand and cost curves.

The exact amount of dead-weight loss measured for each observation varies considerably. As can be seen in the second row of Table 5, the average firm-state observation shows a dead-weight loss of about $34,900, but this value can be essentially 0 or higher than $1 million per firm.[10] This is not surprising, given that many firms have states in which they write relatively few policies.

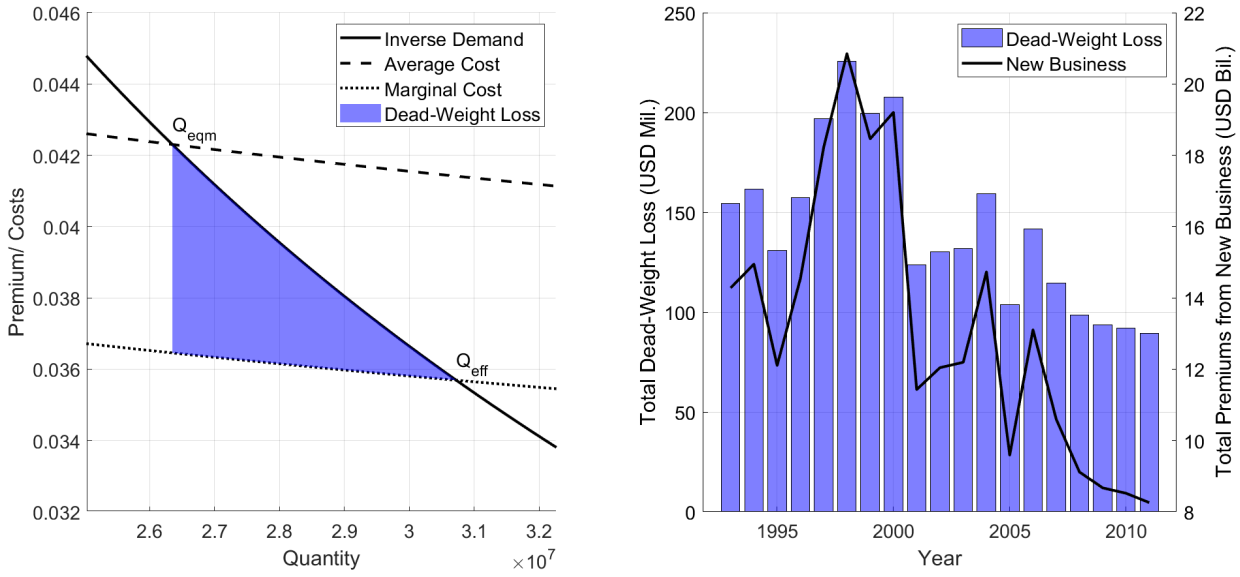
[9] Note that this is the welfare loss for the year that the policy is purchased. We extend the estimates for the full policy timeline below.

[10] Numbers are inflation adjusted to reflect 2023 values and are skewed highly across firm-state observations.

Figure 4: Graphical Display of Welfare Results

(a) Market situation of a sample firm: Aetna Life, in the state of New York in year 1998

(b) Development of annual DWL implied by new business over time



Note: Panel (a) of the figure illustrates the demand, average cost, and marginal cost curves for a sample firm-state-year observation. Curves are calculated according to the assumptions and methods laid out in Section 4.1. Quantities are expressed in US$ policy face values, while premiums and costs are expressed in US$ per US$ policy face value. Panel (b) shows the development of the estimated aggregate dead-weight loss over time. On the secondary y-axis, the panel also reports the development of new business in the form of total annual premiums earned.

Aggregated across all firms, annual dead-weight losses in the initial year of purchase, shown in the third row of Table 5, have an average value of $142.79 million and are not skewed very strongly, showing that this mean value is not driven by outliers.

The fourth and fifth rows of Table 5 are a back-of-the-envelope estimate for the total NPV of the DWL and is calculated by projecting the purchase-year DWL into the future and discounting by both a time discount rate as well as the project rate of policy "ceasing".[11] The total NPV estimates are almost 10 times as large as the purchase, with the average annual aggregate now over a billion dollars a year.

Panel (b) of Figure 4 shows the development of the inflation adjusted dead-weight loss in purchase year over time. The figure shows a slightly declining trend in the absolute numbers. How-

[11]See Appendix A for full details on the calculation.

ever, the figure shows an even stronger decline in the total premiums from new business, mirroring the development of fewer US households having life insurance coverage over the years. Relatively speaking, the problem of adverse selection in life insurance thus seems to have increased rather than decreased.

Table 5: Results of Dead-Weight Loss calculations

	Mean	25^{th} Pctl	Median	75^{th} Pctl
Share of Eff. Q (%)	85.76 [83.57, 89.63]			
DWL in Purchase Year:				
Firm-State-Year ($000s)	34.9 [18.4, 48.7]	0.4 [0.2, 0.5]	3.3 [1.7, 4.5]	19.1 [10.0, 26.6]
Annual Aggregate ($000,000s)	142.79 [75.18, 198.94]	109.15 [57.46, 152.07]	131.88 [69.43, 183.73]	160.47 [84.48, 223.57]
NPV of Total DWL:				
Firm-State-Year ($000s)	316.3 [166.5, 440.6]	3.3 [1.7, 4.6]	29.5 [15.5, 41.1]	172.5 [90.8, 240.3]
Annual Aggregate ($000,000s)	1292.55 [680.49, 1800.79]	988.03 [520.16, 1376.52]	1193.73 [628.46, 1663.10]	1452.58 [764.73, 2023.73]

Notes: Aggregate results from the dead-weight loss calculations according to the methods laid out in Section 4.1. Dollar values are inflation adjusted to reflect prices in 2023. The first row displays the share of policies sold in equilibrium when compared to the efficient market solution: $Q_{\text{eq}}/Q_{\text{ef}}$. There are no summary statistics other than the mean, because this value is the same across all observations due to our assumption of constant elasticities in the demand and cost curves. The second row reports the mean and quartiles for the dead-weight loss in the year of purchase across firms, averaged over all years that they appear in the data. The second row reports the mean and quartiles for the dead-weight loss in the year of purchase of the aggregated welfare losses on the annual level. The last two rows report the estimated NPV of the total dead-weight loss over the life of the contract at again the firm-state-year level and then aggregated to an annual level. 90% confidence intervals for each statistic, derived from block-bootstrapping at the Firm-Year Level, are in brackets.

The values provided in rows two and three of Table 5 seem relatively low. However, this needs to be put into perspective. The numbers refer to annual efficiency losses for the newly instituted policies only. That is, the numbers needs to be multiplied by average contract duration in order to obtain estimates of the total welfare losses. When aggregated to the full contract length, we observe an aggregate net-present value of the dead-weight loss for new business of $1.29 billion. This is a sizable number indicating how problematic the issue of lapsation-based adverse selection in life insurance is. This is the case even though the equilibrium by Einav et al. (2010) commonly

leads to relatively small estimates of absolute efficiency losses. In the original study proposing the equilibrium, the authors themselves only report annual efficiency losses of about $10 per policyholder in health insurance. Moreover, our reduced-form estimates likely underestimate the welfare costs of moral hazard. This is due to the nature of how we calculate costs on the firm-state-year level. Our estimation implicitly assumes that a change in price affects the costs of all insurance contracts in the firm's portfolio. In reality, however, it will only affect the costs of new contracts from this point on. The inertia of the insurance portfolio is not modeled appropriately in our reduced-form model, and thus the price elasticity of costs is likely understated in Table 3 leading to lower estimates of the dead-weight loss. Even with all these associated caveats, the NPV of efficiency losses in new business exceeds one billion dollars a year. In relative terms, the adverse selection reported here leads the market to shrink by about one sixth compared to the first-best solution. This underlines the magnitude of the effect of adverse selection in life insurance for the financial security of American households.

Table 5 provides standard errors for the dead-weight loss estimates which are derived from block-bootstrapping the estimation on the Firm-Year level and using the coefficients from each iteration to calculate the efficiency losses. While the confidence intervals are relatively large, they do not include 0, mirroring the result from Table 3 that average cost are robustly estimated to be increasing in premiums.

4.4 Welfare Robustness

The confidence intervals provided in Table 5 illustrate the robustness of our dead-weight loss estimates against estimation errors. However, they rely on the log-linear specification of our estimation as it is presented in Equations (27) and (28). An alternative way of considering the robustness of welfare estimates is considered by Kang and Vasserman (2022). They raise the important issue that in settings where demand functions are estimated based upon few (or two) points, welfare implications must be extrapolated from average treatment effects through functional form assumptions. Our setting is an especially pronounced example. For every insurance company, in every market, in every year; we infer welfare effects through counterfactual statements extrapolated

from two slope measurements and a firm-market-year specific data point. Our counterfactuals are thus identified at not few, but one data point and rely on the functional form assumption of the log-linear specification.

