Dynamic Competition and Price Regulation When Consumers Have Inertia: Evidence from Medicare Part D

Ariel Bain

October 18, 2018

Abstract

When consumer choices have inertia, firms have incentives to use dynamic pricing by first reducing the price to build a large market share, and then by increasing prices. This strategy may reduce consumer welfare through increases in the prices for incumbents, while also changing the patterns of entry and exit in the market. Although the presence of inertia in health care markets has been well established, little is known about the welfare implications of dynamic pricing in these markets. In order to assess these implications, in this paper I develop and estimate a dynamic model of supply and demand for Medicare Part D prescription drug insurance plans, where multi-product firms consider consumer inertia in their decisions about premiums, offerings of new plans, and exit of plans. Using the model and the estimated parameters, I conduct counterfactual exercises where I explore the welfare effects of a policy that limits dynamic pricing by imposing fixed markups. I find that this policy, given the actual consumer inertia present in this market, would improve consumer welfare by 3.1%, through a reduction in premiums that is partially off-set by a reduction of entry into the market. When the same policy is implemented in a counterfactual scenario without inertia, it has a larger positive effect, increasing consumer welfare by 9.4% relative to the benchmark. This difference indicates that policies that limit dynamic pricing can be more effective in improving consumer welfare in markets with lower levels of consumer inertia, where they are less likely to harm market entry.

Keywords: Health Insurance, Inertia, Dynamic Pricing, Welfare
JEL Classification: L11, I11, D47, I18, H51

1 Introduction

When consumers have inertia, firms have incentives to use dynamic pricing, by first setting low prices to build a large market share, and increasing prices later. This strategy, called "investing-then-harvesting," may reduce consumer welfare by increasing the prices of incumbents and changing the patterns of entry and exit in the market. Dynamic pricing may offer opportunities for new products to enter with relatively low prices to attract consumers, with the hope of being able to increase prices over time. In the theoretical literature, the net effects of these strategies on prices and entry and exit behavior depend on the size of consumer inertia. Understanding the nature of these tradeoffs in equilibrium is thus an empirical question.

The presence of significant inertia in health care markets has been well established. However, we still know very little about the effects of dynamic pricing strategies on prices, and entry and exit in these markets. Medicare Part D, the prescription drug insurance program for Medicare beneficiaries, is an excellent setting to study these effects. Part D has large impacts on both the health and economic welfare of more than 37 million individuals, accounting for more than $76 billions in government spending. It is also a prominent example of the use of market mechanisms for the provision of public programs, which is based on the idea that the combination of supply competition and consumer choice will maximize consumer welfare while keeping costs low. However, these potential benefits would be limited if there are frictions in the market, such as inertia, inducing dynamic pricing. From its inception, press articles on Part D have suggested the presence of behavior consistent with dynamic pricing strategies. Krasner (2006) reports that Humana, one of the largest insurer firms, introduced plans with relatively low premiums in 2006 and then largely increased the premiums of these plans over time. More recently, Ericson (2014) and Ho et al. (2015) provide evidence consistent with dynamic pricing strategies in Medicare Part D.

In this paper, I evaluate the welfare effects of a regulation that limits dynamic pricing strategies in the market for Medicare Part D plans, where multi-product insurers make decisions on the entry and exit of plans based on these dynamic pricing considerations. I make three main contributions. First, using plan-level data, I provide new reduced-form evidence that is consistent with dynamic pricing strategies. I show that firms with higher market shares set higher premiums, and new plans generally enter the market with lower premiums and increase them as they gain market share. Second, this paper is, to my knowledge, the first to estimate a dynamic model of the demand and supply of plans in Part D that incorporates endogenous entry and exit. This model allows me to capture the first order aspects of this setting, namely consumer inertia, heterogeneity of consumer preferences, dynamic pricing by firms, and endogenous market entry and exit. In particular, I provide the first estimates of some of the key parameters of dynamic competition in this market, such as entry costs and scrap values of plans. Finally, I use this model and the estimated parameters to evaluate the effects of a policy that limits dynamic pricing on prices, market entry, and consumer welfare. This is the first paper to provide an evaluation of the full effects of this kind of policy in a setting with consumer inertia, accounting for the dynamic pricing incentives that it creates, and the endogenous entry and exit responses to these incentives.

On the demand side, I estimate separate demand functions for regular and low-income benefi-

ciaries, since they may have differences in preferences and face different subsidies. Additionally, I incorporate more consumer heterogeneity by using random coefficients for premiums (*a la* Berry et al. (1995)). I recognize and estimate inertia by assuming a utility cost for choosing a plan that is different from the plan that a consumer had in the last period.[1] To address the concerns about the endogeneity of premiums, I follow an instrumental variables strategy using a set of instruments based on the location of the products in the product space (BLP-type) and on common cost shocks for plans (Hausman-type). My estimates show that inertia is important for both regular and low-income beneficiaries (with a cost of switching plans of around $700) and that consumers are relatively elastic to premiums (with elasticities between -5.5 and -4). The demand side provides two main contributions to the model. First, it determines the transition between states in the supply model. Second, it allows me to compute consumer welfare, which is necessary to evaluate the counterfactual policies.

On the supply side, I model multi-product firms that compete by making dynamic decisions about offering new plans, the exit of plans, and the pricing of plan premiums. The model is based on Maskin and Tirole (1988) and Ericson and Pakes (1995). However, in my model these three decisions are all simultaneous dynamic decisions,[2] which are based on a vector of state variables that reflect the existence of dynamic competition, particularly the "investing-then-harvesting" dynamics, and private information on marginal costs, exit scrap value, and entry sunk costs.

The estimation of the supply side follows Bajari et al. (2007), and proceeds in two stages. In the first stage, I recover the policy functions on pricing, entry and exit that describe the optimal policies for plans at each point of the state space.[3] Besides controlling for observed characteristics and market fixed effects, a crucial aspect of the identification of these policy functions is the inclusion of the unobserved quality estimated from the demand-side model, since correlation between premiums and lagged market shares can reflect persistent unobserved quality. Additionally, plans with higher unobserved quality are plausibly less likely to exit, and entry can depend on the average unobserved quality of plans available. The policy functions suggest that a 10 percentage point increase in lagged market share increases plan premiums by about $9, and reduces the probability of exit by 2.3%. In the second stage, I impose the complete dynamic model and I recover from it the fundamental parameters of the distribution of marginal costs, entry costs and scrap values for plans. The identification of the marginal cost comes from pricing decisions of plans under different market structures in the different states of the game, while entry costs and exit scrap values are identified by the entry and exit decisions of plans and the duration of plans in the market. My estimate of the mean of the marginal costs is $1,079, which implies an average markup rate of 13%. The estimates of the entry cost and scrap values are $2.5 million and $1, respectively. The marginal costs estimated using the supply-side structural model are in line with previous estimates from

[1] Under the presence of inertia, consumers have incentives to incorporate in their decisions their expectations about future prices and characteristics of plans. However, and based on the literature on consumer choice inconsistencies in Medicare Part D, in this paper I assume myopic consumers that completely discount the future.

[2] In the tradition of Maskin and Tirole (1988) and Ericson and Pakes (1995), competition in prices or quantities is modeled as a static decision and investment is the dynamic decision.

[3] By doing this, this methodology avoids the need for computing the equilibrium of the model as part of the estimation process. This idea is based on the conditional choice probabilities used in Hotz and Miller (1993).

claims data or static demand models. To the best of my knowledge, this is the first attempt to estimate the parameters of entry costs or scrap values.

I use the estimations from the demand and supply model to conduct two counterfactual exercises, where I explore the welfare effects of a policy that limits dynamic pricing by imposing a fixed markup. Unlike in other settings, a fixed-markup policy is feasible to implement in Medicare Part D, because the costs of plans can be forecasted with software that predicts prescription drug expenditures based on enrollee data. The main innovation of my counterfactual exercises it that they account for supply responses to the policy not only in terms of pricing, but also in entry and exit behavior, which affects the overall effects of the policy on premiums and consumer welfare. Because the introduction of a fixed-markup policy induces a completely different game among plans, I cannot directly use the estimated policy functions from the model, but instead have to solve for the new Markov Perfect Nash Equilibrium of the game in each counterfactual scenario. The computational burden of these new games increases geometrically with the number of plans in the market, so I simplify the setting by assuming there are only five firms in each market.[4]

The first counterfactual scenario introduces a fixed-markup policy while keeping the inertia of consumers unchanged. The comparison of this scenario to the benchmark allows me to estimate the consumer welfare gains of implementing the fixed-markup policy, given the actual consumer inertia present in the market for Medicare Part D plans. I find that this policy improves consumer welfare by 3.1%, through a 7% reduction in the premiums paid by enrollees, which also reduces government expenditures in direct subsidies. This net effect is the result of a reduction in premiums that is partially off-set by a reduction of entry into the market. The second counterfactual scenario introduces a fixed-markup policy in a market without consumer inertia. The goal of this exercise is to evaluate how the effects of this policy change according to the degree of inertia in consumer choice present in the market. When implemented in a counterfactual scenario without inertia, this policy has a substantially larger effect on consumer welfare, increasing it by 9.4% relative to the benchmark. The main explanation for this larger welfare effect is that, in the absence of consumer inertia, the policy increases the endogenous entry of new plans instead of harming it, reinforcing the direct effect on premiums.

The findings of this paper underscore the importance of taking into account effects on market entry when evaluating the implementation of policies that limit dynamic pricing behavior. In the presence of substantial consumer inertia, eliminating the possibility of strategic dynamic pricing that exploits this inertia reduces the incentives for potential entrants to enter the market, because it does not allow them to set low entry prices with the expectation of raising them once they have built their market share. This negative effect on market entry partially offsets the effects of this policy in reducing prices. On the other hand, my results also show that without consumer inertia, a policy that limits dynamic pricing has larger positive effects on consumer welfare, through both a reduction in prices and an increase in the incentives for market entry.

This paper contributes to the strands of literature that address the presence of inertia in health care markets, dynamic competition when consumers have inertia, and the increasingly relevant

[4] While this assumption is clearly restrictive, the four largest firms in the market accumulate a market share of 60%.

design of privately-provided public health insurance programs.

This paper is most directly related to the literature that studies the effects of consumer inertia on dynamic competition. The theoretical literature on this topic starts with Klemperer (1987), who shows that in the presence of consumer inertia, firms can engage in dynamic pricing by initially setting low prices in order to increase the market share ("investing") and then increasing prices to "harvest" on inertial consumers. Beggs and Klemperer (1992) show that, in a model with an infinite horizon model, horizontal differentiation and infinite switching costs, the harvesting incentive dominates the investment incentive and that, therefore, the presence of inertia softens competition and increases prices. This conclusion is also reached by most of the theoretical models about competition with switching costs (Farrell and Shapiro, 1988; Klemperer, 1995; Farrell and Klemperer, 2007). However, some recent papers show that it is possible, for certain levels of inertia, for the investment incentive to dominate harvesting and, thus, dynamic pricing can reduce prices (Dubé et al., 2009; Arie and Grieco, 2014). Additionally, inertia may have important effects on entry and exit patterns. Firms can strategically set low entry prices to build a market share, and then increase prices over time (Farrell and Klemperer, 2007). Therefore, this literature highlights the important roles played by the degree of consumer inertia and entry and exit behavior on the net effects of consumer inertia on dynamic competition, underscoring that the effects of consumer inertia on prices and consumer welfare is ultimately an empirical question.

A recent body of empirical work assesses the effects of inertia on strategic pricing behavior in Medicare Part D. Ericson (2014) finds that older plans have approximately 10% higher premiums than comparable new plans, consistent with an "investing-then-harvesting" behavior. Ho et al. (2015) study, in a hedonic reduced-form supply model, the impact of inertia caused by consumer inattention on premiums. My paper is most closely related to the papers by Miller (2014) and Wu (2016), which incorporate dynamic pricing responses to inertia, considering the entry of plans as exogenous to this process.

My main contribution is to account for the endogeneity of the entry of new plans into the market when considering the impacts of dynamic pricing in a context of consumer inertia. I develop and estimate a dynamic model of demand and supply for Medicare Part D plans, and I discuss the welfare implications of dynamic pricing. Based on this model, I simulate counterfactual scenarios that limit dynamic pricing that allow me to decompose the welfare gains of policies limiting the pricing behavior between the effects on premiums and on entry of new plans.

This paper also contributes to a large literature highlighting frictions and different sources of inertia in health care markets, particularly in Medicare. The papers in this literature show that consumers have inconsistencies and learning in choosing plans in Medicare Part D (Abaluck and Gruber, 2011, 2013; Ketcham et al., 2012; Heiss et al., 2013)[5]; document the presence and size of switching costs in Medicare (Nosal, 2012; Miller and Yeo, 2012); analyze its interactions with adverse selection (Handel, 2013; Polyakova, 2016); and discuss the relative importance of inattention and switching costs as sources of consumer inertia (Ho et al., 2015).

Finally, this paper contributes to a growing literature on the supply-side regulations of Medicare

[5] A related literature shows choice inconsistencies around the donut hole, with consumers having hyperbolic discount or myopia (Einav et al., 2015; Dalton et al., 2015).

Part D. Chorniy et al. (2016) analyze the effect of mergers on premiums, while Lavetti and Simon (2016) and Abaluck and Scott Morton (2016) explore firm responses in terms of the design of plans to consumer inertia. Decarolis (2015) discusses the supply-side regulations of the Low Income Subsidy (LIS) in Medicare Part D, showing that, given the design of the auction, insurers game with the subsidy, causing distortions in premiums and raising costs. Additionally, Decarolis et al. (2015) analyze different pricing regulations, finding that the current subsidy mechanism achieves a level of total welfare close to that obtained under an optimal voucher scheme, but is far from the social planner's first best solution.[6]

The paper is organized as follows. Section 2 describes the institutional setting of Medicare Part D. Section 3 presents the data, descriptive statistics and reduced-form evidence of the investing-then-harvesting dynamics. Section 4 introduces the dynamic model of supply and demand for plans. Section 5 describes the empirical approach and identification of the model. Section 6 presents the results of the estimates of the model. Section 7 discusses counterfactual exercises where dynamic pricing is not allowed. Finally, Section 8 concludes.

2 Industry Background

Medicare Part D is a subsidized and privately provided health insurance program for prescription drugs for the elderly and disabled in the United States that was introduced in the year 2006. The design of Medicare Part D has been of particular interest to economists since its beginning (Duggan et al., 2008) for at least three reasons. First, the program has the potential to significantly affect both the health and the economic well-being of the more than 37 million individuals currently enrolled. Second, Part D has substantially increased government spending on health care. In 2014, the government expended more than $76 billion in Medicare Part D. Third, Part D represents an attempt to use market mechanisms in the provision of large-scale public programs. The idea behind the use of private provision is that the combination of supply-side competition and consumer choice would maximize consumer welfare while keeping costs low. Medicare Part D is a privileged example of the private provision of government-subsidized programs, which has been increasingly popular in the last two decades. However, the potential benefits of consumer choice and provider competition could not be achieved if there are important frictions in the market, such us the investing-then-harvesting dynamics studied in this paper.

There are 34 markets[7] in Medicare Part D, where consumers can choose among different plans. Beneficiaries can enroll in two types of private insurance plans: a) Prescription Drug Plans (PDP), which provide coverage exclusively for prescription drug costs and accounted for two thirds of the total number of enrollees by 2013; or b) Medicare Advantage plans (MA-PD), which insure all Medicare covered services (hospital care, physician services and prescription drugs). In this paper, I focus only on the PDP segment, which offers a cleaner setting since the prescription drugs coverage is not tied with other benefits.

[6] Other studies have focused on other aspects or segments of privately provided health insurance markets (Starc, 2014; Ericson and Starc, 2015; Shepard, 2016; Tebaldi, 2016).

[7] Each market is a geographical region. Hereafter, I refer indistinctively to them as *markets* or *regions*.

Also, Part D has two types of beneficiaries that receive different subsidies: regular, and low-income subsidy (LIS) beneficiaries. Regular beneficiaries have the ability of actively choose a plan, and they receive a subsidy on the premium that is determined by CMS through a process explained in the next paragraph. For the vast majority of seniors enrolling is financially favorable (Heiss et al., 2006). Additionally, the design of the program tries to reduce the scope of adverse selection;[8] newly eligible seniors who delay enrolling are required to pay a higher price when they do join. Beneficiaries with low income are eligible for the Low-Income Subsidy (LIS), and they are automatically assigned to a set of plans that only offer the standard benefits and whose premiums are lower than those paid by regular enrollees. Beneficiaries can subsequently change their assignment, becoming what I denominate hereafter *LIS choosers*. LIS beneficiaries receive a higher subsidy from CMS compared to regular enrollees.

In order to determine the amount of the subsidies, the regulator (the Center for Medicaid and Medicare Services, CMS) administers a complex annual mechanism. First, all insurers submit bids for each plan they want to offer, which should reflect how much revenue the insurer "needs" in order to cover an average risk beneficiary in a standard benefit (SDB) plan. CMS calculates a weighted average of all the bids, using the previous enrollment shares as weights, which is called *CMS average bid* and is used as a basis to determine the subsidy. For regular enrollees with the average risk and a plan with standard benefits, the government subsidy for beneficiaries is set at 74.5% of the CMS average bid, and the other 25.5% is set as the premium of the plan. On top of this, CMS adjusts the per capita subsidy by the risk score of each enrollee, and adjusts the premiums for any additional benefits that the plan may offer compared to the standard benefit plan designed by CMS. Thus, the premium payed by the regular enrollee is 25.5% of the *CMS average bid*, plus adjustments for additional benefits. For LIS enrollees, the subsidy is determined using a weighted average of the same bids, but using the previous share of LIS beneficiaries for each plan as weights. Each year, the exact amount of the subsidy for LIS enrollees (*LIS benchmark*) is determined on the basis of fiscal considerations by CMS and its value has an average of 370 dollars in my sample years.

Insurer firms can offer plans with the SDB or an actuarially equivalent one.[9] In addition, firms can also offer "enhanced plans" with additional coverage beyond these levels. The characteristics of these plans include the deductible, extra coverage in the "donut hole," the number and types of drugs in each tier of the plan's drug formulary, and the network of pharmacies.

Multi-product firms compete offering on average less than three plans per year-market.[10] Each

[8]Despite this intention, some adverse selection is still present in the program. Polyakova (2016) explores interactions among adverse selection, switching costs and regulation in Medicare Part D. She finds that switching costs are large and have quantitatively important implications for the sorting of individuals among contracts. In particular, switching costs help sustain an adversely-selected equilibrium with large differences in risks between more and less generous contracts.

[9]For example, in 2014 the structure of the SDB was the following. The SDB plan starts with a deductible of \$310. After the deductible is met, the beneficiary pays 25% of covered costs up to total prescription costs meeting the *initial coverage limit* (\$2,850). Once this amount is reached, the coverage gap ("donut hole") begins. In this coverage gap the beneficiary pays 100% of their prescription costs up to the catastrophic region that is set at \$6455. This level of the catastrophic region implies an aggregate level of out-of-pocket expenditure of \$4,550, which is the sum of the deductible (\$310), the expenditure in the *initial coverage* ((\$2850-\$310)*25%) and the expenditure in the coverage gap ((\$6455-\$2850)*100%). In the catastrophic region consumers pay the maximum between 5% of the price of the drug or a fix amount (\$2.55 for generic drugs and \$6.35 for the others).

