Sinking, Swimming, or Learning to Swim in Medicare Part D*

Jonathan D. Ketcham† Claudio Lucarelli‡ Eugenio J. Miravete§

M. Christopher Roebuck¶

September 5, 2011

Abstract

Under Medicare Part D, senior citizens choose prescription drug insurance offered by numerous private insurers. We examine non-poor enrollees' actions in 2006 and 2007 using panel data. Our sample reduced overspending by $298 on average, with gains by 81% of them. The greatest improvements were by those who overspent most in 2006 and by those who switched plans. Decisions to switch depended on individuals' overspending in 2006 and on individual-specific effects of changes in their current plans. The oldest consumers and those initiating medications for Alzheimer's disease improved by more than average, suggesting that real-world institutions help overcome cognitive limitations.

JEL Codes: D01, D8, H51, I10, I11, I18.

---

* We thank Kosali Simon for generously sharing data, and Kenneth Hendricks, V. Kerry Smith, and Richard Hirth for their comments on earlier versions of the manuscript. Meredith Lewis of CVS Caremark provided invaluable help with the claims data and CMS files. We also thank the editor and three referees for comments that substantially improved the article, as well as those from seminar audiences at Keio University, the University of Texas School of Law, McCombs School of Business conference on "Innovation in Health Care Delivery Systems," the 8th Invitational Choice Symposium, the 2010 ESSET Conference, the 2010 ASHEcon Conference, and the III FTC-Northwestern University Microeconomics Conference. Christopher Roebuck was formerly with CVS Caremark as the director of health economics.

† Arizona State University, W.P. Carey School of Business, Department of Marketing, Box 874106, Tempe, AZ 85287-4106, Phone: (480) 965-5507, E–mail: *Ketcham@asu.edu*

‡ Cornell University, Department of Public Policy and Management, 105 MVR Hall, Ithaca, NY 14853; Leonard Davis Institute of Health Economics, Philadelphia, PA; and Universidad de los Andes, Chile, Phone: (607) 255-1948, E–mail: *Claudio.Lucarelli@cornell.edu*

§ The University of Texas at Austin, Department of Economics, 1 University Station C3100, Austin, Texas 78712-0301; and CEPR, London, UK. Phone: (512) 232-1718. E–mail: *miravete@eco.utexas.edu*; http://www.eugeniomiravete.com

¶ University of Maryland, Baltimore County, Department of Public Policy. E–mail: *mcr@rxeconomics.com*

Abstract

Under Medicare Part D, senior citizens choose prescription drug insurance offered by numerous private insurers. We examine non-poor enrollees' actions in 2006 and 2007 using panel data. Our sample reduced overspending by $298 on average, with gains by 81% of them. The greatest improvements were by those who overspent most in 2006 and by those who switched plans. Decisions to switch depended on individuals' overspending in 2006 and on individual-specific effects of changes in their current plans. The oldest consumers and those initiating medications for Alzheimer's disease improved by more than average, suggesting that real-world institutions help overcome cognitive limitations.

"The new Medicare Part D prescription drug insurance market illustrates that leaving a large block of uninformed consumers to sink or swim, and relying on their self-interest to achieve satisfactory outcomes can be unrealistic."

Nobel Laureate Daniel McFadden

2006 Presidential Address to the American Economic Association

1 Introduction

Medicare Part D stands at a remarkable crossroads of the current political and academic debates. Beginning in 2006, Part D expanded Medicare beneficiaries' access to prescription drug coverage by allowing them to choose among competing private insurance plans. Underlying this approach is the perspective that private competition within a regulated framework would strike the appropriate balance between cost control and providing value to consumers. Within Part D, costs are controlled through private insurers negotiating prices with pharmaceutical manufacturers in conjunction with demand-side cost-sharing such as copays and the "doughnut hole."[1] At the same time, competition for enrollees would incentivize insurers to design plans that were attractive to individuals. To limit adverse selection and encourage participation, the program provided large subsidies even to non-poor beneficiaries and penalties for those who did not enroll when they became initially eligible.[2]

Although many of the widely-publicized opinions of the program were initially pessimistic, the growing evidence from research on Part D is largely positive, with high participation, expanded prescription drug use, lower out-of-pocket prices for drugs, high consumer satisfaction, and total program costs below projections. One contentious remaining aspect of Part D is whether its reliance

[1] Consumers who do not receive federal low-income subsidies enter the "doughnut hole" and pay 100% of drug costs out of pocket once the total drug spending exceeds a threshold, and they exit it once their spending reaches another threshold. In 2006, these thresholds for the standard plan design were $2,250 and $5,100, respectively.

[2] Beneficiaries could avoid this penalty by immediately enrolling in a Part D plan, a Medicare Advantage plan, or an employer-sponsored plan meeting the creditable coverage criteria, which require the plan to have an actuarial value meeting or exceeding the regulated standard plan.

on competition between many private insurers is too complex for beneficiaries to navigate despite the program's regulations and subsidies.

Part D is an important context to study consumer choice due to the large number of lives and dollars involved. It offers an environment rich for testing theories of individual conduct that will inform policy in health care as well as other markets. Because of the age and prevalence of illness of the Medicare population, cognitive limitations may be common among this population. This may be compounded by the fact that prescription drug insurance plans are multiattribute, with some aspects uniform across individuals, such as premiums, while others vary across individuals due to differences in individuals' prescription drug use in conjunction with plans' use of formularies. Finally, because this was a newly created market, all eligible beneficiaries confronted these complex choices for the first time in late 2005. This last feature is econometrically attractive because analyzing data that include 2006 overcomes the initial condition problem raised by Heckman (1981), allowing researchers to separate the effects of aging from the effects of experience.

Behavioral economists have reported a number of biases in consumer decision making, such as inertia and confusion, particularly when cognition is limited by age, illness, or limited attention, e.g., Lusardi, Mitchell, and Curto (2009) and Agarwal, Driscoll, Gabaix, and Laibson (2009a). These contributions have raised a number of academic questions about how to best model and predict decision making. They also evoke calls for a number of regulatory reforms such as stronger consumer protection rules and simplification of credit cards, mortgages, retirement plans and health insurance contracts. In Part D specifically, many observers have called for a reduction in the number of plans available to consumers, which ranged across regions from 27 to 52 plans available in 2006 and 45 to 66 available in 2007.[3]

[3] Liebman and Zeckhauser (2008) perhaps best summarize the mistrust in consumers' ability to choose: "health insurance is too complicated a product for most consumers to purchase intelligently," they state, concluding that, "[i]t is unlikely that most individuals will make sensible decisions when confronted with these choices." Based on these premises they suggest that either a public agency or some private company should mediate consumers' health insurance purchases. Similarly, Hoadley (2008) surveys a panel of medical experts to call for a standardization of plan benefits and formularies to make them easier to compare as well as reducing the number of plans available. Hanoch,

In this article we present evidence about how a subset of consumers' actual plan choices in Medicare Part D evolved from 2006 to 2007. To achieve this, we analyze a large data set from a single insurer that sells Part D plans (PDPs) and administers PDPs sold by other companies. These data report individuals' chosen and available plans, prescription drug use and spending, and other characteristics in 2006 and 2007. Although our data incorporate limitations whose consequences we investigate in detail below, to our knowledge they are the only large-scale panel data made available to researchers to study Part D plan choice and switching decisions.

Two previous articles also rely on large data sets, both finding that consumers inappropriately weight various attributes of the plans in ways that cause them to choose plans that do not minimize their costs for drugs and drug insurance. In the first, Heiss, McFadden, and Winter (2007) conclude that enrollees' choices can be best understood as myopic, relying on static expectations and using only current drug expenditures. They find that beneficiaries appear to be rational in their decisions about whether to participate in Part D, but less so in their plan choices conditional on participation. In the second article, Abaluck and Gruber (2011) conclude that enrollees show inconsistencies with optimizing behavior because they overweight some features of the plan, such as premiums and doughnut hole coverage, while neglecting others. They thus conclude that elders fail to make choices consistent with optimization under full information. Consistent with these conclusions, Kling, Mullainathan, Shafir, Vermeulen, and Wrobel (2011) report that older adults suffer from serious misperceptions of prices and other features of Part D plans. According to their evidence, people's misperceptions result in them choosing prescription plans that are substantially more expensive than the available alternatives. One implication is that greater choice and competition harms rather than improves welfare because plans can flourish by

Rice, Cummings, and Wood (2009) reached the same conclusion after analyzing the experimental evidence of 192 subjects (half of whom were age 65 or older) who performed hypothetical enrollment decisions. Duggan, Healy, and Scott-Morton (2008), Goldman and Joyce (2008), Joyce, Goldman, Vogt, Sun, and Jena (2009) and Heiss, McFadden, and Winter (2009) are notable exceptions to the opinion of the majority of researchers in this area.

promoting confusion rather than by designing products that meet the preferences of fully-informed, unbiased consumers.

With the exception of Kling et al. (2011), these prior articles analyzed only 2006 and did not consider how consumers' actions changed over time. Yet in contrast to the often cross-sectional and typically laboratory-based evidence of common biases in consumer choice, summarized in DellaVigna (2009), other work suggests that markets and market experience ameliorate those biases, e.g., Miravete (2003), List (2003), List (2004), List (2006), and List and Millimet (2008).[4] Considering the dynamic aspects of consumer choice is important to evaluate Part D given the large degree of heterogeneity across consumers and plans as well as the related, but distinct, facts that this was a new market and that consumers had no previous experience in it. Further, using detailed, individual-level panel data of actual choices in a complex, high stakes context will add to economists' understanding of consumer choice more broadly.

Our primary focus in this article is on whether Medicare Part D enrollees improved over time in terms of reducing overspending. We define overspending as the consumers' annual *ex post* out-of-pocket (OOP) costs for insurance and prescription drugs above the cost of the cheapest alternative, where the alternatives include other Part D plans as well as having no coverage. Analyzing changes in this aspect of choice alone is insightful given the high persistence in an individual's drug spending over time,[5] and our use of panel data eliminates the effects of individual-specific, time-invariant attributes, including risk aversion. In this article we exclusively focus on those who did not receive federal low-income subsidies in either year due to differences in Part D's design across income levels. Specifically, those with low-income subsidies can switch plans multiple times throughout the year,

[4] Choi, Laibson, and Madrian (2010) present experimental evidence that individuals pay more attention to some attributes than others when presented with multiattribute financial contracts. In contrast to financial contracts, beneficiaries of Part D may revise their enrollment and choice decisions on a regular basis. We are not aware of any existing evidence, experimental or otherwise, about whether such biases persist over time.

[5] See Pauly and Zeng (2004) for one recent example of this widely-documented fact for the Medicare population.

they receive much larger premium supports, and they face small, regulated copayments for drugs on their plans' formularies.[6]

The results from our sample provide robust evidence of large reductions in overspending from 2006 to 2007, with average reductions in overspending of $298, which is 55% of the 2006 level. In addition to these large average effects, we find substantial heterogeneity, with 81% of the study sample lowering their overspending. The greatest reductions were achieved by those who overspent most in 2006, and this cannot be explained by changes in observed health. Interestingly, the improvements were greatest among those age 85 and above, and those initiating medications for Alzheimer's disease in 2007 improved by more than average. Although we cannot observe the choice process itself, these two results suggest that populations with greater prevalence of cognitive limitations are helped by various sources including family members, health care providers and other private organizations, and decision support tools such as online plan finders.

Among our sample, switching plans was a primary source of improvement, although we also find evidence of improvement among those who did not switch plans. Our analysis of switching decisions reveals that this set of consumers responds to financial incentives in deciding whether to remain in their current plan. Specifically, consumers' likelihood of switching increases greatly with their overspending in their current plan, indicating that their decisions incorporate past performance. Switching was also greater for those whose current plans would become more expensive in 2007 relative to the available alternatives, indicating that consumers factor in forward-looking information as well. Finally, we find evidence that consumers' switching decisions depend on their individual-specific aspects of the levels and changes of their 2006 plans and not simply on readily observable factors that affect everyone equally, such as premiums.

