Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance[†]

By Jason Abaluck and Jonathan Gruber*

We study choice over prescription insurance plans by the elderly using government administrative data to evaluate how these choices evolve over time. We find large "foregone savings" from not choosing the lowest cost plan that has grown over time. We develop a structural framework to decompose the changes in "foregone welfare" from inconsistent choices into choice set changes and choice function changes from a fixed choice set. We find that foregone welfare increases over time due primarily to changes in plan characteristics such as premiums and out-of-pocket costs; we estimate little learning at either the individual or cohort level. (*JEL* G22, H51, I13, I18, J14)

The past five years has seen a sea change in the way that publicly provided health insurance benefits are delivered to the US population. From the introduction of the Medicare and Medicaid programs in 1965, expansions in public health insurance entitlements came through the extension of these monopoly government-run insurance plans. But beginning with the introduction of the Medicare Part D Prescription drug program in 2006, and continuing through the exchanges that are at the center of the Affordable Care Act (ACA) that was passed in 2010, the United States has been moving to a different model: insurance exchanges where the publicly insured choose from a host of subsidized private insurance options.

This privatization of the delivery of public insurance raises a host of interesting policy and research questions. Primary among these is the ability of individuals to make consistent choices across a potentially large array of choices with meaningful differences. In Abaluck and Gruber (2011), we explored this issue in the context of the Part D program. We considered in particular whether elders appeared to be properly weighing the premium and out-of-pocket spending implications of their plan choices. We concluded that they were not, with the typical elder able to save

about 30 percent of total costs (premiums plus out-of-pocket costs) through a more appropriate choice of plan.

A subsequent paper by Ketcham et al. (2012), however, argues that these choice inconsistencies are largely initial errors made by those newly enrolling in Part D and that they are mostly corrected through plan switching from 2006 to 2007. This is an important contention. If in fact choice inconsistency was simply a feature of the initial year of the program, it suggests that this is not a major issue for the long-run welfare evaluation of this new insurance model.

In this paper we revisit the question of evolving choice inconsistencies over time. We do so using newly available data on a 20 percent sample of Part D claims data recently made available by the Centers for Medicare and Medicaid Services (CMS). These data have information on the full set of prescription drug claims under Part D for every senior enrolled in the program, as well as information on the characteristics of the plans that they chose.

Using these improved data, we continue to find that seniors are making inconsistent plan choices in the first year of Part D (2006). We find a similar level of foregone savings and overweighting of premiums relative to expected out-of-pocket costs as in our previous work. However, in contrast to Ketcham et al., we find that choices do *not* improve over time in aggregate. Rather, the amount of "foregone savings" available to consumers choosing Part D plans grows over time, and is larger by 2009 than it was at the start of the Part D program in 2006.

This fact motivates us to develop a structural framework in which to study plan choices over time. This framework allows us to mathematically decompose the "foregone welfare" from inconsistent plan choices into components due to changes in the choice set and product characteristics, changes in the choice function from a fixed choice set, and changes in preferences over time.[1] On the choice function side, we can use data on choices by movers and stayers, as well as by new cohorts over time, to separately identify the effects of inertia, learning from experience, and calendar year or "cohort" learning effects on the quality of choices. On the choice set side, we can use the structural parameters of our model to decompose the total choice set effect into factors such as the change in premiums, the change in out-of-pocket costs, and the set of plan choices available.

Overall, we find that the welfare cost of choice inconsistencies increases over time. These net changes reflect a number of moving pieces and different factors appear to be driving the changes in different years. But the key general conclusion of our analysis is that there is little improvement in the ability of consumers to choose plans over time; we identify and estimate little learning at either the individual or cohort level over the years of our analysis. Inertia does reduce welfare, but even in a world with no inertia we estimate that substantial welfare losses would remain. We conclude that the increased choice inconsistencies over time are driven by changes in plan characteristics that are not offset by substitution both because of inertia and because non-inertial consumers still make inconsistent choices.

Our welfare analyses are sensitive to the assumptions we make about brand preferences; should we regard these as reflecting omitted but desirable features of

---

[1] Earlier drafts of this paper referred to choice set changes as "supply side" and changes in the choice function for a fixed choice set as "demand side."

brands (meaning only within brand errors are possible) or should we regard preferences for particular brands as heuristics which have no actual value to consumers after conditioning on coverage characteristics and quality? In the extreme case in which we assume that only within brand errors are possible because all across brand choices are rationalized by nonparametric brand effects, foregone welfare is on average 60 percent lower. To help assess which normative model is most appropriate, we estimate several random coefficients models to better understand the structure of brand preferences. These models allow brand preferences to rationalize across brand choices but do not require that they do. We find little correlation in preferences for particular brands over time (after conditioning on inertia), but we do find evidence that some beneficiaries consistently prefer more popular brands. We also find that across brand choices are better explained by a large variance in the idiosyncratic error term "ϵ" than by substantial heterogeneity in brand random effects. These results could be interpreted as evidence against the hypothesis that brand choices are rationalized by persistent unobserved quality, although it does not conclusively settle the question of whether estimated brand effects should matter normatively.

Our paper proceeds as follows. In Section I, we review the Medicare Part D program and related previous literature. In Section II, we discuss at length the new CMS Part D data. Section III presents the basic facts on plan choice and plan switching. Section IV describes our empirical strategy for more rigorously modeling choice inconsistencies as well as the reasons for changes in welfare over time, and Section V presents the results of our analysis. Section VI concludes.

I. Background[2]

A. *The Medicare Part D Program*

Medicare, which provides universal health insurance coverage to those over age 65 and to those on the disability insurance program, did not include coverage of outpatient prescription drugs when it was established in 1965; this coverage was added through the Part D program that passed in 2003 and became active in 2006. The most noticeable innovation of the Part D plan is that this new Medicare benefit is not delivered by the government, but rather by private insurers under contract with the government. Beneficiaries can choose from stand-alone plans called Medicare Prescription Drug Plans (PDP) (a plan that just offers prescription drug benefits), Medicare Advantage (MA) plans that provide all Medicare benefits, or existing employer/union plans, so long as coverage is "creditable" or at least as generous (i.e., actuarially equivalent) as the standard Part D plan, for which they would receive a subsidy from the government.

Under Part D, recipients are entitled to basic coverage of prescription drugs by a plan with a structure actuarially equivalent to a standard plan. In 2006 the standard plan offered the following coverage: none of the first $250 in drug costs each year; 75 percent of costs for the next $2,250 of drug spending (up to $2,500 total); 0 percent of costs for the next $3,600 of drug spending (up to $5,100 total, the "donut

---

[2] This section draws heavily on Abaluck and Gruber (2011)—henceforth, AG.

hole"); and 95 percent of costs above $5,100 of drug spending (McClellan 2006). Over 90 percent of beneficiaries in 2006, however, were not enrolled in the standard benefit design, but rather are in plans with low or no deductibles, flat payments for covered drugs following a tiered system, or some form of coverage in the donut hole (McClellan 2006). The main requirement for plans is that they must have equal or greater actuarial value than the standard benefit.

Enrollment in Part D plans was voluntary for Medicare eligible citizens, although Medicare recipients not signing up by May 15, 2006 were subject to a financial penalty if they eventually joined the program (to mitigate adverse selection in the choice of joining the program). One group, however, was automatically enrolled: low-income elders who had been receiving their prescription drug coverage through state Medicaid programs (the "dual eligibles"). These dual eligible were enrolled in Part D plans by default if they did not choose one on their own, and copayments were lowered. There were also special subsidies for lower income elders slightly above Medicaid eligibility.

Despite reluctance voiced before the legislation passed, there was enormous interest from insurers in participating in the Part D program. By November 2006, 3,032 plans were being offered to potential Part D enrollees. Every county in the nation had at least 27 plans available; the typical county had 48 plans, while some counties featured more than 70 choices, primarily due to the high number of MA plans.[3] As of June 2006, there were 10.4 million people enrolled in stand-alone PDP plans, 5.5 million people enrolled in MA plans and about 6 million dual eligibles.[4] Yet 73 percent of people over 65 felt that the Medicare prescription drug benefit was too complicated, while 91 percent of pharmacists and 92 percent of doctors expressed this concern. When asked if they agree with the statement "Medicare should select a handful of plans that meet certain standards so seniors have an easier time choosing," 60 percent of seniors answered "Yes."[5]

Despite these reservations, there were no signs of diminished plan choice in subsequent years. The number of PDPs increased by about 30 percent in 2007, from 1,429 to 1,875 and remained at this level in 2008.[6] By 2012, there were 1,603 PDPs available.

The presentation of options for Part D depends very much on the context in which individuals choose plans. Medicare.gov lays out plan choices in a very clear way and includes a calculator on which our simulator is based. Unfortunately, data in Kling et al. (2012) suggests that only a very small share of enrollees, at least at the time of their survey, use this resource. There is no default plan for initial enrollment, but individuals are auto-reenrolled in their existing plan each year unless they make an affirmative change.

---

[3] Details on number of plans in a median county obtained from Prescription Drug Plan Formulary and Pharmacy Network Files for 2006, provided by CMS.

[4] Enrollment data (rounded) taken from CMS, State Enrollment Data spreadsheet, http://www.cms.hhs.gov/PrescriptionDrugCovGenIn/02_EnrollmentData.asp#TopOfPage. Enrollment numbers also available at http://www.kff.org/medicare/upload/7453.pdf.

[5] Kaiser Family Foundation and Harvard School of Public Health (2006).