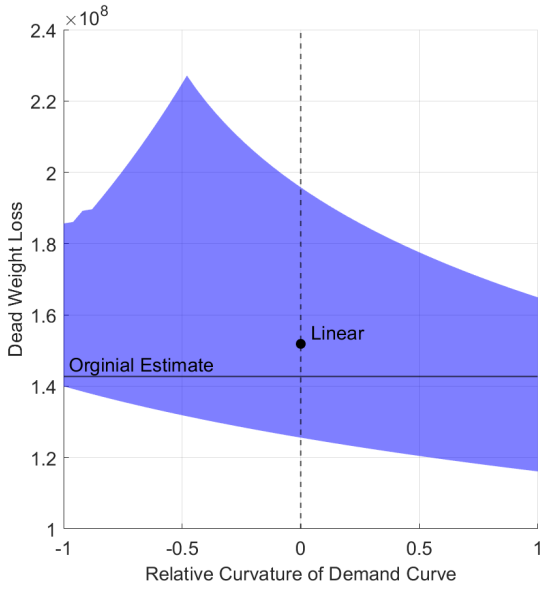
Kang and Vasserman (2022) suggest a procedure for quantifying the robustness of the welfare implications of a limited estimation. In their procedure, they allow for arbitrary functional forms under the restriction that between two observed points, the local slope does not deviate from the average slope between these two points by more that a proportional factor r. Unfortunately, their approach does not directly apply to Akerlovian selection markets. We describe both their approach and the problems in applying it to selection markets in more detail in Appendix C. The most crucial issue is that Akerlovian selection markets require two functional form assumptions – both the demand and cost curve – to determine the welfare consequences, and these functions interact in non-trivial ways (as can be seen Equation (A8)).

We thus propose an alternative robustness measure for functional form assumptions when estimating the dead-weight loss in an EFC framework. In the same spirit as Kang and Vasserman (2022), we stick to the idea that different functional forms lead to different welfare outcomes. However, to solve the issue of the complicated interactions between demand and cost curves, we restrict the functional forms of each curve to second-order polynomials and solve for the welfare bounds numerically. While still allowing us to consider both concave and convex functions for both curves, this assumption considerably simplifies the problem's numerical complexity. The exact procedure is described in Appendix C. Conceptually, we determine the most concave and convex feasible demand and average cost functions given our empirical estimates. We then use a grid between these two functions to calculate the resulting dead-weight loss for various combinations of functional forms. This lets us determine how much the curvature of the demand and cost functions influences our welfare estimates.

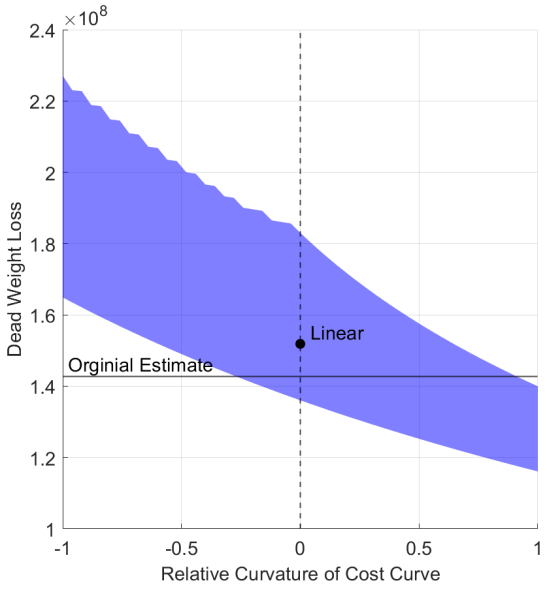
The results of this analysis are summarized graphically in Figure 5. Estimates of the average annual dead-weight loss for new business range between \$120 and \$220 million. This range is smaller than the confidence interval due to potential estimation error as reported in Table 5, but shifted upward due to the conservativeness of the log-linear specification. Even compared to the

Figure 5: Range of Dead-Weight-Loss due to Variations in Functional Assumptions

(a) Range as a function of curvature in the demand curve



(b) Range as a function of curvature in the average cost curve



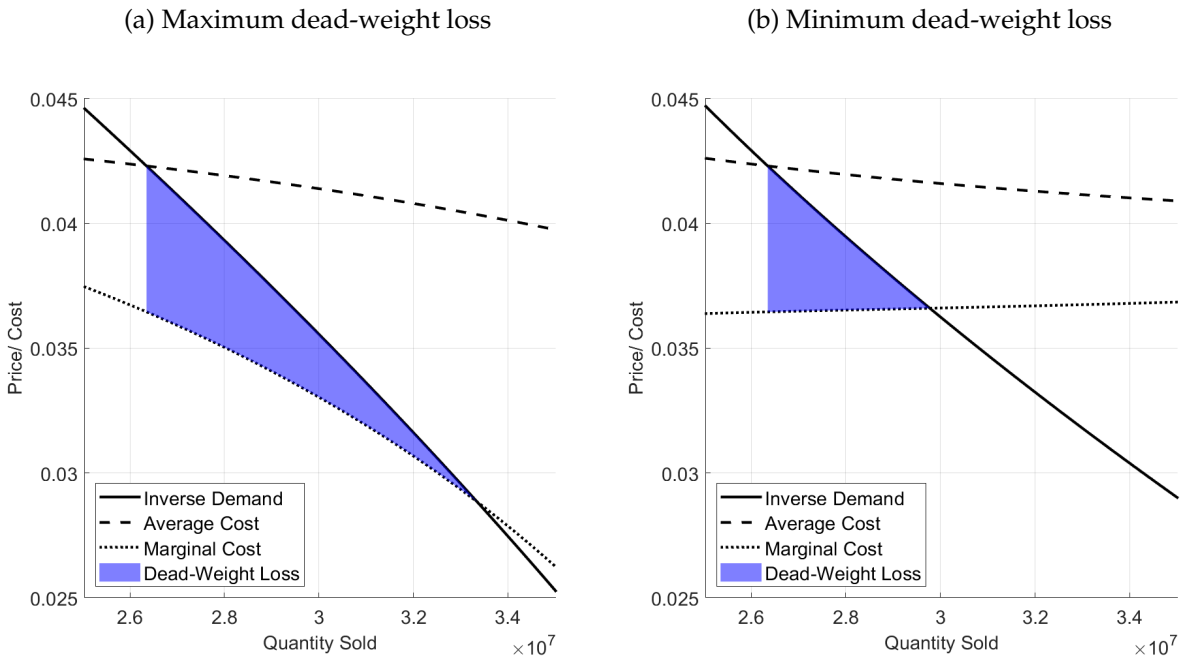
Note: The figure presents the range of possible annual aggregate dead-weight losses due to adverse selection when varying the functional form assumptions on the demand and the average cost curve. For each curve, we assume a second-order polynomial function and vary the curvature of the function between the minimum and the maximum as is described in Appendix C. Each panel varies the curvature of one of the functions on the x-axis and displays the range due to varying the other function on the y-axis. The linear specification where curvature is zero for both curves is indicated in each panel. We also show the dead-weight loss due to our empirical specification described in Table 5.

linear specification that is, for example, used in Einav et al. (2010), our specification underreports the dead-weight loss by about 5%. The figure thus provides confidence in the fact that the functional form assumption made in our estimation is not necessarily decisive for our conclusions about welfare consequences of adverse selection through lapsation. It should be noted that the boundaries on the ranges in both panels are sometimes ragged. This is because some combinations of curvatures for the demand and cost curve do not lead to an efficient market outcome of the form $P(Q_{\text{ef}}) = MC(Q_{\text{ef}})$ for positive prices. In that case, we ignore the combination of functional forms and use the combination with the highest dead-weight loss and positive prices instead. This is particularly likely to appear for very concave demand functions.

Under the caveat of an existing efficient market equilibrium at positive prices, Figure 5 shows

that the most extreme values for the dead-weight loss are achieved by the most extreme curvatures assumed for the demand and cost functions. Specifically, assuming the most concave possible demand and average cost function leads to the maximum possible dead-weight loss, while using the most convex functions leads to the minimum possible value. To show the consequences of these functional forms, Figure 6 displays the two resulting market equilibria for the firm Aetna Life in New York State in 1998, the same observation we used as an example for the original estimate in Figure 4.

Figure 6: Possible extremes of dead-weight loss due to variation in functional form for a sample firm



Note: The figure presents the largest (panel(a)) and smallest (panel (b)) dead-weight loss due to variations in the curvature of the demand and average cost function for the firm Aetna Life in the state of New York in 1998. It thus shows the possible extremes of the sample firm shown in panel (a) of Figure 4. For each curve, we assume a second-order polynomial function with the parameter range chosen as is described in Appendix C. The largest possible dead-weight loss results from the most concave demand and average cost curves, while the smallest results from the most convex curves.

Considering both panels in combination, we can see that changing the functional form does not shift the level of the marginal cost curve, but rather tilts it around the point defined by the observed equilibrium quantity. In other words, our point estimate exactly defines the welfare loss

of the marginal consumer and our functional form assumption and estimated slopes allow us to impute the welfare loss of the inframarginal consumers. In the appendix, we show that this holds generally in our empirical situation and is not dependent on the polynomial approach. In panel (a) we can see that using concave demand and cost curves shifts the efficient market solution outward, leading not only to larger dead-weight losses, but also to a significantly stronger relative contraction of the market due to adverse selection. In contrast, panel (b) shows that using convex demand and average cost curves can lead to marginal cost curves which are even increasing in the quantity sold. This does not remove adverse selection (which is due to the marginal costs being below the average cost), but it does lessen the potential consequences.

