Market Design when Firms Interact with Inertial Consumers: Evidence from Medicare Part D*

Keith M. Marzilli Ericson†

December 22, 2011

Abstract

I use the Medicare Part D insurance market to examine market design when firms interact with inertial consumers. Enrollment data show enrollees face switching frictions leading to inertia in plan choice, and a regression discontinuity design indicates initial defaults have persistent effects. Theory predicts firms respond to inertia by raising prices on existing enrollees, while introducing cheaper alternative plans. The complete set of enrollment and price data from 2006 through 2010 confirms this prediction: older plans have approximately 10% higher premiums than comparable new plans. I then derive optimal dynamic (switching) defaults for individuals, which depend not only on whether inertia results from real switching costs or psychological factors that lead to inaction, but also on the equilibrium responses of firms. A default that switches individuals away from expensive plans can raise the elasticity of demand of existing enrollees and lower the equilibrium price differential between new and existing plans. I show conditions under which the switching default lowers overall switching costs borne and is socially optimal.

*I thank Martin Andersen, Raj Chetty, David Cutler, Stefano DellaVigna, Drew Fudenberg, Andreas Fuster, Larry Katz, David Laibson, Michael Grubb, Oliver Hart, Dan McFadden, Tom McGuire, and Amanda Starc for their thoughtful comments. I thank the Williams College Tyng Committee and the National Science Foundation for research support.

†Address: Boston Univeristy, 595 Commonwealth Ave., Boston, MA 02215. Email: kericson@bu.edu

1 Introduction

Market design decisions shape the functioning of many markets, from health insurance exchanges to electricity auctions, school choice systems, and labor clearinghouses.[1] Public policy often determines the form such markets take–the information available to market participants, the nature of contracts, the defaults individuals face, and the taxes or regulations facing consumers. This paper examines the consequences of market design decisions when firms strategically interact with inertial consumers. In many markets, individuals are subject to switching costs and other frictions that lead to inertia.[2] Rational firms respond to inertia when setting prices, initially pricing low to acquire market share and then raising prices on consumers when they are less responsive to price. However, market design decisions determine the form this response takes. For instance, introductory offers may be optimal for firms, but may be legally prohibited. Moreover, policies that alter the extent to which individuals are inert will change the prices that firms set. Thus, policy makers' choice of defaults may not only have a direct effect on individuals by modifying their switching behavior, but that choice will also have indirect effects through changes in the prices that individuals face.

I examine the consequences of market design decisions in the Medicare Part D prescription drug insurance market, a large and controversial program that receives government subsidies of about $40 billion annually and covers over 24 million people (Duggan, Healy, and Scott Morton 2008). Medicare Part D is the largest change to the Medicare program since its inception. Unlike Medicare's classic fee-for-service components, Medicare Part D established a marketplace in which firms compete to provide prescription drug insurance plans: a competitive heath insurance exchange. It is therefore a model for the insurance exchanges envisioned in the 2011 federal health reform. It began providing coverage in 2006, allowing us to see the market's first year and subsequent evolution. While program costs were initially below expectations, premium growth in recent years has outpaced growth in drug costs (Duggan and Scott Morton 2011). Strategic firm responses to inertia can explain this pattern.

I provide evidence that individuals display inertia in this market and are affected by program defaults. Firms respond to this situation when setting prices by initially offering plans at low prices to attract first-time enrollees. The data show that firms subsequently raise prices in later periods when their plans have a base of enrollees "stuck in place," while

[1] See Wilson (2002) on power auctions, Neal (2002) on school choice, and Roth (2002) on clearinghouses.

[2] Carroll et al. (2009) find that employees typically stay with arbitrary 401(k) savings defaults, but make substantially different decisions when forced to make an explicit choice for themselves. Jones (forthcoming) argues that inertia explains the pattern of over-withholding of income taxes. Chetty et al. (2011) examine labor supply elasticities, and show that observed responses match the pattern predicted by an adjustment cost model: larger tax changes lead to larger estimated elasticities.

new plans are introduced at low prices to attract new individuals entering the market. I examine the different defaults used in this market and derive conditions under which they are optimal.

Inertia in enrollees' choice of plan results from switching *frictions*, which include both switching costs and psychological factors that lead to inaction. Switching costs are the time and effort costs that result from moving between plans, e.g., setting up new paperwork or learning about new plans. Yet psychological factors can also lead individuals to fail to act, even if switching plans is not costly: for instance, inattention (Lacetera, Pope, and Sydnor, forthcoming), procrastination (O'Donoghue and Rabin 2001), and limited memory (Ericson 2011). Switching frictions lead enrollees to be less responsive to price once they have already enrolled in a plan. Existing literature shows that health insurance choices display inertia that can have substantial consequences. Handel (2009) examined insurance choice the year following a large price change and found that individuals may have forgone gains of over $1500 that year to stay in their current plan.[3]

Even though switching prescription drug coverage is arguably easier than switching an entire health insurance plan, changing plans may still be difficult if individuals find it costly to evaluate their options. Abaluck and Gruber (2011) argue that Medicare Part D enrollees have difficulty in making their initial plan choices, while Kling et al. (2009) show that enrollees may not be paying attention to their options in subsequent years. Switching is low in this market, which is consistent with either inertia or preference heterogeneity. While Medicare Part D enrollees have the opportunity to switch plans each year during open enrollment without regard to their health status, only about 10% of enrollees switched between 2006 and 2007 (Heiss, McFadden and Winter 2007). Yet at least some enrollees are attentive: Ketcham, Lucarelli, Miravete, and Roebuck (2010) found that the probability of an enrollee switching plans increased with their potential gain to doing so.

In the presence of inertia, random variations in initial conditions will have persistent effects. I first show suggestive evidence of inertia: higher prices in a plan's first year are associated with lower enrollment in subsequent years, even conditional on subsequent years' prices. I then use a regression discontinuity design in Medicare Part D's low-income subsidy (LIS) program to more credibly identify inertia. LIS recipients, which comprise about half the market, faced an automatic enrollment program set up by policy makers who were

[3] In addition, Samuelson and Zeckhauser (1988) discussed health insurance decisions as an example of status quo bias, though they recognized that inertia might be accounted for by classical explanations such as switching costs. Strombom, Buchmueller and Feldstein (2002) examine plan share sensitivity to health plan premiums at the University of California. They find that new hires have higher premium elasticities than incumbent employees, as predicted by models of inertia. This work has not examined firms' strategic responses to inertia in setting premiums, in part because it has typically examined employer-based health insurance, where firm behavior is constrained by an employer gatekeeper.

concerned they would fail to enroll. Individuals identified as being eligible for the subsidy were automatically defaulted into plans selected at random from the set of plans below a price benchmark. Because the precise level of the benchmark is unknown to firms in advance, a regression discontinuity design can estimate the causal effect of pricing below the benchmark. Pricing below the benchmark in the first year had a strong effect on enrollment: plans priced just below the benchmark had more than twice the market share of plans priced just above. Plans that randomly priced below the benchmark in the market's first year continued to have higher enrollment in later years, indicating that the LIS program's initial assignment of enrollees to plans had persistent effects on later choices.

A large theoretical literature examines the response of firms to switching costs (see Farrell and Klemperer 2007), and predicts a pattern of "bargains-then-ripoffs": products are offered at low prices and then subsequently at high prices. I extend these models to include psychological factors that lead to inaction, and I develop an equilibrium model of firm behavior in the presence of inertia that captures the features of the Medicare Part D market and other health insurance exchanges. The model predicts that inertia leads to a cyclical equilibrium in which plans are at first offered at low prices to attract individuals making initial decisions. Firms raise prices on those plans in subsequent years to take advantage of the lower price sensitivity of enrollees "stuck in place." New plans at low prices are introduced each period to attract individuals entering the market for the first time, as regulations do not allow firms to treat new enrollees differently from existing enrollees. Because some enrollees switch plans as a result of this pricing strategy and expend real resources to do so, the cyclical equilibrium is inefficient compared to a market in which firms could commit to future prices.

One firm provides a stark example of this strategy. In the first year of Medicare Part D, Humana priced its basic plans as loss leaders: about $10 per month on average, substantially below the market's average of $30. Both management and analysts agreed Humana was setting low prices to gain market share in the market's first year. Over the next three years, Humana raised its price on these plans by more than 40% each year, until by 2009 and 2010, the average price was over $40 per month, now above the market average.

While Humana is a particularly extreme example, pricing data confirms that the market as follows the pattern predicted by the model of cyclical equilibrium. I show that firms initially set relatively low prices for newly introduced plans, but then raise prices as plans age while new, low-cost plans are introduced each year. In a given year, plans that have existed for a longer period of time have annual premiums that are 10%, or $50, higher than newly introduced plans. The higher prices of existing firms suggest that many consumers either have switching costs of this amount or face other switching frictions (e.g. procrastination, forgetting) with costs in this range.[4]

I consider optimal dynamic (switching) defaults using my equilibrium model. Defaults determine what happens to individuals who take no action. Although individuals can typically easily opt out of defaults, evidence indicates that defaults can substantially affect individuals' outcomes (Madrian and Shea 2001; Choi, et al. 2004). Well-designed defaults then have the potential to improve welfare (Carroll et al. 2009). I consider the choice between reenrolling individuals in the same plan unless they actively choose to switch ("automatic reenrollment") and switching individuals to a cheaper plan unless they actively choose to stay ("automatic switching"). Automatic reenrollment is the most commonly used default and applies to standard enrollees in Medicare Part D, but LIS recipients face an automatic switching default.[5] The welfare consequences of defaults will depend on whether inertial behavior is a result of real switching costs or of psychological frictions that lead to inaction (e.g. forgetting to change plans). An automatic switching default will lead individuals who take no action to switch to cheaper plans and save premiums. This default can make them better off if they faced low switching costs, but would have failed to opt out of an automatic reenrollment default due to psychological frictions. If instead they face large switching costs but still fail to opt-out of the default, automatic switching can make them worse off.

Existing literature has not considered optimal defaults in contexts where firms strategically interact with individuals subject to the default. Because defaults affect individual behavior, they change the incentives facing firms and thereby alter firms' pricing strategy. Automatic switching can raise the elasticity of demand of existing enrollees and thereby lower the equilibrium price differential between new and existing plans. A lower equilibrium price differential can increase social welfare, as individuals not directly affected by the default switch less, reducing resources expended on switching costs. There is also a reduction in the transfer of resources away from individuals who do not switch plans, which may increase social welfare in the presence of distributional concerns for inattentive individuals. Against these gains are weighed the increased switching costs expended by individuals directly af-

[4] Optimization frictions of this magnitude have implications for what economists can learn from individuals' responses to changes in their environment. Chetty (2011) shows that in the presence of switching costs or other optimization frictions, a range of structural elasticities (i.e. long-run elasticities) is consistent with the observed response to a price change. For policy changes to the Medicare Part D market, such as increased subsidies for more generous coverage, the switching frictions found here would imply that the elasticities estimated from the stock of enrollees would be essentially uninformative about the true long-run elasticity.

[5] Unless they make an active choice, LIS recipients are automatically switched to a new plan if their plan prices above the benchmark in later years. Individuals may opt out of the default and stay with their current plan if switching is costly. When plan prices move from below to above the benchmark, at least half of LIS recipients do move to a new plan, suggesting that this default affects behavior. Yet a substantial fraction (one quarter to one half) of redefaulted LIS recipients make an active choice to stay in their initial plan even though they must pay additional premiums to do so.

fected by the default. I derive conditions under which the automatic switching default is socially optimal.

When firms respond to incentives created by defaults, defaults have externalities and the socially optimal default for the population may not coincide with the privately optimal default for an individual. For instance, automatic switching may be the socially optimal default because it lowers the equilibrium price differential between new and existing plans. Yet a given individual may prefer that an automatic reenrollment default applied to him or her alone, allowing the individual to save on switching costs and leaving others to discipline the market. Thus, having individuals choose their own defaults will not necessarily lead to the socially optimal default being chosen.

The paper is organized as follows. Section 2 describes the structure of the Medicare Part D market. Section 3 discusses the theory of firm pricing when individuals are subject to switching frictions and establishes the existence of the cyclical equilibrium. Section 4 describes the data used in the empirical portions of the paper. Section 5 uses a regression discontinuity design to test for inertia in the LIS program. Section 6 then tests the predictions of the theory for firm pricing. Section 7 discusses how to set optimal defaults when firms strategically interact with individuals subject to the default. Finally, Section 8 discusses the implications of the results and concludes.

2 Basic Structure of the Medicare Part D Market

2.1 Standalone PDPs

Medicare Part D began offering prescription drug insurance in 2006 for seniors over the age of 65 and other Medicare beneficiaries. I focus on the core portion of the program– standalone prescription drug plans (PDPs), which are distinct from other sources of coverage (e.g. Medicare Advantage HMOs or employer/union sponsored PDPs). As in other health insurance exchanges, there is a menu of plans available for purchase at listed prices. Firms must accept all individuals who choose a given plan at a fixed price: the premium enrollees pay does not vary by age or health status. There is free entry of firms, subject to regulatory approval, and many firms compete: from 2006-2010, 92 unique firms offered coverage.

Plan design is constrained by Medicare regulation. Each plan is required to offer at least "basic" coverage, as defined by the Centers for Medicare & Medicaid Services (CMS). Basic plans can come in three different forms: Basic Alternative, Actuarially Equivalent Standard, or Defined Standard Benefit. Each type of basic plan must offer coverage that is actuarially equivalent to the Defined Standard Benefit,[6] with a formulary that covers each

[6] Defined Standard plans have a fixed format: in 2010, the standard benefit has a $310 deductible, 25

therapeutic class of drug.[7] However, "enhanced" plans may offer coverage that is actuarially more generous (e.g. lower deductibles or coverage in the "doughnut hole"). I focus analyses on basic plans, which have fewer unobserved characteristics.

Contracts are annual, with firms committing to a price and formulary for that year.[8] Each year, firms simultaneously submit plan price bids. Then, during an open enrollment period (Nov. to Dec.), individuals observe the new prices and can switch plans. Standard enrollees must initially make an active choice to enroll in Medicare Part D. However, once they are enrolled, they stay with their current plan by default if they take no action. Pricing and plans offered vary by PDP region: each of the 34 PDP regions[9] is either a state or group of states (plus Washington D.C.), and I refer to these regions as "states" throughout.

The prices that enrollees face are a result of firm bids and government subsidies. The subsidies are designed so that enrollees pay the full marginal cost of a more expensive plan; an increase in a firm's bid translates one-for-one into an increase in enrollee premiums. For basic plans and standard enrollees, plan premiums are equal to the plan bid minus a fixed dollar subsidy, which is calculated by CMS based on the national average bid.[10] The payments firms actually receive are risk-adjusted and equal to their bid multiplied by adjustment factors for health risk, as described in Section 2.3. The risk adjustment system is designed so that firms should determine their bids based on the cost of providing coverage to an average individual in the population.

Firms might wish to continually introduce virtually identical cheap plans. However, there are both formal and informal restrictions that make this difficult. CMS requires that firms offering multiple plans demonstrate that there are significant differences among the plans; this regulation only formally applied beginning in 2009, but CMS negotiated with

percent coinsurance up to an initial coverage limit of $2,830 in total drug spending, a coverage gap (the "doughnut hole"), and catastrophic coverage when enrollee out-of-pocket spending exceeds $4,550. Actuarially Equivalent Standard plans have the same deductible, but may use copayments instead of coinsurance and tiered copayments for brand-name and specialty drugs. Basic Alternative allows plans to vary the amount of the deductible.

[7] Formulary variation may be a source of switching costs. Even if drugs within a therapeutic class are close substitutes, individuals face costs of changing their prescription.

[8] While firms can make mid-year changes to the formulary, they must be approved by CMS. Most changes are beneficial from the enrollees' perspective. Approved negative changes most often take the form of swapping a newly-available generic drug for the identical branded drug. See Levinson (2009).

[9] I limit the analysis to plans in 50 United States proper and exclude those in its territories and possessions.

[10] To calculate the subsidy, CMS calculates the national average bid $\bar{p}$. Each plan receives a fixed dollar subsidy, equal to $\frac{0.745-r}{1-r}\bar{p}$, where r is an adjustment factor for the cost of catastrophic reinsurance. The program costs for individuals without the LIS are subsidized 74.5% by the federal government. The premium subsidy is less than 25.5% of a plan's bid, since the government also subsidizes the plans by providing catastrophic reinsurance for expenses above a certain threshold. The next section describes the additional subsidy given to LIS recipients. Heiss, McFadden, and Winter (2007) provide more details on the bidding process and the subsidy calculation for enhanced plans.

firms to enforce this provision earlier. Moreover, for a firm to offer a plan, CMS must approve its bid submission. This bid is required to be tied to the firm's estimate of the revenue it needs to provide the benefit. Thus, firms may not wish to introduce variations in plan prices that they cannot plausibly link to variations in cost of benefit provision. CMS has been progressively increasing the standards that firm bids must meet (see Levinson 2008).

2.2 The Low-Income Subsidy (LIS) Program

Low-income subsidy recipients comprise a large share of the market (52% of PDP enrollees in 2006).[11] LIS recipients enroll in the same plans as standard enrollees, but receive additional premium subsidies and reduced cost-sharing. Medicare beneficiaries become eligible for at least a partial form of the LIS if their incomes are below 150 percent of the federal poverty level and pass an asset test; the exact amount of assistance varies with income and assets. Individuals receiving the full LIS benefit receive a premium subsidy equal to that of the LIS "benchmark" b in that state; if they choose a plan with a premium below the benchmark, they pay no premiums. In a plan with premiums of p, an LIS recipient thus pays $\max\{p-b,0\}$. The benchmark differs in each state and is recalculated each year based on the state's average plan bid; it is not known ex ante to firms.[12] In 2006, the average state's benchmark was about $32 per month.

The LIS program applies defaults in two ways: automatic initial enrollment and automatic switching. First, due to concern about inertia in enrollment behavior, individuals who meet certain eligibility criteria[13] for the full LIS are automatically enrolled into Medicare Part D. They are defaulted into a randomly selected basic PDP with a premium below the benchmark premium. LIS recipients may actively elect to choose another plan; they may do so at any time and are not limited to switching during the open enrollment period.

The mix of plans that price below the benchmark varies between years, as plans change their prices and the benchmark adjusts. The second default–"automatic switching"– is applied if a plan moves from being below the benchmark in one year to above the benchmark

[11] Many individuals not eligible for the LIS do not choose a standalone PDP, but instead choose Medicare Advantage HMOs with prescription drug coverage or receive an employer-sponsored plan.

[12] In 2006-2007, the benchmark was the average bid in that state, with PDPs equal weighted and Medicare Advantage prescription drug (MA-PD) portions enrollment weighted. In subsequent years, the benchmark transitioned to enrollment weighted PDP and MA-PD bids. Appendix Section A.4 gives more detail on the calculation of the benchmark and its evolution over time.