[6] Hoadley et al. (2006). Data on 2008 plans taken from CMS 2008 PDP Landscape Source (v. 09.25.07), http://www.cms.hhs.gov/prescriptiondrugcovgenin/.

B. *Issues of Elder Choice in Part D*

Standard economic theory would suggest that Medicare beneficiaries are better off choosing from a wide variety of plans that meet their needs, rather than constraining them to a limited set of choices being made by the government. But, as reviewed in detail in Iyengar and Kamenica (2010), and summarized in AG, there are a large number of behavioral economics models which suggest that in fact agents may be better off with more restrictive choice sets. Other recent literature has shown that the nature of how choices are presented can have important impacts on choice. And work by Agarwal et al. (2009) shows that issues may be magnified for the elderly.

Our work builds on a small existing literature on elder choice in Part D. Most relevant for the current paper is our previous paper (AG) and the work of Ketcham et al. (2012)—henceforth, KLMR.[7] In AG, we use data from Wolters Kluwer (WK), the largest "switch" operator in the prescription drug market that collects the electronic claims from pharmacies and pass them on to the pharmacy benefit managers (PBMs) and insurance companies that will pay the claims. We find in that paper that only about 12 percent of elders choose the cost minimizing plan (the plan which minimizes the sum of premiums plus expected out-of-pocket costs), and that the average elder could save about 30 percent by choosing the cost minimizing plan. We also estimate plan choice models that document key choice inconsistencies. We conclude with a partial equilibrium welfare analysis which implies that welfare would have been 27 percent higher if patients had all chosen rationally.

KLMR focus on plan switching decisions, highlighting the fact that a number of studies document the amelioration of choice biases through repeated market interactions. They turn to data from CVS Caremark, a PBM for a number of PDPs (9 in 2006 and 18 in 2007). KLMR find large choice inconsistencies in 2006, with most individuals in their sample choosing plans that were significantly more expensive than the cost minimizing option. But they find that these inconsistencies are substantially reduced by plan switching in 2007, with the median amount of foregone savings relative to the cost minimizing choice falling by more than $200, and the number of individuals foregoing savings by less than $100 rises from zero to one-third of the sample.

Our work significantly improves on both of these studies along two dimensions. First, we use administrative data sampled from the full universe of Part D enrollees and claims. AG's data captured by the WK system represent only 31 percent of all third party prescription claims filled in the United States, and there is a sizeable rate of attrition from the data (about one-third of the sample per year), which potentially arises from individuals using pharmacies outside of the WK network. KLMR are

---

[7] Other important studies in this area include Heiss, McFadden, and Winter (2007, 2009), who survey a set of elders about their plans for enrolling in Part D programs, and evaluate whether enrollment intentions in the plan were "rational" given the penalties for delay. They find that, the decision over whether to enroll seems to be made rationally for most enrollees, but that only a minority of enrollees facing a hypothetical plan choice chose the cost-minimizing plan. Another important study is by Kling et al. (2012), who examine how providing people with information about the relative costs of each of the available plans in 2007 computed using their 2006 claims impacts their choices. They find that individuals who receive this intervention are more likely to switch plans, and more likely to end up with lower predicted and realized costs. Most recently, Ericson (2014) documents the importance of inertia in plan choice and how it impacts dynamic firm pricing decisions.

only able to follow individuals who stay within their small set of PDPs, and their limited set of plans do not include any plans with donut hole coverage (which AG highlight as a major source of choice inconsistency) so that they may as a result understate inconsistencies.

Our results confirm the findings of both AG and KLMR that there are sizable choice inconsistencies in 2006. But unlike KLMR, we find that choice inconsistencies *worsen* over time, with a large growth in foregone savings from 2006 to 2009. We are unable to replicate KLMR's results in our data so we cannot say for sure what drives this inconsistency. But we suspect that it arises from the limited nature of the plans in their subsample. In particular, they find that most of this reduction comes from switchers moving to more cost effective plans, but their estimated switching rate of 54 percent is much higher than is reported in virtually every other health insurance study.[8,9] Further, while we do replicate the finding that foregone savings declined from 2006 to 2007, we find that in fact, this is explained by consumers choosing cheaper plans with lower quality ratings and our more comprehensive measure which includes these quality ratings suggests that choices did not improve from 2006 to 2007 and grew substantially worse thereafter. We also find that the large foregone savings in 2006 are driven largely by a single low-cost plan; if we compute foregone savings relative to the second lowest cost plan, we see a monotonic increase over time.

The second major contribution of our approach is the use of a structural model which decomposes the reasons for changes in the nature of choice over time. This allows us to not only document these growing choice inconsistencies, but to explain their sources—and in particular to contrast the (limited) role of learning against the (significant) role of choice set changes.

II. Data

The main source of data for our analysis is a 20 percent sample of the newly released universe of claims for Medicare Part D enrollees; we study the subset of claims for beneficiaries enrolled in stand-alone prescription drug plans not receiving any low-income subsidies. These claims are linked to encrypted plan IDs which are linked to a plan characteristics file containing information on premiums, deductibles, and donut hole coverage. Information on plan formularies is inferred from copayment costs.

---

[8] The classic reference on this topic is Neipp and Zeckhauser (1985), who estimate an annual switching rate among health insurance plans of 3 percent in the mid-1980s.

[9] Two studies written in parallel with ours also use the new CMS Part D data to examine the optimality of Part D plan choice. Heiss et al. (2013) finds significant "overspending" in Part D using these data, replicating the earlier work of AG. Heiss et al. (2013) do not emphasize changes over time; they find that by some measures choices improve over time and by others they get worse. A second study by Ketcham, Lucarelli, and Powers argues that more experienced consumers are more likely to switch plans and save money conditional on switching. Neither article puts their findings in a welfare analytic framework that allows them to incorporate the impacts of variance or plan quality. And neither decomposes their findings into choice set and choice function factors in order to understand why the quality of chosen plans is changing relative to the best available plan. Ketcham et al. (2012) in particular focuses on the determinants of switching behavior; our analysis suggests that inertia accounts for only one-third of foregone welfare.

A. *Construction of Out-of-Pocket Cost Variables*

The total enrollee costs of Part D can be decomposed into premiums, which are known for certain at the time of plan choice, and the distribution of out-of-pocket costs given the information available at the time when plans are chosen. Our focus is on estimating the distribution of costs given all of the information potentially available to individuals at the time when they make their choice. However, we only observe realized out-of-pocket costs for the plan in which an individual is enrolled. We therefore assume that the set of claims is fixed and would remain constant had the individual in question chosen a different plan.[10]

We consider two alternative models of out-of-pocket costs: the first is a "realized cost" model in which we construct out-of-pocket costs using the claims incurred by individuals during that year. This amounts to assuming that individuals chose their Part D plans with perfect foresight as to what their claims would be, which is clearly not fully accurate. The second model we consider is a "rational expectations" model in which we compute expected spending in that year based on either prior year claims or claims in the first month of enrollment. We use this model in our regression analysis because it allows us to quantify in a natural way the riskiness of alternative plans. In AG, we show that our main conclusions regarding choice inconsistencies are not sensitive to the choice of model, and are further robust to allowing consumers to have private information so that they know more than we would predict using the previous year's claims but have less than perfect foresight.

Both models rely on the calculator we construct to determine what out-of-pocket costs would be for a fixed set of claims in each of the plans in a beneficiary's choice set. We can check the accuracy of our calculator by comparing simulated out-of-pocket costs using the calculator to actual out-of-pocket costs in the plan in which beneficiaries were actually enrolled. We find a very high correlation of 0.992. The mean difference is $11.6 with a mean absolute deviation of $33 on a mean of $890 (using our baseline sample of full year enrollees). This is an order of magnitude smaller than the cost savings we find, so we do not believe these errors could be driving our results.

Another issue that must be addressed is that we observe only a single realization of out-of-pocket costs for each individual, so we do not observe the variance in spending across choices. We construct this variance in our rational expectation measure based on the distribution of realized costs among individuals who look ex ante identical. In AG, we defined 1,000 cells based on deciles of total expenditure, day's supply of branded drugs and day's supply of generic drugs in the prior year and called individuals ex ante identical if they were in the same cell. Because we do not observe prior year spending in 2006, in our main specifications we define 10 cells based on expenditures in the first month of enrollment and construct the distribution of realized costs among individuals who enrolled in the same month and were in the

---

[10] This assumption of no moral hazard allows us to use the calculator to determine what each individual's realized costs would be for each plan in their choice set. See Appendix B for a discussion of why this should not materially impact our results.

same cell. We show in Appendix Table A1 that our results do not depend on whether we use the 10 cell or 1,000 cell model in 2007 when both models are feasible.[11]

B. *Sample Selection*

We begin with a 20 percent sample of all Medicare Part D beneficiaries in 2006, 2007, 2008, and 2009. This includes 7,213,755 beneficiaries. We then drop all beneficiaries with employer coverage or who are eligible to receive low-income subsidies; in the former case, we lack sufficient information to characterize their plans and in the latter case, there is little variation in cost-sharing given available subsidies. The omission of those with low-income subsidies does not much limit the broader application of our results, say to ACA exchanges, since everyone in this market is receiving some subsidies; the low-income population simply receives extra subsidies.

We additionally drop beneficiaries who change location during the year, who have claims from multiple plans or are enrolled in plans with brands with fewer than 100 enrollees in a state (this last has little impact on the sample but reduces the number of fixed effects in our model). In the primary sample considered in the text, we also restrict to individuals who were enrolled in a Part D plan for the full year. This restriction makes costs more comparable across years (since more consumers enrolled late in 2006) and leaves 1,762,059 beneficiaries. We show in Appendix Table A2 that the trends we observe in foregone savings over time are not sensitive to this restriction.

