Equilibria in Health Exchanges: Adverse Selection vs. Reclassification Risk

Ben Handel*, Igal Hendel†, and Michael D. Whinston‡

First version: April 26, 2012

This version: July 3, 2012

PRELIMINARY AND INCOMPLETE

Abstract

This paper studies equilibrium and welfare in a class of regulated health insurance markets known as exchanges. We use detailed health plan choice and utilization data to model individual-level (i) projected health risk and (ii) risk preferences. We combine these micro-foundations with a model of competitive insurance markets that generates predictions for plan prices, costs, and market shares under different counterfactual pricing /contract regulations and several equilbrium solution concepts. We investigate the welfare implications of different pricing regulations, with a focus on (i) adverse selection and (ii) premium re-classification risk.

We find that market unravelling from adverse selection is substatial under the proposed pricing rules in the Affordable Care Act (ACA), implying limited coverage for individuals beyond the lowest tier (Bronze) health plan. Though advese selection can be attenuated by allowing (partial) pricing of health status, the welfare loss from re-classification risk is substatially higher than the gains of increasing coverage. Finally, we compute the subsidies / tax penalties that are required to induce different levels of participation in the exchanges.

*Department of Economics, UC Berkeley; handel@berkeley.edu

†Department of Economics, Northwestern University; igal@northwestern.edu

‡Department of Economics, Northwestern University; mwhinston@northwestern.edu

1 Introduction

Health insurance markets in almost all contexts are subject to a variety of regulations that are intended to enhance the efficient and equitable provision of this important product. The recently passed Patient Protection and Affordable Care Act (PPACA) defines a class of regulated markets called *exchanges* where insurers must offer policies that are one year in duration and fall into one of four pre-defined actuarial equivalence (risk protection) categories. In addition, insurers' pricing is subject to restrictions[1], like offering the same premiums to same age individuals. Subject to these and several other restrictions, each of the 50 U.S. states will have flexibility in the way they set up their state-specific exchanges, which are required to begin operation by 2014. While there has been a great deal of public discussion of health exchanges, there has been little formal analysis of the likely outcomes and welfare impact of alternative designs, like alternative pricing rules or longer contracts.

This paper sets up and empirically investigates a model of insurer competition in a regulated marketplace, motivated by the creation of these exchanges. The primary issue that we address is how changes to the rules governing premiums (say pricing pre-existing conditions) would impact consumer welfare.[2] Specifically, in our main analysis we compare equilibrium outcomes and welfare when insurers are prohibited from pricing based on pre-existing conditions to outcomes and welfare when insurers can price on the basis of some, but not all, relevant health information.

Changes to this regulation would directly impact two distinct phenomena that impact consumer welfare: adverse selection and re-classification risk.[3] Adverse selection is present when there is individual-specific information that can't be priced, and insured individuals select their policies based on this information in a manner adverse to insurers' interests. In health insurance markets this information is typically about the individual's health status, and insured individuals then have a tendency to select greater coverage when they are sicker.[4] Reclassification risk, on the other hand, arises when insurance contracts are of limited duration and changes in health status lead to changes in future premiums. In our setting, reductions in the extent to which premiums can be based on pre-existing conditions are likely to increase the extent of adverse selection, but reduce the reclassification risk that insured individuals face.[5] For example, when pricing based on pre-existing conditions is completely prohibited (which is close to the case in the current regulation), reclassification risk is eliminated but adverse selection is likely to be present. At the other extreme, were unrestricted pricing based on health status

[1] Insurers can vary price based on age, up to a ratio of 3:1 across the range of ages, and can vary price to some extent based on smoking status.

[2] In future versions of this work, we also hope to study regulatory changes to the duration and type of contracts that are permitted.

[3] Each of these phenomena is often cited as a key reason why market regulation is so prevalent in this sector in the first place.

[4] See Akerlof (1970) and Rothschild and Stiglitz (1976) for seminal theoretical work.

[5] Note that exchanges are also working on implementing insurer-level transfers to equate differences in the risk pools that insurers attract. To the extent that such transfers are successful, there will be lower adverse selection in the marketplace conditional on the pricing regulation in question.

allowed, adverse selection would be completely eliminated. We would then expect all individuals to obtain full insurance, although at a very high price when they get sick. Thus, in determining the degree to which pricing of pre-existing conditions should be allowed, a regulator needs to consider the potential trade-off between adverse selection and re-classification risk.[6]

While a great deal of attention has been paid to the potential economic implications of different exchange regulations, minimal attention has been paid to exactly how insurers will compete in markets subject to these regulations, and hence, how equilibria in such markets will look like. To study how our counterfactual variation in regulation will impact efficiency, mediated through adverse selection and re-classification risk, we develop a novel stylized model of an insurance exchange that builds on work by Rothschild and Stiglitz (1976), Wilson (1977), Miyazaki (1977), Riley (1985) and Engers and Fernandez (1987) studying competitive markets with asymmetric information. In the model, the population is characterized by a joint distribution of risk preferences and health risk and there is free entry of insurers. Throughout the analysis, we fix two classes of insurance contracts that each insurer can offer: a more comprehensive contract that has 90% actuarial value, and a less comprehensive contract that has 60% actuarial value.[7] These contracts are required to be annual, as in the current legislation. Finally, in our base specification, insurers are in one case prohibited from pricing pre-existing conditions, and in the other are allowed to partially price based on an individual's health status.

To compute both the static equilibrium and long-run welfare in the marketplace, we overcome several challenges. In the static marketplace, insurers can enter freely and offer both policies that the regulator allows. In such a setting, as shown in Rothschild and Stiglitz (1976), a Nash Equilibrium may not exist. As a result, we focus instead on two other concepts developed in the theoretical literature: Riley equilibria [Riley (1979)] and Wilson equilibria [Wilson (1977)]. Under the Riley notion, when an insurer deviates from the potential equilibrium, other insurers can immediately respond to that deviation with a profitable deviation of their own. If this subsequent deviation renders the initial deviation unprofitable, then it is not considered to be a profitable deviation from the equilibrium. In contrast, in the Wilson equilibrium concept, a deviation may lead to policies being withdrawn if the deviation renders the firms offering those policies unprofitable. We develop results showing that in our model both Riley equilibrium and Wilson equilibrium always exist and characterize them. (Moreover, Riley equilibrium is unique.) We also investigate the existence and properties of Nash equilibria in our model.

We then develop a long-run welfare model that integrates premium risk, conditional on the pricing regulation and underlying risk transitions. This model evaluates welfare from the perspective of an ex-ante unborn individual, and follows an individual through many consecutive one-year markets

[6] We refer to this as a potential trade-off at this point because there is no general theoretical result saying that adverse selection must increase as insurers can price discriminate on less information (and vice-versa). However, this seems to be generally true in the empirical settings that we are aware of, including the one studied in this paper.

[7] Actuarial value reflects the proportion of total expenses that an insurance contract would cover if the entire population were enrolled. The two values selected reflect the most and least comprehensive insurance contrats that are allowed under the exchange legislation (in that legislation 70% and 80% actuarial value contracts can also be offered).

characterized by the static model.

To empirically study the impact of price regulation on adverse selection and re-classification risk we need a population of potential insureds for whom we observe both health status and risk preferences. We obtain this information using individual-level health plan choice and health claims data for the employees of a large firm and their dependents. We leverage several unique features of the data to cleanly identify risk preferences including (i) a year where all employees made active, non-default choices, due to a menu change and (ii) the fact that the plans available differ financially, but not in terms of provider availability. To estimate risk preferences and health risk, we develop a structural choice model [closely following Handel (2011)] that jointly quantifies risk preferences and ex ante health risk while netting out switching costs.[8] In the model, consumers make choices that maximize their expected utilities over all plan options conditional on their risk tastes and health risk distributions.[9] We allow for heterogeneity in risk preferences so that we have the richest possible understanding of how consumers select plans.

To model health risk perceived by employees at the time of plan choice, we use the methodology developed in Handel (2011), which characterizes both total cost health risk and plan-specific out-of-pocket expenditure risk. The model incorporates past diagnostic and cost information into individual-level and plan-specific expense projections using both (i) sophisticated predictive software developed at Johns Hopkins Medical School and (ii) a detailed model of how different types of medical claims translate to out-of-pocket expenditures in each plan. The cost model outputs a family-plan-time specific distribution of predicted out-of-pocket expenditures that we incorporate into an expected utility model under the assumption that consumer beliefs about future health expenditures conform to our cost model estimates. This cost prediction framework directly advances our goals of (i) precisely quantifying risk preferences and (ii) understanding how health risk evolves over both the short and long run.

We use the estimates from the model to study market equilibria and long-run welfare in counterfactual market environments with the same stylized features as the theoretical model discussed above. While we realize that our sample, coming from one large firm, is not an externally valid sample on which to base a policy conclusion, the depth and scale of the data present an excellent opportunity to illustrate our methodology. We are in the process re-weighting our sample match a more representative population.

We study market simulations where we vary the pricing regulation that we study as well as the notion of static equilibrium that we employ. Our preliminary results reveal that Riley equilibria completely unravel in the case where pricing pre-existing conditions is prohibited (i.e., the less comprehensive 60% plan has 100% market share in equilibrium). When some pricing of pre-existing conditions is allowed (based on health status quartiles), there is less adverse selection in the sense that both the 60% and

[8] Netting out switching costs allows us to use choice and health risk data for the extent of the panel. For more details on the importance of switching costs in this environment, see Handel (2011).

[9] In a recent survey of the empirical insurance literature, Einav et al. (2010a) refer to this kind of model that directly estimates expected utility function parameters as a 'realized' utility model.

90% plans have positive market share for the two healthiest quartiles (for the sicker two quartiles, equilibrium still unravels to everyone enrolling in the 60% plan). Wilson equilibria, which allow for some cross-subsidzation across the 90% and 60% policies, generally lead to interior allocations, with limited unraveling. But again, unraveling is reduced when pricing can be based on pre-existing conditions. Nash equilibria coincide with the Riley outcomes if firms offer only one policy [as in Rothschild and Stiglitz (1976)], but fail to exist with multi-policy firms.

Of course, in the longer run, pricing pre-existing conditions subjects consumers to re-classification risk as they transition between health quartiles over time. Our initial findings for long-run welfare reveal that the prohibition on pricing pre-existing conditions increases welfare, as the losses from reclassification risk far outweigh any losses due to adverse selection. In addition, we find that the losses due to reclassification risk, even for our fairly limited pricing of health status based on quartiles, are quantitatively large.

We use our estimates, and the simulated equilibria to compute the subsidies necessary guarantee different levels of participation in the exchange. Absent subsidies, around 20% of the population would opt out of the exchange when pricing pre-existing conditions is prohibited. A \$2,700 subsidy per person/year would increase coverage to 90%, while \$3,555 is needed to for 95% participation.

This paper builds on related work that studies the welfare consequences of adverse selection in insurance markets by adding in a long-term dimension, whereby price regulation induces a potential trade-off with re-classification risk. Relevant work that focuses primarily on adverse selection includes Cutler and Reber (1998), Cardon and Hendel (2001), Carlin and Town (2009), Lustig (2010), Einav et al. (2011), and Bundorf at al. (2010). Handel (2011) and Einav et al. (2011) study the welfare consequences of adverse selection with switching costs and moral hazard respectively. These papers all focus on welfare in the context of a short-run marketplace. There is limited work studying long run welfare in insurance markets. Cochrane (2005) studies dynamic insurance from a purely theoretical perspective in an environment where fully contingent long-run contracts are possible. Herring and Pauly (2006) studies guaranteed renewable premiums and the extent to which they effectively provide long-run premium insurance. Hendel and Lizzeri (2003) and Finkelstein at al. (2005) study dynamic insurance contracts with one-sided commitment, while Koch (2010) studies pricing regulations based on age from an efficiency perspective. Finally, Bundorf et al. (2010), while focusing on a static marketplace, also contains some analysis of re-classification risk in an employer setting using a two-year time horizon and different subsidy and pricing regulations. To our knowledge, there is no similar work explicitly studying the long-run welfare impact of re-classification risk and adverse selection as a function of price regulation in an empirical setting.

The rest of the paper proceeds as follows: In Section 2 we present our model of insurance exchanges and characterize Riley, Wilson, and Nash equilibria in the context of our model. Section 3 describes our data, and Section 4 discusses estimation of individual risk preferences. In Section 5 we analyze equilibria in these exchanges both when pre-existing conditions can be (partially) priced, and when they cannot. Section 6 analyzes the welfare properties of these equilibria. Section 7 discusses some extensions of our main analysis, and Section 8 concludes.

2 Model of Health Exchanges

Throughout the paper, we focus on a model of health exchanges in which there are exchanges for two prescribed policies. In our basic specification, these policies will cover roughly 90% and 60% respectively of an insured individual's costs. As such, we will refer to these as the "90 policy" and the "60 policy." Within each exchange, the policies offered by different companies are regarded as perfectly homogeneous by consumers; only their premiums may differ. There is a set of consumers, who differ in their likelihood of needing medical procedures and in their preferences (e.g., their risk aversion). We denote by $\theta \in [\underline{\theta}, \overline{\theta}] \subseteq \mathbb{R}_+$ a consumer's "type," which we take to be the price difference at which he is indifferent between a 90 policy and a 60 policy. That is, if P_{90} and P_{60} are the premiums (prices) of the two policies, then a consumer whose θ is below $P_{90} - P_{60}$ prefers the 60 policy, a consumer with θ above $P_{90} - P_{60}$ prefers the 90 policy, and one with $\theta = P_{90} - P_{60}$ is indifferent. (Note that consumers with a given θ may have different underlying medical risks and/or preferences, but will make identical choices between policies for any prices. Hence, there is no reason to distinguish among them in the model.) Throughout our main specification, we assume that there is an individual mandate that requires that individuals purchase one of the two policies.

We specify the costs of insuring an individual of type θ under policy k to be $C_k(\theta)$ for $k = 90, 60$. Recall that if the price difference is $\Delta P = P_{90} - P_{60}$, those consumers with $\theta < \Delta P$ prefer policy 60, while those with $\theta > \Delta P$ prefer policy 90. Given this fact, we can define the average costs of serving the populations who choose each policy for a given ΔP:

$$AC_{90}(\Delta P) \equiv E[C_{90}(\theta)|\theta \geq \Delta P].$$

and

$$AC_{60}(\Delta P) \equiv E[C_{60}(\theta)|\theta \leq \Delta P].$$

Assumption 1: $C_{90}(\theta)$ and $C_{60}(\theta)$ are continuous increasing functions, with $C_{90}(\theta) > C_{60}(\theta)$ for all θ.

The assumption that $C_{90}(\theta) > C_{60}(\theta)$ for all θ simply says that the 90 policy covers more of a consumer's expenses (in expectation) than does the 60 policy.[10] The first part of Assumption 1, on the other hand, is an adverse selection assumption: since the consumers who choose the 90 policy are those in the set $\{\theta : \theta \geq \Delta P\}$, the assumption implies that the consumers who choose the 90 policy are higher cost under any given policy than those that choose the 60 policy. Moreover,

$$AC_{90}(\Delta P) > C_{90}(\Delta P) > C_{60}(\Delta P) > AC_{60}(\Delta P)$$

[10] In our empirical work, the 90 policy will in fact dominate the 60 policy in its coverage levels.

at any ΔP at which both policies are chosen; i.e., at any $\Delta P \in (\underline{\theta}, \overline{\theta})$, and $AC_k(\Delta P)$ is increasing in ΔP for $k = 60, 90$. It will also be convenient to define for each policy $k = 60, 90$ the largest and smallest possible individual and average costs: $\underline{C}_k \equiv C_k(\underline{\theta})$, $\underline{AC}_k \equiv AC_k(\underline{\theta})$, $\overline{C}_k \equiv C_k(\overline{\theta})$ and $\overline{AC}_k \equiv AC_k(\overline{\theta})$.

To ensure that $AC_k(\Delta P)$ is continuous in ΔP for $k = 90, 60$, we also assume the following:

Assumption 2: θ has a continuous distribution function F.

We next define the profits earned by the firms offering the lowest price for a given policy. For any such lowest price pair (P_{90}, P_{60}) define

$$\Pi_{90}(P_{90}, P_{60}) \equiv [P_{90} - AC_{90}(\Delta P)][1 - F(\Delta P)]$$

and

$$\Pi_{60}(P_{90}, P_{60}) \equiv [P_{60} - AC_{60}(\Delta P)]F(\Delta P)$$

as the aggregate profit from consumers who choose each of the two policies. Let

$$\Pi(P_{90}, P_{60}) \equiv \Pi_{90}(P_{90}, P_{60}) + \Pi_{60}(P_{90}, P_{60})$$

be aggregate profit from the entire population.

Finally, we make a risk-aversion assumption: if the two policies are priced at fair odds for a given individual, then that individual strictly prefers the 90 policy:[11]

Assumption 3: $\theta > C_{90}(\theta) - C_{60}(\theta)$ at any $\theta \in [\underline{\theta}, \overline{\theta}]$

Assumption 3 also means that if the 90 price is set to make a consumer indifferent between the two policies, then the 90 policy is more profitable than the 60 policy.

2.1 Equilibrium Notions and Characterizations

The literature on equilibria in markets with adverse selection started with Rothschild and Stiglitz (1976). Motivated by the possibility of non-existence of equilibrium in their model, follow-on work by Riley (1979) [see also Engers and Fernandez (1987)] and Wilson (1977) proposed alternative notions of equilibrium in which existence was assured in the Rothschild-Stiglitz model. These alternative equilibrium notions each incorporated some kind of dynamic reaction to deviations [introduction of additional profitable policies in Riley (1979), and dropping of unprofitable policies in Wilson (1977)], in contrast to the Nash assumption made by Rothschild and Stiglitz. In addition, follow-on work also allowed for multi-policy firms [Miyazaki (1977)], in contrast to Rothchild and Stiglitz's assumption that each firm offers at most one policy.

[11] The assumption is actually somewhat weaker than this, since it says that this holds for the set of consumers of type θ when pricing is at fair odds for this set.

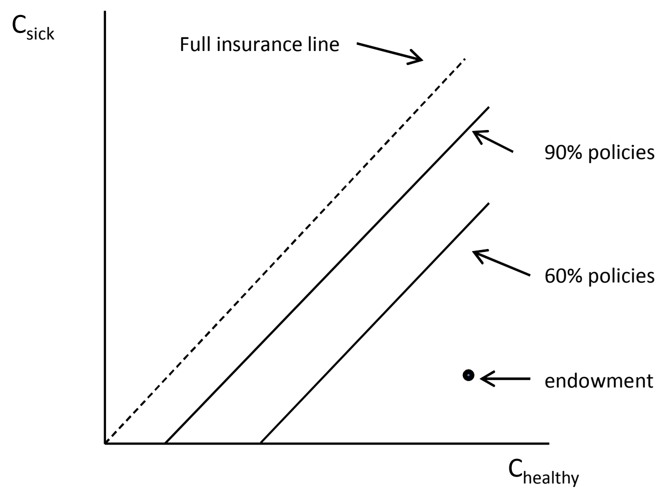


Figure 1: The solid lines with slope equal to 1 indicate the possible consumptions arising with 90% and 60% policies in a two-state (Rothschild-Stiglitz) model of insurance

Our model differs from the Rothschild-Stiglitz setting in three basic ways. First, the prescription of health exchanges limits the set of allowed policies. Figure 1, for example, shows the set of feasible policies in the Rothschild-Stiglitz model (in which each consumer faces just two health states: "healthy" and "sick") with two exchanges, one for a 90% policy and the other for a 60% policy. These lie on lines with slope equal to 1 since a decrease of $1 in a policy's premium increases consumption by $1 in each state. Second, in our model consumers face many possible health states. Third, while the Rothschild-Stiglitz model contemplated just two consumer types, we assume there is a continuum of types of consumers.

In our main analysis we focus on the Riley and Wilson equilibrium notions ("RE" and "WE" respectively).[12] In this section, we characterize these equilibria, which we show always exist in our model. (The Riley equilibrium outcome is also unique.) We also discuss how these compare to Nash equilibria ("NE"), which need not exist.

2.1.1 Riley Equilibria

We use the definition provided in Engers and Fernandez (1987):

[12] The Riley notion is also known as a "reactive equilibrium."

Definition 1. *A **Riley equilibrium** (RE) is a profitable market offering S, such that for any nonempty set S' (the deviation), where $S \cup S'$ is closed and $S \cap S' = \emptyset$, there exists a set S'' (the reaction), disjoint from $S \cup S'$ with $S \cup S' \cup S''$ closed, such that:*

(i) *S' incurs losses when $S \cup S' \cup S''$ is tendered;*

(ii) *S'' does not incur losses when any market offering $\widehat{S}$ containing $S \cup S' \cup S''$ is tendered (we then say S'' is "safe" or a "safe reaction").*

*A deviation S' that is strictly profitable when $S \cup S'$ is offered, and for which there is no safe reaction S'' that makes S' incur losses (with market offering $S \cup S' \cup S''$), is a **profitable Riley deviation**.*

In our setting, a market offering is simply a collection of prices offered for the two policies. Definition 1 says that a set of offered prices is a Riley equilibrium if no firm, including potential entrants, has an incentive to change their offered prices, even if other firms introduce additional price offers, provided that those additional price offers are profitable given the initial deviation and "safe" in the sense that they would remain profitable regardless of any further price offers being introduced.[13]

Our result for Riley equilibria, which we establish in the Appendix, is the following:

Proposition 1. *A Riley equilibrium always exists and involves a unique allocation of consumers to the two policies, and unique prices (P_{90}, P_{60}) for any policies that are purchased. Moreover:*

(i) *If $\Pi_{60}(\underline{AC}_{90}, P_{60}) \leq 0$ for all P_{60} (i.e. if there is no profitable entry into the 60 policy given that the 90 policy is priced to break even), it involves everyone buying the 90 policy at prices $P_{90} = \underline{AC}_{90}$ and $P_{60} \geq \underline{AC}_{90} - \underline{\theta}$;*

(ii) *Otherwise, it involves positive sales of the 60 policy at the lowest price difference $\Delta P > \underline{\theta}$ at which $\Delta P = AC_{90}(\Delta P) - AC_{60}(\Delta P)$ if such a ΔP exists, with prices (P_{90}, P_{60}) that cause both policies to break even, and if such a ΔP does not exist it involves all consumers buying the 60 policy at prices $P_{60} = \overline{AC}_{60}$ and $P_{90} \geq \overline{AC}_{60} + \overline{\theta}$.*

2.1.2 Wilson Equilibria

A Wilson Equilibrium (WE) is a pair of prices (P_{90}, P_{60}) and resulting allocation such that there is no (potentially multi-policy) deviation that is profitable and remains so once any policy offers by firms that become unprofitable after the deviation are withdrawn from the market.[14] (With two policy types, there is at most one such unprofitable policy offer after any profitable deviation, so there is no ambiguity about which policies to withdraw.)