5 Discussion of Equity

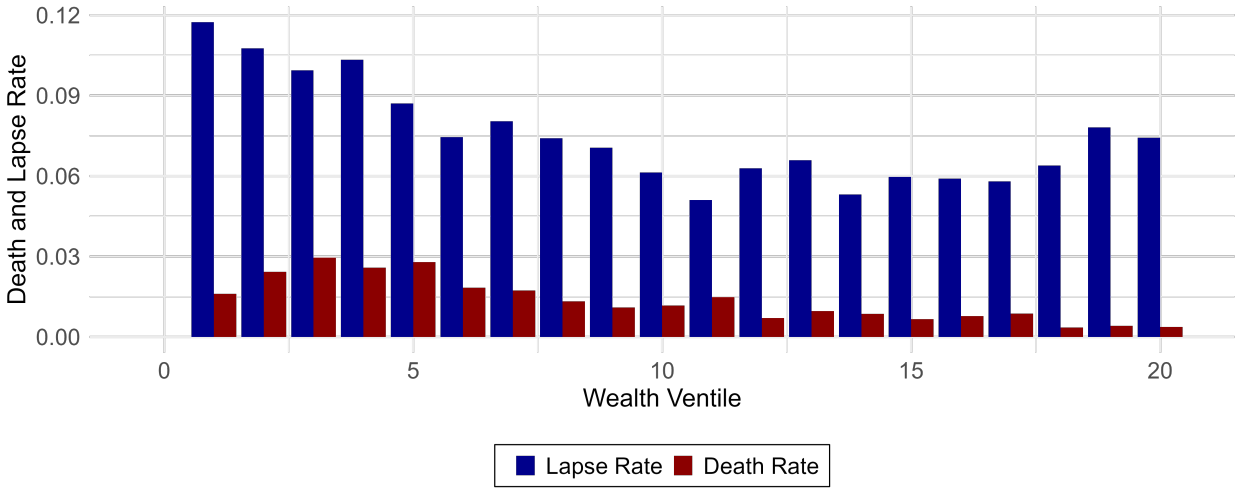
Akerlovian models of adverse selection such as the equilibrium by Einav et al. (2010) have a hard time speaking to the distributional and equity consequences of the market outcome. Our reduced form estimates show a dead-weight loss in total welfare due to lapsation-based adverse selection. This result suggests that measures to make prices contingent on lapse rate or to make prices equal costs for each contract year will increase the size of the market and the overall welfare. As with all results on the population level, however, there is heterogeneity in the individual welfare consequences such measures could have (Jansen et al., 2022; Geruso et al., 2023). As pointed out by Rothschild (2024), this is especially the case in Akerlovian markets of adverse selection. In such markets, ex-post welfare gains cannot be redistributed heterogeneously across the population because, by definition, the relevant dimension of heterogeneity is unknown to the central planner.

Our empirical results show that while there is an adverse selection channel through lapse rates, there is no mortality channel. Any change in the market targeted at increasing efficiency would thus have to use the lapsation channel. One such change would be insurers pricing life insurance contracts contingent on expected individual lapse rates. This is effectively the same as increasing the amount of information used for pricing in the sense of Jansen et al. (2022). The distributional consequences thus depend on the correlation of wealth with the take-up of life insurance coverage

as well as on the correlation of wealth with the tendency to lapse. Previous empirical studies, such as Gropper and Kuhnen (2025) as well as our own analysis of the Survey of Consumer Finances in Figure 1, provide a clear indication that this correlation is positive.

In Figure 7, we plot lapse rates and death rates among insureds in the past six waves of the Health and Retirement Survey by wealth ventile. We can see that the low wealth ventiles are significantly more likely to lapse, while there seems to be no discernible relationship for higher wealth levels.[12]

Figure 7: Empirical correlation of wealth and lapse rates



Note: The figure reports the lapse rates of respondents aged 64 and younger in the Health and Retirement survey for years 2010 through 2020. Only respondents which reported to have life insurance coverage in the last wave of the survey are considered. Wealth is taken as the sum of all assets minus the sum of all debt and the ventiles are constructed according to the prior period's wealth. See Appendix B for details on the variable construction.

To model the equity impacts of selection through lapsation, we employ a Berry et al. (1995) (BLP) style random coefficients discrete choice demand model with micro-moments. We derive the micro-moments on the association of household wealth and lapse, death and purchase rates from the Health and Retirement Survey (HRS).[13] After estimating the base model, we run a coun-

[12]See Appendix B for details on why we use the Health and Retirement Survey rather than the Survey of Consumer Finances and the variable construction.

[13]We merge surrender and lapse together because, though the HRS does report if consumers who have lapsed received a cash value, many of the observations are missing and the reported rates are an order of magnitude lower than the surrender rates in the NAIC data.

terfactual simulation of individually-underwritten lapse pricing and estimate the distributional impact of the resulting change in consumer welfare.

5.1 Utility Specification

We index the S state-level markets by $s = 1, \ldots, S$, with J insurance firms, indexed by $j = 1, \ldots, J$, and I consumers, indexed by $i = 1, \ldots, I$. In each period t, consumers who do not currently have a policy choose one product from the set of available firms or an outside option (denoted by 0).[14] Instead of modeling a dynamic utility function, we define an indirect, instantaneous, utility function for each the individual-firm-market combination in period t. The utility that consumer i derives from purchasing from firm j in state s in period t is specified as:

$$u_{ijst} = -\beta_{p,i} \times p_{jt} + \beta_{w,i} + \xi_{jt} + \tau_{st} + \phi_{js} + \varepsilon_{ijst} \tag{31}$$

where p_{jt} is the price offered by firm j at time t; β_{p_i} is the individual-specific price sensitivity; ξ_{jt} captures unobserved (to the econometrician) firm-year characteristics that influence demand; τ_{st} and ϕ_{js} are state-year and firm-state fixed effects, respectively; and ε_{ijst} is an i.i.d. Type I extreme value error term. $\beta_{w,i}$ represents the impact of consumer i's wealth on the demand for life insurance and can be interpreted as the aggregate effect of "direct" selection (i.e. the consumer's knowledge of their personal lapse rates) and the generally positive correlation of wealth and demand.

The individual-specific price and wealth coefficients are:

$$\beta_{p,i} = exp(\beta_{p,0} + \nu_i) \tag{32}$$

$$\beta_{w,i} = \sum_q \beta_{w,q} \times 1(Q(w_i) \in q) \tag{33}$$

where $\beta_{p,0}$ is the base price sensitivity, and ν_i represents the unobserved heterogeneity in the consumer's price elasticity and is assumed to be distributed with mean zero and standard deviation

[14] We define the total potential market size as the dollars of personal income minus the face value of individual insurance in force.

σ_p. We model the impact of wealth as a series of intercept terms, where q is a collection of equally spaced sets that partition the unit interval, $Q(w_i)$ is the consumer's wealth quantile, and $1()$ is an indicator function.

5.2 Estimation

To estimate the model, we follow the Taylor approximation method developed by Salanié and Wolak (2019) to estimate our discrete-choice demand model efficiently. Specifically, we linearize the random coefficient model around a small variance (σ^2) to circumvent computational difficulties associated with numerical integration over consumer heterogeneity and multiple micro-moments. We first define a function F, which maps observed market shares s_{jst} and parameters (β, σ) to unobserved product characteristics ξ_{jt}. We then approximate F with a second order Taylor approximation around $\sigma = 0$

$$F(s_{jst}, \beta, \sigma) \approx f_0(s_{jst}) - f_1(s_{jst})\beta - \frac{1}{2} f_2(s_{jst})\sigma^2. \tag{34}$$

Explicitly, the approximating terms become:

$$f_0(s_{jst}) = \ln\left(\frac{s_{jst}}{s_{0,st}}\right) \tag{35}$$

$$f_1(s_{jst}) = [p_{jt}, \tau_{st}, \phi_{js}] \tag{36}$$

$$f_2(s_{jst}) = [p_{jt}^2] \tag{37}$$

Where the first-order term captures average impact (at the firm-state-year) of each part of the utility function and the second-order terms (f_2), approximate the non-linear impacts of the randomness of individual-level wealth and price impacts. We then simply run a 2SLS model regressing f_0 on f_1 and the "artificial regressors" of f_2 to recover β and σ.[15] Following the estimation of the 2SLS model, we then set each $\beta_{w,q}$ such that each micro-moment condition is exactly met.

[15] We address price endogeneity through the same instrument (group-market premiums) as in Section 4.1. We additionally include the square of the group-market premiums to instrument for p_{jt}^2.

We then refine these estimates through a Newton-Raphson correction. This involves explicitly solving for structural errors ξ_{jt} through the Berry inversion, incorporating the exact structural errors into our dependent variable, and iteratively re-estimating the model parameters. This correction substantially reduces approximation biases from the linearized first-stage estimation.[16]

5.3 Counterfactual

Our primary counterfactual considers level premiums with individual underwriting of lapsation risk.[17] For each insurer (firm) j and simulated average wealth quantile ω, we compute the firm's expected discounted cost per unit of death benefit as:

$$\text{ExpectedCost}_j(\omega) = \sum_{t=1}^{T} \frac{\text{DeathProb}_j(\omega) \cdot \text{DeathPayout}_t}{(1+r)^t}, \tag{38}$$

where the death payout is normalized to $\$1$ per unit of coverage and discounting is applied at an annual rate $r = 3\%$.

We then express the firm's cost at each hypothetical wealth level relative to its cost at the firm's actual wealth level, ω_{jt}:

$$\text{CostAdjustment}_{jt}(\omega) = \frac{\text{ExpectedCost}_{jt}(\omega)}{\text{ExpectedCost}_{jt}(\omega_{jt})}. \tag{39}$$

To derive counterfactual prices, we adjust the observed equilibrium premium by incorporating the relative cost differences. Specifically, letting p_{jt} denote the residualized prices and p_{jt}^{raw} the raw observed price, the counterfactual price is given by:

$$p_{jt}(\omega) = p_{jt} + \Big(\text{CostAdjustment}_{jt}(\omega) - 1\Big) p_{jt}^{\text{raw}}. \tag{40}$$

Given the estimated mean utility parameters δ_{jt}, the base price sensitivity $\beta_{p,0}$, and wealth

[16] See Salanié and Wolak (2019) for further details.