In our structural models, we consider two samples. The first, which we refer to as the "full sample," includes a 20 percent sample of our data in all years, or a 4 percent sample of Part D enrollees because our data is itself a 20 percent sample (we impose this further restriction because it is computationally more convenient and more than adequate for statistical significance in all of our analysis). The second, which we refer to as the "panel sample," restricts to a 20 percent sample of beneficiaries who are present in our data in 2006, 2007, 2008, and 2009, or a 4 percent sample of all such Part D beneficiaries. Our structural model is estimated using the full sample (new enrollees entering in each year help identify several parameters), but we report simulation results using the panel sample so that changes over time are not confounded by individuals entering and leaving the sample.

Appendix Tables A3 and A4 analyze the reasons why beneficiaries enter and exit the sample in each year for our full sample. The most common reasons for entering the sample are having enrolled after January in the previous year, having claims from multiple plans in the previous year, or being enrolled in a non-PDP plan the previous year (either an MA plan or a plan through one's employer). The most common reasons for exiting are switching from a non-PDP plan, death or having claims from multiple plans, or—in 2009—simply disappearing from the data entirely. Some identification in our analysis below comes from comparing new entrants with returning beneficiaries; new entrants are defined based on their first appearance in

---

[11] This is consistent with our findings in AG, in which the variance measures did not much impact plan choice, and our basic findings based on means were not much impacted by conditioning on the variance of spending across plans as well.
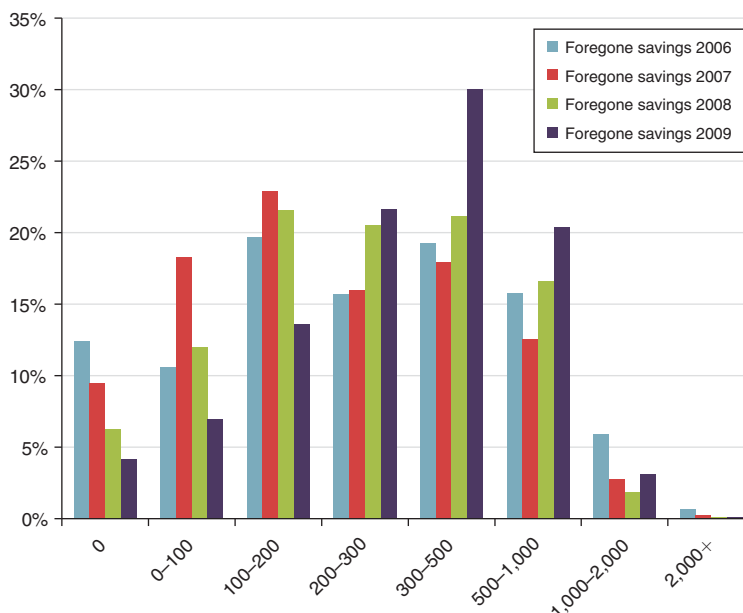
Figure 1. Perfect Foresight Foregone Savings

the beneficiary file, not based on their year of entry into our sample. For example, a beneficiary who turned 65 in 2007 would count as a new entrant; a beneficiary who entered our final sample in 2007 because they enrolled in a month after January in 2006 would count as a returning beneficiary.

III. Plan Choices and Switching: The Facts

We begin our analysis by presenting the basic facts on plan choice in Figure 1. For each individual in the data, we estimate the total cost of enrolling in each PDP plan in their county, adding both premiums and expected out-of-pocket costs. We then estimate the difference in total costs between the plan chosen by that individual and the lowest cost plan in their county, which we call "foregone savings." This corresponds to KLMR's concept of "overspending." Figure 1 conducts this exercise using the perfect foresight model of expectations, while Figure 2 reports the results with our rational expectations model.

As Figure 1 shows, fewer than 15 percent of individuals choose the lowest realized cost plan in their choice set in 2006. On average, individuals could save 26–33 percent of their total Part D spending by choosing the lowest cost plan rather than the plan they chose. We find that half of beneficiaries could have saved more than $330 by enrolling in a different plan. These findings are very similar to those in AG, although they appear smaller than for KLMR for 2006, where mean foregone savings is $550. Figure 2 replicates Figure 1 using our "ex ante" predicted cost measure. Potential savings are smaller according to this measure (as one would expect, since this measure mechanically reduces the variation in out-of-pocket cost across plans) but nonetheless remain substantial with more than half having foregone savings of $230 or more.
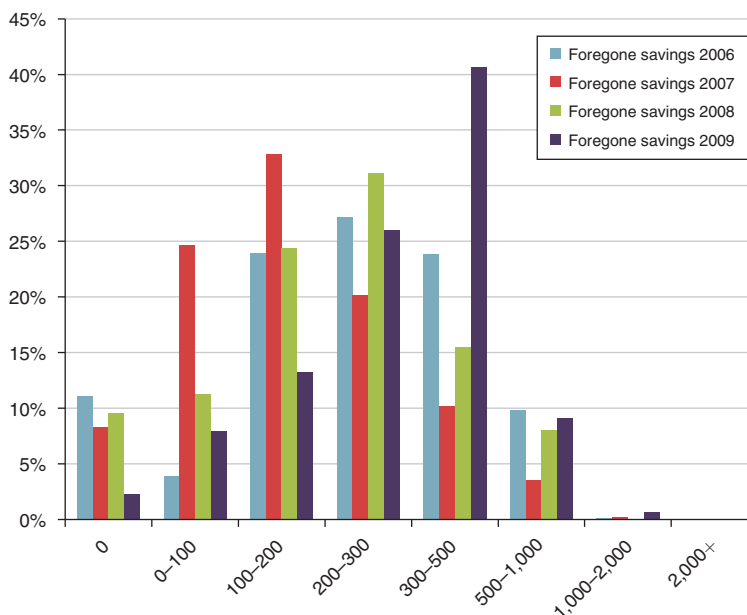
Figure 2. Rational Expectations Predicted Foregone Savings

Notes: Figure 1 shows the distribution of foregone savings per year in the full sample using our perfect foresight measure while Figure 2 gives the distribution using our predicted costs measure. The x-axis gives dollars of foregone savings, and the y-axis the fraction of the total population in that bin.

Unlike KLMR, the CMS data show that the share of individuals making cost minimizing choices does not improve over time. We further find that an ever-falling share of individuals choose the cost minimizing plan. On net, we find that 11 percent of individuals in 2006, 8 percent in 2007, 9 percent in 2008, and just 2 percent in 2009 chose the low-cost plan according to our predicted costs measure.

One reason consumers might not choose the cost minimizing plan is because they are willing to pay more for plans with better risk protection. To deal with this issue, we use an "efficient frontier" measure: we ask what cost savings are possible if consumers are restricted to choosing a plan with weakly better risk protection, as measured by a weakly lower variance of costs. If a plan lies on the efficient frontier, it implies that no alternative plan is both lower in expected costs and better in terms of the risk protection it provides. While plans are required to be actuarially equivalent to the standard plan, for any given individual, a plan might lie off the efficient frontier because it offers financial characteristics which are not valuable to that individual. For healthy individuals, a plan with poor coverage in the initial coverage range and high premiums due to donut hole coverage may be dominated in an efficient frontier sense by a plan with lower premiums, no donut hole coverage, but more generous coverage prior to the donut hole. This point is depicted graphically in Figure 3 where we show the average mean and standard deviation of costs for each plan in the CA choice set. Because we are averaging across all consumers, this does not literally show the efficient frontier for any given beneficiary. Nonetheless, one can see that the majority of plans lie off the frontier demarcated by the lower-left envelope of the scatter plot.
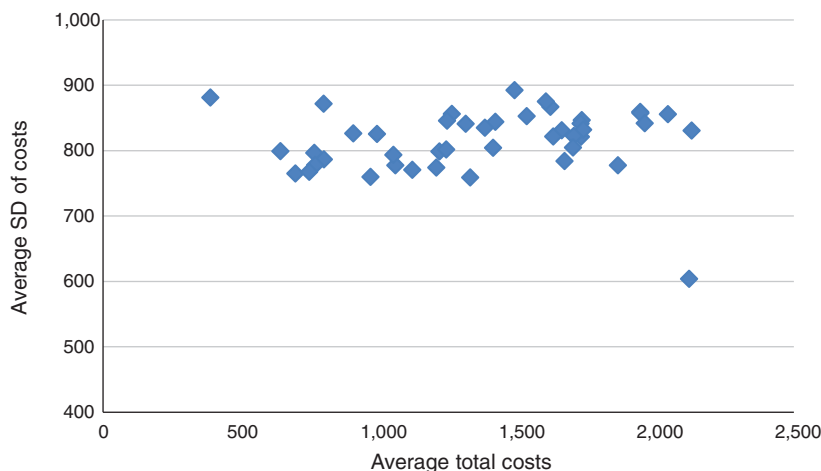
FIGURE 3. CALIFORNIA CHOICE SET, 2006

Notes: Each point represents a single plan. Average total costs are the average of perfect foresight total costs (premiums plus realized out-of-pocket costs).

The results from this efficient frontier measure are summarized in Table 1. We report these results using our predicted cost measure because the variance of costs as a measure of risk only makes sense in that model (in the perfect foresight model, there is no uncertainty and thus no risk). Because we are restricting the set of plans to which one can move to save money, efficient frontier savings are smaller but most beneficiaries could still have saved several hundred dollars in all years without sacrificing risk protection.