[10]CMS uses the term *parent organization* to refer to an insurer firm. In the rest of the paper, I refer to parent

year, plans can have one of the following three statuses: *renewed*, *terminated*, or *consolidated*, and also *new plans* can be introduced. *Renewed plans* are plans that continue from the previous year and its enrollees are carried-over. The characteristics of these plans could change over time. *Terminated plans* are plans that exit the market and stop being offered in the year. When a plan is terminated, its enrollees have to actively choose a new plan. This is the concept of exit of plans that is used in this paper. *New plans* are introduced to the market for the first time and they have no enrollees from the previous calendar year. This is the concept of entry of plans used in this paper. Finally, *consolidated plans* combine two or more plans and the enrollees of both plans are carried over to the consolidated plan. Consolidation of plans is done by merging firms or by firms that do not participate in any merger or acquisition. There has been a wave of mergers in Medicare Part D during the period of my sample (Chorniy et al., 2016), but in this paper I consider these consolidations as exogenous and I do not model them in the supply model.

3 Data and Reduced-Form Evidence

The database used in this paper comes from a variety of statistics released annually by the CMS at the aggregate level of plan by market and year. In particular, I use information on enrollment (for regular and LIS beneficiaries), bids, premiums and plan financial characteristics, at the plan level for years 2007 to 2012, for the 34 Medicare regions (markets). In addition, the crosswalk files allow me to link plans across years and determine their status. I also use CMS publicly available (at a cost) data on pharmacies and drugs included in the formulary for each prescription drug plan. Finally, I use demographic population characteristics from the Integrated Public Use Microdata Series (IPUMS) that I merge to my dataset at the region-year level.

Table 1 provides summary statistics for the Medicare Part D plans that I use in my sample. Using data for 2009 for reference, about 45.5 million individuals were Medicare beneficiaries and therefore eligible to purchase prescription drug coverage via Medicare Part D. 16.5 million of these Medicare beneficiaries bought the PDP plans that are in my sample. Out of these individuals, 8.6 million are regular (non-LIS) beneficiaries, and 7.9 are LIS beneficiaries. Beneficiaries can also obtain coverage via other sources: 9.7 million bought Medicare Advantage plans (prescription drug plans bundled with Medicare Advantage), 6.5 million had employer sponsored coverage, and 6.4 million had other forms of coverage. Finally, 6.4 million individuals did not have any form of creditable coverage. In the demand model subsection I explain how I use these data to construct the market shares and outside options of my demand models. Figures 3 and 4 show the level and the variation of the outside option for regular and LIS-eligible enrollees, across each of the 34 geographic markets defined in Medicare Part D. On average, the outside option represents about 65% of regular beneficiaries and around 75% of LIS beneficiaries.

There are on average 1,450 stand-alone prescription drug plans per year during the span of my sample. Each market has an average of 50 plans, which are run by insurer firms (denominated *parent organizations* by CMS). Table 1 shows that there are, on average, about 55 firms per year,

organizations (or insurer firms) as *firms*, and to each product (prescription drug plan) offered by a firm as a *plan*.

and each firm offers an average of 2 to 3 plans per market each year. Table 2 shows that these averages hide a vivid dynamic of entry and exit of plans in each market, with plans entering and exiting the market almost every year in each region. This table also shows that consolidations among plans are the main reason for the reduction in the number of plans per year, while there is also exit of plans every year. Panel (b) in Figure 1 presents the overall dynamics of entry and exit, with Panel (a) presenting the effect of this dynamic on the total number of plans. These graphs show a small but steady reduction in the number of plans during the period, with a larger number of consolidations in year 2010.[11] Figure 2 shows that in the sample period, new entries and exits generally respond to firms' decisions to introduce or withdraw one plan in a market at a time. Panel (a) in this graph presents the histogram of the number of plans that firms introduce per year and region for the whole sample. This graph shows that about 95% of the time, firms decide either not to introduce a new plan, or to introduce only one new plan in a market. Panel (b) shows that firms decide each period either to maintain all their plans in the market, or withdraw one plan. Based on this evidence, in the empirical analysis I consider that firms make a discrete choice regarding the entry or exit of each plan, which for simplification I refer to as plan entry and exit decisions.

There is a large heterogeneity in market shares attained by single plans and firms both within and across markets. Figure 3, Panel (a) presents this heterogeneity in plan market shares for regular beneficiaries in year 2009. While many plans only obtain a market share close to zero, some other obtain about 20% of the market. This picture is even more clear when we consider the market shares by firm, as shown in Panel (b) of Figure 3. While some firms have market shares of about 25% of the market, others have market shares that are very close to zero. Figure 4 shows that the previous remarks are qualitatively also true for LIS beneficiaries.

Table 1 shows that the average premium for standard benefit plans has increases and decreases over time, around an average of $390 dollars during the period of study. This behavior contrasts with the behavior of the unweighted average annual premium paid by regular enrollees for all plans, which increased quite substantially, going from $442 in 2007 to $648 in 2012. The average of the *CMS average bid* is $1010 and the *CMS base consumer premium* is on average $361, which represents 36% of the average bid. On the other hand, the evolution of the LIS subsidy is similar to the evolution of the premiums for plans with standard benefits because they have increased more slowly during the period, going from $341 in 2007 to $390 in 2012.

The difference between the evolution of standard benefits plans and the average premium is driven by the composition of plans available in each period. This fact is reflected in the increased dispersion in plan premiums, and in particular in a higher number of very expensive plans. Panel (a) in Figure 5 shows considerable variation of annual premiums by plan and market for year 2007, and Panel (b) presents considerably more variation for year 2012. Table 3 shows the effects of different characteristics on the annual premiums of plans for regular beneficiaries. In general, plans with extra coverage in the donut hole, more top drugs in the formulary, and more drugs in the

[11] In year 2011 CMS began publishing guidance to encourage insurers to consolidate low enrollment plans. The idea of the regulator is that plans should have a "meaningful difference" with plans sufficiently differentiated in characteristics from existing plans by the same insurer. These regulations can be found in the website of CMS in the following link `https://goo.gl/0nb1rA`

more accessible tiers (tiers 1 and 2) of the formulary, have higher premiums, while plans with higher deductibles have lower annual premiums.

Additionally, part of the explanation for the dispersion of premiums are the differences in market power exercised by insurers in different markets. Panel (a) in Figure 6 documents a clear downward sloping relationship between the level of plan premiums and the number of competing firms in a market, which is consistent with the variation in market power across regions.

Finally, the variation in premiums is also related to the investing-then-harvesting strategy. Panel (b) in Figure 6 shows the positive unconditional correlation between the annual premium and the lagged market share of the firm, for incumbent plans. In other words, plans offered by firms with a higher market share in the last period exhibit higher prices than other plans. Obviously, these unconditional correlations do not control for other possible factors affecting annual premium, most notably market power or characteristics of the plans (observed and unobserved quality). Table 3 projects the premiums on the observed characteristics of the plans and the lagged market share of the firm. The premium has a statistically significant correlation with the lagged share of the firm in the specifications without (Column 1) and with (Column 2) plan fixed effects.

Additionally, the prices of new plans present a behavior that is consistent with these investing-then-harvesting strategies. Columns 3 and 4 in Table 3 show that new plans generally have a lower premium than the incumbent plans in the specifications with and without plan fixed effects. Figure 7 shows the pattern of prices for new plans in the years that follow the year of entry. The graph shows the estimated value of a dummy variable indicating the number of years since entry for the plan in a regression that includes plan characteristics and plan and year fixed effects. The pattern exhibited in this graph is consistent with the idea that after entry, new plans start systematically increasing the premiums once they gain market share. However, the increase in premiums can also be due to selection, since we only observe in this graph the plans that survive in the market. My theoretical model with endogenize entry, exit and pricing in order to analyze the dynamic incentives faced by these plans.

4 Model

My theoretical model represents dynamic competition in the market for Medicare Part D PDP plans. The main focus of the model is to capture the "investing-then-harvesting" dynamics in which multi-product insurers are involved in this market. On the demand side, the model represents consumer choices among available plans. On the supply side, I model the decisions that each period these forward-looking multi-product insurers make, about which new plans to offer and which plans to withdraw from the market, as well as about premiums for incumbent and new plans.

The model builds on Maskin and Tirole (1988) and Ericson and Pakes (1995). However, in the typical model in this tradition incumbents make two decisions at each period: a dynamic decision on how much to invest, which affects the state variable (e.g. capacity or quality), and a static decision about prices or quantities determined in the static competition with other firms. In my model, the main policy variables are premiums and entry and exit of plans, and their determination

affects the revenue and profits for the current period as well as the evolution of the market share for the next period through the demand equation.

To focus on decisions about dynamic pricing, and the endogenous offering of new plans and exit of plans that they generate, I make three simplifying assumptions: i) non-price characteristics of the plans change exogenously, ii) consumers have inertia but they are myopic (having a discount factor equal to zero), and iii) insurers set premiums and decide on entry and exit for each plan without gaming on the design of the LIS. I discuss these assumptions in more depth below, where they are formally presented in the model.

I simplify the state space by summarizing it using four state variables. The state vector for plan j, in market m and time t ($\vec{S}_{jmt}$) is composed by the market share of the firm in the last period (s_{jmt-1}), the number of firms in the market in the last period (N_{mt-1}), the average number of plans by firm in the market the last period (D_{mt-1}), and the size of the market (M_{mt}), which I assume evolves exogenously and at a fixed rate. I assume that firms know the demand function and use it to form expectations about transitions between state variables. Therefore, the state variables evolve according to pricing decisions and entry and exit of plans via the estimated demand system.

The timing of the model proceeds as follows. First, plans are endowed with a set of characteristics. Then, simultaneously, both incumbent and potentially entrant plans make two decisions each. Incumbent plans receive a draw from the scrap value distribution and make two decisions: a) whether to stay in the market or exit, and b) which price to set if they stay. On the other hand, potential entrants (endowed with a set of characteristics) also make two decisions: a) whether or not to enter the market, and b) which price to set if they enter. After these decisions are made, consumers choose plans given plan characteristics and premiums, and therefore state variables evolve.

4.1 Demand Side

My demand model captures the demand for plans for different types of consumers. The demand model is based on Berry (1994) and Berry et al. (1995) but it incorporates dynamic elements of consumer choice (Gowrisankaran and Rysman, 2012; Cullen and Shcherbakov, 2010; Nosal, 2012). The model assumes that each consumer faces a choice set of plans and chooses the plan that maximizes her utility. I assume that consumers have linear utility functions and, therefore, I describe the plans by their characteristics and their premiums[12] Given the existence of inertia in consumer decisions, I incorporate in the model a cost that consumers have to pay if they decide to change to a plan different to the one they were enrolled in last period. Additionally, I use random coefficients (as in Berry et al. (1995)) on the price coefficient to incorporate heterogeneity in the consumer side. Following this literature, I recover the structural parameters for demand using only information about each plan's market share and characteristics.

Because of the institutional setting of Medicare Part D, regular and LIS consumers have different

[12]The linear form of the utility function may appear in contradiction with the usual assumption that consumers have risk-averse preferences over expected health risk. However, it is possible to think about the coefficients of these financial characteristics (other than premiums) in the linear utility function as reduced-form parameters that capture the revealed valuation of different financial characteristics of the plans.

demand functions and, therefore, I estimate their demand functions separately. I follow the strategy of Decarolis et al. (2015) to determine the outside options for each of these segments of the market. Regarding regular enrollees, they choose their plans and face the total amount of the premium and pay the full cost-sharing structure of Medicare Part D plans through deductibles, co-insurance and co-pays. For this segment, the market is composed by all the non-LIS Medicare beneficiaries who choose to enroll in a stand-alone prescription drug plan (PDP), enroll in a Medicare Advantage prescription drug plan (MA-PD), or do not have any Part D coverage.[13] Therefore, I consider the choice of not enrolling in any part D plan or enrolling in a Medicare Advantage prescription drug plan (MA-PD) as the outside option.

There are two main complications for the estimation of demand for LIS enrollees, which are related to the random assignment and the characteristics that enrollees face. First, although LIS enrollees are allowed to choose the plan they want, they are randomly assigned to eligible plans by the government as soon as they become Medicare Part D beneficiaries. The problem is that the data do not allow me distinguish among enrollees who are in LIS-eligible plans due to random assignment or by choice. I make the assumption that all LIS enrollees in LIS-eligible plans are there by random assignment. Therefore, I estimate the parameters of the preferences for LIS enrollees using the choices of LIS choosers that are enrolled in plans not eligible for random assignment. In this sense, I define the outside option for LIS enrollees as staying in a randomly assigned plan in Part D PDP. Second, LIS beneficiaries face different characteristics of plans, since their cost sharing is covered in large part by the government. The premium for LIS beneficiaries is the difference between the insurer's bid and the region-level LIS subsidy (LIPSA), which is higher than the subsidy for regular enrollees. Additionally, LIS population does not face a deductible, gap in coverage, or copayments above certain thresholds because the government covers these costs for them.

The presence of switching costs and other sources of inertia (remarkably inattention) is well known in the literature of Medicare Part D (Polyakova, 2016; Ho et al., 2015) and in other health markets (Handel, 2013). In this model of demand, I incorporate switching costs and other sources of inertia as a cost that individuals pay if they switch plans, and which they do not have to pay if they remain enrolled in the same plan. A similar strategy for the estimation of switching costs has been used in previous literature (Nosal, 2012; Fleitas, 2016). Under this strategy, I remain agnostic about the sources of inertia. These costs can be related to both a financial burden for switching plans, or a psychological cost, related to other behavioral factors.

It is also well documented in the literature that consumers have problems and inconsistency in the choice of Medicare plans (Abaluck and Gruber, 2011; Ketcham et al., 2012; Heiss et al., 2013; Einav et al., 2015; Dalton et al., 2015). In order to reflect these inconsistencies, on top of the presence of consumer inertia, I assume that consumers are myopic. Formally, I assume that individuals consider only the current-period utility to make decisions, and do not take into account expectations over the future value of payoff-relevant variables. The assumption of myopic consumers simplifies the model, since in a model with dynamic consumers and firms, the expectations over the future of consumers and firms have to be compatible. In a dynamic supply and demand model

[13] Note that this represents all Medicare beneficiaries that are not eligible for low income subsidies, and did not receive their Part D coverage through their employer or through special groups like Veteran Affairs.

with myopic consumers, the only expectations that are taken into account in the model are firms' expectations.

4.1.1 Flow Utility and Value Functions

Let's now present the model formally. I begin by discussing the flow utility for each enrollee, to later construct the value functions and, finally, the choice probabilities that are used for estimation of the model and to describe how the state variables evolves. Let the consumer linear utility function for plans be as follows:

$$f_{ijmt} = \begin{cases} X_{jmt}\Pi + \alpha_i p_{jmt} + \xi_{jmt} + \mu_{ijmt}, & \text{if} \quad j \neq 0 \\ \mu_{i0mt}, & \text{if} \quad j = 0 \end{cases}$$

where for individual i in plan j, market m and time t, X_{jmt} are characteristics of the plan, p_{jmt} is the observed premium of the plan, ξ_{jmt} is the unobserved quality of the plan, and μ_{ijmt} are $i.i.d$ Type 1 Extreme Value idiosyncratic shocks. Note that Π is a vector of parameters that is not defined at the individual level but α_i is random coefficient. I assume that α_i is normally distributed with mean $\bar{\alpha}$ and variance e_α.

Because consumers have inertia in their plan choice, the value function will depend on the plan the individual had in the last period (j_{it-1}). I will assume that in order to change plans a consumer has to pay a utility cost of η. Also, because some plans exit the market each period, we need to define the value function for those consumers whose plan is still in the market and for those whose plan is not in the market anymore. I define a function Γ that takes a value of 1 if plan j_{it-1} is still in the market ($\Gamma(j_{it-1}) = 1$), and a value of 0 otherwise.

With this notation, the value function for a consumer whose plan j in period $t-1$ is still in the market in period t can be defined as:

$$V(j_{it-1}, \mu_{ijmt}|\Gamma(j_{it-1}) = 1) \quad = \quad max\left\{ f_{ij_{it-1}mt}, \max_{\substack{j \in \mathbb{J}_{mt} \\ j \neq j_{it-1}}} \left\{ -\eta + f_{ijmt} \right\} \right\}$$

where the value function represents the trade-off for the consumer of staying in the same plan and receiving the utility for that plan, or picking the best plan available in her choice set and paying the cost of switching.

On the other hand, the value function for a consumer whose previous plan is not in the market anymore or who had chosen the outside option in the last period ($\Gamma(j_{it-1}) = 0$) can be written as:

$$V(j_{it-1}, \mu_{ijmt}|\Gamma(j_{it-1}) = 0) \quad = \quad \left\{ \max_{\substack{j \in \mathbb{J}_{mt} \\ j \neq j_{it-1}}} \left\{ f_{ijmt} \right\} \right\}$$

where the value function represents that this consumer is only picking the best option among the available plans. Note that consumers do not pay the cost of switching if the plan is not in the choice set anymore or if they had chosen the outside option last period.

4.1.2 Choice Probabilities

Since I estimate the model using data on market shares for each plan, I need to derive the market shares from my model. The choice probabilities are calculated according to the standard formula in Logit models. Let $Pr_j^i(j'$ be the probability of choosing plan j in period t having chosen plan j' in period $t-1$. The details about the calculation these probabilities are explained in Appendix A. The choice probabilities can be used to express the expected market share in the current period for a plan j ($s_{\hat{jmt}}$) as follows:

$$s_{\hat{jmt}} = \int \Big[\sum_{j' \in \mathbb{J}_{mt-1}} s_{ij'mt-1} Pr_j^i(j') \Big] dF_i$$

where $s_{ij'mt-1}$ is the period $t-1$ market share for a plan j' in market m for consumer type i, which is multiplied by the probability of choosing plan j in period t having chosen plan j' in period $t-1$, and then added across all plans in the same market, and integrated over all consumer types. Note that this demand model is used not only for estimation but also to compute the transition among states for the dynamic model of supply.

4.2 Supply Model

My supply model captures the dynamic competition of multi-product insurers that each period decide on offering new plans and/or withdrawing some plans, and on which premiums to set for the plans in the market. Modeling the supply side of Medicare Part D also has several challenges because of the regulations involving the profit functions of the plans. In this section, I develop the model I use for estimation and I explicitly discuss the assumptions that I make to simplify the problem, keeping the first-order aspects of the "investing-then-harvesting" dynamics. I proceed by first describing the per-period profits without considering the possibility of entry and exit, and then I incorporate these possibilities in the profit function. To construct the per-period profits of the plans without the possibility of entry and exit I follow the exposition of Decarolis et al. (2015). After this, I derive the value function for each incumbent including the possibility of exit. Finally, I discuss the value function of a potential entrant.

4.2.1 Per-period profits

I begin by discussing the revenues of a plan offered by a firm in a Medicare Part D market m in a year t. For simplicity of exposition, I omit the market and time subindices. On the revenue side, a plan collects an enrollee premium (p_j) from each individual that the plan enrolls. At the same time, the plan collects an individual-specific subsidy ($z_i(\bar{b}, r_i)$) from the government, which is composed by the baseline subsidy (that depends on the average bid $\bar{b}$) and an adjustment for the enrollee's ex-ante health risk (r_i). For the average-risk beneficiary, the sum of the premium and government subsidy is equal to the bid that the firm submitted for the plan. This mechanism is implemented by CMS to reduce incentives for risk-based selection, ensuring that all consumers look equally profitable to plans.

The cost side is also subject to regulations. The ex-post costs of a plan are different for each enrollee, since they depend on individual prescription drug expenditures. The government mitigates part of the costs of the individuals with very high expenditures via catastrophic reinsurance provisions, which cover 80 percent of an individual's drug spending for these individuals. Also, the cost for the plan of a individual with a given drug consumption will depend on the characteristics of the plan, such as the deductible level, co-payments and co-insurance or potential coverage in the donut hole. Therefore, the cost for a individual is a function of the cost-sharing characteristics of the plan (ρ_j) and the individual's measure of health risk: $c_{ij}(r_i, \rho_j)$.

In addition to these regulations on revenues and costs, Medicare Part D implements risk corridors, which are transfers between insurers and the federal government to reduce the effects on benefits of unexpected costs for the basic Part D benefit.[14] Let G be a function that adjusts a plan's ex-post profit.