[17] This approach avoids the complications of income dynamics associated with time-varying premiums while addressing the core equity tradeoff in selection.

ventile parameters $\beta_{w,q}$, we simulate consumers at different wealth quantiles and with different price elasticities and integrate over choice probabilities under the observed and counterfactual price regimes. Under observed prices, the choice probabilities are defined as:

$$\widehat{s}_{jt}^{\text{orig}}(\omega) = \frac{\exp\Big(\delta_{jt} + \beta_{p,i}\, p_{jt} + \beta_{w,i}\Big)}{1 + \sum_k \exp\Big(\delta_{kt} + \beta_{p,i}\, p_{kt} + \beta_{w,i}\Big)}, \tag{41}$$

and under the counterfactual prices:

$$\widehat{s}_{jt}^{\text{cf}}(\omega) = \frac{\exp\Big(\delta_{jt} + \beta_{p,i}\, p_{jt}(\omega) + \beta_{w,i}\Big)}{1 + \sum_k \exp\Big(\delta_{kt} + \beta_{p,i}\, p_{kt}(\omega) + \beta_{w,i}\Big)}. \tag{42}$$

Consumer surplus is computed using the log-sum formulation:

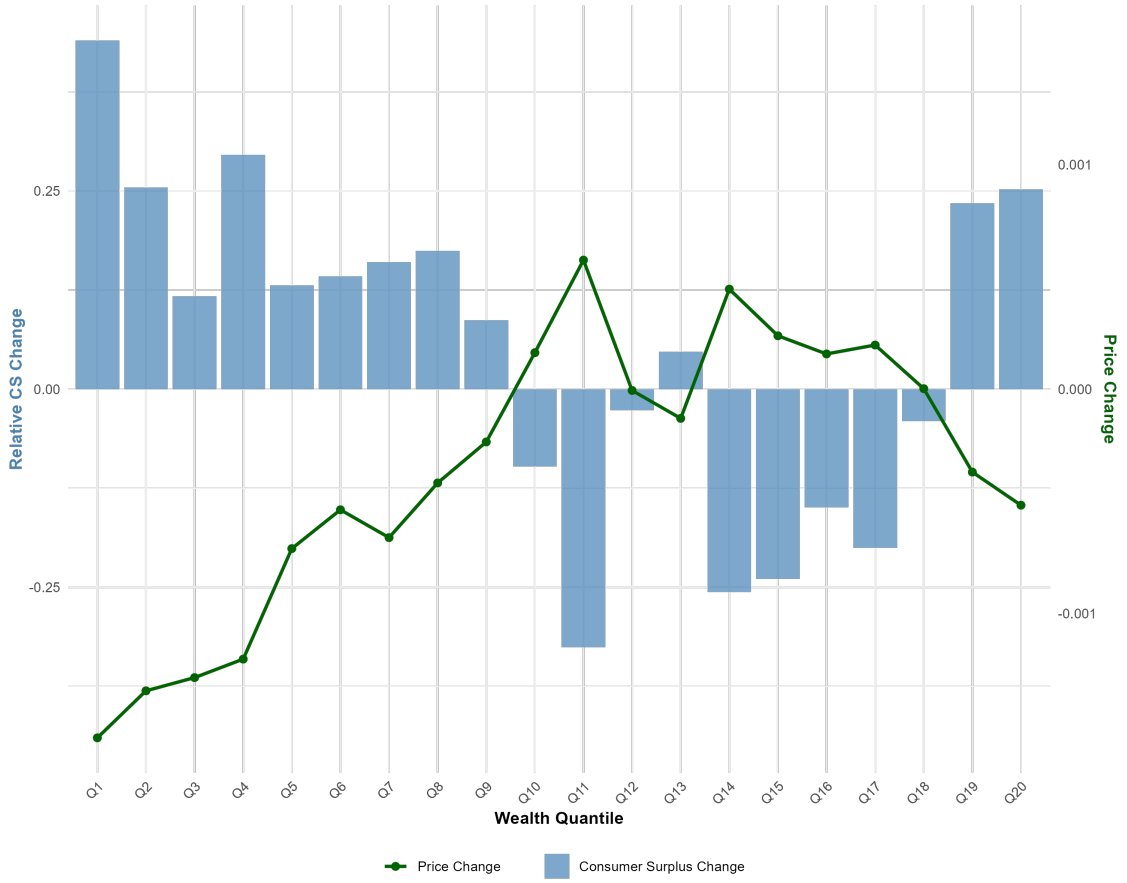
$$CS(\omega) = \frac{1}{|\beta_p|} \log\left(1 + \sum_j \exp\Big(\delta_{jt} + \beta_{p,i}\, p_{jt}(\omega) + \beta_{w,i}\Big)\right). \tag{43}$$

Finally, we aggregate predicted consumer choices across wealth quantiles—weighted by market sizes—and summarize changes in consumer surplus, prices paid, and market participation, as indicated by shifts into or out of the outside option.

Figure 8 illustrates the relative distributional impact of implementing individual lapse-based pricing in the life insurance market. The findings reveal a largely progressive redistribution of consumer surplus across wealth groups, with significant gains for lower-wealth households and corresponding losses for higher-wealth households.

The most prominent result is the substantial increase in consumer surplus for households in the lowest wealth quantiles. Specifically, the bottom five wealth ventiles (Q1–Q5) account for approximately 80% of the overall positive change in consumer surplus, with the lowest ventile receiving the largest transfer. This progressive redistribution arises directly from addressing adverse selection caused by policy lapsation, as lower-wealth households, typically characterized by higher lapse rates, would experience significantly reduced premiums under risk-based pricing. As wealth grows, the relative impact of price changes on consumer surplus becomes larger, as

Figure 8: Wealth-Based Distributional Impacts of Risk-Based Pricing



Note: This figure presents the relationship between wealth and the distributional impacts of risk-based pricing in the life insurance market. The bars (left axis) show the relative consumer surplus change by wealth quantile, with each quantile's change expressed as a proportion of the total consumer surplus change. The line (right axis) displays the corresponding average price-paid change for consumers. Wealth quantiles range from Q1 (lowest wealth) to Q20 (highest wealth). The results are derived from an analytical choice model that accounts for wealth-specific price elasticities and product preferences. Price changes reflect modifications based on wealth-specific lapse behavior, while consumer surplus captures the overall welfare impact incorporating both price effects and potential switching to alternative products or the outside option.

more people as in the market.

As household wealth increases, this pattern gradually shifts. Upper middle wealth ventiles (Q10–Q18) experience negative changes in consumer surplus, effectively subsidizing many of the gains received by lower-wealth households. This aligns with the observed trend in price changes across wealth quantiles—lower-wealth households would pay significantly less under a risk-based system, whereas wealthier households would face higher premiums. Interestingly, the

highest wealth households would also see a net gain. This is because the lapse rates, as reported in Figure 7, is non-monotonic.

These empirical results support our theoretical framework regarding adverse selection driven by policy lapsation. Under the existing flat-premium structure, a regressive subsidy exists where lower-wealth households, characterized by higher lapse probabilities, implicitly subsidize wealthier households, who tend to retain policies for longer periods. Introducing individual lapse-based pricing would correct this distortion, making life insurance more affordable and accessible for financially vulnerable households.

6 Conclusions

We examine lapsing behavior as a channel for selection in life insurance markets. Using insurance regulatory data, we find strong evidence of adverse selection. Further, we find that this selection is driven solely by lapse behavior, with mortality having no influence. We find that a 1% increase in premiums leads to a 0.15% decrease in lapse rates, while having no effect on mortality. This decrease in lapse rates leads to a 0.13% increase in future costs. Adverse selection reduces market size by 17% and creates an annual net present value for the dead-weight losses of $1.3 billion for each new generation of policyholders.

Our findings challenge the conventional wisdom that life insurance markets exhibit little (or potentially advantageous) selection. By demonstrating that lapse behavior is a primary, and previously ignored, selection channel, we help to explain conflicting results in the literature. We also show that sophisticated consumer knowledge about their own risk type is not necessary for this adverse selection to occur and detail that, unlike other selection markets, the distributional effects are also negative.

Our analysis does have limitations. The reduced-form approach likely understates welfare costs by assuming price changes affect all contracts rather than just new ones. Additionally, our lack of individual contract data limits our ability to model to other contract features that would

more accurately allow us to estimate demand.

Our theoretical contribution extends beyond life insurance markets. By showing how selection can operate through multiple channels with opposing effects, we provide a framework for analyzing complex selection patterns in other insurance markets where contracts span multiple periods and early termination is possible, such as long-term care insurance or annuities. We suggest two potential policy solutions: incorporating individual-based lapse estimates into pricing or shifting from flat premiums to rates that reflect contemporaneous mortality risk. Both changes would benefit low-income consumers by reducing cross-subsidization of wealthy policyholders who maintain long-term coverage. These reforms would also reshape the secondary life insurance market, particularly affecting life settlements and stranger-originated life insurance schemes.

References

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge (2023). When should you adjust standard errors for clustering? The Quarterly Journal of Economics 138(1), 1–35.

Ai, M. and P. Steinorth (2024, March). Revisiting selection in life insurance. Unpublished manuscript. German Socio-Economic Panel (SOEP).