Table 1 also reports foregone savings and the efficient frontier measure as a percentage of total costs. The change in this measure overtime is confounded by the fact that better choices reduce both the numerator (foregone savings) and denominator (realized costs). Nonetheless, the magnitude gives a sense of the stakes involved, between 10 percent and 21 percent of total costs for the efficient frontier measure and between 25 and 33 percent for the perfect foresight measure. It is therefore clear that choices did not materially improve over the first three years of the Part D program, and clearly worsened substantially in 2008 and 2009.

We also investigate whether the apparent foregone savings arise because individuals are failing to find a single best plan by comparing choices to the cost of the second least cost plan. We find that in 2006, the foregone savings from the single best plan are almost twice as large as the foregone savings from the second best plan. In subsequent years, savings from the second best plan are 70–85 percent as large as the best plan. These results suggest that there are substantial returns to choosing the best available plan, but also that the results are not driven by a particular low cost plan.[12] Measured relative to the second best plan, foregone savings is monotonically

---

[12] Another important issue is whether there is demographic heterogeneity in foregone savings. The small amount of demographic information available in Medicare claims limits our analysis here. We have examined heterogeneity

TABLE 1—REALIZED OVERSPENDING, PREDICTED OVERSPENDING, AND EFFICIENT FRONTIER OVERSPENDING

	2006	2007	2008	2009
Foregone savings PF ($)	351	277	308	373
Foregone savings PF (%)	26.6	24.9	29.3	32.9
Foregone savings pred ($)	262	176	231	313
Foregone savings pred (%)	18.8	14.7	19.4	24.7
Efficient frontier pred ($)	213	130	174	250
Efficient frontier pred (%)	15.1	10.3	14.6	20.6
Foregone savings—Second best plan ($)	173	198	245	314
Foregone savings pred—Second best plan ($)	128	147	184	269
Number of beneficiaries	538,807	1,344,636	1,316,396	1,179,665

Notes: Table shows various measures of choice quality from 2006 through 2009, both in absolute terms and as a percentage of total costs (computed as the sum of premiums paid and out-of-pocket costs) in the full sample described in the text (as opposed to the full panel). Foregone savings PF gives our perfect foresight measure: realized total costs relative to the plan which minimizes realized total costs (ex post). Foregone savings pred compares predicted costs in the chosen plan to predicted costs in the cost minimizing plan, where predicted costs are computed as average costs among all individual in the same decile of costs in January of that year. Efficient frontier pred gives the same measure, but compares the chosen plan only with plans which have weakly lower variance (computed as the variance in simulated out-of-pocket costs for that plan among 200 beneficiaries in the same decile of January expenditure). Percentages are of total costs in the chosen plan. Savings from the bottom quartile are average savings from choosing a plan in the bottom quartile of total costs. The number of beneficiaries is lower in 2006 because we restrict to beneficiaries enrolled for the entire year, and this restriction excludes a larger share of the sample in 2006 when open enrollment stretched from January to June.

increasing over time and nearly doubling from 2006 to 2009. The drop in foregone savings relative to the best plan is due to a single very low cost plan which was available in 2006 but exited the market in 2007.

These reduced form facts are compelling—but limited. In particular, we have no way of accounting for strong brand preferences. What appear to be choices off the efficient frontier could just be preferences for other aspects of brands that are incorporated into premiums or out-of-pocket costs. This motivates our move toward structural estimation in the next section.

IV. Modeling Plan Choice and Switching

A. *Restrictions on Preferences*

To move from foregone savings to a more comprehensive welfare metric, we consider a structural model of plan choice. We begin by specifying a CARA utility model with a normally distributed cost distribution,

$$(1) \qquad U(C) \; = \; -\exp\!\big(-\gamma(W - C)\big) \quad \text{where} \quad C \sim N(\mu, \sigma^2).$$

We show in our earlier paper that this specification leads to a conditional logit model of plan choice where the utility of individual i from choosing plan j in year t is given by

$$(2) \quad u_{ijt} \; = \; \pi_{jt}\beta_{0it} + \mu^{*}_{ijt}\beta_{1it} + \sigma^2_{ijt}\beta_{2it} + \mathbf{x}_{jt}\gamma_{it} + \boldsymbol{\xi}_{b(j)t} + \xi_{ij=C_{ij}(t-1)}(\mathbf{x}_i) + \epsilon_{ijt}.$$

_____

by age, and we find that foregone savings are fairly similar in percentage terms by age.

In this equation, π_{jt} gives the annual premium of plan j, μ_{ijt}^* gives expected out-of-pocket costs, σ_{ijt}^2 gives the variance of costs, $\mathbf{x}_{jt}$ represents any financial plan characteristics which impact choice, $\boldsymbol{\xi}_{b(j)t}$ represents brand fixed effects, $\xi_{ij=C_{ij}(t-1)}$ is a dummy variable which is 1 if and only if plan j was chosen by consumer i during the previous year, and ϵ_{ijt} are i.i.d. type I extreme value random variables. We allow the inertia dummy to depend on plan characteristics $\mathbf{x}_i$ in two ways: first, the dummy is interacted with the change in characteristics of the previously chosen plan and second, the dummy is interacted with the change in characteristics of the minimum cost plan.

The financial plan characteristics include the deductible of the plan; a dummy for whether the plan covers all donut hole expenditures; a dummy for whether the plan covers generic expenditures in the donut hole only; and a cost-sharing index. The cost sharing index is calculated for each plan as the average percentage of expenditures covered by the plan between the deductible and the donut hole. This variable differs from expected out-of-pocket costs in that it has the same value for everyone in the sample for each plan, and because it is not directly impacted by whether plans have deductibles or donut hole coverage. To control for other aspects of plan quality, we include a full set of brand dummies. This will capture the many features of plans that are common within brand, such as consumer support and pharmacy access.

Identification is a natural concern in this context. All of the plan characteristics included in our model may be endogenous due to unobserved demand factors, and they may be biased by correlation with unobserved plan characteristics. To address this concern, we observe and include in our model much of the publicly available information that might be used by individuals to make their choices—premiums, deductible information, donut hole coverage dummies, as well as variables capturing formulary completeness and cost sharing. We also control for a full set of brand dummies, so that we are only comparing choices of plans with different cost-sharing structure within a given insurer.

Including brand dummies also raises a normative question regarding whether these reflect additional value that consumers receive from plans. One way of interpreting these coefficients is to assume that the brand dummies represent heuristic shortcuts taken by boundedly rational individuals who cannot properly evaluate financial comparisons across plans. In this case, given that we (the econometrician) *can* directly evaluate the financial consequences, brand dummies should *not* count in the normative welfare function. Alternatively, one might interpret the brand dummies as capturing some omitted feature of brands which consumers do value, such as familiarity with the logistics of plans from an earlier experience with a given brand. In this case, brand dummies *should* count normatively in our appraisal of plans. We consider both possibilities in our analysis below. In our baseline model, we do not count brand dummies in the normative welfare function but we do include a quality rating (which summarizes features such as customer service) whose normative weight is recovered by an OLS regression of the brand dummies on the quality variable (that is, we assume that $\xi_{b(j)t} = q_{b(j)t}\delta + v_{jt}$ where only the $q_{b(j)t}\delta$ term counts in the normative welfare function). In an alternative model, we do count brand dummies as part of normative welfare.

A similar normative question arises regarding the inclusion of the inertia term in our normative welfare function. Our measured inertia term may reflect several

different underlying phenomena. Actively choosing a plan may be costly, leading individuals to prefer (all else equal) to just remain in the same plan. Alternatively, our inertia term could represent an adjustment cost that is incurred by consumers if they switch plans in order to learn how the new plan functions. In the adjustment cost story, everything else held equal, consumers may be made worse off if they are forced to enroll in a new plan. In the setting of prescription drug insurance we believe that adjustment costs are small (unlike the setting of health insurance where they might include needing to switch doctors) and so we omit the inertia term from welfare. Once again we consider an alternative treatment of the inertia term in the welfare analysis below.

The model laid out above suggests three natural restrictions on preferences which extend the efficient frontier concept to the discrete choice setting.

Restriction 1: $\beta_0 = \beta_1$.—This restriction states that the coefficient on premiums should equal the coefficient on expected out-of-pocket costs. Controlling for the risk characteristics of plans, individuals should be willing to pay exactly one dollar in additional premiums for coverage which reduces expected out-of-pocket costs by one dollar. If this restriction fails to hold, individuals may choose off the efficient frontier: they could switch to alternative plans with comparable risk characteristics but lower total costs.[13]

Restriction 2: $\gamma = 0$.—This restriction states that financial plan characteristics other than premiums, expected out-of-pocket costs and the variance of out-of-pocket costs do not impact choices. Individuals should not care about deductibles, donut hole coverage, or copays per se; they should only care about these factors to the extent that they impact the distribution of out-of-pocket costs. Once we control for this distribution, these factors should be redundant.

Restriction 3: $\beta_2 < 0$.—This restriction states that individuals should be risk averse.