To place our results in the broader context, it is worth emphasizing that our data come from only a single firm that sells and administers a subset of Part D plans. As such, our results

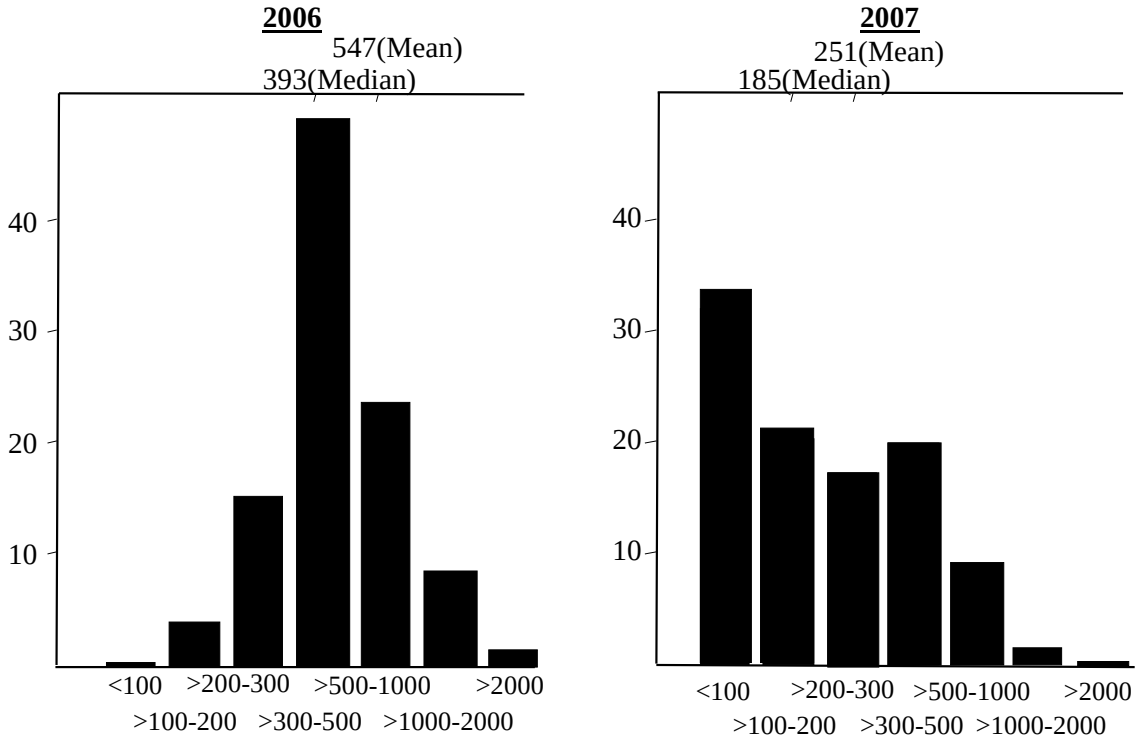
[6] See Duggan et al. (2008) for a further discussion of these issues.

are informative and suggest what may have occurred among the non-poor population enrolled in PDPs overall, but they represent only a single piece of evidence in a burgeoning area of research. At various points in this article we provide additional insight about these limitations and their potential effects on the generalizability our results. Most notably, compared to the average across all PDPs, our in-sample plans improved more on average along some dimensions but worsened on average on other dimensions. On net, a simple analysis suggests that these changes are equivalent to 15% of the total average improvement that we observe among our sample, amounting to 33% of the improvement among those who did not switch between the plans in our sample but only 10% of the improvement of those who did switch. Similarly, we estimate that our sample's demographics may have led us to overestimate the average mean improvement of the overall population by 4% ($12). In addition to leaving open questions about the changes among enrollees in other plans in 2006 to 2007, our analysis also cannot provide insights to how consumers' choices in Part D have evolved since then.

Using data described below, Figure 1 provides descriptive results for the changes in overspending from 2006 to 2007. [7] During the initial year of the program most beneficiaries overspent by $300 to $500 dollars. The distribution of overspending has a long right tail, with 9.8% of our sample overspending by more than $1,000. This results in the mean overspending, $547 dollars, being almost 40% larger than the median overspending of $393 dollars. More important, however, is the contrast between the two panels of this figure. In 2007, just one year into the program, the distribution shifts left, with substantially more beneficiaries closer to the cost-minimizing choice and only 1.7% overspending more than $1000. Both mean and median potential savings in 2006 are cut in half in 2007, i.e., $251 and $185, respectively. Further, the variation in overspending within our sample fell sharply from 2006 to 2007, with the standard deviation dropping from $747 to $246.

[7] This analysis does not control for additional sources of heterogeneity such as medical conditions, age, or others; however, the subsequent analyses in this article do. All variables are defined in the next section.

Figure 1. Overspending by Year



In the remainder of this article we further explore these descriptive results and how they varied across sub-populations within our sample. Section 2 describes our data and how the key variables are defined. Section 3 evaluates how overspending changed between 2006 and 2007 overall, as well as the degree of heterogeneity in those changes across individuals in our sample. Section 4 reports how the changes vary with observed individual characteristics, such as overspending in 2006, age, sex, and levels and changes in health. In Section 5 we analyze the role of switching in the improvement we observe, and we examine the switching decision itself in Section 6. Section 7 follows with an analysis of the robustness under various alternative approaches. In Section 8 we discuss whether the experience of our sample is likely to generalize to the population of non-poor PDP enrollees as a whole. We conclude in Section 9.

2 Data

In this section we describe our data and provide some summary statistics of the individuals' and plans in our sample. We then summarize how we generated each consumer's spending in each plan available to them in each year. With these estimates, we construct overspending, the key measure of plan choice in our analysis. Finally, we discuss the use of *ex ante* versus *ex post* drug consumption to evaluate changes in overspending.

2.1 Study sample

We use data from the Centers for Medicare and Medicaid Services (CMS), CVS Caremark (a pharmacy benefits manager, hereafter "the PBM"), and ancillary sources to estimate the cost of each plan to each patient in both 2006 and 2007. The CMS data provide details about how each plans' design changed over time as well as which plans exited and entered each region. The PBM provided us with a large data set of its enrollees, for whom we observe their chosen plan with certainty, the universe of their claims, whether and which subsidy level they received. This information on subsidies allows us to exclude all of those who ever received a federal low-income subsidy. The PBM's enrollment information data allow us to identify which individuals switched plans, which plans they switched to, which they switched from, and for those who switched, how they would have fared in 2007 if they had stayed in their 2006 plans.

Our study sample includes 71,399 individuals who were enrolled in stand-alone Prescription Drug Plans (PDPs) provided or administered by the PBM for all of 2006 and 2007 and who did not receive a federal low income subsidy in either year. This sample excludes those not enrolled for the full year, as well as those who switched plans within the year, which this non-subsidy population could do if they moved between regions, among other reasons. Because our analysis required both years of data, we also cannot examine the decisions of those who switched into or out of the set

of plans offered by the PBM. Likewise, we exclude those who switched into or out of a Medicare Advantage (MA) plan.[8]

The data provide a number of individual-level characteristics including age, sex and time-varying health measures, as summarized in Table 1. The health measures are defined by a proprietary system known as Ingenix Pharmacy Risk Group variables (PRGs) and their corresponding calculated risk scores. The PRGs are a vector of 127 indicator variables for whether drugs were consumed for specific clinical indications. These indicators do not vary with the quantity or cost of drugs taken for a condition, but rather they simply indicate whether the individual took any drugs for each of the conditions. From these, we generated dummy variables indicating whether the individual took medications in each of the 15 most common PRGs in our sample. These fall into 9 broader illness categories that we include as dummy regressors in our analysis. They are, from more to less frequently observed: hypertension, cholesterol and other cardiovascular, pain, mental health, antibiotics, anticoagulants, thyroid conditions, diabetes and osteoporosis. We also include controls for a tenth illness, Alzheimer's disease, because of its link with cognitive impairment. The Ingenix risk scores are used by insurers to predict individuals' future prescription drug expenditures based on their prescription drug claims history and demographics. In our sample, the risk score has a mean of 4.4 in 2006 and 4.6 in 2007. To give a sense of scale, taking medications for Alzheimer's disease increases the risk score by 2.8, while taking hypertension medications increases it by 1.4.

One notable feature of our data is that we observe individuals enrolled in PDPs that the PBM sells directly, under the SilverScript brand, and those in PDPs administered by the PBM but sold under different names. These administrative agreements between insurers and the

[8] This is necessary because we cannot measure an individual's costs for prescription drugs and prescription drug insurance in those plans because the premiums include both drug coverage and coverage for other types of health care. Among the population of 214,896 individuals in either MA plans or PDPs that meet the other criteria for inclusion in our study, only 0.07% (i.e., 7 per 10,000) switched from an MA to a PDP from 2006 to 2007, while only 0.2% switched from a PDP to an MA plan from 2006 to 2007. Thus among this population, PDP and MA plans appear to function as separate markets, and excluding these few individuals from the sample of PDP enrollees will not meaningfully change the results.

Table 1. Characteristics of Individuals in the Sample

	Individuals in Study Sample (N=71,399)			20% Sample of All PDP Enrollees Meeting Criteria (N=1,168,712)			Difference in the Changes	Difference in the 2006 Levels
	2006	2007	Change	2006	2007	Change		
Female(%)	70.1	70.1	0.0	67.0	67.0	0.0	0.0	3.1
Age(%)								
65-69	20.4	16.2	-4.2	22.7	17.4	-5.3	1.1	-2.3
70-74	19.8	20.1	0.3	23.7	24.5	0.8	-0.5	-3.9
75-79	20.8	20.6	-0.2	21.3	21.7	0.4	-0.6	-0.5
80-84	20.2	20.6	0.4	16.4	17.6	1.2	-0.8	3.8
85 up	18.7	22.5	3.8	16.0	18.8	2.8	1.0	2.7
Gross drug spending	$2,568	$3,002	$434	$2,731	$2,978	$247	$186.7	-$162.7
Ingenix Pharmacy Risk Score	4.4	4.6	0.2	--	--	--	--	--
Took medication for(%)								
Hypertension	77.4	78.9	1.5	--	--	--	--	--
Cholesterol and other cardiovascular	51.8	53.6	1.8	--	--	--	--	--
Pain	27.2	29.8	2.6	--	--	--	--	--
Mental health	28.0	26.2	-1.7	--	--	--	--	--
Antibiotics	40.6	42.9	2.3	--	--	--	--	--
Anticoagulants	21.7	23.0	1.3	--	--	--	--	--
Thyroid	17.5	18.4	0.9	--	--	--	--	--
Diabetes	19.1	20.0	1.0	--	--	--	--	--
Osteoporosis	17.8	17.0	-0.8	--	--	--	--	--
Alzheimer's	4.1	5.5	1.4	--	--	--	--	--

NOTE: The data for all PDP enrollees meeting the study eligibility criteria were acquired through personal communication with the CMS.

PBM are rarely known to individuals and are likely to be unrelated to any important individual characteristics. As a result, our study sample incorporates not only those who chose a SilverScript plan but also those who in essence were randomly assigned to be administered by the PBM.[9] The right side of Table 1 shows comparable statistics provided by the CMS for a random 20% sample of all individuals that met our study inclusion criteria, i.e. the non-poor that were continuously enrolled in a PDP for all of 2006 and 2007. The results show some differences between our sample and the overall population. Our sample has some small differences from the overall population, with a higher prevalence of females, a lower share age 65-69 and 70-74 and a higher share age 80-84 and 85 and above. Our sample also had slightly lower gross drug spending in 2006 but slightly higher

[9] Additionally, the PBM is prohibited from designing these other plans or negotiating prices for them; they may only administer the claims.

gross spending in 2007. In the second to last column, the table also shows some small differences in the changes from 2006-2007. Because the CMS does not use the Ingenix PRG system, we cannot derive comparable measures of health for the full population.