Aizawa, N. and Y. S. Kim (2018). Advertising and risk selection in health insurance markets. American Economic Review 108(3), 828–867.

Ameriks, J., A. Caplin, S. Laufer, and S. Van Nieuwerburgh (2011). The joy of giving or assisted living? using strategic surveys to separate public care aversion from bequest motives. The Journal of Finance 66(2), 519–561.

Bernheim, B. D. (1991). How strong are bequest motives? evidence based on estimates of the demand for life insurance and annuities. Journal of Political Economy 99(5), 899–927.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. Econometrica: Journal of the Econometric Society, 841–890.

Bissonnette, L., M. D. Hurd, and P.-C. Michaud (2017). Individual survival curves comparing subjective and observed mortality risks. Health Economics 26(12), e285–e303.

Carson, J. M., C. M. Ellis, R. E. Hoyt, and K. Ostaszewski (2020). Sunk costs and screening: two-part tariffs in life insurance. Journal of Risk and Insurance 87(3), 689–718.

Carson, J. M. and M. D. Forster (2000). Suitability and life insurance policy replacement: An analytical tool. Journal of Insurance Regulation 18(4), 427–447.

Cawley, J. and T. Philipson (1999). An empirical examination of information barriers to trade in insurance. American Economic Review 89(4), 827–846.

Chen, R., K. A. Wong, and H. C. Lee (2001). Age, period, and cohort effects on life insurance purchases in the us. Journal of Risk and Insurance, 303–327.

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler (2016). The association between income and life expectancy in the United States, 2001-2014. Jama 315(16), 1750–1766.

Cohen, A. and P. Siegelman (2010). Testing for adverse selection in insurance markets. Journal of Risk and insurance 77(1), 39–84.

Collier, B. L. and M. A. Ragin (2020). The influence of sellers on contract choice: Evidence from flood insurance. Journal of Risk and Insurance 87(2), 523–557.

Crail, C. (2023). Life insurance statistics, data and industry trends 2023. Online report, Forbes Advisor. https://www.forbes.com/advisor/life-insurance/life-insurance-statistics/#facts-about-life, accessed 11/9/2023.

Cutler, D. M., A. Finkelstein, and K. McGarry (2008). Preference heterogeneity and insurance markets: Explaining a puzzle of insurance. The American Economic Review 98(2), 157–162.

Einav, L., A. Finkelstein, and M. R. Cullen (2010). Estimating welfare in insurance markets using variation in prices. The Quarterly Journal of Economics 125(3), 877.

Eling, M. and M. Kochanski (2013). Research on lapse in life insurance: What has been done and what needs to be done? The Journal of Risk Finance 14(4), 392–413.

Fang, H., M. P. Keane, and D. Silverman (2008). Sources of advantageous selection: Evidence from the medigap insurance market. Journal of Political Economy 116(2), 303–350.

Fang, H. and E. Kung (2010). How does life settlement affect the primary life insurance market? NBER Working Paper Series.

Fang, H. and E. Kung (2021). Why do life insurance policyholders lapse? the roles of income, health, and bequest motive shocks. Journal of Risk and Insurance 88(4), 937–970.

Fei, W., C. Fluet, and H. Schlesinger (2015). Uncertain bequest needs and long-term insurance contracts. Journal of Risk and Insurance 82(1), 125–148.

Fier, S. G., K. A. McCullough, and J. M. Carson (2013). Internal capital markets and the partial adjustment of leverage. Journal of Banking & Finance 37(3), 1029–1039.

Finkelstein, A. and J. Poterba (2004). Adverse selection in insurance markets: Policyholder evidence from the uk annuity market. Journal of Political Economy 112(1).

Gao, J., S. Ge, L. Schmidt, and C. J. T. Trillo (2023). How do health insurance costs affect firm labor composition and technology investment? Working Paper.

Ge, S. (2022). How do financial constraints affect product pricing? evidence from weather and life insurance premiums. The Journal of Finance 77(1), 449–503.

Geruso, M., T. Layton, and A. Leive (2023). The incidence of adverse selection: Theory and evidence from health insurance choices. Technical report, National Bureau of Economic Research.

Giesbert, L., S. Steiner, and M. Bendig (2011). Participation in micro life insurance and the use of other financial services in ghana. Journal of Risk and Insurance 78(1), 7–35.

Gottlieb, D. and K. Smetters (2021). Lapse-based insurance. American Economic Review 111(8), 2377–2416.

Gropper, M. J. and C. M. Kuhnen (2025). Wealth and insurance choices: Evidence from US households. The Journal of Finance. forthcoming.

Hackmann, M. B., J. T. Kolstad, and A. E. Kowalski (2015). Adverse selection and an individual mandate: When theory meets practice. American Economic Review 105(3), 1030–1066.

Harris, T. and A. Yelowitz (2014). Is there adverse selection in the life insurance market? evidence from a representative sample of purchasers. Economics Letters 124(3), 520–522.

Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In The Economics of New Goods. University of Chicago Press.

He, D. (2009). The life insurance market: Asymmetric information revisited. Journal of Public Economics 93(9-10), 1090–1097.

He, D. (2011). Is there dynamic adverse selection in the life insurance market? Economics Letters 112(1), 113–115.

Hedengren, D. and T. Stratmann (2016). Is there adverse selection in life insurance markets? Economic inquiry 54(1), 450–463.

Hurd, M. D. and K. McGarry (2002). The predictive validity of subjective probabilities of survival. The Economic Journal 112(482), 966–985.

Jansen, M., F. Nagel, C. Yannelis, and A. L. Zhang (2022). Data and welfare in credit markets. Technical Report 2022-88, University of Chicago, Becker Friedman Institute for Economics.

Jayachandran, S. and A. Lleras-Muney (2009, 02). Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines*. The Quarterly Journal of Economics 124(1), 349–397.

Kang, Z. Y. and S. Vasserman (2022). Robust bounds for welfare analysis. NBER Working paper (w29656).

Kinge, J. M., J. H. Modalsli, S. Øverland, H. K. Gjessing, M. C. Tollånes, A. K. Knudsen, V. Skirbekk, B. H. Strand, S. E. Håberg, and S. E. Vollset (2019). Association of household income with life expectancy and cause-specific mortality in Norway, 2005-2015. Jama 321(19), 1916–1925.

Koijen, R. S. and M. Yogo (2015). The cost of financial frictions for life insurers. American Economic Review 105(1), 445–475.

Koijen, R. S. and M. Yogo (2016). Shadow insurance. Econometrica 84(3), 1265–1287.

Kuo, W., C. Tsai, and W.-K. Chen (2003). An empirical study on the lapse rate: The cointegration approach. Journal of Risk and Insurance 70(3), 489–508.

Lai, G. C., H. Nakamura, S. Yamamoto, and T. Yoneyama (2021). Adverse retention: Strategic renewal of guaranteed renewable term life insurance policies. Journal of Risk and Insurance 88(4), 1001–1022.

Liebenberg, A. P., J. M. Carson, and R. E. Dumm (2012). A dynamic analysis of the demand for life insurance. Journal of Risk and Insurance 79(3), 619–644.

LIMRA (2021). Facts about life 2021. Technical report.

Linton, M. (1932). Panics and cash values. Transactions of the Actuarial Society of America 33, 265–394.

Lockwood, L. M. (2018). Incidental bequests and the choice to self-insure late-life risks. American Economic Review 108(9), 2513–2550.

McCarthy, D. and O. S. Mitchell (2010). International adverse selection in life insurance and annuities. Springer.

Mulholland, B. S. and M. S. Finke (2014). Does cognitive ability impact life insurance policy lapsation? Available at SSRN 2512076.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. Econometrica 69(2), 307–342.

Nolte, S. and J. C. Schneider (2017). Don't lapse into temptation: A behavioral explanation for policy surrender. Journal of Banking & Finance 79, 12–27.

Outreville, J. F. (1990). Whole-life insurance lapse rates and the emergency fund hypothesis. Insurance: Mathematics and Economics 9(4), 249–255.

Pesando, J. E. (1974). The interest sensitivity of the flow of funds through life insurance companies: an econometric analysis. Journal of Finance 29(4), 1105–1121.

Rothschild, C. (2024). Welfare analysis in insurance markets. The Geneva Risk and Insurance Review, 1–23.

Salanié, B. and F. A. Wolak (2019). Fast," robust", and approximately correct: estimating mixed demand systems. Technical report, National Bureau of Economic Research.

Schott, F. H. (1971). Disintermediation through policy loans at life insurance companies. The Journal of Finance 26(3), 719–729.

Shapiro, B. T. (2020). Advertising in health insurance markets. Marketing Science 39(3), 587–611.

Shefrin, H. (2002). Beyond Greed and Fear: Understanding Behavioral Finance and the Psychology of Investing. Oxford University Press on Demand.

Truett, D. B. and L. J. Truett (1990). The demand for life insurance in mexico and the united states: A comparative study. Journal of Risk and Insurance, 321–328.

Wenning, D. (2024). Equal prices, unequal access: The effects of national pricing in the us life insurance industry. Working Paper.

Wuppermann, A. C. (2017). Private information in life insurance, annuity, and health insurance markets. The Scandinavian Journal of Economics 119(4), 855–881.

Appendix A Description of Back-of-the-Envelope Estimation of NPV of Deadweight Loss.