In this paper, we remain agnostic about the reasons we might observe violations of these restrictions. Consumers might underweight out-of-pocket costs relative to premiums because they construct a mistaken forecast of what their claims will be in the coming year, because they fail to adequately price these claims given each plan's formularies and cost-sharing features, or because they simply do not consider anything beyond the nominal characteristics of alternative plans (such as whether a plan has donut hole coverage). In any of these cases, it may be that consumers are simply erring (for example, not understanding what a deductible is) or it may be that they are behaving rationally given high search costs. All of these cases imply that consumers could be made better off were they enrolled in different plans,

---

[13] Of course, this condition will not hold if expected out-of-pocket costs suffer from measurement error and premiums do not. We consider this concern in detail in Appendix B. In Abaluck and Gruber (2009), we used our perfect foresight measure of expected out-of-pocket spending, instrumented by our rational expectations measure, which is a function only of the category of previous year spending (tantamount to instrumenting by previous year spending category times plan dummy). In this paper we simply estimate the reduced form of that equation, which gives very similar results to the IV due to a first stage coefficient of roughly 1.

and this is the normative judgment we seek to make. There is a legitimate concern that some of these stories might imply structural specifications different from the one we estimate. In Abaluck and Gruber (2009) we consider a model in which out-of-pocket costs receive less weight than premiums because there is a component of out-of-pocket costs which is observable to the econometrician but not observable to consumers. This leads to a model quite close to the one we estimate here, but with some additional random coefficients; in that paper, we find that omitting those random coefficients has little impact on our results.

Appendix D of AG shows how to evaluate welfare in conditional logit models when positive and normative utility functions fail to coincide. Unlike the money metric of foregone savings, the welfare metric we compute takes into account risk aversion and plan quality. That is, for each plan, we compute

$$(3) \qquad W_{ijt} = \frac{1}{\beta_{0it}} \left(\beta_{0it} (\pi_{jt} + \mu_{ijt}^*) + \sigma_{ijt}^2 \beta_{2it} + q_{b(j)t} \delta_{it} \right).$$

This is the welfare measure taking into account total costs, risk protection, and plan quality variables and scaling by the marginal utility of income so that it is expressed in a money-metric. This metric omits from welfare other plan financial characteristics, nonfinancial brand characteristics, the inertia dummy, and the error term. This is the appropriate metric if one believes that the financial characteristics of plans and the plan quality rating capture everything of importance about plans and that—if they matter at all—other factors matter only heuristically because consumers are unable to calculate, or unwilling to spend the time to calculate, the welfare metric above. As noted above, omitting the inertia term from our welfare evaluation is consistent with a story in which the true adjustment costs of switching plans are small. We also consider below models in which brand characteristics matter for welfare.

Let $W_{it}^* = \max_j W_{ijt}$, welfare for the best plan. We define foregone welfare for individual i at time t in plan j as

$$(4) \qquad FW_{it} = \Delta W_{it} = W_{ijt} - W_{it}^*.$$

This is the welfare analogue of our foregone savings measure.

B. *Modeling the Dynamics of Foregone Welfare*

Conceptually, changes in foregone welfare can be driven by changes in how consumers choose from a fixed choice set, choice set factors such as changing plan characteristics, changes in the underlying claims (which determine the welfare consequences of choosing a given plan) or changes in the normative parameters (i.e., how consumers value risk protection and plan quality ratings).

We can further decompose the choice function factors into three effects, which we label inertia, individual learning, and cohort learning. Inertia is the tendency of consumers to remain in the same plan regardless of changes in the plan choice environment. Learning is the tendency for consumer choices to change as they gain experience in the market, individually or collectively. Individual learning concerns whether individuals with experience in the market choose differently conditional

on choosing a new plan than individuals with less experience, while cohort learning reflects calendar year effects—given a fixed choice set, do we see differences in the choice function of each cohort of individuals over time? (controlling for their individual experience in the market)—so we are comparing across years beneficiaries with a given amount of experience. We allow nearly all the structural parameters in our model to vary both by year (cohort learning) and with experience (individual learning)—thus, our model is capable of capturing a large range of ways in which choice functions might change as consumers learn. We seek to understand specifically whether these changes lead consumers to make better choices, so we summarize the impact of each type of learning by evaluating how the estimated changes in structural parameters with experience impact foregone welfare given a fixed choice set.

It is important to note that given our definition, individuals could save money by switching plans in every year but this is not necessarily evidence of learning; that is, learning is not just the complement of inertia. If, for example, all consumers switched whenever foregone welfare exceeded $400 and chose plans with foregone welfare of $300 conditional on switching, this would suggest that they saved money by switching and it would suggest that inertia is making consumers worse off (assuming that inertial consumers would choose as well as switchers were they to switch). It would *not* suggest that consumers who switched plans were learning—because in every year their behavior is the same—they switch plans and choose plans averaging $300 in foregone welfare if their current plan becomes sufficiently unsuitable. It *would* be evidence of learning if either (i) Consumers with more experience in the market systematically chose better plans so switchers did better than new enrollees (individual learning), or (ii) Controlling for choice set differences, consumers in 2007 chose systematically better than consumers in 2006 (cohort learning).

Even if there is no change in consumers' ability to choose over time from a given choice set, plan characteristics may change and choice sets might become more dangerous; for example, suppliers may learn to better conceal costs and otherwise take advantage of consumers' biases. We can decompose choice set changes into changes in which plans are available in different years (which we separate into the impact of plans exiting and entering), changes in plan premiums and changes in plan characteristics which impact out-of-pocket costs.

Formally, we can define each of these effects in terms of changes in the parameters of our structural model (equation (1)). Above, we allowed the structural coefficients β, γ, δ, and ξ to vary flexibly from year-to-year for each individual. We will now place additional structure on this variation. Let $\beta_{it} = \alpha_{1t}^{\beta} + \alpha_2^{\beta} E_{it}$, $\gamma_{it} = \alpha_{1t}^{\gamma} + \alpha_2^{\gamma} E_{it}$, and $\delta_{it} = \alpha_{1t}^{\delta} + \alpha_2^{\delta} E_{it}$ where E_{it} denotes individual i's years of experience in the market in calendar year t. Cohort learning can then be identified with changes in the α_1 coefficients or brand fixed effects, individual learning with changes in the α_2 coefficients, and inertia with the inertial dummy ξ. Choice set changes can be straightforwardly equated with changes in π_{jt}, μ_{ijt}^*, σ_{ijt}^2, and $\mathbf{x}_{jt}$ and $q_{b(j)t}$ over time.[14] Out-of-pocket costs in a given year are a function both of an individual's claims and of the characteristics of the plan in which they are enrolled.

---

[14] Note that conditional on the inclusion of brand-fixed effects, changes in the quality variable impact the normative utility function but not the positive utility function in our baseline specification.

To separately identify the impact of changes in plan characteristics *holding fixed* an individual's claims data, we use our calculator to simulate how out-of-pocket costs would have changed if the characteristics of plans changed from year t to year $t + 1$ but the underlying claims remained fixed at their year t level. The changes in welfare we attribute to the choice set all hold fixed these underlying claims.

We will consider decomposing the change in welfare over time for the panel sample so that all individuals are present in all years; this is not substantively important but it simplifies the exposition. Individuals not in this sample—such as beneficiaries who appear for the first time in 2007—will nonetheless be of use in identifying the structural parameters as we describe below.

Formally, let $\Delta^n_m W^d_s$ denote foregone welfare for individual i, evaluated from the standpoint of the normative preference parameters estimated in year n using that individual's claims in year m. The subscript gives the year from which choice set characteristics are taken, including the plans in the choice set and the premium and cost-sharing features of those plans. The superscript gives the year in which choice function factors are evaluated (i.e., which structural parameters of the choice function are used); these include cohort effects, experience terms, and the inertial tendency.[15] The change in foregone welfare from year t to year $t + 1$ is given by

$$(5) \qquad \Delta FW_{it} = \Delta^t_t W^t_{it} - \Delta^{t-1}_{t-1} W^{t-1}_{i(t-1)}.$$

We can decompose this into choice set (CS), choice function (CF), underlying changes in claims, and normative terms:

$$(6) \qquad CS_{it} = \Delta^{t-1}_{t-1} W^{t-1}_{it} - \Delta^{t-1}_{t-1} W^{t-1}_{i(t-1)}$$

$$CF_{it} = \Delta^{t-1}_{t-1} W^t_{it} - \Delta^{t-1}_{t-1} W^{t-1}_{it}$$

$$C_{it} = \Delta^{t-1}_t W^t_{it} - \Delta^{t-1}_{t-1} W^t_{it}$$

$$N_{it} = \Delta^t_t W^t_{it} - \Delta^{t-1}_t W^t_{it}$$

$$\Delta FW_{it} = CS_{it} + CF_{it} + C_{it} + N_{it}.$$

We further decompose the choice function factors into the welfare impacts of cohort learning, inertia, and individual learning. We decompose choice set welfare effects into the change in welfare induced by changes in premiums, changes in out-of-pocket costs and plan characteristics for a fixed set of claims, and changes in the choice set itself. The claims change in welfare describes how welfare changes due to changes in

---

[15]Note that $I(i,t)$ differs from $I(i,t-1)$ in two ways: first, the plan to which inertia applies may differ if the beneficiary switched plans (in 2007, inertia would apply to the plan chosen in 2006, while in 2008, inertia would apply to the plan chosen in 2007). Second, the weight attached to that plan may differ if the tendency to choose the same plan as in the previous year changes from year to year.

the underlying claims observed. This term is not completely independent of choice set effects—it may for example be driven by utilization responses to changes in plan characteristics. Nonetheless, we find it useful to separate out the changes in welfare due to changes in plan characteristics which impact premiums and out-of-pocket costs for a fixed set of claims and changes in the observed claims. The normative change in welfare, N_{it}, asks how welfare changes due to the fact that preferences may change overtime (e.g., individuals may become more risk averse or the revealed preference value placed on plan quality may change).

In order to estimate each of the above effects, we need to identify the associated coefficients in the structural model. We now discuss the intuition for the identification of each of these effects.