[13] Approximately 84% of LIS recipients in 2010 were deemed automatically eligible for the full LIS by their Medicaid, Supplemental Security Income (SSI), or Medicare Savings Program (MSP) status. Other potential LIS recipients must apply for the subsidy. CMS reserves the term "automatic enrollment" for Medicare and Medicaid dual-eligibles, and uses a similar "facilitated enrollment" process for individuals who were not dual-eligible but otherwise deemed eligible for the full LIS. Since the processes are virtually identical, I use the term "automatically eligible" to refer to both groups.

in the next. If an auto-enrolled LIS recipient in such a plan had never made an active choice, they are automatically switched to a different plan below the price benchmark, unless they take action to stay in their current plan. LIS recipients who actively enrolled themselves, or who were auto-enrolled but then chose to move from their default plan, are notified that they will pay a higher premium if they do not switch, but they are not re-defaulted into a new plan.

Concerned with the difficulties of switching LIS recipients away from plans that previously priced below the benchmark, CMS instituted a "de minimis" policy for LIS recipients for 2007 and 2008. De minimis plans were those whose premium exceeded the benchmark by less than \$2 (2007) or \$1 (2008) per month. Under the policy, LIS enrollees in de minimis plans would not be automatically switched by default. However, no new LIS enrollees would be defaulted into such plans, and de minimis plans would not receive any additional premiums over the benchmark amount from any of their LIS recipients.[14] While this policy reduced the need to switch LIS recipients between plans, it also had the effect of making LIS recipients less profitable for firms, as they could yield \$12-\$24 less per year in revenue than a standard enrollee.

2.3 Risk Adjustment

Because premiums are community-rated (all enrollees pay the same price) and guaranteed-issue (plans must take all comers), a risk adjustment scheme was designed to reduce the incentives for firms to select a healthier or lower-cost population. Firms receive higher payments from CMS for enrollees with higher expected costs, with payments determined by enrollees' risk adjustment factors. These factors are based on demographic characteristics and diagnostic history, with additional adjustment made for low-income subsidy status and institutionalization status. For more detail, see Robst, Levy and Ingber (2007) and Appendix Section A.4.

Effective risk adjustment implies that as an enrollee ages, they do not become more costly to a firm. Evidence indicates that this is indeed the case. Risk adjustment is based on diagnostic history, but when that information is unavailable, a simpler model based on age and sex is used. This simple model is effective in accounting for how costs rise with age. Appendix Section A.4 shows that as the population ages by five years, risk adjusted payments to firms rise by 3.1%. This is roughly consistent with data from the Medical Expenditure Panel Survey, which shows that the population's average prescription drug spending would rise by about 2.6% in the same time.

[14] That is, all LIS recipients who were eligible for the full subsidy. Partial subsidy recipients were not automatically enrolled or switched, and so the de minimis policy did not apply to them.

However, risk adjustment for LIS recipients is insufficient. When designing the risk adjustment scheme, CMS had limited data on the relative costs of LIS recipients. Hsu et al. (2010) show that while CMS risk adjustment scheme assumes that full-subsidy LIS recipients are only 8% more expensive than comparable standard enrollees, they are in fact 21% more expensive. LIS recipients are therefore less profitable for firms than standard enrollees.

3 Theory: Inertia and Firm Responses

3.1 Introduction

If consumers display inertia in their health insurance choices, firms will rationally respond. In setting prices, firms have two motives: an investment motive, to acquire market share for the future, and a harvesting motive, to maximize profits this period on new and existing customers. Farrell and Klemperer (2007) review the theoretical literature on how inertia affects equilibrium under imperfect competition. In a variety of contexts, it finds a "bargains-then-ripoffs" pattern, in which products are initially sold at low (perhaps below marginal) cost, but sold at higher prices in later periods.

I adapt the insights of these models to the Medicare Part D context and additionally consider the effects of psychological frictions and defaults. I model individual behavior as subject to both classical switching costs and psychological factors that lead to inaction: these two sources of inertia differ in their implications for welfare and the effect of defaults. I then model the incentives facing firms when setting prices and show that a plan's price will depend on whether a plan is newly introduced and has no attached consumers, or if it has a customer base "stuck in place". If inertia leads the demand of existing enrollees to be more inelastic, as suggested by evidence in the next sections, then firms should optimally raise price on existing plans.[15]

Finally, I examine equilibrium in the limiting case of perfect competition with overlapping generations of consumers. New plans enter each period offering low prices, as they invest in future market share. Existing plans with market share have higher prices to extract money from consumers stuck in place. The equilibrium is similar to that in Farrell and Shapiro (1988), who model a duopoly with overlapping generations and perfect substitutability between goods: they find an "alternating equilibrium" in which firms cycle between selling to new consumers only or selling to old consumers only. However, the Medicare Part D market allows for free entry, and the bargains-then-ripoffs pricing pattern provides an

[15] The demand curve faced by a plan depends on its past market share, individuals' preferences, and the probability individuals will switch for a given gain. Without further assumptions, a plan's previous market share can have an ambiguous effect on optimal price.

entry motive for new firms or new plans from existing firms.

3.2 Modeling Individuals: Switching Frictions

Switching frictions lead to inertia in individuals' choice of plan. I model two classes of switching frictions: 1) real switching costs that result from moving between plans and reduce welfare, and 2) psychological frictions that affect whether an individual acts, but not their welfare conditional on the action taken. For instance, when an individual switches plans, they need to learn the rules of the new insurance plan, may need to do paperwork at their pharmacy, and may experience disutility from negative emotions (e.g. confusion, loss aversion)– these are real switching costs that reduce welfare. On the other hand, an individual may wish to switch plans but forget (Ericson 2011) or procrastinate (O'Donoghue and Rabin 2001)– these are psychological frictions that lead them not to act and simply take the default option. Evidence suggests that both classes of switching frictions affect behavior in many contexts.

Both types of switching frictions result in similar individual behavior and induce similar responses by firms, so it is difficult to distinguish them using data from the Medicare Part D market. However, they differ in the welfare implications of defaults, and can be distinguished by giving sophisticated individuals their choice of default.[16] In the current Medicare Part D market, the government sets the defaults: for standard enrollees, the default option is to stay in their current plan, while LIS recipients are switched by default if their plan becomes too expensive. Section 7 shows that the optimal default will depend on the source of switching frictions, highlighting the importance of research quantifying the sources of switching frictions.

To capture real switching costs , I assume that every period, individuals each draw a real switching cost ω_{it} that must be paid if and only if the individual changes insurance plans, where ω_{it} is drawn i.i.d. from the cumulative distribution function $G(\omega)$.[17] Individuals bear these costs regardless of whether the switching results from their choice, or from them being switched by default. For instance, regardless of how they are switched between plans, enrollees must learn their new plan rules and set up new billing information at a pharmacy; such real switching costs are likely to be even larger for full service health insurance plans, as prescription drug plan enrollees can switch plans without switching doctors, but many

[16] However, Section 7 shows that because defaults have externalities via firm pricing behavior, the socially optimal default does not necessarily coincide with the default individuals would choose for themselves.

[17] The i.i.d. assumption implies that there are no persistent heterogeneity in individual propensities to switch. This assumption substantially simplifies the calculation of equilibrium, but could be relaxed. In the presence of persistent heterogeneity in switching costs, firms would set price taking into account that the mix of individuals that would enroll is endogenous to the price.

health insurance plans have limited provider networks.

However, psychological frictions can lead individuals to fail to act, even though switching plans would not be costly for them (Carroll et al. 2009). When individuals fail to act, the default option determines their outcome. I model these psychological frictions as a tolerance for inaction, and assume that the probability a psychological friction will lead an individual to take the default is decreasing in the gain to action. Hence, I assume that each period an individual has maximum tolerable loss from taking the default λ_{it}, where λ_{it} is an i.i.d. draw λ_{it} from the cumulative distribution function $H(\lambda)$. Thus, making switching the default option would lead more people to switch, even if opting out of that default was costless, because psychological frictions sometimes lead people to take the default when they would gain by switching. For instance, they may forget to send back the appropriate form. Individuals who face no psychological frictions may be affected by defaults through classical channels: individuals may bear a real resource "opt-out" cost if they do not take the default (i.e. the cost of sending back a form). This cost is likely to be small relative to the other real switching costs and psychological frictions, and so for simplicity, the main paper assumes this cost to be zero; Appendix Section A.2.3 gives a full treatment of positive opt-out costs.

Resulting behavior is as follows. Individuals seek to maximize their discounted expected utility over their lifetime. I assume linear utility for money in the region of premiums, an approximation that is reasonable given the range of premiums at stake. Individuals perceive a gain ΔU in lifetime utility from switching plans, from which is subtracted switching costs ω_{it}. In the baseline model, I assume individuals are sophisticated about future firm behavior and their own switching frictions and so correctly forecast ΔU. Under the automatic reenrollment default faced by standard enrollees, an individual switches if $\Delta U - \omega_{it} > \lambda_{it}$. The net gain of switching is the utility from the better plan choice minus switching costs and opt-out costs. The individual only switches if the gain to doing so is greater λ_{it}, the maximum tolerable loss to staying with the default. Under an automatic switching default, the individual switches more often for a given gain: whenever $\Delta U - \omega_{it} > -\lambda_{it}$, tolerating a loss up to λ_{it} from staying with the default and switching.

When setting prices, firms care only how individuals behave, not the source of the switching frictions. Individual behavior can be summarized as follows: the probability an individual switches for a gain of ΔU under the automatic reenrollment default is given by the summary distribution $F(\Delta U) = \int_0^\infty H(\Delta U - \omega)\, dG(\omega)$, where F is a c.d.f. that is continuous, differentiable, and bounded with derivative $f(\cdot)$. I use this summary function F when describing the firms' decision, and distinguish between the sources of switching frictions in Section 7's examination of optimal defaults.

3.3 A General Model of Firm Price Setting

I model insurer behavior in the Medicare Part D market, which is regulated as described in Section 2. Insurers must issue a policy to anyone who requests it, and must charge all enrollees the same price for a given plan. Risk adjustment implies that individuals do not vary in cost by age.[18] I make the simplifying assumption that the form of the insurance contract (e.g. copays, drugs covered) is fixed, which is a good approximation to government regulation of basic plans. Keeping with the way Medicare Part D and other insurance markets are regulated, firms offer policies for one period, without the possibility for commitment to future premium levels.

Each firm j offers one plan,[19] and sets its price p_{jt} in each period. Quantity sold this period s_{jt} is a function of this price and past market share.[20] The expected cost of each enrollee, net of risk adjustment, to the firm is c_j. Firms are infinitely lived with discount factor δ, and seek to maximize the expected discounted present value of profits V_{jt}. The value of the firm V_{jt} is given by flow profits and future profits in the recursive equation:

$$V_{jt} = (p_{jt} - c_j)\, s_{jt} + \delta V_{jt+1}\,(s_{jt})$$

where the second term captures that future firm value may depend on its current market share.[21] The firm's first order condition for optimal pricing is thus:

$$p_{jt} - c_j = \frac{s_{jt}}{-ds_{jt}/dp_{jt}} - \delta \frac{dV_{jt+1}\,(s_{jt})}{ds_{jt}} \tag{1}$$

where ds_{jt}/dp_{jt} is the firm's demand curve. Factors that make demand more inelastic, such as switching frictions, raise markups. The demand curve ds_{jt}/dp_{jt} that a firm j faces when setting prices is the sum of the demand curves for three different types of individuals: 1) potential repeat customers, 2) potential switchers from other plans, and 3) new enrollees

[18] Even if risk adjustment were imperfect and older enrollees cost more than existing enrollees, in the absence of switching frictions, firms that have existed longer should not disproportionately attract older individuals. Section 3.4 shows that in a competitive market, imperfect risk adjustment does not lead to the cyclical equilibrium without switching costs. Section 2.3 shows that risk adjustment based on age seems accurate.

[19] While firms may offer more than one plan so long as they are sufficiently distinct, for simplicity, I examine the case where one plan only is offered.

[20] The demand of sophisticated consumers for a plan will depend on both its price and its market share, as market share may predict firm's future behavior. In this discussion, I ignore this effect, which is equivalent to assuming individuals cannot observe market share or are myopic. The equilibrium model in Section 3.4 allows for sophisticated consumers.

[21] This model could be generalized in a number of ways. Switching costs or attachment to the firm could depend on the length of time an enrollee has been in a plan. Furthermore, type of consumer might matter: older individuals may be less valuable since they will not live as long.

entering the market unattached to any plan.

In this general model, switching frictions and previous market share s_{jt-1} can have an ambiguous effect on optimal prices, depending on the relative elasticities of these three groups. However, it is likely that potential repeat customers have relatively inelastic demand, compared to the other groups, since new choosers and potential switchers can choose from many close substitutes. In such a case, older plans will face more inelastic demand and optimally set prices higher than comparable newer plans. Indeed, the next section examines the limiting case when plans are perfect substitutes. Consistent with the predictions of other models of equilibrium under imperfect competition (Farrell 1986; Farrell and Klemperer 2007), it shows that new entrants will have lower prices than comparable existing plans.

3.4 Cyclical Equilibrium Results From Inertia

I now consider how inertia affects firm behavior in equilibrium by examining the limiting case of perfect competition: in the Medicare Part D market, many firms are competing to offer very similar products. In the model, when individuals initially enter the market, all products are perfect substitutes and individuals simply choose the cheapest plan. Hence, new firms without a customer base face a perfectly elastic demand curve. In later periods, switching frictions give a plan market power over enrollees that previously chose it, and the market transitions away from perfect competition for existing firms and enrollees.

In the presence of inertia, firms will have an incentive to raise prices on existing plans. Yet because firms are involved in an infinitely-repeated game, many possible collusive equilibria may exist. I consider a simple Markov-perfect equilibrium in which firms' prices will depend only on whether its plan is newly introduced or existed in the past.[22] In this equilibrium, new plans enter the market each period and offer prices below marginal cost to attract new enrollees. In later periods, these plans raise prices on enrollees stuck in place.

I assume each individual must purchase exactly one insurance plan in every period[23] and that plans are identical in all aspects except price. Utility-maximizing individuals therefore seek to minimize their discounted expected premiums paid and switching costs borne, subject to the switching frictions they face. As discussed in Section 3.2, the probability an individual switches for a utility gain of ΔU is given by $F(\Delta U)$. In the baseline model, I assume individuals are sophisticated about future firm behavior and their own switching frictions and so correctly forecast the lifetime utility consequences of switching plans.

[22] In a simpler model in which the market ends after two periods, an equilibrium similar to that described below is the unique equilibrium.

[23] Nothing in the equilibrium would qualitatively change if individuals had the option to opt-out of the market if the cost of the plan exceeded their reservation price. For simplicity, I eliminate this decision from the model.

There is a continuum of individuals, normalized to measure one, with a constant hazard $\rho \in (0,1)$ of dying each period. Thus, fraction $1-\rho$ of the population survives from the last period. Each period, measure ρ of new individuals that are unattached to any plan enter the population, and so population size remains constant. Individuals discount future utility by $\delta < 1$ each period, in addition to the discounting that results from the probability of death.

Firms are infinitely lived with discount factor $\delta < 1$, and seek to maximize the present discounted value of profits. The marginal cost to the firm of an enrollee is their expected spending net of risk adjustment, which I assume is a constant c. Firms compete via Bertrand competition on premiums only. Keeping with the structure of the Medicare Part D market, firms do not have the ability to commit to future prices. Each firm receives equal share of all unattached consumers who choose a plan with that premium, and keeps its existing enrollees if they do not die or switch plans. Firms can only offer one plan at a time. Each period, $N \geq 2$ firms have the opportunity to enter the market; they do so with no previous market share. Bertrand competition implies that the market is perfectly competitive for new firms, and so firms in the first period compete away the profits they will later make on enrollees "stuck in place".

Proposition 1 shows that a simple pure-strategy Markov-perfect equilibrium (Maskin and Tirole 2001) exists in which a firm's strategy depends solely on whether it is a newly introduced plan or a continuing plan that has enrollees "stuck in place" who are attached to the plan.[24] The core prediction of this model is that new plans charge lower prices than existing plans. The difference in price, Δp, between newly introduced plans and continuing plans is determined by the elasticity of repeat demand. Distributions of switching frictions F that lead to more inelastic demand of stuck-in-place enrollees lead to higher price differentials.[25] (In contrast, when there are no switching frictions, all plans with positive enrollment charge the same price, regardless of plan age, as individuals would simply choose the cheapest plan each period.[26])

Proposition 1. *A pure-strategy Markov-perfect equilibrium exists and takes the following form. New firms ($N \geq 2$) enter each period and all set price p_L. Plans that continue from the previous period with stuck-in-place enrollees charge higher premiums $p_H > p_L$. Define the enrollment of new firms as $s_0 = \frac{1}{N}\left[\rho + (1-\rho) F(p_H - p_L)\right]$. Then, prices are given by $p_L = c - \delta \frac{V((1-\rho)s_0)}{s_0}$, and $p_H(s) = c + \frac{1-F(p_H-p_L)}{f(p_H-p_L)} - \delta(1-\rho) V'(s)$. The value of a firm*

[24] Depending on the distribution F, there may be multiple equilibria having the specified form. At least one such equilibrium exists.

[25] The intuition that uniformly increasing switching frictions for all individuals should increase Δp is incorrect: whether Δp increases or decreases depends on the elasticity of the switching function. As in other monopoly price setting contexts, a uniform shift in willingness-to-pay does not always lead to more inelastic demand.

[26] This holds even if risk adjustment were incomplete and older individuals were more costly to the firm.

with measure s of enrollees is $V(s) = s\frac{[1-F(p_H-p_L)]^2}{f(p_H-p_L)}$. *The price differential between new and continuing plans is* $(p_H - p_L) = \frac{1-F(p_H-p_L)}{f(p_H-p_L)}$.

Compared to a situation in which firms could commit to future prices or simply charged the same price each period (lifetime average cost), this equilibrium is inefficient: switching uses real resources, and switching is higher without commitment. These results also suggest other potential inefficiencies. Because switching is higher, firms and individuals may have reduced incentives to invest in relationship-specific investments (e.g. insurer investments in enrollees' future health, or enrollee investments in learning their plan structure).[27]

The existence of this cyclical equilibrium is robust to various assumptions regarding the sophistication of individuals and their ability to predict future firm pricing.[28] Here, sophisticated individuals are able to fully predict the path of firm prices over time. They choose the cheapest plan when they enter, correctly predicting that its price will increase in the future but taking advantage of the low price in the present period. In later periods, sophisticates will switch if the price differential Δp is greater than their switching friction, correctly anticipating that the gain from switching is a one time event: in the future they will pay p_H every period until they switch again. However, the same form of equilibrium results if instead consumers are myopic and incorrectly believe that firms will maintain their current prices in all future periods. Myopes will choose the cheapest plan available when they enter the market, incorrectly believing price will remain constant at p_L in future periods. In later periods they are surprised when their plan charges p_H and will wish to switch plans. Because they are myopic, they will overestimate the benefits of switching and may switch too often. However, the probability a myopic individual switches will still be described by some function that increases in the difference between the price of their current plan and that of the cheapest available plan, which is all that is necessary for the proof of Proposition 1.

[27] Proposition 1 describes an equilibrium in which competition implies that firms do not make excess profits as a result of inertia, even if individuals are myopic. For models of imperfect competition, there is an active debate about whether switching costs raise or lower the average level of markups: compare Farrell and Klemperer (2007) and Dubé, Hitsch & Rossi (2009), who find that the effect of switching costs on average markups are non-monotonic and depend on the setting. Markups are transfers from enrollees to firms and so affect the distribution of income. Higher markups would also lead to added deadweight loss for the increased taxes to pay for higher premiums (consumers only pay about 25% of the premiums), and from individuals substituting out of the market.