We have the following characterization of WE in our model (proof in the Appendix):

[13] In fact, it suffices to restrict attention to deviations by potential entrants.

[14] Wilson (1977) assumed firms offered a single policy; the multi-policy extension is due to Miyazaki (1977).

Proposition 2. *Suppose that* $\overline{AC}_{60} \neq \underline{AC}_{90} - \underline{\theta}$ *. Then* (P_{90}^w, P_{60}^w) *with* $\Delta P^w = P_{90}^w - P_{60}^w \in [\underline{\theta}, \overline{\theta}]$ *is a Wilson equilibrium iff it is a solution to*[15]

$$\begin{array}{lll} \min_{(P_{90},P_{60})} & & P_{60} \\ s.t. & & (i)\ \Pi_{90}(P_{90}, P_{60}) \leq 0 \\ & & (ii)\ \Pi(P_{90}, P_{60}) = 0 \\ & & (iii)\ \Delta P \in [\underline{\theta}, \overline{\theta}] \end{array} \tag{1}$$

Thus, the WE is the price pair (P_{90}, P_{60}) that has the lowest price for the 60 policy among all price pairs that yield zero aggregate profit and non-positive profit for the 90 policy. This characterization can be understood intuitively as follows: First, any WE must have $\Pi_{90}(P_{90}, P_{60}) \leq 0$ since otherwise there would be a profitable deviation in only the 90 policy, slightly undercutting P_{90} (note that any dropping of a 60 policy in response would only increase the profits from this deviation). Next, given that $\Pi_{90}(P_{90}, P_{60}) \leq 0$, the deviation that must be prevented is a deviation that attracts the best risks currently in the 90 policy to a new lower-priced 60 policy. This deviation will make the insurer offering the current 90 policy lose money, causing it to drop its 90 policy. Once the 90 policy is dropped, though, if the deviator hasn't offered a 90 policy, everyone will go to the new 60 policy, causing it to lose money. So the deviator must offer a pair of policies: the new 60 policy that attracts some individuals currently buying the 90 policy, and a new 90 policy to keep everyone from going to the 60 policy. At the solution to (1) no such profitable deviation exists. Note in particular that a WE may involve cross-subsidization of the 90 policy by the 60 policy (the 90 policy may lose money on its own), something that cannot happen in either a RE or (as well will see) a NE.

We also note the following result, indicating that if a Riley equilibrium has a positive share of the 60 policy, then any Wilson equilibrium has (weakly) greater coverage than the Riley equilibrium:[16]

Corollary 1. *Any Wilson equilibrium must have a* ΔP *no greater than the lowest* $\Delta P > \underline{\theta}$ *at which* $\Delta P = AC_{90}(\Delta P) - AC_{60}(\Delta P)$*, if such a* ΔP *exists.*

Proof. Let $\widehat{\Delta P}$ be the lowest $\Delta P > \underline{\theta}$ at which $\Delta P = AC_{90}(\Delta P) - AC_{60}(\Delta P)$ and suppose there is a Wilson equilibrium at which $\Delta P^w > \widehat{\Delta P}$. Since both policies break even at $\widehat{\Delta P}$, and the share of the 90 policy is greater at $\widehat{\Delta P}$ than at ΔP^w [where $\Pi_{60}(P_{90}^w, P_{60}^w) \geq 0$], it must be that $\widehat{P}_{60} = AC_{60}(\widehat{\Delta P}) < AC_{60}(\Delta P^w) \leq P_{60}^w$ – a contradiction. □

Finally, note that we have not established that WE is unique [in contrast to Wilson (1977)], although in practice we have always found it to be so in our data.

[15] If a WE (P_{90}^w, P_{60}^w) has $\Delta P^w < \underline{\theta}$, so that the 60 policy is not purchased, then there is an equivalent WE with $\widetilde{P}_{60} = P_{90}^w - \underline{\theta}$ [the price pair $(P_{90}^w, \widetilde{P}_{60})$ results in the same measures of sales, the same expected utilities for each type, and the same profits, and is also a WE]. Similarly if $\Delta P^w > \overline{\theta}$.

[16] In the next subsection we will see that if the RE is all-in-90, the RE is also a WE.

2.1.3 Nash Equilibria

The original Rothschild-Stiglitz model instead examined Nash equilibria with single-policy firms. The following result characterizes NE in our model for both single- and multi-policy firms:

Proposition 3. *With either single- or multi-policy firms, any NE must have firms break even on all policies that are sold in equilibrium. The unique equilibrium has all consumers buying the 90 policy if* $\Pi_{60}(\underline{AC}_{90}, P_{60}) \leq 0$ *for all* P_{60}*. If this condition does not hold, a NE must have positive sales of the 60 policy, must involve the lowest* ΔP *such that* $\Delta P = AC_{90}(\Delta P) - AC_{60}(\Delta P)$ *if such a* ΔP *exists, and must have all consumers buying the 60 policy if such a* ΔP *does not exist. Such a price pair is a NE for:*

(i) *single-policy firms if there is no profitable entry opportunity in the 90 policy; i.e., if* $\Pi_{90}(\widehat{P}_{90}, P_{60}) \leq 0$ *for all* $\widehat{P}_{90} \leq P_{90}$*.*

(ii) *multi-policy firms if there is no profitable entry opportunity that slightly undercuts* P_{60} *and undercuts* P_{90}*: i.e., if* $\sup_{\widehat{P}_{90} \leq P_{90}} \Pi(\widehat{P}_{90}, P_{60}) = 0$*. This is impossible if all consumers buy the 60 policy at* (P_{90}, P_{60})*.*

Propositions 1, 2, and 4 imply that all consumers buying the 90 policy is an equilibrium (and the unique one for RE and NE) under the exact same circumstances with all three concepts: when $\Pi_{60}(\underline{AC}_{90}, P_{60}) \leq 0$ for all P_{60}. Where they differ is in what happens when this is not true. Comparing RE and NE, we see that when there are prices that break even and have both policies purchased, with either concept only the one with the lowest ΔP can be an equilibrium. However, such a ΔP can be a RE when it fails to be a NE because under the RE concept a profitable Nash deviation can be rendered unprofitable by additional profitable (and "safe") entry once the initial deviation occurs.

Observe in Proposition 4 that with multi-policy firms there must be positive sales of the 90 policy in any NE. This stems from the fact that, otherwise, with risk-aversion (Assumption 2) it is always worthwhile to enter offering the worst consumers more insurance (in the present case, the 90 policy). However, under RE this need not be the case, because this deviation may be rendered unprofitable by additional profitable (safe) entry once this deviation occurs. Observe, though, that if a multi-policy NE does exist, then it is also a RE and WE:

Corollary 2. *If a multiproduct Nash equilibrium exists, it is an RE, a WE, and a single-policy NE.*

Proof. Any multi-policy NE is a single-policy NE because fewer deviations are possible. It is both an RE and a WE because there are no profitable deviations, even before considering any reactions. □

3 Estimation

To simulate equilibria in health exchanges we need a population of insurees, their preferences and health status. More specifically, we need the following ingredients: i. the distribution of health states (risk) in the population, ii. the distribution of health expanses for each health state, iii. the distribution of risk preferences, iv. the transition across health states over time. In terms of the model, risk preferences and the distribution of health expenses (given health status) are needed to compute θ of eachperson in the sample, i.e., their willingness to pay for extra coverage. The willingness to pay is determined by the risk faced and risk aversion. The joint distribution of health status and preferences are needed to figure out $F(\theta)$. Recall $F(\theta)$ determines both market shares and costs of each policy, given premiums. Finally, transition are useful to study long term coverage.

To obtain these ingredients we use detailed data on health insurance choices and medical utilization of employees (and their dependents) at a large U.S. based firm over the time period from 2004 to 2009. These proprietary panel data include the health insurance options available in each year, employee plan choices, and detailed, claim-level employee (and dependent) medical expenditure and utilization information.

The parameters used in the simulation of the exchanges are recovered through the following three procedures. The first step is to use patient history (diagnostics, expenses and demographics) to define an index of current health status for each individual in the sample. The procedure is based on a commercial software used by insurance companies (more below). The second step, denoted the cost model, entails estimating the distribution of expenses faced by individuals in each health state. Basically, by grouping individuals of similar initial health, we use data on the -health expense- realizations to estimate the risk they faced. Finally, risk preferences are estimated fitting an insurance choice model. Each individual, given her health status, faces uncertain out of pocket expenses. Risk preferences affect their willingness to pay for coverage. A discrete choice model that explains plan choice in our sample is fitted to recover preferences.

All the estimates are taken from Handel (2012). The next three section briefly explain how the estimation works, as well as the data used, for further details see Handel (2012).

3.1 Health status and its distribution

We use detailed medical information together with medical risk prediction software developed at Johns Hopkins Medical School to create individual-level measures of predicted future medical utilization at each point in time. These measures are generated using past diagnostic, expense, and demographic information[17] and allow us to precisely gauge medical expenditure risk prior to the plan choice. The program, known as the Johns Hopkins ACG (Adjusted Clinical Groups) Case-Mix System, is one of the most widely used and respected risk adjustment and predictive modeling packages in the health care sector. It was specifically designed to use diagnostic claims data, such as the individual-level ICD-9 codes we observe, to predict future medical expenditures in a sophisticated manner. In addition,

[17] We observe this detailed medical data for all employees and dependents enrolled in one of several PPO options, which is the set of available plans our analysis focuses on. For a further discussion see the sample composition section. These data include detailed claim-level diagnostic information, such as ICD-9 codes and NDC drug codes, as well as provider information and a detailed payment breakdown (e.g. deductible paid, coinsurance paid, plan paid).

the program takes into account the NDC pharmaceutical drug utilization codes we observe as well as individual age and gender.

By plugging the diagnostic codes and medical expenses of prior health utilization, as well as demographic information, the software generates an AGC index that summarizes expected expenses for the coming year for every individual in the sample. Denote an individual's past year of medical diagnoses and payments by ξ_{it} and the demographics age and sex by ζ_{it}. We use the ACG software mapping, denoted A, to map these characteristics into a predicted mean level of health expenditures (including pharmaceuticals) for the upcoming year, denoted λ:

$$A : \xi \times \zeta \rightarrow \lambda$$

3.2 The Cost Model

The ACG index predicts health expenses. However, to evaluate the expected utility from different coverage options we need the distribution of expenses, not just its mean. We utilize the cost model developed in Handel (2012) to estimate the distribution of health expenditure for different values of the ACG index, λ.

The cost model makes several advances relative to the recent literature that uses micro-level claims data to quantify individual health risk. It offers a parsimonious method to non-parametrically link health risk to expected future expenditures by combining the predictive ACG health risk output with observed cost data.[18]

In order to predict expenses, medical claims are categorized into mutually exclusive categories (e.g., hospital, pharmacy, physician office visit). We use the prediction λ, of each individual (described above) to categorize them into expenditure cells. Then for each group of individuals in each claims category, the actual ex post realized claims for that group is used to estimate the ex ante distribution for each individual under the assumption that this distribution is identical for all individuals within the cell. The minimum number of individuals in any cell is 73 while almost all cells have over 500 members. Since there are four categories of claims, each individual can belong to one of approximately 10^4 or 10,000 combination of cells.

Denote an arbitrary cell within a given category d by z. Denote the population in a given category-cell combination (d, z) by I_{dz}. Denote the empirical distribution of ex-post claims in this category for

[18] The cost model assumes that there is no moral hazard and that there is no private information. While both of these phenomena have the potential to be important in health care markets, and are studied extensively in other research, we believe that these assumptions do not materially impact our results. One primary reason is that both effects are likely to be quite small relative to the magnitude of the overall money at stake, which is what the risk preference estimates are sensitive to. For private information, we should be less concerned than prior work because our cost model combines detailed individual-level prior medical utilization data with sophisticated medical diagnostic software. This makes additional selection based on private information much more unlikely than it would be in a model that uses coarse demographics or aggregate health information to measure health risk.

this population $G_{\hat{I_{dz}}}(\cdot)$. Then we assume that each individual in this cell has a distribution equal to a continuous fit of $G_{\hat{I_{dz}}}(\cdot)$, which we denote G_{dz}:

$$\varpi : G_{\hat{I_{dz}}}(\cdot) \rightarrow G_{dz}$$

The above process generates a distribution of claims for each d and z but does not model correlation over -expense categories- D. It is important to model correlation over claim categories because it is likely that someone with a bad expenditure shock in one category (e.g. hospital) will have high expenses in another area (e.g. pharmacy). We model correlation at the individual level by combining marginal distributions G_{idt} $\forall$ d with empirical data on the rank correlations between pairs (d, d'). Here, G_{idt} is the distribution G_{dz} where $i \in I_{dz}$ at time t. Since correlations are modeled across d we pick the metric λ to group people into cells for the basis of determining correlations (we use the same cells that we use to determine group people for hospital and physician office visit claims). Denote these cells based on λ by z_λ. Then for each cell z_λ denote the empirical rank correlation between claims of type d and type d' by $\rho_{z_\lambda}(d, d')$. Then, for a given individual i we determine the joint distribution of claims across D for year t, denoted $H_{it}(\cdot)$, by combining i's marginal distributions for all d at t using $\rho_{z_\lambda}(d, d')$:

$$\Psi : G_{iDt} \times \rho_{z_{\lambda_{it}}}(D, D') \rightarrow H_{it}$$

Here, G_{iDt} refers to the set of marginal distributions $G_{idt} \forall d \in D$ and $\rho_{z_{\lambda_{it}}}(D, D')$ is the set of all pairwise correlations $\rho_{z_{\lambda_{it}}}(d, d') \forall (d, d')$. In estimation we perform Ψ by using a Gaussian copula to combine the marginal distribution with the rank correlations.

3.3 Risk Preferences: Choice Model

Risk preference have been estimated using choice data. Each household k in the sample faces a choice set, and a distribution of health expenses $F_{kjt}(\cdot)$ conditional on state λ and coverage j (described above). We assume (CARA) preferences:

$$u_k(x) = -\frac{1}{\gamma_k(X_k^A)} e^{-\gamma_k(X_k^A)x}$$

The distribution of γ_k is estimated fitting predicted expected utility maximizing choices to observed one. Where the expected utility is given by:

$$U_{kjt} = \int_0^\infty f_{kjt}(OOP) u_k(W_k, OOP, P_{kjt}, \mathbf{1}_{kj,t-1}) dOOP$$

OOP is a realization of medical expenses from $F_{kjt}(\cdot)$. W_k denotes wealth and P_{kjt} is premium contribution for plan j, which as described earlier depends both on how many dependents are covered and on employee income. γ_k is a household-specific risk preference parameter that is known to the family but unobserved by the econometrician. The random coefficients γ_k is assumed normally distributed.

Consumption x conditional on a draw OOP from $F_{kjt}(\cdot)$ is:

$$x = W_k - P_{kjt} - OOP + \epsilon_{kjt}(Y_k)$$

3.4 Data

The data is from a firm that employs approximately 9,000 people per year. The first column of Table 1 describes the demographic profile of the 11,253 employees who work at the firm for some stretch within 2004-2009. These employees cover 9,710 dependents, implying a total of 20,963 covered lives. 46.7% of the employees are male and the mean employee age is 40.1 (median of 37). The age distribution of the firm is very uniform:

Quantile	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
Age	26	29	33	38	43	47	50	54	58	64	68

Returning to Table 1, we observe income grouped into five tiers, the first four of which are approximately \$40,000 increments, increasing from 0, with the fifth for employees that earn more than \$176,000. Almost 40% of employees have income in tier 2, between \$41,000 and \$72,000, with 34% less than \$41,000 and the remaining 26% in the three income tiers greater than \$72,000. 58% of employees cover only themselves with health insurance, with the other 42% covering a spouse and/or dependent(s). 23% of the employees are managers, 48% are white-collar employees who are not managers, and the remaining 29% are blue-collar employees. 13% of the employees are categorized as 'quantitatively sophisticated' managers. Finally, the table presents information on the mean and median characteristics of the zip codes the employees live in.

A key feature of the data for our study is that, in the middle of our observational period, the firm substantially changed the menu of health plans that it offered to employees, in a year that we denote t_0.[19] At the time of this change, the firm forced all employees to leave their prior plan and actively re-enroll in one of five options from the new menu, with no default option. This is important to be able to recover preference in a period with active choice (clean of switching costs or inertia).

Health Risk Descriptives.

Table 3.4 describes health status transitions in the population over one and two year time horizons. For the table, we group employees into ex ante health quartiles using the Johns Hopkins ACG program referenced earlier and described in more detail in the cost model section. These quartiles represent employees grouped by mean expected expenditure in each year, where the expectation is determined using prior cost and diagnostic information. The table reveals that there are real transition risks even for the fairly short one and two year time horizons: for example 32% of the individuals in the healthiest quartile in year $t-1$ transition to one of the other three quartiles at year t. The table also presents

[19] This change had the two stated goals of (i) encouraging employees to choose new, higher out-of-pocket spending plans in order to help control total medical spending and (ii) providing employees with a broader choice of different health insurance options (e.g. a consumer driven health plan with a linked health savings account (HSA)).

Sample Demographics	All Employees	PPO Ever	Final Sample
N - Employee Only	11,253	5,667	2,023
N - All Family Members	20,963	10,713	4,544
Mean Employee Age (Median)	40.1	40.0	42.3
	(37)	(37)	(44)
Gender (Male %)	46.7%	46.3%	46.7%
Income			
Tier 1 (< $41K)	33.9%	31.9%	19.0%
Tier 2 ($41K-$72K)	39.5%	39.7%	40.5%
Tier 3 ($72K-$124K)	17.9%	18.6%	25.0%
Tier 4 ($124K-$176K)	5.2%	5.4%	7.8%
Tier 5 (> $176K)	3.5%	4.4%	7.7%
Family Size			
1	58.0 %	56.1 %	41.3 %
2	16.9 %	18.8 %	22.3 %
3	11.0 %	11.0 %	14.1 %
4+	14.1 %	14.1 %	22.3 %
Staff Grouping			
Manager (%)	23.2%	25.1%	37.5%
White-Collar (%)	47.9%	47.5%	41.3%
Blue-Collar (%)	28.9%	27.3%	21.1%
Additional Demographics			
Quantitative Manager	12.8%	13.3%	20.7%
Job Tenure Mean Years (Median)	7.2	7.1	10.1
	(4)	(3)	(6)
Zip Code Population Mean (Median)	42,925	43,319	41,040
	(42,005)	(42,005)	(40,175)
Zip Code Income Mean (Median)	$56,070	$56,322	$60,948
	($55,659)	($55,659)	($57,393)
Zip Code House Value Mean (Median)	$226,886	$230,083	$245,380
	($204,500)	($209,400)	($213,300)

average and median ex post cost by quartile grouping, indicating an increase in expected expenditures from $\$1,812$ for quartile 1 to $\$15,199$ for quartile 4.

Preference Estimates

Since the CARA coefficients are difficult to interpret, table 7 shows the implication of the estimates. The table presents the value X that would make an individual with our estimated risk preferences indifferent between inaction and accepting a gamble with a 50% chance of gaining \$100 and a 50% chance of losing $\$X$.[20] Thus, a risk neutral individual will have $X = \$100$ while an infinitely risk averse individual will have X close to zero. The top section of the table presents the results for the primary specification. X is \$94.6 for the mean / median individual, implying a moderate amount of risk aversion relative to other results in the literature, which we present at the bottom of the table. X is \$92.2 for the 95th percentile of γ and \$91.8 for the 99th, so preferences don't exhibit substantial heterogeneity in the context of the literature.

4 Equilibria

4.1 No pricing of preexisting conditions

The main results of section 2.1 guide the way we find equilibria. We start with the whole population, namely, in an exchange where premiums are independent of pre-existing conditions and demographics. We assume that the 60 policy has a 20% co-pay, a \$3000 deductible, and a \$1500 out-of-pocket maximum, while the 90 policy has the same 20% co-pay, a \$3000 deductible, and a \$5950 out-of-pocket maximum. Over the whole population these policies cover roughly 60% and 90% of expenses.

The first step towards finding equilibria involves checking whether pooling at 90, the highest level of coverage, is an equilibrium. If all consumers are in 90, for such a policy to break even it must be the case that $P_{90} = \underline{AC}_{90}$. For that policy to be an equilibrium it is necessary and sufficient that $\Pi_{60}(\underline{AC}_{90}, P_{60}) \leq 0$ for all P_{60}. If the condition holds, propositions 1, 2, and 4 guarantee that all-in-90 is an equilibrium under all concepts, Nash, Wilson and Reactive, and the equilibrium is unique.

Figure 2 displays $\Pi_{60}(\underline{AC}_{90}, P_{60})$, more precisely, it shows $\Pi_{60}(\underline{AC}_{90}, \underline{AC}_{90} - \Delta P)$ as a function of ΔP. For $\Delta P = 0$ all consumers prefer the 90 policy, thus the market share of the 60 policy as well as its profits are 0. As ΔP increases the best types opt out of 90 and start buying 60. The graph shows a whole range of P_{60} for which $\Pi_{60}(\underline{AC}_{90}, \underline{AC}_{90} - \Delta P) > 0$, making it profitable for 60 policies to poach the best risks out of the 90 policy. Thus, in our population all-in-90 is not an equilibrium. The equilibrium must involve purchases of the 60 policy.

The second step towards finding either Riley or Nash equilibria involves finding price pairs (P_{90}, P_{60}) at which both policies break even. This can be found by identifying the ΔP at which $\Delta P = AC_{90}(\Delta P) - AC_{60}(\Delta P)$. Among them (if any exist), according to Proposition 1 the one with the lowest ΔP, namely,

[20] These figures are computed for an individual with mean age and mean income.