Consider first "individual learning" effects, α_2. Individual learning is identified by comparing the choices of returning consumers *conditional on choosing a new plan* with the choices of new consumers in a given year. A confounding factor is that returning consumers who choose a new plan may be a selected sample of the broader pool of returning consumers—in other words, those consumers who choose to switch might do so in part because they are better at choosing plans. To control for this, we also attempt to identify individual learning by comparing the choices of new beneficiaries and "forced switchers," consumers whose choice in the prior year is no longer available meaning they had no choice but to choose a new plan. This solution is not perfect, however, as forced switchers are not randomly chosen: consumers who choose plans in year $t - 1$, which are no longer available in year t, tend to perform worse than average in year $t - 1$. The comparison between new beneficiaries and active switchers should provide an upper bound on the change in foregone welfare from learning (because active switchers are better than average choosers) and the comparison between new beneficiaries and forced switchers should provide a lower bound on this learning effect (because forced switchers are worse than average choosers).

Consider next inertia, the $\xi_{ij=C_{ij(t-1)}}$ dummies. These are identified by asking how much more likely beneficiaries are to choose the plan they were enrolled in last year than we would expect given the characteristics of that plan. The welfare impact of inertia is identified by comparing the choices of switchers and non-switchers; it depends on whether non-switchers would have been better off had they actively chosen a plan. This is not necessarily the case even if they are not already enrolled in the best plan: the choice set might become more dangerous and active choices might lead them to choose even more poorly than if they just remained in the same plan. We can assess whether they would be made better off by switching by assuming that had they switched, they would have chosen as well as switchers (either active or forced).

Consider finally the "cohort learning" effects, α_1. Cohort learning is the impact of market experience on choice functions controlling for any direct individual experience with the market. Some channels for this type of learning might be a greater abundance of tools which help consumers choose better or increasing knowledge on the part of healthcare providers such as pharmacists and doctors who consumers turn to for advice. These effects can be identified by comparing beneficiaries in different years with the same amount of experience: for example, new beneficiaries in 2006 and 2007 or beneficiaries with a single year of experience in 2007 and 2008.

In the estimates we report here, the model uses both sources of variation to identify cohort learning.[16]

Identification of the choice set factors is straightforward. Given the estimated structural parameters, we can use the model to simulate how choices (and ultimately welfare) differ when premiums, out-of-pocket costs, or the available plans change.[17] Intuitively, one can think of the choice set estimates as analogous to the usual "area under the demand curve" welfare measures. If premiums of Plan A increase by $100, welfare does not fall by $100 because consumers can substitute toward alternative plans: the amount by which welfare falls depends on the degree of substitution via the price elasticity.

V. Structural Results

We begin by reporting the estimated coefficients in the structural model. Table 2 reports our baseline specification as described above. The first column shows our results for 2006, the year studied in AG. The coefficients are the structural coefficients in a conditional logit model and not marginal effects. They can be roughly interpreted as the impact of a one unit increase in the variable of interest on the probability that a plan is chosen; a premium coefficient of -0.68 implies that a $100 increase in premiums decreases the probability that a plan is chosen by 68 percent (this interpretation holds exactly for plans which are a negligible share of the overall market).

We find that even with this improved data, the choice inconsistencies described above and documented in AG persist. First, there is a sizable gap between the estimated premium and out-of-pocket cost coefficients, with the former often two to three times as large as the latter. That is, consumers are substantially overweighting premiums relative to out-of-pocket cost exposure in choosing their insurance plans.

We also find that there are significant coefficients on several plan characteristics: the deductible, donut hole coverage variables, and formulary variables all matter even after controlling for out-of-pocket cost variables. The magnitude of these coefficients is substantial: we observe that, *controlling for out-of-pocket cost consequences*, individuals are willing to pay $350 to obtain donut hole coverage in 2006 and substantially more in 2007 (the ratio of the donut hole coefficient to the premium coefficient).

The coefficient on the variance is insignificant in most specifications, and even where significant it is close to zero in magnitude: foregone welfare changes by less than $6 relative to a world where the variance coefficient is identically zero. As we saw in our efficient frontier analysis, we cannot rationalize consumers' foregone savings as arising because consumers prefer plans with superior risk protection.

While these restrictions follow naturally from utility maximization with full information and standard preferences, the model from which they are derived makes several important functional form assumptions. Even in models which considerably

---

[16] Any estimated cohort learning effects could also potentially incorporate differential advertising to new versus existing Part D enrollees.

[17] As noted above, this involves using our plan calculator tool to compute how out-of-pocket costs would change counterfactually were the underlying claims held fixed.

TABLE 2—CONDITIONAL LOGIT MODEL COEFFICIENTS WITH BRAND FIXED EFFECTS

Brand dummies	2006	2007	2008	2009
Premium (hundreds)	−0.68	−0.69	−0.96	−0.80
	(0.03)	(0.06)	(0.07)	(0.06)
OOP (hundreds)	−0.20	−0.38	−0.35	−0.58
	(0.03)	(0.04)	(0.04)	(0.09)
Variance ($\times 10^6$)	0.41	0.44	−0.34	−2.70
	(0.52)	(0.29)	(0.52)	(0.74)
Inertia	X	3.90	6.48	7.07
		(0.43)	(0.07)	(0.11)
Deductible (hundreds)	−0.96	−0.25	−0.81	−0.40
	(0.09)	(0.04)	(0.03)	(0.05)
Full donut hole coverage	2.39	5.11	X	X
	(0.17)	(0.41)		
Generic coverage	0.69	0.71	1.28	1.60
	(0.04)	(0.07)	(0.08)	(0.10)
Cost sharing	−7.88	0.79	−0.63	−1.41
	(0.88)	(0.35)	(0.33)	(0.42)
Number of top 100 on form	0.08	0.33	0.44	−0.03
	(0.01)	(0.02)	(0.02)	(0.01)
Quality rating	0.63	0.27	0.76	0.60
	(0.02)	(0.01)	(0.02)	(0.03)

Notes: Table shows conditional logit results from estimating equation (2) by maximum likelihood. Standard errors are in parentheses. In addition to the coefficients reported here, all specifications include brand fixed effects, separate coefficients for active and forced switchers (only those for active switchers are reported), interactions between all the reported coefficients and experience variables, and interactions between the inertia coefficient and (demeaned) values of the plan characteristics for the minimum cost plan, deciles of expenditure in the previous year, and characteristics of the chosen plan in the previous year. In later years, separate inertia dummies are included for the plan one was enrolled in during each preceding year. The average quality variable is a normalized version of the average rating index provided by CMS, recovered by auxiliary regression of estimated brand fixed effects on the quality rating variable.

loosened these assumptions, however, we would still not expect to see these rejections of consistency. Appendix Table A5 carries out simulations where we assume that consumers maximize expected utility given CARA or CRRA utility, the empirically estimated distribution of costs, and different levels of risk aversion. We then take the simulated choices and estimate our conditional logit model as if the simulated choices were the true choices observed in the data. In these simulations, we sometimes see significant coefficients on plan characteristics, but they are several orders of magnitude smaller. This suggests that the significant coefficients we find in the data are not due to the linearization or normality assumptions we make in our theoretical model, but rather due to the fact that consumers attach special weight to these characteristics beyond their personalized out-of-pocket cost consequences.

Importantly, our findings do *not* imply that consumers err primarily by choosing plans with low premiums but poor coverage. When we omit plan characteristics entirely from our model (Table 3), we find that the resulting coefficients on out-of-pocket costs are substantially *larger* than the coefficients on premiums. This suggests that if anything, beneficiaries may be choosing plans whose coverage is too generous (in other words, they are paying higher premiums to choose plans which provide more generous coverage on average, but not for their particular mix of drugs). Taken together, the results with and without plan characteristics imply

TABLE 3—CONDITIONAL LOGIT MODEL COEFFICIENTS WITH BRAND FIXED EFFECTS, OTHER
PLAN CHARACTERISTICS OMITTED

Brand dummies	2006	2007	2008	2009
Premium (hundreds)	−0.21	−0.39	−0.58	−0.52
	(0.04)	(0.05)	(0.04)	(0.07)
OOP (hundreds)	−0.31	−0.50	−0.93	−0.93
	(0.06)	(0.06)	(0.09)	(0.09)
Variance ($\times 10^6$)	1.12	−1.03	1.72	−2.91
	(0.30)	(0.55)	(1.30)	(0.69)
Inertia	X	3.64	5.63	6.66
		(0.10)	(0.04)	(0.14)
Quality	0.59	0.25	0.55	0.61
	(0.02)	(0.01)	(0.03)	(0.03)

Note: Table shows conditional logit results from estimating equation (2) by maximum likelihood; the specification is identical to Table 2, but all financial plan characteristics are omitted.

that enrollees are making two sets of "mistakes": overweighting plan characteristics that do not apply to them, while underweighting relative to premiums more subtle coverage characteristics that do reduce the cost of the particular drugs they consume.

Of course, this means that our finding of choice inconsistency in this respect is generated by the joint finding of a premium-OOP gap in tandem with the fact that plan characteristics enter the model conditional on OOP costs. This raises the question of whether our model with plan characteristics included is really separately identified from the simpler model that excludes these (theoretically irrelevant) broader plan characteristics. Indeed, we do find that the two models make very different predictions. In Appendix B, we show that 12 percent of choices are predicted to change if we drop plan characteristics from our model and the average change in the probability that a plan is chosen is 20 percent.