The net present value (NPV) of deadweight losses is determined over a 30-year time horizon. This calculation accounts for policy exits—such as lapses, deaths, and surrenders—and applies a 3% discount rate. The first step is to compute the average exit rate, λ_{exit}, which is the sum of the individual exit rates:

$$\lambda_{exit} = \lambda_{lapse} + \lambda_{death} + \lambda_{surrender} \tag{A1}$$

Next, the survival probability, S_t, which represents the probability that a policy remains active up to period t, is given by:

$$S_t = (1 - \lambda_{exit})^{t-1}. \tag{A2}$$

The discount factor, D_t, which adjusts future values to their present value, is calculated as:

$$D_t = \left(\frac{1}{1.03}\right)^{t-1}. \tag{A3}$$

The combined NPV factor—representing the expected present value of \$1 received annually over 30 years—is obtained by summing the product of the survival probability and the discount factor:

$$\text{Combined Factor} = \sum_{t=1}^{30} S_t \cdot D_t. \tag{A4}$$

Finally, the NPV of deadweight losses is calculated by multiplying the annual deadweight loss (DWL), as determined from the structural model, by the combined factor:

$$\text{NPV_DWL} = \text{DWL} \cdot \text{Combined Factor}. \tag{A5}$$

Appendix B Description of Consumer Finance Data

B.1 Survey of Consumer Finances Data

In panel (a) of Figure 1, we consider the share of households with life insurance by income relative to the federal poverty line. Income is measured in the Survey of Consumer Finances (SCF) as variable x5729, the total income of the household before taxes and other deductions. The federal poverty line, adjusted by the number of people in the household, was taken from the Office of the Assistant Secretary for Planning and Evaluation. For 2016, the adjustment by family members was variable and we used the average value of $4,140. For this and all other figures using data from the SCF, we limit teh analysis to households whose repsondent was between 18 and 64 to limit the influence of tax considerations on life insurance demand.

In panel (b) of Figure 1, we consider the share of households with life insurance by wealth quintile. For the figure, we use the last four waves (2013, 2016, 2019, and 2022) of the SCF. Wealth is defined as the sum of all assets (both liquid and illiquid) minus the sum of all debt. We list all variables used in the construction of the different types of assets and debt in Table B1.

The general trends taken from Figure 1 are robust to alternative definitions of wealth or poverty. In Figure B1 below, we report the same diagram, but use two alternative definitions. Panel (a) uses the answer to question x7362 in the SCF for income. This question asks what the total income of the household in a "normal" year would be, thus correcting for unexpected shocks to income. Panel (b) uses only liquid assets as a definition of wealth. Both panels show the same general trends as the figure in the main text. Insurance demand is increasing in both wealth and income and decreasing over time.

B.2 Health and Retirement Survey Data

To analyze how consumers' tendencies to lapse correlate with wealth in Figure 7, we use data from the Health and Retirement Survey (HRS) rather than data from the SCF. With the exception of a special wave during the financial crisis, the SCF is not a panel and does not ask consumers whether

Table B1: Variables used for constructing wealth in the Survey of Consumer Finances

Wealth type	Variables in the SCF
Liquid Assets	
Checking accounts, including prepaid payment cards	x3506, x3510, x3514, x3518, x3522, x3526, x3529, x7596
IRAs and Keogh accounts	x6551, x6559, x6567, x6552, x6560, x6568, x6553, x6561, x6569, x6554, x6562, x6570, x6756, x6757, x6758
Savings and money market accounts	x3730, x3736, x3742, x3748, x3754, x3760
Mutual funds	x3822, x3824, x3826, x3828, x3830, x7787
Bonds	x3902, x7635, x7637, x7636, x7638, x7639
Stocks	x3915
Certificates of deposit	x3721
Illiquid Asset	
Real Estate	x513, x526, x604, x614, x623, x716, x1306, x1325, x1339, x1705/10000*x1706, x1805/10000*x1806, x2002, x2012
Businesses	x3125, x3224, x3129, x3229,x3335, x3408, x3412, x3452, x3416, x3420, x3428, x8452
Vehicles	x8166, x8167, x8168, x8188, x2422, x8430, x2506, x2606, x2623, x8435
Other assets	x3930, x6577, x3960, x4018, x4022, x4026, x4030
Debt	
Mortgages and real estate loans	x805, x905, x1005, x1044, x1108, x1119, x1130, x1136, x1215, x1219, x1715, x1815, x2006, x2016
Vehicle loans	x2218, x2318, x2418, x7169, x2424, x8432, x2519, x2619, x2625, x8437
Consumer debt (incl. credit cards)	x413, x421, x427, x443, x7575, x2723, x2740, x2823, x2840, x2923, x2940, x7183, x8443
Student loans	x7824, x7847, x7870, x7924, x7947, x7970, x7179, x8440
Other debt	x3932, x4032

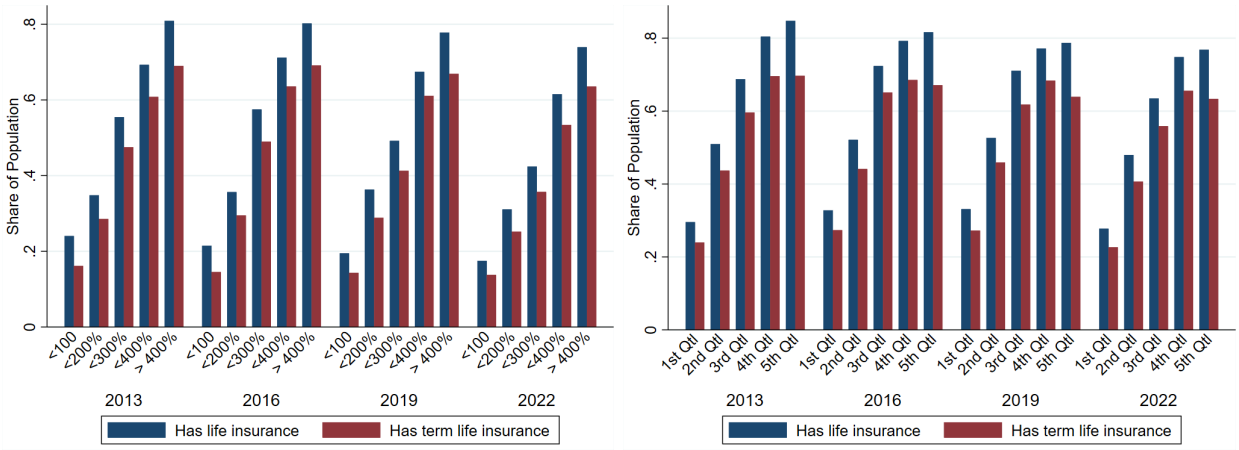
Notes: The table lists all wealth categories and the corresponding variables in the SCF used for the construction of the wealth variable in Figure 1.

they have lapsed a life insurance policy in some given time period. This makes an analysis of life insurance lapse decisions impossible except for doing it in a very specific and probably not representative economic environment. In the HRS, respondents are surveyed in a panel structure and asked whether they have lapsed a life insurance policy, such that the data are more suited for the analysis. The biggest downside of the HRS for our purpose is that, by design, it oversamples older individuals. We deal with this limitation, by restricting our analysis to respondents who are

Figure B1: Life Insurance Demand by Income and Wealth - Alternative Measures

(a) Normal Income relative to the federal poverty line

(b) Quintile of liquid assets



Note: Figure repeats the analysis of panel (b) in Figure 1 using alternative definitions of wealth and poverty. The graphs show the amount of households with life insurance coverage for any type of policy (blue) or term life (red) in the US by quintile of liquid assets (panel (a)) or by income relative to the federal overty line (panel (b)). Data is taken from four waves of the Survey of Consumer Finances, see Table B1 for details.

64 and younger. As in the analysis of the SCF, this limits the relevance of tax considerations in the life insurance demand and lapse decisions of the respondents.

For the construction of the sample, we utilize the last six waves of the RAND version of the HRS, which are 2010, 2012, 2014, 2016, 2018 and 2020. Because we use lagged variables, we also use data from 2008. For wealth, we use the variable "atotb", which is the primary wealth proxy provided by RAND. Ventiles are defined on this variable for all respondents with age younger or equal 64. For the display in Figure 7, we use the wealth ventile of the previous wave to sort respondents, because the lapse decision was made between the previous and the current wave. Note that data on life insurance lapses are not part of the RAND HRS panel and need to be merged into the panel from the raw HRS files. Because lapsing is only possible if respondents already had a life insurance policy and because life insurance ownership is influenced by wealth (see Figure 1), we restrict the sample to those individuals which reported having a life insurance policy in the previous wave.

Appendix C Robustness to Functional Form Assumptions in Akerlovian Selection Markets

In this appendix, we describe the approach we employ to assess the robustness to functional assumptions when calculating the dead-weight loss in Akerlovian selection markets. We begin by describing the problem generally and then exemplify the procedure using the original estimates of Einav et al. (2010, henceforth EFC). Note that some of the details stated below also appear in the main text, but are repeated here to allow for a stand-alone introduction of the method.

Kang and Vasserman (2022, henceforth KV) raise an important issue about the robustness of welfare inference in settings where demand functions are estimated based upon few (or two) points: when researchers can only measure a few points on the demand curve, welfare implications must be extrapolated from average treatment effects through functional form assumptions. Our setting is an especially pronounced example. For every insurance company, in every market, in every year; we infer welfare effects through counterfactual statements extrapolated from two slope measurements and a firm-market-year specific data point. Our counterfactuals are thus identified at not few, but one data point.