1 Year Transtion				
t-1 / t	ACG Quartile 1	ACG Q2	ACG Q3	ACG Q4
ACG Quartile 1	0.68	0.18	0.08	0.06
ACG Quartile 2	0.12	0.46	0.25	0.17
ACG Quartile 3	0.05	0.15	0.39	0.41
ACG Quartile 4	0.03	0.05	0.17	0.75

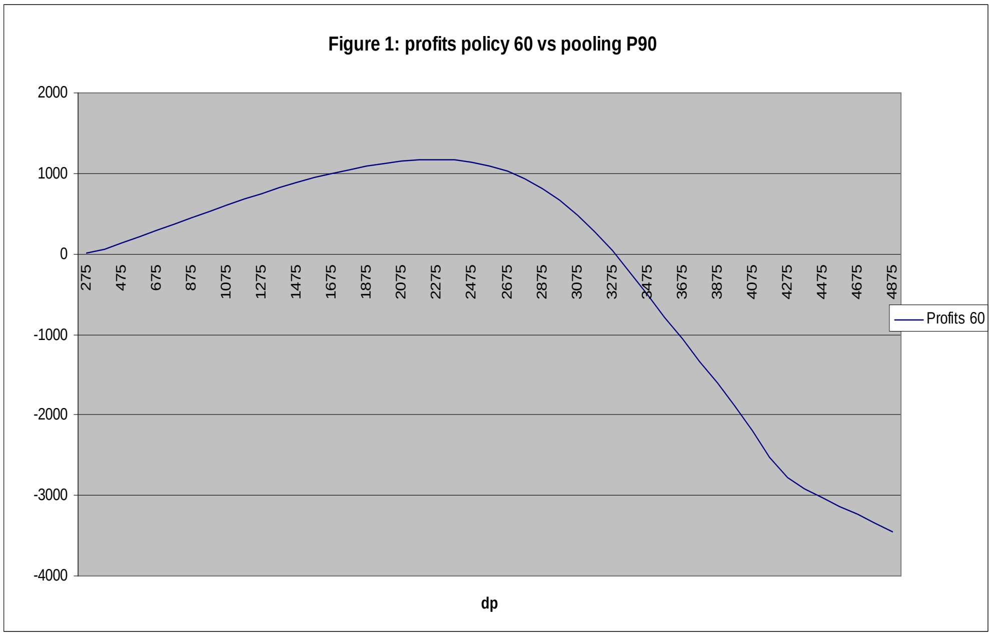
2 Year Transtion				
t-1 / t	ACG Quartile 1	ACG Q2	ACG Q3	ACG Q4
ACG Quartile 1	0.58	0.22	0.09	0.11
ACG Q2	0.10	0.35	0.29	0.26
ACG Q3	0.03	0.14	0.28	0.56
ACG Q4	0.03	0.05	0.10	0.82

Cost Profile ($)			
	Quartile	Avg. Cost	Med. Cost
	ACG Q1	1812	302
	ACG Q2	3544	1107
	ACG Q3	5543	2542
	ACG Q4	15199	6831

Table 2: This table describes health status transitions in the population over one and two year time horizons. For the table, we group employees into ex ante health quartiles using the Johns Hopkins ACG program referenced earlier and described in more detail in the cost model section. The top two sections describe these transitions, while the final sections profiles costs as a function of quartile.

Risk Preference Analysis		
	Absolute Risk Aversion	Interpretation
Normal Heterogeneity		
Mean / Median Individual	$4.22 * 10^{-4}$	94.6
25th percentile	$2.95 * 10^{-4}$	96.1
75th percentile	$5.49 * 10^{-4}$	93.8
95th percentile	$7.31 * 10^{-4}$	92.2
99th percentile	$8.59 * 10^{-4}$	91.8
Log normal Heterogeneity		
Mean	$9.82 * 10^{-4}$	91.0
25th percentile	$1.53 * 10^{-4}$	97.2
Median	$3.85 * 10^{-4}$	95.0
75th percentile	$9.72 * 10^{-4}$	91.1
95th percentile	$3.70 * 10^{-3}$	72.8
99th percentile	$9.30 * 10^{-3}$	51.1
Comparable Estimates		
[?] Benchmark Mean	$3.1 * 10^{-3}$	76.5
[?] Benchmark Median	$3.4 * 10^{-5}$	99.7
[?]	$3.1 * 10^{-4}$	97.0
[?]	$3.2 * 10^{-2}$	21.0
[?]	$2.0 * 10^{-3}$	83.3

Table 3: This table examines the risk preference estimates from the empirical results presented in table **??**. The first section of the table is for the normally distributed risk preference estimates in the Primary specification, where the age and income coefficients are evaluated at the median values of those variables. The second section is for the model with log-normally distributed preferences studied in column 4 of table **??**. The interpretation column is the value X that would make someone indifferent about accepting a 50-50 gamble where you win \$100 and lose X versus a status quo where nothing happens. Our estimates are similar under both specifications with the exception that the log normal model predicts a fatter tail with higher risk aversion. These estimates are in the middle of the (wide) range found in the literature and show moderate risk aversion except at the tails in the log-normal model where consumers are quite risk averse.



with the highest 90 share, is the unique Riley equilibrium. Moreover, it is the only candidate for Nash equilibrium, should a Nash equilibrium exist.

Figure 3 plots $\Delta AC = AC_{90}(\Delta P) - AC_{60}(\Delta P)$, against ΔP. The figure shows that there is no interior equilibrium candidate, namely, no pair of premiums at which both policies have positive shares and both break even. Thus, the only break-even candidate is a ΔP high enough to have all consumers in the 60 policy. Proposition 1 guarantees existence of a Riley equilibrium, so having ruled out all other candidates, we know all-in-60 must be the Riley equilibrium. However, the existence of a Nash equilibrium is not guaranteed, and depends on the profitability of 90 deviations. Notice that by Proposition 4 we know that all-in-60 does not survive a double deviation with multi-policy firms: there is always a profitable way to offer coverage to the worst risks and catch the improved risks left in the 60 pool via $P_{60} = \underline{AC}_{60} - \varepsilon$.

The third, and final, step involves checking whether all-in-60 is a single-policy Nash equilibrium. Figure 4 shows profits from a single and a double deviation, namely, $\Pi_{90}(P_{90}, \underline{AC}_{60})$ and $\Pi(P_{90}, \underline{AC}_{60} - \varepsilon)$. For high P_{90} all consumers purchase 60, so profits for the 90 policy (and total profits, since 60 breaks even) are 0. As the blue line shows, $\Pi_{90}(P_{90}, \underline{AC}_{60})$ is always negative. Attracting the worst risks into 90 coverage with a single-policy deviation is not profitable at any price that would manage to attract them. Thus, pooling at 60 is a single-policy Nash equilibrium as well as a Riley equilibrium.

The pink (higher) line confirms that a double deviation from all-in-60 is profitable. As found in Proposition 4, while the 90 customers are not profitable by themselves, the pool left in 60, which can be attracted with $P_{60} = \underline{AC}_{60} - \varepsilon$, more than compensates for the losses in 90.

Wilson equilibria, as in the Rothschild-Stiglitz framework, may but need not coincide with the Riley and Nash equilibria. We actually know that since the Riley/single-policy Nash equilibrium premiums

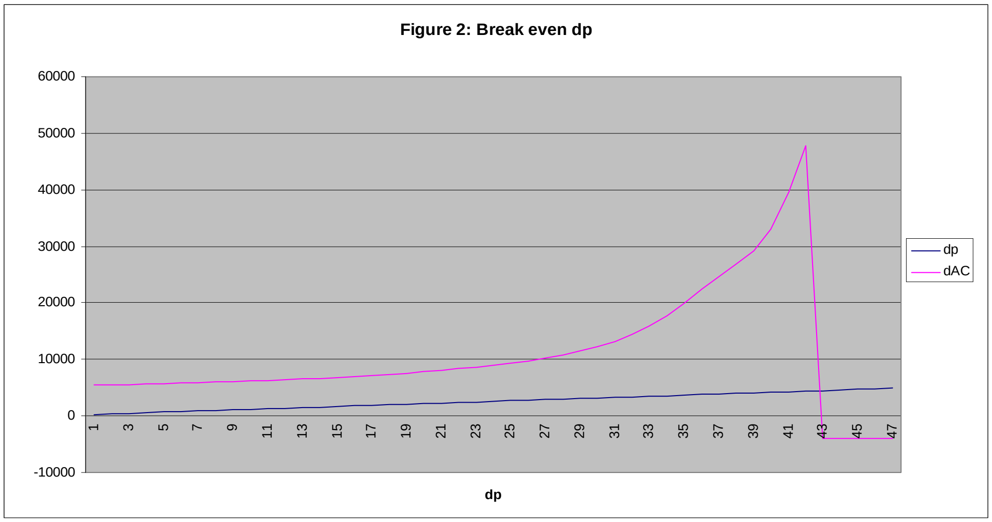


Figure 2: Break even dp
60000
50000
40000
30000
20000
10000
0
-10000
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
35
37
39
41
43
45
47
dp
dp
dAC

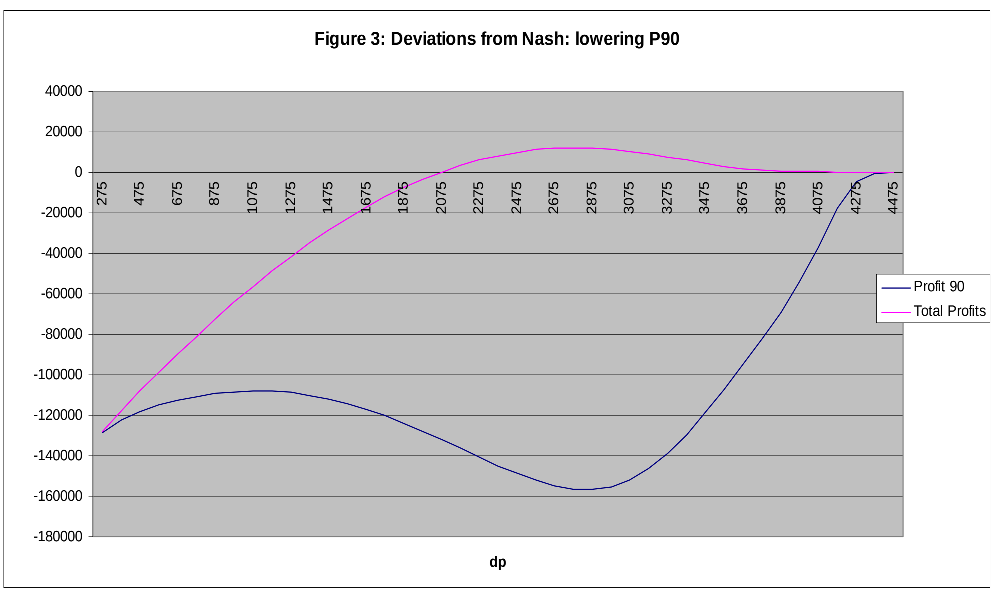


Figure 3: Deviations from Nash: lowering P90
40000
20000
0
-20000
-40000
-60000
-80000
-100000
-120000
-140000
-160000
-180000
275
475
675
875
1075
1275
1475
1675
1875
2075
2275
2475
2675
2875
3075
3275
3475
3675
3875
4075
4275
4475
dp
Profit 90
Total Profits

we found do not survive a double deviation, they are not a Wilson equilibrium (since double deviations are unaffected by existing policies being dropped).

Wilson policies break even in total, but they do so allowing the 60 policy to cross-subsidize the 90 policy. Following Proposition 2 we need to find the lowest P_{60} that paired with a P_{90} delivers non-negative total profits and as well as non-positive 90 profits.

Table 5.1 shows the premiums that solve (1). They indeed involve a cross subsidy from 60 to 90 customers, and zero total profits. The subsidy to 90 leads to 36% of the market getting high coverage, in contrast to Nash where all of the population is in 60. Table 5.1 summarizes the predictions for the case with a prohibition on pricing pre-existing conditions.

Table 4.1: Equilibria Without Pre-existing Conditions

	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley	3974	100	3974	–	0	–
Wilson	3852	64	1347	6627	36	11125
Single Nash	3974	100	3974	–	0	–
Double Nash			Does not exist			

4.2 ACG Pricing

We now turn consider equilibria when insurers can price observables. In this section we assume insurers can separate the population into quartiles based on ACGs, and price each segment of the market separately. We follow the same steps as in the previous subsection to find equilibria, but now for each market segment separately.

The most interesting subpopulation is quartile 1 (the healthiest quartile). In this subsample there is an interior Nash equilibrium that survives double deviations. Thus, it coincides with Riley, but also with Wilson (as it survives the double deviation).

The first step, as above, is to check whether all-in-90 is an equilibrium. There is, as in all cases below, a whole range of P_{60} for which $\Pi_{60}(\underline{AC}_{90}, \underline{AC}_{90} - \Delta P) > 0$. Thus, all-in-90 is not an equilibrium.

Figure 5 shows that there are two interior break-even ΔP, which are the RE/NE equilibrium candidates. Propositions 1 and 4 guarantee that only the one with the highest share of customers in 90 (lowest ΔP) can be either a Riley or Nash equilibrium. Figure 6 displays profitability of single and double deviations, starting at the lower of these ΔP. As neither profit is positive at any such ΔP, this ΔP is an equilibrium under all of our equilibrium notions.

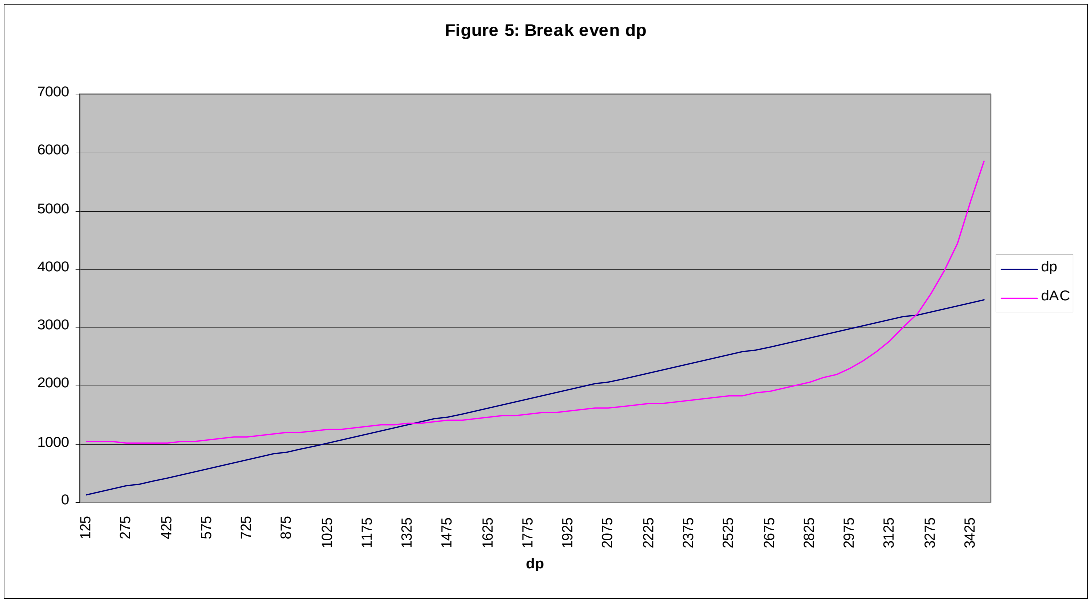


Table 4.2: Equilibria with ACG based Pricing (quartile 1)						
	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley	229	61	229	1601	39	1601
Wilson	229	61	229	1601	39	1601
Single Nash	229	61	229	1601	39	1601
Double Nash	229	61	229	1601	39	1601

Table 5.2 summarizes the equilibria for quartile 2, which is similar to the whole population. The only distinction is that the single policy Nash is interior, but a double deviation Nash equilibrium does not exist. The latter explains why it differs from Wilson, which involves more customers in 90 (consistent with Corollary1).

Table 4.3: Equilibria with ACG based Pricing (quartile 2)						
	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley	1176	82	1175	3850	18	3856
Wilson	1065	52	886	3140	48	3336
Single Nash	1176	82	1175	3850	18	3856
Double Nash			Does not exist			

Figure 9 confirms, the candidate survives single deviation but not double.

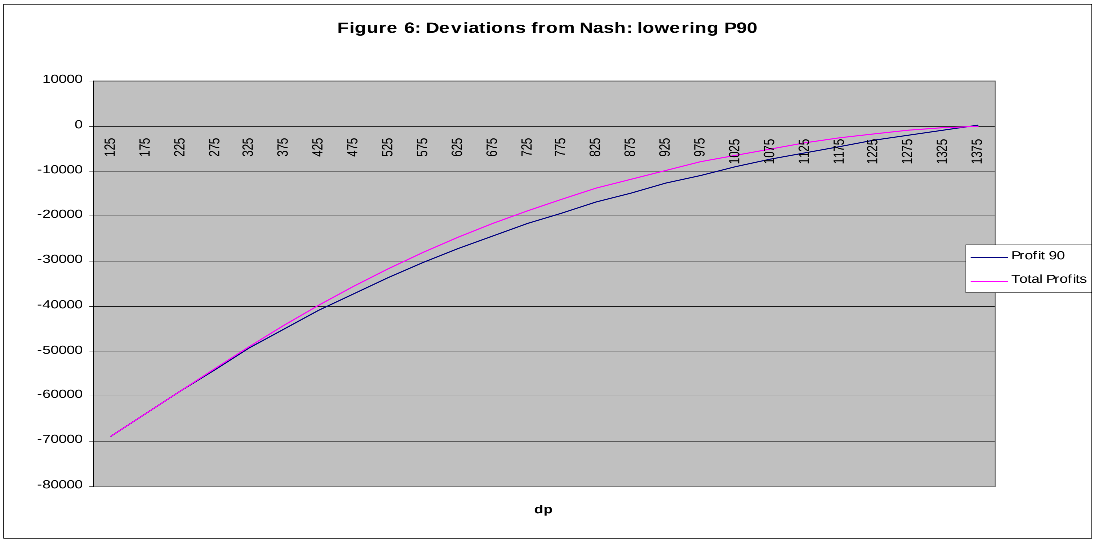


Figure 6: Deviations from Nash: lowering P90
10000
0
-10000
-20000
-30000
-40000
-50000
-60000
-70000
-80000
125
175
225
275
325
375
425
475
525
575
625
675
725
775
825
875
925
975
1025
1075
1125
1175
1225
1275
1325
1375
dp
Profit 90
Total Profits

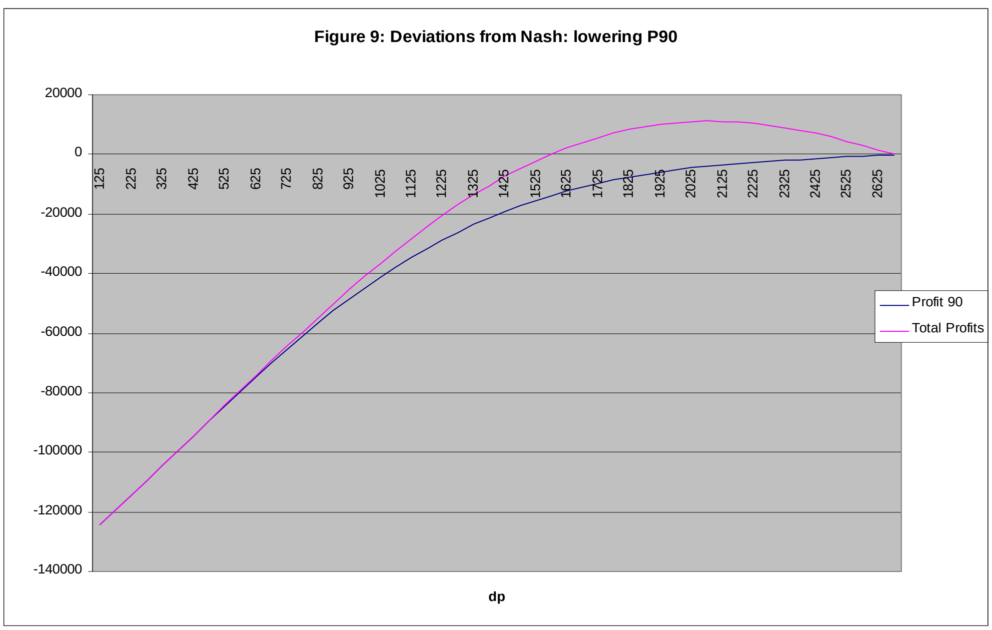


Figure 9: Deviations from Nash: lowering P90
20000
0
-20000
-40000
-60000
-80000
-100000
-120000
-140000
125
225
325
425
525
625
725
825
925
1025
1125
1225
1325
1425
1525
1625
1725
1825
1925
2025
2125
2225
2325
2425
2525
2625
dp
Profit 90
Total Profits

Equilibria in quartiles 3 and 4 are summarized in Tables 5.3 and 5.4. The equilibria are qualitatively identical to the one for the whole population. The single policy Nash equilibrium is all-in-60 and coincides with Riley, while Wilson involves positive 90 sales, and cross-subsidization. A multiple policy Nash equilibrium does not exist. We omit the graphs, which look exactly as for the population as a whole.

Table 4.4: Equilibria with ACG based Pricing (quartile 3)

	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley	4408	100	4408	–	0	–
Wilson	4147	51	2389	6872	49	8732
Single Nash	4408	100	4408	–	0	–
Double Nash			Does not exist			

Table 4.5: Equilibria with ACG based Pricing (quartile 4)

	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley	9670	100	9670	–	0	–
Wilson	9553	50	4987	12803	50	17314
Single Nash	9670	100	9670	–	0	–
Double Nash			Does not exist			

5 Welfare

Our aim is to evaluate the expected utility of an individual starting at age 25 from an ex ante ("unborn") perspective; that is, before he knows his health status. From the standpoint of an unborn individual with a given risk aversion, the individual faces uncertainty about how his health status will transition from one year to the next, and thus what policies he will purchase and what premiums he will pay. Since individuals differ in their risk aversion, we will calculate this expected utility separately for different risk aversion levels.