The remaining columns of Table 2 show that the stylized facts from the 2006 results persist in 2007, 2008, and 2009. We observe that the premium and out-of-pocket cost coefficients move closer together. The specifications include a flexible set of interactions between the inertia dummy and plan characteristics both in the present year and in the previous year. These interactions allow the decision of whether to switch to depend flexibly both on the current year choice set and specifically on changes to the plan in which the consumer was previously enrolled. With these interactions included, the reported coefficients are identified only by the choices of switchers *conditional on switching*. The fact that the premium-OOP coefficient gap narrows in later years relative to 2006 reflects the fact that these coefficients are identified only by the 10 percent of individuals who switched plans.

Despite the fact that we are focusing on switchers, however, the premium and out-of-pocket cost coefficients remain significantly different. Moreover, the other choice inconsistencies persist: other plan characteristics are highly significant and the variance term is insignificant or small in magnitude. The inertia term is extremely large in magnitude, reflecting the fact that the vast majority of consumers remain in the same plan they chose in the previous year. The inertia term can be interpreted to mean that consumers are 500–700 percent more likely to choose a plan if it is the plan they chose last year; comparing the inertia term to the premium coefficient, we

Table 4—Conditional logit Model Coefficients
with Brand Fixed Effects and Forced Switchers

Brand dummies and forced switchers	2007	2008	2009
Premium (hundreds)	−0.70	−0.90	−0.87
	(0.07)	(0.04)	(0.06)
OOP (hundreds)	−0.41	−0.62	−0.60
	(0.08)	(0.06)	(0.05)
Variance ($\times 10^6$)	1.05	−0.56	−0.87
	(0.52)	(0.62)	(0.41)
Inertia	X	X	X
Deductible (hundreds)	−0.06	−0.67	−0.30
	(0.10)	(0.04)	(0.03)
Full donut hole coverage	5.01	X	X
	(0.59)		
Generic coverage	1.82	1.78	1.41
	(0.08)	(0.19)	(0.09)
Cost sharing	2.35	−3.26	1.16
	(0.96)	(0.72)	(0.58)
Number of top 100 on form	0.96	0.81	0.06
	(0.11)	(0.05)	(0.01)
Quality rating	0.72	0.03	0.61
	(0.04)	(0.07)	(0.02)

Notes: Table shows conditional logit results from estimating equation (2) by maximum likelihood. Standard errors are in parentheses. This table differs from Table 2 in that we report the coefficients on each variable interacted with whether you were a forced switcher (meaning that the plan in which you were enrolled the previous year is no longer available). In addition to the coefficients reported here, all specifications include brand fixed effects, interactions between all the reported coefficients and experience variables, and interactions between the inertia coefficient and (demeaned) values of the plan characteristics for the minimum cost plan, deciles of expenditure in the previous year, and characteristics of the chosen plan in the previous year. In later years, separate inertia dummies are included for the plan one was enrolled in during each preceding year. The average quality variable is a normalized version of the average rating index provided by CMS, recovered by auxiliary regression of estimated brand fixed effects on the quality rating variable.

find that consumers are willing to give up more than $600 worth of premiums to remain in the same plan.

Table 4 considers the results when we condition on forced switchers. In this case we obtain slightly higher coefficients on out-of-pocket costs but the sizable gap between the premium and out-of-pocket cost coefficients remains, as does the significance of the coefficients on other plan characteristics (note that forced switchers only exist from 2007 onward because they are defined as Part D enrollees whose choice in the previous year was discontinued). The similarities between the choices of switchers and forced switchers suggest that it is not unreasonable to use the observed choices of switchers to model how inertial consumers would choose were they forced to switch.

A. *Heterogeneous Preferences*

An additional implication of our model is that beneficiaries sacrifice a great deal of money to choose particular brand names. The range of brand fixed effects (expressed in dollars of premiums) is more than $1,100. The panel nature of our data

Table 5—Restricted Sample Random Coefficient Model Coefficients

Brand dummies	2006	2007	2008	2009
Premium (hundreds)	−0.97	−0.92	−0.86	−1.03
	(0.05)	(0.06)	(0.05)	(0.07)
OOP (hundreds)	−0.20	−0.51	−0.26	−0.64
	(0.03)	(0.07)	(0.08)	(0.11)
Variance ($\times 10^6$)	−0.16	−0.26	−0.59	−2.33
	(0.45)	(0.73)	(0.84)	(1.10)
Inertia (plan)	X	1.78	3.96	5.85
	X	(2.28)	(0.32)	(0.42)
Inertia (brand)	X	2.07	2.69	1.82
	X	(0.13)	(0.15)	(0.14)
Deductible (hundreds)	−1.35	−0.37	−0.73	−0.62
	(0.09)	(0.08)	(0.07)	(0.09)
Full donut hole coverage	3.55	5.88	X	X
	(0.33)	(0.58)	X	X
Generic coverage	1.27	1.52	1.35	3.06
	(0.18)	(0.17)	(0.22)	(0.34)
Cost sharing	−10.66	0.96	0.54	−3.70
	(0.84)	(0.99)	(0.89)	(1.07)
Number of top 100 on form	0.09	0.48	0.37	−0.19
	(0.01)	(0.14)	(0.09)	(0.07)

Notes: Table shows coefficients results from estimating equation (2) by simulated maximum likelihood given the correlational structure specified in Appendix C. Standard errors are in parentheses. In addition to the coefficients reported here, all specifications include brand random effects, separate coefficients for active and forced switchers (only those for active switchers are reported), interactions between all the reported coefficients and experience variables, and interactions between the inertia coefficient and (demeaned) values of the plan characteristics for the minimum cost plan, deciles of expenditure in the previous year, and characteristics of the chosen plan in the previous year. In 2007–2009, separate inertia dummies are included for the plan one was enrolled in during each preceding year as well as separately for the brand one was enrolled in.

allows us to investigate in some detail the structure of these preferences—are they persistent over time or are they idiosyncratic, and are they correlated across similar types of plans? Additionally, by allowing for persistent brand preferences, we prevent these factors from loading onto our estimated inertial term.

To address these issues, we estimate an extension of our baseline model in which brand fixed effects are heterogeneous and normally distributed and can be decomposed into brand and time components: $\xi_{i,b,t} = \xi_{ib}^B + \xi_{it}^T$ where we allow for arbitrary correlations between ξ_{ib}^B and $\xi_{ib'}^B$ and between ξ_{it}^T and $\xi_{it'}^T$. Estimation of this model is described in more detail in Appendix C. Because it is much more computationally demanding than our baseline model, we use a 2.5 percent random sample and restrict to beneficiaries who chose amongst the 11 most popular brands; this yields 9,127 beneficiaries. In addition to our usual plan-level inertial terms, we add inertial terms at the brand level to aid in separate identification of brand preferences. In the Appendix we show the results of estimating our baseline model on this more restricted sample; the results are very similar to the full sample.

Table 5 shows the results from estimating this richer model that incorporates heterogeneity. The stylized facts regarding choice inconsistencies reported in the previous section continue to hold (in this case, the inertial coefficients from our original model can be compared to the sum of the plan and brand coefficients in the more flexible specification). Firstly, we note that—comparing the inertial coefficients in

Table 5 with Appendix Table A6—the inertial coefficients are hardly impacted; this suggests that rather than reflecting persistent unobserved factors of chosen plans, they reflect either adjustment costs or inattention.

Table 6A reports the estimation correlation matrix for the time and brand components of the estimated random effects. In each case, the random effects are identified relative to an omitted brand (the most popular brand) and an omitted time (2006). The time random effects show little evidence of either heterogeneity (the diagonal terms) or persistence (the off-diagonal terms). In other words, any changes across time in brand preferences are well-captured by shifts in means. This does not conclusively tell us whether these brand preferences should count normatively—uncorrelated brand fixed effects could represent some changing aspect of customer service not captured by our quality rating, or they could represent changes in advertising that increase demand in some years but not others. Nonetheless, the lack of correlation across beneficiaries over time does weigh against the possibility that brands which look bad on observable dimensions have some persistent compensating factor.

The across brand random effects show more evidence of both heterogeneity and correlation. The standard deviation of the brand random effects ranges from $125 to $181 (this is the square root of the non-normalized covariance coefficient divided by the premium coefficient); this suggests that beneficiaries vary in the weight they attribute to different plans, but the range of the variation is substantially less than the $1,100 difference in average valuation between the most and least desirable plan. We also find some evidence of correlation amongst brand preferences; while only a few of the off-diagonal terms are significant, all are positive, implying that beneficiaries are correlated in their preference for the most popular plan (which is normalized to 0).

To investigate this correlation further, we estimate a version of the model which classifies all brands into terciles based on overall popularity across all years (Table 6B). This model suggests both that preferences for popular plans are correlated (so beneficiaries who prefer brands in the most popular tercile are also more likely to prefer the second tercile to the least popular brands) and that preferences for popular brands are persistent, at least in 2008 and 2009. Thus, while there is little tendency for beneficiaries who idiosyncratically preferred a given brand to prefer that same brand in later years, we do find that beneficiaries who idiosyncratically preferred popular brands have a consistent preference for popular brands over time.

B. *Welfare Implications*

We next use our structural model to move from the results for foregone savings presented above to a welfare metric which includes not only expected spending but also risk protection and plan quality based on the revealed preference weights estimated by the model.