The study sample includes 95 different PDPs in 2006 and 154 in 2007, where only 19% of the sample was enrolled in the SilverScript plans.[10] Similar to all available plans, those in our sample vary greatly in their number of enrollees. The average attributes of our in-sample plans, and the standard deviation across plans in our sample, are shown in Table 2. For comparison, the right hand side of Table 2 shows the levels and changes among PDPs overall, indicating some differences from our sample plans in terms of premiums, deductibles and formulary coverage. The changes in the means indicate that the plans in our study sample became relatively more generous in some aspects with relatively larger reductions in average premiums and deductibles, and greater increase in the prevalence of doughnut hole coverage for generic drugs. At the same time, our study plans became relatively less generous on average with greater relative increases in the share of common drugs on the formulary, requiring prior authorization, and with OOP prices above $20. The standard deviations in Table 2 also indicates divergence among the in-sample plans for deductibles, doughnut hole coverage and other enhancements, but convergence in terms of premiums and formulary coverage. In contrast, PDPs overall diverged from 2006 to 2007 along all dimensions excluding formulary coverage. After we present our results, in Section 8 we return to these differences in our sample's individuals and plans to consider their implications for the generalizability of our results.

[10] As is the norm for PDPs, the sample sizes reflect the number of plan-region combinations. The sample includes 9 different plan names in 2006 and 18 different plan names in 2007, but a given plan name may have different attributes in each region.

Table 2. Part D Plan Characteristics

	Plans in Study Sample		All Part D Plans		
	Mean	Standard Deviation	Mean	Standard Deviation	Difference
2006					
Deductible	$161.58	$74.18	$92.23	$116.06	$69.35
Annual Premium	$542.33	$164.84	$446.10	$151.82	$96.23
Number of the Top 100 Drugs					
On the Formulary	95.74	2.91	93.44	6.63	2.30
Requiring Prior Authorization	5.45	4.01	9.61	9.09	-4.16
With OOP Prices <$20	70.30	17.69	61.38	13.19	8.92
"Doughnut hole" coverage for generics	0.00	0.00	0.13	0.34	-0.13
"Doughnut hole" coverage for brands	0.00	0.00	0.02	0.15	-0.02
Enhanced plan	0.00	0.00	0.43	0.50	-0.43
Observations	95		1,431		
2007					
Deductible	$99.77	$108.35	$88.83	$120.20	$10.94
Annual Premium	$452.38	$138.09	$436.91	$183.33	$15.47
Number of the Top 100 Drugs					
On the Formulary	89.53	2.69	91.46	5.68	-1.93
Requiring Prior Authorization	4.96	1.35	2.35	2.98	2.61
With OOP Prices <$20	61.59	1.73	64.24	5.13	-2.65
"Doughnut hole" coverage for generics	0.31	0.46	0.25	0.43	0.06
"Doughnut hole" coverage for brands	0.00	0.00	0.05	0.21	-0.05
Enhanced plan	0.52	0.50	0.49	0.50	0.03
Observations	154		1,804		

NOTE: The plan is identified by the plan ID, which is unique for each region. The values for all Part D plans are from those that had positive enrollment.

2.2 Measuring overspending

Our focus in this article is to evaluate how the chosen plan compares to the person's cheapest alternative, where we refer to the difference in the individual's costs between the two as "overspending". To calculate this, we use data from the PBM on the individual's prescription drug claims, enrollment decisions, region of residence, and subsidy level in conjunction with CMS data on the characteristics of each plan available in each person's region in each year. These characteristics include premiums, deductibles, doughnut hole coverage, and formularies, which indicate what the beneficiaries' costs would be for every possible drug.

The total cost of each alternative for each person is the sum of the OOP prescription drug costs and the plan premiums. We provide extensive details about how the OOP prescription drug costs were calculated for each alternative in Appendix A. Briefly, we relied on the CMS formulary files to determine the copayment each individual would have paid for each drug he consumed under every PDP available in his region. When the formulary file provided a coinsurance rate, or when the individual was below the deductible or in the doughnut hole, we determined the relevant price in each plan from Wolters Kluwer Health data or from data collected from the CMS plan finder website.

Throughout the article, we consider "no insurance" as one of the options that individuals may choose, although we validate the robustness of the results when this option is excluded in Section 7. For the OOP prescription drug costs without insurance, we relied on the usual, customary price provided by CVS. For the main analysis presented in this article we assume a price elasticity of demand for prescription drugs of -0.54 to allow the total amount of drugs consumed by a beneficiary to vary under different marginal drug prices.[11] To check the sensitivity of the results to our elasticity assumption, we alternatively assumed perfectly inelastic demand, holding constant the total prescription drug consumption across all options at the level reported in our PBM data itself. These results are reported in Section 7.

2.3 *Ex ante* and *ex post* costs

To analyze how consumers' choices change over time in a context with uncertainty, researchers must adopt a perspective on what information consumers use to make their decisions. Conceptually, consumers may rely on several different sets of information about future drug consumption. At

[11] This is the drug price arc-elasticity estimate for Medicare beneficiaries obtained by Shea, Terza, Stuart, and Briesacher (2007). This is likely the most accurate estimate for this population, and it is more elastic than the other estimates. Using the most elastic estimate is likely to yield the lowest reductions in overspending from 2006 to 2007, as it reduces the amount of overspending. Our results using an assumption of perfectly inelastic demand confirm this, as shown in Section 7. We elaborate on the reasoning in the appendix.

one extreme, they may be fully myopic and utilize only their current consumption, ignoring the possibility of future changes when making choices about which plans to enroll in. At the other end, they may have perfect information and no uncertainty, anticipating precisely how future drug consumption will change and incorporating that into their choices of plans. Between those are intermediate cases, such as consumers being aware of probabilities and expenses of various potential illnesses and purchasing plans accordingly but not knowing precisely whether or when they will acquire those illnesses.

All three of these cases are plausible. For both practical and conceptual reasons, in the primary analysis reported here we adopt the fully-informed approach in which consumers' choices for a given year are evaluated based on their actual drug consumption in that same year.[12] We refer to this as the *ex post* approach, while we refer to the other extreme as the *ex ante* approach. Conceptually important in our context, the *ex ante* approach eliminates the potential for learning, as consumers may have changed their drug consumption as they learned the intricacies of PDPs or gained familiarity with their own plans' formularies. These assumptions of the *ex ante* approach are rejected by Tchernis, Normand, Pakes, Gaccione, and Newhouse (2006), who present evidence consistent with consumers having private information regarding the evolution of their own health status. They find that consumers choose plans that are more generous for treatments which they might likely need in the near future. Practically speaking, we cannot rely on the *ex ante* benchmark to study within-person changes over time because we do not observe individuals' drug consumption in 2005. This precludes us from generating the 2006 cross-sectional estimates under this alternative approach, which are essential for evaluating the within-person changes over time. However, we compare the cross-sectional overspending results for 2007 from both the *ex post* and the myopic *ex ante* approaches in Section 7, and we find no notable differences between them.

[12] This is the same approach taken by Fang, Keane, and Silverman (2008, §5) when comparing the *ex post* total medical expenditures under basic Medicare (parts A and B) with and without Medigap supplemental insurance.

3 Within-Person Changes in Overspending

We begin by estimating first difference regressions of overspending for 2006 and 2007. Define ΔO_i as the within-person change in overspending for individual i, ΔH_i denotes a vector of variables for the within-person changes in the measures of health, specified above, and Δu_i are the changes in the idiosyncratic error. We estimate the following equation:

$$\Delta O_i = \alpha + \Gamma \Delta H_i + \Delta u_i. \tag{1}$$

The main coefficient of interest is α, which is interpreted as the average within-person change in overspending. In Table 3 we present estimates from four models to determine whether the results can be explained by changes in health. The first model does not control for changes in health, i.e., we exclude ΔH_i, but the second model does. In the third and fourth models, we analyze only the subset of consumers with stable health, which we define in two alternate ways. In the third column, we define people in stable health if their total risk score changed by less than 0.5. In the fourth column, we further required that they did not change for any of the ten individual conditions that we include. In the most basic model, average overspending was reduced in 2007 by almost $300, i.e., a reduction of 54% of the average overspending of $546 in the regression sample in 2006.

To show the degree of heterogeneity in the change in overspending, we also report various points in the distribution of the unadjusted within-person change for the sample included in each regression model. The bottom half of the table demonstrates the large amount of heterogeneity in these improvements. Overspending fell for 81% of the sample, with improvements ranging from a few dollars up to amounts exceeding $1,000. While 19% overspent more in 2007 than in 2006, the absolute value of their increases ($224) amounted to only slightly over half of the average decreases experienced by the 81% who improved ($419).

Table 3. First Difference Models of Within-Person Change in Overspending 2006-2007

	Health Controls		*Stable Health Only*	
	No	*Yes*	*Inclusive Definition*	*Narrower Definition*
Intercept	-295.97 [3.890] ***	-298.29 [4.131] ***	-266.01 [6.924] ***	-255.39 [9.268] ***
Observations	71,399	71,399	30,149	15,247
Mean Overspending in 2006	546.9	546.9	515.2	504.5
Within-person change in Overspending				
Mean	-296.0	-296.0	-266.0	-255.4
5th Percentile	-1,136.0	-1,136.0	-1,044.0	-991.3
10th Percentile	-766.4	-766.4	-682.0	-642.5
25th Percentile	-409.4	-409.4	-381.4	-364.8
50th Percentile	-236.7	-236.7	-210.6	-189.1
75th Percentile	-44.1	-44.1	-38.9	-38.7
90th Percentile	98.9	98.9	77.0	72.7
95th Percentile	235.8	235.8	188.3	147.6

NOTE: Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. The second, third and fourth models include controls for within-person changes in health.

The remaining columns of Table 3 indicate that these results are largely invariant to changes in individual health, with similar means and distributions of improvements across all four columns. The results in the second column are conditional on changes in individual's health. These show slightly larger reductions than the unconditional changes, with a mean improvement of $298, or 55% of the 2006 mean. Because of concerns about these results being due to changes in health, in the remainder of this article we continue to analyze both the full population conditional on changes in health, and the subset of those in stable health. Given the similarity of results with our two definitions of stable health, as shown in the third and fourth columns of Table 3, we rely on the more inclusive one (in column three) because of its larger sample size.

These results highlight the importance of analyzing how choices evolve over time, particularly in a new market, which itself was evolving to provide better information to consumers, and in which consumers have no prior experience. The (unconditional) mean overspending in 2006 by our regression sample was $546, i.e. a mean overspending of 36.5% percent of the total out of pocket costs. However, in 2007 overspending averaged 21.6% of total out of pocket costs, amounting to a 41% reduction in only one year. The results show that reductions in overspending are independent

of changes in individual medical conditions, but even among those in stable health the improvements vary substantially.

4 Heterogeneity of Improvement across Individual Characteristics

To investigate what explains the substantial degree of heterogeneity in improvement, we first expand the basic model used in the previous section to consider how the within-person changes vary by observed individual characteristics. Specifically, we estimate

$$\Delta O_i = \alpha + \Gamma \Delta H_i + \beta X_i + \Delta u_i, \tag{2}$$

where X_i includes a vector of observed time-invariant characteristics for individual i, and the remaining variables are as in equation (1). By including X_i in these models we allow the within-person change in overspending to vary with observed individual characteristics, including age in 2006 (65-69, 70-74, 75-79, 80-84, and 85 and above), health (the 2006 risk score, and indicators for whether the person took medication in 2006 for each of the 10 individual illnesses), and sex. In general, these coefficients provide evidence about the equity of Part D's current design by indicating whether some groups improved less, or even worsened, from 2006 to 2007. For example, the coefficients on age and on the levels and changes in taking medications for Alzheimer's disease provide evidence about whether the improvements were smaller among populations where cognitive impairment was more prevalent.[13] In a second version of the model, we also include in X_i a set of categorical variables for the extent of overspending in 2006 (less than \$100, \$100-\$200, \$200-\$300,

[13]To substantiate this interpretation, a recent review article on Alzheimer's disease in *The New England Journal of Medicine* characterizes it as "a deterioration of memory and other cognitive domains" that is "the most common form of dementia, accounting for 50-56% of cases at autopsy and in clinical series." The article continues by citing previous work showing that, "The incidence of the disease doubles every 5 years after 65 years of age ...," and, "... the odds of receiving the diagnosis of Alzheimer's disease after 85 years of age exceed one in three," (Querfurth and LaFerla (2010)).