[28] Moreover, although there are many possible collusive equilibria, myopes are drawn to the firms that follow the loss leader strategy, since myopes believe initial low prices will persist. Thus, so long as there is a positive measure of myopes in the population, there is no Nash equilibrium in which all types of firms charge a constant price each period (see Appendix Section A.2).

4 Describing the Medicare Part D Market

4.1 Data Source

Data from the Medicare Part D market show both that individuals display inertia and that firm prices display the pattern predicted by the model above. I use data from CMS on plan premiums, characteristics, and aggregate enrollment. Data on PDP premiums and characteristics for each year are available from 2006 (the first year of the market) through 2010. I divide the 2,464 plans into cohorts based on the year they were first offered. Enrollment data is available for July 1 of each calendar year from the monthly enrollment reports. The Data Appendix provides more details.

For each plan, I observe its premium, deductible, and benefit type,[29] along with the firm and plan name. Table 1 gives descriptive statistics of the Medicare Part D plans, by year of plan introduction (cohort). States vary in the number of plans offered and average premiums. Moreover, a given firm may price essentially the same plan quite differently in different states. For example, in 2006 Humana offered the "Humana PDP Complete" plan for $767 per year in Ohio and only $575 in New York.

There is substantial variation in premiums, even for basic plans. Figure 1 shows the distribution of premiums in 2010 for basic plans, split between older cohorts of plans (plans introduced in 2006 and 2007) and newer cohorts (plans introduced 2008 and later). Though the peaks of the distributions are similar (around $400/year), the older cohorts have a larger tail of high premium plans, consistent with the predictions of Section 3 that plans raise premiums as they age. However, the variance in prices indicates that there is heterogeneity in firm strategies or costs. Variation in pricing can come from firm-specific costs of providing coverage, price strategies (e.g. firm estimates of demand elasticity, or whether firms recognize the investment value of acquiring market share), and perceived quality of firms (firm-specific demand shocks).

New plans come from one of three sources: existing firms offering sufficiently distinct plans, existing firms expanding into different geographical regions, or new firms entering the market. Table 1 indicates that for the first five years of the market, it was primarily existing firms expanding in both ways. Most new plans were offered by firms who already offered plans somewhere else in the country, while about two-thirds were introduced by firms already offering a plan in the same state.

The number of individuals choosing plans for the first time was largest in 2006, since this was the first year Medicare Part D offered coverage, and the stock of all people eligible for Medicare could choose in that year. The initial enrollment period ended May 15, 2006,

[29] Basic alternative, actuarially equivalent standard, defined standard benefit, or enhanced.

after which individuals faced a late enrollment penalty fee if they did not have a qualifying form of prescription drug coverage. Immediate enrollment was optimal for most seniors, and most seniors did in fact enroll: by May 2006, Medicare had met its target that 90% of the eligible population have some form of prescription drug coverage (Heiss, McFadden, and Winter 2007). In subsequent years, new entrants to the PDP market come from individuals newly eligible for Medicare and from individuals leaving another source of coverage (e.g. Medicare Advantage plans).

Figure 2 shows total enrollment over time, broken down by plan cohort (the year in which a plan was introduced). The 2006 cohort of plans captured most of the market, as most of the inflow into the PDP market took place in the market's first year; inertia implies that enrollees are likely to stay with their initial plan. This cohort has an aggregate enrollment[30] of 15.4 million in 2006, a number that drops over time, as enrollees leave these plans (by death or switching) or as plans attrit from the sample. Subsequent cohorts of plans have much lower enrollment, consistent with the predictions of the model in Section 3: there are fewer new enrollees after the first year of the market.[31] After 2006, the number of new choosers is small relative to the size of market: I estimate that newly eligible individuals comprise less than 10% of new PDP enrollees in each year.[32]

I examine the behavior of standard choosers (non-LIS enrollees) separately from that of LIS recipients, since LIS recipients face different prices and are not necessarily making an active choice even when they first enroll. I subtract estimates of LIS enrollment from total enrollment to get estimates of standard enrollment.[33] I construct plan market shares of total enrollment in each state, and then market shares of standard enrollees: a plan's non-LIS enrollment over the state's total non-LIS enrollment. Plan shares of total enrollment in 2006 range from less than $\frac{1}{1000}$% to 38%; the median plan share is 0.4%. The median plan's share of standard enrollment is also 0.4%. Appendix Figures A.1 and A.2 plot LIS enrollment and standard enrollment by cohort of plan. The fraction of enrollees receiving the LIS among the 2006 cohort is initially high (52%), but falls to 41% by 2009. Newer plans have a higher

[30] These numbers differ from the aggregate numbers released by Medicare by about 1 million, as my numbers exclude Employer/Union Only Direct Contract PDPs and PDP enrollment outside the 50 U.S. states.

[31] Other factors could also contribute to the observed pattern of lower enrollment in subsequent cohorts. For instance, fewer plans are introduced in later years. Yet this is unlikely to explain the full story: the number of plans introduced in 2007 was over half the number introduced in 2006, but the 2007 cohort's enrollment is substantially below half of that in the 2006 cohort.

[32] From 2007 to 2010, about 2 million Americans turned 65 each year and become eligible for Medicare; less than half of them chose a standalone PDP.

[33] Since CMS has not released LIS enrollment figures regularly, I have LIS enrollment data from July of 2006 and 2007, but from February of 2008 and 2009; they were unavailable for 2010. Hence, these data slightly underestimate the share of LIS enrollees in later years. The Data Appendix gives more details.

fraction of LIS enrollment in 2009 (70% to 89% depending on cohort), which is expected, since new plans have lower prices.

4.2 Correlation between Enrollment and Past Prices

I begin with standard enrollees and provide suggestive evidence that this half of the market displays inertial behavior. Using aggregate enrollment data, I test whether past prices predict market share (conditional on present prices and characteristics). I estimate regressions of the following form:

$$\ln s_{jtm} = x_{jtm}\beta_1 + \alpha_1 p_{jtm} + x_{jt-1m}\beta_2 + \alpha_2 p_{jt-1m} + v_{tm}$$

where $\ln s_{jtm}$ is plan j's log market share in market m at time t, p_{jtm} is the plan's premium, and x_{jtm} contains its observed characteristics. State fixed effects v_{tm} capture factors that vary among states, including the number of plans offered. Of course, firms set prices endogenously to unobserved quality, with the expectation of price increasing in quality in most models. If firm price-setting is subject to random noise (e.g. information shocks), then even conditional on present prices, the expectation of quality should increase in lagged price p_{jt-1m}, giving $\alpha_2 > 0$ in the absence of inertia.[34] Inertia predicts that $\alpha_2 < 0$: higher past prices induce lower enrollment which persists into later periods.

I estimate this regression using standard (non-LIS) enrollment,[35] limiting the sample to basic plans: these plans offer similar actuarial value and have little flexibility in plan design, reducing unobserved heterogeneity.[36] I run regressions both with and without firm fixed effects. Each specification is useful: using variation in pricing among firms is valuable because such variation may be less endogenous to market conditions (e.g. if firms are subject to information shocks), but controlling for firm fixed effects reduces unobserved heterogeneity.

Table 2 examines the association between 2007 enrollment and 2006 prices for the cohort of plans introduced in 2006. It shows that past prices strongly and negatively predict enrollment. Column 1 regresses 2007 log plan shares on 2006 and 2007 prices. It finds that premiums in 2006 still predict enrollment in 2007, with a coefficient on past premiums nearly as large as that on current premiums. Column 2 runs the "naive" regression of 2007 log plan shares on 2007 prices only and shows that the coefficient on 2007 premiums is 50% larger in magnitude when lagged prices are omitted, due to the correlation of past and present prices.

[34] Other models of unobserved heterogeneity can lead to biases in either direction; hence this evidence is only suggestive.

[35] LIS recipients face different defaults and prices. I include controls for whether the plan is below the benchmark to capture any effect of the LIS program on the plan.

[36] Ideally, I would like to separate out new enrollees from existing enrollees, but this is not possible using aggregate data.

For comparison, column 3 examines initial choices in 2006, regressing log plan shares on price for the same sample. The coefficient on contemporaneous price is larger in magnitude for the first year of the market (column 3) than for 2007 (column 1): premiums that are \$1 higher are predict a plan share that is 14% lower in 2006, compared to 9.7% lower in 2007. Columns 4-6 present analogous regressions with firm fixed effects included and show that the results are similar.

The association between enrollment and past prices is a robust phenomenon. Similar regressions for 2009 data shows that even three years later, premium in 2006 is still negatively associated with enrollment (Appendix Table A.1). Moreover, in 2009, there is a series of previous prices that can be included as controls. Of all the past prices, the 2006 premiums should have the largest effect, since that was when the largest cohort of individuals made its initial choices. Indeed, Appendix Table A.2 shows that premiums in the year of introduction have the largest association with enrollment when all the lags of premiums and plan characteristics are included.

5 Low-Income Subsidy: Defaults and Inertia

5.1 Regression Discontinuity Design

While the above analysis suggests standard enrollees display inertia, this section provides more precisely identified evidence on inertia from the other half of the market: LIS recipients. The LIS program only automatically enrolls individuals into plans that set their price below a price benchmark. Because the benchmark is not known ex ante, but is a random variable, firms cannot precisely choose whether to set prices above or below the benchmark. Hence, a regression discontinuity strategy can identify the causal effect of being randomly assigned LIS enrollees. I compare the subsequent enrollment and pricing strategies of plans that randomly fell just above the benchmark in 2006 to those that fell just below. The identification assumption is that pricing directly above or below the benchmark is as good as random, so that plan characteristics do not change discontinuously around the benchmark.

The regression discontinuity approach is particularly credible in 2006, as it was the first year of the Medicare Part D market. Because the benchmark in 2006 is an equal-weighted average of PDP bids in each state, even a large number of firms colluding could not precisely predict the benchmark level. Define the variable "relative premiums" to be a plan's premiums minus that state's benchmark level; this is the forcing variable. Appendix Table A.3 supports the identification assumption that there are no discontinuous changes in covariates at the benchmark. The observed characteristics of PDPs (type of basic plan, and deductible level) are similar on either side of the benchmark for the bandwidths used here, though in some bandwidths, the mix of basic plans differs slightly. I show regressions with and without controls for these characteristics; results are similar.[37]

Plans attrit from the sample overtime. Attrition can occur because firms cease offering a plan, or if they merge with or are acquired by another firm. Attrition, of course, has no effect on the estimates of 2006 enrollment, but may affect estimates of enrollment and price responses in subsequent years. Attrition between 2006 and 2007 is negligible: Appendix Table A.4 shows that less than 5% of plans attrit by 2007 in the regression discontinuity windows used here. Attrition by 2008 is similarly small. Yet by 2009 and 2010, more than 20% of plans in the regression discontinuity windows have attrited, and plans that price below the benchmark in 2006 are more likely to attrit. I present estimates for 2009 and 2010, but they should be viewed as conditional on remaining in the data.

5.2 Effect of Pricing Below Benchmark on Enrollment

Figure 3 confirms that pricing below the benchmark leads to a substantial increase in enrollment. This figure plots 2006 premiums relative to the LIS subsidy amount against 2006 log enrollment share, and plots predicted enrollment, controlling for premiums relative to the benchmark in linear and quartic polynomial specifications. The first two panels in Table 3 confirm the visual effect: Panel 1 shows a regression that controls for relative premiums linearly, while Panel 2 uses a quadratic polynomial of relative premiums, plus plan characteristic controls. The Imbens and Kalyanaraman (2009) optimal bandwidth for log plan shares is approximately $4,[38] but the effect is robust to the use of other bandwidths. Regardless of specification, the coefficient in column 1 for being below the benchmark indicates that pricing just below the benchmark leads to market shares that are approximately 200 log-points (150%) higher than other plans. Average plan market shares in the $4 window above the benchmark are just under 1%, while below the benchmark the average is about 5.5%. A placebo test using only the enrollment of non-LIS individuals finds effects that are small in magnitude and not significantly different than zero, supporting the identification

[37] Although it is not necessarily for the validity of the design, McCrary (2008) suggests testing for discontinuities in the density of the forcing variable. A discontinuous density at the cutoff may suggest firms were able to manipulate whether they are above or below the benchmark. In the absence of collusion with CMS, this seems implausible. Applying the test suggests there may be a discontinuity in the density at the cutoff, but these seems to be a result of the density not being smooth in general. Appendix Figure A.3 graphically displays the result of the density discontinuity test at the cutoff, which finds a log difference in density height at the cutoff of 0.317 (standard error 0.14), giving a t-statistic of 2.21. Yet rather than firms sorting around the cutoff, further tests suggest the density is not smooth: testing for discontinuities at one dollar intervals around the cutoff gives t-statistics above 1.6 at four of ten locations. Appendix Figure A.4 displays the histogram of relative premiums and shows that there are spikes at a number of points in the histogram, including one near zero.

[38] The optimal bandwidth varies slightly by year; I use a consistent cutoff for each year.

strategy: the benchmark does not appear to affect non-LIS enrollment.

These initial defaults have a persistent effect on subsequent enrollment. Additional columns in Table 3 show that pricing below the benchmark in 2006 predicts enrollment not only in 2006, but in later years as well: plans below the benchmark in 2006 have market shares that are 130 log points higher in 2007. The effect decays over time, but is still substantial in 2008. Appendix Figure A.5 shows this visually for 2008 enrollment. For 2009 and 2010, the local linear regressions indicate a large effect, but not the polynomial regressions. The estimated effect on enrollment in these later years is conditional on not attriting from the data.

The persistent effect of random variation in initial conditions comes from two sources: plans that continue to price below the benchmark hold on to the enrollees they have acquired by default, and individuals make active choices to stay with plans that subsequently price above the benchmark. Panel 3 of Table 3 regresses log plan shares in each year on indicators for being a benchmark plan in 2006 interacted with being a benchmark or de minimis plan in the current year. Focus on 2007, in which the three indicator variables control for each possible history of pricing below the benchmark: below the benchmark in both years, below in 2006 only, or below in 2007 only, compared to never having been a benchmark plan. The first row indicates that plans that priced below the benchmark both years had market shares that were 209 log points higher than plans that were below the benchmark in neither year. The coefficient in third row shows that pricing below the benchmark in 2007 alone leads to market shares that were only 15 log points higher than plans never below the benchmark. Comparing these two coefficients shows that pricing below the benchmark has a larger effect on enrollment if the plan was previously a benchmark plan, as such plans keep their previously acquired LIS recipients by default.[39]

Thus, inertia in LIS enrollment comes both from the effect of defaults as well as from active choices to avoid switching costs. Being below the benchmark in 2006 is associated with higher enrollment in 2007 even if the plan is not a benchmark plan in 2007: such plans have market shares that are 62 log points higher than plans that were never below the benchmark. These estimates indicate that about a quarter to one half of LIS recipients chose to stay with their plan even after it priced above the benchmark.

5.3 Effect of Pricing Below Benchmark on Subsequent Pricing

Firms that receive LIS recipients have a relatively larger base of existing enrollees, which may affect firms' pricing in later periods. However, LIS recipients behave differently from

[39] Appendix Table A.5 shows that results are similar if controls are included, including premium in the current year.

standard enrollees, as they are automatically switched if the plan raises its price over the benchmark. Because they face different defaults and prices, Appendix Section A.2.1 shows that the effect of acquiring LIS recipients on a plan's future prices is theoretically ambiguous.

To examine whether falling above or below the benchmark in 2006 had any effect on average premiums in the subsequent year, Figure 4 plots monthly relative premiums in 2007 against relative premiums in 2006 (horizontal axis). In contrast to the enrollment results, visual inspection indicates no obvious discontinuity in average firm behavior above or below the cutoff. This is confirmed in Appendix Table A.6, which finds that being below the benchmark had an insignificant effect on 2007 pricing using a bandwidth of $6 (approximately the optimal bandwidth for premiums). Similarly, for later years (2008 - 2010) the effect is noisily estimated, never significantly different from zero. The sign of the point estimate is not stable across years or specifications.

Even if acquiring LIS recipients did not have an effect on average prices, the desire to hold on to auto-enrollees could create an incentive to keep prices below the benchmark or the de minimis amount in subsequent years. Average prices could remain the same, even as firms were more likely to price below the benchmark. Yet Appendix Figure A.6 shows that plans that were below the benchmark in 2006 are no more likely to be below the benchmark or to be a de minimis plan in 2007. The absence of an effect is confirmed by Appendix Table A.7, which shows that the point estimate is insignificant and in fact negative in most specifications: the point estimates indicate plans are slightly less likely to fall below the benchmark in subsequent years if they did so in the first year, with the local linear regressions indicating a 6 percentage point decrease. Thus the evidence suggests little effect on firm pricing behavior of having acquired LIS recipients.

6 Cyclical Equilibrium Observed in Firm Pricing Behavior

The core prediction of switching frictions for firm behavior is that older plans should charge higher prices. Figure 5 confirms graphically the prediction of the cyclical equilibrium. It plots the average premium charged by each basic plan in each year, separating out plans by cohort. As predicted, we see that premiums in each cohort rise over time. Plans are introduced each year, with new plans generally having lower premiums than existing plans. The pattern is not perfect, as premiums for the 2006 cohort declined slightly from 2006 to 2007; afterwards, the 2006 and 2007 cohorts appear to act similarly. CMS, along with other commentators, noted the drop in premiums from 2006 to 2007 and suggested it was the result of lower than expected prescription drug costs, more substitution into generic drugs than anticipated, and higher than expected competitive pressures. It is likely that substantial firm learning occurred between 2006 and 2007.

Recall that Figure 1 compares the distribution of basic plan premiums in year 2010, for the earlier cohorts (2006 and 2007) and later cohorts (2008+) of plans. It shows that the higher premiums of older cohorts is not due to a few outliers, but to the behavior of many plans. Moreover, in addition to having a higher mean, the 2010 distribution of premiums in the older cohorts is more right skewed than the distribution of premiums for the newer cohorts. This suggests heterogeneity in the extent to which firms are raising prices on existing plans.

Table 4 regresses log premiums on plan age with various controls for observable plan characteristics.[40] It includes year fixed effects (interacted with state fixed effects) in all specifications and so identifies the effect of plan age on price by comparing plans of different ages in a given year. The regressions cluster standard errors at the firm level, to account for the fact that premiums are serially correlated at the plan level and to allow for the possibility that plans offered by the same firm experience common shocks. These analyses show that the observed association between plan age and premiums is not merely due to changes in composition of plans toward cheaper plan types.

Column 1 gives the association between plan age and premiums, confirming the visual results of Figure 5 among basic plans when controlling for state by year fixed effects. Older plans have higher premiums than new plans, about 6% higher in their fourth year and 18% higher in their fifth year.[41] Column 2 adds controls for the form of the basic benefit type, interacted with year fixed effects. These regressions indicate that plans in their fourth year cost 12% more than comparable newly introduced plans, while five-year-old plans cost 15% more than comparable new plans. This column also includes an indicator for whether the firm offering the plan also offered a Medicare Advantage (M.A.) plan, as firms may strategically attract Part D enrollees in an attempt to also enroll them in a Medicare Advantage plan.[42] Firms that offer a Medicare Advantage plan are cheaper, by about 15% per year, suggesting firms may be using PDPs as loss leaders. The cyclical equilibrium of Section 3 can result from both new firms entering and existing firms introducing new plans. Column 3 includes firm fixed effects and identifies the effect of plan age on price using variation within firms over

[40] Individual plan fixed effects regressions are not estimated due to the well-known inability to separately identify cohort, age (i.e. year of plan existence), and year fixed effects.