KV suggest a procedure for quantifying the robustness of the welfare implications of a limited estimation. In their procedure, they allow for arbitrary functional forms under the restriction that between two observed points, the local slope does not deviate from the average slope between these two points by more that a proportional factor r. The researcher then finds the highest r^* such that, within this limited set of functions, the welfare implications of their preferred functional form could flip from positive to negative (or vice-versa). Higher values of r^* then indicate more robustness, with an $r^* = 0$ implying any deviation of the researcher's assumed gradient between identified points could reverse the welfare implications.

Unfortunately, this approach does not apply to Akerlovian selection markets. As long as the demand curve slopes down and the cost curve slopes up, dead-weight loss (DWL) will be positive and only the magnitude will vary, regardless of r.[18] While KV do extend their functional

[18]Note that we define the cost curve as a function of premiums. When plotted against quantity, both curves slope

bounds on the demand curve to bounding DWL, Akerlovian selection markets instead require two functional forms—both the demand and cost curve—to determine DWL, and these functions interact in non-trivial ways (see Equation (A8) below).

We thus propose an alternative robustness measure for functional form assumptions when estimating DWL in an EFC framework. In the same spirit as KV, we stick to the idea that different functional forms lead to different welfare outcomes. However, we seek only to bound the possible outcomes, rather than find a critical value that flips the sign. To solve the issue of the complicated interactions between demand and cost curves, we restrict the functional forms of each curve to second-order polynomials and solve for the bounds numerically. While still allowing us to consider both concave and convex functions for both curves, this assumption considerably simplifies the problem's numerical complexity.

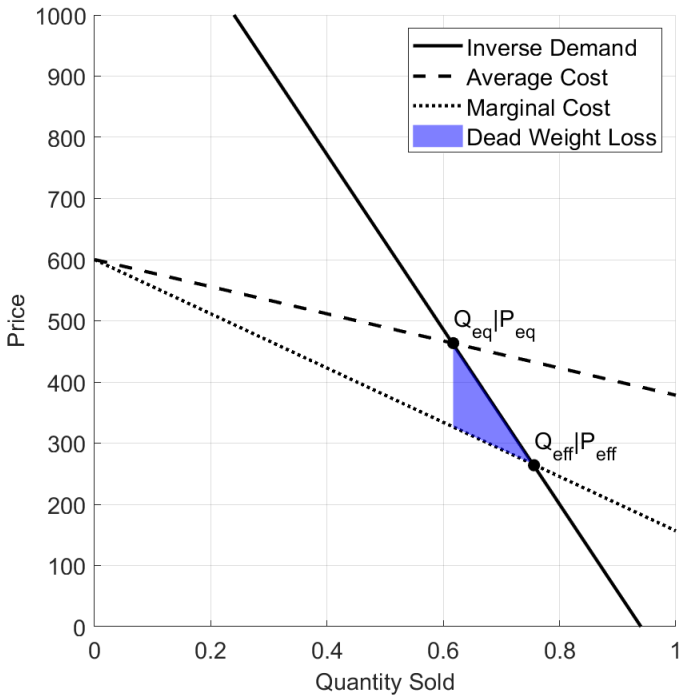
In what follows, we assume that, for a given market, we observe the current competitive equilibrium of price (P_{eq}) and quantity sold (Q_{eq}). We also assume that we observe the slopes of the demand curve ($D(P)$) and the average cost curve ($AC(P)$) at this point, and only at this point. For simplicity and in congruence with the notation in EFC, we denote the slope of the demand curve as $\frac{\partial D(P)}{\partial P} = \beta$ and the slope of the average cost curve as $\frac{\partial AC(P)}{\partial P} = \delta$. Note that this deviates from the definition of β and δ in Section 4.1. We discuss how to translate to our specific setting at the end of this appendix.

To exemplify our procedure, we will use the estimated slopes and equilibrium point of EFC. Figure C2 shows the health insurance market analyzed by EFC under the linearity assumptions used in their study. We denote the (inverse) demand function, that is the function that maps quantities to prices, as $P(Q)$. While the actual estimated cost function maps prices to cost, we follow custom and show cost functions mapping quantities to cost, which we denote $AC(Q)$ and $MC(Q)$.

As long as demand is decreasing and average costs are increasing in price (and thus decreasing in quantity, given the demand assumption), Figure C2 makes clear that the relevant price range for calculating DWL is given by all positive prices below the equilibrium price. From this, the

down.

Figure C2: Market situation and dead-weight loss estimate in EFC



Note: Figure shows the inverse demand curve and average cost curve from the estimates in EFC. Marginal costs and dead-weight loss are inferred from the estimated curves. The marked points are the competitive equilibrium and the efficient market outcome.

observed equilibrium and the measured slope, we have three conditions on the demand function:

(i) $\frac{\partial D(P)}{\partial P} < 0; \forall\ P \in]0, P_e q]$,

(ii) $D(P_{\text{eq}}) = Q_{\text{eq}}$ and

(iii) $\frac{\partial D(P)}{\partial P}\Big|_{P=P_{\text{eq}}} = \beta$.

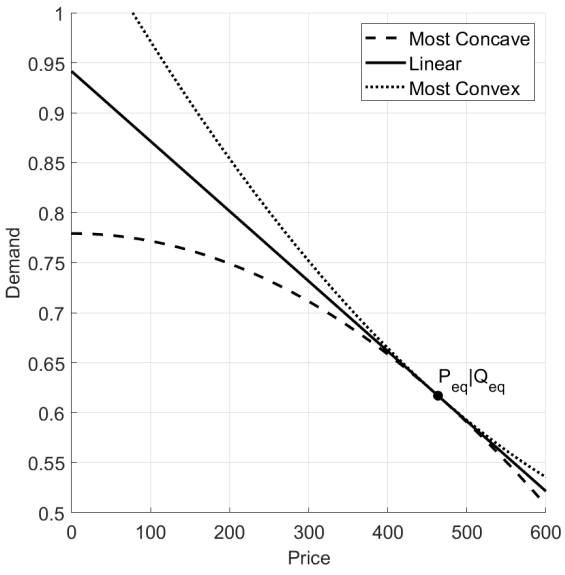
We consider second-order polynomial demand functions of the form $D(P) = a_D + b_D P + c_D P^2$. The slope is given by $\frac{\partial D(P)}{\partial P} = b_D + 2c_D P$. From condition (iii), we thus know that $b_D = \beta - 2c_D P_{\text{eq}}$. Because $\beta < 0$, condition (i) will never be binding when $c_D > 0$, that is when $D(P)$ is convex. When $c_D < 0$, the most concave admissible demand curve is going to have its highest slope at $P = 0$. Condition (iii) can thus be used to find the smallest admissible value of c_D, which we will denote $\underline{c_D}$. Substituting from above, the condition implies that $\underline{c_D} = \frac{\beta}{2P_{\text{eq}}}$. From this, we can

identify the values of a_D and b_D for the most concave function, using $b_D(\underline{c_D}) = \beta - 2\underline{c_D}P_{\text{eq}}$ and $a_D(\underline{c_D}) = Q_{\text{eq}} - b_D(\underline{c_D})P_{\text{eq}} - \underline{c_D}P_{\text{eq}}^2$.

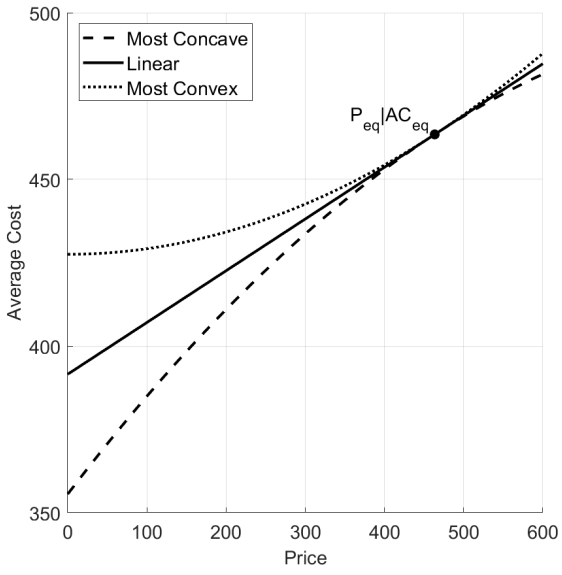
Because demand functions can be infinitely convex and still fulfill conditions (i) through (iii), we need a sensible upper bound for the range of functions to consider. Given that curvature of the polynomial is given by c_D and that the most concave function has curvature $\underline{c_D}$, we implement the upper curvature bound of $\overline{c_D} = -\underline{c_D}$. $a_D(\overline{c_D})$ and $b_D(\overline{c_D})$ are defined analogously to their most concave counterparts. These two functions then define our space of admissible demand functions and are displayed in panel (a) of Figure C3. Functions between the two extremes are constructed by a grid between $\underline{c_D}$ and $\overline{c_D}$ and the implied values for a_D and b_D are calculated for each value in the grid.

Figure C3: Ranges of demand and cost curves used in robustness analysis

(a) Range of demand curves



(b) Range of cost curves



Note: Figures show the range of possible demand curves and cost curves that are used in the robustness analysis. The procedure developed here varies the curvature of the the demand curves between the two extremes in a discrete grid.

We use a similar procedure as for the demand curves to define the admissible average cost curves. All cost curves are of the form $AC(P) = a_C + b_C P + c_C P^2$. The three conditions for the

functions are given by

(i) $\frac{\partial AC(P)}{\partial P} > 0; \forall\ P \in]0, P_{\text{eq}}]$,

(ii) $AC(P_{\text{eq}}) = P_{\text{eq}}$ and

(iii) $\frac{\partial AC(P)}{\partial P}\Big|_{P=P_{\text{eq}}} = \delta$.