$300-$500, $500-$1,000, $1,000-$2,000, more than $2,000). This provides evidence about whether overspending in 2006 persisted through 2007, or whether those who overspent most in 2006 also improved most in 2007. In a third version of the model, we also control for the level and change in gross drug spending in the actual plan, defined as the total of the amount spent on prescription drugs by the individual and the plan on the individual's behalf. Including these two additional variables provides evidence about whether any difference by 2006 overspending or the health measures are simply due to differences in total drug consumption.

Table 4 presents the results from these three first difference models. As before, we also estimated models on the subset in stable health but we do not report them given their similarity with the other results shown in Table 4. The results show that the change in overspending varies substantially with these observed characteristics. Improvements were greater for females and for those who were older, with those age 85 and above reducing overspending by $93 to $108 more than those age 65-69. Given the inclusion of categorical dummy variables in these models, the intercepts in these models indicate the expected within-person change among those in the omitted groups. For example, the results from the second model show increases in overspending of $295 for females age 65-69 who overspent by less than $100 in 2006, did not take or initiate medications for any of the 10 most common illnesses, and who had risk scores of 0 in both years.

The results also show that initiating medications for some of the most common medical conditions in 2007 led to significant reductions in overspending. Results from the second column indicate that beneficiaries who began treatment for hypertension, anticoagulants, osteoporosis and Alzheimer's reduced their overspending from 2006 to 2007 by $22, $37, $27 and $50 more than average, respectively. Further, the within-person changes in overspending among those who were taking medication for Alzheimer's in 2006 did not significantly differ from those not taking these medications. This, in conjunction with the results by age and the results showing that those who initiated medication for Alzheimer's in 2007 improved by more than average indicates that other

Table 4. First-Difference Models of Within-Person Change in Overspending 2006-2007, by Observed Individual Characteristics

2006-2007 Change Allowed to Vary with:	*Age, Sex, Levels and Changes in Health*	*And 2006 Overspending*	*And Levels and Changes in Gross Drug Spending*
Overspending Level in 2006 ($)			
less than 100		*Reference Category*	*Reference Category*
between 100 and 200		-163.51 [17.406] ***	-206.70 [46.436] ***
between 200 and 300		-254.67 [21.333] ***	-301.89 [50.344] ***
between 300 and 500		-408.61 [17.003] ***	-457.76 [53.532] ***
between 500 and 1000		-632.37 [16.969] ***	-644.07 [38.039] ***
between 1,000 and 2,000		-1298.78 [17.988] ***	-1229.90 [42.816] ***
more than 2,000		-3172.82 [205.489] ***	-2953.45 [111.632] ***
Age in 2006			
Age 65-69	*Reference Category*	*Reference Category*	*Reference Category*
Age 70-74	-42.49 [10.561] ***	-26.39 [9.083] ***	-31.02 [9.314] ***
Age 75-79	-63.07 [16.328] ***	-39.02 [15.701] **	-50.03 [16.336] ***
Age 80-84	-113.53 [9.332] ***	-87.49 [7.733] ***	-91.02 [11.234] ***
Age 85 up	-108.47 [8.668] ***	-94.01 [7.284] ***	-93.30 [13.535] ***
Male	13.36 [11.429]	26.90 [10.172] ***	26.96 [10.718] **
Risk Score in 2006	-40.85 [4.325] ***	-1.21 [3.159]	5.61 [34.246]
Took medication in 2006 for			
Hypertension	23.33 [10.227] **	11.21 [9.743]	16.02 [9.388] *
Cholesterol and other cardiovascular	-72.76 [11.853] ***	-22.05 [10.881] **	2.28 [14.430]
Pain	36.64 [11.117] ***	7.53 [9.930]	7.17 [9.953]
Mental health	20.54 [13.165]	19.08 [11.413] *	24.34 [12.737] *
Antibiotics	9.72 [9.417]	-3.72 [8.143]	5.61 [7.082]
Anticoagulants	-43.55 [10.891] ***	-16.16 [9.100] *	-15.31 [10.458]
Thyroid	0.60 [9.013]	11.65 [7.591]	11.38 [11.751]
Diabetes	-2.45 [13.535]	-0.99 [10.349]	-2.82 [17.528]
Osteoporosis	-14.36 [9.775]	-23.73 [7.617] ***	-13.63 [11.942]
Alzheimer's	17.88 [17.760]	-6.39 [13.276]	-37.98 [25.713]
Change in Risk Score	5.25 [5.685]	14.91 [5.241] ***	-14.10 [28.768]
Change in takes medication for			
Hypertension	-16.62 [13.402]	-22.12 [11.747] *	-38.77 [11.376] ***
Cholesterol and other cardiovascular	-14.91 [18.389]	1.62 [17.136]	-3.10 [18.785]
Pain	2.68 [8.382]	-7.33 [7.411]	-9.83 [7.421]
Mental health	2.42 [12.040]	3.56 [10.667]	23.71 [14.991]
Antibiotics	-5.60 [8.516]	-11.12 [7.726]	-9.96 [7.386]
Anticoagulants	-51.36 [15.017] ***	-36.72 [12.975] ***	-54.74 [15.437] ***
Thyroid	22.46 [13.256] *	14.59 [10.377]	1.86 [11.761]
Diabetes	-27.32 [39.774]	-43.22 [37.173]	-28.50 [29.665]
Osteoporosis	-27.07 [12.925] **	-26.57 [10.585] **	-32.20 [17.213] *
Alzheimer's	-11.59 [19.951]	-54.63 [16.813] ***	-84.81 [19.043] ***
2006 Gross Drug Spending			-0.04 [0.052]
Change in Gross Drug Spending			0.10 [0.070]
Intercept	-39.78 [9.761] ***	294.66 [19.681] ***	339.78 [50.496] ***
Observations	71,395	71,395	71,395

NOTE: Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1.

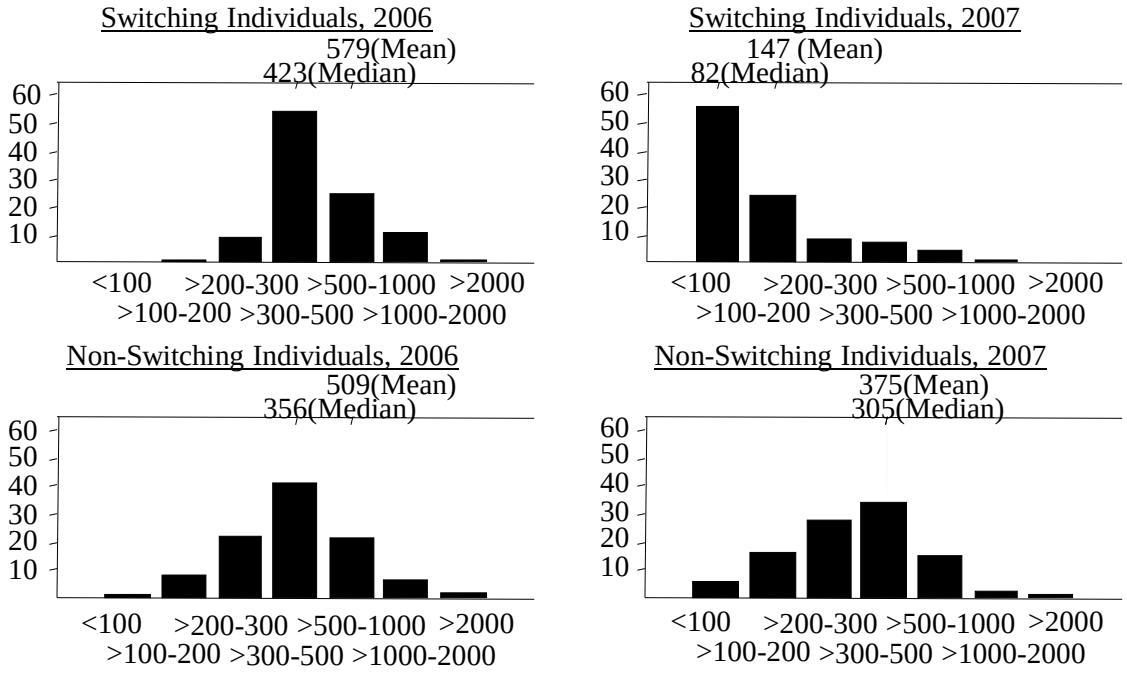
people are likely helping these populations choose plans. Although we cannot distinguish between the various plausible explanations, among them is that the new diagnosis of a condition such as Alzheimer's disease triggers greater involvement of family members and health care providers in the choice of prescription drugs and drug insurance. As shown in the third column, these results are not simply reflecting levels or changes in total drug consumption, as they persist even conditional on that, and in fact the estimated improvements by those who initiated medication for Alzheimer's disease are even larger.

The last two columns of Table 4 show a final set of important differences in within-person changes in overspending. Relative to those with overspending below $100 in 2006, all other groups of beneficiaries reduced their overspending substantially more. These reductions increased monotonically with the size of the 2006 overspending, with those overspending by more than $2000 reducing their overspending by an average of $2953 to $3173 more than those who overspent least in 2006. This is consistent with the results in Figure 1, which shows a substantial reduction in the share of the population overspending by more than $1000. This shift, as well as the large reductions in the mean and standard deviation of overspending, indicates that improvements in 2007 among those with the highest initial levels in overspending were not simply offset by increases in overspending among those who overspent least in 2006.

5 Switching and Improvement

In this section, we explore whether consumers' decisions to switch plans explains the large average reductions in overspending and the heterogeneity across individuals in these reductions. Figure 2 illustrates and compares the gains by switchers and non-switchers, using the same approach as Figure 1. The top two panels report the overspending by year for switchers. These show that switching was a primary contributor to improvement: in 2006, the distribution of overspending

Figure 2. Overspending by Year and Switching



has a thick right tail with a mean of $579. In 2007, however, those who switched overspent by an average of $147, and the distribution becomes highly concentrated on the left, with half of switchers overpaying by $82 or less in 2007. The two bottom panels of Figure 2 make evident that even non-switchers overpaid less in 2007 than in 2006. In 2007, beneficiaries who stayed in their 2006 plans reduced their mean overspending from $509 to $375. Finally, the figure also indicates that those who chose to switch had higher overspending in 2006 than those who did not. We explore this issue more in detail in Section 6.

To further analyze the effects of switching, we estimate two models. The first is identical to equation (1) except that we also allow the change from 2006 to 2007 to vary with whether the person switched or not, denoted as S_i,

$$\Delta O_i = \alpha + \Gamma \Delta H_i + \theta S_i + \Delta u_i. \tag{3}$$

The second is identical to equation (2) except that it includes S_i,

$$\Delta O_i = \alpha + \Gamma \Delta H_i + \beta X_i + \theta S_i + \Delta u_i. \tag{4}$$

As reported in Table 5, the results conditional on individual characteristics are similar to the changes in the unconditional means shown in Figure 2. Non-switchers reduced their overspending from 2006 to 2007 by \$137 on average, while switchers averaged an additional \$299 reduction for a total decrease of \$435. As the results from the full models in second and third columns indicate, even when the change is allowed to vary with a number of other observed characteristics, the difference between switchers and non-switchers is \$232-233. Thus, switching was a primary source of improvement.

While the coefficients on the switching dummy indicate the incremental savings due to switching, the models also impose the assumption that switchers, had they not switched, would have improved the same amount as non-switchers conditional on the other variables in the model. We implemented an additional analysis to explore the incremental savings due to switching without imposing this assumption. Specifically, we compared the amount that switchers would have spent in their 2006 plan had they remained in it for 2007 with the amount they actually spent in their new 2007 plan. These results show that 96.5% of switchers improved by doing so, with mean incremental savings of \$306.80, a median of \$279.90 and 5th and 95th percentiles of \$91.37 and \$605.13, respectively.