[41] Taken literally, the model in Section 3.4 predicts that firms simply raise price once to a new high level. The empirical results show that the increase is more gradual. This may result from a number of factors. Sharp raises may draw unwelcome publicity and attention from policy makers; Humana was criticized for its extreme strategy. Switching costs may also develop over time: if a person joins a plan in November and has the opportunity to switch beginning in January, he may not have learned enough about his current plan to make learning about another plan more costly. Finally, firms may experiment over time to find the optimal price.

[42] Because this variable is collinear with firm, it must be dropped in regressions that use firm fixed effects.

time. The pattern persists, indicating that the observed effect is not due to new firms entering at lower prices but not raising them; the pattern persists even controlling for firm quality. Although I do not observe the detailed fomulary characteristics of each plan, controlling for firm fixed effects should remove most of the variation in plan formularies.

While the regressions in Columns 1-3 equally weight all plans and therefore describe the experience of the average plan, enrollment-weighted regressions provide a better description of the experience of the average enrollee. Columns 4-6 weight each plan observation by its total enrollment in that year. The estimated effect of offering a Medicare Advantage plan shrinks, but the age effects become somewhat larger in magnitude when regressions are enrollment-weighted. Compared to new plans, premiums are statistically significantly higher for all plans at least three years old. Because the strategy of introducing plans at lower prices is successful at attracting higher enrollment, the price increase experienced by the average enrollee is larger than the average plan's price increase.

The enrollment-weighted regressions also indicate that the results are not being driven by attrition of plans from the sample. Plans can leave the sample either because the firm discontinues the plan or because the plan is merged with another plan (e.g. when firms merge). Relatively few plans are discontinued (less than 8%; see Appendix Table A.8). Dropping such plans from the regressions does not affect the results. An additional 28% of plans leave the sample because they merge with another plan. The new, larger plan receives additional weight in the enrollment-weighted regressions. These regressions indicate the age effect remains robust.

Appendix Table A.9 shows that these results are robust to a number of changes in the regression specification. When regressions include firm interacted with year fixed effects, they identify the effect of age on pricing using variation in a given year at a particular firm. Similarly large effects of plan age on pricing are found using equally weighted regressions. An enrollment-weighted regression find noisy to zero effects within firms, suggesting that larger firms do not vary their prices within a given year based on plan age, consistent with the potential regulatory constraints described in Section 4. The age effect also persists when enhanced plans are included in the sample: the percentage increase with age is larger, albeit measured with more noise. (Recall, we do not capture all the features of enhanced plans). Finally, when the dependent variable is the absolute premium in dollars rather than logs, the results are similar and show that plans that are five years older cost about \$50 more than comparable newly introduced plans.

These results from the Medicare Part D market show the pattern of firm pricing predicted by the model in Section 3. Plans in their fifth year charge an additional 10%, or about \$50, per year than equivalent, newly introduced plans. Although we do not know the distribution of switching frictions, these results are quantitatively consistent with the model as well: it seems reasonable that seniors may not switch for gains as small as \$50.[43]

The model of Section 3 predicts that firms are sophisticated and vary prices in response to variation in the price elasticity of demand they face. An alternative explanation (not supported by the data) supposes that firms price on lagged average costs, and that older plans have enrollees who are older and more costly. Then, older cohorts of plans would charge more. However, Section 2.3 showed Medicare Part D's risk-adjustment scheme implies that even though older individuals will have more drug spending, they will not be more costly to firms. Thus, age-related costs do not account for the observed pattern of firm pricing.

Nor is the observed pattern of older plans due to plans charging more because their LIS recipients are more costly. The results in Section 5 indicate that there is no consistent effect of acquiring a large number of LIS recipients on subsequent premiums; the preferred specification finds a negligible negative effect (about 68 cents per month). Risk-adjustment for LIS recipients is insufficient to cover their higher costs (Hsu et al. 2010). This could contribute to an incentive to raise premiums among plans that disproportionately attract LIS recipients. Yet it is new cohorts of plans that have a higher fraction of their enrollees receive the LIS, a result of their lower prices. In 2009, 40% of enrollees in the 2006 cohort of plans receive the LIS, compared to 70-89% of later cohorts (Appendix Figures A.1 and A.2). Hence, incomplete risk adjustment for LIS recipients implies that the estimated effects of plan age actually underestimate the increases in prices that would occur if risk adjustment were perfect. Thus, the data suggest that firms are relatively sophisticated in setting prices, and that the patterns of increasing prices lead to higher profits in later periods than earlier periods.

7 Optimal Defaults When Firms Respond to Inertia

Firms respond to the level of inertia that individuals display. Thus, defaults that change individuals' switching behavior not only have a direct effect on individuals' outcomes, but an indirect effect: defaults can alter the pricing strategy that firms use. This section considers how to set optimal defaults when firms strategically interact with individuals subject to the default. Previous research (Carroll et al. 2009) has considered optimal defaults only in cases where firms are not strategic actors (e.g. a benevolent employer). I examine dynamic defaults for enrollees who are already enrolled in a plan and who now face a new open enrollment period in which firms have changed their prices. I consider the government's choice between two defaults: if enrollees take no action, should they stay with their current plan ("automatic

[43] This is a pure utility gain if plans have perfect substitutes; to the extent plans are imperfect substitutes, the utility gain from switching would be attenuated.

reenrollment"), or be switched to the most inexpensive plan ("automatic switching")?[44] Standard enrollees in the Medicare Part D market face the automatic reenrollment default, but LIS recipients face automatic switching. The analysis below shows that the source of inertia (switching costs v. psychological frictions) matters for setting optimal defaults. In contrast to Section 3, where firms simply cared about *whether* individuals switch plans if they raised their prices, setting optimal defaults requires understanding *why* individuals behave the way they do.

I make some simplifying assumptions for clarity in the following discussion. As in Section 3.4, products are homogenous. Recall that individuals face real switching costs ω_{it}, and psychological frictions λ_{it}. I assume λ takes on only two values: $\lambda_{it} = 0$ with probability $1 - \psi$, and $\lambda_{it} = \bar{\lambda}$ with probability ψ; this assumption is not crucial, and Appendix Section A.2.3 allows lets λ follow an arbitrary distribution.[45] As in the baseline model, I continue to let individuals be sophisticated about future firm strategy and let ω_{it} be distributed according to c.d.f. G, which I assume is continuous, differentiable, and bounded with derivative g.

Thus, with probability $1 - \psi$ individuals are "attentive" and not affected by defaults: they simply switch if the premium savings outweighs the switching cost: $\Delta p - \omega_{it} > 0$. But when $\lambda_{it} = \bar{\lambda}$, the individual stays with the default unless the gain to making an active decision exceeds $\bar{\lambda}$. I call these individuals "inattentive," though they could be forgetful, etc. Under an automatic reenrollment default, an inattentive individual switches only if $\Delta p > \omega_{it} + \bar{\lambda}$. I assume $\bar{\lambda} \geq \Delta p$ in all cases below, so that inattentive individuals do not switch under an automatic reenrollment default. In contrast, under an automatic switching default, the individual switches if $\Delta p > \omega_{it} - \bar{\lambda}$. This is because an inattentive individual will switch even if the cost ω_{it} exceeds the gain Δp, so long as it does not exceed the gain by more than $\bar{\lambda}$. Throughout this section, I assume that firms play the equilibrium strategies described in Proposition 1, so Δp gives the equilibrium difference in price between continuing plans and the inexpensive new plans.

The privately optimal default for an individual may differ from the socially optimal

[44] Other defaults are also possible, such as probabilistic defaults or non-participation defaults. An automatic switching default that is applied only if Δp is greater than some threshold can improve on the all-or-nothing automatic switching default. However, such a default would essentially replicate price regulation: instead of a fine, the punishment firms face if they do not set the government's selected price would be to have their enrollees defaulted into another plan. I focus on the choice between the simpler automatic reenrollment and automatic switching defaults, which are both actually used in Medicare Part D.

[45] Appendix Section A.2.3 also examines the case in which there are real resource costs of opting-out of the default (e.g. sending back a form announcing your preference) that are distinct from the costs switching plans. The results are similar, with real opt-out costs creating an additional motivation to choose a default that matches the modal switching behavior of the population. Similarly, the amount of time an individual spends thinking about whether to switch (a real cost) may vary depending on which default applies. The framework in Section A.2.3 can also be applied to that case.

default. When choosing a default for herself, an individual weighs the change in premiums paid against the change in switching costs borne.[46] I assume the social welfare function attaches equal welfare weights to all individuals, so that premiums are simply transfers from one individual to another and the socially optimal default minimizes total switching costs borne. Distributional concerns for inattentive individuals could be added to the model and would place additional weight on the privately optimal default for such individuals. Proposition 2 below considers when automatic switching versus automatic reenrollment would be the optimal default for an individual or small group (formally, the case where the default affects a measure-zero subset of the population.)

Proposition 2. *Suppose firms play the strategies in Proposition 1. The privately optimal default for a measure-zero subset of the population is automatic switching if* $\int_0^{\bar{\lambda}+\Delta p} \omega dG(\omega) < \Delta p \cdot G\left(\bar{\lambda}+\Delta p\right)$, *and is otherwise automatic reenrollment.*

The intuition behind Proposition 2 is as follows: the default only matters for individuals when they are inattentive, so ψ does not enter the expression. Individuals compare the expected additional switching costs borne under automatic switching to the premiums saved. Under automatic switching, inattentive individuals switch whenever the switching cost is below the threshold $\bar{\lambda}+\Delta p$, which occurs with probability $G\left(\bar{\lambda}+\Delta p\right)$. They save Δp in premiums when they switch, but expected switching costs are the integral of ω up to that threshold. If the price differential is small but $\bar{\lambda}$ is high, then automatic switching will not typically be optimal, since inattentive individuals would bear large switching costs (up to $\bar{\lambda}$) for little gain. Conversely, when Δp is large because individuals are inattentive, but individuals do not face many real switching costs, automatic switching will be optimal.

When defaults are chosen for the entire population of enrollees, the response of firms to the default must be considered. Moving from an automatic reenrollment default to an automatic switching default alters the elasticity of demand of existing enrollees, and so the equilibrium price differential between new and continuing plans Δp will differ. Let this differential take the value Δp_{Sw} under an automatic switching default and Δp_{Re} under an automatic reenrollment default. Proposition 3 shows that whether automatic switching or automatic reenrollment is optimal will depend on the difference between Δp_{Sw} and Δp_{Re}.

Proposition 3. *Suppose firms play the strategies in Proposition 1. The socially optimal default is automatic switching if* $\psi \int_0^{\bar{\lambda}+\Delta p_{Sw}} \omega dG(\omega) < (1-\psi) \int_{\Delta p_{Sw}}^{\Delta p_{Re}} \omega dG(\omega)$, *and is otherwise automatic reenrollment.*

[46] Individuals seek to maximize their lifetime utility. When firms follow the strategy in Proposition 1, relative prices and expected switching costs paid are constant in each period after an enrollee enters the market. Thus, the proof of Proposition 2 shows that it is sufficient to examine the tradeoff between premiums saved and switching costs in a given period.

Proposition 3 shows that the socially optimal default compares the two effects of automatic switching. First, automatic switching increases the probability the ψ inattentive individuals will switch, increasing their switching costs borne. This increases the elasticity of demand firms face[47] and so lowers the equilibrium price differential between new and existing plans: $\Delta p_{Sw} < \Delta p_{Re}$. As a result, we have our second effect: the $1-\psi$ attentive individuals are less likely to switch under an automatic switching default. Social welfare counts as a gain the reduction in switching costs borne by attentive individuals who draw ω in the region between Δp_{Sw} and Δp_{Re}. Automatic reenrollment may be optimal if inattentive individuals bear large switching costs but automatic switching has only a small effect on prices. Conversely, automatic switching is optimal if inattentive individuals drive firms' prices but do not bear large switching costs.

Defaults have externalities, as illustrated by the difference between the optimal default from a given individual's perspective and the optimal default for the entire population. In some cases, automatic reenrollment may be optimal from the social perspective (i.e. it leads to lower switching costs), but a given individual may prefer that he or she (alone) faced an automatic switching default that leads to savings in premiums paid. In other cases, automatic switching may be the optimal population default because it raises the elasticity of demand and leads to lower price differentials; nonetheless, a given individual may prefer that his or her own default was automatic reenrollment to save his or her own switching costs, leaving other people to discipline the market.

This analysis shows a new consideration that must be taken into account when setting defaults: the effect defaults have on firms. While defaults are not the only tool governments have at their disposal to change firm behavior (e.g. firms may be regulated or taxed), market designers typically must either choose a default or allow decentralized choice of default; the results of the latter may differ from the socially optimal default. Although automatic reenrollment is a commonly used default, automatic switching may be the privately optimal default (if premium savings outweigh switching costs) and/or the socially optimal default (when it lowers switching costs by altering firm behavior).

To determine the socially optimal default, information on the source of switching frictions is needed. The two classes of switching frictions can be distinguished in a number of ways. First, researchers can examine switching behaviors in contexts where individuals make an active decision (Carroll et al. 2009); in such cases, psychological frictions such as memory, procrastination, and inattention are not likely to play a large role. Second, they can be distinguished by giving individuals a choice between different defaults: if individuals

[47] In this simple setting in which λ_{it} is either 0 or $\bar{\lambda}$, it is always the case that $\Delta p_{Sw} < \Delta p_{Re}$. However, for some distributions of λ, $\Delta p_{Sw} > \Delta p_{Re}$. See below for more detail.

choose an automatic reenrollment default despite large price differences, that indicates they perceive real switching costs to be high (assuming that they are sophisticated about the psychological frictions they face).

This model could be extended in a number of ways. For this simple two-point distribution of λ, automatic switching always leads to a lower Δp, but Appendix Section A.2.3 shows that for some distributions of λ, automatic switching may actually raise Δp by causing the firm to lose its relatively price elastic customers, leaving it with more inelastic demand. More complex switching defaults could be considered, in which the probability a person is automatically switched increases in Δp, or people are only switched if Δp exceeds some threshold. However, setting such defaults effectively requires the government to have even more detailed knowledge of the shape of the switching friction distributions. This context considered homogenous products, but unobservable quality differences will reduce the desirability of automatic switching. Moreover, if products are heterogeneous, automatic switching is less likely to be optimal, as it would disrupt the match between person and a product. However, automatic switching could be applied within a product category, by switching them to the cheapest product in that category.

8 Discussion and Conclusion

Inertia and firms' responses to it have implications for researchers and policy makers. Since firms predictably raise prices on plans in later years, analysis of this market should consider the lifecycle price of an insurance product. Total premiums paid will depend on an enrollee's ability and willingness to switch plans. Enrollees who switch to inexpensive plans will effectively receive transfers from enrollees stuck in place at relatively expensive plans, which may raise equity concerns; automatic switching defaults can reduce these transfers.

Inertia limits how enrollees will respond to changes in their environment, and so enrollees who face switching frictions will respond to a policy change differently than individuals making initial decisions. Even moderate switching frictions can limit what can be learned about long-run population responses from existing enrollees. The results in Table 4 suggest an approximate magnitude of switching frictions: \$50, or about 10% of annual premiums.

Chetty (2011) shows that in the presence of switching costs or other optimization frictions, a range of structural elasticities (i.e. long-run elasticities) is consistent with the observed response to a price change. Consider a hypothetical large policy change that puts a 50% subsidy on premiums paid, replacing the current arrangement in which individuals pay the full marginal cost of choosing a more expensive plan. How would this subsidy affect total expenditure on premiums? Suppose a researcher examined existing enrollees and precisely identified the change in their premium spending that resulted from the policy, estimating a price elasticity of spending of -0.07 (similar to that measured in other contexts.)[48] Appendix Section A.5 uses the results of Chetty (2011) to show that with switching frictions of 10% of premiums, an observed elasticity of -0.07 would be consistent with long-run elasticities that range from virtually zero (-9.0×10^{-4}) to very large (-5.0), a rather uninformative range.

The LIS program was one of the first major attempts to use defaults applied to individual behavior to alter firm incentives. Evaluating the optimality of this default requires knowing the relative contribution of real switching costs and psychological frictions to inertia, as well as the counterfactual firm pricing that would have occurred if LIS recipients were automatically reenrolled in the same plan. Whether the automatic switching default was privately optimal for a given LIS recipient requires comparing the premiums saved versus switching costs borne. Proposition 2 shows that whether automatic switching is privately optimal depends on Δp. When Δp is very small, the savings from switching is small. Hence, the de minimis policy was likely privately optimal from a given LIS recipient's perspective, since it automatically reenrolled LIS recipients in their current plan when the price differential is less than \$1-\$2 per month; real switching costs paid by inattentive individuals could easily exceed that amount. Yet in later years, some plans that initially received LIS auto-enrollees would have cost (net of subsidy) over \$200 per year, while alternative free plans were available. Automatic switching in such cases is likely to be privately optimal.

Contract restrictions play a major role in determining the form equilibrium takes. Under current regulations, plans must charge all enrollees the same price. If firms were instead allowed to charge "introductory prices" for first-time enrollees, they would choose to do so (see Taylor 2003). Such a policy would still lead to inefficient switching between plans. However, it would weaken incentives for firm entry, since existing firms could simultaneously offer attractive prices to new enrollees while charging enrollees stuck in place a higher price.[49]

The current contracting structure makes it difficult for firms to commit to future prices, but commitment to future prices (e.g. by allowing multi-year bids) could reduce inefficient switching.[50] Some rough calculations give a sense of the potential welfare gain to flat pricing. Heiss, McFadden and Winter (2007) find that 10% of enrollees switch between 2006 and

[48] The response of interest here is the percentage change in total spending for a percentage change in price. This differs from the plan share elasticity estimated in Section 4.2, which measures substitutability among plans. Gruber and Washington (2005) observe an elasticity of total premiums spent on health insurance of about -0.07 for employer provided health insurance.

[49] Similarly, if firms were allowed to offer multiple, identical plans at different prices, they would desire to do so, as this would essentially replicate introductory pricing.

[50] Firms submit annual bids. Because final prices are determined by a subsidy amount that is unknown to firms when submitting their bid, firms cannot easily communicate future pricing intentions to enrollees (e.g. a firm cannot advertise that their plan will cost \$30 month for the next five years). Other barriers to commitment include uncertainty about future costs and inability to commit to unobserved quality.

2007. We do not know how much of this switching is induced by price changes, as opposed to consumer learning or preference change. Suppose that only half of the observed switching would have occurred if firms had set constant prices, so there are about 0.8 million excess switches per year. If the average switching cost borne, conditional on actually switching, is $25 (recall, switchers can save about $50), then about $20 million per year in real costs are expended on switching that would not have occurred if firms committed to constant prices.