Adding across all individuals enrolled in the plan, the ex-post profit for plan j as a function of its bid b_j, and the bids of other plans in the market b_{-j} can be written as follows:

$$\Pi_j(b_j, b_{-j}) = G\Big[\sum_{i|j_{it}=j} p_j(\bar{b}, b_j) + z_i(\bar{b}, r_i) - c_{ij}(r_i, \rho_j)\Big]$$

In order to simplify the plan level profits, the subsidy and the cost can be expressed as individual-specific deviations from the baseline subsidy ($z_i = z + \tilde{z}_i$) and from the average plan-specific cost of coverage ($c_{ij} = c_j + \tilde{c}_{ij}$), respectively. Furthermore, we can denote the individual-specific difference between the cost and the subsidy as $\eta_{ij} = \tilde{c}_{ij} - \tilde{z}_i$. It is important to notice that η_{ij} captures the adverse or advantageous selection occurring at the plan level. Previous literature (Polyakova, 2016) has shown that individual-specific risk is mainly a function of the characteristics of the plan, and in particular related to whether or not a plan offers coverage in the gap. Using this finding, I define the individual-specific risk as a function of plan characteristics and, therefore, the profit function can be written as:

$$\Pi_j(b_j; b_{-j}) = G\Big[\Big(\sum_{i|j_{it}=j} p_j(\bar{b}, b_j) + z(\bar{b}) - c_j(\bar{r}_j, \rho_j)\Big) + H_j(\rho)\Big]$$

where $H_j(\rho) = \sum_{i \in J} \eta_{ij}(\rho_j)$. Therefore, I can obtain a profit function for the plan that does not have any individual specific term. Moreover, the sum of the premium and the baseline level subsidy is, by construction, equal to the bid submitted by the insurer to Medicare: $p_j(\bar{b}, b_j) + z(\bar{b}) = b_j$.

Finally, I make three assumptions in order to simplify the per-period profits. First, I assume that firms do not game the subsidy structure of Medicare Part D when they make the bids. Although Decarolis (2015) shows that firms play strategies taking advantage of the institutional setting of the subsidy, I think this type of games are an additional factor for premium increase and not a first order

[14] CMS describes the risk corridors as: "Specified risk percentages above and below the target amount. For each year, CMS establishes a risk corridor for each Part D plan. Risk corridors will serve to decrease the exposure of plans where allowed costs exceed plan payments for the basic Part D benefit." See 42 Code of Federal Regulations 423.336(a)(2).

important factor for the "investing-then-harvesting" mechanism studied in this paper. Note that this assumption is important in my paper, because it allows me to consider the decisions of pricing, entry and exit for each plan as independent from the other plans that the firm has in the market. Second, I assume that firms only take into account the ex-ante risk of beneficiaries in their decisions, and not the adjustments implied in the H function or the risk corridors. Note that previous evidence suggests that the individual-specific risk component is related to the characteristics of the plan and therefore it is partially included in marginal costs. Finally, I assume that the cost for regular and LIS enrollees is the same and that there are no risk-corridor adjustments.

Given these assumptions, and bringing back market and time subindices, now I can re-write the profit of plan j in period t as:

$$\Pi_{jmt}(b_{jmt}, b_{-jmt}) = (b_{jmt} - c_{jmt})s_{jmt}M_{mt}$$

where b_{jmt} is the bid for plan j in market m at time t, b_{-jmt} represents the bids for the other plans in the same market and time, s_{jmt} is the market share of plan j, and M_{mt} is the market size. Note that s_{jmt} incorporates all the regulatory details and the dynamic elements of the demand for plans for regular and LIS beneficiaries that were discussed in Section 4.1. Also, because characteristics are assumed to be exogenous, I use the non-characteristics-related part of the bid (the projection of the bid on the state variables) in the per-period pay-off function. Finally, the costs are a random shock ($c_{jmt} = \varepsilon^c_{jmt}$) that is private information and has a normal distribution such that $\varepsilon^c_{jmt} \sim \mathcal{N}(\bar{c}, e^2_c)$.

Finally, plans face fixed costs unrelated to enrolled beneficiaries, represented by the function Λ, which vary depending on the current status and action (a_{jmt}) that a plan takes:

$$\Lambda(a_{jmt}, \kappa_{jmt}, \phi_{jmt}) = \begin{cases} -\kappa_{jmt}, & \text{if the plan is a new entrant,} \\ \phi_{jmt}, & \text{if the plan exits the market.} \end{cases}$$

To enter the market, plans have to pay a fixed cost, κ_{jmt}, which is private information draw from a common distribution of entry costs, F_κ. On the other hand, when a plan exits the market, it receives a payment of ϕ_{jmt}, which can be positive or negative, and represents a scrap value of shuttering a plan. This payment is i.i.d. private information and it is drawn each period from the common distribution, F_ϕ. Finally, note that these parameters are assumed to be independent of the status of the plan in the market.

Putting together the costs and revenues of each plan for different strategic decisions, the per-period payoff function is:

$$\Pi_{jmt}(b_{jmt}, b_{-jmt}, \vec{a}_{mt}, \vec{\varepsilon}_{mt}) = (b_{jmt} - c_{jmt})s_{jmt}(\vec{b}_{mt}, \vec{a}_{mt})M_{mt} - \Lambda(a_{jmt}, \kappa_{jmt}, \phi_{jmt}) \tag{1}$$

. where $\vec{a}_{mt}$ is the vector of actions for all plans in market m at time t, $\vec{\varepsilon}_{mt}$ is the vector of random cost shocks for all the plans in the market, and $\vec{b}_{mt}$ is the vector of bids for the plans in the market.

In the following subsections I discuss what the value functions are for plans and the equilibrium of this game.

4.2.2 Value Functions

In each discrete time period, the plan j makes entry, exit and pricing decisions, which I denote as a_{jmt}. A dynamic game as the one proposed has many possible Nash Equilibria. Following previous literature (Maskin and Tirole, 1988; Ericson and Pakes, 1995; Ryan, 2012), I restrict plan strategies to be anonymous, symmetric and Markovian, meaning that plans only condition on the current state vector and their private shocks when making decisions. A plan's strategy is a mapping from states and shocks into actions ($\sigma_j : (\vec{S}_{jmt}, \varepsilon_{jmt}) \rightarrow a_{jmt}$), where $\vec{S}_{jmt}$ is the vector of state variables for plan j in market m and time t (its lagged market share s_{jmt-1}, the lagged number of firms in the market N_{mt-1}, the lagged average number of plans per firm in the market D_{mt-1}, and the current market size M_{mt}). Note that here ε_{jmt} represents the plan's private information about the cost of entry, exit and enrollees. In this model, σ_j is a set of policy functions which describes plan j's pricing, entry and exit behavior as a function of the present state vector.

In a Markovian setting with an infinite horizon, bounded payoffs, and a discount factor less than the unity, the value function for an incumbent at the time of the exit decision is:

$$W(\vec{S}_{jmt}, \vec{\varepsilon}_t) = max\left\{\phi_{jmt}, \max_{b_{jmt}}\left\{E\Big[(b_{jmt} - c_{jmt})s_{jmt}M_{mt} + \beta E\big[W(\vec{S}_{jmt+1}, \vec{\varepsilon}_{mt+1}|\vec{S}_{jmt})\big]\Big]\right\}\right\} \tag{2}$$

where the second term takes the expectation over the cost shocks and exit decisions in this period, and the second expectation also includes the expectations about how the state vector evolves.

Potential entrants consider entry evaluating the benefits of entry at an optimal premium and their draws of enrollee costs and entry costs. I assume that potential entrants receive an endowment of characteristics (including which firm is offering the plan and the unobserved quality of the plan) and that they are short-lived: if they do not enter in this period they disappear and take a payoff of zero forever. I also assume that plans that enter cannot leave the market in the same period. The value function for a potential entrant is:

$$W^e(\vec{S}_{jmt}, \vec{\varepsilon}_{mt}) = max\left\{0, \max_{b_{jmt}}\left\{E\Big[(b_{jmt} - c_{jmt})s_{jmt}M_{mt} + \beta E\big[W(\vec{S}_{jmt+1}, \vec{\varepsilon}_{mt+1}|\vec{S}_{jmt})\big]\Big]\right\} - \kappa_{jmt}\right\} \tag{3}$$

where the second term takes expectations over the cost shocks and the second expectation also includes the expectations about how the state vector evolves.

The Markov Perfect Nash Equilibrium (MPNE) requires each plan's strategy profile to be optimal given the strategy profiles of other plans. Note that this means that the value function evaluated at the optimal strategy should be larger than the value function evaluated at any other alternative strategy. The introduction of private information makes the best response functions continuous and, therefore, it guarantees the existence of at least one pure strategy equilibrium (Doraszelski and Satterthwaite, 2010), although there are no guarantees that this equilibrium is unique.

5 Empirical Strategy and Identification

In this section, I present the methods I use for the estimation of the structural model. I begin by discussing the estimation of the demand systems and, after that, I present the estimation of the supply side in two steps. In each of these subsections I discuss the methods involved as well as the aspects related to the sources of identification of the parameters.

5.1 Demand Estimation and Identification

The estimation of the demand model follows Berry et al. (1995). For comparison purposes, I also report estimates using the logit model of Berry (1994). While the latter can be estimated in a linear fashion, the estimation of the former requires to solve for the Berry (1994) inversion nested in the general method of moments estimation.

In my demand models I include, in addition to the annual premium, a rich set of variables to describe plans' financial and non-financial characteristics. First, I include the deductible and the extra coverage in the donut hole, to characterize the financial risks associated to each plan. Second, I include a set of variables to describe the access to and cost for different drugs. This includes the number of drugs in the plan formulary, and also the number of top 100 drugs that are included in the formulary. Additionally, I include the number of top drugs that are in the first and second tiers of the formulary. Finally, I include the size of the pharmacy network, and the number of preferred pharmacies in the network. I also use firm, year, and region indicator variables. As in the rest of the paper, I assume that the characteristics of the plans are exogenous.

Even when I include a rich set of controls for observable characteristics, there are concerns about the identification of the demand models. For example, there may be room for other unobservable characteristics at the plan-specific level to be correlated with premiums and generate endogeneity issues. These unobserved factors can be any aspect of unobserved quality of the plans, such as advertising or customer services. Because of these factors, I need to instrument for the premium. Note that here the source of endogeneity in plan premiums is that plans with higher shocks of unobserved quality will set higher prices, implying a correlation between the unobserved quality term (ξ_{jmt}) and the plan premium (p_{jmt}).

For the identification of the model for regular enrollees I rely on four instrumental variables. Three of them are BLP-type instruments (Berry et al., 1995): the number of PDP plans in the market for each insurer firm, the average number of plans with extra coverage in the gap in the market, and the average, across plans in the market, of the number of top 100 drugs included in the plan formulary, where the last two averages are computed without using information of the plan being instrumentalized. The fourth instrument is a Hausman-type instrument (Hausman, 1996). The idea behind this type of instrument is that the prices in different markets share some common cost shocks that are not correlated with market specific demand shocks. As an instrument for the premium, I use the average of the premiums of plans offered by the same firm in other regions.[15]

[15] To compute this instrument, I use the division of US in four regions: East, South, Midwest, West. My instrument is the average premium of plans offered by the same firm in the other three regions that do not include the market of reference. The results are robust to alternative ways to compute this Hausman-type instrument.

The idea of this instrument seems to be particularly correct for Medicare Part D markets due to its regulatory structure, where CMS treats each market separately in order to determine the premiums of plans.

For the identification of the demand model for LIS enrollees, I only use the instrumental variables described above that are relevant for LIS enrollees, whose out-of-pocket expenditures are subsidized by the government. Thus, I do not use the instrument related to extra coverage in the gap. The instruments used in the first stage of the LIS demand model are the firm's number of PDP plans in the market, the Hausman-type instrument (average premium of plans offered by the same firm in other regions), and the average number of top 100 drugs in the plan formulary for plans in the market.[16]

5.2 Estimation of Policy Functions

To estimate the supply side model, the first step is the estimation of the policy functions that characterize the pricing and entry and exit behavior of plans conditional on the state variables.

Pricing policy function I use a linear model to estimate the pricing equation of plans, conditional on being in the market in this period. The optimal pricing policy function depends on the vector of state variables. As discussed before, the state variables in my model are the lagged firm market share s_{jmt-1}, the lagged number of firms in the market N_{mt-1}, the lagged average number of plans by firm in the market D_{mt-1}, and the current size of the market M_{mt}. When a plan is new to the market, I assume that the lagged market share of the plan is zero. Thus, the investing-then-harvesting behavior is linear in lagged market shares, with new entrants having no incentives to harvest. In addition to the state variables, I include all the characteristics included in X_{jmt}, and also the estimated unobserved quality from the demand model ξ_{jmt}. The equation for a plan's bid is therefore as follows:

$$b_{jmt} = \gamma_0^b + \gamma_1^b s_{jmt-1} + \gamma_2^b N_{mt-1} + \gamma_2^b D_{mt-1} + \gamma_3^b M_{mt} + \nu_1^b \xi_{jmt} + \nu_2^b X_{jmt} + \varepsilon_{jmt}^b$$

Exit policy function I characterize the probability of exit using the following Probit regression:

$$\Pr(\text{Exit} = 1|\vec{S}_{jmt}) = \Phi(\gamma_0^E + \gamma_1^E s_{jmt-1} + \gamma_2^E N_{mt-1} + \gamma_2^E D_{mt-1} + \gamma_3^E M_{mt} + \nu_1^E \xi_{jmt} + \nu_2^E X_{jmt})$$

where Φ is the Standard Normal CDF. Like the pricing equation, the exit probability depends on the vector of state variables and the exogenous observed and unobserved characteristics of plans.

Entry policy function The entry policy function only depends on the state variables related to the market. The firm's lagged market share and other plan characteristics are not considered in this policy function. To estimate the effect of this three state variables (lagged number of firms in the market, lagged number of plans per firm, and market size) on entry decisions, I assume that each period up to 10 plans can potentially enter the market. Based on this assumption, I create 10

[16] I use the estimates from the full model with inertia and random coefficients as my main estimates of demand for both regular and LIS enrollees, but I check for robustness of the estimates to different types of instruments and specifications. I discuss these issues in more depth in the results section.

observations per market-year that correspond to each of these potential entrants. For each of these observations I create a dummy variable indicating whether the potential entrant enters the market, which takes the value of 1 for the same number of observations per market-year as the number of plans that actually enter the market. Also, for each of these observations I create variables for the average, across all plans in the market, of the observed characteristics and estimated unobserved quality of plans. Therefore, the entry policy function is modeled using a Probit regression, where the binary decision of entry is regressed on the state variables and the average characteristics of the market in each period:

$$\Pr(\text{Entry} = 1|\vec{S}_{jmt}) = \Phi(\gamma_0^N + \gamma_1^N N_{mt-1} + \gamma_2^N D_{mt-1} + \gamma_3^N M_{mt})$$

I assume that once the plan enters the market, it is endowed with a complete vector of characteristics, which I randomly draw from the distribution of characteristics of 100 plans that entered Medicare Part D at some point during my sample period.

5.3 Estimation of Structural Parameters

The estimation proceeds in two steps. The first step provides the policy functions, which determine the optimal actions of plans at each state and how the state variables evolve. The second step finds parameters that make these observed policy functions optimal, given the underlying theoretical model. The parameters that I estimate in this second stage are the structural parameters of the model: the mean and standard deviation of distributions of the cost, entry and exit shocks. In this subsection, I present the intuition for the estimation of these parameters. I leave the details for Appendix B, where I explain the construction of the function for the minimum distance estimators.

The main insight for the estimation of the parameters of the marginal cost and exit scrap values distributions is that, under a Markov Perfect Nash Equilibrium, the value function for plans evaluated at the true policy should be, in every period, larger than the same value function evaluated at any fake policy function. However, because of simulation error and errors in the policy function, in the computed value of the value function evaluated at an alternative policy can be higher than the value at the actual policy. The intuition of the estimation is to take these situations and choose the value of the parameters that minimize the square sum of these deviations. In order to implement this estimation, I construct the value function for the actual policy function and for 500 fake policy functions. Since the value function for any policy function depends on the future states, via the continuation value, these value functions are constructed using forward simulation of the actions for all competitors. In order to do that, I take draws for the private shocks for the next 100 periods and use the actual optimal policy functions estimated in the first step of the supply estimation. I repeat this process for 500 initial values in the state space, which I take from the observed values in the data. Note that the linearity of the unknown parameters is useful during the minimization, since I do not have to recomputed separate outcome paths for each set of parameters.

Having recovered the policy functions and the parameters needed for the construction of the payoffs for incumbents, it is now possible to find the parameters of the distribution of entry costs.

The main intuition behind the estimation of these parameters is to match the probabilities of entry in a given state of the market, obtained using the policy function, with the probabilities predicted by the model. In order to implement this, I consider 1000 initial states and compute the probabilities of entry in each of them using the policy function. Using the value of the parameters for marginal costs and exit costs, I compute the continuation value that the potential entrant would get if it enters the market. I then estimate the parameters of the distribution of the entry costs by minimizing the difference between these two probabilities, which is a function of the value of the unknown parameters. Finally, the standard errors of all these parameters can be computed by bootstrapping over the different market histories (Bajari et al., 2007; Ryan, 2012).

5.4 Identification of the Supply Model

The first stage of the supply-side model recovers the policy functions that describe the optimal policy of firms in each particular state. The identification of the policy functions is crucial, since the second stage uses these policy functions, as well as fake policy functions, in order to estimate the structural parameters. The fundamental source of identification for the parameters of these policy functions is the observed actions of firms about entry, exit and pricing under different points of the state space. The policy functions are estimated by controlling for the observed characteristics of plans and also for market fixed effects. In addition, a crucial aspect of the identification of these policy functions is the inclusion of the unobserved quality estimated from the demand side model. This is important because the fact that premiums are correlated with lagged market shares could be explained by the existence of persistent unobserved quality of plans. It is also possible to think that plans with higher unobserved quality are less likely to exit or that entry depends on the average unobserved quality of plans already in the market. Therefore, without controlling for unobserved quality, the effects of the lagged market share are confounding with persistent quality of plans. The availability of an estimate of the unobserved quality of each plan from the demand model allows me to identify these policy functions, controlling for these unobserved aspects in addition to the usual controls for observed characteristics and fixed effects.

The second stage of the supply-side model rationalizes the choices of plans about premiums, entry and exit, by estimating the fundamental parameters of the distribution of marginal costs, entry sunk costs and exit scrap values. The identification of the marginal cost comes from pricing decisions of plans under different market structures in the different states of the game. In the estimation of the structural parameters, plans make pricing decisions in many different states, both because profits are forward-simulated in different market histories, and because many states are used to initialize these simulations. Additionally, the entry costs and exit scrap values are identified from the entry and exit decisions of plans and the duration of plans in the market. For example, the variation in the observed probability of entry in states that exhibit high and low profits upon entry, helps to identify the variance of sunk costs.

6 Results

6.1 Estimates of the Demand Model

Table 4 presents the results of the demand estimates for regular enrollees. Column (1) presents the OLS estimates of the Berry (1994) logit model. As it was discussed, characteristics are assumed to be exogenous but there are concerns about potential endogeneity of premiums. Because of these concerns, the estimates in Column (2) and (3) are obtained using the instrumental variables for premiums that were discussed before. Finally, Column (4) presents the estimates for the full model, which includes inertia and random coefficients, and also corrects for the endogeneity of premiums using instruments.