To be more specific, suppose we have pricing rule x (e.g., no pre-existing conditions) and equilibrium concept e (e.g., Riley equilibrium). The analysis in the previous section tells us what policy each individual will choose as a function of their health status (acg) and risk aversion (γ), and what premium they will pay. From this, we can get the certainty equivalent mapping $CE_{x,e}(acg, \gamma)$ giving the certainty equivalent of the uncertain consumption that an individual of type (acg, γ) will face in a year because of uncertainty over his health realization in that year. We can also derive a benchmark certainty equivalent level $CE_{90}(acg, \gamma)$ which corresponds to the case in which the individual gets the 90 policy at premium $P_{90} = \underline{AC}_{90}$ (this is the closest to first-best we can get given the restriction to 90 and 60 policies).

We want to compute expected utility starting at age 25 from an ex ante perspective. We assume that our sample represents a steady state population in which the distribution of health types at each age corresponds to the ex ante distribution that any (unborn) individual faces.[21] For each risk aversion level γ, we define the fixed yearly payment starting at age 25 that makes the individual have the same expected utility starting at age 25 under regime x (with equilibrium concept e) as when assured of always getting the 90 policy at the fixed premium $P_{90} = \underline{AC}_{90}$:

$$\sum_{t=25}^{\infty} \delta^t E[-e^{-\gamma\{CE_{x,e}(acg_t,\gamma)+y_{x,e}(\gamma)\}}] = \sum_{t=25}^{\infty} \delta^t E[-e^{-\gamma\{CE_{90}(acg_t,\gamma)\}}],$$

or

$$y_{x,e}(\gamma) = -\frac{1}{\gamma} \ln \left(\frac{\sum_{t=25}^{\infty} \delta^t E[-e^{-\gamma\{CE_{90}(acg_t,\gamma)\}}]}{\sum_{t=25}^{\infty} \delta^t E[-e^{-\gamma\{CE_{x,e}(acg_t,\gamma)\}}]} \right). \tag{2}$$

For a given discount factor $\delta \leq 1$, we calculate $\sum_t \delta^t E[-e^{-\gamma\{CE_{x,e}(acg_t,\gamma)\}}]$ as follows: We first generate the value of $\delta^t e^{-\gamma\{CE_{x,e}(acg_t,\gamma)\}}$ for each individual of age t in our randomly-generated pseudo-population whose risk-aversion is in a band around γ. We then derive the value of $\delta^t E_{x_t}[-e^{-\gamma\{CE_{x,e}(acg_t,\gamma)\}}]$ by calculating the sample mean of those values for those individuals of age t.[22] We then add these values over t to get $\sum_t \delta^t E_{x_t}[-e^{-\gamma\{CE_{x,e}(x_t)\}}]$. We follow a similar procedure to calculate $\sum_t \delta^t E[-e^{-\gamma\{CE_{90}(acg_t,\gamma)\}}]$ and we then form the ratio in (2) to calculate $y_{x,e}(\gamma)$. This gives us the amount of yearly income (starting at age 25) the individual would need to compensate for facing the extra risk in pricing regime x (under equilibrium concept e) compared to the risk if the individual was assured of getting the 90 policy at the fixed premium $P_{90} = \underline{AC}_{90}$. This is our measure of welfare loss in regime x (under equilibrium concept e).

Our main result compares the welfare loss when premiums cannot be priced based on pre-existing conditions – which eliminates reclassification risk but exacerbates adverse selection, to the welfare loss when health risk can be partially priced based (for now) on acg-quartiles. This latter scenario reduces the presence of adverse selection but at the cost of introducing reclassification risk. We also will see what the absolute magnitude of these losses are, which is of independent interest.

Table 6.1 shows the values of $y_{x,e}(\gamma)$ under the Riley equilibrium notion for the cases of no pre-existing conditions ($x =$ "$no-pre$") and pricing based on acg-quartiles ($x =$ "$acg4$"). We take $\delta = 0.975$. The table shows that the welfare gains from eliminating reclassification when pricing on pre-existing conditions is prohibited far exceed any losses this rule introduces due to adverse selection. Moreover, the losses from reclassification with acg-quartile pricing are large, amounting to (at a minimum) between \$800 and \$4600 per year, depending on an individual's level of risk aversion. Recall that the median risk aversion parameter we estimate is 0.0004. The annual loss with acg-quartile pricing at that risk

[21] Recall that the age distribution in our sample is close to uniform, as it should be in a steady state population.

[22] In practice we actually do this for groups of individuals in 5-year age bands and adjust the discount factor accordingly. We also do this only for ages $t \in [25, 65]$.

aversion level is $3044 is about 60% of the size of the average total expenses in the population, which equals $5582 (see Table 7.2 in the Section 7). Table 6.2 shows similar results for Wilson equilibrium. Here the welfare gains from prohibiting pricing based on pre-existing conditions are even larger, as the additional adverse selection losses from prohibiting pricing based on health status are smaller under the Wilson concept than under the Riley concept.

Table 6.1: Welfare Loss in Riley Equilibrium Relative to First-Best

γ	$y_{no-pre,Riley}(\gamma)$	$y_{acg4,Riley}(\gamma)$
0.0002	248	1089
0.0003	445	2166
0.0004	621	3044
0.0005	820	3782
0.0006	998	4345
0.0007	1158	4843
0.0008	1309	5215
0.0009	1448	5553
0.001	1572	5829
0.0011	1681	6061
0.0012	1783	6279
0.0013	1863	6440
0.0014	1939	6607

Table 6.2: Welfare Loss in Wilson Equilibrium Relative to First-Best

γ	$y_{no-pre,Wilson}(\gamma)$	$y_{acg4,Wison}(\gamma)$
0.0002	223	893
0.0003	264	1953
0.0004	308	2823
0.0005	364	3541
0.0006	426	4084
0.0007	483	4563
0.0008	549	4903
0.0009	614	5212
0.001	682	5455
0.0011	745	5661
0.0012	813	5827
0.0013	871	5968
0.0014	932	6099

6 Extensions

6.1 Age-based Pricing

The Affordable Care Act legislation stipulates that insurers offering plans in any state based exchange can vary prices on the basis of age by up to a 3:1 ratio: that is, older cohorts cannot be charged more than three times the premiums of younger buyers. States may further restrict this ratio to be lower, but cannot raise it. This is one of the few exceptions to the case of pure community rating, studied in sections four and five allowed in the current regulations. In this section we use our framework to study market equilibrium and consumer welfare when insurers are allowed to vary prices with age and comment on (i) whether age-based pricing reduces adverse selection and (ii) whether we can expect the 3:1 ratio to be binding in price-setting. We note that age-based pricing inherently does not impact the extent of re-classification risk since the evolution of age is a deterministic process.

Table 7.1 describes the results for our age-based pricing analysis. In order to have a sizable number of observations per group, we group consumers into five year age bins and use that categorization as the basis for differential pricing. The first column shows mean total medical expenses by age cohort: those age 30-34 have a mean of \$3,357 while those age 60-64 have a mean of \$9,413. Thus, if there were only one available plan, the 3:1 age ratio would be non-binding, given $P = AC$ conditioal on age. If we include those age 25-29, with mean expenses \$2,756 the age restriction is binding: however, in both the Affordable Care Act and Massachusetts legislation individuals up to age 30 can buy into a special catastrophic insurance pool for just that age group. As many of the most healthy individuals sign up for the catastrophic plan in Massachusetts, this suggests still that the age restriction of 3:1 might not be binding in practice.[23] Column 2 shows mean expected expenses as a function of the ACG predicted resource indicator. This reveals, similarly a lightly less than 3:1 ratio in projected costs for those 60-64 relative to those 30-34.[24]

The final column studies equilibrium with age-based pricing. Interestingly, allowing for age-based pricing does not change the unraveling of equilibrium that occurs in the case of pure community rating, though allowing for pricing based on ACG quartiles did reduce adverse selection to some extent. For each age group, the RE/sp-NE has 100% of market share in the 60% AV plan, while the market share of that plan is decreasing with age in WE. Age-based pricing undoes some of the transfers from the younger, healthier age groups to the older, sicker groups that occur in pure community rating. However, the distributions of health risk still have substantial enough tails in the younger age group that full unraveling occurs in equilibrium for that group (in quartile based pricing, the distributional tails are substantiall reduced for healthier quartiles).

[23] In Massachusetts, the age rating restriction is more restrictive than the ACA, and doesn't allow age-based pricing at higher than a 2:1 ratio.

[24] This restriction may be binding or not depending on the specific population studied, e.g. the specific state considering setting up an exchange.

Table 7.1: **Population Risks by Age Cohort**

	Expenses		ACG		Sh60	
Ages	Mean	S.D.	Mean	S.D.	WE	RE/sp-NE
All	5582	6495	1.08	1.42	63	100
25-29	2756	4657	0.56	0.79	73	100
30-34	3357	5189	0.71	1.23	75	100
35-39	3762	5029	0.78	1.05	66	100
40-44	4560	5642	0.87	1.04	62	100
45-59	5778	6575	1.06	1.32	60	100
50-54	6722	6831	1.32	1.82	55	100
55-59	8394	7134	1.53	1.58	54	100
60-64	9413	7268	1.77	1.73	53	100

6.2 Preference Heterogeneity

The simulation results reported thus far were based on the assumption that consumers evaluate plans exclusively on the basis of their health risk and risk aversion. In this section we introduce a shock to preferences, to explore the role of preference heterogeneity on market equilibrium. We implement the shock as an additive factor, on top of premiums, that increases or decreases utility for one plan relative to the other. The added shock, taken literally, reflects noisy preferences towards coverage, but it can also be interpreted as capturing partial information about the choice set, infrequent optimization or mistakes. We simulate the functioning of the exchange for different standard deviations of this shock: note that in our baseline estimates, we estimate probit shocks that have standard deviations of 204 and 502 for two of the three plans, relative to the baseline plan.

The exercise serves a dual purpose. First, it helps us evaluate the robustness of our simulation should consumers' behavior depart from our model. Increasing the variance of the shock, interpreted as noise, amounts to allowing larger and larger departures from our prescribed behavior. Second, the shock can be construed as an additional source of gains from trade, in which case the extent of the shock can be interpreted as a comparative static as we boost a trade motive that is orthogonal to risk, thus, potentially reducing adverse selection in the spirit of [10].[25]

Table 7.2 shows market share for the population as a whole as a function of the shock standard deviation. We also investigate age-based pricing in the presense of the shock.

[25] Consider RE/sp-NE when preference shocks are introduced. Profits in 60 (90) become negative (positive) as the original pool gets mixed with health risks from the other policy, reducing ΔAC at each ΔP. Since each policy breaks even in an RE/NE, one would expect an increase in P_{60} and a decline in P_{90}; and higher coverage associated with the decline in ΔP.

Table 7.2: Equilibria with Preference Shocks								
	No Shock Sh60		sd=1000 Sh60		sd=1500 Sh60		sd=2500 Sh60	
Ages	WE	RE/sp-NE	WE	RE/sp-NE	WE	RE/sp-NE	WE	RE/sp-NE
All	63	100	72	100	72	98	62	62
25-29	73	100	72	100	74	79	56	56
30-34	75	100	73	100	74	90	58	58
35-39	66	100	70	100	71	78	55	55
40-44	62	100	70	100	71	84	58	58
45-59	60	100	69	100	71	93	60	60
50-54	55	100	68	100	71	87	57	57
55-59	54	100	69	100	72	87	59	59
60-64	53	100	69	100	72	86	60	60

The RE/sp-NE is fully unraveling, all-in-the 60% AV plan, for both the no shock and the lower shock ($1000 standard deviation) cases. This is true for both the age-based pricing and no age-basd pricing. The low shock is already quite substantial relative to total expenses by group, shown in the first column of Table 7.1. The medium shock ($1500 standard deviation) generates some, limited, coverage in the 90% AV plan. We take both, the low and medium shock predictions, which are similar to the no shock case, as a sign of the robustness of the main results. After introducing some additional trade motives equilibrium predictions remain qualitatively similar. The largest shock ($2500 standard deviation), leads to aproxiamtely 40% market share in comprehensive insurance for the different population groupings, suggesting that very large preference heterogeneity orthoogonal to health risk can help reduce adverse selection and increase coverage. Though not listed, for this large shock there exists a multi-policy NE that is equivalent to the WE, RE, and sp-NE.

Under WE, Market share for the 90% AV plan is non-monotonic as the standard deviation of the preference shock increases. For some age bins these market shares actually decrease when moving from no preference shock to the small and medium preference shocks, while for some they increase. These market shares for all groups are still higher for the largest shock, with $2500 standard deviation. This generally reflects the fact that the equilibrium difference in plan pries in WE, ΔP^w, is not equal to ΔAC (and is weakly less than ΔAC), as discussed in Corollary 1.[26] Note that the WE for the largest

[26] Wilson equilibrium premiums always involve a weakly lower ΔP than the lowest break-even ΔP with a positive 60 share. In fact, if (P^*_{90}, P^*_{60}) are the prices at this lowest break-even ΔP, it can be seen that the WE ΔP^w is exactly $\arg\max_{\Delta P \leq \Delta P^w} \Pi(P^*_{60} + \Delta P, P^*_{60})$. In general, it appears that an increase in the shock could change ΔP^w in either direction as it can either lower or raise the profit effect of a small change in P_{90}. (The marginal profit effect has two components – the reduced revenue from individuals choosing the 90 policy, and the marginal difference between policies for the marginal individual; both components can change in either direction with an increase in the shock.) This is consistent with what can be seen in Table 7.2, where introduction of the smallest shock increases coverage for some age cohorts and reduces it for other in the WE.

preference shock does not involve cross-subsidization, as is the same as the other equilibria. At that size shock, a multi-policy deviation to a lower ΔP is not profitable because marginal cross-subsidization is weaker given that the prefernce shocks already imply a large extent of cross-subsidization.

6.3 Health-Status Based Pricing

Our analysis in sections 4 and 5 studied the cases of pricing based on no pre-existing conditions, and pricing based on ACG quartiles (to represent pricing based on pre-existing conditions in a stylized setting). As an extension, we investigate how the welfare trade-off between adverse selection and re-classification risk manifests when we allow different levels of coarseness or fineness based on health status. We perform the analysis of sections 4 and 5 for a range of cases, from the case of no pricing based on pre-existing conditions to the case where insurers can price discriminate at the individual level, i.e. they can price based on all available information and price each individual at expected cost.

Table 7.3 summarizes aggregate equilibrium market share and corresponding welfare for counterfactual pricing regimes that vary with the health status information that insurers can price based on. The table reveals that, as pricing becomes finer and finer, the market share of comprehensive coverage (90% AV) increases: for RE/sp-NE with quartile based pricing 15% of consumers enroll in this plan, with 20 distinct categories 53% of people enroll, and with individual-level pricing 100% enroll (by construction). The welfare columns illustrate the impact of finer pricing on the trade-off between adverse selection and re-classification risk. Finer pricing increases the market share of the 90% AV plan, reducing adverse selection, but leaves insurees subject to repricing as their health status evolves. The welfare loss from increased reclassification risk swamps the welfare gain from less adverse selection: the welfare loss from pricing on 20 categories is almost three times that from pricing on four categories.

Table 7.3: Health-Status Based Pricing

# ACG groups	Riley/sp-NE Sh90	Riley/sp-NE Median Welfare Loss ($/year)	Wilson Sh90	Wilson Median Welfare Loss ($/year)
0	0	621	0	203
2	3	1926	43	1635
4	15	3044	46	2823
6	22	3900	53	3716
8	22	4346	52	4158
10	23	4992	55	4849
20	53	8450	55	8349
∞	100	16580	100	16580

6.4 Participation and Subsidies

While the individual mandate will be a component of all exchanges to be implemented under the Affordable Care Act, in reality the mandate is a tax that is paid to the IRS when someone who can afford insurance does not posses it. It is plausible that certain individuals, especially healthy ones, will decide to pay the mandate penalty rather than pay a more expensive insurance premium. In Massachusetts, where an individual mandate has been in place since 2006, the penalty has been 50% of the cost of the least generous (Bronze) plan available through the exchange (Commonwealth Connector).[27] This is on average slightly large than the initial penalty under the Affordable Care Act, which is the maximum of $695 per household member (up to three) and 2.5% of household income. In Massachusetts, only 3% of the population did not comply with the individual mandate in 2008, with many of those people receiving exemptions from the mandate penalty due to low income.[15]

We use our framework to the subsidy (or equivalently tax) that would be required to get a certain proportion of the population to comply with the mandate and join the exchange market. To do this we compare (i) the expected utility of a consumer from joining their favorite insurance plan in the market to (ii) their expected utility from being uninsured. In our analysis, uninsured means that the consumer pays zero premium and is at full risk for total cost of their health expenditures. Note that this computation is quite stylized: we do not plan options with less than 70% actuarial equivalence value in our data: we extrapolate preferences from our observed setting and continue to assume zero moral hazard, which are assumptions that can both be questioned.

Conditional on the pricing regulation environment and equilibrium concept, we calculate the proportion of individuals that would opt out of the market for a given subsidy. Figure 4 shows the proportion of people who opt out of the exchange for the case of pricing based on no pre-existing conditions, similiar to the legislation in the Affordable Care Act. Approximately 20% of individuals opt out with no penalty / subsidy under both equilibrium concepts. The required subsidy to lower the proportion of people optin out to 15% is $1576 per year under RE/sp-NE, for 10% $2706 is required and for 5% 3556 is required. For 0%, the premium of the lowest cost plan has to, essentially, be fully funded. The numbers for Wilson equilibrium are similar, but slightly lower. The $2706 required to have only 5% opting out is similar in magnitude to what we see in the Massachusetts exchange, though this may be a coincidence since our environment is quite stylized.

6.5 Rebalancing of the Population

The analysis to this point relied on health choice and utilization data from a large firm with approximately 10,000 employees and 20,000 covered lives. While these data have a lot of depth on dimensions that are essential to model health risk and risk preferences, they represent a specific population working for a specific large employer. Our results thus represent the case of exchange design as if our population

[27] In 2010, in the 02138 zip code in Cambridge, MA, this penalty would have been $5,500 for a family with two 40 year old parents, $3,300 for a couple with two 35 year olds, and $1,434 for a single 31 year old.[15]

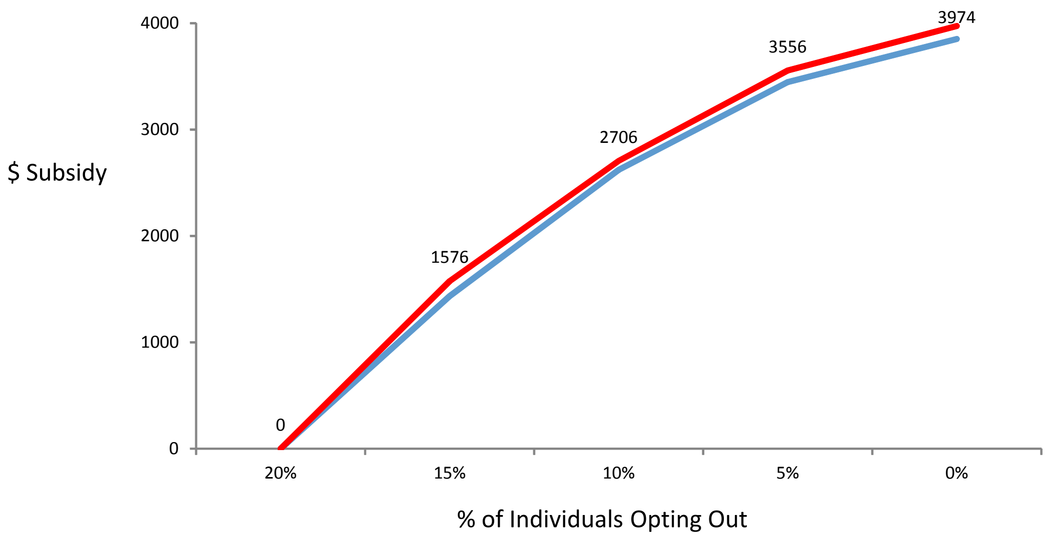


Figure 2: Figure 4 shows the number of individuals that would opt out of the exchange market with no pricing based on pre-existing conditions. The blue line represents the subsidy required under Wilson equilibrium, the red line under Riley / single-policy Nash equilibrium.

were the population of interest. While some large employers may strive to set up insurance choice environments with a high degree of insurer competition and speific pricing restrictions (using the leeway given by ERISA), the market structure and regulator decisions we study are meant to mimic those that national policymakers and state regulators have faced and will face when setting up the health exchanges tied to the Affordable Care Act. In this section we extend the analysis by applying our framework to a more externally relevant sample from the Medical Expenditures Panel Survey (MEPS), a survey that was specifically created to study medical care decisions for a nationally representative population. This analysis servers two purposes: (i) to study whether our broad conclusions are robust to changes in the population composition and (ii) to apply our analysis to a nationally representative sample.

We study individuals in the MEPS data from 2004 to 2008. The data are structured as overlapping two-year panels where each individual is in the panel for two consecutive years, and a new panel of individuals enters in each year. Table 8 describes the number of individuals in the data over the sample period. Note that because of the panel structure, an individuals 'counts' twice, once in each year they appear.

Year	Number of Individuals
2004	34,403
2005	33,961
2006	34,145
2007	30,964
2008	33,066

We base our analysis on two distinct sample that may be of interest to an exchange regulator, described in Table 9, along with statistics for the full MEPS sample. Column 1contains the summary statistics for the entire sample, with no sample cuts. Column 2 contains summary statistics for the sample of individuals between ages 25-65. In this table and in the analysis we split individuals within the same family into distinct individuals and run our simulations as if the entire market is an individual market.[28] Column 3 describes the sample of individuals between 25-65 who are uninsured, unemployed, or are employed for an employer that does not offer coverage (implying that if the individual has coverage, it is from the individual market). We run our exchange market analysis for the samples in Column 2 and Column 3. Column 2 is of interest if all individuals 25-65 participate in an exchange (e.g. the exchange insurers everybody). This sample is about half of the overall MEPS sample. Column 3 is of interest because it covers the population that will enter exchanges immediately when they are set up (uninsured and insured on individual market). This sample is about 15% of the overall MEPS sample, similar in age, lower in income, and more likely to come from the South. We note that Table 9 describes the unweighted samples: MEPS also provides sample weights, which we use in our final analysis, as described shortly.