A comparison of the results for the other variables in the models reported in Tables 4 and 5 shows how switching moderates the other observed differences in the magnitude of the within-person changes from 2006 to 2007. Interestingly, the large differences across the 2006 overspending level remain, indicating that they exist for reasons in addition to differences in the choice to switch plans itself. Similarly, the larger reductions in overspending by those who initiated medication

Table 5. First-Difference Models of Within-Person Change in Overspending 2006-2007, by Switching and Other Observed Individual Characteristics

	Full Sample		*Subset with Stable Health Only*
2006-2007 Change Allowed to Vary with:	*Switching Plans and Changes in Health*	*And Other Characteristics*	*And Other Characteristics*
Switched plans	-298.46 [8.256] ***	-232.98 [7.279] ***	-231.97 [12.884] ***
Overspending Level in 2006 ($)			
less than 100			
between 100 and 200		-174.57 [17.424] ***	-170.37 [21.042] ***
between 200 and 300		-222.99 [21.514] ***	-196.72 [36.930] ***
between 300 and 500		-313.81 [17.135] ***	-291.67 [21.145] ***
between 500 and 1000		-547.74 [17.285] ***	-517.15 [21.687] ***
between 1,000 and 2,000		-1195.86 [18.646] ***	-1175.86 [25.325] ***
more than 2000		-3103.24 [206.809] ***	-2394.51 [434.065] ***
Age in 2006			
Age 65-69			
Age 70-74		-2.89 [9.216]	-8.58 [6.898]
Age 75-79		16.19 [16.810]	27.64 [34.832]
Age 80-84		-12.16 [8.519]	-11.10 [10.106]
Age 85 up		-3.66 [8.227]	0.48 [9.520]
Male		-3.91 [9.684]	14.23 [17.709]
Risk Score in 2006		0.91 [3.140]	2.91 [4.261]
Took medication in 2006 for			
Hypertension		12.65 [9.717]	17.21 [15.611]
Cholesterol and other cardiovascular		-20.69 [10.836] *	-42.26 [20.047] **
Pain		3.80 [9.887]	3.49 [14.172]
Mental health		13.81 [11.358]	2.15 [13.619]
Antibiotics		-10.13 [8.093]	-13.84 [13.390]
Anticoagulants		-18.37 [9.053] **	-31.80 [18.110] *
Thyroid		2.34 [7.549]	-7.36 [11.622]
Diabetes		5.24 [10.219]	-21.75 [17.385]
Osteoporosis		-25.41 [7.518] ***	-56.37 [11.554] ***
Alzheimer's		-18.18 [13.155]	-44.38 [22.236] **
Change in Risk Score	22.61 [6.416] ***	16.47 [5.245] ***	70.33 [39.000] *
Change in takes medication for			
Hypertension	-8.97 [12.528]	-19.72 [11.680] *	-0.02 [12.340]
Cholesterol and other cardiovascular	27.17 [18.239]	-1.58 [17.115]	-33.61 [14.589] **
Pain	-18.63 [6.255] ***	-12.28 [7.378] *	-5.56 [10.136]
Mental health	-25.19 [12.789] **	-16.77 [10.753]	0.36 [15.646]
Antibiotics	-9.04 [7.973]	-13.75 [7.719] *	-42.13 [17.352] **
Anticoagulants	-30.57 [13.929] **	-34.99 [12.880] ***	-24.04 [21.959]
Thyroid	6.86 [12.671]	9.72 [10.020]	-3.98 [21.133]
Diabetes	-37.35 [40.805]	-40.66 [36.886]	-4.69 [39.925]
Osteoporosis	-6.20 [12.621]	-26.19 [10.402] **	-48.30 [20.475] **
Alzheimer's	-44.92 [19.897] **	-58.64 [16.600] ***	-56.39 [27.195] **
Intercept	-136.90 [7.654] ***	295.04 [19.648] ***	280.84 [28.118] ***
Observations	71,399	71,395	30,145

NOTE: Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1.

for hypertension, anticoagulation, osteoporosis and Alzheimer's disease remain even conditional on switching. Finally, the results suggest that switching decisions fully explain the differences observed by age and sex.

6 The Switching Decision

To analyze individuals' decisions about whether to switch plans, we estimate probit models, where the probability of switching is written as

$$P(S_i = 1|\Delta H_i, X_i) = \Phi(\alpha + \Gamma \Delta H_i + \beta X_i) \tag{5}$$

where Φ is the normal cumulative distribution function and as above, ΔH_i is the change in health status and X_i is a vector of time-invariant individual characteristics including health status and age in 2006, among others. X_i also includes the category of the individual's overspending in 2006, which indicates whether those who overspent more in the previous year responded to these financial incentives by switching to other plans. Also within X_i is a variable that indicates how the person's 2006 plan changed in 2007 relative to the alternatives, i.e., it is defined as the difference in the person's 2006 plan's percentile ranking in 2006 and 2007. By capturing how much worse the individual's 2006 plan would become in 2007 relative to the available alternatives, this variable provides insights about whether people were forward looking in their plan choices or whether inertia causes consumers to remain in their plans despite those plans growing relatively worse over time. We estimate this model on the full sample and on the sample in stable health.

To provide additional insights to the switching decisions and reassurance that our observed switching choices are not driven by the relative improvements of the plans in our sample, we also

estimate these switching models with an additional vector of 2006 plan fixed effects, denoted by $P06_i$,

$$P(S_i = 1|\Delta H_i, X_i, P06_i) = \Phi(\alpha + \Gamma \Delta H_i + \beta X_i + \Psi P06_i). \quad (6)$$

.

The 2006 plan fixed effects eliminate any plan-specific attributes that affect the average probability of switching out of the plan. These plan attributes include both the levels and changes of the 2006 plan to the extent that they influence switching decisions. For example, these models exclude switching driven by either the levels or the changes in premiums, as a given plan's premiums affect all non-poor enrollees the same. Thus, by evaluating only the within-plan differences in switching, the estimates from these models assuage concerns that our switching results may be driven by non-generalizable aspects of the plans in our study. Further, the results from these models indicate whether individuals' switching decisions incorporated their plans' attributes that had heterogeneous, individual-specific effects. For example, in these models with plan fixed effects, the backwards-looking 2006 overspending levels as well as the forward-looking change in the 2006 plan's percentile ranking in 2007 are identified only off of individual-specific changes in the plans' performance.

Table 6 presents the average marginal effects, or partial effects for the dummy variables, from these probit estimates. The unconditional probability of switching is 0.54, and as suggested by the results in Table 5, the switching decisions differ with observed characteristics, with males 14 percentage points less likely than females, and the oldest group 39 percentage points more likely to have switched plans than those age 65-69. These results are conditional on the 2006 level of health, changes in health, and the magnitude of overspending in 2006. Consequently, they reflect differences by sex and age in how the decision for 2007 was made, rather than differences in incentives or past performance *per se*. The effects of the level of health status on switching suggest that people with higher volumes of prescription drugs, as a proxy for experience incorporated in

Table 6. Average Marginal and Partial Effects from Probit Models of Switching

	Full Sample	*Subset with Stable Health only*	*Full Sample with 2006 Plan Fixed Effects*
Overspending Level in 2006 ($)			
less than 100	*Reference Category*	*Reference Category*	*Reference Category*
between 100 and 200	-0.08 [0.039] **	-0.12 [0.067] *	-0.03 [0.062]
between 200 and 300	0.21 [0.032] ***	0.21 [0.057] ***	-0.09 [0.057]
between 300 and 500	0.49 [0.029] ***	0.51 [0.051] ***	-0.06 [0.059]
between 500 and 1000	0.50 [0.020] ***	0.49 [0.036] ***	0.01 [0.059]
between 1,000 and 2,000	0.48 [0.010] ***	0.49 [0.017] ***	0.14 [0.059] **
more than 2000	0.43 [0.007] ***	0.45 [0.011] ***	0.19 [0.060] ***
Change in 2006 Plan's Percentile Ranking	0.77 [0.009] ***	0.84 [0.015] ***	0.13 [0.012] ***
Age in 2006			
Age 65-69	*Reference Category*	*Reference Category*	*Reference Category*
Age 70-74	0.12 [0.007] ***	0.14 [0.010] ***	0.13 [0.008] ***
Age 75-79	0.25 [0.006] ***	0.28 [0.009] ***	0.30 [0.008] ***
Age 80-84	0.33 [0.006] ***	0.36 [0.008] ***	0.38 [0.007] ***
Age 85 up	0.39 [0.005] ***	0.41 [0.008] ***	0.43 [0.007] ***
Male	-0.14 [0.005] ***	-0.15 [0.007] ***	-0.13 [0.006] ***
Risk score in 2006	0.01 [0.001] ***	0.00 [0.002] *	0.01 [0.001] ***
Took medication in 2006 for			
Hypertension	-0.02 [0.006] ***	-0.01 [0.009]	0.00 [0.007]
Cholesterol and other cardiovascular	-0.03 [0.005] ***	-0.04 [0.008] ***	-0.02 [0.006] ***
Pain	-0.00 [0.006]	-0.00 [0.011]	0.04 [0.007] ***
Mental health	-0.03 [0.006] ***	-0.03 [0.010] ***	0.01 [0.007] **
Antibiotics	-0.04 [0.006] ***	-0.04 [0.009] ***	-0.06 [0.007] ***
Anticoagulants	-0.04 [0.006] ***	-0.04 [0.010] ***	-0.02 [0.007] **
Thyroid	-0.06 [0.006] ***	-0.06 [0.009] ***	-0.01 [0.007] *
Diabetes	0.01 [0.006]	0.02 [0.011]	0.02 [0.008] **
Osteoporosis	-0.02 [0.006] ***	-0.03 [0.010] ***	-0.03 [0.007] ***
Alzheimer's	-0.06 [0.012] ***	-0.02 [0.021]	-0.07 [0.014] ***
Change in Risk Score	0.01 [0.001] ***	0.03 [0.016]	0.01 [0.002] ***
Change in takes medication for			
Hypertension	-0.00 [0.009]	0.03 [0.021]	0.02 [0.011] *
Cholesterol and other cardiovascular	-0.02 [0.008] ***	-0.03 [0.024]	-0.02 [0.009] **
Pain	-0.01 [0.005] **	-0.01 [0.009]	0.02 [0.006] ***
Mental health	-0.10 [0.007] ***	-0.14 [0.018] ***	-0.05 [0.008] ***
Antibiotics	-0.02 [0.005] ***	-0.02 [0.008] *	-0.02 [0.005] ***
Anticoagulants	-0.01 [0.009]	-0.02 [0.024]	0.01 [0.010]
Thyroid	-0.04 [0.015] ***	-0.02 [0.030]	-0.00 [0.018]
Diabetes	-0.01 [0.014]	0.04 [0.054]	-0.01 [0.017]
Osteoporosis	0.00 [0.009]	-0.00 [0.024]	-0.00 [0.011]
Alzheimer's	-0.02 [0.014]	0.02 [0.060]	-0.04 [0.017] **
Observations	71,391	30,145	70,914

NOTE: Robust standard errors in brackets. *** p<0.01, **p<0.05, * p<0.1.

the risk score variable and the specific illness indicators, were not more likely to switch plans. The results show that the probability of switching varied with the presence of a medical condition by at most 6 percentage points, with people with any of 8 individual conditions significantly less likely to switch plans. In contrast, greater prevalence of other illnesses incorporated in the risk score was associated with higher probability of switching, but those effects are small as well. Individuals who initiated medication for new conditions in 2007 were slightly less likely to have switched plans from 2006 to 2007. Although these effects are generally small, they range up to 10 percentage points for mental health and five of the ten achieve a significance level of $p < 0.05$.