Medicare Part D is a large, functioning exchange that is important to study in its own right, and also gives insights into the design of other health insurance exchanges. Yet firms' strategic responses to inertia are relevant for market design in domains other than health insurance: for instance, governments organize school voucher programs and private social security accounts. Choice of defaults and contracting constraints should take into account the inertial behavior of individuals, real switching costs individuals face, and the strategic responses of firms to both.

REFERENCES

Abaluck, Jason and Jonathan Gruber. "Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review*, 101 (2011), 1180-1210.

Carroll, Gabriel, James Choi, David Laibson, Brigitte Madrian, and Andrew Metrick. "Optimal Defaults and Active Decisions." *Quarterly Journal of Economics* 124, (2009),1639-1674.

Chetty, Raj. "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." NBER Working Paper 15616, 2011.

Chetty, Raj, John Friedman, Tore Olsen, and Luigi Pistaferri. "Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records." *Quarterly Journal of Economics* 126 (2011), 749-804.

Choi, James, David Laibson, Brigitte Madrian, and Andrew Metrick. "For Better or For Worse: Default Effects and 401(k) Savings Behavior." David Wise, editor. *Perspectives in the Economics of Aging*. Chicago, IL: University of Chicago Press, (2004), 81-121.

Dubé, Jean-Pierre, Günter Hitsch, and Peter Rossi. "Do Switching Costs Make Markets Less Competitive?." *Journal of Marketing Research* 46 (2009), 435-445.

Duggan, Mark, Patrick Healy, and Fiona Scott Morton. "Providing Prescription Drug Coverage to the Elderly: America's Experiment with Medicare Part D." *Journal of Economic Perspectives* 22 (2008), 69-92.

Duggan, Mark, and Fiona Scott Morton. "The Medium-Term Impact of Medicare Part D on Pharmaceutical Prices." *American Economic Review*103 (2011):387-392.

Ericson, Keith M. "Forgetting We Forget: Overconfidence and Memory." *Journal of the European Economic Association.* 9 (2011), 43-60.

Farrell, Joseph. "A Note on Inertia in Market Share." *Economics Letters* 21 (1986), 77-79.

Farrell, Joseph, and Klemperer, Paul. "Coordination and Lock-In: Competition with Switching Costs and Network Effects." *Handbook of Industrial Organization*, Elsevier 3 (2007), 1967-2072.

Farrell, Joseph, and Carl Shapiro. "Dynamic Competition with Switching Costs." *RAND Journal of Economics* 19 (1988), 123-137.

Gruber, Jonathan, and Ebonya Washington. "Subsidies to Employee Health Insurance Premiums and the Health Insurance Market." *Journal of Health Economics* 24 (2005), 253-276.

Handel, Benjamin. "Adverse Selection and Switching Costs in Health Insurance Markets: When Nudging Hurts." Working Paper, 2009.

Heiss, Florian, Daniel McFadden, and Joachim Winter. "Mind the Gap! Consumer Perceptions and Choices of Medicare Part D Prescription Drug Plans." NBER Working Paper 13627, 2007.

Hoadley, Jack, Juliette Cubanski, Elizabeth Hargrave, Laura Summer, and Tricia Neuman. "Medicare Part D Spotlight: Part D Plan Availability in 2010 and Key Changes Since 2006." Kaiser Family Foundation Issue Brief 7986, 2009.

Hsu, John, Vicki Fung, Jie Huang, Mary Price, Richard Brand, Rita Hui, Bruce Fireman, William Dow, John Bertko, and Joseph Newhouse. "Fixing Flaws In Medicare Drug Coverage That Prompt Insurers To Avoid Low-Income Patients." *Health Affairs* 29 (2010), 2335-2343

Imbens, Guido and Karthik Kalyanaraman. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." Working Paper, 2009.

Jones, Damon. Forthcoming. "Inertia and Overwithholding: Explaining the Prevalence of Income Tax Refunds." American Economic Journal: Economic Policy.

Ketcham, Jonathan, Claudio Lucarelli, Eugenio Miravete, and M. Christopher Roebuck. "Sinking, Swimming, or Learning to Swim in Medicare Part D." Working Paper, 2010.

Kling, Jeffrey R., Sendhil Mullainathan, Eldar Shafir, Lee Vermeulen, and Marian V. Wrobel. "Misperception in Choosing Medicare Drug Plans." Working Paper, 2009.

Lacetera, Nicola, Devin Pope, and Justin Sydnor. Forthcoming. "Heuristic Thinking and Limited Attention in the Car Market." *American Economic Review.*

Levinson, Daniel. "Centers for Medicare & Medicaid Services Audits of Medicare Part D Bids." Office of Inspector General, Department of Health and Human Services, 2008.

Levinson, Daniel. "Midyear Formulary Changes in Medicare Prescription Drug Plans." Office of Inspector General, Department of Health and Human Services, 2009.

Madrian, Brigitte and Dennis Shea. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116 (2001), 1149-1187.

Maskin, Eric, and Jean Tirole. "Markov Perfect Equilibrium: I. Observable Actions." *Journal of Economic Theory* 100 (2001), 191-219.

McCrary, Justin. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2008), 698-714.

Neal, Derek. "How Vouchers Could Change the Market for Education." *Journal of Economic Perspectives* 16 (2002), 25-44.

O'Donoghue, Ted, and Matthew Rabin. "Choice and Procrastination." *Quarterly Journal of Economics* 116 (2001) 121-160.

Robst, John, Jesse Levy, and Melvin Ingber. "Diagnosis-Based Risk Adjustment for Medicare Prescription Drug Plan Payments." *Health Care Financing Review* 28 (2007), 15.

Roth, Al. "The Economist as Engineer: Game Theory, Experimental Economics and Computation as Tools of Design Economics." *Econometrica* 70 (2002), 1341-1378.

Samuelson, William, and Richard Zeckhauser. "Status Quo Bias in Decision Making." *Journal of Risk & Uncertainty* 1 (1988), 7-59.

Strombom, Bruce, Thomas Buchmueller, and Paul Feldstein. "Switching Costs, Price Sensitivity and Health Plan Choice." *Journal of Health Economics* 21 (2002), 89-116.

Taylor, Curtis. "Supplier Surfing: Competition and Consumer Behavior in Subscription Markets." *RAND Journal of Economics* 34 (2003), 223-246.

Wilson, Robert. "Architecture of Power Markets." *Econometrica* 70 (2002),1299–1340.

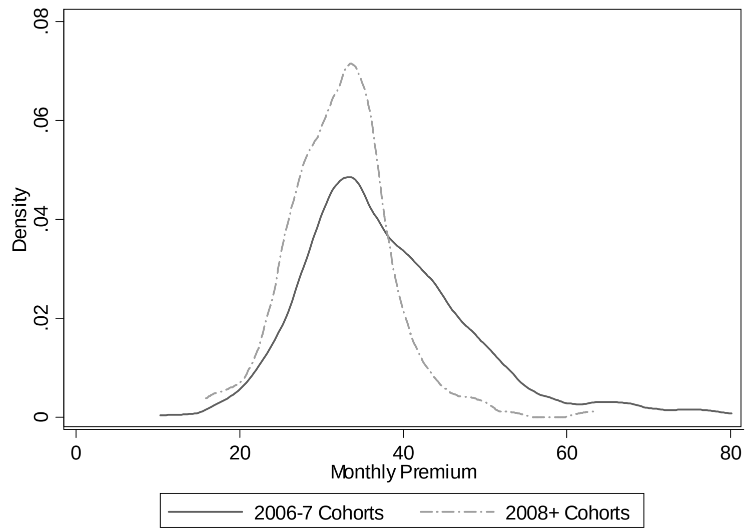

Figure 1: Distribution of Basic PDP Plan Premiums in 2010, by Year of Plan Introduction. Epanechnikov kernel density.

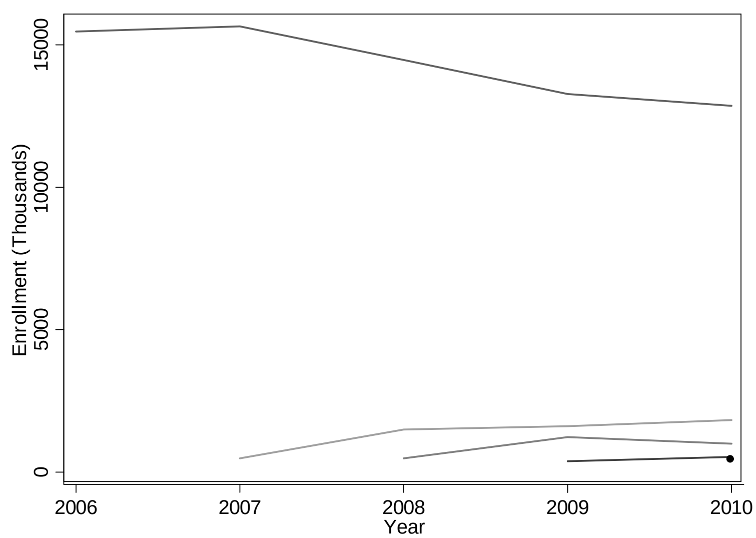

Figure 2: Total PDP Enrollment, by Year and Cohort of Plan. Each line traces the total enrollment of each cohort of plans over time. The enrollment of the 2010 cohort is indicated by a circular marker. Total enrollment includes both standard enrollees and LIS recipients, and is taken as of July 1 of each year. See Appendix Section A.3 for details on data construction.

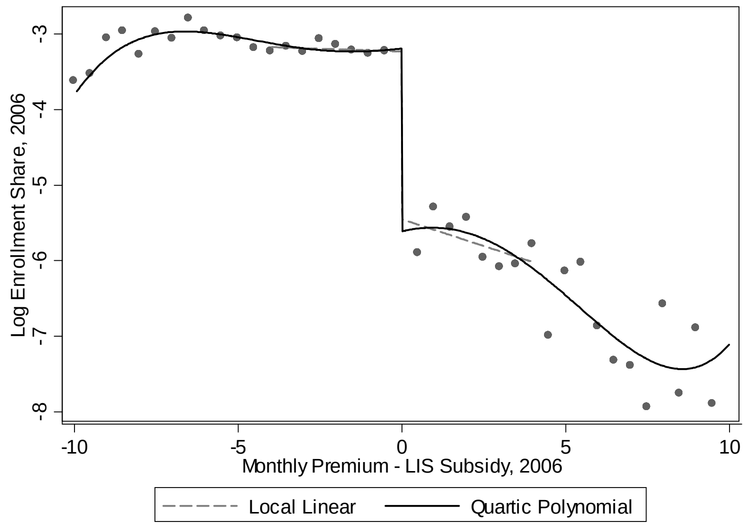

Figure 3: The Effect of 2006 Benchmark Status on 2006 Enrollment. Dots are local averages with a binsize of \$0.50. Dashed lines are predictions from local linear regressions with bandwidth of \$4. Solid lines are predictions from regressions with a quartic polynomial with a bandwidth of \$10.

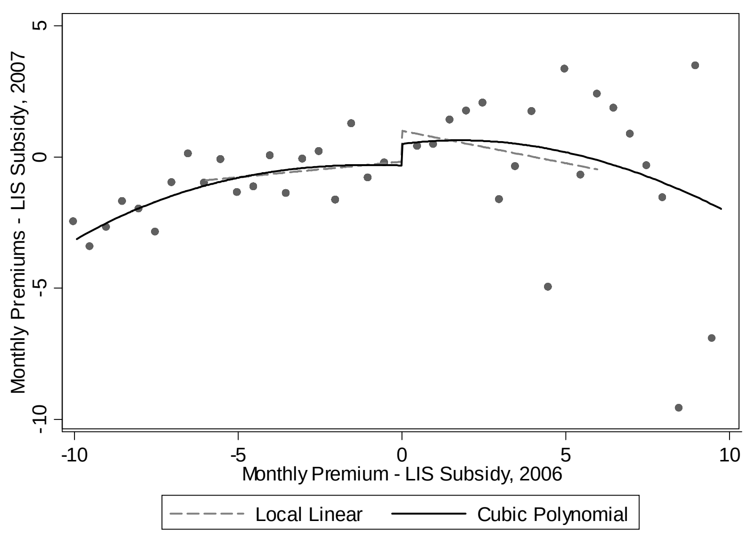

Figure 4: The Effect of 2006 Benchmark Status on 2007 Premiums. Dots are local averages with a binsize of \$0.50. Dashed lines are predictions from local linear regressions with bandwidth of \$6. Solid lines are predictions from regressions with a cubic polynomial with a bandwidth of \$10.

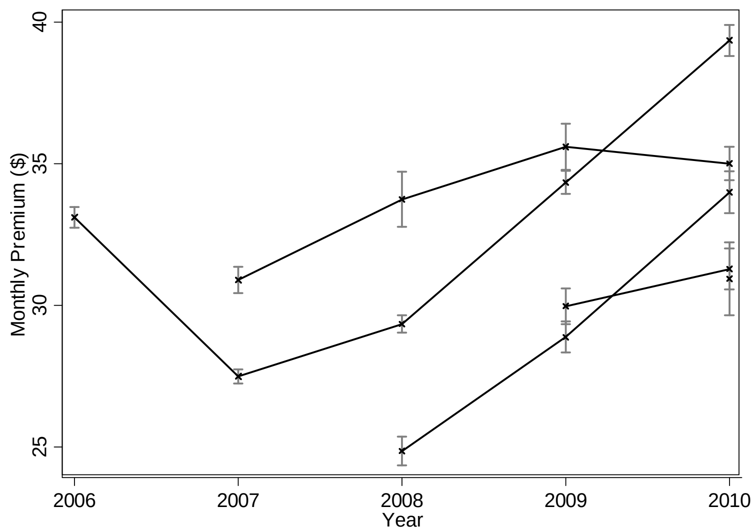

Figure 5: Evolution of Cohort Premiums over Time. Average monthly premiums for basic PDP plans, by plan cohort and year. Each line traces gives the annual premium over time of a given cohort. Standard errors are in grey.

Table 1: Descriptive Statistics of Medicare Part D Plans

	Cohort (Year of Plan Introduction)				
	2006	2007	2008	2009	2010
Mean monthly premium	$ 37	$ 40	$ 36	$ 30	$ 33
	(13)	(17)	(20)	(5)	(9)
Mean deductible	$ 92	$ 114	$ 146	$ 253	$ 118
	(116)	(128)	(125)	(102)	(139)
Fraction enhanced benefit	0.43	0.43	0.58	0.03	0.69
Fraction of plans offered by firms already offering a plan...					
...in the U.S.	0.00	0.76	0.98	1.00	0.97
...in the same state	0.00	0.53	0.91	0.68	0.86
N Unique Firms	51	38	16	5	6
N Plans	1429	658	202	68	107

Source: Author's calculations from CMS Landscape Source Files. Plan characteristics are taken from the year the plan was introduced (e.g. premium in plan's first year). Standard deviations in parentheses.

Table 2: Response of Enrollment to Contemporaneous and Past Prices: 2007

	(1)	(2)	(3)	(4)	(5)	(6)
	$\ln s_{2007}$	$\ln s_{2007}$	$\ln s_{2006}$	$\ln s_{2007}$	$\ln s_{2007}$	$\ln s_{2006}$
Premium in 2007	-0.0971***	-0.146***		-0.0899***	-0.105***	
	(0.0308)	(0.0447)		(0.0285)	(0.0335)	
Premium in 2006	-0.0773***		-0.140***	-0.0694***		-0.173***
	(0.0185)		(0.0281)	(0.0222)		(0.0254)
Type of Basic Plan	Yes	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	No	No	No	Yes	Yes	Yes
N	560	560	553	560	560	553
R^2	0.648	0.484	0.552	0.827	0.800	0.757

OLS regression. Dependent variable: log of plan market share for non-LIS enrollees in a year. Sample: basic PDP plans that were introduced in 2006, and that do not attrit or switch to or from enhanced benefit type before 2007. Plans are dropped from the regression if they have fewer than 10 total enrollees or if estimated enrollment net of LIS is negative. See Appendix Section A.3 for more details. In all columns, state fixed effects and benefit type indicators (Defined Standard, Actuarially Equivalent Standard, or Basic Alternative) are included, and for Basic Alternative plans, deductible bins of $0, $1 to $50,$51 to $100 ..., are included. In columns 1 and 4, controls are included separately for type of basic plan and deductible in both 2006 and 2007. Indicators for pricing below the LIS benchmark are also included, separately for 2006 and 2007. Heteroskedasticity robust standard errors, clustered at the firm level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Effect of LIS Benchmark Status in 2006 on Plan Enrollment

$\ln s_t$	2006	2007	2008	2009	2010
	Panel 1: Local linear, bandwidth $4				
Below Benchmark, 2006	2.224***	1.332***	0.902***	0.803**	0.677
	(0.283)	(0.267)	(0.248)	(0.362)	(0.481)
Premium - Subsidy, 2006					
... Below Benchmark	-0.0141	-0.0774	-0.0731	-0.170	-0.215**
	(0.0322)	(0.0882)	(0.116)	(0.105)	(0.0878)
... Above Benchmark	-0.142*	-0.0331	0.0494	0.0737	0.0488
	(0.0783)	(0.110)	(0.163)	(0.170)	(0.202)
N	306	299	298	246	212
R^2	0.576	0.325	0.131	0.141	0.124
	Panel 2: Polynomial with controls, bandwidth $4				
Below Benchmark, 2006	2.464***	1.364***	0.872***	0.351	-0.277
	(0.222)	(0.321)	(0.246)	(0.324)	(0.301)
Premium - Subsidy, 2006	Quadratic	Quadratic	Quadratic	Quadratic	Quadratic
N	306	299	298	246	212
R^2	0.794	0.576	0.472	0.535	0.685
	Panel 3: Past interactions, local linear, bandwidth $4				
Below Benchmark or de minimis in:					
...2006 and current year	2.224***	2.089***	2.377***	2.633***	2.443***
	(0.283)	(0.364)	(0.275)	(0.257)	(0.309)
...2006 but not current year		0.628**	0.892**	1.068**	0.967
		(0.293)	(0.329)	(0.446)	(0.625)
...current year but not 2006		0.148	1.356***	2.107***	2.281***
		(0.290)	(0.293)	(0.242)	(0.259)
Premium - Subsidy, 2006	Linear	Linear	Linear	Linear	Linear
N	306	299	298	246	212
R^2	0.576	0.480	0.426	0.498	0.467

Each panel is a separate regression. Dependent variable: log of total plan market share (including LIS enrollees) in a year. Sample: basic PDP plans with premiums within the bandwidth window ($4 on either side of the benchmark) in 2006. In "Polynomial with controls", regressions include state and firm fixed effects, and benefit type indicators (Defined Standard, Actuarially Equivalent Standard, or Basic Alternative). For Basic Alternative plans, deductible bins of $0, $1 to $50, $51 to $100 ..., are included. Premium minus subsidy is included as a polynomial separately above and below the benchmark. Heteroskedasticity robust standard errors, clustered at the firm level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Medicare Part D Premiums by Plan Age

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(Monthly Premium)					
	Equal Weighted			Enrollment Weighted		
Year of Plan Existence						
...2nd Year	-0.0167	-0.0103	0.0129	0.0183	-0.0229	0.0139
	(0.0508)	(0.0597)	(0.0511)	(0.0478)	(0.0446)	(0.0593)
...3rd Year	0.0290	0.0585	0.0785	0.128**	0.0795**	0.133***
	(0.0808)	(0.0699)	(0.0519)	(0.0528)	(0.0326)	(0.0358)
...4th Year	0.0690	0.117*	0.148***	0.199***	0.112**	0.191***
	(0.0660)	(0.0617)	(0.0496)	(0.0647)	(0.0522)	(0.0684)
...5th Year	0.177**	0.147**	0.0960*	0.320***	0.154***	0.152*
	(0.0871)	(0.0593)	(0.0551)	(0.0861)	(0.0530)	(0.0764)
Firm Offers M.A. Plan		-0.145**			-0.0390	
		(0.0653)			(0.0350)	
Type of Basic Plan	No	Yes	Yes	No	Yes	Yes
Firm Fixed Effects	No	No	Yes	No	No	Yes
N	4,276	4,276	4,276	4,123	4,123	4,123
R^2	0.189	0.396	0.405	0.364	0.632	0.683

Dependent variable: log monthly PDP premium or monthly premium. Sample: basic PDP plans. All regressions include state fixed effects interacted with year fixed effects. Controls for type of basic plan include benefit type indicators (Defined Standard, Actuarially Equivalent Standard, or Basic Alternative) interacted with year fixed effects. For Basic Alternative plans, deductible bins of \$0, \$1 to \$50,\$51 to \$100 ..., are also included and interacted with year fixed effects. Enrollment weighted regressions are weighted using the plan's total enrollment in July of each year. Plans with fewer than 10 enrollees are dropped from weighted regressions. See Appendix Section A.3 for more details. Heteroskedasticity robust standard errors, clustered at the firm level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

For Online Publication: Web Appendix

For "Market Design when Firms Interact with Inertial Consumers: Evidence from Medicare Part D"

Keith M Marzilli Ericson

A.1 Proof of Propositions in the Text

Proposition 1. *A pure-strategy Markov-perfect equilibrium exists and takes the following form. New firms* ($N \geq 2$) *enter each period and all set price* p_L. *Plans that continue from the previous period with stuck-in-place enrollees charge higher premiums* $p_H > p_L$. *Define the enrollment of new firms as* $s_0 = \frac{1}{N}\left[\rho + (1-\rho) F(p_H - p_L)\right]$. *Then, prices are given by* $p_L = c - \delta \frac{V((1-\rho)s_0)}{s_0}$, *and* $p_H(s) = c + \frac{1-F(p_H-p_L)}{f(p_H-p_L)} - \delta(1-\rho) V'(s)$. *The value of a firm with measure* s *of enrollees is* $V(s) = s\frac{[1-F(p_H-p_L)]^2}{f(p_H-p_L)}$. *The price differential between new and continuing plans is* $(p_H - p_L) = \frac{1-F(p_H-p_L)}{f(p_H-p_L)}$.