Table 10 describes the insurance coverage for each of the three samples described in Table 9. For individuals in Column 2 (age 25-65), 64% have some form of private insurance, 12% are on Medicard, and 22% are uninsured. 76% of families have employer insurance offered at some point, while 62% always have an offer of employer insurance. Column 3, by design, has 83% of people uninsured and 17% of people insured on the individual market (with no employer insurance offer).

Our analysis matches individuals in the employer data used throughout our analysis to the two MEPS populations of interest (Coluns 2 and 3) and creates two new samples with demographic weights similar to the MEPS samples. We match individuals in our data to those in the MEPS data based on three demographics: age, income, and gender. To do this, we probabalistically model cells of age,

[28] Once the structure of family based premium setting is clear, we could run an additional analysis taking that into account, to the extent that it differs from aggregating up individual premiums into family premiums, which our current approach is essentially equivalent to.

	Entire MEPS (1)	All Ind. 25-65 (2)	25-65 Unins/Ind (3)
N - Individual-Year Obs.	166,539	81,733	21,856
N - Individuals in Panel	105,353	51,922	13,804
N - Family-Year Obs.	58,647	-	-
N - Families in Panel	36,317	-	-
Avg. Family Members	2.90	-	-
Age-Individual			
Mean	33.82	43.15	42.6
10th Qtile	5	28	27
25th Qtile	14	34	32
Median	32	43	42
75th Qtile	51	52	52
90th Qtile	66	59	60
Gender-Individual			
Male %	47.7%	46.6%	50.2%
Total Income-Family-Year* **			
Mean	53613	64058	42746
10th Qtile	9240	12733	8000
25th Qtile	19000	26000	17068
Median	39080	50000	31114
75th Qtile	72375	85584	54995
90th Qtile	115086	131080	89600
Wage Income-Family-Year**			
Mean	44583	59945	38882
10th Qtile	0	7348	300
25th Qtile	8000	24000	14280
Median	32000	48300	30000
75th Qtile	65000	83753	52000
90th Qtile	104438	124996	82680
Region-Individual			
Northeast	14.5%	15.0%	10.1%
Midwest	19.2%	19.6%	15.0%
South	38.3%	38.7%	46.3%
West	26.9%	26.8%	28.7%

Table 4: This table describes demographic data for key samples of interest in the MEPS data for the pooled data from 2004-2008. A more detailed description of each column's sample is contained in the text.
Missing for 60% of family observations, not sure why.
*In individual samples can count same family income twice since two individuals can be from same family.

	Entire MEPS (1)	All Ind. 25-65 (2)	25-65 Unins/Ind (3)
Family-Year: Coverage Type*			
Private (Employer or Ind.)	66.3%	73.3%	41.0%
Medicaid (someone)	30.7%	33.4%	45.4%
Medicare (someone)	29.01%	14.0%	16.4%
Uninsured** (someone)	26.7%	35.0%	84.7%
Only Public in Fam	22.5 %	15.1%	0%
Always Offered Employer (someone)	48.8 %	62.1%	–
Offered Employer Sometimes (someone)	62.0%	76.1%	–
Family Member Emp. Always	69.7%	84.7%	76.2%
Family Member Emp. Once	77.5%	92.3%	87.4%
Individual-Year: Coverage Type*			
Private (Employer or Ind.)	54.5%	64.0%	16.8%
Medicaid	25.4%	12.4%	0.72%
Medicare	13.4%	3.9% 1	.25%
Uninsured**	16.6%	22.3%	83.2%
Only Public	27.6%	12.7%	0%
Always Offered Employer	21.3 %	38.9%	–
Offered Employer Sometimes	32.5%	55.0%	–
Individual Emp. Always	37%	65.4%	37.5%
Individual Emp. Once	48%	78.3%	48.0%

Table 5: This table describes insurance coverage, expenditures, and other stats in the MEPS data for the pooled data from 2004-2008. A more detailed description of each column's sample is contained in the text.

*Coverage type is family ever have this kind of coverage (for any member) throughout the year, so should add to more than 100%.

**Uninsured variable occurs when none of other coverage types are held, so is uninsured for whole year, while others show if you ever hold that coverage.

gender, and income in the MEPS samples, and then draw randomly from individuals in those bins in our data with weights proportional to the MEPS cell weights. We note that before we construct the MEPS cell weights, we incorporate the sample weights in the MEPS data, which are intended to correct for sampling and response issues. Table 11 describes the age, income, and gender cell multivariate cell weights for each of the two MEPS samples. We model the multivariate distribution of age and income fully non-parametrically, and assume that the probability of gender conditional on income is the same for all ages. Note that, unsurprisingly, the uninsured / individual market population has lower income, is younger, and is more likely to be male than that the full 25-65 sample. We note that we use matching from the MEPS data to our data, and then use the expenses and risk preferences from our data to simulate equilibrium as opposed to explicitly using the MEPS data in our framework. This is because (i) our data provide a clean way to identify risk preferences (which the MEPS data does not) and (ii) we have richer health status and expense distribution information for our population.

We note that in this analysis, at this point, we do not match our sample to MEPS using health expenditure data, though we have expenditure data for MEPS as well. In this version of the analysis we match only on the three demographics under the assumption that health status in the nationally representative sample will be similar to health status on our sample conditional on those demographics. We are currently exploring additional matching algorithms that do incorporate expenses. Table 12 present a profile of health expenses for both MEPS samples.[29]

Given this population weighting procedure, we study market equilibrium for each weighted sample for each equilibrium concept and for the pricing regimes of (i) pure community rating and (ii) ACG quartile based pricing. Table 13.1 presents the equilibrium simulation results for both matched samples (All 25-65 and those who are uninsured / individual market insured). This is analagous to Table 4.1 presented earlier for our own data. We find that our results our robust to re-weighting by the MEPS demogrpahic weights: for both samples the Riley equilibrium and single-policy Nash have full unraveling while the Wilson equilibrium has some market share in the 60% and 90% due to cross-subsidization. These results are quite similar to those in table 4.1, suggesting that full unraveling, everyone in the 60% AV plan, is a robust result in the context of our model.

[29] Future work can at least compare expenditure distributions for the weighted sample from our data with those from the MEPS data.

MEPS Weights Incorporated					
All 25-65 Sample					
Age Bucket / Fam. Wages	0-$35,000	$35,000-$70,000	$70,000-$105,000	$\geq$ 105,000	**Total**
25-29	4.1%	4.5	2.7	1.9	13.1%
30-34	3.3%	4.4	2.6	1.9	12.3%
35-39	3.5%	4.2	2.8	2.3	12.9%
40-44	3.6%	4.5	3.0	2.8	13.9%
45-49	3.5%	4.2	3.0	3.1	13.9%
50-54	3.5%	3.8	2.8	2.9	13.1%
55-59	3.8%	3.2	2.3	2.3	11.7%
60-64	4.4%	2.3	1.3	1.2	9.2%
Total	29.7%	31.1%	20.5%	18.4%	100%
% Male by Income*	45.6%	49.9%	50.3%	51.4%	

25-65 Unins./ Private					
Age Bucket / Fam. Wages	0-$35,000	$35,000-$70,000	$70,000-$105,000	$\geq$ 105,000	**Total**
25-29	7.4%	5.0	1.9	1.6	15.9%
30-34	6.0%	4.4	1.3	0.7	12.4%
35-39	6.4%	3.5	1.1	0.6	11.6%
40-44	6.1%	4.0	1.4	0.8	12.2%
45-49	6.2%	3.1	1.6	0.9	10.8%
50-54	5.9%	2.9	1.1	0.9	10.8%
55-59	7.0%	2.5	1.1	0.8	11.4%
60-64	10.1%	2.3	0.8	0.8	14.0%
Total	55.1%	27.7%	10.3%	7.1%	100%
% Male by Income*	51.4%	56.2%	55.4%	56.8%	

Table 6: This table describes the discrete age probabiliities for different age / gender / income categories in the sample of all individuals in MEPS 25-65 in age and in the sample of individuals who are uninsured or have individual market private coverage. This is the version of the chart that incorporates MEPS weights, so is the one we should use in our final analysis.

*Percentages of gender across age are essentially constant conditional on income, which is why those figures are not presented here.

MEPS Weights Incorporated All 25-65 Sample							
Age Bucket / Quantile	10th	25th	50th	75th	90th	95th	**Mean**
25-29	0 (0)	0 (203)	125 (843)	620 (2833)	2109 (7638)	4155 (12007)	**997 (2820)**
30-34	0 (0)	0 (241)	224 (940)	922 (3179)	2815 (9040)	5582 (13122)	**1376 (3146)**
35-39	0 (0)	0 (239)	331 (925)	1314 (2928)	3499 (8158)	6333 (13595)	**1696 (3126)**
40-44	0 (0)	25 (258)	450 (967)	1669 (2955)	4513 (7844)	9099 (13843)	**2235 (3544)**
45-49	0 (0)	115 (365)	703 (1342)	2425 (3827)	6423 (9143)	12125 (15505)	**3016 (3838)**
50-54	0 (90)	221 (563)	1114 (1860)	3385 (4744)	8562 (10683)	16271 (17135)	**4187 (4551)**
55-59	0 (102)	410 (781)	1837 (2437)	4953 (5820)	11929 (13615)	21069 (22741)	**5315 (6129)**
60-64	71 (255)	707 (1109)	2337 (2906)	5916 (6771)	15261 (14493)	27033 (24997)	**6790 (6666)**

25-65 Unins./ Private							
Age Bucket / Quantile	10th	25th	50th	75th	90th	95th	**Mean**
25-29	0 (0)	0 (0)	0 (166)	173 (758)	819 (2959)	1824 (5502)	**391 (952)**
30-34	0 (0)	0 (0)	0 (180)	254 (852)	1062 (3234)	2024 (6095)	**608 (1322)**
35-39	0 (0)	0 (0)	0 (174)	328 (1024)	1650 (3187)	3164 (5748)	**744 (1223)**
40-44	0 (0)	0 (0)	50 (308)	750 (1459)	2929 (3966)	4500 (6908)	**1381 (2449)**
45-49	0 (0)	0 (0)	120 (425)	857 (1846)	3108 (4566)	6719 (9658)	**2089 (1967)**
50-54	0 (0)	0 (144)	340 (798)	1576 (2866)	5590 (7462)	11851 (12952)	**2474 (3085)**
55-59	0 (0)	24 (176)	1076 (1312)	3565 (3996)	9290 (9990)	16419 (19459)	**3898 (4941)**
60-64	0 (60)	449 (732)	1966 (2398)	5166 (5730)	13749 (12017)	24157 (21839)	**6003 (6043)**

Table 7: This table describes the expenditure quantiles for all individuals in MEPS 25-65 in age and in the sample of individuals who are uninsured or have individual market private coverage. Female numbers presented in parantheses.

Table 13.1: Equilibria Without Pre-existing Conditions

Sample: All 25-65						
	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley / sp-NE	3920	100		-	0	-
Wilson	3803	64	1328	6553	36	11007
Double Nash	Does note exist					

Sample: Uninsured / Ind. Market 25-65						
	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley / sp-NE	4203	100	4203	-	0	-
Wilson	4106	66	1587	6906	34	11819
Double Nash	Does not exist					

Table 13.2 presents the results when we allow for health status based pricing, using ACG quartiles. Table 13.2 specifically shows the results for those in the healthiest quartile. The results are again similar to those based solely on our employer data presented in Table 4.2. The All 25-65 weighted sample has 42% market share in the 90% AV plan in the RE/sp-NE, indicating that there is less adverse selection under health status based pricing. For the uninsured / individual market sample 36% enroll in the 90% AV plan. This table again shows the robustness of our initial results with respect to weighting on demographics form a more representative sample. Further, it reveals that the equilibrium for the All

25-65 population is not particularly different from the uninsured / private market sample. These results are subject to the assumption of equal health status conditional on age, gender and income across these two groups. In the even that uninsured people are likely to be healthier, as Table 12 suggests, these results would change. In future specifications with matching based on health status, we will contorl for this.

Table 13.2: Equilibria: ACG Quartile 1 (Healthy)

Sample: All 25-65

	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley / sp-NE	215	58	215	1515	42	1515
Wilson	215	58	215	1515	42	1515
Double Nash	215	58	215	1515	42	1515

Sample: Uninsured / Ind. Market 25-65

	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley / sp-NE	231	64	231	1581	36	1590
Wilson	231	64	231	1581	36	1590
Double Nash	231	64	231	1581	36	1590

Table 13.3 presents the results for both samples for ACG quartile two, which is the second healthiest quartile. Again, the results are similar across the two different weighted samples and the original data, results shown in table 4.3.

Table 13.3: Equilibria ACG Quartile 2 (2nd Healthiest)

Sample: All 25-65

	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley / sp-NE	1167	79	1167	3780	21	3780
Wilson	1079	54	924	3179	46	3368
Double Nash	Does not exist					

Sample: Uninsured / Ind. Market 25-65

	P_{60}	Sh_{60}	AC_{60}	P_{90}	Sh_{90}	AC_{90}
Riley / sp-NE	1236	84	1236	3886	16	3886
Wilson	1131	55	965	3181	45	3384
Double Nash	Does not exist					

Overall, the analysis of MEPS data in this section suggests that, at a first pass, our main results are robust to different weighting of demographics to reflect a more nationally representative sample. Extending the analysis in the section to allow matching based on health status (into the All 25-65 and uninsurd/individual market groups) will provide further checks that the results are robust. Further, the analysis here illustrates that close to full unraveling is likely to occur under pure community rating if there is ineffective risk adjustment and there are no additional preference considerations.

6.6 Analysis with Moral Hazard

6.7 Longer-term Contracts

6.8 Risk Adjustment

7 Conclusion

TBA

References

[1] Bundorf, K. and J. Levin, and N. Mahoney (2010), "Pricing and Welfare in Health Plan Choice," Stanford University working paper.

[2] Cardon, J. and I. Hendel (2001), "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey," *RAND Journal of Economics* 32: 408-27.

[3] Carlin, C. and R. Town, R. (2009), "Adverse Selection, Welfare, and Optimal Pricing of Employer Sponsored Health Plans," University of Minnesota working paper.

[4] Cochrane, J. (1995), "Time-Consistent Health Insurance," *Journal of Political Economy* 103: 445-73.

[5] Cutler, D. and S. Reber, S. (1998), "Paying for Health Insurance: The Tradeoff Between Competition and Adverse Selection," *Quarterly Journal of Economics* 113: 433-66.

[6] Einav, L., A. Finkelstein, and J. Levin (2010a), "Beyond Testing: Empirical Models of Insurance Markets," *Annual Review of Economics* 2: 311-36.

[7] Einav, L., A. Finkelstein, A., and P. Schrimpf, P. (2010b), "Optimal Mandates and the Welfare Cost of Asymmetric Information: Evidence from the U.K. Annuity Market," Econometrica 78: 1031-92.

[8] Einav, L., A. Finkelstein, S. Ryan, P. Schrimpf, and M. Cullen (2011), "Selection on Moral Hazard in Health Insurance," Stanford University working paper.

[9] Engers, M. and L. Fernandez (1987), "Market equilibrium with Hidden Knowledge and Self-selection," *Econometrica* 55: 425-39.

[10] Handel, B. (2011), "Adverse Selection and Switching Costs in Health Insurance Markets: When Nudging Hurts," NBER working paper no. 17459.

[11] Hendel, I. and A. Lizzeri (2003), "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance," *Quarterly Journal of Economics* 118: 299-327.

[12] Herring, B. and M. Pauly (2006), "Incentive-Compatible Guaranteed Renewable Health Insurance Premiums," *Journal of Health Economics* 25: 395-417.

[13] Finkelstein, A., K. McGarry, and A. Sufi, A. (2005), "Dynamic Inefficiencies in Insurance Markets: Evidence from Long-Term Care Insurance," MIT working paper.

[14] Koch, T. (2011), "One Pool to Insure Them All? Age, Risk, and the Price(s) of Medical Insurance," UC-Santa Barbara Working Paper.

[15] Kolstad, J. and A. Kowalski (2012), "Mandate-Based Health Reform and Evidence from the Labor Market: Evidence from the Massachusetts Reform," Wharton working paper.

[16] Lustig, J. (2010), "Measuring Welfare Losses from Adverse Selection and Imperfect Competition in Privatized Medicare," Boston University working paper.

[17] Miyazaki, H. (1977), "The Rat Race and Internal Labor Markets," *Bell Journal of Economics* 8: 394-418.

[18] Riley, J. G. (1985), "Informational Equilibrium," *Econometrica* 47: 331-59.

[19] Rothschild, M. and J. E. Stiglitz (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics* 90: 629-49.

[20] Wilson, C. (1977), "A Model of Insurance Markets with Imperfect Information," *Journal of Economic Theory* 16: 167-207.

8 Appendix I

8.0.1 Proof of Proposition 1

We begin by considering which price offers are "safe" in the sense that they remain profitable regardless of any additional offers being introduced.

Lemma 1. *If* (P_{90}, P_{60}) *are the lowest prices offered for the 90 policy and the 60 policy, respectively, then any* $P''_{60} < P_{60}$ *such that* $\Pi_{60}(P_{90}, P''_{60}) \geq 0$ *is safe.*

Proof. Any prices $\widehat{P} = (\widehat{P}_{90}, \widehat{P}_{60})$ with a $\widehat{P}_{60}$ lower than P''_{60} gives the firm offering P''_{60} a profit of zero. Any prices $\widehat{P}$ with $\widehat{P}_{90} \geq P_{90}$ and $\widehat{P}_{60} \geq P''_{60}$ cannot make the firm offering P''_{60} incur losses. Finally, any prices $\widehat{P}$ with $\widehat{P}_{90} < P_{90}$ and $\widehat{P}_{60} \geq P''_{60}$ weakly lowers the sales of the firm offering P''_{60}. If that firm makes no sales at $(\widehat{P}_{90}, P''_{60})$, then its profit is zero. If it has positive sales at $(\widehat{P}_{90}, P''_{60})$, then it must have $\Pi_{60}(\widehat{P}_{90}, P''_{60}) \geq 0$ since $AC_{60}(\widehat{P}_{90}, P''_{60}) \leq AC_{60}(P_{90}, P''_{60}) \leq P''_{60}$. □

Definition 2. *The* ***lowest safe 60 price given*** P_{90} *is* $\underline{P}_{60}(P_{90}) \equiv \min\{P''_{60} : \Pi_{60}(P_{90}, P''_{60}) \geq 0\}$.

Remark 1. *Define the price* $\widetilde{P}_{60}(P_{90}) \equiv \{\ \widetilde{P}_{60} : \widetilde{P}_{60} = AC_{60}(P_{90} - \widetilde{P}_{60})\}$. *Note that this equality has a unique solution, which is continuous and weakly increasing in* P_{90} *and strictly increasing at any* P_{90} *at which* $P_{90} - \widetilde{P}_{60}(P_{90}) \in (\underline{\theta}, \overline{\theta})$ *(so that there are sales of both policies), which occurs when* $P_{90} \in (\underline{AC}_{60} + \underline{\theta}, \overline{AC}_{60} + \overline{\theta})$. *This can be seen in Figure 1. Also,* $\widetilde{P}_{60}(\underline{AC}_{60} + \underline{\theta}) = \underline{AC}_{60}$ *and* $\widetilde{P}_{60}(\overline{AC}_{60} + \overline{\theta}) = \overline{AC}_{60}$. *Moreover, one can see in the figure that* $AC_{60}(P_{90} - \widetilde{P}_{60}(P_{90}))$ *is strictly increasing for this range of* P_{90}, *which also means that* $P_{90} - \widetilde{P}_{60}(P_{90})$ *is strictly increasing, and these are weakly increasing at all* P_{90}. *Moreover, observe that*

$$\underline{P}_{60}(P_{90}) = \left\{ \begin{array}{cc} P_{90} - \underline{\theta} & \text{if } P_{90} \leq \underline{AC}_{60} + \underline{\theta} \\ \widetilde{P}_{60}(P_{90}) & \text{if } P_{90} \in (\underline{AC}_{60} + \underline{\theta}, \overline{AC}_{60} + \overline{\theta}) \\ \overline{AC}_{60} & \text{if } P_{90} \geq \overline{AC}_{60} + \overline{\theta} \end{array} \right\}.$$

When $P_{90} \leq \underline{AC}_{60} + \underline{\theta}$, *all consumers buy the 90 policy at prices* $(P_{90}, \underline{P}_{60}(P_{90}))$; *when* $P_{90} \in (\underline{AC}_{60} + \underline{\theta}, \overline{AC}_{60} + \overline{\theta})$ *there are positive sales of both policies at prices* $(P_{90}, \underline{P}_{60}(P_{90}))$; *and when* $P_{90} \geq \overline{AC}_{60} + \overline{\theta}$ *all consumers buy the 60 policy at prices* $(P_{90}, \underline{P}_{60}(P_{90}))$. *Finally, observe that* $\Pi_{60}(P_{90}, \underline{P}_{60}(P_{90}))$ *is continuous in* P_{90}.

Remark 2. *The lowest safe 60 price given* P_{90} *is the lowest* single policy *reaction in* P_{60} *that is safe. However, observe that if a two-policy reaction* (P''_{90}, P''_{60}) *is safe and makes a deviating firm offering* P'_{90} *earn strictly negative profits, then the single-policy reaction* P''_{60} *is also safe and makes the deviating firm earn strictly negative profits. (This follows because the only case in which the two differ in their outcome is if* $P''_{90} = P'_{90}$ *and the two firms split the 90 consumers, each making losses on them, while* $\Pi_{60}(P'_{90}, P''_{60}) > 0$, *which by Lemma 1 means that the single-policy reaction* P''_{60} *is safe.) Hence, in looking at safe reactions to single-policy deviations in* P_{90}, *we can restrict attention to single-policy safe reactions in* P_{60}.

Lemma 2. *If at* $(P_{90}, \underline{P}_{60}(P_{90}))$ *we have positive sales of the 90 policy and* $\Pi_{90}(P_{90}, \underline{P}_{60}(P_{90})) \geq 0$, *then* $\Pi_{90}(P_{90}, P_{60}) > 0$ *at all* $P_{60} > \underline{P}_{60}(P_{90})$.