One bias that is commonly cited in the behavioral economics literature is inertia. The estimated differences across the 2006 overspending categories show that the probability of switching plans jumps quite significantly for those whom 2006 overspending was between $200 and $300 a year, and switching is substantially higher yet at levels above $300. Previous studies have documented poor choices during the first year of Part D, leading researchers to conclude that consumers could not choose well in this context and additional reform was needed, such as reducing the number of plans available. Due to the dynamic aspects that we observe in our sample, however, the partial-equilibrium static approach underlying such proposed reforms could be incomplete: overspending as small as $25 per month ($300 per year) in 2006 resulted in a 49 percentage point increase in an individual's likelihood of switching into a new plan for 2007.

These results provide additional insights about inertia by showing how switching depends on the changes in the person's 2006 plan in 2007. The large positive effect of this variable indicates that beneficiaries tend to switch out of their 2006 plans if those plans would become more expensive in 2007 relative to the alternatives available to them. Specifically, the results show a 7.7 percentage point increase in the likelihood of switching for each additional ten percentage points that the 2006 plan worsened in the individual's distribution of options available in 2007. In addition to mollifying concerns that inertia is dominant in this market, this result indicates that consumers in our sample

incorporated some forward-looking information in their choice of plans, as these results are based on within-plan changes in performance and are conditional on the level of overspending in 2006.

The results can also be naturally interpreted as showing who was more likely to remain in their 2006 plans. Specifically, the results indicate that those most likely to remain in their 2006 plans, conditional on other factors, were precisely those who overspent least in 2006, and those for whom their 2006 plans would improve most in 2007 relative to the alternatives. For example, those whose 2006 plans would improve by ten percentile points in their distribution of alternatives in 2007 were 7.7 percentage points less likely to switch out of that plan.

The results in the final column of Table 6 indicate that switching decisions depended not only on plan-level factors but also on the individual-specific factors that affected overspending and future changes in the plan's relative costs. The number of observations for these estimates is slightly smaller because the estimates exclude individuals in plans that had no within-plan variation in the switching decisions for the individuals in our sample. These results show that overspending levels above $1,000 led to significantly higher levels of switching relative to those in the same plan who overspent less in 2006. Similarly, those for whom the plan would become relatively worse in 2007 were also more likely to switch out than others in the same plan. The results for the other coefficients also maintain their general size and significance, showing that they, too, are not driven by plan-level factors. For example, the lower switching rates by males appear due to differences in the decision process and not simply due to males being disproportionately enrolled in 2006 plans that enrollees as a whole were more likely to switch out of.

7 Robustness of Results

In this section we consider whether the results for the mean within-person changes we have reported are robust to alternative approaches. First, we replicate the analysis but adopt the assumption that

Table 7. First Difference Models of Within-Person Changes in Overspending Using Alternative Approaches

	Controlling for Changes in Health (Identical to Table 3 Column 2)		*And Switching (Identical to Table 5 Column 1)*	
A. Main results reported in Tables 1 and 3				
Intercept	-295.97	[3.890] ***	-136.90	[7.654] ***
Switched plans			-298.46	[8.256] ***
Mean Overspending in 2006	546.9		546.9	
B. Assuming perfectly inelastic demand				
Intercept	-368.60	[4.991] ***	-158.25	[8.864] ***
Switched plans			-389.00	[9.876] ***
Mean Overspending in 2006	794.0		794.0	
C. Using actual rather than simulated cost for actual plan				
Intercept	-273.46	[4.677] ***	-107.1	[8.582] ***
Switched plans			-307.66	[9.329] ***
Mean Overspending in 2006	586.0		586.0	
D. Excluding no insurance as an option				
Intercept	-303.02	[4.123] ***	-139.25	[7.637] ***
Switched plans			-302.87	[8.236] ***
Mean Overspending in 2006	538.5		538.5	

NOTE: Robust standard errors in brackets. *** p<0.01, **p<0.05, * p<0.1. N = 71,399 for all models.

demand for prescription drugs is perfectly inelastic, rather than assuming an elasticity of -0.54 as in our main results. Second, we replicate the results but use the actual spending for the actual plan. In the main results we relied on the simulated spending for the actual plan to be consistent in our approach across the entire choice set. Third, we evaluate within-person changes in overspending when we exclude no insurance as an option from the choice set.

For each of these versions, we estimate two models. The first is identical to that reported in Table 3 column 2, showing the average within-person change conditional on changes in health. The second is identical to that in Table 5 Column 1, which are also conditional on changes in health but allow the within-person change to differ between those who switched plans and those who did not.

For comparison, the main results are reported in Panel A of Table 7. The approach behind these results allowed for a quantity response to the prices of drugs under each available alternative.

To test the sensitivity of the results to this assumption, we replicate many of the analyses assuming perfectly inelastic demand for drugs. The results, reported in Table 7 Panel B, indicate somewhat larger average improvements than under the assumption of elastic demand. Specifically, here the reductions of overspending are $369. As a percent of the mean 2006 overspending, this amounts to a 46% reduction, which is slightly below the 54% found for the elastic results. The results by switching show only somewhat larger improvements for those who did not switch but $112 ($547-$435) greater improvements for those who did switch, compared with the main results. [14]

Table 7 Panel C shows the results when we use the actual spending in the actual plan instead of the simulated spending in the actual plan. These results are largely similar with our main results, showing slightly smaller reductions in overspending from 2006 to 2007. In Panel D we show the results when no insurance is excluded as an option. These results show just slightly larger reductions in overspending than in our main results, which include no insurance.

As a final consideration of the robustness of our results, we compare the 2007 cross-sectional results using both our primary *ex post* approach with those from the alternative *ex ante* approach. All of the preceding analyses adopt an *ex post* benchmark by considering overspending based on the actual drug consumption for the year following consumers' decisions about which plan to join. To determine whether the results depend on this assumption, here we alternatively assume that consumers' only relied on their 2006 drug consumption when making their enrollment choices for 2007. To do so, we generate the total *ex ante* 2007 spending by combining the patient OOP costs from their 2006 drugs with the 2007 attributes of each available alternative such as the plans' formularies and premiums. Because we lack information on individuals' 2005 drug claims, this analysis is limited to comparing the cross-sectional results for 2007 from these two alternative assumptions about consumers' information sets. The results in Table 8 indicate great similarity

[14] Although we do not report them in the table, the remainder of the key implications of the main results also persist.

Table 8. Comparing 2007 Overspending Using *Ex Ante* and *Ex Post* Prescription Drug Claims

	Ex Post *Using 2007 Claims ($)*	*Ex Ante* *Using 2006 Claims ($)*
Mean	251.0	298.4
Median	184.8	197.8
5th Percentile	0.0	0.0
10th Percentile	1.7	14.0
25th Percentile	65.0	79.1
75th Percentile	184.8	345.9
90th Percentile	515.7	526.8
95th Percentile	682.9	700.5

NOTE: The *ex ante* approach defines the total spending in each available plan in 2007 using the claims filled by the person in 2006. The *ex post* approach uses the claims filled by the person in 2007. Both rely on the plans available and their attributes (e.g., premiums and formularies) in 2007.

between the two approaches, both in terms of the mean as well as at various other points in the distribution of 2007 overspending. In addition to this high degree of similarity, the results show that overspending is somewhat lower using the *ex post* benchmark rather than the *ex ante* approach.

8 Generalizability of Our Estimated Mean Improvement

Given that we have access to data from only a subset of plans, it is important to consider whether our estimated mean improvement may generalize to the broader population of non-poor who were enrolled in a PDP for all of 2006 and 2007. Specifically, concerns may arise that differences between the individuals and plans in our sample and the population of PDP enrollees and plans overall may cause the average improvement of the population to be above or below our sample's average improvement of $298. In this section we present several analyses that evaluate how our results may be influenced by the differences shown in Tables 1 and 2. First, we evaluate the extent to which the differences in the individual demographics shown in Table 1 may cause the estimated average improvement for our sample to be above or below the average mean improvement of the population as a whole. We then examine how the differences in the levels and changes in the attributes of

our in-sample plans versus PDP plans overall may influence our results. The concerns about plan redesign are most relevant in the context of understanding why even those who did not switch plans improved on average, and consequently, we implement several analyses to evaluate that sub-sample specifically. In contrast, these concerns are less relevant to the results that come from differences in improvement between individuals within our study sample, such as the larger gains among those who switched, the greater likelihood of switching among those who overspent more in 2006, and the result that the sub-sample with cognitive limitations had gains that were at least as large as those of our sample as a whole.

As shown in Table 1, our study sample has a greater prevalence of females and of those age 85 and above. Because our results show that both of these groups improved by more than average, this suggests that our results overstate the average mean improvement due to demographic differences alone. To estimate the degree to which our mean improvement is biased upwards, we calculate what our average mean improvement would have been if our sample had demographics that were identical to the whole population in terms of age, sex and gross spending in 2006. Specifically, we multiply the coefficients on these variables from Table 4 column 3 (the model that includes gross spending in 2006) by the differences in these 2006 demographic characteristics between our sample and the full population. These differences are reported in the last column of Table 1. The result from this calculation shows that our estimated mean improvement would have been $12 lower, or $286. Thus, based on individual demographics, our average mean improvement appears to incorporate only minimal bias.

As our first consideration of the role of plan redesign, we implement a simple cross-sectional analysis to quantify the net effects of the average changes shown in Table 2. The purpose of this analysis is to assign the consumers' costs, in dollars, of the changes in plan attributes such as formulary coverage so that they can be compared with the countervailing changes in premiums. To do this, we use the 2007 cost simulation file that reports the simulated OOP cost for drugs for each

person under each plan available to him, assuming a price elasticity of -0.54 as before. From the CMS data, we link on the plan attributes reported in Table 2 and then estimate the regression

$$C_{ij} = \Gamma A_j + \xi_i + u_{ij} \tag{7}$$

where C_{ij} is the simulated OOP cost of drugs for individual i on plan j and A_j is the vector of plan attributes other than premiums shown in Table 2. In addition, in our composite error term we allow for individual specific unobserved heterogeneity ξ_i and an idiosyncratic error u_{ij}. The coefficients from this model show the average within-person change in OOP spending due to a one unit change in each plan attribute in Table 2. By multiplying these coefficients by the relative changes in attributes in Table 2 and adding the relative change in premiums, we are able to simulate the magnitude of the effect of the changes of the average attributes of our in-sample plans relative to the changes in the average attributes of all PDPs.

The results of this calculation indicate that we expect a net relative reduction of $45 due to these different changes in average plan attributes. This amounts to 15% of our estimated average reduction of $298, or one third of the reduction among non-switchers but only one tenth of switchers' improvement. This $45 is the difference between the $23 reduction we estimate for the changes in our in-sample plan's average attributes and the $22 increase for the changes in the average attributes of all PDPs. An examination of the roles of the specific plan attributes shows that the decreased generosity of our in-sample plans' formulary coverage increased consumers' costs by $50 relative to plans overall, largely offsetting the $81 relative reduction due to premiums. Thus on net, the changes in average plan attributes shown in Table 2 have relatively little relevance for understanding our estimated improvement overall and the estimated improvement among switchers in particular. These average plan changes have some relevance for understanding the improvements among non-switchers, but even among that sub-sample, two thirds of the gains remains unexplained.

As a second consideration, we evaluate improvement by non-switchers enrolled in the in-sample plans that improved by less than the national average in terms of the sum of the premium and deductible. Table 2 shows that on average, the sample plans improved by $152 in terms of the sum of the premium and deductible, while the overall PDP average improvement was $13. Some of the plans in our study sample, however, improved by less than this nationwide average. Specifically, 9,670 individuals who did not switch from 2006 to 2007 remained in plans whose sum of deductibles and premiums improved by less than $13. Re-estimating equation (1) for only this sub-sample shows that they experienced an average conditional reduction in overspending of $182. This is larger than the average conditional improvements by our full sample of non-switchers of $137, as shown in the first column of Table 3. Thus, the relative reductions in the average premiums and deductibles by our study plans do not appear to explain the average improvements by non-switchers.