Proof. I show the proposed equilibrium exists by construction. Take the proposed value function for a firm with market share s:

$$V(s) = s\frac{[1-F(\Delta p^*)]^2}{f(\Delta p^*)}$$

where we define $\Delta p^* = p_H^* - p_L^* = \frac{1-F(\Delta p^*)}{f(\Delta p^*)}$. This object exists, as it is the solution to $\max_{\Delta p} \Delta p (1 - F(\Delta p))$. Note that Δp^* does not depend on either p_H or p_L directly.

This value function is linear in s and hence has a constant derivative. Since the value function of a firm with positive enrollment is given by

$$V(s) = \max_p s(p-c)[1-F(p-p_L)] + \delta V(s(1-\rho)[1-F(p-p_L)])$$

the first order condition determining the optimal price is

$$p_H(s) = c + \frac{1-F(p_H-p_L)}{f(p_H-p_L)} - \delta(1-\rho) V'(s)$$

This does not depend on s and is constant given the linearity of V. Substituting in for V' gives

$$p_H^* - c = \frac{1-F(\Delta p^*)}{f(\Delta p^*)}\left[1 - \delta(1-\rho)[1-F(\Delta p^*)]\right]$$

which defines p_H. It is easy to see that a strategy of setting p_H^* yields the proposed value function, as

$$\begin{aligned} V(s) &= s(p_H - c)[1 - F(\Delta p^*)]\left\{1 + \delta(1-\rho)[1 - F(\Delta p^*)] + \delta^2(1-\rho)^2[1 - F(\Delta p^*)]^2 + ...\right\} \\ &= s(p_H - c)[1 - F(\Delta p^*)]\frac{1}{1 - \delta(1-\rho)[1 - F(\Delta p^*)]} \end{aligned}$$

Now consider newly introduced plans. The market is competitive, and so new firms must have zero expected value. Thus, $s_0(p_L - c) = -\delta V((1-\rho)s_0)$, or by the linearity of V:

$$\begin{aligned} p_L^* &= c - \delta(1-\rho)V'(s) \\ &= c - \delta(1-\rho)\frac{[1 - F(\Delta p^*)]^2}{f(\Delta p^*)} \end{aligned}$$

It is easily checked that these values of p_L^* and p_H^* satisfy the definition of Δp^*. Thus, this method solves for p_H and p_L.

Now I show that neither type of firm has an incentive to deviate from the proposed strategy. First, note that the optimal strategy of an existing firm depends solely on its own enrollment and p_L. This is true since a plan's enrollment is a function only of the plan's own price and the lowest price in the market. Hence the behavior of firms other than new firms does not matter. Consider deviations of new firms, setting price to p'. If $p' > p_L$, the firm gets no enrollment, and makes zero profit. If $p' < p_L$, the firm makes negative discounted profits. Hence there are no profitable deviations for new firms.

Finally, consider deviations of existing firms. Given p_L, p_H is defined as profit maximizing and so there is no incentive to deviate to any other $p'_H > p_L$. If $p'_H < p_L$, the firm makes negative discounted profits, just as a new firm pricing below p_L would. The other potential deviation is to $p'_H = p_L$. But this would give zero profits: such a firm would get higher enrollment $s' > s_0$ since it would attract unattached individuals as well as keep all its own enrollees. But the value of such a firm is $s'(p_L - c) + \delta V((1-\rho)s') = 0$ (assuming future optimal action) which is invariant to s' and equal to zero by construction. Hence the proposed $\{p_L, p_H\}$ strategy is an equilibrium.

Proposition 2. *Suppose firms play the strategies in Proposition 1. The privately optimal default for a measure-zero subset of the population is automatic switching if* $\int_0^{\bar{\lambda}+\Delta p} \omega dG(\omega) < \Delta p \cdot G\left(\bar{\lambda} + \Delta p\right)$, *and is otherwise automatic reenrollment.*

Proof. Maximizing lifetime utility here is the same as minimizing total costs borne (premiums plus switching costs). When an enrollee enters the market, they pay p_L regardless of the default. Afterwards, each period the enrollee faces the same distribution of plan prices, switching costs, and psychological frictions. Hence, we simply need to consider expected welfare in a single period. The expected cost is p_H if the individual does not switch and p_L plus the switching cost paid if the individual switches. Then, under an automatic reenrollment default, expected total costs (ETC_{Re}) in a period are :

$$\begin{aligned} ETC_{Re} &= p_H \left[1 - (1-\psi) G(\Delta p)\right] + p_L (1-\psi) G(\Delta p) + (1-\psi) \int_0^{\Delta p} \omega dG(\omega) \\ &= p_H - \Delta p \left[(1-\psi) G(\Delta p)\right] + (1-\psi) \int_0^{\Delta p} \omega dG(\omega) \end{aligned}$$

since an individual switches only with probability $(1-\psi)\Pr(\Delta p > \omega_{it})$. Under an automatic switching default, expected total costs in a period are

$$\begin{aligned} ETC_{Sw} &= p_H - \Delta p \left[(1-\psi) G(\Delta p) + \psi G\left(\bar{\lambda} + \Delta p\right)\right] \\ &\quad + (1-\psi) \int_0^{\Delta p} \omega dG(\omega) + \psi \int_0^{\bar{\lambda}+\Delta p} \omega dG(\omega) \end{aligned}$$

since individuals switch 1) if they are attentive and $\Delta p > \omega_{it}$, and 2) if they are inattentive and $\bar{\lambda} + \Delta p > \omega_{it}$. We have the automatic switching default optimal if $ETC_{Re} > ETC_{Sw}$, or

$$\Delta p \cdot G\left(\bar{\lambda} + \Delta p\right) > \int_0^{\bar{\lambda}+\Delta p} \omega dG(\omega)$$

as asserted.

Proposition 3. *Suppose firms play the strategies in Proposition 1. The socially optimal default is automatic switching if* $\psi \int_0^{\bar{\lambda}+\Delta p_{Sw}} \omega dG(\omega) < (1-\psi) \int_{\Delta p_{Sw}}^{\Delta p_{Re}} \omega dG(\omega)$*, and is otherwise automatic reenrollment.*

Proof. The optimal default from the social welfare perspective simply minimizes switching costs paid, when transfers to all individuals are equally weighted. (Benefit provision costs of c per period are invariant to the default.) By the same argument as in the proof of Proposition 2, we need only consider the expected switching costs paid in each period. Switching costs borne per period under the automatic switching default are equal to

$$(1-\psi) \int_0^{\Delta p_{Sw}} \omega dG(\omega) + \psi \int_0^{\bar{\lambda}+\Delta p_{Sw}} \omega dG(\omega)$$

since individuals optimally switch when attentive (probability $1-\psi$), and switch by default (probability ψ) so long as ω is not too large. Similarly, switching costs borne under automatic reenrollment are equal to $(1-\psi) \int_0^{\Delta p_{Re}} \omega dG(\omega)$ per period. Hence, switching costs are lower under automatic switching if

$$\psi \int_{0}^{\bar{\lambda}+\Delta p_{Sw}} \omega dG\left(\omega\right) < \left(1-\psi\right) \int_{\Delta p_{Sw}}^{\Delta p_{Re}}$$

as asserted.

A.2 Theory Appendix

A.2.1 Optimal Pricing with LIS Recipients

This section considers the theoretical predictions for firm pricing that result from acquiring LIS recipients. Consider two firms that are identical, except that one firm has acquired a number of LIS enrollees by pricing just below the price benchmark. This is the situation analyzed in the regression discontinuity design in Section 5. Both types of firms share a common component to their demand curves $s_{jt}\left(p_{jt}\right)$, which captures the behavior of the standard (non-LIS) enrollees, as well as the potential to capture new LIS auto-enrollees if the firm prices below the benchmark in subsequent years. However, LIS recipients and standard enrollees face different prices and defaults, so acquiring LIS auto-enrollees will alter a firm's incentives when setting prices in subsequent years. The effect of acquiring an LIS recipient on subsequent pricing is theoretically ambiguous.

Firms that price below the benchmark have an additional component to their demand curve: the effective demand of the LIS auto-enrollees, $\tilde{L}\left(p_j - b\right)$. The LIS demand curve $\tilde{L}$ is a composite of individual preferences and the automatic switching default, as LIS auto-enrollees are defaulted into a different plan if $p_j > b$, but some of these auto-enrollees may actively choose to stay with their current firm even if defaulted elsewhere. Note than when $p_j < b$, the plan is free to LIS recipients who receive the full subsidy, and their enrollment is relatively insensitive to changes in the firm's price in that region. Hence the demand curve $\tilde{L}$ is relatively flat below b, falls discontinuously at b, and then is more price sensitive above b. However, because the benchmark is unknown when setting prices, firms with LIS recipients face an expected LIS demand curve. Write the demand curve from the firm's perspective as $L\left(p_j\right) = E\left[\tilde{L}\left(p_j - b\right)\right]$.

Imperfect risk adjustment implies that the costs to the firm of non-LIS and LIS recipients may differ. However, for simplicity, I set the costs of both types of individuals to be the same in the condition below. Modifying Equation 1 from Section 3.3, a firm with LIS recipients sets prices to maximize:

$$V_{jt} = \left(p_{jt} - c_N\right) s_{jt} + \left(p_{jt} - c_L\right)\left[L_{jt}\left(p_{jt}\right)\right] + \delta V_{jt+1}\left(s_{jt}, L_{jt}\right)$$

Then, the firm with LIS recipients has the first order condition:

$$p_{jt} - c = \frac{s_{jt} + L_{jt}}{-\left(s'_{jt} + L'_{jt}\right)} - \frac{\delta}{s'_{jt} + L'_{jt}} \left[\frac{\partial V_{jt+1}\left(s_{jt}, L_{jt}\right)}{\partial s_{jt}} s'_{jt} + \frac{\partial V_{jt+1}\left(s_{jt}, L_{jt}\right)}{\partial L_{jt}} L'_{jt} \right]$$

where the derivatives s'_{jt} and L'_{jt} are taken with respect to price.

Comparing this condition to that for firms without LIS recipients (Equation 1) shows that acquiring LIS recipients has an ambiguous effect on firm pricing.[51] Consider the simplest case in which firms are concerned only about this present period's profit ($\delta = 0$) and continue to assume the costs of the two types of individuals are the same. Then, theory predicts that the price will be higher for the firms with LIS recipients if the *expected* LIS enrollment is less price elastic than the enrollment of non-LIS individuals ($\frac{-L'_{jt}}{L_{jt}} > \frac{-s'_{jt}}{s_{jt}}$), and lower if the expected LIS enrollment is more elastic. This elasticity depends not only on the behavior of LIS recipients, but also on the firm's (unobserved) subjective probability distribution of the location of the benchmark.

The costs of LIS recipients are higher, due to a failure of risk adjustment, and so prices for firms with LIS recipients can be higher even if expected LIS enrollment is more price elastic (and vice versa). Moreover, firms may believe that acquiring an LIS recipient has less of an effect on their future profits, due to policies (such as the de minimis policy) that may make them less profitable in the future. In such a case, $\frac{\partial V_{jt+1}}{\partial L_{jt}} < \frac{\partial V_{jt+1}}{\partial s_{jt}}$ and optimal pricing is also higher.

A.2.2 Sophisticates and Myopes

This section formalizes claims and extensions to the model in Section 3.4 that were discussed in the text.

Claim 1. *Switching if $p_H - p_L > \sigma_{it}$ for σ_{it} drawn from some distribution F is an optimal strategy for both sophisticates and myopes given the equilibrium proposed in Proposition 1.*

Proof. Sophisticates recognize that the gain from switching is a one time event: in the future they will pay p_H every period, regardless of what they pay today. In this case, they will switch if Δp is greater than their switching cost is today.

[51] This discussion considers the case in which the first order condition is satisfied at a single point. However, the first order condition may be satisfied at multiple points, at prices above and below the benchmark, as the expected LIS demand curve is relatively flat above and below the region in which the benchmark is likely to be. Consider the limiting case in which the benchmark is known perfectly ex ante and a firm's existing LIS recipients will stay with their current plan if and only if its price is below benchmark. Then a firm may choose between setting a price just below the benchmark and keeping its LIS recipients, or setting a price substantially above the benchmark and maximizing profits on standard enrollees.

Myopes believe the future prices of all plans will remain the same as they are in this period. Let myopes draw a switching cost ω_{it} from G each period, but let the psychological friction λ_{it} be equal to zero. They compare switching this period at cost ω_{it} (and paying p_L forever), to the option value of staying in their current plan this period and potentially switching in the future. They then have the Bellman equation

$$V_t(\omega_{it}) = \max\left\{-\omega_{it} - \frac{1}{1-\delta(1-\rho)}p_L, -p_H + \delta(1-\rho)EV_{t+1}(\omega)\right\}$$

The solution to this optimal stopping problem is a threshold ω^*, which decreases in Δp. The probability that an individual switches is the probability that $\omega_{it} < \omega^*(\Delta p)$, which can be rewritten as $F(\sigma)$, as claimed in the text. When psychological frictions are added ($\lambda > 0$), myopes only switch if $-\omega_{it} - \frac{1}{1-\delta(1-\rho)}p_L - \lambda_{it} > -p_H + \delta(1-\rho)EV_{t+1}(\omega)$, and take this lack of switching into account when calculating the continuation value of not switching.

Claim 2. *So long as there is a positive measure of myopes, there is no Nash equilibrium in which all firms charge the same price $\tilde{p}$ each period.*

Proof. Let there be measure m of myopic individuals and let all others be sophisticates. We only need consider $\tilde{p} > c$. It cannot be that $\tilde{p} < c$, or firms would make negative profits and could deviate by exiting the market. If $\tilde{p} = c$, then firms make zero profits in the proposed equilibrium. In this case, a firm with positive market share could deviate and make positive profits, by setting $p' > \tilde{p}$ in the present period and then leave the market in later periods.

Now suppose $\tilde{p} > c$. I show that new firms will have an incentive to deviate to a lower price. Under the proposed pricing, new firms make lifetime discounted profits of $\frac{1}{1-\delta(1-\rho)}(\tilde{p}-c)$ per acquired enrollee. Since firms make positive profits, new firms will choose to enter when they have the opportunity. Denote the total number of firms in the market in period t as n_t. Since all firms set the same price, they all receive an equal share of new enrollees. Since the number of firms in the market is increasing, a new firm expects to receive less than $\frac{\rho}{n_t}$ enrollees each period in the future. Hence, the present discounted profits of a new firm following this strategy is less than $\frac{1}{1-\delta}\frac{\rho}{n_t}\frac{(\tilde{p}-c)}{1-\delta(1-\rho)}$. A firm deviating to price $p' = \tilde{p}-\varepsilon$, for some $\varepsilon > 0$, would capture all unattached myopic consumers, totaling measure m. (It may or may not attract sophisticated customers, depending on the anticipated response of other firms.) It therefore receives a profit of $m[\rho(\tilde{p}-c)-\rho\varepsilon]$ in that period alone. The deviation is certainly profitable if $m[\rho(\tilde{p}-c)-\rho\varepsilon] > \frac{1}{1-\delta}\frac{\rho}{n_t}\frac{(\tilde{p}-c)}{1-\delta(1-\rho)}$. Since new firms enter every period, n_t will eventually be large enough to make this condition hold. Hence, there is no equilibrium in which all firms charge the same price each period for $m > 0$.

A.2.3 Optimal Defaults with General Distribution H

This appendix generalizes the discussion in Section 7. I now consider optimal defaults when λ_{it}, the tolerable losses from inaction resulting from psychological frictions, is drawn from an arbitrary distribution H that is continuous, bounded and differentiable with p.d.f. h. As in Section 7, individuals also draw a switching cost ω from a continuous, bounded and differentiable distribution G.

Furthermore, I allow for an "opt-out" cost κ that must be borne when an individual does not take the default option, where κ is a real resource cost and lowers utility. The opt-out cost κ represents the cost of actively expressing preference (e.g. sending back a form), and is distinct from the cost of switching plans (e.g. setting up prescription to be billed to a new insurer). Thus, κ is borne when people switch under an automatic reenrollment default, and when people do not switch under an automatic switching default. The cost κ acts similarly to λ in how it affects individual and firm behavior, but creates an additional motivation to choose a default that matches the modal behavior of the population: if most people switch plans each period, then an automatic switching default might raise welfare by saving most people the cost of opting out of the default.