As expected, the estimates for the premium coefficient are always negative, and smaller (in absolute value) in the non-instrumented specifications. The estimated coefficient for the OLS model is -0.0024 and the coefficient turns more negative in Column (2) with the IV estimates (-0.0039). In Column (3) I include a variable that counts the number of years the plan has been in the market (Plan Vintage), as used in Decarolis et al. (2015), which is a way to capture the inertia of consumers. The results in Column (3) show that the coefficient for plan vintage is positive and statistically significant, suggesting that plans that entered earlier capture a larger enrollee pool, which is consistent with the presence of consumer inertia. However, the premium coefficient is virtually unchanged by the inclusion of this variable.

The estimates of the first stage for regular enrollees are presented in the first two columns of Table 5). The four instruments used for the premium are significant. The F-statistic of the joint hypothesis of all the coefficients for the instruments being zero has a value of 79 for the specification in Column (1) and 88 for the specification in Column (2). Premiums are positively correlated with the average premium of plans offered by the same firm in other regions (Hausman-Type Instrument), suggesting that this instrument is recovering cost factors that are present in all regions. The BLP-type instruments also have the expected sign; premiums are lower for plans offered by firms with more product diversification in the market (Number of PDP from Same Firm), and for plans offered in markets with more plans with gap coverage (Average Nr PDP with Extra Coverage), while they are higher for plans offered in markets with a higher average number of top 100 drugs in the formulary of plans (Average Nr Top 100 Drugs). This discussion of the first stage is important not only to understand the results of Columns (2) and (3) of Table 4, where I correct for the potential endogeneity of premiums, but also because the same instruments are used to create moments in the GMM estimation of the full model.

The estimates for the full specification of the demand model for regular enrollees, which includes random coefficients for the premium and consumer inertia, are reported in Column (4) of Table 4. The inclusion of a parameter for the cost of switching plans allows the model to rationalize observations for which premiums increase and market shares tend to respond very little, because consumers have inertia to choose the same plan they chose in the previous period. In this sense, the estimated price elasticity in this model is expected to be larger (in absolute value) compared to the previous estimates. Indeed, the estimated coefficient for the premium is more negative in this

specification (-0.0106). The price elasticities implied by the different specifications are -1.29, -2.09 and -5.24 in the OLS, 2SLS and the full specifications, respectively. These price elasticities are very similar to estimates found in previous literature (Lucarelli et al., 2012; Decarolis et al., 2015). The estimated standard deviation of the normal distribution of the random coefficients, although small, is significant with a value of 0.0013. Finally, the coefficient for inertia is also significant and it represents a dollar value of $743 per year (-7.8724/-0.0106= 743). The the cost of switching is similar in order of magnitude to most previous works for Medicare Part D (Abaluck and Gruber, 2013; Polyakova, 2016), although it is lower than the estimates of Miller and Yeo (2012), who estimate a dynamic model of consumer demand in Medicare Part D.

Table 6 presents the results of the demand estimates for LIS choosers. The main differences between regular and LIS enrollees are the way I construct the outside option. The key assumption is that all individuals that I observe in plans that are eligible for LIS random assignment are considered to have chosen the outside option. In addition, the demand model for LIS choosers does not include the deductible and the gap coverage as characteristics of plans, since these enrollees receive support from the government to cover out-of-pocket expenditures. Finally, the premiums that LIS choosers face are not the same as those faced by regular enrollees. LIS enrollees pay a premium that is the bid minus the LIPSA subsidy.

The estimates of the demand for LIS enrollees are qualitatively similar to those for regular enrollees. Column (1) of Table 6 presents the OLS estimates of the Berry (1994) logit model. Again, I rely on instruments to address the concerns about endogeneity and to identify the premium coefficient. Columns (2) and (3) present the 2SLS estimates, with the difference between them being that the latter includes plan vintage as a control variable. As expected, in both specifications the premium coefficient becomes more negative compared to the OLS estimate, although the change is relatively small. The first stage estimates for the 2SLS are reported in the last two columns of Table5. The estimated coefficients for all the instruments have the expected signs and they are all significant, with joint F-tests as strong as in the case of the regular enrollees.

Column (4) in Table 6 presents the estimates of the full model. The premium coefficient turns more negative, since this model also accounts for inertia of consumers. The coefficient for the premium is -0.0081, significantly larger in absolute value than the IV estimates. The price elasticities for these enrollees are -1.23, -1.39 and -4.12, in the OLS, 2SLS and full specifications, respectively. In the full specification, the standard deviation of the premium random coefficient is very small and not statistically significant. LIS consumers have a relatively more inelastic demand and they are more homogeneous in their preferences about premiums, reflected by a statistically insignificant standard deviation of the premium coefficient. Additionally, the coefficient for the cost of switching is significant and it represents an equivalent dollar value of $608 per year (-4.9241/-0.0081= 608). Overall, inertia for LIS enrollees is similar but smaller than for regular consumers.

Across all specifications for regular and LIS enrollees, in general the coefficients for other plan characteristics have the expected signs. Regular consumers derive negative utility from higher plan deductibles and, and positive utility from extra coverage in the donut hole. All consumers enjoy other measures of coverage: broader coverage of common drugs, more benefits on top drugs, and

larger pharmacy networks. Also, preferred pharmacy network are disliked by consumers.

The results are robust to the inclusion or exclusion of different characteristics and to the definition of these characteristics, in particular in the case of coverage for drugs or descriptions of the pharmacy networks. The results are also robust to the use of alternative instruments. The estimates of the premium coefficient are similar in models that include the means of all the exogenous characteristics, or where the instruments proposed by Gandhi and Houde (2015) are used. In general, all the robustness tests confirm the results discussed above.[17]

6.2 First Stage of the Supply Model

The first stage of the supply model recovers the best response functions of the players in the game. My estimation is based on the method proposed by Bajari et al. (2007) to estimate dynamic games. The main idea of this method is to recover the conditional choice probabilities, in the sense of Hotz and Miller (1993), from the actions actually observed in the data. In the first stage of the supply-side estimation, I recover the pricing policy (conditional on being in the market), and the entry and exit policies. Therefore, the identification of the parameters of these policy functions comes from the observed actions.

Pricing Policy Estimates for different specifications of the pricing policy conditional on being in the market are presented in Table 7. The main difference between these specifications and the ones presented in Section 3 is that this policy function not only controls for the observed characteristics of the plans but also for the estimated unobserved quality of each plan (ξ_{jmt}). Controlling for unobserved quality is extremely important here, since the correlation between lagged market shares and bids could be reflecting the fact that these plans have higher persistent unobserved quality and, therefore, a higher price next period. The inclusion of the unobserved quality helps to identify the causal effect of lagged market share on the bidding behavior, holding unobserved quality constant. Previous work documenting this relationship between lagged market shares and premiums uses a structure of fixed effects to control for unobservables that could be correlated with the bids ((Ho et al., 2015; Wu, 2016)). In this paper, I leverage the fact that I estimate the unobserved quality in the demand model and I include this estimated value in the equation directly.

Column (1) presents the results of regressing the bids on the lagged firm market share, plan characteristics, and region fixed effects. The estimated coefficient is 182, which means that a 10 percent point increase in the lagged market share implies an increase in the bid of \$18.3 per year. This increase represents the 3.4% of the average bid in my sample (\$540). In order to show the effect of the inclusion of the unobserved quality, Column (2) incorporates to the previous specification the unobserved quality estimated from the demand for regular plans. The effect of the lagged market share is strongly significant but its value is reduced to 150. Columns (3) and (4) present estimates where all the state variables are included. The difference between these two columns is that the latter includes the unobserved quality of plans from demand estimation. The effect of the

[17] The results for these robustness checks are not included in the paper, but they are available from the author upon request.

lagged plan market share is reduced to 122 when all state variables are included and it is reduced once again when unobserved quality is included. In all the estimations, this effect is statistically significant.

Additionally, bids are negatively related to the number of firms in the market, positively related to the market size and it is positively (though not statistically significantly) related to the number of plans by firm in last period. The F-stat of the significance of all state variables is 128. Additionally, it is important to notice that the estimated policy function captures the variation in the bids very well, with a very high R-Square (and Adjusted R-Square) for all the specifications. Finally, for each specification I also present the standard deviation of the residuals, which will be used to forward-simulate the pricing policies in the estimation of the structural parameters.

Exit Policy Estimates for different specifications of the exit policy are presented in Table 8. Panel (A) presents the estimated coefficients and Panel (B) presents the average marginal effects. Analogously to the case of the pricing policy, the correlation between lagged market shares and bids could be reflecting the fact that plans that have higher persistent unobserved quality have, for this reason, a lower probability of exiting next period. Therefore, controlling for the unobserved quality of plans is extremely important to estimate the exit policy functions.

The estimated effect of the lagged market share on the probability of exit is negative and strongly significant in all specifications. Columns (1) and (2) present the estimates of a Probit model that includes the lagged firm market share, observed characteristics and region fixed effects. The difference between these two specifications is that the latter includes the estimated unobserved quality from the demand model. The estimates show that plans with higher unobserved quality have a lower probability of exiting the market. However, the magnitude of the effect of the lagged market share is very similar in both specifications: a 10 percentage point increase in the lagged market share decreases the probability of exit by about 2.6 percentage points.

Columns (3) and (4) present the estimates of a Probit model when the whole set of state variables are included. As expected, the lagged number of firms in the market increases the probability of exit while the market size reduces this probability. The lagged number of plans by firm has a small positive effect on exit, but this coefficient is not statistically significant. The effect of the lagged market share is very similar, with a 10 percent point increase in the lagged market share decreasing the probability of exit by about 2.4 percentage points.

Entry Policy Table 9 presents estimates for different specifications of the entry policy. As in the previous table, Panel (A) presents the estimated coefficients and Panel (B) presents the average marginal effects. Column (1) presents the estimates of a Probit model with the lagged number of firms in the market, the averages across plans in the market of observed plan characteristics, and region fixed effects. The model estimated in Column (2) also includes the average estimated unobserved quality. The estimated effect of the lagged number of firms on the probability of entry is negative and strongly significant in both specifications. In particular, the estimated average marginal effects suggest that the presence of one additional competitor in the market last period reduces the probability of entry by 4.2 percentage points. Column (2) shows that the average

estimated unobserved quality has a negative but not statistically significant estimated coefficient. The positive sign means that the higher the average unobserved quality of plans in the market, the lower the probability of entry.

Columns (3) and (4) present the estimates when all the state variables are included in the model, including the lagged average number of plans per firm and the size of the market. The specification in column (4) additionally includes the average unobserved quality of plans in the market. The effect of the lagged number of firms is similar to the previous specifications, and the effects of market size and lagged number of plans per firm have coefficients and average marginal effects very close to zero. Finally, it is important to note that the entry policy function describes the entry decision based on the state variables; however, the entry decision for a plan is endogenous to the "investing-then-harvesting" dynamics, since in the model the plan compares the entry cost shock to the expected value of being in the market, conditional on its pricing decision.

6.3 Second Stage of the Supply Side

The second stage of the supply model estimates the parameters that rationalize plans' actions according to the optimal policy functions estimated in the first stage. All estimates from the second stage are presented in Table 10.

In order to compute the marginal cost and the entry and exit costs, I have to standardize all the plans to have the same characteristics. Otherwise, the variations in the parameters will recover not only the randomness of the cost shocks but also the different characteristics of the plans. Because of this, in my estimation I standardize the characteristics of all plans to be the characteristics of the average standard benefit plan (SDB) in my sample. However, the characteristics of plans, which are assumed to remain unchanged in the forward simulations, are used for the transition equation to predict the market shares over time.

The first two rows of Table 10 show the estimates of the marginal costs. The distribution of marginal costs has a mean of \$1079 and a variance of \$367. Both parameters are statistically significant at a one percent significance level. These costs imply an average markup rate of 13% for the plans. As discussed before, the identification of the marginal costs comes from the effect on premiums of the different market structures present in different forward simulations. The estimates of the marginal costs from my dynamic structural model are similar to other existing estimates by the Congressional Budget Office (CBO) and previous literature. In 2014, the CBO (Congressional Budget Office, 2014) estimated that the average net drug spending for a beneficiary in a basic benefit plan was \$1,382. My estimate is also similar to the estimates that Decarolis et al. (2015) finds by inverting the first order condition of a static demand model.[18]

The structural model estimates also show entry and exit costs that are consistent with plan decisions about entry and exit. The distribution of the scrap value has a mean of \$987,765 and a standard deviation of \$381,987. However, none of these numbers are very precisely estimated. The mean of the scrap value is statistically significant only at a 10% level, while the standard

[18] Using my estimates of demand for a model that includes the vintage of plans and following the methodology of Decarolis et al. (2015), I estimate marginal costs, inverting the first order conditions, and I find similar results to the ones in my dynamic structural model.

deviation is not statistically significant. The distribution of entry costs is estimated to have a mean of $2,426,733 dollars and a $843,536 standard deviation. Both parameters are statistically significant at one and five percent levels, respectively.

It is more difficult to evaluate the size of the entry and exit costs since, to the best of my knowledge, there are no previous studies of these type of costs. The estimation of the exit cost or scrap values is very imprecise, so it is not possible to reject the hypothesis of zero scrap value. This fact is consistent with the idea that at the time of exiting the market these plans do not receive a large amount of money for their investment, although they will certainly recover some part of the materials and other capital investments.

On the other hand, the estimates of the sunk costs of new plans by market seem large, with an average of $2.5 million per plan. Two points are import to notice here. First, there are actually large costs involved in offering a new plan. In order to set up a new plan, a firm has to design the product, which involves creating a formulary and a system for formulary management, study the actuarial and pricing characteristics of the plan, conduct the negotiations with pharmacies and negotiate drug rebates, and conduct the marketing efforts. The marketing of these plans involve large expenditures in television, radio and print ads, and it is expected that new plans bear most of the cost of this advertising as a entry cost. For example, industry reports and other sources inform that insurers offering Medicare Part D plans incur in high advertising costs.[19] The second point is that, since the dynamic model does not distinguish between the entry of new plans by existing or new firms, the estimates of the sunk cost of entry in my model are an average of the costs of the entry of new plans and the cost of entry of new firms. The sunk cost of offering new plans by new firms should be clearly larger than the cost of offering new plans for existing firms. Therefore, this fact can increase my estimates of the sunk costs.

7 Counterfactuals

The goals of the counterfactual analyses discussed in this section are twofold. The first goal is to assess the effects that a policy limiting dynamic pricing would have on consumer welfare in Medicare Part D markets, when the entry and exit of plans is endogenous to supply responses to consumer inertia. The second goal is to evaluate how these effects change according to the degree of inertia in consumer choice present in the market. Importantly, and unlike previous literature, my counterfactual exercises not only incorporate the direct effects on premiums of a policy that limits dynamic pricing, but they also allow for the entry and exit behavior of plans to endogenously respond to the implementation of this policy, indirectly affecting premiums as well.

To achieve these two goals, I compute two counterfactual exercises where dynamic pricing is not allowed, by imposing a fixed-markup policy on plans and solving for the new Markov Perfect Nash Equilibrium. Fixing expected markup rates is a feasible policy in this setting, where the expected

[19] Kaiser (2008) estimates that between October 1, 2007 and December 31, 2007, Medicare plan sponsors expended $76 million in television, print and radio ad occurrences. The report establishes this as a conservative estimate of marketing expenditures because they do not include other potentially significant costs associated with other marketing activities, such as direct mail, seminars, agent fees, billboard and other open air placements, or ads placed on websites.

drug costs from consumers can be forecasted using standard software packages that predict health conditions and drug costs for different types of consumers, based on their demographics and past health consumption.[20]

The first counterfactual scenario reflects the effects of implementing the fixed-markup policy while keeping the inertia of consumers unchanged. I denominate this counterfactual scenario "Fixed Markup with Inertia." In the second counterfactual scenario, the fixed-markup policy is implemented in a market without inertia. As explained in more detail in the next subsection, this is achieved by removing inertia from consumers and computing the supply-side responses to the game without inertia. For simplicity of exposition, I denominate this counterfactual scenario "Fixed Markup without Inertia."

7.1 Implementation

Because of the computational complexity of solving a dynamic game among the firms, I have to simplify the setting in such a way that captures the first order considerations about dynamic pricing and endogenous entry and exit, but reduces its computational complexity. Solving the new Markov Perfect Equilibrium of these games is a highly demanding computational task, and the time required crucially depends on the number of competitors in the market. The size of the state space increases geometrically with the number of competitors in the game, since the value function in equilibrium has to be solved for every state of the game. For this reason, to compute the counterfactual scenarios I assume that there are only five single-product firms in each market: four large firms and a composite firm that represents all the other inside options. This assumption is restrictive and is made to simplify the computational burden. However, the largest four firms represent about 60% of the total market in each market-year in my sample.

To be able to assess the welfare effects of the two counterfactual scenarios, I compute a baseline scenario that serves as a benchmark for the purpose of comparison with the counterfactual scenarios. The baseline scenario is obtained by computing the equilibrium of this simplified version of the market (under the assumption of five single-product firms per market) under the actually observed pricing policy. Thus, the benchmark scenario maintains two key aspects of the market, namely, inertia in consumer decisions, as estimated in my demand models, and dynamic pricing behavior of firms, as captured by the estimated policy functions of the supply model. Since the policy functions are best response functions of the game induced by the current incentives, I can solve the game using forward simulation based on the policy functions estimated in Section 6.2. To compute this baseline counterfactual, I start from a point in the state space and simulate 1000 periods ahead. As a starting point in the state space, I use the average market share of these firms across all markets in my sample.

In addition to the baseline scenario, I compute two counterfactual scenarios where dynamic pricing is not allowed, by imposing a fixed markup rate. The first counterfactual scenario ("Fixed

[20] Examples of these software packages are the Adjusted Clinical Groups (ACG) Case-Mix System by Johns Hopkins University, and the Hierarchical Condition Categories (HCC) model made by CMS. Some examples of the discussion and use of these softwares in the literature can be found in Carlin and Town (2009), Handel (2013), Einav et al. (2016), and Fleitas et al. (2016).

Markup with Inertia") reflects the effects of implementing the fixed-markup policy, while keeping the inertia of consumers unchanged. The comparison of this first counterfactual scenario to the benchmark, allows me to estimate the consumer welfare gains of implementing the fixed-markup policy, given the actual consumer inertia present in the market for Medicare Part D prescription drug plans.

In the second counterfactual scenario ("Fixed Markup without Inertia."), the fixed-markup policy is implemented in a market without inertia. More specifically, I first remove inertia from consumers to compute the supply-side responses to the game without inertia. Then, I compute consumer welfare using the new entry and exit behavior that is obtained from the game without inertia, and the pricing that follows from the implementation of the fixed-markup policy, while also taking into account the effect of consumer inertia on welfare.[21] The comparison of this second counterfactual scenario to the "Fixed Markup with Inertia" scenario, allows me to estimate how the consumer welfare gains of implementing a fixed-markup policy change according to the consumer inertia present in the market. In particular, this comparison provides an assessment of the difference of implementing a fixed-markup policy in Medicare Part D with the actual consumer inertia, to its effects on a similar market without consumer inertia.

In the two counterfactual exercises, I assume that the markup rate is fixed at 13%, which is the average markup rate estimated form my sample, and I use the model and estimated parameters presented in the previous section. With the estimated parameters of the distribution of the marginal cost, entry and exit, I compute profits using a simplified demand function, which uses the estimates of the fixed effects for each firm and the price coefficient. In the counterfactual exercise that maintains the inertia of the consumers, I also use the estimated inertia from the previous estimates. I use all these parameters from the demand model for the regular enrollees, who are the majority of the beneficiaries.