We conduct one additional analysis to understand the extent of improvement among non-switchers that was due to changes in plan design, plan availability, and drug prices as opposed to individual-specific changes in drug consumption from factors such as changes in unobserved health or individual's learning how to navigate their formularies. To accomplish this, we re-estimate equation (1) but hold each non-switching person's drug consumption constant at their 2006 levels. Specifically, we first simulate each person's 2007 spending using his 2006 claims in conjunction with each option's 2007 attributes and drug prices. We then define the within-person change in overspending for this analysis as the difference between this value and the 2006 overspending, which we define as before. Under this approach, any changes in overspending from 2006 to 2007 are due entirely to changes in the premiums and cost of drugs in the individuals' actual plans and their minimum-cost options.[15]

[15] As before, we assume a price elasticity of demand for drugs of -0.54, although the results are virtually identical when we instead assumed perfectly inelastic demand.

The results from this model show that when only plan design and availability and drug prices are allowed to change from 2006 to 2007, non-switchers reduced their overspending by $117, which is 85% of their average total reduction conditional on changes in health shown in Table 5 column 1.[16] Thus, improvement in their own plans that they chose to remain with, relative to the changes in the minimum-cost options available to them, explains a large share of improvement among non-switchers. The remaining 15% is due to changes in drug consumption not accounted for by changes in the measures of health included in the model. One potential source of such changes is consumers learning how to navigate their plans' formularies in ways that reduced their spending on their actual plan more than it reduced their simulated spending on their minimum-cost plans, which could occur if the formularies differ.

A final observation from Table 2 is that our sample does not include any plans with doughnut hole coverage for either brands or generics in 2006. Abaluck and Gruber (2011) argue that an important reason why beneficiaries overspent substantially in 2006 was that they held harmful, biased preferences for plans that included doughnut hole coverage. Our sample plans greatly expanded generic doughnut hole coverage in 2007, yielding far higher levels and changes in that coverage among our in-sample plans than among PDPs as a whole. If the results found by Abaluck and Gruber (2011) apply to our sample, our estimates of the improvements would be biased downwards overall and for both those who switched into plans adding such coverage and those who remained in plans that added such coverage. For those who switched, this follows from the idea that the in-sample plans presented the in-sample individuals with above-average opportunities to act on this harmful bias in 2007 but below-average opportunities in 2006. Among non-switchers, this expansion in gap coverage with its concomitant increase in premiums would have caused their improvements to be below that of individuals who remained in plans that did not add such coverage,

[16]The effects of changes in the minimum cost plans' designs are ameliorated by our inclusion of no insurance in the choice set, which had a stable design by definition. However, this option is particularly affected by secular increases in drug prices, since without insurance the individual would pay all of those higher costs.

which were more prevalent outside of our sample. According to the results of Abaluck and Gruber (2011), our results would also represent an underestimate because those who chose doughnut hole coverage in 2006 would have overspent most, where we have found that those who overspent most also improved most. Similarly, the fact that individuals in our sample did not select such plans in 2006 may indicate that they are less imbued with such biases and overspent less in 2006, again leading us to underestimate the overall average improvements from 2006 to 2007.

In sum, these analyses suggest that attributes of our particular sample have not introduced any important systematic bias to our estimated average improvements. The differences in individual demographics suggest that we overestimated the true average improvement of the whole population by 4% ($12). Some differences in plan attributes suggest that we may have underestimated the true reduction experienced by the population overall, while others suggest that we may have overestimated it. On net, the differences in the changes in average plan characteristics in Table 2 have some salience relative to the improvement among non-switchers but are fairly small relative to our estimated overall improvement. Instead, we find that relative changes between the actual and minimum-cost plans are important in explaining the improvements among non-switchers. We lack the ability here to determine what explains this improvement among this sub-sample of non-switchers, and whether the explanations generalize to the non-poor PDP population overall. Explanations that may generalize to the overall population include the idea that individuals learn about their specific plans' formularies resulting in lower spending, as well as the idea that individuals making active choices to remain in plans that were set to improve. One alternative explanation that may be specific to the subset of non-switchers in our sample is that they enjoyed passive gains from their cheapest plans exiting the market or being redesigned to become more expensive for them in 2007.

9 Conclusions

We analyze changes in Medicare Part D plan choices of older consumers who faced a potentially dizzying array of options for complex, multiattribute products. Despite these features raising concerns for potential widespread consumer confusion, we find evidence that overspending fell substantially on average among our study sample even after less than a full year of experience in this new market. Specifically, we find that the mean cost difference between individuals' actual choices and their cheapest options fell by about $300 from 2006-2007, and the median difference fell by about $240. Further, our results show that overspending fell for a large majority (81%) of our sample. The greatest reductions were by those who overspent the most in 2006, resulting in a substantial reduction in the distribution of overspending among our sample in 2007 despite a divergence in PDP plan attributes overall.

We find evidence that this average reduction in overspending is due at least in part to choices by individuals in our sample. Specifically, those who chose to switch plans reduced their overspending far more than those who did not. The likelihood of switching plans for 2007 increased substantially with the amount of overspending in 2006. This result contrasts with well-known evidence of inertia from other contexts, in which the levels of prior overspending were not associated with differences in the likelihood of changing gym membership choices (DellaVigna and Malmendier (2006)). However, our result is consistent with evidence of learning in another new market, where consumers faced new options for local telephone service (Miravete (2003)). The possibility of learning from experience has also been documented by Agarwal, Driscoll, Gabaix, and Laibson (2009b) in the context of the credit card industry, another complex and competitive market similar to Medicare Part D.

In addition to suggesting that consumers incorporate information about past performance in their current plan choices, our results show that their choices to stay or switch incorporated

the plans' announced future changes. Specifically, we found that the likelihood of switching out of a plan was higher when the current plan's relative costs were set to increase in the next year. Although we cannot observe consumers' decision processes, consumers' ability to anticipate these changes is likely due to plans publicizing their upcoming changes prior to the open enrollment period. Interestingly, consumers appear to respond not only to readily-visible plan changes that affected everyone equally, such as premiums, but also to how changes in the plan's traits would affect them individually, relative to their average effects on all enrollees. For example, these results predict that if a plan announced it was going to drop formulary coverage for a drug, those who took that drug would be more likely to switch out.

The richness of our data also allowed us to consider how the within-person changes varies across a number of individual characteristics. Notably, our results provide evidence that even the sub-populations with greater prevalence of cognitive limitations, as indicated by taking medications for Alzheimer's disease or being age 85 or above, experienced these average improvements. In fact, the oldest consumers improved by the most, those initiating medications for Alzheimer's disease in 2007 improved by more than average, and those taking such medications in both years did not significantly differ from the average improvement. The improvements among the oldest appears to be fully explained by their higher likelihood of switching plans, but the results show that the above-average improvements from those who initiated medication for Alzheimer's disease derives from sources other than switching. These results suggest that these populations are assisted in their choices of plans and medications by various real-world institutions such as children and other relatives, medical personnel, social networks, and other organizations and decision support tools. By evaluating actual choices, our analysis incorporates these various institutions in ways that the previous survey- and laboratory-based research has not (e.g., Hanoch et al. (2009), Hsu, Fung, Price, Huang, Brand, Hui, Firema, and Newhouse (2008)). The benefits of such institutions likely extend beyond those with cognitive limitations and contributed to the improvement among our

sample as a whole. As the Part D market has evolved, consumers have had greater access to information through plan ratings, user-friendly websites, and software applications of pharmacy chains and other institutions to help people choose well-matching plans. However, these sources of information did not appear overnight when Medicare Part D was first implemented.[17]

Omitted from our analysis are risk aversion and uncertainty, although evaluating within-person changes over time accounts for them as long as their individual-specific effects are time-invariant. While the PBM data offer a number of strengths, they can provide insights about only a limited scope of choices related to Part D. First, because our data include only enrollees, we cannot analyze the participation decision itself, as done by Heiss et al. (2007). Second, we analyze the choice of prescription drug insurance in isolation, where Medicare beneficiaries face a number of decisions about how to acquire insurance for the full range of medical care. Thus, some individuals may improve in their choices in Part D plans but not improve in terms of health insurance coverage overall. Unfortunately, it is currently infeasible for researchers to analyze these broader decisions. Third, our results cannot speak to the extent of switching, or the gains from switching, between plans that were not sold or administered by the PBM that we study. Of particular interest is the experience of those outside of our sample who enrolled in plans that offered brand gap coverage in 2006, especially their overspending and switching decisions after 2007 as insurers rapidly eliminated such coverage.

We also find average improvements even among those in our sample who did not switch plans, although the specific sources of these gains are unknown. Our results show that 15% of this improvement was due to changes in drug consumption in ways that lowered costs in the actual plan relative to the minimum cost plan. One plausible cause of such changes is consumers learning

[17] See `http://www.medicalnewstoday.com/articles/57318.php` for one example of a private sector initiative that was not available in 2005 for the 2006 open enrollment but was available in the fall of 2006 for the 2007 open enrollment period. The plan finder offered through the CMS was also overhauled between the two periods to be more user friendly. Similarly, the CMS added "star ratings" to summarize plan quality in 2008, which has not been studied but may have affected choices subsequent to our study period.

about the formulary coverage in their actual plans. One implication of such plan-specific learning is that switching costs may exist even beyond the cost of searching for new plans. The large majority (85%) of the improvements among non-switchers was due to changes in their actual plans relative to their minimum cost plans. These include changes in attributes such as premiums, formulary coverage, and underlying drug prices. Given the limitations of our data and our approach, it is unclear whether the majority of these improvements among non-switchers reflect active decisions or passive gains. Supporting the idea of improvements due to active choices, our estimates of the decisions to remain in their 2006 plan show that those who chose to remain were those who overspent least in 2006, and those for whom their 2006 plans improved the most in 2007 relative to the alternatives. In contrast, supporting the view that these may reflect passive gains that were enjoyed despite non-switchers' inertia is the observation that our sample plans improved more on average in terms of premiums and deductibles. However, our plans worsened in other ways such as formulary coverage, and an investigation of the net effects of these average changes indicate that they were equivalent to 33% of the average improvement among non-switchers. Further, an analysis of the sub-sample of non-switchers in plans that improved by less than the national average in terms of premiums and deductibles showed that they actually improved by more than our overall sample of non-switchers.

Another potential sources for passive gains by non-switchers is random variations over time, which could arise from changes in health that we cannot observe, from randomness in individuals' plan choices, or from other sources. A final potential source for passive gains is supply-side decisions about plan offerings, although standard theory predicts that those are made in response to consumers' choices. While some of our results suggest that consumers respond to their past performance, particularly those from our estimates of the switching decisions, we lack the ability to quantify the contributions of each of these potential active and passive causes of the improvement among our study sample. Additional research that incorporates suppliers' decisions and consumers'

decisions, search costs, and switching costs such as plan-specific learning would provide a deeper understanding of the precise mechanisms that explain these dynamic changes. Such an analysis would shed light on the sources of the gains observed among the non-switchers in our sample in particular. In addition to the need for alternative analytical approaches, research using broader, longer individual-level panel data would also be instrumental in resolving these remaining questions. Ultimately, an analysis of the full CMS data covering all PDP enrollees and plans beyond 2007 is likely to be required to provide convincing answers. At this time, however, the CMS has not permitted outside researchers to identify which plans individuals are enrolled in, precluding any independent analysis of plan choice or switching on the full population (Goldman and McFadden (2008)). Although these questions about the performance of the PDP population overall will remain until such analyses are conducted, our investigation suggests that our estimates of the average improvement introduced only a small degree of bias due to specific features of our sample plans and individuals.

While incorporating these strengths and limitations, our results provide insight to how Part D plan choices changed from 2006 to 2007 for a large sample of PDP enrollees. Our results add to the accumulating evidence that Part D represents a successful implementation of a market-based approach to deliver a large scale entitlement program (Goldman and McFadden (2008)). In addition to the high enrollment found by Joyce et al. (2009), Duggan et al. (2008) report that Part D has reduced pharmaceutical prices, increased the utilization of prescription drugs, and reduced medical expenditure risk, all at a substantially lower cost than initially expected. Furthermore, although Heiss, McFadden, and Winter (2006) initially questioned whether a government run program could have more effectively controlled for the cost of drugs, they later concluded that Part D has been a tactical success that has induced high enrollment levels, ensured competition among private insurance sponsors, and kept drug prices and rates of consumer deception low (Heiss et al. (2007)).