When the default is automatic reenrollment, individuals switch if the gain to doing so, net of switching costs, exceeds $\lambda_{it} + \kappa$, the psychological friction and the real cost of opting out of the default. This occurs with probability $q_{Re} \equiv \int_0^\infty H\left(\Delta p - \omega - \kappa\right) dG\left(\omega\right)$: the probability that $\Delta p - \omega_{it} > \lambda_{it} + \kappa$, integrated over draws of ω. Similarly, when the default is automatic switching, individuals always switch when $\Delta p - \omega_{it} > -\kappa$, since they compare the gain of switching to paying κ if they opt out of the default. However, when $\lambda > 0$, they also switch so long as $\omega_{it} - \Delta p - \kappa < \lambda_{it}$, since they are willing to tolerate a loss of λ to stay with the default of switching. Thus, the probability they switch is given by $q_{Sw} \equiv 1 - \int_0^\infty H\left(\omega - \Delta p - \kappa\right) dG\left(\omega\right).$

Now, we have the analogues of Propositions 2 and 3. Proposition A.1 again shows that the privately optimal default for an individual weighs the premiums saved against the increased switching costs and opt-out costs borne. Similarly, Proposition A.2 shows that the socially optimal default for the entire population is the default that minimizes switching costs and opt-out costs.

Proposition A.1. *Suppose firms play the strategies in Proposition 1. The privately optimal default for a measure-zero subset of the population is automatic switching if*

$$\kappa\left[\left(1 - q_{Sw}\right) - q_{Re}\right] + \int_0^\infty \int_{\Delta p - \lambda - \kappa}^{\Delta p + \lambda + \kappa} \omega dG\left(\omega\right) dH\left(\lambda\right) < \Delta p\left(q_{Sw} - q_{Re}\right)$$

and is otherwise automatic reenrollment.

Proof. The proof follows that of Proposition 2. Under automatic reenrollment, expected total costs in a period are given by

$$ETC_{Re} = \kappa q_{Re} + p_H (1 - q_{Re}) + p_L (q_{Re}) + \int_0^{\infty} \int_0^{\Delta p - \kappa - \lambda} \omega dG(\omega) dH(\lambda)$$

where the last term is switching costs borne. For each value of λ, the switching costs borne are those between $\omega = 0$ and $\omega = \Delta p - \lambda - \kappa$. Similarly, under automatic switching, expected total costs in a period are given by

$$ETC_{Sw} = \kappa (1 - q_{Sw}) + p_H (1 - q_{Sw}) + p_L (q_{Sw}) + \int_0^{\infty} \int_0^{\Delta p + \lambda + \kappa} \omega dG(\omega) dH(\lambda)$$

where again, the last term is switching costs borne. For each value of λ, we take the integral of switching costs from $\omega = 0$ to $\omega = \Delta p + \lambda + \kappa$, since the latter switching cost gives the maximal tolerable loss from switching. Now, automatic switching is privately optimal if $ETC_{Sw} < ETC_{Re}$, which requires

$$\kappa [(1 - q_{Sw}) - q_{Re}] + \int_0^{\infty} \int_{\Delta p - \lambda - \kappa}^{\Delta p + \lambda + \kappa} \omega dG(\omega) dH(\lambda) < \Delta p (q_{Sw} - q_{Re})$$

as asserted.

Proposition A.2. *Define* $z = \frac{\Delta p_{Re} - \Delta p_{Sw} - 2\kappa}{2}$. *Suppose firms play the strategies in Proposition 1. The socially optimal default is automatic switching if*

$$\kappa [(1 - q_{Sw}) - q_{Re}] + \int_z^{\infty} \int_{\Delta p_{Re} - \lambda - \kappa}^{\Delta p_{Sw} + \lambda + \kappa} \omega dG(\omega) dH(\lambda) < \int_0^{z} \int_{\Delta p_{Sw} + \lambda + \kappa}^{\Delta p_{Re} - \lambda - \kappa} \omega dG(\omega) dH(\lambda)$$

and is otherwise automatic reenrollment.

Proof. The proof follows that of Proposition 3: the socially optimal default minimizes real switching costs and opt-out costs borne. Using the logic of Proposition A.1, note that total switching costs and opt-out costs borne under automatic reenrollment are given by

$$\int_0^{\infty} \int_0^{\Delta p_{Re} - \lambda - \kappa} \omega dG(\omega) dH(\lambda) + \kappa q_{Re}$$

where Δp_{Re} is the equilibrium price differential between new and old plans, given the reenrollment default. Similarly, switching costs and opt-out costs borne under automatic switching are given by

$$\int_0^\infty \int_0^{\Delta p_{Sw}+\lambda+\kappa} \omega dG(\omega)\, dH(\lambda) + \kappa(1-q_{Sw})$$

Note that when $\Delta p_{Sw} \geq \Delta p_{Re}$, total switching costs borne are certainly higher under automatic switching, but opt-out costs may be lower. In general, automatic switching is optimal if total switching and opt-out costs are higher under automatic reenrollment

$$\int_0^\infty \left[\int_0^{\Delta p_{Re}-\lambda-\kappa} \omega dG(\omega) - \int_0^{\Delta p_{Sw}+\lambda+\kappa} \omega dG(\omega)\right] dH(\lambda) + \kappa\left[q_{Re} - (1-q_{Sw})\right] > 0$$

Now, we can break the integrals apart, noting that $\Delta p_{Re} - \lambda - \kappa > \Delta p_{Sw} + \lambda + \kappa$ when $z \equiv \frac{\Delta p_{Re}-\Delta p_{Sw}-2\kappa}{2} > \lambda$. Hence we have automatic switching optimal when

$$\int_0^z \int_{\Delta p_{Sw}+\lambda+\kappa}^{\Delta p_{Re}-\lambda-\kappa} \omega dG(\omega)\, dH(\lambda) > \int_z^\infty \int_{\Delta p_{Re}-\lambda-\kappa}^{\Delta p_{Sw}+\lambda+\kappa} \omega dG(\omega)\, dH(\lambda) + \kappa\left[(1-q_{Sw}) - q_{Re}\right]$$

as asserted.

A.3 Data Appendix

The Medicare Part D Landscape Source File lists premiums and characteristics for all PDP plans. Plans can be linked from year to year using a contract and plan identifier assigned by the Centers for Medicare & Medicaid Services (CMS). Each contract identifier is linkable to a particular firm, but a firm may have multiple contract identifiers. Crosswalk files describe whether plan is merged into another plan, or is terminated.

Total enrollment data, combining both standard enrollees and LIS recipients, is taken from the Monthly Enrollment by Plan file as of July 1 of each calendar year. The July date is chosen because for 2006 and 2007, only the July Monthly Enrollment by Plan was made public by CMS. CMS has also released figures for the enrollment of LIS recipients by plan, but not at regular intervals. Data on LIS enrollment by plan was available for July of 2006 and 2007 and February of 2008 and 2009.

A plan's market share is simply a plan's total enrollment over total enrollment in the state, dropping plans with less than 10 enrollees. Plans with less than 10 enrollees have their enrollment suppressed by Medicare and may not be active. To construct plan market shares of standard (non-LIS enrollment), I subtract each plan's LIS enrollment from its total enrollment, even if these are not taken in the same month (i.e. for 2008 and 2009). Plans with less than 10 LIS enrollees have their LIS enrollment suppressed as well. In these cases, I impute an LIS enrollment of 5 and subtract that from the total enrollment. The resulting estimates of standard enrollment are negative in a small number of cases; these plan-year observations are dropped in the regressions using standard enrollment data.

Firm names are coded as follows. I take the enrollment files, which include PDPs as well as Medicare Advantage plans. For each contract identifier in the enrollment files, I identify the firm name as the CMS "organizational parent" listed for that contract in 2010, or the last year that the contract exists if it attrits before 2010. This system treats subsequently merged firms as one firm, since mergers may be anticipated in pricing. I then hand code the data to combine all forms of a given firm name (e.g. "Universal American Corp.", "Universal American Corporation", and "Universal American Financial Corporation" are all the same firm). These codings are available upon request from the author. Blue Cross and Blue Shield plans act individually to offer Medicare Advantage Plans, but act in alliance to offer PDP plans (e.g., one PDP parent is listed as "BCBS RI & BCBS MA & BCBS VT"). In these cases, I code individual Blue Cross plans in those states as part of that alliance.

The results are not sensitive to the method of firm codings. In regressions using firm fixed effects, coding each contract identifier as a separate firm gives similar results. I have also explored alternative firm codings based on CMS fields for "organizational marketing name" and a variety of treatments for the Blue Cross and Blue Shield plans; results are similar.

A.4 Additional Detail on Medicare Part D

This section gives additional detail on two features of the Medicare Part D program: the calculation of the LIS benchmark amount and the risk-adjustment system.

LIS program recipients receive a regionally determined premium subsidy amount. For each region (state or group of states), the subsidy amount is the greater of a weighted average firm bid or the lowest monthly beneficiary premium for a prescription drug plan that offers basic prescription drug coverage in that PDP region. (Due to the inclusion of Medicare Advantage plans in the calculation of the benchmark, it is possible that no standalone PDP could be below the benchmark). Appendix Figure A.7 shows the evolution of the average benchmark across time. The average firm bid used in the calculation of the benchmark amount is a weighted average of the monthly beneficiary premiums in each region. For 2006, Medicare Advantage prescription drug (MA-PD) bids were assigned a weight based upon prior enrollment, while PDPs were all assigned equal weight as actual enrollment was not yet available. The same approach was used in 2007. In 2008, CMS began to transition to enrollment-weighting the PDP premiums, so PDP premiums were 50% weighted based on prior year's enrollment and 50% equal weighted. Beginning in 2009, the PDP premiums were all enrollment weighted. In 2010, the bids for MA-PD were used before they have been reduced by any applicable MA rebates.[52] This had the effect of raising the subsidy slightly.

I assess the accuracy of Medicare Part D risk adjustment by comparing the age-related adjustment factors to average prescription drug spending in the 2007 Medical Expenditure Panel Survey.[53] Because I do not have access to enrollee claim history, I do not evaluate the diagnosis-related risk adjustment model used for existing Medicare Part D enrollees. Instead, I evaluate the age-related adjustment factors for new enrollees who were not originally disabled.

Medicare sets adjustment factors for each sex separately, and for the following age categories: 65, 66, 67, 68, 69, 70-74, 75-79, 80-84, 85-89. (It also sets age adjustments for individuals above age 90, but I ignore these as the MEPS does not report spending for individuals above age 90). For each age category, I combine the male and female adjustment factors using a weight of 0.5747 for women and 0.4253 for men, which are the relative fractions of women and men over age 65 in the 2007 MEPS data.

To produce estimates of prescription drug spending, I use the 2007 MEPS data ("Table 2: Prescription Medicines"). I construct population figures and mean prescription drug spending per person in each of the same age categories used for Medicare Part D's age-related risk adjustment factor. These estimates are imperfect measures of insurer costs, as they give total prescription drug spending, rather than the total prescription drug spending covered by the insurer. However, these two quantities are likely to move together.

These data indicate that the average prescription drug spending for the population aged 65-89 is $2122, and that the average risk adjustment factor for a firm enrolling a representative sample of the non-disabled population aged 65-89 is 0.9425.

Now consider a firm that has only enrollees aged 70 to 89, in their population relative weights. This would be the experience of a firm that initially introduced a plan 5 years prior, enrolled a representative fraction of the population aged 65-89, and subsequently acquired no new enrollees. Assuming average mortality, the firm's population distribution 5 years later would mirror that of the population, except it would have no individuals aged 65-69. The average prescription drug spending for that population (aged 70-89) is $2177, and the average risk adjustment factor is 0.9719.

Thus, these results indicate that as a population ages by five years and experiences

[52] For more details, see "Medicare Prescription Drug Benefit Manual: Chapter 13 - Premium and Cost-Sharing Subsidies for Low-Income Individuals." Rev. 9, Feb. 5, 2010. http://www.cms.gov/PrescriptionDrugCovContra/Downloads/R7PDB.pdf

[53] Agency for Healthcare Research and Quality. Prescription Medicines-Mean and Median Expenses per Person With Expense and Distribution of Expenses by Source of Payment: United States, 2007. Medical Expenditure Panel Survey Household Component Data. Generated interactively. (August 25, 2010)

average mortality, average prescription drug spending increases by 2.6%. This is matched by age-related risk adjustment that increases by 3.1%. There is no evidence that the age-related risk adjustment in Medicare Part D is insufficient.

A.5 Bounds on Elasticities

Chetty (2011) shows that in the presence of optimization frictions such as adjustment costs, elasticities are not point identified, but bounded. An optimization friction leads to some deviation from an individual's optimal choice. Chetty shows that if the utility loss of the deviation is bounded, bounds on structural elasticities can be derived from observed behavior. Consider an optimization friction that has utility costs of fraction γ of spending on health insurance. In this context, take $\gamma = 0.1$ since switching costs of \$50 are about 10% of annual premiums.

Given an observed elasticity $\hat{\eta} < 0$ and optimization friction δ, the lower and upper bound elasticities η_L, η_U consistent with the observed elasticity are given by (Chetty 2011)[54]:

$$\begin{aligned} \eta_L &= \hat{\eta} - \frac{4\gamma}{(\Delta \ln p)^2}(1-\rho), \\ \eta_U &= \hat{\eta} - \frac{4\gamma}{(\Delta \ln p)^2}(1+\rho) \end{aligned}$$

with

$$\rho = \left(1 + \frac{1}{2}\frac{-\hat{\eta}}{\gamma}(\Delta \ln p)^2\right)^{1/2}$$

For a 50% tax, $\Delta \ln p = 0.41$. Take $\hat{\eta} = -0.07$. Then bounds are given by $\eta_L = -9.02 \times 10^{-4}$ and $\eta_U = -5.01$.

[54] These formulas differ slightly from those in Chetty (2011), as he uses η to represent the negative of the elasticity.

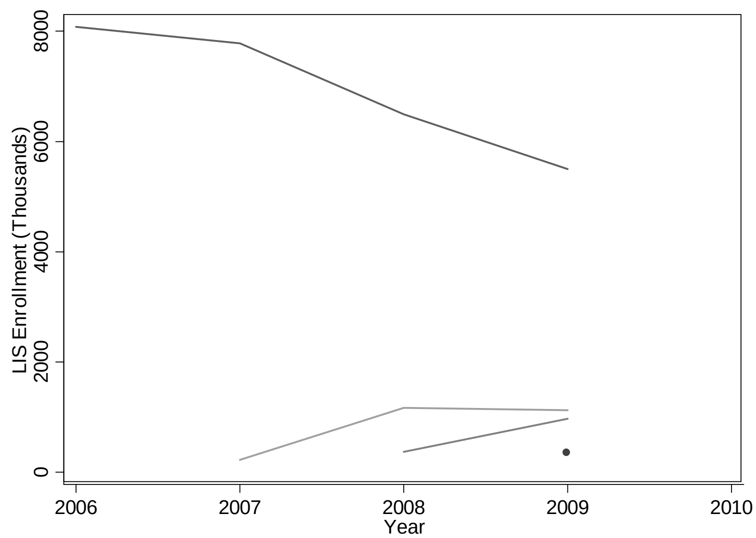

Figure A.1: Aggregate LIS Enrollment, by Year and Cohort of Plan. 2009 cohort indicated by circular marker. See Appendix Section A.3 for details on data construction.

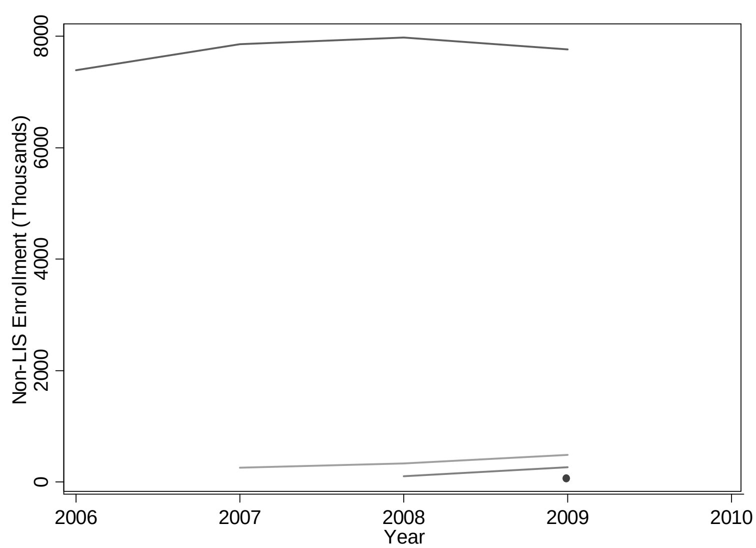

Figure A.2: Aggregate Non-LIS Enrollment, by Year and Cohort of Plan. 2009 cohort indicated by circular marker. See Appendix Section A.3 for details on data construction.

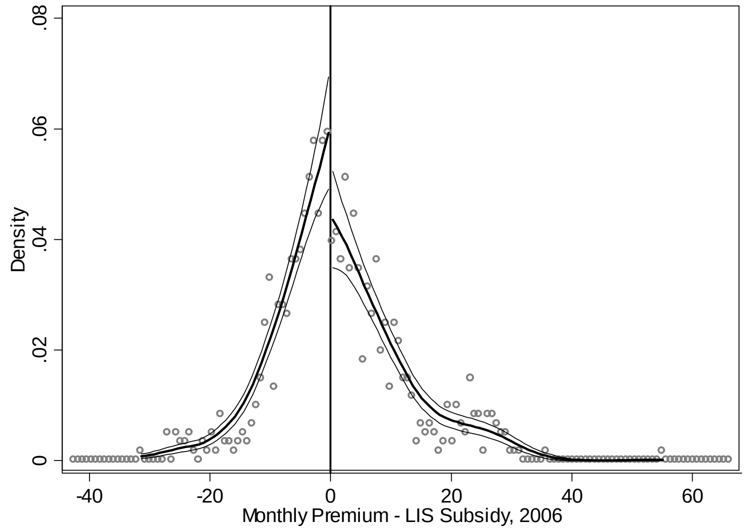

Figure A.3: Test for Density Discontinuity of the Forcing Variable. Dots are density with binsize of 0.74. Lines show smoothed density and standard errors as calculated in McCrary (2008).

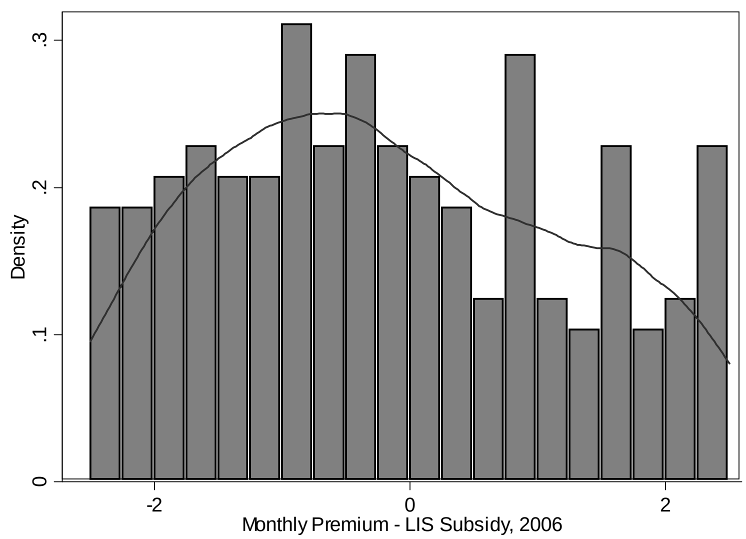

Figure A.4: Histogram of Forcing Variable. Bin width is 0.25. Overlaid with Epanechnikov kernel density. Sample: Basic Plans in 2006.