In order to solve for these counterfactuals, I cannot use the same policy functions that I estimated in the paper, because these policy functions are the best response functions for the game that is being played in my sample. Therefore, I first have to completely solve for the Markov Perfect Nash Equilibrium that these new games induce. In order to do that, I have to solve for the equilibrium value functions, basically writing them as Bellman Equations and iterating them until convergence. The state space should include all the information that firms use to make decisions about pricing, entry and exit. Since I assume a fixed markup, the information needed for pricing is reduced. The number of possible states in the scenario without inertia are the number of firms to the power of the two possible states for each firm (incumbent or potential entrant), which adds up to 32 (5^2) states. However, in the case of the counterfactual with inertia, the market shares of firms are also state variables. To make the computation feasible, I divide the market share of the firms in a grid of ten values and I interpolate the value of the Bellman Equation with the corresponding

[21] It is important to note that this second counterfactual exercise provides an estimation of the welfare gains that would be derived from a fixed markup policy in a context without inertia, while leaving unchanged the direct effects of consumer inertia on welfare. Thus, to calculate the welfare changes in this second scenario, it would not be correct to compute welfare when all the inertia is removed. Because I remain agnostic about the sources of inertia, I do not evaluate the effects of a policy that would remove consumer inertia completely. For this reason, this counterfactual exercise only considers the elimination of the non-welfare relevant part of inertia for the welfare computations.

proportions for the market shares between grid points. In this scenario, the size of the state space is 3.2 million ($10^5 \times 2^5$). After solving for the values of the Value Functions in equilibrium, I start with a point in the state space and simulate 1000 periods via forward simulation, solving for the actions of the firm in each period.

7.2 Results

The results of the counterfactual exercises are presented in Table 11. Column (1) presents the results of the baseline scenario, where firms actually play the same game they play in my sample with dynamic pricing and consumer inertia. Column (2) presents the scenario where dynamic pricing is not allowed and consumers still have inertia. Finally, Column (3) presents the results where dynamic pricing is not allowed in a market without consumer inertia. Panel (A) shows the average premiums under the different scenarios. Panel (B) presents the behavior of endogenous entry and exit, showing the distribution of the number of firms in the market in the simulations, and some descriptive statistics of these simulations, in particular the number of periods with firms entering and exiting the market. Finally, Panel (C) shows the value and percent change in consumer welfare of each counterfactual scenario, compared to the benchmark.

Dynamic Pricing and Premiums Given the policy functions estimated in Section 6.2, the investing-then-harvesting behavior implies that firms with larger market shares charge higher premiums. An additional question is whether this mechanism implies average premiums that are higher or lower than premiums without the investing-then-harvesting mechanism. The conventional wisdom by Klemperer and others is that the harvesting mechanism dominates and that the premiums are higher under these incentives (Farrell and Klemperer, 2007). On the other hand, Dubé et al. (2009) and Arie and Grieco (2014) present situations where the investing-then-harvesting dynamics reduce average premiums, with the investment phase dominating. As discussed before, the result depends heavily on the size of consumer inertia.

In my counterfactual results, the absence of dynamic pricing reduces premiums. Panel A of Table 11 shows that average premiums in the counterfactuals that impose a fixed-markup policy are 7% lower than the premiums in the current policy scenario. This implies that, when dynamic pricing is eliminated, the effect on the reduction of premiums of high-market share firms is larger than the increase of premiums of plans offered by low-market share firms, even when the average markup rate is the same by construction.[22] Therefore, the evidence presented in this paper is in line with the original works of Klemperer and others, where dynamic pricing has a predominantly anticompetitive role, and it is also in line with previous reduced-form empirical evidence for Medicare Part D (Ericson, 2014; Ho et al., 2015).

The reduction in premiums found in the counterfactuals also has implications for government subsidies. Both in reality and in my model, premiums are calculated as a fixed percentage of the bids submitted by firms, and the difference between the bids and premiums are direct government

[22] This has a similar implication about the effects on market entry. Because dynamic pricing creates incentives for new plans to enter the market with low prices, eliminating dynamic pricing increases the prices of new entrants. However, this effect is offset by the decrease in the prices of incumbent firms with large market shares.

subsidies. Thus, a 7% decrease in premiums implies the same percent reduction in bids and direct subsidies. Therefore, according to the results of my counterfactuals, the dynamic pricing behavior increases the expenditures of the government, of about 7% of the average bid for each enrollee. However, this change in the direct subsidies should be taken with caution, since there are factors other than the investing-then-harvesting behavior, such as the LIS subsidy structure, that are potentially affecting the bidding behavior of firms.

I now turn the focus of the discussion to the effects of dynamic pricing on entry and exit patterns, and the overall effects on welfare. The effects of limiting dynamic pricing on entry patterns are expected to be different in a scenario where consumers have inertia than in a scenario without consumer inertia. Intuitively, the presence of inertia makes entry less attractive to the potential entrant, since they are competing for consumers that do not switch plans very often. In this context, since incumbent firms increase their premiums in proportion to their market share, dynamic pricing opens opportunities for new entrants that choose to enter with a lower premium to get involved in an investing behavior. Therefore, it is expected that a policy that removes dynamic pricing while consumers still have inertia could harm entry in the market, because new entrants have to enter with a pricing comparable to that of the incumbents and therefore receive a very low market share. On the other hand, in a market without inertia, the policy of limiting dynamic pricing is expected to increase entry, since now consumers are more likely to choose one of the new plans.

Current Policy vs. Fixed Markup with Inertia Panel B of Table 11 shows the distribution of the number of firms in the 1000 counterfactual simulations for each scenario. The effects of the policy of fixing the markup while keeping inertia can be seen by comparing the results for the current policy equilibrium (presented in Column (1)), and the counterfactual with fixed markup and consumer inertia (presented in Column (2)). This comparison shows that the average number of firms in the market under the current policy (2.87 firms) is larger than the average number of firms in the counterfactual scenario where dynamic pricing is not allowed but consumers still have inertia (2.54). This lower number of firms is the result of the entry and exit patterns. Panel B also shows the percentage of periods with entry and exit in these two counterfactual scenarios. The number of periods with entry is substantially reduced with respect to the current policy (52%) in the counterfactual with fixed markup and inertia (38%). As mentioned before, firms find it less attractive to enter a market where consumers with inertia are tied to incumbent firms.

Overall, the comparison of these two scenarios in terms of average premiums and entry and exit behavior exhibits the two main aspects of the welfare effects of policies limiting dynamic pricing. On one hand, a policy limiting dynamic pricing can increase consumer welfare through a reduction the average premiums that consumers pay. On the other, in a context of inertia, a fixed-markup policy reduces the incentives for firms to enter the market, and therefore can reduce consumer welfare by limiting the options available for consumers and increasing the market size of the inside options. The net effects of the decreases in average premiums and reduction of entry are summarized in the consumer welfare calculations presented in Panel C. The average consumer welfare under the

current policy is \$360,[23] and it is \$372 under the counterfactual scenario with fixed markups and consumer inertia. Therefore, a policy of not allowing dynamic pricing increases consumer welfare by 3.1%.

Fixed Markup without Inertia A possible way to disentangle the effects of inertia and dynamic pricing in this setting is to compute the counterfactual scenario where dynamic pricing is not allowed in a context without consumer inertia. This exercise sheds light on what would be the effects of policies limiting dynamic pricing in a setting where entry is not harmed because inertia is not discouraging entry. Also, comparing the two scenarios with dynamic pricing, with and without inertia, can help us understand the burden that inertia imposes in a context of a policy that limits dynamic pricing.

The full effect of removing dynamic pricing and inertia at the same time on entry and exit can be obtained by comparing Columns (1) and (3) of Table 11, Panel B.This comparison shows that the average number of firms in the market under the current policy (2.87 firms) is lower than the average number of firms in the counterfactual scenario where dynamic pricing is not allowed and consumers do not have inertia (3.18). This increase in the average number of firms in the market comes from the increase in the entry of firms. Under this counterfactual, entry is present in 61% of the total periods of the simulation, while is present only in 52% of the periods under the current policy.

In addition to this effect, it is important to compare the effects of inertia on entry and exit patterns and consumer welfare from comparing the two scenarios with fixed markups. The results for these effects can be found comparing Columns (2) and (3) of Table 11, Panel B. While both scenarios limit dynamic pricing, Column (3) shows much more entry (61%) than the scenario with consumer inertia (38%). Therefore, removing inertia substantially increases entry in a market with a policy that limits dynamic pricing in place.

The overall effect of a fixed-markup policy in a context without inertia on consumer welfare[24] is shown in Panel C, Column (3). This scenario represents an increase of 9.4% in welfare over the scenario with the current policy (consumer welfare increases from \$360 dollars to \$394). These results imply that the positive welfare effects of a fixed-markup policy are larger in markets with less consumer inertia. When comparing the welfare effects in the counterfactuals with a fixed markup policy with and without inertia, the policy in the scenario without inertia increases consumer welfare by \$22 more, which represents a 5.9% increase in consumer welfare compared to the situation in which dynamic pricing is not allowed and consumers have the actual inertia estimated for the

[23] This estimate is similar, although a little higher, than the estimate of consumer welfare by (Decarolis, 2015), who estimate a consumer welfare of about \$280 per enrollee

[24] As explained in the previous subsection, computing consumer welfare in a situation where inertia is removed is not trivial because in this paper I remain agnostic about the sources of inertia, for example without taking a stance on whether they are actual costs that can be removed via policy or other factors. For my computation of consumer welfare I do not want to compute the reduction of inertia (making $\eta = 0$) as a welfare increase, because inertia is not a policy variable and because this reduction would directly increase welfare, but for reasons other than the competition in the market. To assess this counterfactual, I want the non-welfare-relevant change in inertia given by supply-side changes. Therefore, to compute welfare I use the demand function of the consumer with inertia but I impute the supply-side responses computed from a game where consumers do not have inertia.

Medicare Part D market.

The main takeaways from these counterfactuals are twofold. First, in markets with consumer inertia, a policy that limits dynamic pricing improves consumer welfare through lower average premiums, but this effect is partially offset by a negative effect on market entry. Second, the welfare effects of limiting dynamic pricing are sensitive to the level of consumer inertia present in the market; in particular, this type of policy is more effective in terms of increasing consumer welfare in markets with lower levels of consumer inertia.

8 Conclusion

When consumer decisions have inertia, firms have incentives to use dynamic pricing by first reducing the price to build a large market share, and increasing the price later. This strategy may reduce consumer welfare by increasing the prices of incumbents in the market and by changing the patterns of market entry and exit. This paper evaluates the effects on consumer welfare of policies that limit the capacity of firms to engage in dynamic pricing behavior, while taking into account that market entry and exit of plans are endogenous to these dynamic incentives. It also evaluates how sensitive the welfare effects are to the degree of consumer inertia present in the market.

Using data for prescription drug plans in Medicare Part D, I estimate a dynamic model of demand and supply for plans that accounts for dynamic pricing as well as endogenous entry and exit of plans. On the demand side, I recover the inertia of consumers as well as other parameters of their preferences, including the unobserved quality of plans. On the supply side, I first document that incumbents engage in dynamic pricing behavior. I find that a 10 percentage point increase in a firm's lagged market share increases its bid by about \$9 or 1% of the average bid, after controlling for plan observed and unobserved quality. Second, I model and estimate a game where multi-product firms make decisions on market entry, exit, and dynamic pricing, which allows me to recover the distribution of marginal costs, entry costs, and scrap values for plans. With the estimates of the demand and supply model, I evaluate the consumer welfare effects of a policy that fixes markup rates, by computing two counterfactual scenarios where this policy is implemented, with and without consumer inertia. For each scenario, I solve for the new Markov Perfect Nash Equilibrium induced by the regulation and the behavior of consumers.

The counterfactual analysis shows two main results. First, a fixed-markup policy would improve consumer welfare by 3.1% in Medicare Part D, through a reduction of premiums. This policy would also help to decrease the amount paid by the government in direct subsidies. Second, the counterfactuals show that under the presence of inertia observed in this market, this effect is the result of a reduction in premiums that is partially off-set by a reduction of entry into the market. When the fixed-markup policy is implemented in a counterfactual scenario without inertia, this policy has a substantially larger effect on consumer welfare, increasing it by 9.4%, because without consumer inertia the policy makes it more attractive for new firms to enter the market.

The results of this paper shed light on the role of consumer inertia in mitigating the positive impacts of limiting dynamic pricing on consumer welfare, through its negative effects on market

entry. This paper is the first to show that, after accounting for endogenous entry responses, the impacts of a policy that limits dynamic pricing are larger in markets with lower levels consumer inertia, because the dynamics of pricing create incentives for new products to enter the market. This highlights the importance of accounting for endogenous entry and exit when specifying and estimating models of dynamic pricing behavior. These results are in line with the theoretical models of Farrell and Klemperer (2007), which show that dynamic pricing responses of suppliers to consumer inertia reduce consumer welfare, and that this result depends fundamentally on the size of consumer inertia.

There are some aspects of the market for Medicare Part D plans that this paper does not explicitly deal with. First, I focus on the entry and exit of plans, and do separately model the entry, exit or mergers of insurer companies. While there was practically no entry of new insurers into this market, the consolidation via mergers was important in the period of my sample. Arguably, some part of this consolidation process may be endogenous to the investing-then-harvesting dynamics. Second, in my model I abstract from the games that firms may play with the low income subsidy (Decarolis, 2015). The empirical implications of this gaming in the entry behavior of plans are different from the pattern observed in this paper, but it has an effect on premiums that goes in the same direction as the behavior studied here. Finally, this paper focuses on the evaluation of policies that fix markups. However, there are other policies that can reduce the incentives for dynamic pricing by firms. Two examples of interesting extensions to be explored in future work are to analyze the effects of policies that set a cap in the changes of premiums over time, allowing firms to optimally choose premiums each period, and policies that increase the length of the contracts, requiring firms to commit to a schedule of premiums into the future.

References

Abaluck, J. and Gruber, J. (2011). Heterogeneity in choice inconsistencies among the elderly: Evidence from prescription drug plan choice. *American Economic Review*, 101(3):377.

Abaluck, J. and Gruber, J. (2013). Evolving choice inconsistencies in choice of prescription drug insurance. Technical report, National Bureau of Economic Research.

Abaluck, J. and Scott Morton, F. (2016). The impact of consumer inattention on insurance plan generosity: evidence from medicare part d.

Arie, G. and Grieco, P. L. (2014). Who pays for switching costs? *Quantitative Marketing and Economics*, 12(4):379–419.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, pages 1331–1370.

Beggs, A. and Klemperer, P. (1992). Multi-period competition with switching costs. *Econometrica*, pages 651–666.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, pages 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics*, pages 242–262.

Carlin, C. and Town, R. (2009). Adverse selection, welfare and optimal pricing of employer-sponsored health plans.

Chorniy, A., Miller, D., and Tang, T. (2016). Mergers in medicare part d: decomposing market power, cost efficiencies, and bargaining power.

Congressional Budget Office, C. (2014). Competition and the cost of medicare's prescription drug program.

Cullen, J. and Shcherbakov, O. (2010). Measuring consumer switching costs in the wireless industry.

Dalton, C. M., Gowrisankaran, G., and Town, R. (2015). Myopia and complex dynamic incentives: Evidence from medicare part d. Technical report, National Bureau of Economic Research.

Decarolis, F. (2015). Medicare part d: Are insurers gaming the low income subsidy design? *American Economic Review*, 105(4):1547–80.

Decarolis, F., Polyakova, M., and Ryan, S. P. (2015). The welfare effects of supply-side regulations in medicare part d. Technical report, National Bureau of Economic Research.

Doraszelski, U. and Satterthwaite, M. (2010). Computable markov-perfect industry dynamics. *The RAND Journal of Economics*, 41(2):215–243.

Dubé, J.-P., Hitsch, G. J., and Rossi, P. E. (2009). Do switching costs make markets less competitive? *Journal of Marketing Research*, 46(4):435–445.

Duggan, M., Healy, P., and Morton, F. S. (2008). Providing prescription drug coverage to the elderly: America's experiment with medicare part d. *Journal of Economic Perspectives*, pages 69–92.

Einav, L., Finkelstein, A., Kluender, R., and Schrimpf, P. (2016). Beyond statistics: the economic content of risk scores. *American Economic Journal: Applied Economics*, 8(2):195–224.

Einav, L., Finkelstein, A., and Schrimpf, P. (2015). The response of drug expenditure to non-linear contract design: evidence from medicare part d. *Quarterly Journal of Economics*, 130(2):841.

Ericson, K. M. (2014). Consumer inertia and firm pricing in the medicare part d prescription drug insurance exchange. *American Economic Journal: Economic Policy*, 6(1):38–64.

Ericson, K. M. M. and Starc, A. (2015). Pricing regulation and imperfect competition on the massachusetts health insurance exchange. *Review of Economics and Statistics*, 97(3):667–682.

Ericson, R. and Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62(1):53–82.

Farrell, J. and Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects. *Handbook of Industrial Organization*, 3:1967–2072.

Farrell, J. and Shapiro, C. (1988). Dynamic competition with switching costs. *RAND Journal of Economics*, pages 123–137.

Fleitas, S. (2016). The costs of lock-in in health care markets: evidence from hospital choice in uruguay.

Fleitas, S., Gowrisankaran, G., and LoSasso, A. (2016). Reclassification risk in the small group health insurance market.

Gandhi, A. and Houde, J.-F. (2015). Measuring substitution patterns in differentiated products industries.

Gowrisankaran, G. and Rysman, M. (2012). Dynamics of consumer demand for new durable goods. *Journal of Political Economy*, 120(6):1173–1219.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7):2643–2682.

Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The economics of new goods*, pages 207–248. University of Chicago Press.

Heiss, F., Leive, A., McFadden, D., and Winter, J. (2013). Plan selection in medicare part d: evidence from administrative data. *Journal of Health Economics*, 32(6):1325–1344.

Heiss, F., McFadden, D., and Winter, J. (2006). Who failed to enroll in medicare part d, and why? early results. *Health Affairs*, 25(5):w344–w354.

Ho, K., Hogan, J., and Morton, F. S. (2015). The impact of consumer inattention on insurer pricing in the medicare part d program. Technical report, National Bureau of Economic Research.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *Review of Economic Studies*, 60(3):497–529.

Ketcham, J. D., Lucarelli, C., Miravete, E. J., and Roebuck, M. C. (2012). Sinking, swimming, or learning to swim in medicare part d. *American Economic Review*, pages 2639–2673.

Klemperer, P. (1987). Markets with consumer switching costs. *Quarterly Journal of Economics*, pages 375–394.

Klemperer, P. (1995). Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade. *Review of Economic Studies*, 62(4):515–539.

Krasner, J. (2006). Insurer hits millions of seniors with drug cost hike; accessed: 06.15.2016. `https://goo.gl/UkdQPA`.

Lavetti, K. and Simon, K. (2016). Strategic formulary design in medicare part d plans.

Lucarelli, C., Prince, J., and Simon, K. (2012). The welfare impact of reducing choice in medicare part d: A comparison of two regulation strategies. *International Economic Review*, 53(4):1155–1177.

Maskin, E. and Tirole, J. (1988). A theory of dynamic oligopoly, i: Overview and quantity competition with large fixed costs. *Econometrica*, pages 549–569.

Miller, D. P. and Yeo, J. (2012). Estimating dynamic discrete choice models of product differentiation: An application to medicare part d with switching costs.

Miller, K. (2014). Do private medicare firms have lower costs.

Nosal, K. (2012). Estimating switching costs for medicare advantage plans.

Polyakova, M. (2016). Regulation of insurance with adverse selection and switching costs: Evidence from medicare part d. *American Economic Journal: Applied Economics*, 8(3):16530.

Ryan, S. P. (2012). The costs of environmental regulation in a concentrated industry. *Econometrica*, 80(3):1019–1061.