Repeated evidence that choice overwhelms consumers and leads to persistent, poor matches between consumers and products will provide a partial economic rationale for policies such as strengthening consumer protection rules or mandating simplified retirement plans, credit cards, mortgages, or health insurance contracts. Alternatively, evidence that consumers quickly adapt to evolving information and markets and capitalize on the heterogeneity of available products would indicate that policies that reduce barriers to entry and promote competition will enhance welfare (Economides, Seim, and Viard (2008)). An additional economic rationale for greater regulatory intervention grows out of concerns that unsophisticated consumers do not enjoy these gains, or that they are even harmed by the success of others (Gabaix and Laibson (2006)). The existing evidence suggesting that consumers in general, and the elderly in particular, cannot appropriately choose financial products such as Part D Plans has been obtained almost entirely from cross-sectional data that often abstracts from the actual choice process, e.g., by conducting surveys of financial literacy or familiarity with the products' attributes. By construction, such approaches rule out the possibilities that consumers improve their choices, rely on support systems or utilize information available in the market. Our results, from an analysis of how consumers' actual choices evolved in a new market, provide some contrast by showing large average gains over time among our sample overall and even among sub-populations with high prevalence of cognitive limitations. This contrast highlights the importance of incorporating market dynamics and various sources of decision support in analyzing consumer choices.

References

Abaluck, J. and J. Gruber (2011): "Choice Inconsistencies Among the Elderly: Evidence From Plan Choice in the Medicare Part D Program." *American Economic Review*, 101, 4, 1180–1210.

Agarwal, S., J. Driscoll, X. Gabaix, and D. Laibson (2009a): "The Age of Reason: Financial Decisions over the Life-Cycle and Implications for Regulation." *Brookings Papers*

on Economic Activity, , 2, 51–117.

Agarwal, S., J. Driscoll, X. Gabaix, and D. Laibson (2009b): "Stimulus and Response: The Path from Naivite to Sophistication in the Credit Card Market." Working Paper.

Choi, J., D. Laibson, and B. Madrian (2010): "Why Does the Law of One Price Fail? An Experiment with Mutual Funds." *Review of Financial Studies*, 23, 4, 1405–1432.

DellaVigna, S. (2009): "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47, 315–372.

DellaVigna, S. and U. Malmendier (2006): "Paying Not to Go to the Gym." *American Economic Review*, 96, 694–719.

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

Economides, N., K. Seim, and B. V. Viard (2008): "Quantifying the Benefits of Entry into Local Phone Service." *RAND Journal of Economics*, 39, 3, 699–730.

Fang, H., M. P. Keane, and D. Silverman (2008): "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market." *Journal of Political Economy*, 116, 303–350.

Gabaix, X. and D. I. Laibson (2006): "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, 121, 2, 505–540.

Goldman, D. and G. Joyce (2008): "Medicare Part D: A Successful Start with Room for Improvement." *Journal of the American Medical Association*, 299, 1954–1955.

Goldman, D. and D. McFadden (2008): "Want to Monitor Medicares New Drug Benefit Program? Start by Sending a Check for $120,000." *Economists' Voice, The Berkeley Electronic Press*, 1–4.

Hanoch, Y., T. Rice, J. Cummings, and S. Wood (2009): "How Much Choice is Too Much? The Case of the Medicare Prescription Drug Benefit." *Health Services Research*, 44, 1157–1168.

Heckman, J. J. (1981): "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time - Discrete Data Stochastic Process." In C. F. Manski and D. L. McFadden (eds.): *Structural Analysis of Discrete Data with Applications*. Cambridge, MA: MIT Press.

Heiss, F., D. McFadden, and J. Winter (2006): "Who Failed to Enroll in Medicare Part D, And Why? Early Results." *Health Affairs*, 25, w344–w354.

Heiss, F., D. McFadden, and J. Winter (2007): "Mind the Gap! Consumer Perceptions and Choices of Medicare Part D Prescription Drug Plans." NBER Working Paper No. 13627.

Heiss, F., D. McFadden, and J. Winter (2009): "Regulation of Private Health Insurance Markets: Lessons from Enrollment, Plan Type Choice, and Adverse Selection in Medicare Part D." NBER Working Paper No. 15392.

Hoadley, J. (2008): "Medicare Part D: Simplifying the Program and Improving the Value of Information for Beneficiaries." *The Commonwealth Fund Issue Brief*.

Hsu, J., V. Fung, M. Price, J. Huang, R. Brand, R. Hui, B. Firema, and J. P. Newhouse (2008): "Medicare beneficiaries knowledge of Part D prescription drug program benefits and responses to drug costs." *Journal of the American Medical Association*, 299, 16, 1929–1936.

Joyce, G., D. Goldman, W. Vogt, E. Sun, and A. Jena (2009): "Medicare Part D After 2 Years." *The American Journal of Managed Care*, 15, 536–544.

Kling, J., S. Mullainathan, E. Shafir, L. Vermeulen, and M. Wrobel (2011): "Comparison Friction: Experimental Evidence from Medicare DrugPlans." Quarterly Journal of Economics, forthcoming.

Liebman, J. and R. Zeckhauser (2008): "Simple Humans, Complex Insurance, Subtle Subsidies." NBER Working Paper No. 14330.

List, J. A. (2003): "Does Market Experience Eliminate Market Anomalies?" *The Quarterly Journal of Economics*, 41–71.

List, J. A. (2004): "Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace." *Econometrica*, 72, 2, 615–625.

List, J. A. (2006): "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions." *Journal of Political Economy*, 114, 1, 1–37.

List, J. A. and D. L. Millimet (2008): "The Market: Catalyst for Rationality and Filter of Irrationality." *B.E. Journal of Economic Analysis & Policy*, 8, 1, 1–55.

Lusardi, A., O. S. Mitchell, and V. Curto (2009): "Financial Literacy and Financial Sophistication Among Older Americans." NBER Working Paper No. 15469.

Miravete, E. J. (2003): "Choosing the Wrong Calling Plan? Ignorance and Learning." *American Economic Review*, 93, 1, 297–310.

Pauly, M. V. and Y. Zeng (2004): "Adverse Selection and the Challenges to Stand-Alone Prescription Drug Insurance." *Frontiers in Health Policy Research*, 7, 3, 1–22.

Querfurth, H. W. and F. M. LaFerla (2010): "Alzheimer's Disease." *The New England Journal of Medicine*, 4, 362, 329–44.

Shea, D. G., J. V. Terza, B. C. Stuart, and B. Briesacher (2007): "Estimating the Effects of Prescription Drug Coverage for Medicare Beneficiaries." *Health Services Research*, 43, 933–949.

Simon, K. I. and C. Lucarelli (2006): "What Drove First Year Premiums in Stand-alone Medicare Drug Plans?" NBER Working Paper No. 12595.

Tchernis, R., S.-L. Normand, J. Pakes, P. Gaccione, and J. Newhouse (2006): "Selection and Plan Switching Behavior." *Inquiry*, 43, 10–22.

APPENDIX

A Individual's Out-of-Pocket Costs for Each Available Plan and Year

For each person in each year, we estimated the total out-of-pocket (OOP) costs (the sum of plan's premiums and OOP prescription drug costs) for all PDPs available in the person's market. The Geographic Locator File from CMS was used to determine each person's PDP region code at each point in time. We then used the CMS Plan Information file to determine which plans were available in each person's market. To generate the OOP and gross prescription drug costs for each person in each available PDP, we used the PBM claims data, which include all claims paid by the PBM as well as all of the claims submitted but rejected. We combined these with the CMS formulary files. Across all plans we held constant whether the prescription was filled via mail service or retail, and for retail prescriptions we assumed that they were filled at a preferred network pharmacy. This was essential because within a given plan, prices for a given drug can vary by pharmacy type (mail *vs.* retail) and whether the pharmacy is in the plan's network.

In the main estimates reported in the article, we allowed the consumption bundle to vary with the average price of the person's drugs in each plan. Starting from the amount consumed under the actual, chosen plan, we compute the amount of each drug consumed by applying the assumed arc-elasticity of -0.54 (from Shea et al. (2007)) to the difference in prices of each drug under the different formularies.[18] We then multiply this adjusted quantity by the individual's average price of prescriptions in each plan and add the appropriate premium for each plan.[19] This creates each plan's total OOP cost to each patient for each year, which we use to analyze how the chosen plan compares to the cheapest alternative.

One challenge to researchers examining consumers' actual Part D plan choices is that the CMS formulary file does not provide all required information in cases where patients paid the full cost, as under the deductible or doughnut hole, or some fraction of it, as with coinsurance. Specifically, the formulary files do not provide the underlying prices of each drug in each available

[18] Because we define each option's costs based only on the drugs that were purchased under the actual plan, we are not able to fully incorporate substitution patterns among drugs, which may perhaps led to measurement error. For example, assume that drugs A and B are perfect substitutes, and in the person's actual plan they pay \$10 for drug A, which they consume because the copay for drug B is \$30. If in an alternative plan drug A's copay is \$30 but drug B's is \$10, the individual's total drug consumption and total costs would have remained unchanged. Under the assumption of perfectly inelastic demand, we would estimate the person's expenditures to be \$20 higher per prescription under the alternative plan. Allowing for somewhat elastic demand incorporates some but not all of the demand response and thus gets closer to the true counterfactual drug spending. Unfortunately it is not feasible for us to estimate the impacts of specific formulary designs on consumption patterns in a more precise way.

[19] One useful characteristic of our data is that they identify which of the four federal low income subsidy (LIS) levels, if any, was received by each person in each year. This allowed us to eliminate the population who received any of the levels of low income subsidies in either 2006 or 2007.

plan to which the coinsurance would be applied.[20] We relied on a range of data sources to provide these underlying prices. First, we used Wolters Kluwer Health Source LX claims data to generate an average price per unit (e.g., day's supply) for each of the Food and Drug Administration's National Drug Code (NDC) in each plan by quarter, region, and pharmacy type (retail or mail service). To do this, indicators in the Wolters Kluwer Health data were cleaned and used to generate these measures at the level of the PDP parent company. Individual plan identifiers are not available in these data, but the prices we need do not vary across plans within a parent organization, which we confirmed by examining "scraper" data from the Medicare Prescription Drug Plan finder website for 2006 as in Simon and Lucarelli (2006). These scraper data captured the price per unit for 400 common drugs for each plan in 2006.

Because the WKH data did not provide all of these underlying prices needed to estimate costs for each person in each available plan, we relied on the scraper data for most of the additional prices. Where needed, we multiplied the coinsurance rate by the average unit price from either of these two sources for the given NDC, quarter, region and pharmacy type.[21] To validate this approach of determining each person's total OOP spending in each available plan, we compared the total spending for the person's actual plan derived from this simulation method with that directly observed from the PBM data. The correlation coefficient between these was 0.81 in 2006 and 0.77 in 2007. This suggests that our method for estimating spending in counterfactual plans was highly accurate.

We also included not enrolling as a possible option to the choice set. One benefit of including this alternative is that it limits the extent to which any observed improvements in plan choice can be ascribed to exit or redesign of the lowest-cost plans. For this "no insurance" option, premiums were $0 and drug costs were determined by the usual and customary price, which is what the pharmacy would have charged a cash-paying customer.

[20] CMS began reporting such prices for a subset of National Drug Codes (NDCs) for the 2009 data but does not plan to release them for earlier years.

[21] Where necessary, for a given NDC we imputed these values by relying first on adjacent quarters, then on adjacent regions, and moved farther away in geographic or temporal space. For mail service pharmacies, where necessary we also imputed from retail pharmacies in the same region and quarter. Any remaining needed prices were imputed from the PBM data by dividing the price for the person's actual PBM plan by the average relative price for a given plan in the scraper data.