Because these curves need to have a positive slope, the most convex function has to be bound by condition (iii) implying that $\overline{c_C} = \frac{\delta}{2P_{\text{eq}}}$. It then follows that $b_C(\overline{c_C}) = \delta - 2\overline{c_C}P_{\text{eq}}$ and $a_C(\overline{c_C}) = (1 - b_C(\overline{c_C}))P_{\text{eq}} - \overline{c_C}P_{\text{eq}}^2$. Having defined the average cost curve with the highest admissible curvature, we again define the opposite of the range via $\underline{c_C} = -\overline{c_C}$. Corresponding parameters and the grid of functions are defined as for the demand functions. The range of average cost curves is displayed in panel (b) of Figure C3.

EFC use the linear functional form assumption to measure the dead-weight loss as $DWL = \frac{1}{2}(Q_{\text{eff}} - Q_{\text{eq}})(P_{\text{eq}} - MC(P_{\text{eq}})$. Without the luxury of linear functional forms, we instead use

$$DWL = \int_{Q_{\text{eq}}}^{Q_{\text{eff}}} P(Q)dQ - \int_{Q_{\text{eq}}}^{Q_{\text{eff}}} MC(Q)dQ. \tag{A6}$$

Q_{eq} is known, so we require the derivation of Q_{eff}, $P(Q)$, and $MC(Q)$. $P(Q)$ is simply given by the inverse of $D(P)$. We can use the polynomial form to derive the inverse demand curve. It is given by

$$P(Q) = \begin{cases} \frac{-b_D}{2c_D} - \frac{\sqrt{b_D^2 - 4a_D c_D + 4c_D Q}}{2c_D} & \text{for } c_D \neq 0 \\ \frac{-a_D}{b_D} + \frac{1}{b_D}Q & \text{for } c_D = 0 \end{cases} \tag{A7}$$

Because $AC(Q)$ maps quantities to cost, we can write $AC(Q) = AC(P(Q))$. Marginal cost as a function of quantities is given by

$$MC(Q) = \frac{\partial AC(Q) \cdot Q}{\partial Q} = \frac{\partial AC(P(Q))}{\partial P(Q)} \cdot \frac{\partial P(Q)}{\partial Q} \cdot Q + AC(P(Q)). \tag{A8}$$

Finally, Q_{eff} is given by the implicit definition $P(Q_{\text{eff}}) = MC(Q_{\text{eff}})$.

An important observation from Equation (A8) is that by the guidelines of our robustness

analysis, namely that β, δ, P_{eq} and Q_{eq} are known, the point $MC(Q_{\text{eq}})$ will not change due to changes on the functional forms of D and AC. We know that $P(Q_{\text{eq}}) = P_{\text{eq}}$ and thus that $\frac{\partial AC(P(Q))}{\partial P(Q)}\Big|_{Q=Q_{\text{eq}}} = \frac{\partial AC(P)}{\partial P}\Big|_{P=P_{\text{eq}}} = \delta$. Further, from the chain rule and the fact that $P(Q)$ is the inverse function of $D(P)$, we know that $\frac{\partial P(Q)}{\partial Q}\Big|_{Q=Q_{\text{eq}}} = 1/\beta$. Thus

$$MC(Q_{\text{eq}}) = \frac{\delta}{\beta}Q_{\text{eq}} + AC(Q_{\text{eq}}) = \frac{\delta}{\beta}Q_{\text{eq}} + P_{\text{eq}}. \tag{A9}$$

The important implication is that a change in the magnitude of DWL due to a change in the functional form of D or AC cannot come from a shift of the marginal cost curve. Changing these function can only tilt the marginal cost curve as is shown in Figure 6 of the main text.

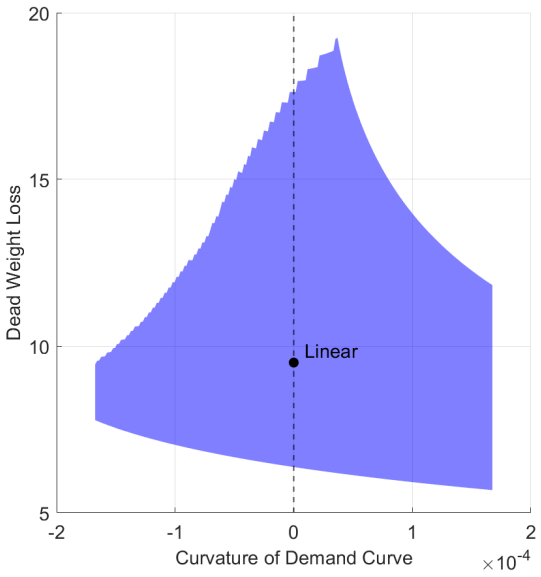
The result of the robustness analysis is shown in Figure C4. The results show that the dead-weight loss can vary widely with the functional form of the cost and demand curve. From the original value of \$9.55 per person reported in EFC, the value can increase by almost 100% or decrease by over 40% allowing for a total range between \$5.67 and \$19.24 per person. From the figure, we can see that the curvature of the cost curve seems to be more influential than that of the demand curve. In fact, the range of possible DWL values is highest when the demand curve is linear or slightly convex. For a linear demand curve, for example, changing the curvature of the cost function can lead to DWL estimates between \$6.34 and \$17.62. No range of this magnitude can be achieved by varying the curvature of the demand curve for any fixed curvature of the cost curve.

The boundary of the DWL range is somewhat ragged in certain areas of both graphs in Figure C4. This is because not all curvature combinations between demand and cost curve support an efficient solution of the market in the sense that there is no intersection of $P(Q)$ and $MC(Q)$ for a positive price and a quantity smaller 1. While one could assume that the market simply clears with $Q = 1$, this assumption seems unrealistic especially when considering the low prices implied by this situation. We thus refrain from reporting a DWL for these situations and restrict the DWL range to those markets in which a solution of the form $P(Q) = MC(Q)$ exists.

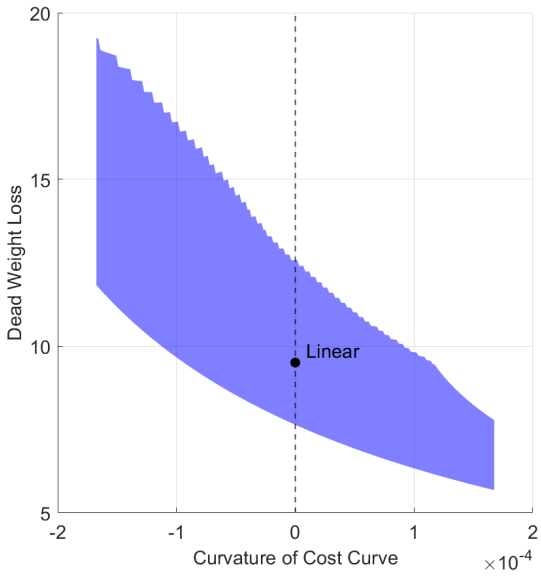
Having described our procedure and exemplified it using the EFC estimates, we now consider

Figure C4: Possible range of the dead-weight loss in EFC by curvature of cost and demand function

(a) Range by curvature of the demand function



(b) Range by curvature of the cost function



Note: Analysis uses the measured slopes and market outcomes of EFC. Curvature of the function is set by allowing the function to be a second degree polynomial and varying the quadratic term within feasible ranges. The linear solution (red dot) is the dead-weight loss implied by linear cost and demand curves as is specified in EFC.

how it can be applied to other estimates of DWL due to adverse selection. The application of the procedure to our own estimates demonstrates that the original functional forms used to estimate the slopes of the demand and cost curves are immaterial beyond their impact on the local estimates of the slope of demand and average cost in P at point $P_{\text{eq}}|Q_{\text{eq}}$. Our estimates of α_{ist}, β, γ_{ist} and δ from Equations (14) and (15) as well as the backed out time-frame and profit-corrected value of $\tilde{\gamma}_{ist}$ imply

$$\left.\frac{\partial Q_{ist}}{\partial P}\right|_{P=P_{\text{eq}}} = e^{\alpha_{ist}} \beta P_{\text{eq},ist}^{\beta-1} \tag{A10}$$

$$\left.\frac{\partial \tilde{AC}_{ist}}{\partial P}\right|_{P=P_{\text{eq}}} = e^{\tilde{\gamma}_{ist}} \delta P_{\text{eq},ist}^{\delta-1}. \tag{A11}$$

Using these slopes and the observed prices and quantities we can apply the robustness procedure to our estimates along the steps of our procedure given above. The results are shown in Figure 5

of the main text.

A bigger caveat for applications to other potential settings is whether the assumptions of an observed equilibrium point ($P_{\text{eq}}|Q_{\text{eq}}$) and slopes measured at exactly this point are actually fulfilled. While that assumption fits our particular setting very well, it can be questioned in other settings. In fact, even EFC observe the slope of the demand and cost function for a set of different prices and infer the equilibrium point from their estimate. While one can make the argument that for the majority of their data, this range is not particularly large, their estimates are certainly not made at only one point. We nevertheless use their estimates for the demonstration of the procedure due to their prominence and the simplicity of the linear estimation. In theory, more points with measured slopes allow for higher order polynomials in the cost and demand curves. We leave the exploration of this possibility for future research.