Proof. Since there are positive sales of the 90 policy at $(P_{90}, \underline{P}_{60}(P_{90}))$, for any $P_{60} > \underline{P}_{60}(P_{90})$ we have $\Pi_{90}(P_{90}, P_{60}) > 0$ since increases in P_{60} increase sales of the 90 policy and lower AC_{90}. □

Remark 3. *Lemma 2 implies that a single-policy deviation to* P'_{90} *can be rendered unprofitable by a safe reaction only if it is rendered unprofitable by a single-policy reaction to* $\underline{P}_{60}(P'_{90})$.

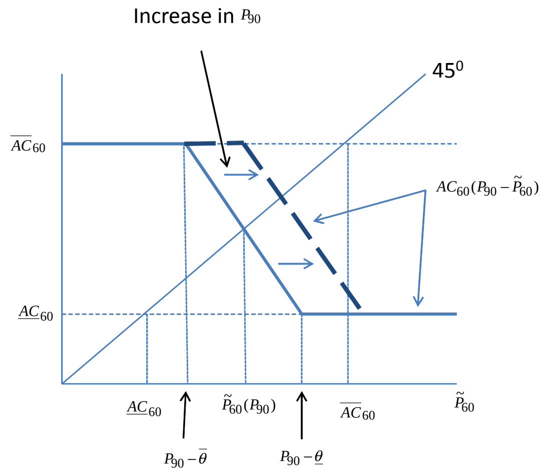


Figure 3: Graphical Description of Riley Equilibrium

8.1 Characterization of Equilibria

Lemma 3. *If* (P^*_{90}, P^*_{60}) *is a Riley equilibrium, then there can be no* $P'_{60} < P^*_{60}$ *such that* $\Pi_{60}(P^*_{90}, P'_{60}) > 0$*.*

Proof. Suppose by contradiction that (P^*_{90}, P^*_{60}) is a Riley equilibrium and there is a $P'_{60} < P^*_{60}$ such that $\Pi_{60}(P^*_{90}, P'_{60}) > 0$. Then by Lemma 1 a single-policy deviation that offers P'_{60} is safe, and hence there is no reaction that can render it unprofitable. □

Corollary 3. *If* (P^*_{90}, P^*_{60}) *is a Riley equilibrium, then* $\Pi_{60}(P^*_{90}, P^*_{60}) \leq 0$*.*

Proof. If $\Pi_{60}(P^*_{90}, P^*_{60}) > 0$, then for small $\varepsilon > 0$ we have $\Pi_{60}(P^*_{90}, P^*_{60} - \varepsilon) > 0$, in contradiction to Lemma 3. □

Lemma 4. *If* (P^*_{90}, P^*_{60}) *is a Riley equilibrium, then* $\Pi_{90}(P^*_{90}, P^*_{60}) \leq 0$*.*

Proof. The result is immediate if $\Delta P^* \geq \overline{\theta}$ so that the 90 policy makes no sales. So suppose that $\Delta P^* < \overline{\theta}$ (the 90 policy has positive sales) and that contrary to the claim (P^*_{90}, P^*_{60}) is a Riley equilibrium with $\Pi_{90}(P^*_{90}, P^*_{60}) > 0$. Consider a single-policy deviation to $P'_{90} = P^*_{90} - \varepsilon$ for small $\varepsilon > 0$ such that $\Pi_{90}(P^*_{90} - \varepsilon, P^*_{60}) > 0$. Now, no reaction that has $P''_{90} < P'_{90}$ can make the deviator incur a loss. As noted in Remarks 2 and 3, $P^*_{90} - \varepsilon$ can be rendered unprofitable by a safe reaction only if it can be rendered unprofitable by a single-policy reaction to $\underline{P}_{60}(P^*_{90} - \varepsilon)$. By Corollary 3, we know that

$\Pi_{60}(P^*_{90}, P^*_{60}) \leq 0$. If $\Pi_{60}(P^*_{90}, P^*_{60}) < 0$, then $P^*_{60} < \underline{P}_{60}(P^*_{90})$. Since $\underline{P}_{60}(P_{90})$ is continuous in P_{90}, for small enough ε there is no safe reaction that is below P^*_{60}, so $P^*_{90} - \varepsilon$ cannot be rendered unprofitable by a safe reaction. If instead $\Pi_{60}(P^*_{90}, P^*_{60}) = 0$, then $P^*_{60} = \underline{P}_{60}(P^*_{90})$, which implies that as $\varepsilon \to 0$ we have $\underline{P}_{60}(P^*_{90} - \varepsilon) \to P^*_{60}$. Since this implies $\Pi_{90}(P^*_{90} - \varepsilon, \underline{P}_{60}(P^*_{90} - \varepsilon)) \to \Pi_{90}(P^*_{90}, P^*_{60}) > 0$, for small enough ε the devaition cannot be rendered unprofitable by a safe reaction. This contradicts that (P^*_{90}, P^*_{60}) is a Riley equilibrium. □

Corollary 4. *If (P^*_{90}, P^*_{60}) is a Riley equilibrium, then $\Pi_{90}(P^*_{90}, P^*_{60}) = \Pi_{60}(P^*_{90}, P^*_{60}) = 0$.*

Proof. By Lemma 3 and Lemma 4 we have $\Pi_{60}(P^*_{90}, P^*_{60}) \leq 0$ and $\Pi_{90}(P^*_{90}, P^*_{60}) \leq 0$. But if either is strictly negative, then some firm must be earning strictly negative profits, and would do better by dropping all of its policies. □

Corollary 5. *There is a Riley equilibrium in which all consumers buy the 90 policy if and only if there is no P'_{60} such that $\Pi_{60}(\underline{AC}_{90}, P'_{60}) > 0$.*

Proof. By Corollary 4, $P^*_{90} = \underline{AC}_{90}$. Necessity follows then by Lemma 3. For sufficiency, suppose that there is no P'_{60} such that $\Pi_{60}(\underline{AC}_{90}, P'_{60}) > 0$ but that $P' = (P'_{90}, P'_{60})$ is a profitable Riley deviation. If the deviator makes no sales of the 90 policy, then its sales would be the same with a single policy deviation to P'_{60}, so we would have $\Pi_{60}(\underline{AC}_{90}, P'_{60}) > 0$ – a contradiction. Suppose, instead, that the deviator does make positive sales of the 90 policy (which requires that $P'_{90} \leq \underline{AC}_{90}$). Since it is a profitable deviation it must also make positive sales of the 60 policy (since $P'_{90} \leq \underline{AC}_{90}$, it can't make a positive profit selling only the 90 policy). But, in this case $P'_{90} \leq \underline{AC}_{90} < AC_{90}(\Delta P')$, so we must have $\Pi_{60}(P'_{90}, P'_{60}) > 0$ since it is a profitable deviation. In turn, letting $\widehat{P}_{60} \equiv \underline{AC}_{90} - \Delta P' \geq P'_{60}$, this implies that $\Pi_{60}(\underline{AC}_{90}, \widehat{P}_{60}) \geq \Pi_{60}(P'_{90}, P'_{60}) > 0$ – a contradiction. □

Remark 4. *This Corollary shows that a Riley equilibrium in which all consumers buy the 90 policy exists if and only if such a Nash equilibrium (with either single or multi-policy deviations) exists.*

We also have the following result:

Lemma 5. *Suppose $P^* = (P^*_{90}, P^*_{60})$ with positive sales of the 60 policy (so $\Delta P^* \in (\underline{\theta}, \bar{\theta}]$) and both policies break even. Then P^* is a Riley equilibrium iff there are no single-policy Riley profitable deviations in P_{90}.*

Proof. Since both policies are breaking even and there are positive sales of the 60 policy at P^* a single-policy deviation in only P_{60} is unprofitable, because

$$0 = \Pi_{60}(P^*_{90}, P^*_{60}) \geq \Pi_{60}(P^*_{90}, P'_{60}), \quad \text{for any } P'_{60} \leq P_{60}.$$

Thus, any profitable deviation involves a deviation in P_{90} (and maybe also P_{60}) to a $P'_{90} < P^*_{90}$ and has positive sales of the 90 policy.[30]

[30] Note that if $P'_{90} = P^*_{90}$ and $P'_{60} < P^*_{60}$, then the deviation is unprofitable.

Consider a multi-policy deviation $P' = (P'_{90}, P'_{60})$ that is profitable and does not incur losses when any safe reaction is added. We will show that we necessarily have $\Pi_{90}(P'_{90}, P^*_{60}) > 0$ and $\Pi_{90}(P'_{90}, \widetilde{P}_{60}) \geq 0$ for all $\widetilde{P}_{60} \in [\underline{P}_{60}(P'_{90}), P^*_{60}]$. Thus, a single-policy deviation to P'_{90} would be profitable and no safe single-policy reaction in P_{60} could make it unprofitable. This will imply (see Remark 2) that in looking for Riley profitable deviations when the 60 policy has positive sales, we can restrict attention to single-policy deviations in P_{90}.

Observe that it must be the case that $\Pi_{90}(P'_{90}, \widetilde{P}_{60}) \geq 0$ for all $\widetilde{P}_{60} \in [\underline{P}_{60}(P'_{90}), P^*_{60}]$: First, if there are no sales of the 90 policy at prices $(P'_{90}, \underline{P}_{60}(P'_{90}))$, then (see Remark 1) $\underline{P}_{60}(P'_{90}) = \overline{AC}_{90} \geq P^*_{60}$ so the claim is trivially true. Suppose instead that there are sales of the 90 policy at prices $(P'_{90}, \underline{P}_{60}(P'_{90}))$. We show that $\Pi_{90}(P'_{90}, \underline{P}_{60}(P'_{90})) \geq 0$, from which the claim follows by Lemma 2. If, contrary to the claim we have $\Pi_{90}(P'_{90}, \underline{P}_{60}(P'_{90})) < 0$, then the safe single-policy reaction to $\underline{P}_{60}(P'_{90})$ makes the deviator incur losses if $\underline{P}_{60}(P'_{90}) < P'_{60}$, while if $\underline{P}_{60}(P'_{90}) \geq P'_{60}$ then the deviation is unprofitable (either the 60 policy has no sales at P' or it makes no profit).

The rest of the proof shows that $\Pi_{90}(P'_{90}, P^*_{60}) > 0$. First, if there are positive sales of only the 90 policy at prices P', then this is also true at prices (P'_{90}, P^*_{60}) and we have $\Pi_{90}(P'_{90}, P^*_{60}) > 0$ since it is a profitable deviation. So, we shall henceforth assume that there are positive sales of both policies at P'.

Next, observe that if $P'_{60} \leq \underline{P}_{60}(P'_{90})$, then $\Pi_{60}(P'_{90}, P'_{60}) \leq 0$, so since the deviation is profitable we would have $\Pi_{90}(P'_{90}, P'_{60}) > 0$, which implies $\Pi_{90}(P'_{90}, P^*_{60}) > 0$. So, henceforth we shall assume that $P'_{60} > \underline{P}_{60}(P'_{90})$. This implies that there are positive sales of the 60 policy at prices $(P'_{90}, \underline{P}_{60}(P'_{90}))$ since there are positive sales of the 60 policy at P'.

Now, if there are no sales of the 90 policy at prices $(P'_{90}, \underline{P}_{60}(P'_{90}))$ then $P'_{60} > \underline{P}_{60}(P'_{90}) = \overline{AC}_{60} \geq P^*_{60}$ (the equality follows by Remark 1 and the weak inequality by the fact that teh 60 policy has positive sales and breaks even at prices P^*), implying that $\Pi_{90}(P'_{90}, P^*_{60}) > 0$ since the deviation is profitable. So, henceforth we assume that there are positive sales of both policies at $(P'_{90}, \underline{P}_{60}(P'_{90}))$.

Since $\Pi_{90}(P'_{90}, \underline{P}_{60}(P'_{90})) \geq 0$, there are positive sales of both policies at $(P'_{90}, \underline{P}_{60}(P'_{90}))$, and (by Remark 1) we have $\underline{P}_{60}(P'_{90}) < \underline{P}_{60}(P^*_{90}) = P^*_{60}$, we have $\Pi_{90}(P'_{90}, P^*_{60}) > 0$. □

Lemma 6. *Among all price pairs (P_{90}, P_{60}) at which both policies break even and there are positive sales levels of the 60 policy, only the one(s) with the lowest sales of the 60 policy (and highest sales of the 90 policy) can be a Riley Equilibrium.*

Proof. Suppose there are two price pairs that are Riley Equilibria $P' = (P'_{90}, P'_{60})$ and $P'' = (P''_{90}, P''_{60})$ with both having positive and differing 60 shares. We can assume that $\Delta P' < \Delta P''$. Thus, there are positive sales of the 90 policy at P'. Also, $P'_{90} < P''_{90}$ and $P'_{60} < P''_{60}$. Consider the single-policy deviation to P'_{90} from the equilibrium (P''_{90}, P''_{60}). Since $\Pi_{90}(P'_{90}, P'_{60}) = 0$ (by Corollary 4) and there are positive sales of 90 at P', we have $\Pi_{90}(P'_{90}, P''_{60}) > 0$. Now, observe that the lowest safe 60 price given P'_{90} is P'_{60}; i.e., $\underline{P}_{60}(P'_{90}) = P'_{60}$, so $\Pi_{90}(P'_{90}, \underline{P}_{60}(P'_{90})) = 0$. By Lemma 2, there are no safe reactions that make the deviator incur a loss. This implies that (P''_{90}, P''_{60}) is not a Riley equilibrium, which is a contradiction. $\square$

Lemma 7. *A Riley equilibrium exists.*

Proof. Suppose otherwise. It is immediate that if at prices $(P_{90}, P_{60}) = (\underline{AC}_{90}, \overline{AC}_{60})$ all consumers buy policy 60, then by Lemma 5 this is a Riley equilibrium as there is no single-policy deviation in P_{90} that can earn a strictly positive profit, yielding a contradiction. Thus, we henceforth assume that there are positive sales of the 90 policy at $(P_{90}, P_{60}) = (\underline{AC}_{90}, \overline{AC}_{60})$.

Given Lemma 6, we need to show that it cannot be that both the everyone-in-90 break-even configuration and the break-even price pair with the lowest ΔP among those with a positive 60 share both fail to be Riley Equilibrium. Let $P^{**} = (P^{**}_{90}, P^{**}_{60})$ be the break-even price pair with the lowest $\Delta P > \underline{\theta}$. Let $P^{*}_{90} = \underline{AC}_{90}$ be the break-even 90 price in the everyone-in-90 outcome. Note also that $P^{**}_{60} = \underline{P}_{60}(P^{**}_{90})$. We will suppose that neither is a Riley equilibrium.

If P^{**} is not a Riley Equilibrium, by Lemma 5 there is a profitable single-policy deviation $P'_{90} < P^{**}_{90}$ with $\Pi_{90}(P'_{90}, P^{**}_{60}) > 0$ that does not incur losses when facing the reaction of the lowest safe price $\underline{P}_{60}(P'_{90}) \leq P^{**}_{60}$. If $\Pi_{90}(P'_{90}, \underline{P}_{60}(P'_{90})) = 0$, then we have found a price pair $(P'_{90}, \underline{P}_{60}(P'_{90}))$ with $\underline{P}_{60}(P'_{90}) < P^{**}_{60}$ and a positive share of 60 (since the change from P^{**}_{60} to $\underline{P}_{60}(P'_{90})$ lowered Π_{90}) that breaks even on both policies. Note also that there must then be positive sales of the 90 policy at prices $(P'_{90}, \underline{P}_{60}(P'_{90}))$ since otherwise (by Remark 1) we have $\underline{P}_{60}(P'_{90}) = \overline{AC}_{90} \geq P^{**}_{60}$. This in turn (by Remark 1) implies that $\Delta P' = P'_{90} - \underline{P}_{60}(P'_{90}) < \Delta P^{**}$, contradicting P^{**} being the break-even price pair with the lowest ΔP among those with positive sales of the 60 policy. So if P^{**} is not a Riley Equilibrium, it must be the case that $\Pi_{90}(P'_{90}, \underline{P}_{60}(P'_{90})) > 0$.

Now, suppose that everyone-in-90 is not a Riley Equilibrium. Then by Corollary 5 there is a $\widehat{P}_{60}$ such that $\Pi_{60}(\underline{AC}_{90}, \widehat{P}_{60}) > 0$. But this means that $\underline{P}_{60}(\underline{AC}_{90}) < \widehat{P}_{60}$ and at prices $(\underline{AC}_{90}, \underline{P}_{60}(\underline{AC}_{90}))$ there are positive sales of the 60 policy. Moreover, there must also be positive sales of the 90 policy at prices $(\underline{AC}_{90}, \underline{P}_{60}(\underline{AC}_{90}))$: if not, by Remark 1 we would have $\underline{P}_{60}(\underline{AC}_{90}) = \overline{AC}_{60}$, but then there must be (recall the first paragraph of the proof) positive sales of the 90 policy at $(P_{90}, P_{60}) = (\underline{AC}_{90}, \overline{AC}_{60})$. Because there are positive sales of both policies at prices $(P_{90}, P_{60}) = (\underline{AC}_{90}, \underline{P}_{60}(\underline{AC}_{90}))$, we have $\Pi_{90}(\underline{AC}_{90}, \underline{P}_{60}(\underline{AC}_{90})) < 0$. Continuity of the function $\Pi_{90}(P_{90}, \underline{P}_{60}(P_{90}))$ in P_{90} implies that there is a $\widetilde{P}_{90} \in (\underline{AC}_{90}, P'_{90})$ at which $\Pi_{90}(\widetilde{P}_{90}, \underline{P}_{60}(\widetilde{P}_{90})) = 0$. Hence, both policies break even at price pair $(\widetilde{P}_{90}, \underline{P}_{60}(\widetilde{P}_{90}))$. Moreover, $\Delta\widetilde{P} < \Delta P' \leq \Delta P^{**}$ (by Remark 1), contradicting P^{**} being the break-even price pair with the lowest ΔP among those with positive sales of the 60 policy. $\square$

Definition 3. *A Riley equilibrium (P^{*}_{90}, P^{*}_{60}) at which there are only sales of the 90 policy is a **strict Riley equilibrium** if $P_{60} < AC_{60}(\underline{AC}_{90}, P_{60})$ for all $P_{60} \leq \underline{AC}_{90} - \underline{\theta}$.*

Remark 5. *Generically, any Riley equilibrium at which there are only sales of the 90 policy is a **strict Riley equilibrium.***

Lemma 8. *If there is a strict Riley Equilibrium at which all consumers buy the 90 policy, then there is no Riley Equilibrium with positive sales of the 60 policy.*

Proof. Let $P^* = (P^*_{90}, P^*_{60})$ be a strict Riley Equilibrium in which all consumers buy the 90 policy. Then $P^*_{60} \geq \underline{AC}_{90} - \underline{\theta}$ and by Corollary 4 we must have $P^*_{90} = \underline{AC}_{90}$. Suppose there is another Riley equilibrium $P^{**} = (P^{**}_{90}, P^{**}_{60})$ in which $\Delta P^{**} > \underline{\theta}$ (so there are positive sales of the 60 policy). If so, then we know that $P^{**}_{60} = AC_{60}(P^{**}_{90}, P^{**}_{60})$ by Corollary 4.

We first argue that it cannot be that everyone buys the 60 policy at both prices P^{**} and prices $(\underline{AC}_{90}, P^{**}_{60})$. If they did, then we must have $P^{**}_{60} = \overline{AC}_{90} = AC_{60}(\underline{AC}_{90}, P^{**}_{60})$. But then for small $\varepsilon > 0$ we would have positive sales of the 60 policy at prices $(\underline{AC}_{90}, P^{**}_{60}+\varepsilon)$ and $\Pi_{60}(\underline{AC}_{90}, P^{**}_{60}+\varepsilon) > 0$, which contradicts P^* being a Riley equilibrium by Lemma 3.

Next, we argue that we must have $P^{**}_{90} > \underline{AC}_{90}$ and $P^{**}_{60} > AC_{60}(\underline{AC}_{90}, P^{**}_{60})$. Suppose, first, that there are sales of the 90 policy at P^{**}. Then from Corollary 4 we must have $P^{**}_{90} = AC_{90}(P^{**}_{90}, P^{**}_{60}) > \underline{AC}_{90}$, which in turn implies that $P^{**}_{60} > AC_{60}(\underline{AC}_{90}, P^{**}_{60})$. If, instead, all consumers buy policy 60 at prices P^{**} but some consumers buy the 90 policy at prices $(\underline{AC}_{90}, P^{**}_{60})$, then it must be that $P^{**}_{90} > \underline{AC}_{90}$ and $P^{**}_{60} > AC_{60}(\underline{AC}_{90}, P^{**}_{60})$.