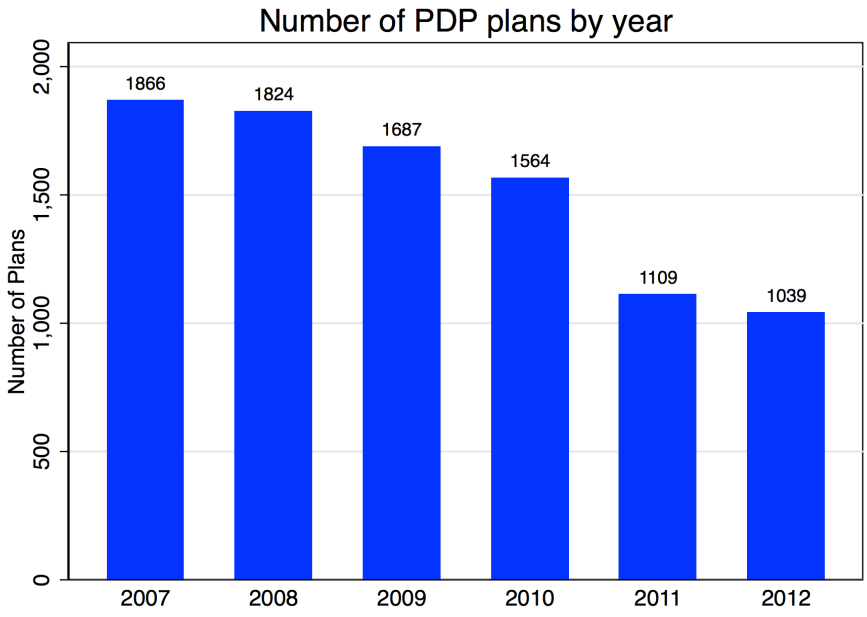
Shepard, M. (2016). Hospital network competition and adverse selection: Evidence from the massachusetts health insurance exchange.

Starc, A. (2014). Insurer pricing and consumer welfare: Evidence from medigap. *RAND Journal of Economics*, 45(1):198–220.

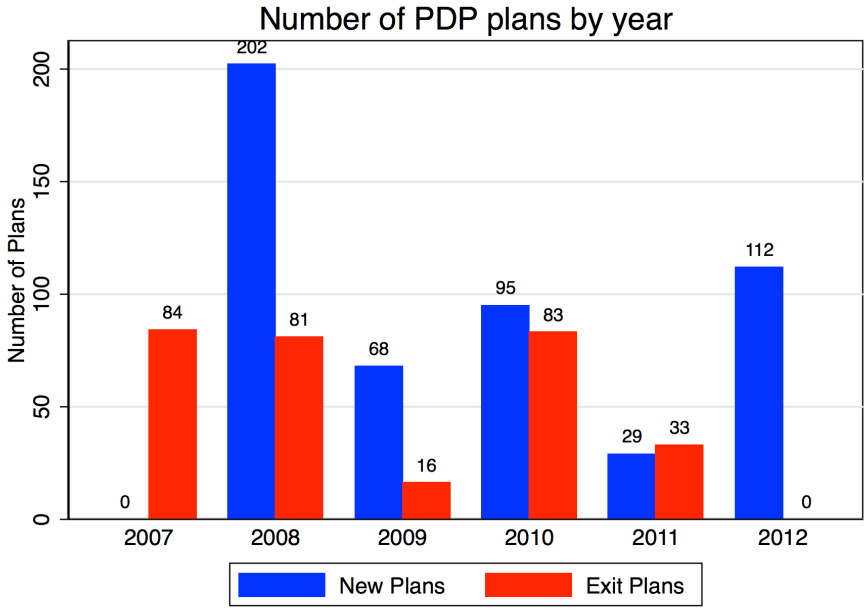
Tebaldi, P. (2016). Estimating equilibrium in health insurance exchanges: Price competition and subsidy design under the aca.

Wu, Y. (2016). Supply response to consumer inertia: Strategic pricing in medicare part d.

Figure 1: Demography of Plans by Year

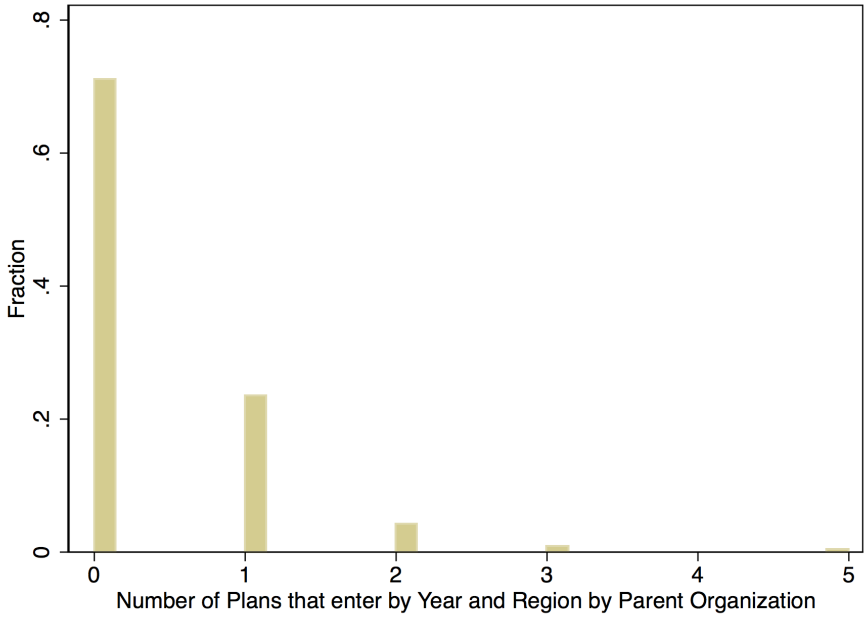


(a) Number of Plans by Year

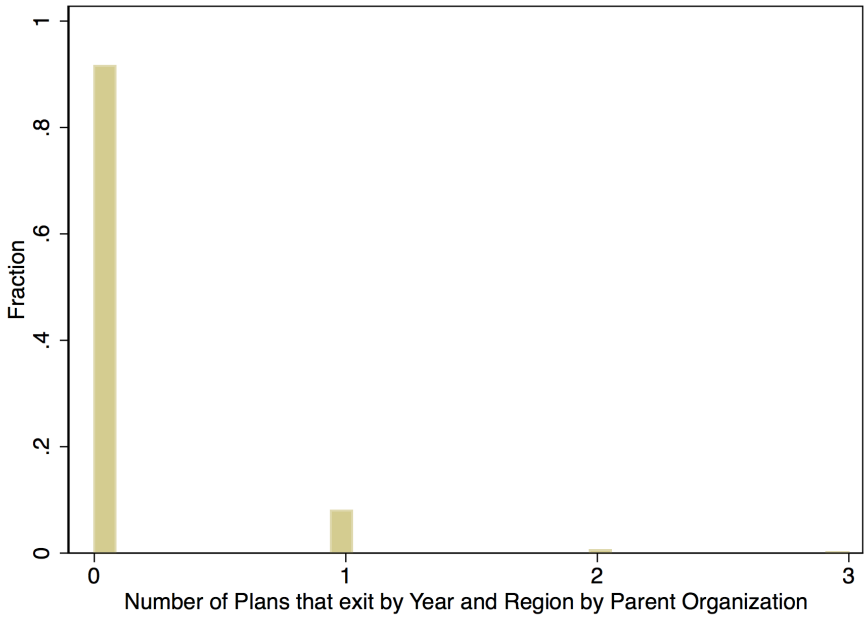


(b) Entry and Exit of Plans by Year

Figure 2: Demography of Plans by Insurer Firm

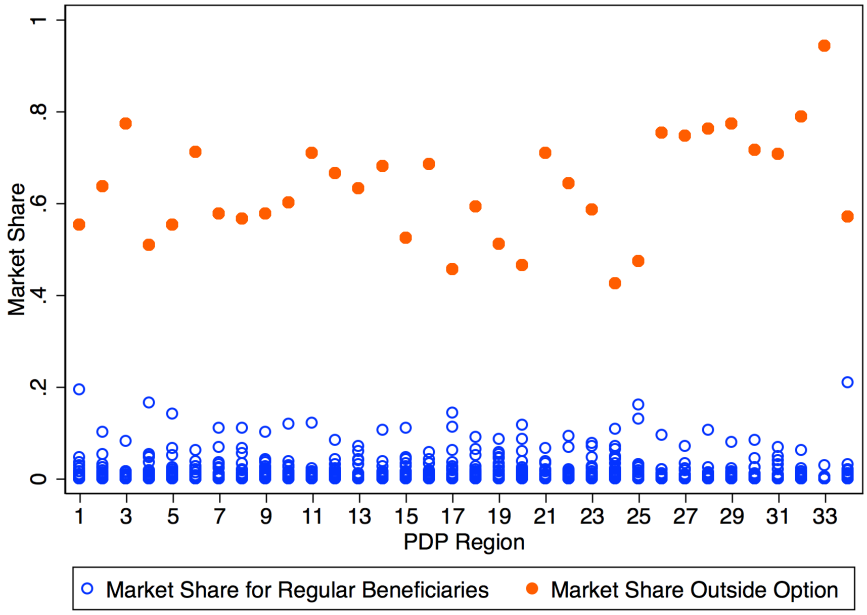


(a) Number of Plans entering by Year and Region by Firm

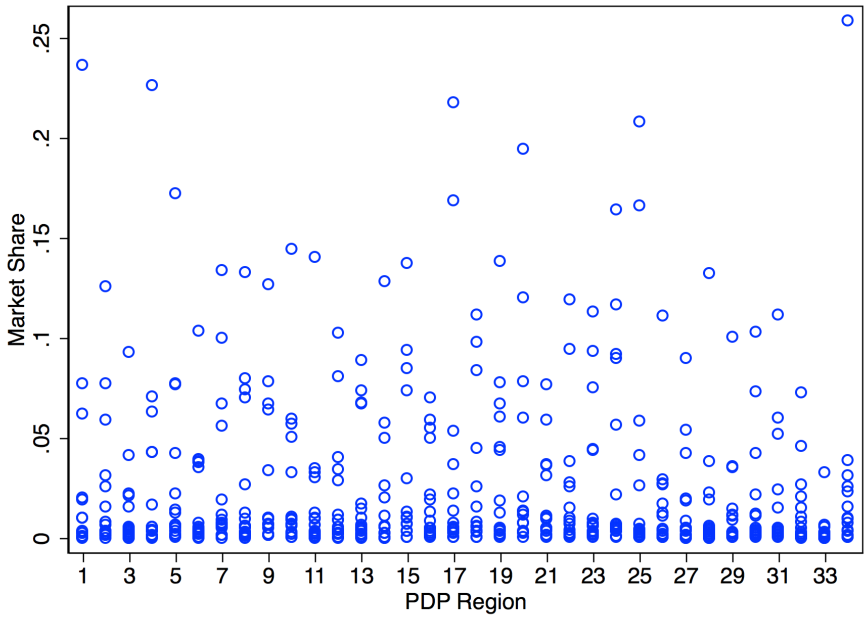


(b) Number of Plans exiting by Year and Region by Firm

Figure 3: Variation in Market Share by Makets for Regular Beneficiaries

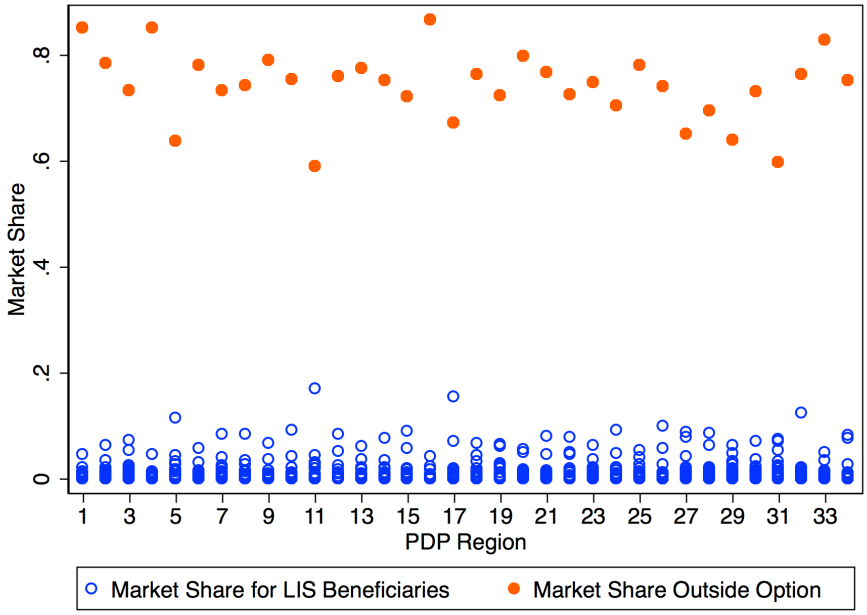


(a) Market Share by Market for each Plan (Year 2009)

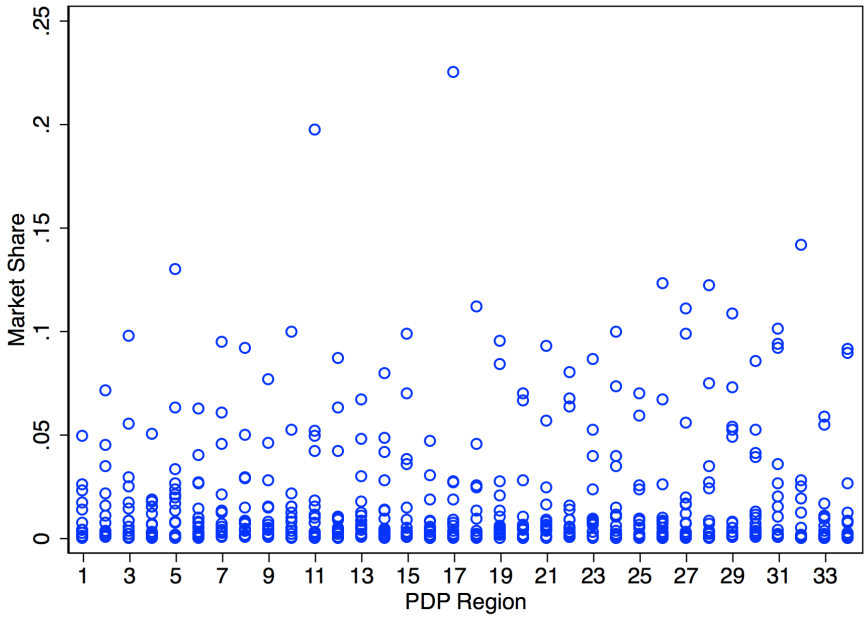


(b) Market Share by Market by Firm (Year 2009)

Figure 4: Variation in Market Share by Markets for LIS Beneficiaries

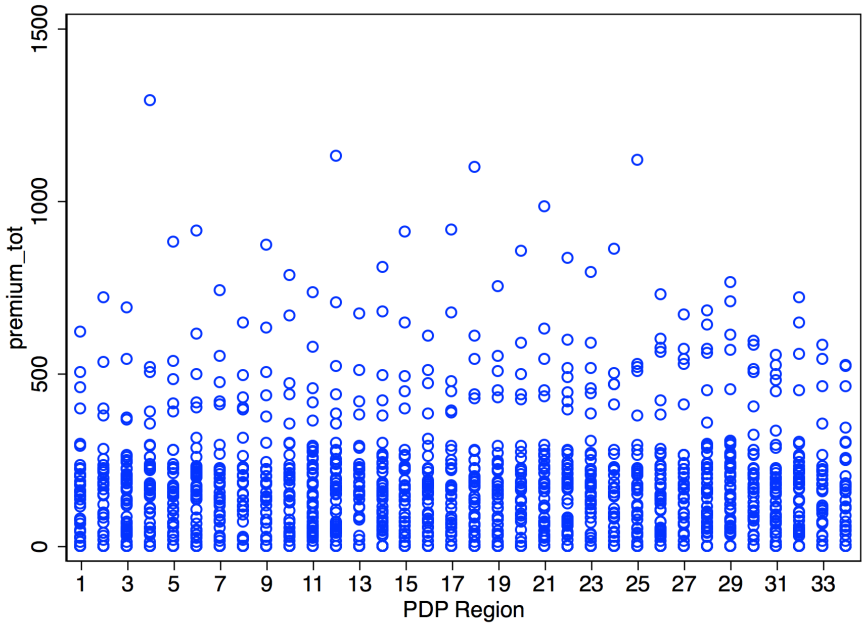


(a) Market Share by Market for each Plan (Year 2009)

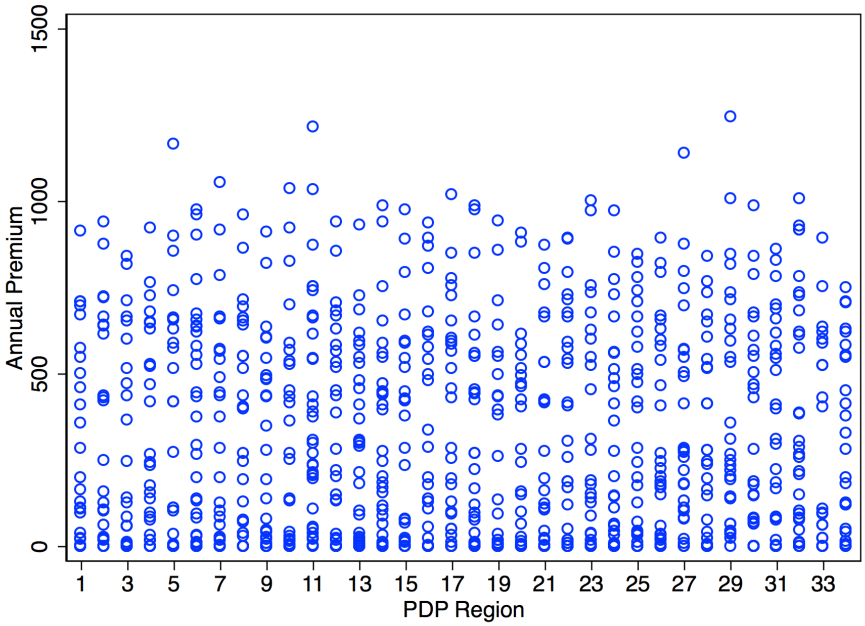


(b) Market Share by Market by Firm (Year 2009)

Figure 5: Variation on Premiums over Markets

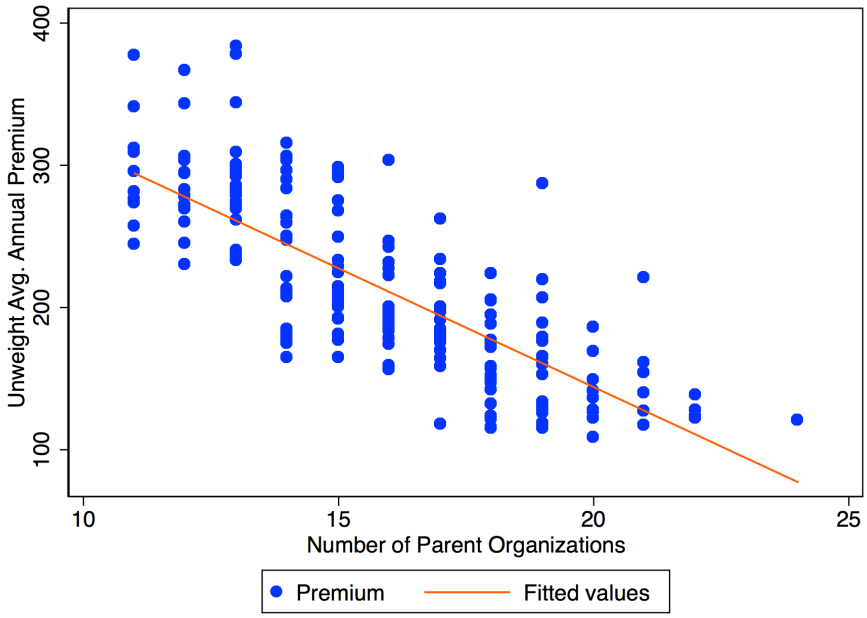


(a) Cohort of plans (Year 2007)