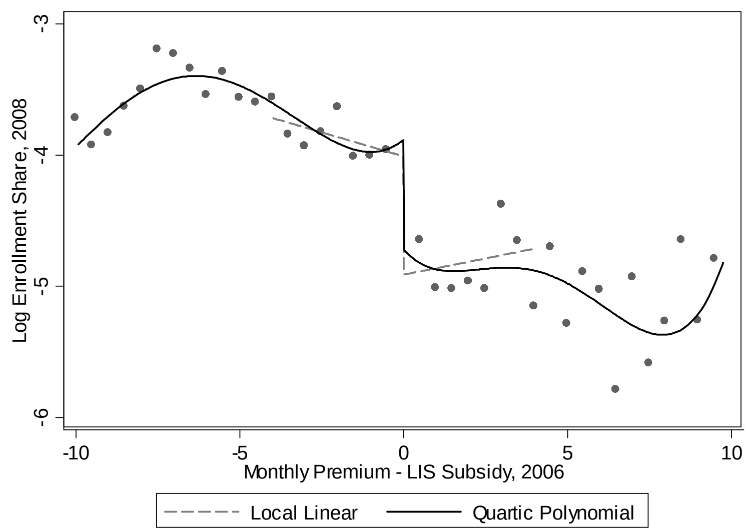

Figure A.5: The Effect of 2006 Benchmark Status on 2008 Enrollment. Dots are local averages with a binsize of \$0.50. Dashed lines are predictions from local linear regressions with bandwidth of \$6. Solid lines are predictions from regressions with a cubic polynomial with a bandwidth of \$10.

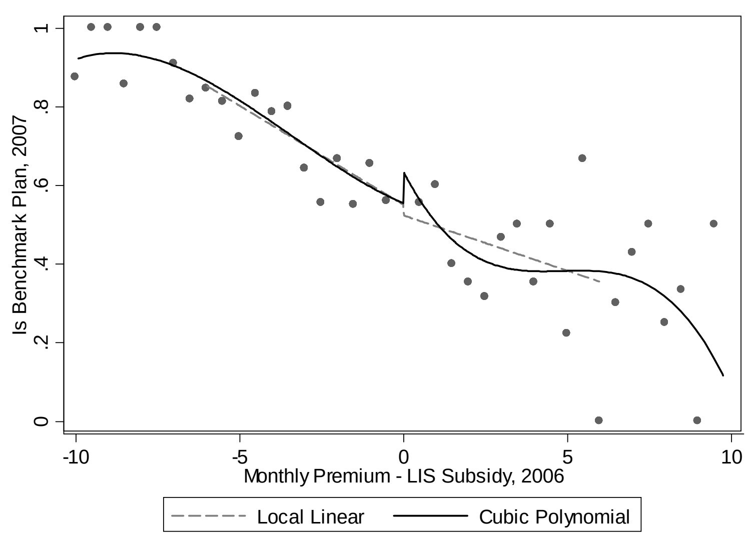

Figure A.6: The Effect of 2006 Benchmark Status on 2007 Benchmark Status. Dependent variable equals 1 if plan is below benchmark or is a de minimis plan in 2007. Dots are local averages with a binsize of \$0.50. Dashed lines are predictions from local linear regressions with bandwidth of \$6. Solid lines are predictions from regressions with a cubic polynomial with a bandwidth of \$10.

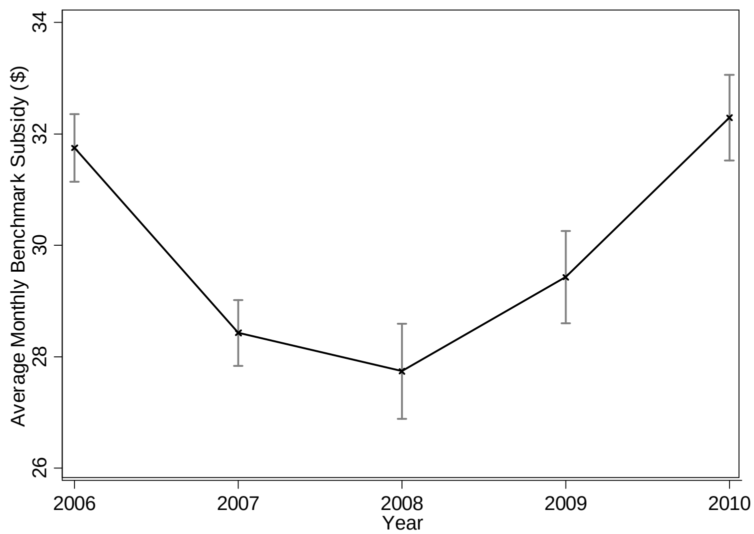

Figure A.7: Average LIS Benchmark Subsidy Level. Equal weighted average over each PDP region (state or group of states). Standard errors are in grey. Source: Author's calculations from CMS data.

Table A.1: Response to Contemporaneous and Past Prices: 2009

	(1) $\ln s_{2009}$	(2) $\ln s_{2009}$	(3) $\ln s_{2006}$	(4) $\ln s_{2009}$	(5) $\ln s_{2009}$	(6) $\ln s_{2006}$
Premium in 2009	0.0120	-0.0103		-0.0628***	-0.0516**	
	(0.0211)	(0.0233)		(0.0106)	(0.0190)	
Premium in 2006	-0.0703***		-0.155***	-0.0620**		-0.233***
	(0.0151)		(0.0276)	(0.0245)		(0.0332)
Type of Basic Plan	Yes	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	No	No	No	Yes	Yes	Yes
N	308	308	301	308	308	301
R^2	0.707	0.460	0.639	0.888	0.803	0.848

OLS regression. Dependent variable: log of plan market share for non-LIS enrollees in a year. Sample: basic PDP plans that were introduced in 2006, and that do not attrit or switch to or from enhanced benefit type before 2009. Plans are dropped from the regression if they have fewer than 10 total enrollees or if estimated enrollment net of LIS is negative. See Appendix Section A.3 for more details. In all columns, state fixed effects and benefit type indicators (Defined Standard, Actuarially Equivalent Standard, or Basic Alternative) are included, and for Basic Alternative plans, deductible bins of \$0, \$1 to \$50,\$51 to \$100 ..., are included. In columns 1 and 4, controls are included separately for type of basic plan and deductible in both 2006 and 2009. Indicators for pricing below the LIS benchmark are also included, separately for 2006 and 2009. Heteroskedasticity robust standard errors, clustered at the firm level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.2: Response to Contemporaneous and Previous Prices

	(1) $\ln s_{2009}$	(2) $\ln s_{2009}$	(3) $\ln s_{2009}$	(4) $\ln s_{2009}$
Premium in 2009	-0.0142	-0.0557***	-0.0137	-0.0382***
	(0.0233)	(0.0190)	(0.0126)	(0.00847)
Premium in 2008	0.0201	-0.0447	-0.00242	-0.00259
	(0.0242)	(0.0289)	(0.0267)	(0.00873)
Premium in 2007	0.0168	0.0155	0.0308	0.0319*
	(0.0403)	(0.0462)	(0.0253)	(0.0173)
Premium in 2006	-0.0630***	-0.0603**	-0.0706***	-0.0437**
	(0.0179)	(0.0290)	(0.0126)	(0.0164)
Type of Basic Plan: all lags	Yes	Yes	Yes	Yes
Firm Fixed Effects	No	Yes	No	Yes
Enhanced Plans Included?	No	No	Yes	Yes
N	308	308	878	878
R^2	0.798	0.893	0.576	0.831

OLS regression. Dependent variable: log of plan market share for non-LIS enrollees in a year. Sample in columns 1 and 2: basic PDP plans that were introduced in 2006, and that do not attrit or switch to or from enhanced benefit type before 2009. Sample in columns 3 and 4: all plans that do not attrit from the sample before 2009. Plans are dropped from the regression if they have fewer than 10 total enrollees. See Appendix Section A.3 for more details. In all columns, state fixed effects and benefit type indicators (Defined Standard, Actuarially Equivalent Standard, or Basic Alternative) are included for a plan's characteristics in each year, and for Basic Alternative plans, deductible bins of \$0, \$1 to \$50,\$51 to \$100 ..., are included for each year. In columns 3 and 4, indicators for enhanced plan and level of deductible are also included. Indicators for pricing below the LIS benchmark are also included, separately for each year. Heteroskedasticity robust standard errors, clustered at the firm level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Balanced Covariates at Cutoff

	Is AE	Is DS	Deductible	Is AE	Is DS	Deductible
	Bandwidth \$10			Bandwidth \$4		
Below Benchmark, 2006	-0.0776	0.0343	-2.038	-0.148	0.137*	-2.673
	(0.15)	(0.095)	(44.8)	(0.10)	(0.079)	(34.2)
Premium - Subsidy, 2006						
... Below Benchmark	0.0334	-0.0111	2.536	-0.0197	0.0286	-5.439
	(0.025)	(0.010)	(5.07)	(0.030)	(0.026)	(5.46)
... Above Benchmark	0.00829	-0.00154	5.920	0.00420	0.0229	10.40
	(0.024)	(0.016)	(5.72)	(0.045)	(0.023)	(6.64)
N | R^2	593 | 0.08	593 | 0.02	593 | 0.04	306 | 0.01	306 | 0.01	306 | 0.01
	Bandwidth \$6			Bandwidth \$2.50		
Below Benchmark, 2006	-0.101	0.0805	2.926	-0.182*	0.0560	-29.44
	(0.11)	(0.072)	(35.7)	(0.10)	(0.074)	(32.3)
Premium - Subsidy, 2006						
... Below Benchmark	0.0231	0.00122	0.383	-0.0928	0.0525	-10.37
	(0.028)	(0.013)	(4.14)	(0.058)	(0.058)	(13.5)
... Above Benchmark	-0.00129	0.00926	8.340*	0.0308	-0.0821	-16.05
	(0.030)	(0.020)	(4.51)	(0.067)	(0.049)	(16.5)
N | R^2	421 | 0.03	421 | 0.01	421 | 0.02	193 | 0.02	193 | 0.03	193 | 0.01

OLS regression. Dependent variables: Is AE =1 if plan is an Actuarially Equivalent basic plan. Is DS =1 if plan is a Defined Standard basic plan. Deductible: is each plan's yearly deductible. Sample: basic PDP plans with premiums within the bandwidth window in 2006. Heteroskedasticity robust standard errors, clustered at the firm level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.4: Attrition by Benchmark Status in 2006

	Fraction of Plans with Last Year of:				
		2006	2007	2008	2009
$10 Window	Below Benchmark 2006:	0.197	0.204	0.253	0.305
	Above Benchmark 2006:	0.034	0.034	0.231	0.383
$6Window	Below Benchmark 2006:	0.077	0.082	0.142	0.219
	Above Benchmark 2006:	0.004	0.004	0.235	0.370
$4 Window	Below Benchmark 2006:	0.045	0.053	0.120	0.211
	Above Benchmark 2006:	0.006	0.006	0.254	0.382
$2.50 Window	Below Benchmark 2006:	0.048	0.048	0.120	0.217
	Above Benchmark 2006:	0.000	0.000	0.236	0.336

A plan attrits if it is terminated or merged into another plan.

Table A.5: LIS Benchmark Status in 2006 Interacted with Subsequent Benchmark Status

$\ln s_t$	2007	2008	2009	2010
	Polynomial with controls, bandwidth \$4			
Benchmark or de minimis in:	(omitted category: not in 2006 or current year)			
...2006 and current year	2.301***	3.279***	2.198***	1.379***
	(0.409)	(0.674)	(0.270)	(0.383)
...2006 but not current year	0.906**	1.000**	0.728*	-0.0454
	(0.349)	(0.438)	(0.359)	(0.244)
...current year but not 2006	0.623**	2.187***	1.628***	1.471***
	(0.262)	(0.391)	(0.283)	(0.225)
Premium in Current Year	-0.0529**	0.0254*	-0.0561**	-0.0496***
	(0.0245)	(0.0144)	(0.0219)	(0.00752)
Premium - Subsidy, 2006	Quadratic	Quadratic	Quadratic	Quadratic
N	299	298	246	212
R^2	0.743	0.697	0.815	0.896

OLS regression. Dependent variable: log of total plan market share (including LIS enrollees) in a year. Sample: basic PDP plans with premiums within the bandwidth window ($4 on either side of the benchmark) in 2006. In "Polynomial with controls", regressions include state and firm fixed effects, and 2006 benefit type indicators (Defined Standard, Actuarially Equivalent Standard, or Basic Alternative). For Basic Alternative plans, deductible bins of $0, $1 to $50, $51 to $100 ..., are included. Premium minus subsidy is included as a polynomial separately above and below the benchmark. Heteroskedasticity robust standard errors, clustered at the firm level, are in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

Table A.6: Effect of LIS Benchmark Status in 2006 on Premiums in Later Years

Premium - Subsidy	2007	2008	2009	2010
	Local linear, bandwidth \$6			
Below Benchmark, 2006	-1.172	1.202	-0.0503	-1.737
	(0.907)	(1.891)	(1.841)	(2.486)
Premium - Subsidy, 2006				
... Below Benchmark	0.119	0.162	0.159	-0.185
	(0.282)	(0.282)	(0.218)	(0.332)
... Above Benchmark	-0.247	0.787	0.735	0.541
	(0.452)	(0.577)	(0.567)	(0.869)
N	329	277	203	182
R^2	0.013	0.017	0.046	0.024
	Polynomial with controls, bandwidth \$6			
Below Benchmark, 2006	-0.676	-1.307	-0.505	-4.767
	(1.371)	(1.305)	(1.480)	(3.313)
Premium - Subsidy, 2006	Quadratic	Quadratic	Quadratic	Quadratic
N	329	277	203	182
R^2	0.627	0.684	0.638	0.601

OLS regression. Dependent variable: monthly PDP premiums minus state-specific subsidy. Sample: basic PDP plans with premiums within the bandwidth window ($6 on either side of the benchmark) in 2006. In "Polynomial with controls", regressions include state and firm fixed effects, and 2006 benefit type indicators (Defined Standard, Actuarially Equivalent Standard, or Basic Alternative). For Basic Alternative plans, deductible bins of $0, $1 to $50, $51 to $100 ..., are included. Premium minus subsidy is included as a polynomial separately above and below the benchmark. Heteroskedasticity robust standard errors, clustered at the firm level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.7: Effect of LIS Benchmark Status in 2006 on Probability of Pricing Below the Benchmark in Later Years

Pr(Below Benchmark or De Minimis)	2007	2008	2009	2010
	Local linear, bandwidth \$2.50			
Below Benchmark, 2006	-0.0689	-0.110	-0.165	-0.0984
	(0.0881)	(0.154)	(0.101)	(0.0765)
Premium - Subsidy, 2006				
... Below Benchmark	-0.0232	-0.0502	-0.0607	-0.00564
	(0.0447)	(0.0546)	(0.0645)	(0.0420)
... Above Benchmark	-0.123*	-0.106	-0.128	-0.108*
	(0.0708)	(0.0856)	(0.0873)	(0.0550)
N	189	189	157	138
R^2	0.027	0.020	0.033	0.019
	Polynomial with controls, bandwidth \$2.50			
Below Benchmark, 2006	0.0105	-0.0146	-0.118	-0.00212
	(0.0858)	(0.0514)	(0.160)	(0.204)
Premium - Subsidy, 2006	Quadratic	Quadratic	Quadratic	Quadratic
N	189	189	157	138
R^2	0.768	0.778	0.679	0.584

OLS regression. Dependent variable: =1 if plan prices below the benchmark or is classified as a de minimis plan by CMS, =0 if else. Sample: basic PDP plans with premiums within the bandwidth window (\$2.50 on either side of the benchmark) in 2006. In "Polynomial with controls", regressions include state and firm fixed effects, and 2006 benefit type indicators (Defined Standard, Actuarially Equivalent Standard, or Basic Alternative). For Basic Alternative plans, deductible bins of \$0, \$1 to \$50, \$51 to \$100 ..., are included. Premium minus subsidy is included as a polynomial separately above and below the benchmark. Heteroskedasticity robust standard errors, clustered at the firm level, are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.8: Fraction of Plans that Attrit, by Cohort and Year

Fraction Merged or Terminated					
	Year of Plan Introduction (Cohort)				
	2006	2007	2008	2009	2010
Year merged or terminated					
2006	0.15				
2007	0.07	0.23			
2008	0.06	0.12	0.18		
2009	0.09	0.10	0.09	0.06	
Available until 2010	0.62	0.56	0.73	0.94	1.00

Fraction Terminated					
	Year of Plan Introduction (Cohort)				
	2006	2007	2008	2009	2010
Year terminated:					
2006	0.001				
2007	0.008	0.114			
2008	0.005	0.062	0.173		
2009	0.008	0.011	0.015	0.029	
Never Terminated as of 2010	0.98	0.81	0.81	0.97	1.00

Unit of observation is a plan offered in a PDP-region in a year (state or group of states). A plan is merged if its unique identifier leaves the data, but is combined into another plan.

Table A.9: Medicare Part D Premiums By Plan Age: Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(Monthly Premium)				Monthly Premium ($)	
Year of Plan Existence						
...2nd Year	0.0382	-0.0324	-0.00244	0.0250	1.201	0.480
	(0.0234)	(0.0226)	(0.0491)	(0.0766)	(1.627)	(0.897)
...3rd Year	0.0925*	-0.00625	0.0340	0.167**	2.678	2.562***
	(0.0505)	(0.0341)	(0.0723)	(0.0738)	(2.064)	(0.711)
...4th Year	0.112	-0.00734	0.0820	0.219**	5.290***	4.711**
	(0.0739)	(0.0618)	(0.120)	(0.102)	(1.573)	(1.786)
...5th Year	0.143**	0.0640	0.119	0.229*	4.495**	3.999*
	(0.0618)	(0.0506)	(0.163)	(0.123)	(1.706)	(2.057)
Type of Plan	Yes	Yes	Yes	Yes	Yes	Yes
Additional Fixed Effects	Firm x Year	Firm x Year	Firm	Firm	Firm	Firm
Weighting	Equal	Enrollment	Equal	Enrollment	Equal	Enrollment
Includes Enhanced Plans	No	No	Yes	Yes	No	No
N	4,276	4,123	8,382	8,185	4,276	4,123
R^2	0.782	0.863	0.475	0.609	0.418	0.695

All regressions include state fixed effects interacted with year fixed effects. Controls for type of basic plan include benefit type indicators (Defined Standard, Actuarially Equivalent Standard, or Basic Alternative) interacted with year fixed effects. For basic alternative and enhanced plans, controls for deductible in bins of $0, $1 to $50,$51 to $100 ..., are also included and interacted with year fixed effects. In regressions with enhanced plans, indicators for enhanced benefit type are also included. Heteroskedasticity robust standard errors, clustered at the firm level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.