Finally, we argue that if P^* is a strict Riley equilibrium, and if $P^{**}_{90} > \underline{AC}_{90}$ and $P^{**}_{60} > AC_{60}(\underline{AC}_{90}, P^{**}_{60})$, then P^{**} cannot be a Riley equilibrium. Since $P^{**}_{60} > AC_{60}(\underline{AC}_{90}, P^{**}_{60})$ and P^* is a Riley equilibrium, Lemma 3 implies that there must be that there are no sales of the 60 policy at prices $(\underline{AC}_{90}, P^{**}_{60})$. But, if so, then for small $\varepsilon > 0$ the single-policy deviation to $P'_{90} = \underline{AC}_{90} + \varepsilon$ from P^{**} has $\Pi_{90}(\underline{AC}_{90} + \varepsilon, P^{**}_{60}) > 0$. Moreover, if P^* is a strict Riley equilibrium, there is no single-policy safe reaction in P_{60} to P'_{90} that earns nonnegative profit. Hence, by Remark 2 there is no safe reaction that renders the deviation to $P'_{90} = \underline{AC}_{90} + \varepsilon$ unprofitable. □

Remark 6. *Strictness is used is to insure that there are no safe reactions to P'_{90} at the end of the proof, the need for which arises because of the difference between all-in-90 being a Riley equilibrium (no strictly profitable single-policy deviation in P_{60}) and the zero profits that are allowed for safe single-policy reactions in P_{60}.*

8.1.1 Proof of Proposition

First, observe that any WE (P^w_{90}, P^w_{60}) must satisfy constraints (i) and (ii). Constraint (ii) is obvious. If (i) was not satisfied, a single policy deviation of $P^w_{90} - \varepsilon$ for small $\varepsilon > 0$ would be profitable if the 60 policy still remained, and would be even more profitable if it did not. We next consider some properties of the solution of problem (1). Note first that in any solution constraint (ii) must bind – otherwise we could do better lowering both prices by the same small amount $\varepsilon > 0$. This also implies that if (P^*_{90}, P^*_{60}) is a solution, then $\Pi_{60}(P^*_{90}, P^*_{60}) \geq 0$. Second, in any solution to problem (1), policy 90 must be purchased. Suppose otherwise, and let (P^*_{90}, P^*_{60}) be a solution to problem (1), where $P^*_{60} = \overline{AC}_{60}$ and $P^*_{90} \geq P^*_{60} + \overline{\theta}$. Clearly, it must then be that $\overline{AC}_{60} < \underline{AC}_{90} - \underline{\theta}$, for otherwise the (feasible) price pair $(\underline{AC}_{90}, \underline{AC}_{90} - \underline{\theta})$ in which everyone buys policy 90 would achieve a lower value in problem (1) (recall that we have assumed that $\overline{AC}_{60} \neq \underline{AC}_{90} - \underline{\theta}$). Next, observe that, from the same argument as in Proposition 4(i), there would be a $\widehat{P}_{90} = P^*_{90} - \varepsilon$ for small $\varepsilon > 0$ such that $\widehat{P}_{90} - P^*_{60} \in (\underline{\theta}, \overline{\theta})$ (both policies are purchased), $\Pi(\widehat{P}_{90}, P^*_{60}) > 0$, and $\Pi_{60}(\widehat{P}_{90}, P^*_{60}) > 0$. In this case, we can then find a $P'_{90} < \widehat{P}_{90}$ such that $\Pi(P'_{90}, P^*_{60}) = 0$ and $\Pi_{60}(P'_{90}, P^*_{60}) > 0$, and hence, $\Pi_{90}(P'_{90}, P^*_{60}) < 0$,[31] which implies that there exists a $\widehat{P}_{60} < P^*_{60}$ and a $\widehat{P}_{90}$ such that $\Pi(\widehat{P}_{90}, \widehat{P}_{60}) = 0$ and $\Pi_{90}(\widehat{P}_{90}, \widehat{P}_{60}) < 0$. Thus, we then have a contradiction to (P^*_{90}, P^*_{60}) being a solution to problem (1). Third, observe that any solution has $P^*_{60} < \overline{AC}_{60}$. To see this, observe that $(P_{90}, P_{60}) = (\overline{AC}_{60} + \overline{\theta}, \overline{AC}_{60})$ is feasible. But then the argument of the previous paragraph implies that we can find a feasible $(\widehat{P}_{90}, \widehat{P}_{60})$ with $\widehat{P}_{60} < \overline{AC}_{60}$. We next argue that any (P_{90}, P_{60}) that satisfies the constraints but is not a solution to this problem cannot be a WE. Let (P^*_{90}, P^*_{60}) be a solution (a solution clearly exists). By definition, $P^*_{60} < P_{60}$. If $(P^*_{90}, P^*_{60}) << (P_{90}, P_{60})$ then a deviation to $(P^*_{90} + \varepsilon, P^*_{60} + \varepsilon)$ for small $\varepsilon > 0$ is strictly profitable regardless of whether prices P_{90} and P_{60} are withdrawn. So suppose that $P^*_{90} \geq P_{90}$. Then, if the original P_{90} were still offered, a deviation to $(P^*_{90} + \varepsilon, P^*_{60} + \varepsilon)$ for small $\varepsilon > 0$ attracts any former purchasers of policy 60 and also the best purchasers of policy 90. So $\Pi_{90}(P^*_{90}, \widehat{P}_{60} + \varepsilon) < 0$ and therefore P_{90} will be dropped and the deviation will be strictly profitable; i.e., the deviation "Wilson-breaks" (P_{90}, P_{60}). Finally, we argue that any solution (P^*_{90}, P^*_{60}) to problem (1) is a WE. First, it is not possible to enter in just the 60 policy with $\widehat{P}_{60} < P^*_{60}$ and make a strictly positive profit: This is clearly true if everyone buys the 60 policy after the deviation since (as noted above) $P^*_{60} < \overline{AC}_{60}$. If some consumers still buy the 90 policy, then doing so results in the 90 policy being dropped because it earned nonpositive profit to start with (and had positive sales) and the deviation takes its best consumers, resulting in $\Pi_{90}(P^*_{90}, \widehat{P}_{60}) < 0$. But once the 90 policy is dropped, the deviation policy with price $\widehat{P}_{60} < P^*_{60}$ loses money, since $P^*_{60} < \overline{AC}_{60}$. Second, it is not possible to enter in just the 90 policy and make a strictly positive profit: This is clearly true if all consumers buy policy 90 at (P^*_{90}, P^*_{60}). On the other hand, if some buy policy 60 at (P^*_{90}, P^*_{60}), then it must be that $\overline{AC}_{90} - \underline{\theta} \geq P^*_{60}$, or equivalently, $\overline{AC}_{90} \geq P^*_{60} + \underline{\theta}$, for otherwise $(\overline{AC}_{90}, \overline{AC}_{90} - \underline{\theta})$ would be feasible and better than (P^*_{90}, P^*_{60}) in problem (1). Now, the deviation never gets rid of P^*_{60} since it takes the worst consumers from policy 60. Given this fact, if there were a $\widehat{P}_{90} < P^*_{90}$ such that $\Pi_{90}(\widehat{P}_{90}, P^*_{60}) > 0$, we would have $\Pi_{60}(\widehat{P}_{90}, P^*_{60}) \geq 0$, so $\Pi(\widehat{P}_{90}, P^*_{60}) > 0$. But then P^*_{60} cannot be the value of the solution to problem (1): by the same logic as above, there is a $P'_{90} < \widehat{P}_{90}$ such that $\Pi(P'_{90}, P^*_{60}) = 0$ and $\Pi_{90}(P'_{90}, P^*_{60}) < 0$, which leads to a contradiction.[32] Thus, any deviation that Wilson-breaks (P^*_{90}, P^*_{60}) must be a two-policy deviation $(\widehat{P}_{90}, \widehat{P}_{60})$ that attracts all consumers, so it must have $\widehat{P}_{90} < P^*_{90}$, $\widehat{P}_{60} < P^*_{60}$, and $\Pi(\widehat{P}_{90}, \widehat{P}_{60}) > 0$. The best such deviation must either have $\widehat{P}_{90} = P^*_{90} - \varepsilon$ or $\widehat{P}_{60} = P^*_{60} - \varepsilon$ and result

[31] A P'_{90} such that $\Pi(P'_{90}, P^*_{60}) = 0$ exists, since the fact that $P^*_{60} + \underline{\theta} = \overline{AC}_{60} + \underline{\theta} < \underline{AC}_{90}$ implies that $\Pi(P^*_{60} + \underline{\theta}, P^*_{60}) < 0$, and $\Pi(\cdot, \cdot)$ is continuous. Moreover, this implies that there are some purchases of policy 60 at (P'_{90}, P^*_{60}) since $P'_{90} > P^*_{60} + \underline{\theta}$, hence $\Pi_{60}(P'_{90}, P^*_{60}) > 0$.

[32] Once again, note that $\Pi(P^*_{60} + \underline{\theta}, P^*_{60}) < 0$, so such a P'_{90} exists and it has $P'_{90} > P^*_{60} + \underline{\theta}$, so some consumers still buy the 60 policy.

in profits of $\max\{\Pi(P_{90}^*, \widehat{P}_{60}), \Pi(\widehat{P}_{90}, P_{60}^*)\}$. We first argue that $\Pi(P_{90}^*, \widehat{P}_{60}) \leq 0$. Since $\widehat{P}_{60} < P_{60}^*$, we have $\Pi_{90}(P_{90}^*, \widehat{P}_{60}) \leq 0$ which implies that $\Pi(P_{90}^*, \widehat{P}_{60}) \leq \Pi_{60}(P_{90}^*, \widehat{P}_{60})$ – i.e., a single policy deviation is at least as profitable. But we have already seen that it cannot be profitable, so $\Pi(P_{90}^*, \widehat{P}_{60}) \leq 0$. We next argue that $\Pi(\widehat{P}_{90}, P_{60}^*) \leq 0$. If all buy policy 90 at $(\widehat{P}_{90}, P_{60}^*)$, then it is equivalent to a single policy deviation, which we have seen cannot be profitable. If some buy policy 60 at $(\widehat{P}_{90}, P_{60}^*)$, then $\Pi_{60}(\widehat{P}_{90}, P_{60}^*) > 0, \Pi(\widehat{P}_{90}, P_{60}^*) > 0$, and $\Pi_{90}(\widehat{P}_{90}, P_{60}^*) \leq 0$. Once again, we can then find a $P_{90}' < \widehat{P}_{90}$ such that $\Pi(P_{90}', P_{60}^*) = 0$ and $\Pi_{90}(P_{90}', P_{60}^*) < 0$, which leads to a contradiction.

8.1.2 Nash Equilibria

Note that, in general, if an active firm's best deviation from a set of price offers is strictly profitable, then there exists a srictly profitable deviation for an entrant: The entrant can offer the same deviating prices less ε, and make a post-deviation profit that is arbitrarily close to that of the active firm. (Note that a firm's best deviation can without loss of generality be taken to capture either all consumers or no consumers who purchase each policy.) Since the entrant's pre-deviation profit is zero, it must find the deviation at least as attractive as did the active firm. Hence, we can restrict attention to deviations by entrants.

8.1.3 Single-policy firms

A Nash equilibrium with single-policy firms is a pair of prices (P_{90}, P_{60}), and resulting allocation of consumers to the two contracts, such that no firm (including potential entrants) has a profitable deviation. We assume that each equilibrium price is offered by at least 2 firms; any NE price pair can be sustained in this way since this reduces the number of possible deviations (which must then involve price reductions).

Lemma 9. *Suppose that* (P_{90}^*, P_{60}^*) *is a NE. Then both policies must break even: i.e.,* $\Pi_{90}(P_{90}^*, P_{60}^*) = \Pi_{60}(P_{90}^*, P_{60}^*) = 0$.

Proof. Suppose policy k is profitable. A new firm could enter and offer price $P_k - \varepsilon$ for that contract and make a positive profit: This deviation attracts all consumers who are purchasing policy k, and earns a positive profit on them, and for small ε, attracts very few others. Taking $\varepsilon \to 0$ yields the result. □

Lemma 10. *Among all price pairs* (P_{90}, P_{60}) *at which both policies break even and* $\Delta P > \underline{\theta}$ *(there are positive sales of the 60 policy), only the ones with the lowest sales of the 60 policy can be NE.*[33]

Proof. Let $(P_{90}^{'}, P_{60}')$ and (P_{90}'', P_{60}'') lead both policies to break even, have $\min\{\Delta P', \Delta P''\} > \underline{\theta}$, and have differing purchase patterns (note that the 90 policy might not be purchased at the price pair with

[33] The statement is intended to allow for the fact that we can have multiple price pairs that have all consumers buying the 60 policy and both policies breaking even.

the larger ΔP). Without loss of generality, let $\Delta P' < \Delta P''$ (note that these must be different in the two equilibria if the purchasing is different). Also, note that there must be positive sales of the 90 policy with price difference $\Delta P'$. Then, since $AC_k(\Delta P)$ is increasing in ΔP for $k = 90, 60$, we must have $P'_{60} < P''_{60}$. So, at price pair $(P''_{60} + \Delta P', P''_{60})$ we get the same sales levels as at prices (P'_{90}, P'_{60}), and the 90 policy makes a strictly positive profit since $P''_{60} + \Delta P' > P'_{90}$. Hence, (P''_{90}, P''_{60}) cannot be a NE: an entrant would profit by offering the 90 policy at price $\widehat{P}_{90} \equiv P''_{60} + \Delta P' < P''_{90}$. □

The above result tells us that we can narrow down the possible equilibria to two: (i) everyone in the 90 policy, (ii) the lowest $\Delta P \in (\underline{\theta}, \overline{\theta})$ at which $AC_{90}(\Delta P) - AC_{60}(\Delta P) = \Delta P$, or if no such interior ΔP exists, everyone in the 60 policy. The former is an equilibrium only if there is no profitable deviation to a $P_{60} < \underline{AC}_{90} - \underline{\theta}$. Moreover, observe the following:

Lemma 11. *If there is a NE at which all consumers buy the 90 policy, then this is the unique NE.*

Proof. Let $P^*_{90} = \underline{AC}_{90}$ be the equilibrium price in this Nash equilibrium and note that we must have $P^*_{60} \geq P^*_{90} - \underline{\theta}$. Suppose as well that there is another NE $(P^{**}_{90}, P^{**}_{60})$ in which policy 60 is purchased by a positive measure of consumers. Note that we then must have $P^{**}_{90} > P^{**}_{60} + \underline{\theta}$ and $P^{**}_{60} = AC_{60}(\Delta P^{**})$, where $\Delta P^{**} \equiv P^{**}_{90} - P^{**}_{60} \geq \underline{\theta}$. Thus, $P^{**}_{60} > \underline{MC}_{60}$. We first argue that we must have $P^{**}_{60} > P^*_{90} - \underline{\theta}$: if not, then at prices (P^*_{90}, P^*_{60}), an entrant could attract a positive measure of consumers close to type $\underline{\theta}$ by offering $\widehat{P}_{60} \equiv P^*_{90} - \underline{\theta} - \varepsilon$ for small $\varepsilon > 0$ and would make a strictly positive profit since $\widehat{P}_{60} > \underline{MC}_{60}$, contradicting (P^*_{90}, P^*_{60}) being a NE. But if $P^{**}_{60} > P^*_{90} - \underline{\theta}$, then $(P^{**}_{90}, P^{**}_{60})$ cannot be a NE, as an entrant could attract all consumers by offering $\widehat{P}_{90} \equiv P^{**}_{60} + \underline{\theta} > P^*_{90} = \underline{AC}_{90}$ (note that $\widehat{P}_{90} < P^{**}_{90}$). □

These results suggest the following procedure: First, see if all buying the 90 policy is a NE by seeing if there is profitable policy 60 deviation. If it is, then that is the unique NE. If it is not an equilibrium, identify all break-even price pairs in which the 60 policy is purchased and then see if the one with the lowest ΔP is a NE. In doing this, the following result tells us that we can restrict attention to deviations in P_{90}:

Lemma 12. *Suppose* (P_{90}, P_{60}) *have* $\Pi_{90}(P_{90}, P_{60}) = \Pi_{60}(P_{90}, P_{60}) = 0$ *and* $\Delta P \in (\underline{\theta}, \overline{\theta}]$. *Then no single-policy deviation by an entrant in* P_{60} *is profitable.*

Proof. The deviation $\widehat{P}_{60}$ must be below P_{60} to attract any consumers (when at least one other firm offers P_{60}). But any such deviation loses \$ on the consumers who were choosing policy 60, and by Assumption 1 loses \$ on any consumers it attracts away from policy 90. □

8.1.4 Multi-policy firms (with multi-policy deviations)[34]

A Nash equilibrium with multi-policy firms and multi-policy deviations is a pair of prices (P_{90}, P_{60}), and resulting allocation of consumers to the two contracts, such that no firm (including potential entrants) has a profitable multi-contract deviation. Clearly, lemmas 1-3 still hold and our conclusions above still hold (all previous deviations are still feasible). The new result is as follows:

Proposition 4. *If (P_{90}^*, P_{60}^*) is a NE with multi-policy firms and mutli-policy deviations, then*

(i) *some consumers must be buying the 90 policy;*

(ii) *if all consumers are buying the 90 policy, it is a NE if there is no profitable single policy deviation by an entrant in P_{60};*

(iii) *if some consumers are buying the 60 policy, it is a NE iff $\Pi(P_{90}^*, P_{60}^*) = 0 = \sup_{\widehat{P}_{90}<P_{90}^*} \Pi(\widehat{P}_{90}, P_{60}^*)$, that is, if there is no profitable multi-policy deviation by an entrant that reduces P_{90} and lowers P_{60} by ε to capture all consumers.*

Proof. For part (i), suppose all consumers were purchasing the 60 policy. Then $P_{60} = \overline{AC}_{60}$ and $P_{90} \geq P_{60} + \overline{\theta}$. Now consider a deviation to $(\widehat{P}_{90}(\varepsilon), P_{60})$ where

$$\widehat{P}_{90}(\varepsilon) = P_{60} + \overline{\theta} - \varepsilon.$$

We will show that for small $\varepsilon > 0$, aggregate profits are strictly positive, which implies that there is a δ such that $(\widehat{P}_{90}(\varepsilon), P_{60} - \delta)$ is a profitable deviation. Specifically,

$$\psi(\varepsilon) \equiv \Pi(\widehat{P}_{90}(\varepsilon), P_{60}) = \int_{\overline{\theta}-\varepsilon}^{\overline{\theta}} [\widehat{P}_{90}(\varepsilon) - C_{90}(\theta)] f(\theta) d\theta + \int_{\underline{\theta}}^{\overline{\theta}-\varepsilon} [P_{60} - C_{60}(\theta)] f(\theta) d\theta.$$

Now

$$\psi'(\varepsilon) = [\widehat{P}_{90}(\varepsilon) - C_{90}(\overline{\theta} - \varepsilon)] f(\overline{\theta} - \varepsilon) - [P_{60} - C_{60}(\overline{\theta} - \varepsilon)] f(\overline{\theta} - \varepsilon) - [F(\overline{\theta}) - F(\overline{\theta} - \varepsilon)],$$

so

$$\begin{aligned} \psi'(0) &= [\widehat{P}_{90}(\varepsilon) - C_{90}(\overline{\theta})] f(\overline{\theta}) - [P_{60} - C_{60}(\overline{\theta})] f(\overline{\theta}) \\ &= f(\overline{\theta}) \{\overline{\theta} - [C_{90}(\overline{\theta}) - C_{60}(\overline{\theta})]\} > 0, \end{aligned}$$

where the inequality follows by Assumption 3. For part (ii), let (P_{90}^*, P_{60}^*) be the candidate equilibrium in which only the 90 policy is purchased. Consider any entrant multi-policy deviation $(\widehat{P}_{90}, \widehat{P}_{60}) \leq (P_{90}^*, P_{60}^*)$. To be profitable, some consumers must buy policy 60 in the deviation, so $\widehat{P}_{60} < P_{60}^*$ and $\Delta \widehat{P} < \Delta P$. But then the most profitable such deviation must have $\widehat{P}_{90}$ equal

[34] Note that when firms can offer multiple policies but can deviate in only one policy at a time there is no change from the analysis of the previous subsection since we only need to consider entrant's deviations and the profitability of these is unaffected by whether firm's offer more than one policy.

to or arbitrarily close to P^*_{90}. (Otherwise, both $\widehat{P}_{90}$ and $\widehat{P}_{60}$ could be raised by a small and equal amount (since $\Delta\widehat{P} < \Delta P$). But, since the reduction in P_{60} makes the 90 policy at price P^*_{90} unprofitable, this deviation is less profitable than the profitability of an entrant's single-policy deviation to $\widehat{P}_{60}$. For part (iii), we already know that an entrant's single-policy deviation in P_{60} is unprofitable. A single-policy deviation by an entrant in P_{90} to $\widehat{P}_{90}$, since it raises the profitability of the 60 policy, is less profitable than a multi-policy by the entrant to $\widehat{P}_{90}$ which also lowers P_{60} by ε to capture the entire market. This deviation has profits equal to $\sup_{\widehat{P}_{90}<P^*_{90}} \Pi(\widehat{P}_{90}, P^*_{60})$. □

Appendix A: Cost Model Setup and Estimation

This appendix describes the details of the cost model, which is summarized at a high-level in section 4. The output of this model, F_{kjt}, is a family-plan-time specific distribution of predicted out-of-pocket expenditures for the upcoming year. This distribution is an important input into the choice model, where it enters as a family's predictions of its out-of-pocket expenses at the time of plan choice, for each plan option. We predict this distribution in a sophisticated manner that incorporates (i) past diagnostic information (ICD-9 codes) (ii) the Johns Hopkins ACG predictive medical software package (iii) a novel non-parametric model linking modeled health risk to total medical expenditures using observed cost data and (iv) a detailed division of medical claims and health plan characteristics to precisely map total medical expenditures to out-of-pocket expenses.[35] The unique level of precision we gain from the cost model leads to more credible estimates of the choice parameters of primary interest (e.g. switching costs).

In order to most precisely predict expenses, we categorize the universe of total medical claims into four mutually exclusive and exhaustive subdivisions of claims using the claims data. These categories are (i) hospital and physician (ii) pharmacy (iii) mental health and (iv) physician office visit. We divide claims into these four specific categories so that we can accurately characterize the plan-specific mappings from total claims to out-of-pocket expenditures since each of these categories maps to out-of-pocket expenditures in a different manner. We denote this four dimensional vector of claims C_{it} and any given element of that vector $C_{d,it}$ where $d \in D$ represents one of the four categories and i denotes an individual (employee or dependent). After describing how we predict this vector of claims for a given individual, we return to the question of how we determine out-of-pocket expenditures in plan j given C_{it}.

Denote an individual's past year of medical diagnoses and payments by ξ_{it} and the demographics age and sex by ζ_{it}. We use the ACG software mapping, denoted A, to map these characteristics into a predicted mean level of health expenditures for the upcoming year, denoted θ:

[35] Features (iii) and (iv) are methodological advances. We are aware of only one previous study that incorporates diagnostic information in cost prediction for the purposes of studying plan choice ([?]) in a structural setup. [?] use this type of framework in ongoing work.

$$A : \xi \times \zeta \rightarrow \theta$$

In addition to forecasting a mean level of total expenditures, the software has an application that predicts future mean pharmacy expenditures. This mapping is analogous to A and outputs a prediction λ for future pharmacy expenses.