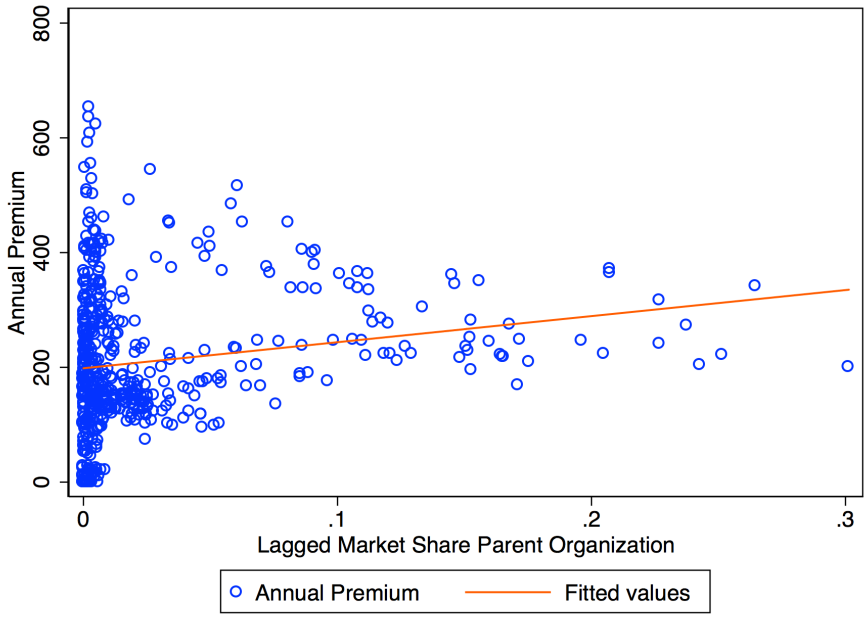


(b) Cohort of plans (Year 2012)

Figure 6: Plan premiums and market structure: Investing-Harvesting and Competition Effect

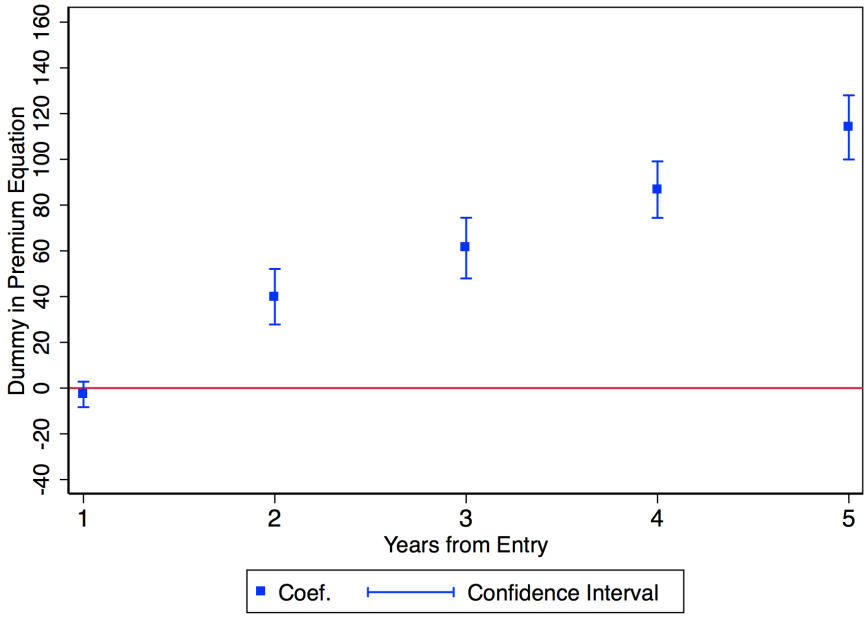


(a) Plan premiums and Number of Firm



(b) Premium and Lagged Market Share by Firm (Year 2009)

Figure 7: Plan premiums and market structure: Competition and Entry



(a) Changes in Premiums by Years of Entry

Table 1: Summary Statistics

	2007	2008	2009	2010	2011	2012
Plans						
Number of Plans	1866	1824	1687	1564	1109	1039
Average Number of PDP plans per market	55	54	50	46	33	31
Firms						
Total Number of Firms	68	61	53	52	43	40
Average Nr. of Firms per Market	21	19	16	16	13	12
Average Nr. of PDP per firm and market	2.7	2.8	3.2	3	2.4	2.5
Enrollment (in millions)						
All Part D Eligible	43.3	44.4	45.5	46.6	47.7	49.7
PDP Enrollment non-LIS	8.1	8.4	8.6	8.7	8.9	9.3
PDP Enrollment LIS	8.0	8.0	7.9	7.9	8.2	8.3
Medicare Advantage Enrollment	7.8	8.9	9.7	9.8	10.4	11.5
Employer Sponsored Coverage	7.0	6.6	6.5	6.7	6.2	5.6
Other Coverage	5.8	5.9	6.4	7.3	7.0	7.9
No Creditable Coverage	6.6	6.5	6.4	6.2	7.1	7.2
Premiums						
Average Premiums of SDB plans	336	348	420	396	432	408
Unweighted avg. annual PDP Premium	442	480	546	560	664	648
Subsidies						
CMS national average bid (annual)	965	966	1012	1060	1045	1014
CMS base consumer premium (annual)	328	335	364	383	386	373
Low income (LIS) benchmark threshold	341	332	353	387	400	390

Notes: All columns show descriptive statistics of Medicare Part D over different years. The sample includes all PDP Plans in Medicare Part D during 2007-2012.

Table 2: Demography of Entry, Exit and Number of Plans

	t=2007			t=2008				t=2009				t=2010				t=2011				t=2012	
Region	TOT	DIS	EXI	NEW	TOT	DIS	EXI	NEW	TOT	DIS	EXI	NEW	TOT	DIS	EXI	NEW	TOT	DIS	EXI	NEW	TOT
1	53	5	3	5	53	8	4	1	46	5	1	2	43	13	1	0	30	5	1	3	28
2	51	5	3	5	51	6	2	2	47	4	0	5	48	14	2	0	34	7	3	3	30
3	61	11	5	5	55	4	2	0	51	4	0	2	49	18	2	2	33	7	2	3	29
4	57	9	2	9	57	6	2	1	52	7	0	2	47	15	3	1	33	6	1	3	30
5	55	8	2	5	52	6	2	2	48	7	0	4	45	12	2	0	33	5	1	3	31
6	66	9	2	6	63	7	3	1	57	7	0	4	54	16	4	0	38	6	2	4	36
7	53	6	3	5	52	6	2	2	48	6	0	2	44	13	2	1	32	5	1	3	30
8	51	6	2	7	52	5	2	2	49	5	0	3	47	15	4	1	33	6	2	3	30
9	59	9	2	6	56	7	2	2	51	9	0	4	46	13	2	1	34	5	0	3	32
10	55	6	3	5	54	6	2	2	50	9	0	3	44	13	2	1	32	5	1	3	30
11	57	8	2	9	58	5	2	1	54	8	1	3	49	18	7	1	32	4	1	5	33
12	56	9	2	6	53	6	2	2	49	7	0	4	46	12	1	0	34	5	1	3	32
13	54	7	2	8	55	5	2	1	51	8	2	3	46	14	1	3	35	3	1	2	34
14	60	7	3	5	58	10	6	1	49	5	0	1	45	14	2	3	34	6	3	5	33
15	53	6	3	5	52	6	2	2	48	6	0	2	44	13	2	1	32	4	1	3	31
16	54	5	3	8	57	6	2	2	53	7	2	2	48	17	4	1	32	9	2	6	29
17	56	8	2	5	53	6	2	2	49	7	0	4	46	13	2	2	35	5	1	3	33
18	53	6	3	5	52	6	2	2	48	5	0	1	44	13	2	1	32	5	1	3	30
19	58	8	2	5	55	5	2	2	52	6	0	3	49	16	4	1	34	6	0	2	30
20	52	8	2	5	49	5	2	3	47	6	0	3	44	13	2	1	32	5	0	3	30
21	52	7	2	5	50	5	2	2	47	6	0	3	44	12	1	0	32	6	0	4	30
22	60	9	3	5	56	6	2	3	53	6	0	3	50	18	5	1	33	3	0	3	33
23	56	9	2	5	52	5	2	2	49	6	0	3	46	16	2	3	33	6	0	3	30
24	53	6	2	5	52	6	2	2	48	5	0	2	45	12	2	0	33	5	1	3	31
25	53	6	2	5	52	7	4	3	48	4	0	2	46	14	2	1	33	3	0	3	33
26	57	7	2	5	55	7	2	2	50	7	1	4	47	16	2	1	32	5	1	3	30
27	55	7	3	7	55	6	2	4	53	7	2	2	48	17	5	0	31	9	2	6	28
28	53	7	2	5	51	5	2	3	49	6	2	3	46	17	3	1	30	4	1	4	30
29	54	6	3	5	53	6	2	2	49	6	1	2	45	14	1	0	31	8	1	6	29
30	57	7	2	5	55	9	6	2	48	7	0	3	44	12	2	0	32	5	0	3	30
31	56	7	2	5	54	6	2	3	51	5	0	2	48	14	1	1	35	5	0	3	33
32	55	7	3	8	56	6	2	1	51	7	0	2	46	13	2	0	33	5	1	3	31
33	46	7	3	10	49	5	2	3	47	9	2	3	41	13	3	0	28	4	0	1	25
34	45	6	2	8	47	5	2	3	45	9	2	4	40	11	1	0	29	5	1	1	25
Total	1866	244	84	202	1824	205	81	68	1687	218	16	95	1564	484	83	29	1109	182	33	112	1039

Notes: This table presents the demography of Medicare Part D PDPs from 2007-2012. For each year and region, the following columns represent: *TOT* the number of plans at the beginning of the period, *DIS* the number of plans that are consolidated, *ENTRY* represents the number of plans that enter the market and *EXIT* the number of plans that exit the market.

Table 3: Reduced-Form Evidence of Biding Behavior

	(1) Bids	(2) Bids	(3) Bids	(4) Bids
Lagged Firm Market Share	146.36**	258.25***		
	(61.83)	(70.78)		
Part D Drug Deductible	-0.49***	-0.46***	-0.49***	-0.46***
	(0.01)	(0.02)	(0.01)	(0.02)
Extra Coverage in Donut Hole	331.22***	273.32***	334.91***	277.73***
	(3.79)	(4.11)	(3.79)	(4.09)
Enhanced Plan	21.01***	30.14***	20.55***	29.79***
	(4.53)	(4.71)	(4.52)	(4.78)
Number of Drugs in Formulary	-7.13***	-9.51***	-8.38***	-10.51***
	(1.50)	(1.54)	(1.45)	(1.48)
Number of Top Drugs in Formulary	10.80***	17.51***	8.56***	16.07***
	(1.37)	(1.29)	(1.46)	(1.31)
Number of Top Drugs Tiers 1-2	2.51***	0.76***	2.71***	0.99***
	(0.16)	(0.13)	(0.17)	(0.14)
Size of Pharmacy Network	-31.78***	-17.25***	-33.25***	-19.86***
	(6.40)	(6.61)	(6.19)	(6.47)
Preferred Pharmacies in Network	-19.60***	-2.65	-15.11***	-0.78
	(3.29)	(3.59)	(3.41)	(3.69)
New Plan in the Market			-62.51***	-42.02***
			(3.62)	(3.55)
Characteristics	Yes	Yes	Yes	Yes
Plan FE	No	Yes	No	Yes
Firm FE	Yes	No	Yes	No
Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	8866	8866	8866	8866

Notes: All columns show estimates of the bidding decisions for incumbents and new plans of the first-stage of the supply model. The first two columns use the lagged market share of plans as the main independent variable (new plans are assigned to zero lagged market share) and the last two columns use a dummy variable indicating if the plan is new. The dependent variable for all columns are the bids of the plans. The sample includes all PDP Plans in Medicare Part D during 2007-2012. All columns include plan characteristics (Part D Drug Deductible, Extra Coverage in the Donut Hole, Enhanced Plan, Number of Drugs in Formulary, Number of Top Drugs in Formulary, Number of Top Drugs in Tiers 1-2, Size of Pharmacy Network, Preferred Pharmacies in Network). All columns also include fixed effects detailed below the estimates. Robust standard errors (clustered by market) in parentheses. + p<0.1, ** p<0.05, *** p<0.01.

Table 4: Parameter Estimates for the Demand System for Regular Enrollees

	(1)	(2)	(3)	(4)
Annual Premium	-0.0024***	-0.0039***	-0.0038***	-0.0106***
	(0.0001)	(0.0004)	(0.0004)	(0.0002)
Part D Drug Deductible	-0.0034***	-0.0039***	-0.0036***	-0.0051***
	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Extra Coverage in Donut Hole	-0.2952***	0.1237	0.1203	0.2563**
	(0.0410)	(0.1155)	(0.1196)	(0.1124)
Enhanced Plan	0.0532	0.1732***	0.1712***	-0.0149
	(0.0445)	(0.0523)	(0.0530)	(0.0467)
Number of Drugs in Formulary	0.0840***	0.0847***	0.1027***	0.0596***
	(0.0233)	(0.0230)	(0.0224)	(0.0120)
Number of Top Drugs in Formulary	0.2439***	0.2528***	0.1942***	0.2623 ***
	(0.0153)	(0.0147)	(0.0135)	(0.0232)
Number of Top Drugs Tiers 1-2	0.0046**	0.0075***	0.0036**	0.0090
	(0.0018)	(0.0020)	(0.0017)	(0.0003)
Size of Pharmacy Network	0.0989	0.0578	0.1019	0.0137
	(0.0878)	(0.0879)	(0.0840)	(0.0027)
Preferred Pharmacies in Network	-0.2434***	-0.2644***	-0.2172***	-0.2874***
	(0.0423)	(0.0405)	(0.0387)	(0.0512)
Plan Vintage			0.2591***	
			(0.0209)	
Switching Costs (η)				7.8724***
				(0.8756)
Standard Deviation of Premium ($\sigma_{premium}$)				0.0013***
				(0.0003)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	8866	8866	8866	8866

Notes: All columns show estimates of the demand model for regular enrollees. The sample includes all PDP Plans in Medicare Part D during 2007-2012. All columns include plan characteristics (Part D Drug Deductible, Extra Coverage in the Donut Hole, Enhanced Plan, Number of Drugs in Formulary, Number of Top Drugs in Formulary, Number of Top Drugs in Tiers 1-2, Size of Pharmacy Network, Preferred Pharmacies in Network). Column (3) includes a variable that indicates the number of years the firm is on the market. All columns also include a full set of dummy variables for all (34) Medicare regions, for year and for firms (Parent Organizations). Column (1) is estimated via Ordinary Least Squares (OLS) estimation. Column (2) and (3) are estimated using Two Stage Least Squares (2SLS) and the First Stage of these estimates are shown in Columns (1) and (2) of Table 5. Column 4 is estimated using the full model, which includes a parameter for switching costs (η) and a parameter for the standard error of the premium random coefficient ($\sigma_{premium}$). Robust standard errors (clustered by market) in parentheses. + p<0.1, ** p<0.05, *** p<0.01.

Table 5: First Stage for Both Types of Beneficiaries

	(1) Regular	(2) Regular	(3) LIS	(4) LIS
Number PDP from Same Firm	-7.1682***	-4.8099+	-11.7787***	-9.0853***
	(2.4647)	(2.4382)	(2.4265)	(2.5274)
Hausman-Type Instrument	0.2619***	0.2661***	0.2700***	0.2784***
	(0.0242)	(0.0238)	(0.0306)	(0.0318)
Average Nr Top 100 Drugs	28.1593***	30.4227***	34.8641***	37.4376***
	(2.4788)	(2.4893)	(3.7418)	(3.7669)
Average Nr PDP with Extra Coverage	-55.7580***	-47.3561**		
	(18.1795)	(17.8921)		
Part D Drug Deductible	-0.4010***	-0.3799***		
	(0.0118)	(0.0124)		
Extra Coverage in Donut Hole	278.8484***	280.5126***		
	(3.4834)	(3.5615)		
Enhanced Plan	78.9309***	79.5513***	236.6834***	239.4763***
	(3.5478)	(3.5519)	(4.0259)	(3.9842)
Number of Drugs in Formulary	5.0956***	7.0026***	35.2880***	38.5187***
	(1.2423)	(1.2029)	(2.9808)	(2.9443)
Number of Top Drugs in Formulary	-8.2991***	-13.5039***	8.3711***	-1.3938
	(1.6954)	(1.7607)	(2.8129)	(2.8861)
Number of Top Drugs Tiers 1-2	1.5178***	1.3115***	-2.4058***	-2.1111***
	(0.1447)	(0.1542)	(0.3055)	(0.3189)
Size of Pharmacy Network	-29.2774***	-26.1519***	-62.9155***	-55.1077***
	(5.8872)	(5.6093)	(7.9970)	(7.4157)
Preferred Pharmacies in Network	-22.8885***	-19.9983***	-32.5373***	-22.7952***
	(3.7334)	(3.6731)	(5.3882)	(5.6567)
Plan Vintage		19.7365***		33.8145***
		(1.0027)		(1.9240)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
N	8866	8866	6004	6004
F-Test First Stage	79	88	86	88

Notes: All columns show estimates of the first stage of the Two Stage Least Squares estimates of the demand models. Columns (1) and (2) show the first stage for Columns (2) and (3) of Table 4. Columns (3) and (4) show the first stage for Columns (2) and (3) of Table 6. The sample includes all PDP Plans in Medicare Part D during 2007-2012. All columns include a set of plan characteristics (Enhanced Plan, Number of Drugs in Formulary, Number of Top Drugs in Formulary, Number of Top Drugs in Tiers 1-2, Size of Pharmacy Network, Preferred Pharmacies in Network). Columns (1) and (2) also include Part D Drug Deductible, Extra Coverage in the Donut Hole. Column (3) includes a variable that indicates the number of years the firm is on the market. All columns also include a full set of dummy variables for all (34) Medicare regions, for year and for firms (Parent Organizations). All Columns are estimated via Ordinary Least Squares (OLS) estimation. F-test First State represents the joint F statistic for the instruments included in each Column. Robust standard errors (clustered by market) in parentheses. + p<0.1, ** p<0.05, *** p<0.01.

Table 6: Parameter Estimates for the Demand System for LIS Enrollees

	(1)	(2)	(3)	(4)
Annual Premium with LIS	-0.0023***	-0.0026***	-0.0025***	-0.0081***
	(0.0001)	(0.0004)	(0.0004)	(0.0012)
Enhanced Plan	-0.9725***	-0.8854***	-0.8876***	-0.2372**
	(0.0440)	(0.0938)	(0.0927)	(0.1185)
Number of Drugs in Formulary	-0.1273***	-0.1166***	-0.0877**	-0.0953***
	(0.0377)	(0.0413)	(0.0404)	(0.0532)
Number of Top Drugs in Formulary	0.4372***	0.4477***	0.3320***	0.2132***
	(0.0280)	(0.0312)	(0.0269)	(0.0227)
Number of Top Drugs Tiers 1-2	0.0015	0.0006	0.0042	0.0013
	(0.0033)	(0.0035)	(0.0030)	(0.0065)
Size of Pharmacy Network	-0.2067***	-0.2290***	-0.1345**	-0.1246
	(0.0655)	(0.0648)	(0.0579)	(0.0793)
Preferred Pharmacies in Network	-0.2528***	-0.2590***	-0.1282***	-0.1352***
	(0.0458)	(0.0465)	(0.0402)	(0.0398)
Plan Vintage			0.4129***	
			(0.0224)	
Switching Costs (η)				4.9241**
				(0.5756)
Standard Deviation of Premium ($\sigma_{premium}$)				0.0009
				(0.0005)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Regions FE	Yes	Yes	Yes	Yes
Observations	6004	6004	6004	6004

Notes: All columns show estimates of the demand model for LIS choosers enrollees. The sample includes all PDP Plans in Medicare Part D during 2007-2012 that are not eligible for random assignment. All columns include plan characteristics (Enhanced Plan, Number of Drugs in Formulary, Number of Top Drugs in Formulary, Number of Top Drugs in Tiers 1-2, Size of Pharmacy Network, Preferred Pharmacies in Network). Column (3) includes a variable that indicates the number of years the firm is on the market. All columns also include a full set of dummy variables for all (34) Medicare regions, for year and for firms (Parent Organizations). Column (1) is estimated via Ordinary Least Squares (OLS) estimation. Column (2) and (3) are estimated using Two Stage Least Squares (2SLS) and the First Stage of these estimates are shown in Columns (3) and (4) of Table 5. Column 4 is estimated using the full model, which includes a parameter for switching costs (η) and a parameter for the standard error of the premium random coefficient ($\sigma_{premium}$). Robust standard errors (clustered by market) in parentheses. + p<0.1, ** p<0.05, *** p<0.01.

Table 7: Supply Side First Stage: Bidding Decisions for Incumbents and New Plans

	(1)	(2)	(3)	(4)
	Bid Amount ($)	Bid Amount ($)	Bid Amount ($)	Bid Amount ($)
Lagged Firm Market Share	182.2309***	150.0126***	122.3674***	91.1387**
	(40.3318)	(40.4659)	(39.3519)	(39.4766)
Demand-side Unobserved Quality		10.9417***		10.6959***
		(1.5232)		(1.4826)
Lagged Number of Firm by Market			-13.0549***	-13.1103***
			(0.8328)	(0.8304)
Market Size (in 1000)			0.4185***	0.4139***
			(0.0308)	(0.0308)
Lagged Number of Plans by Firm			0.3481	0.3860
			(0.8604)	(0.8579)
Characteristics	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
R-Square	0.92	0.92	0.93	0.93
Adjusted R-Squared	0.92	0.92	0.93	0.93
Std. Dev. Residuals	166	166	162	161
F-test State Variables	20	14	129	128
Obs.	8866	8866	8866	8866

Notes: All columns show estimates of the bidding decisions for incumbents and new plans of the first-stage of the supply model. The dependent variable for all columns are the bids of the plans. The sample includes all PDP Plans in Medicare Part D during 2007-2012. All columns include plan characteristics (Part D Drug Deductible, Extra Coverage in the Donut Hole, Enhanced Plan, Number of Drugs in Formulary, Number of Top Drugs in Formulary, Number of Top Drugs in Tiers 1-2, Size of Pharmacy Network, Preferred Pharmacies in Network). All columns also include fixed effects for all (34) Medicare regions. F-test State Variables represents the joint F statistic for the state variables included in each column. Robust standard errors (clustered by market) in parentheses. + p<0.1, ** p<0.05, *** p<0.01.

Table 8: Supply Side First Stage: Probability of Exit

	(1)	(2)	(3)	(4)
Panel A				
Probability of Exit: Coefficients				
Lagged Firm Market Share	-8.3255***	-8.0067***	-7.7000***	-7.3365***
	(1.3389)	(1.3443)	(1.3416)	(1.3448)
Demand-side Unobserved Quality		-0.0682**		-0.0700***
		(0.0270)		(0.0272)
Lagged Number of Firm by Market			0.0414**	0.0450**
			(0.0182)	(0.0184)
Lagged Number of Plans by Firm			0.0038	0.0034
			(0.0153)	(0.0153)
Market Size (in 1000)			-0.0013**	-0.0013+
			(0.0007)	(0.0007)
Panel B				
Probability of Exit: Average Marginal Effects				
Lagged Firm Market Share	-0.2732***	-0.2599***	-0.2502***	-0.2354***
	(0.0407)	(0.0404)	(0.0407)	(0.0405)
Demand-side Unobserved Quality		-0.0022**		-0.0022**
		(0.0009)		(0.0009)
Lagged Number of Firm by Market			0.0013**	0.0014**
			(0.0006)	(0.0006)
Lagged Number of Plans by Firm			0.0001	0.0001
			(0.0005)	(0.0005)
Market Size (in 1000)			-0.0000**	-0.0000+
			(0.0000)	(0.0000)
Characteristics	Yes	Yes	Yes	Yes
Region FE	No	No	Yes	Yes
Observations	7830	7830	7830	7830

Notes: All columns show estimates of the exit decisions for incumbents plans of the first-stage of the supply model. The dependent variable for all columns is a dummy variable that indicates exit. The sample includes all PDP Plans in Medicare Part D during 2007-2011. All columns include plan characteristics (Part D Drug Deductible, Extra Coverage in the Donut Hole, Enhanced Plan, Number of Drugs in Formulary, Number of Top Drugs in Formulary, Number of Top Drugs in Tiers 1-2, Size of Pharmacy Network, Preferred Pharmacies in Network). All columns also include dummies for all (34) Medicare regions fixed effects. *Panel A* presents the coefficients of the Probit Model and *Panel B* presents the Average Marginal Effects. Robust standard errors (clustered by market) in parentheses. + p<0.1, ** p<0.05, *** p<0.01.

Table 9: Supply Side First Stage: Probability of Entry

	(1)	(2)	(3)	(4)
Probability of Entry: Coefficients				
Lagged Number of Firms per Market	-0.1287***	-0.1270**	-0.1302**	-0.1276**
	(0.0494)	(0.0495)	(0.0509)	(0.0510)
Demand-side Unobserved Quality		-0.0296		-0.0290
		(0.0367)		(0.0369)
Market Size (in 1000)			-0.0016	-0.0015
			(0.0011)	(0.0011)
Lagged Number of Plans by Firm			0.0142	0.0156
			(0.0225)	(0.0226)
Probability of Entry: Average Marginal Effects				
Lagged Number of Firms per Market	-0.0420***	-0.0415**	-0.0425**	-0.0417**
	(0.0161)	(0.0162)	(0.0166)	(0.0166)
Demand-side Unobserved Quality		-0.0097		-0.0095
		(0.0120)		(0.0121)
Market Size (in 1000)			-0.0005	-0.0005
			(0.0004)	(0.0004)
Lagged Number of Plans by Firm			0.0046	0.0051
			(0.0073)	(0.0074)
Mean Characteristics	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	1700	1700	1700	1700

Notes: All columns show estimates of the entry decisions for plans. The depend variable for all columns is a dummy variable that indicates entry. Each period 10 firms are assumed to potentially enter the market. The sample includes these ten potential firms per each market-year and the associated mean characteristics of each market. The entry variable is constructed using the information about entry decisions for all PDP Plans in Medicare Part D during 2007-2012. All columns include mean characteristics for each market (Part D Drug Deductible, Extra Coverage in the Donut Hole, Enhanced Plan, Number of Drugs in Formulary, Number of Top Drugs in Formulary, Number of Top Drugs in Tiers 1-2, Size of Pharmacy Network, Preferred Pharmacies in Network). All columns also include dummies for all (34) Medicare regions fixed effects. *Panel A* presents the coefficients of the Probit Model and *Panel B* the Average Marginal Effects. Robust standard errors (clustered by market) in parentheses. + p<0.1, ** p<0.05, *** p<0.01.

Table 10: Estimates from Structural Model

	(1)	(2)
	Mean Coefficient	Standard Error
Marginal Cost		
Mean	1,079***	141
Std.Dev	367**	178
Exit Cost		
Mean	987,765+	548,766
Std.Dev	381,987	302,844
Entry Cost		
Mean	2,426,733***	376,673
Std.Dev	843,536**	409,287

Notes: This table shows the estimates of the structural parameters. To estimate the parameters of the marginal costs and exit costs, I follow the methodology of Bajari et al. (2007). I start with 500 initial values in the state space, which I take from the the observed values in the data (all PDP Plans in Medicare Part D during 2007-2012). For each of these states I forward simulate 100 periods following the actual and also 500 alternative optimal policies. To estimate the entry cost parameters I followed a minimum distance estimator minimizing the difference between the prediction of the model and the probability of entry of entry episodes in the data estimated using the policy policy function. Column (1) presents the estimated value of the parameter while Column (2) present the standard error of each point estimate. Standard errors (clustered by market history) in parentheses. + p<0.1, ** p<0.05, *** p<0.01.

Table 11: Simulated Counterfactuals: Welfare and Entry Results

	(1)	(2)	(3)
		Counterfactual Policies	
	Current Policy	Fixed Markup with Inertia	Fixed Markup without Inertia
Panel (A): Premiums			
Average Premiums	404	378	378
% Change		-7%	-7%
Panel (B): Entry, Exit and Distribution of Firms			
Number of firms	% of times in simulations		
0	1	3	2
1	16	18	8
2	23	34	24
3	28	21	31
4	19	15	21
5	13	9	17
% Periods with Entry	52	38	61
% Periods with Exit	40	31	42
Panel (C): Consumer Welfare			
Consumer Welfare	$360	$372	$394
% Change		3.1%	9.4%

Notes: This table present the results of the counterfactual analysis. In these counterfactual analysis, I solve the dynamic game for five single product firms in the market. I use the estimates of the demand side and supply side parameters previously shown for these counterfactuals. Column (1) present the Current Policy scenario, where the firms play the game they play in the actual competition in Medicare Part D. Column (2) present the counterfactual where the dynamic pricing is limited by a policy of fixed markups and consumer have inertia. Column (3) present the counterfactuals where dynamic pricing is limited by a policy of fixed markups but the market works without the inertia of consumers.

Appendix A Choice Probabilities in the Demand-Side Model

Since I estimate the model using data on market shares for each plan, I need to derive the market shares from my model. First, let's define the expected value of the utility-maximizing plan among all available plans, excluding the plan the individual had in the last period. Given the assumption of i.i.d. Type One Extreme Value distributed shocks, this expected value can be written as:

$$\delta^{imt}(j_{i,t-1}) = ln\Bigg(\sum_{j\in\mathbb{J}_{mt}, j\neq j_{i,t-1}} \exp\Big(\delta_j^{imt}\big(j_{i,t-1}, \Gamma(j_{it-1})\big)\Big)\Bigg)$$

where $\delta_j^{imt}(j_{i,t-1}, \Gamma(j_{it-1}))$ is the mean utility for changing to plan j. Note that this function depends on the function $\Gamma(j_{it-1})$ in the following way:

$$\delta_j^{imt}(j_{i,t-1}, \Gamma(j_{it-1}) = 1) = -\eta + \overline{f_{ijmt}}$$
$$\delta_j^{imt}(j_{i,t-1}, \Gamma(j_{it-1}) = 0) = \overline{f_{ijmt}}$$

where $\overline{f_{ijmt}}$ is the expectation of the flow utility (f_{ijmt}) with respect to the Type One Extreme Value shock.

Now, let's define the probability of choosing a particular plan. For simplicity of notation, let δ be the option to stay with the same plan the individual had in the last period. Assume that consumer i had chosen plan j' in the previous period; thus, because of the properties of the Type One Extreme Value distribution, her probability of switching is:

$$\begin{aligned} Pr^i_{\text{switch}}(j') &= Pr^i(j_{it} \neq j' | j_{it-1} = j') \\ &= \frac{exp(\delta^{imt}(j'))}{exp(\delta^{imt}(j')) + exp(\delta(j'))} \end{aligned}$$

Likewise, if consumer i chose plan j' in the previous period, her probability of not switching is:

$$\begin{aligned} Pr^i_{\text{notswitch}}(j') &= Pr^i(j_{it} = j' | j_{it-1} = j') \\ &= \frac{exp(\delta(j'))}{exp(\delta^{imt}(j')) + exp(\delta(j'))} \end{aligned}$$

Moreover, conditional on switching and given that j' was the plan chosen in the previous period,

the probability of choosing plan j is:

$$\begin{aligned} Pr^i_{j|\text{switch}}(j') &= Pr^i(j_{it} = j | j_{it-1} = j', j_{it} \neq j') \\ &= \frac{exp(\delta^{imt}_j(j'))}{exp(\delta^{imt}(j'))} \end{aligned}$$

Therefore, and using the Law of Iterated Expectations, the total probability of choosing plan j in period t having chosen plan j' in period $t-1$ is:

$$\begin{aligned} Pr^i_j(j') &= Pr^i(j_{it} | j_{it-1} = j') \\ &= 1\{j = j'\} Pr^i_{\text{notswitch}}(j') + 1\{j \neq j'\} Pr^i_{j|switch}(j') Pr^i_{\text{switch}}(j') \end{aligned}$$

These probabilities can be used to express the expected market share in the current period as a function of the last period's market share for a given consumer type. Let $s_{ij'mt-1}$ be the period $t-1$ market share for plan j' in market m for consumer type i. Then, the expected market share for a plan j for consumer type i is:

$$s_{i\hat{j}mt} = \sum_{j' \in \mathbb{J}_{mt-1}} s_{ij'mt-1} Pr^i_j(j')$$

Integrating the market shares over consumer types yields predicted market share for plan j:

$$s_{\hat{j}mt} = \int \hat{s}_{ijmt} dF_i$$

Appendix B Estimation of Structural Parameters

After the first step, which provides policy functions and therefore how the state vector evolves over time, the second step of the estimation finds parameters that make these observed policy functions optimal, given the underlying theoretical model. The parameters that I estimate in this stage (the structural parameters of this model) are the mean and standard deviation of the cost, entry and exit shocks. Given the per-period profit function in Equation 1, and integrating out the private shocks, I can re-write the expected value of the pay-off function as follows:

$$E_{\vec{\varepsilon}_{mt}}\Pi_{jmt}(\tilde{b}_{jmt}, \tilde{b}_{-jmt}, \vec{S}_{jmt}) = (\tilde{b}_{jmt} - c_{jmt})s_{jmt}M_{mt} + \Pr(\text{Exit} = 1|\vec{S}_{jmt})\phi_{jmt}$$

where the probability of exit replaces a dummy variable that indicates if the plan exits the market in that period. Also, the bids for the plan (b_{jmt}) and other plans in the market (b_{-jmt}) have been replaced with the associated equilibrium expected bids ($\tilde{b}_{jmt}$ and $\tilde{b}_{-jmt}$).

The computation of these objects via simulation of future shocks requires a high amount of computational power, in that it has to be solved every time in the search for the optimal value of the parameters. Therefore, it is important to reduce the computational burden using the insight proposed by Bajari et al. (2007). The idea is that, since all the unknown parameters enter linearly into the payoffs of the plan, the previous equation can be written as the inner product of a row vector and a column vector:

$$E_{\vec{\varepsilon}_{mt}}\Pi_{jmt}(\tilde{b}_{jmt}, \tilde{b}_{-jmt}, \vec{S}_{jmt}) = \begin{pmatrix} \tau & \psi(\vec{S}_{jmt}) \end{pmatrix} \begin{pmatrix} 1 \\ \theta \end{pmatrix}$$

where $\tau = \tilde{b}_{jmt}s_{jmt}M_{mt}$, θ is the vector of all the parameters to be estimated, and $\psi(\vec{S}_{jmt})$ is the vector of all the terms that depend on the parameters.

This allows me to compute the following function $Z(\cdot)$, which does not depend on the parameters and has to be computed only once, using forward simulation, in the estimation process:

$$Z(\vec{S}_{jmt}, \sigma_j, \sigma_{-j}) = E_{\sigma(\vec{S}_{mt}, \vec{\varepsilon}_{mt})} \sum_{t'=0}^{\infty} \beta^{t'} \begin{pmatrix} \tau & \psi(\vec{S}_{jmt}) \end{pmatrix}$$

where σ_j is a strategy vector for plan j, and σ_{-j} is a strategy vector for the rest of the plans in the

same market as plan j, as defined in Section 4.2.2.

Imposing the Markov Perfect Equilibrium (MPE) condition implies that each period the value function evaluated at the true optimal policy function (σ_j^*) should have a larger value than when it is evaluated at any other alternative policy ($\tilde{\sigma}_j^k$).

$$Z(\vec{S}_{jmt}, \sigma_j^*, \sigma_{-j}) \begin{pmatrix} 1 \\ \theta \end{pmatrix} \geq Z(\vec{S}_{jmt}, \tilde{\sigma}_j^k, \sigma_{-j}) \begin{pmatrix} 1 \\ \theta \end{pmatrix}$$

Now, for a given state vector $\vec{S}^r$, I can re-write the above equation in terms of profitable deviations from the optimal policy in the following way:

$$g(\vec{S}^r, \tilde{\sigma}_j^k, \theta) = \left[Z(\vec{S}^r, \tilde{\sigma}_j^k, \sigma_{-j}) - Z(\vec{S}^r, \sigma_j^*, \sigma_{-j})\right] \begin{pmatrix} 1 \\ \theta \end{pmatrix}$$

Now, I can construct the objective function to minimize in order to estimate the parameters. Note that the objective function is based on the condition of MPE. To estimate the parameters, I start with 500 initial values in the state space (denoted by the supraindex r in $\vec{S}^r$), which I take from the observed values in the data. Then, I forward-simulate the function $Z(\cdot)$, taking draws for the private shocks for the next 100 periods and using the actual optimal policy functions estimated in the first step of the supply estimation. I also compute 500 alternative policy functions (denoted by the supraindex k in $\tilde{\sigma}_j^k$) by changing one or more of the values of the parameters estimated for the state variables. With these draws of the private shocks, I compute the actions, the per-period payoffs and how the state space evolves for each of these 501 policy functions. The estimator then searches for parameters such that profitable deviations from the optimal policies are minimized:

$$min_\theta Q(\theta) = \frac{1}{K}\frac{1}{R}\sum_{k=1}^{K}\sum_{r=1}^{R} 1\{g(\vec{S}^r, \tilde{\sigma}_j^k, \theta) > 0\} g(\vec{S}^r, \tilde{\sigma}_j^k, \theta)^2$$

Note that the linearity of the unknown parameters become useful during the minimization, as I do not have to recomputed separate outcome paths for each set of parameters. The idea of the estimation is that, because of simulation error and errors in the policy function, in the estimation computation the value function evaluated at an alternative policy can be higher than the value at

the actual policy. The intuition of the estimation is to take these situations and choose the value of the parameters that minimize the square sum of these deviations. Finally, the standard errors can be computed doing bootstrap over the different market histories (Bajari et al., 2007; Ryan, 2012).

Having recovered the policy functions and the parameters needed for the construction of the payoffs for incumbents, it is now possible to find the distribution of the entry costs. Consider Equation 3. After the estimation of the distributions of marginal costs and scrap values, all the terms in that equation are known, except for the parameters of the distribution of κ_j. Assuming the entry costs are distributed normally with mean μ_κ and variance e_κ^2, the probability of a plan entering the market is the probability that κ_j is less than or equal to the expected value from being in the market:

$$\Pr(\text{Entry}|\vec{S}_{jmt}) = F_\kappa\left(\max_{b_{jmt}}\left\{E\left[(b_{jmt}-c_{jmt})s_{jmt}M_t + \beta E[W(\vec{S}_{jmt+1},\vec{\varepsilon}_{mt+1}|\vec{S}_{jmt})]\right]\right\}\right)$$

where F_κ is the CDF of the entry cost shock.

Note that the parameters of F_κ are unknown, but the term inside the CDF can be computed through forward simulation, starting from different draws of points in the state space:

$$\min_{\mu_\kappa, e_\kappa^2} = \frac{1}{R}\sum_{r=1}^{R}\left[\Pr(\text{Entry}|\vec{S}^r) - F_\kappa\left(\max_{b_{jmt}}\left\{E\left[(b_{jmt}-c_{jmt})s_{jmt}M_{mt} + \beta E[W(\vec{S}^r_{jmt+1},\vec{\varepsilon}_{t+1}|\vec{S}^r)]\right]\right\}\right)\right]^2$$

The intuition behind the estimation technique is to match as well as possible the observed probabilities of entry.