We use the predictions θ and λ to categorize similar groups of individuals across each of four claims categories in vector in C_{it}. Then for each group of individuals in each claims category, we use the actual ex post realized claims for that group to estimate the ex ante distribution for each individual under the assumption that this distribution is identical for all individuals within the cell. Individuals are categorized into cells based on different metrics for each of the four elements of C:

$$\begin{aligned} \text{Pharmacy:} &\quad \lambda_{it} \\ \text{Hospital / Physician (Non-OV):} &\quad \theta_{it} \\ \text{Physician Office Visit:} &\quad \theta_{it} \\ \text{Mental Health:} &\quad C_{MH,i,t-1} \end{aligned}$$

For pharmacy claims, individuals are grouped into cells based on the predicted future mean pharmacy claims measure output by the ACG software, λ_{it}. For the categories of hospital / physician (non office visit) and physician office visit claims individuals are grouped based on their mean predicted total future health expenses, θ_{it}. Finally, for mental health claims, individuals are grouped into categories based on their mental health claims from the previous year, $C_{MH,i,t-1}$ since (i) mental health claims are very persistent over time in the data and (ii) mental health claims are uncorrelated with other health expenditures in the data. For each category we group individuals into a number of cells between 8 and 10, taking into account the tradeoff between cell size and precision. The minimum number of individuals in any cell is 73 while almost all cells have over 500 members. Thus since there are four categories of claims, each individual can belong to one of approximately 10^4 or 10,000 combination of cells.

Denote an arbitrary cell within a given category d by z. Denote the population in a given category-cell combination (d, z) by I_{dz}. Denote the empirical distribution of ex-post claims in this category for this population $\hat{G_{I_{dz}}}(\cdot)$. Then we assume that each individual in this cell has a distribution equal to a continuous fit of $\hat{G_{I_{dz}}}(\cdot)$, which we denote G_{dz}:

$$\varpi : \hat{G_{I_{dz}}}(\cdot) \rightarrow G_{dz}$$

We model this distribution continuously in order to easily incorporate correlations across d. Otherwise, it would be appropriate to use $G_{I_{dz}}$ as the distribution for each cell.

The above process generates a distribution of claims for each d and z but does not model correlation over D. It is important to model correlation over claim categories because it is likely that someone with a bad expenditure shock in one category (e.g. hospital) will have high expenses in another area (e.g. pharmacy). We model correlation at the individual level by combining marginal distributions G_{idt} $\forall$ d with empirical data on the rank correlations between pairs (d, d').[36] Here, G_{idt} is the distribution G_{dz} where $i \in I_{dz}$ at time t. Since correlations are modeled across d we pick the metric θ to group people into cells for the basis of determining correlations (we use the same cells that we use to determine group people for hospital and physician office visit claims). Denote these cells based on θ by z_θ. Then for each cell z_θ denote the empirical rank correlation between claims of type d and type d' by $\rho_{z_\theta}(d, d')$. Then, for a given individual i we determine the joint distribution of claims across D for year t, denoted $H_{it}(\cdot)$, by combining i's marginal distributions for all d at t using $\rho_{z_\theta}(d, d')$:

$$\Psi : G_{iDt} \times \rho_{z_{\theta_{it}}}(D, D') \rightarrow H_{it}$$

Here, G_{iDt} refers to the set of marginal distributions $G_{idt} \forall d \in D$ and $\rho_{z_{\theta_{it}}}(D, D')$ is the set of all pairwise correlations $\rho_{z_{\theta_{it}}}(d, d') \forall (d, d'^2$. In estimation we perform Ψ by using a Gaussian copula to combine the marginal distribution with the rank correlations, a process which we describe momentarily.

The final part of the cost model maps the joint distribution H_{it} of the vector of total claims C over the four categories into a distribution of out of pocket expenditures for each plan. For each of the three plan options we construct a mapping from the vector of claims C to out of pocket expenditures OOP_j:

$$\Omega_j : C \rightarrow OOP_j$$

This mapping takes a given draw of claims from H_{it} and converts it into the out of pocket expenditures an individual would have for those claims in plan j. This mapping accounts for plan-specific features such as the deductible, co-insurance, co-payments, and out of pocket maximums listed in table 2. I test the mapping Ω_j on the actual realizations of the claims vector C to verify that our mapping comes close to reconstructing the true mapping. Our mapping is necessarily simpler and omits things like emergency room co-payments and out of network claims. We constructed our mapping with and without these omitted categories to insure they did not lead to an incremental increase in precision. We find that our categorization of claims into the four categories in C passed through our mapping Ω_j closely approximates the true mapping from claims to out-of-pocket expenses. Further, we find that it is important to model all four categories described above: removing any of the four makes Ω_j less accurate. Figure 6 shows the results of one validation exercise for PPO_{250}. The top panel reveals that actual employee out-of-pocket spending amounts are quite close to those predicted by Ω_j, indicating the precision of this mapping. The bottom panel repeats this mapping when we add out of network expenses as a fifth category. The output in this case is similar to that in the top panel without this category, implying that including this category would not markedly change the cost model results.

[36] It is important to use rank correlations here to properly combine these marginal distribution into a joint distribution. Linear correlation would not translate empirical correlations to this joint distribution appropriately.

Once we have a draw of OOP_{ijt} for each i (claim draw from H_{it} passed through Ω_j) we map individual out of pocket expenditures into family out of pocket expenditures. For families with less than two members this involves adding up all the within family OOP_{ijt}. For families with more than three members there are family level restrictions on deductible paid and out-of-pocket maximums that we adjust for. Define a family k as a collection of individuals i_k and the set of families as K. Then for a given family out-of-pocket expenditures are generated:

$$\Gamma_j : OOP_{i_k,jt} \rightarrow OOP_{kjt}$$

To create the final object of interest, the family-plan-time specific distribution of out of pocket expenditures $F_{kjt}(\cdot)$, we pass the claims distributions H_{it} through Ω_j and combine families through Γ_j. $F_{kjt}(\cdot)$ is then used as an input into the choice model that represents each family's information set over future medical expenses at the time of plan choice. Eventually, we also use H_{it} to calculate total plan cost when we analyze counterfactual plan pricing based on the average cost of enrollees. Figure 7 outlines the primary components of the cost model pictorially to provide a high-level overview and to ease exposition.

We note that the decision to do the cost model by grouping individuals into cells, rather then by specifying a more continuous form, has costs and benefits. The cost is that all individuals within a given cell for a given type of claims are treated identically. The benefit is that our method produces local cost estimates for each individual that are not impacted by the combination of functional form and the health risk of medically different individuals. Also, the method we use allows for flexible modeling across claims categories. Finally, we note that we map the empirical distribution of claims to a continuous representation because this is convenient for building in correlations in the next step. The continuous distributions we generate very closely fit the actual empirical distribution of claims across these four categories.

Cost Model Identification and Estimation. The cost model is identified based on the two assumptions of (i) no moral hazard / selection based on private information and (ii) that individuals within the same cells for claims d have the same ex ante distribution of total claims in that category. Once these assumptions are made, the model uses the detailed medical data, the Johns Hopkins predictive algorithm, and the plan-specific mappings for out of pocket expenditures to generate the the final output $F_{kjt}(\cdot)$. These assumptions, and corresponding robustness analyses, are discussed at more length in the main text.

Once we group individuals into cells for each of the four claims categories, there are two statistical components to estimation. First, we need to generate the continuous marginal distribution of claims for each cell z in claim category d, G_{dz}. To do this, we fit the empirical distribution of claims $G_{I_{dz}}$ to a Weibull distribution with a mass of values at 0. We use the Weibull distribution instead of the lognormal distribution, which is traditionally used to model medical expenditures, because we find that the lognormal distribution overpredicts large claims in the data while the Weibull does not. For each d and z the claims greater than zero are estimated with a maximum likelihood fit to the Weibull distribution:

$$\max_{(\alpha_{dz},\beta_{dz})} \Pi_{i\in I_{dz}} \frac{\beta_{dz}}{\alpha_{dz}} (\frac{c_{id}}{\alpha_{dz}})^{\beta_{dz}-1} e^{-(\frac{c_{id}}{\alpha_{dz}})^{\beta_{dz}}}$$

Here, $\hat{\alpha_{dz}}$ and $\hat{\beta_{dz}}$ are the shape and scale parameters that characterize the Weibull distribution. Denoting this distribution $W(\hat{\alpha_{dz}}, \hat{\beta_{dz}})$ the estimated distribution $\hat{G_{dz}}$ is formed by combining this with the estimated mass at zero claims, which is the empirical likelihood:

$$\hat{G_{dz}}(c) = \begin{cases} G_{I_{dz}}(0) & \text{if } c = 0 \\ G_{I_{dz}}(0) + \frac{W(\hat{\alpha_{dz}},\hat{\beta_{dz}})(c)}{1-G_{I_{dz}}(0)} & \text{if } c > 0 \end{cases}$$

Again, we use the notation $\hat{G_{iDt}}$ to represent the set of marginal distributions for i over the categories d: the distribution for each d depends on the cell z an individual i is in at t. We combine the distributions $\hat{G_{iDt}}$ for a given i and t into the joint distribution H_{it} using a Gaussian copula method for the mapping Ψ. Intuitively, this amounts to assuming a parametric form for correlation across $\hat{G_{iDt}}$ equivalent to that from a standard normal distribution with correlations equal to empirical rank correlations $\rho_{z_{\theta_{it}}}(D, D')$ described in the previous section. Let $\Phi^i_{1|2|3|4}$ denote the standard multivariate normal distribution with pairwise correlations $\rho_{z_{\theta_{it}}}(D, D')$ for all pairings of the four claims categories D. Then an individual's joint distribution of non-zero claims is:

$$\hat{H_{i,t}}(\cdot) = \Phi_{1|2|3|4}(\Phi_1^{-1}(\hat{G_{id_1t}}), \Phi_2^{-1}(\hat{G_{id_2t}}), \Phi_3^{-1}(\hat{G_{id_3t}}), \Phi_4^{-1}(\hat{G_{id_4t}}))))$$

Above, Φ_d is the standard marginal normal distribution for each d. $\hat{H}_{i,t}$ is the joint distribution of claims across the four claims categories for each individual in each time period. After this is estimated, we determine our final object of interest $F_{kjt}(\cdot)$ by simulating K multivariate draws from $\hat{H}_{i,t}$ for each i and t, and passing these values through the plan-specific total claims to out of pocket mapping Ω_j and the individual to family out of pocket mapping Γ_j. The simulated $F_{kjt}(\cdot)$ for each k, j, and t is then used as an input into estimation of the choice model.

Table 8 presents summary results from the cost model estimation for the final choice model sample, including population statistics on the ACG index θ, the Weibull distribution parameters $\hat{\alpha_{dz}}$ and $\hat{\beta_{dz}}$ for each category d, as well as the across category rank correlations $\rho_{z_{\theta_{it}}}(D, D')$. These are the fundamentals inputs used to generate F_{kjt}, as described above, and lead to very accurate characterizations of the overall total cost and out-of-pocket cost distributions (validation exercises which are not presented here).

Final Sample				
Cost Model Output				
	Overall	PPO_{250}	PPO_{500}	PPO_{1200}
Individual Mean (Median)				
Unscaled ACG Predictor				
Mean		1.42	0.74	0.72
Median		0.83	0.37	0.37
Pharmacy: Model Output				
Zero Claim Pr.	0.35 (0.37)	0.31 (0.18)	0.40 (0.37)	0.42 (0.37)
Weibull α	1182 (307)	1490 (462)	718 (307)	596 (307)
Weibull β	0.77 (0.77)	0.77 (0.77)	0.77 (0.77)	0.77 (0.77)
Mental Health				
Zero Claim Pr.	0.88 (0.96)	0.87 (0.96)	0.90 (0.96)	0.90 (0.96)
Weibull α	1422 (1295)	1447 (1295)	1374 (1295)	1398 (1295)
Weibull β	0.98 (0.97)	0.99 (0.97)	0.98 (0.97)	0.98 (0.97)
Hospital / Physician				
Zero Claim Pr.	0.23 (0.23)	0.21 (0.23)	0.26 (0.23)	0.26 (0.23)
Weibull α	2214 (1599)	2523 (1599)	1717 (1599)	1652 (1599)
Weibull β	0.58 (0.55)	0.59 (0.55)	0.55 (0.55)	0.55 (0.55)
($> \$40,000$) Claim Pr.	0.02 (0.01)	0.02 (0.01)	0.01 (0.01)	0.01 (0.01)
Physician OV				
Zero Claim Pr.	0.29 (0.20)	0.26 (0.20)	0.33 (0.46)	0.34 (0.46)
Weibull α	605 (553)	653 (553)	517 (410)	529 (410)
Weibull β	1.15 (1.14)	1.15 (1.14)	1.15 (1.14)	1.14 (1.14)
Correlations				
Rank Correlation Hospital-Pharm.	0.28 (0.34)	0.26 (0.32)	0.31 (0.34)	0.32 (0.34)
Rank Correlation Hospital-OV	0.73 (0.74)	0.72 (0.74)	0.74 (0.74)	0.74 (0.74)
Rank Correlation Pharm.-OV	0.35 (0.41)	0.33 (0.37)	0.38 (0.41)	0.39 (0.41)

Table 8: This table describes the output of the cost model in terms of the means and medians of individual level parameters, classified by the plan actually chosen. These parameters are aggregated for these groups but have more micro-level groupings, which are the primary inputs into our cost projections in the choice model. Weibull α, Weibull β, and Zero Claim Probability correspond to the cell-specific predicted total individual-level health expenses as described in more detail in Appendix A.

Appendix B: Choice Model Estimation Algorithm Details

This appendix describes the details of the choice model estimation algorithm. The corresponding section in the text provided a high-level overview of this algorithm and outlined the estimation assumptions we make regarding choice model fundamentals and their links to observable data.

We estimate the choice model using a random coefficients probit simulated maximum likelihood approach similar to that summarized in [**?**]. The simulated maximum likelihood estimation approach has the minimum variance for a consistent and asymptotically normal estimator, while not being too computationally burdensome in our framework. Since we use panel data, the likelihood function at the family level is computed for a *sequence* of choices from t_0 to t_2, since switching costs imply that the likelihood of a choice made in the current period depends on the choice made in the previous period. The maximum likelihood estimator selects the parameter values that maximize the similarity between actual choices and choices simulated with the parameters.

First, the estimator simulates Q draws from the distribution of health expenditures output from the cost model, F_{kjt}, for each family, plan, and time period. These draws are used to compute plan expected utility conditional on all other preference parameters. It then simulates S draws for each family from the distributions of the random coefficients γ_k and δ_k, as well as from the distribution of the preference shocks $\epsilon_j(Y_k)$. We define the set of parameters θ as the full set of ex ante model parameters (before the S draws are taken):

$$\theta \equiv (\mu, \beta, \sigma^2_\gamma, \mu_\delta(Y_k), \sigma_\delta(Y_k), \alpha, \mu_{\epsilon_J}(Y_k), \sigma_{\epsilon_J}(Y_k), \eta_0, \eta_1, \eta_2).$$

We denote θ_{sk} one draw derived from these parameters for each family, including the parameters constant across draws:

$$\theta_{sk} \equiv (\gamma_k, \delta_k, \alpha, \epsilon_{JT}, \eta_0, \eta_1, \eta_2)$$

Denote θ_{Sk} the set of all S simulated draws for family k. For each θ_{sk} the estimator then uses all Q health draws to compute family-plan-time-specific expected utilities U_{skjt} following the choice model outlined in earlier in section 4. Given these expected utilities for each θ_{sk}, we simulate the probability of choosing plan j in each period using a smoothed accept-reject function with the form:

$$Pr_{skt}(j = j^*) = \frac{\left(\frac{\frac{1}{-U_{skj^*t}}(\cdot)}{\Sigma_J \frac{1}{-U_{skjt}}(\cdot)}\right)^\tau}{\Sigma_{\hat{j}}\left(\frac{\frac{1}{-U_{sk\hat{j}t}}(\cdot)}{\Sigma_J \frac{1}{-U_{skjt}}(\cdot)}\right)^\tau}$$

This smoothed accept-reject methodology follows that outlined in [**?**] with some slight modifications to account for the expected utility specification. In theory, conditional on θ_{sk}, we would want to pick the j that maximizes U_{kjt} for each family, and then average over S to get final choice probabilities. However, doing this leads to a likelihood function with flat regions, because for small changes in the estimated parameters θ, the discrete choice made does not change. The smoothing function above mimics this process for CARA utility functions: as the smoothing parameter τ becomes large the smoothed Accept-Reject simulator becomes almost identical to the true Accept-Reject simulator just described, where the actual utility-maximizing option is chosen with probability one. By choosing τ to be large, an individual will always choose j^* when $\frac{1}{-U_{kj^*t}} > \frac{1}{-U_{kjt}} \forall j \neq j^*$. The smoothing function is modified from the logit smoothing function in [?] for two reasons (i) CARA utilities are negative, so the choice should correspond to the utility with the lowest absolute value and (ii) the logit form requires exponentiating the expected utility, which in our case is already the sum of exponential functions (from CARA). This double exponentiating leads to computational issues that our specification overcomes, without any true content change since both models approach the true Accept-Reject function.

Denote any sequence of three choices made as j^3 and the set of such sequences as J^3. In the limit as τ grows large the probability of a given j^3 will either approach 1 or 0 for a given simulated draw s and family k. This is because for a given draw the sequence (j_1, j_2, j_3) will either be the sequential utility maximizing sequence or not. This implicitly includes the appropriate level of switching costs by conditioning on previous choices within the sequential utility calculation. For example, under θ_{sk} a choice in period two will be made by a family k only if it is optimal conditional on θ_{sk}, other preference factors, and the switching costs implies by the period one choice. For all S simulation draws we compute the optimal sequence of choices for k with the smoothed Accept-Reject simulator, denoted j^3_{sk}. For any set of parameter values θ_{Sk} the probability that the model predicts j^3 will be chosen by k is:

$$\hat{P_k^{j^3}}(\theta, F_{kjt}, X_k^A, X_k^B, H_k, Y_k) = \Sigma_{s \in S} \mathbf{1}[j^3 = j^3_{sk}]$$

Let $\hat{P_k^{j^3}}(\theta)$ be shorthand notation for $\hat{P_k^{j^3}}(\theta, F_{kjt}, X_k^A, X_k^B, H_k, Y_k)$. Conditional on these probabilities for each k, the simulated log-likelihood value for parameters θ is:

$$SLL(\theta) = \Sigma_{k \in K} \Sigma_{j^3 \in J^3} d_{kj^3} \ ln \hat{P_k^{j^3}}$$

Here d_{kj^3} is an indicator function equal to one if the actual sequence of decisions made by family k was j^3. Then the maximum simulated likelihood estimator (MSLE) is the value of θ in the parameter space Θ that maximizes $SLL(\theta)$. In the results presented in the text, we choose $Q = 100$, $S = 50$, and $\tau = 4$, all values large enough such that the estimated parameters vary little in response to changes.

Appendix C: Moral Hazard Robustness analysis

In the text we discuss a robustness specification that investigates the cost model assumption of no moral hazard. To do this we necessarily make some substantial simplifying assumptions: for a full structural treatment of moral hazard in health insurance utilization see, e.g., [?], [?] or [?]. We implement the moral hazard robustness check by adjusting the output of the cost model to reflect lower total utilization in the less comprehensive plans (and vice-versa). The intent is to show that even when including price elasticities that are quite large relative to those found in the literature, the model output for switching costs and risk preferences does not change substantially. While the specification addresses moral hazard, it also sheds light on whether our estimates are sensitive to consumers having a reasonable level of private information above and beyond the detailed medical data we observe.

We operationalize this test in the following steps. First, we find the implied total spending changes across plans in the population for a price elasticity of -1.3, well higher than the range of -0.1 to -0.4 that represents most of the literature (for further discussion, see [?]). To do this we perform a back of the envelope calculation for the arc-elasticity of demand with respect to price, following the prior work of [?] and [?]. For this calculation, we use the average share of out-of-pocket spending for each plan as the price, and total medical expenditures as the quantity. For the three plans we study, the empirical shares of out-of-pocket spending are 15.5%, 20.9%, and 23.4% going from most to least comprehensive. We use these prices together with the average total spending of $13,331 in PPO_{-1} at t_{-1} as the basis for this calculation. The formula for the arc-elasticity is:

$$Elasticity = \frac{(q_2 - q_1)/(q_2 + q_1)}{(p_2 - p_1)/(p_2 + p_1)}$$

We use the conservative elasticity of -1.3 and solve for the corresponding total cost changes this price response implies (here, p_j are the empirical shares of out-of-pocket spending for each plan and q_j is implied total spending in plan j). Solving for q_2 as a percentage of q_1 implies an, approximately, 25% reduction in total spending moving from PPO_{250} to PPO_{500}, a 33% reduction in total spending moving from PPO_{250} to $PP0_{1200}$, and a 10% reduction in total spending moving from PPO_{500} to PPO_{1200}. We then apply these reductions (or increases when moving into more comprehensive plans) and adjust the output of the cost model F_{kjt} according to the potential plan being chosen and the previous plan the family actually enrolled and incurred costs in. These new 'moral hazard adjusted' F_{kjt} then are input into the choice model, which is otherwise specified and estimated as in the text. The results from the analysis are presented in column 3 of table 7 and suggest that the initial assumption of no moral hazard does not markedly change our estimates of switching costs or risk preferences.[37]

It is important to point out that this exercise does not explicitly model the value of additional health spending, which the literature does through a non-linear budget constraint model where the family trades off the value of extra spending with the price of medical care (see [?], [?] or [?]). Moral hazard here is captured purely by differences in spending. For a family that chooses PPO_{500} or PPO_{1200} in the data and is considering switching to PPO_{250}, our moral hazard cost wedge serves as an upper bound for the expected utility difference between these two plans and the more comprehensive plan, since their actual choice implies they value the increase in medical services less than the corresponding overall utility gain from a different financial lottery. The reverse is not true: for consumers considering switching to a less comprehensive plan our out-of-pocket cost wedge may not be an upper bound value

[37] In prior work, [?] presented evidence that the combined effect of moral hazard and selection on private information is not large in our setting. This suggests that the elasticity choice here of -1.3 represents more moral hazard than actually exists in our environment, implying this is a conservative approach.

of utility differences between two plans. However, we view this as a conservative approach in this case, since the high elasticity we've chosen together with the differences in the marginal prices of care between the plans (and the resulting implications for value of medical care foregone) make it unlikely that the value wedge between prospective plans is larger in reality than in this specification.