Table 7 shows the results of this welfare analysis in our baseline model. For comparison purposes, the first two rows replicate in the panel sample our earlier results using the predicted spending measure, with foregone savings falling from 2006 to 2007 but then rising from 2007–2009, even relative to plans with similar or lower variance (Table 1 reports the analogous results in the full sample). The third row

TABLE 6A—BRAND RANDOM COEFFICIENT COVARIANCE STRUCTURE

	2007	2008	2009
Panel A. Covariance matrix of ξ_{it}^{T} (in units of premiums)			
2007	13.72	−6.53	−2.98
	(134.79)	(32.15)	(18.45)
2008	−6.53	9.57	−7.09
	(32.15)	(27.64)	(42.78)
2009	−2.98	−7.09	21.37
	(18.45)	(42.78)	(118.38)

Panel B. Covariance matrix of ξ_{ib}^{B} (in units of premiums)

	Brand 2	Brand 3	Brand 4	Brand 5	Brand 6
Brand 2	159.40	75.74	93.31	100.74	80.68
	(42.47)	(37.37)	(52.36)	(180.34)	(94.58)
Brand 3	75.74	316.67	73.41	140.66	109.01
	(37.37)	(61.26)	(70.75)	(452.23)	(75.13)
Brand 4	93.31	73.41	151.74	107.34	129.55
	(52.36)	(70.75)	(85.86)	(51.09)	(52.06)
Brand 5	100.74	140.66	107.34	138.52	130.95
	(180.34)	(452.23)	(51.09)	(285.31)	(264.54)
Brand 6	80.68	109.01	129.55	130.95	187.55
	(94.58)	(75.13)	(52.06)	(264.54)	(169.85)
Brand 7	75.61	195.83	104.50	132.95	155.66
	(72.52)	(93.92)	(63.49)	(102.41)	(120.46)
Brand 8	82.03	60.97	155.33	116.93	170.06
	(59.56)	(51.24)	(89.52)	(190.84)	(87.28)
Brand 9	149.90	70.59	99.89	96.77	89.41
	(39.44)	(73.06)	(102.20)	(120.50)	(64.14)
Brand 10	139.84	123.86	91.02	116.70	81.24
	(44.67)	(73.84)	(88.34)	(85.85)	(69.44)
Brand 11	84.68	137.76	124.47	144.11	145.79
	(46.88)	(137.35)	(119.16)	(133.59)	(92.91)

	Brand 7	Brand 8	Brand 9	Brand 10	Brand 11
Brand 2	75.61	82.03	149.90	139.84	84.68
	(72.52)	(59.56)	(39.44)	(44.67)	(46.88)
Brand 3	195.83	60.97	70.59	123.86	137.76
	(93.92)	(51.24)	(73.06)	(73.84)	(137.35)
Brand 4	104.50	155.33	99.89	91.02	124.47
	(63.49)	(89.52)	(102.20)	(88.34)	(119.16)
Brand 5	132.95	116.93	96.77	116.70	144.11
	(102.41)	(190.84)	(120.50)	(85.85)	(133.59)
Brand 6	155.66	170.06	89.41	81.24	145.79
	(120.46)	(87.28)	(64.14)	(69.44)	(92.91)
Brand 7	199.41	116.96	87.16	94.33	129.94
	(125.04)	(76.89)	(82.16)	(43.35)	(78.51)
Brand 8	116.96	193.88	92.66	70.96	138.40
	(76.89)	(70.32)	(63.02)	(89.22)	(123.68)
Brand 9	87.16	92.66	163.15	132.57	74.14
	(82.16)	(63.02)	(44.21)	(43.73)	(87.11)
Brand 10	94.33	70.96	132.57	161.03	114.09
	(43.35)	(89.22)	(43.73)	(42.89)	(62.96)
Brand 11	129.94	138.40	74.14	114.09	186.01
	(78.51)	(123.68)	(87.11)	(62.96)	(108.98)

Notes: Panel A reports the estimated covariance matrix for the time random effects and panel B reports the estimated covariance matrix for brand random effects, as specified in Appendix C. The time random effects are defined relative to 2006—the diagonal terms are the variance of the difference between time component of the random effect in the specified year and the time component in 2006. The off-diagonal terms are the correlation between these time components. The brand random effects are defined relative to the omitted brand (which is chosen to be the most popular brand). All coefficients are expressed in dollar terms by dividing the estimated covariance parameters by the premium coefficient in 2006.

TABLE 6B—TERCILE OF BRAND POPULARITY RANDOM COEFFICIENT COVARIANCE STRUCTURE

	2007	2008	2009
Panel A. Covariance matrix of time tercile random effects			
2007	0.08	−0.25	−0.37
	(0.12)	(0.19)	(0.23)
2008	−0.25	1.08	1.23
	(0.19)	(0.30)	(0.33)
2009	−0.37	1.23	1.84
	(0.23)	(0.33)	(0.65)

Panel B. Covariance matrix of brand tercile random effects	Tercile 2	Tercile 3
Tercile 2	81.99	76.17
	(45.23)	(40.28)
Tercile 3	76.17	78.74
	(40.28)	(39.30)

Notes: Table 6B is identical to 6A, but for the model where brand random effects are replaced with random effects for terciles of brand popularity. Panel A reports the estimated covariance matrix for the time component of these random effects and panel B reports the estimated covariance matrix for the brand component of these random effects, as specified in Appendix C. The brand random effects are defined relative to the omitted tercile (which is chosen to be the least popular brands). All coefficients are expressed in dollar terms by dividing the estimated covariance parameters by the premium coefficient in 2006.

TABLE 7—FOREGONE SAVINGS, EFFICIENT FRONTIER SAVINGS, AND FOREGONE WELFARE

Foregone welfare	2006	2007	2008	2009
Foregone savings PF ($)	351	277	308	373
Efficient frontier pred ($)	213	130	174	250
Welfare (baseline)	161	170	211	251
Number of beneficiaries: baseline	55,585	55,585	55,585	55,585
		2006–2007	2007–2008	2008–2009
Inertia welfare loss (baseline model)		12	66	60

Notes: First panel shows various measures of choice quality from 2006 through 2009. The first two rows replicate the corresponding rows of Table 1 in our full panel (beneficiaries who appear in the data for a full year in all years from 2006–2009) and all subsequent rows report results in the full panel sample. Welfare (baseline) reports our welfare adjusted measure, computed as total costs plus the dollar equivalent value of risk protection and quality ratings estimated in our logit model with brand fixed effects as described in equation (3). The second panel uses the model to simulate the impact of inertia on welfare relative to a world with no inertia in which inertial consumers choose from among all plans given the estimated structural coefficients identified primarily from the choices of switchers. As noted in the text, the model is estimated on the full sample but the results are reported in the full panel. The reported value is the welfare loss in a world with inertia relative to a world with no inertia (so a negative number implies a welfare gain from inertia).

shows the welfare equivalent in our baseline model: foregone welfare increases over time, rising from about $161 in 2006 to $251 in 2009. This occurs despite the fact that mean expenditures in our final sample are lower in 2009 than in 2006 ($2,150 in 2006 compared to $2,100 in 2009). In contrast to foregone savings, welfare is increasing monotonically over time because the one low-cost plan in 2006 is rated poorly according to the plan quality measure. Thus, while we find as in Ketcham et al. (2012) that foregone savings fall from 2006 to 2007, our more comprehensive

welfare measure suggests that this arises mainly because some consumers in 2007 enrolled in cheap, low quality plans.

The bottom row in Table 7 shows the contribution of inertia to foregone welfare from 2007 onwards. In all years, the contribution of inertia is less than one-fourth of foregone welfare. In 2007, those who stayed behind in the same plan rather than switching lost only $12 relative to had they switched and chosen as well as switchers. In 2008 and 2009 in the base model, the inertia estimates become more sizable, but still remain only a small fraction of the total foregone welfare. This is striking evidence that, in general, switching never fully offset the choice inconsistencies among Part D consumers. We find that switchers did save $245 relative to had they not switched, but our model suggests that this was due to the fact that their prior year plan was particularly unsuitable; comparable savings was not available to inertial consumers if they had chosen as well as the switchers.

C. *Alternative Normative Assumptions*

There are a number of alternative normative assumptions we consider to understand better what is driving our welfare results. The results of these exercises are also reported in Table 8. The first row of Table 8 replicates our above computation of foregone welfare in the restricted sample used to estimate the model with heterogeneous brand effects and subsequent rows consider alternative normative assumptions in that model. Foregone welfare tends to be slightly lower in the restricted sample since there are fewer alternative plans from which to choose the best plan.

A natural intuition is that the fewer restrictions we impose on the normative model—in other words, the more factors we allow to matter for welfare—the more choices we will be able to rationalize and the better choices will look. Our analysis suggests that this need not be the case—in a second best world in which there are multiple potential sources of mistakes, allowing additional factors to matter in the normative utility function can worsen the welfare cost of the remaining departures from full rationality.

Firstly, to understand better the role played by omitted plan characteristics, we compute foregone welfare in a model in which all observable characteristics (including the deductible and donut hole) are assumed to matter for welfare; this contrasts with our baseline model in which these characteristics only enter welfare via the distribution of out-of-pocket costs. In this model, reported in row 2 of Table 8, foregone welfare actually tends to be higher than in our baseline model (row 1). This counterintuitive result arises because, while chosen plans tend to have high values of desirable plan characteristics, the best among the many unchosen plans tend to have even higher values of these characteristics.

One critical driver of our results is the assumption that brand preferences do not count normatively. The model with heterogeneous brand preferences (reported in Tables 5 and 6) allows us to explore this assumption further. In particular, we consider several alternative cases:

(i) Our baseline model—with fixed brand preferences—and in which brand preferences are not relevant for welfare (row 1 of Table 8);

TABLE 8—ALTERNATIVE NORMATIVE ASSUMPTIONS

	2006	2007	2008	2009
Welfare (restricted sample)	135	122	154	243
Welfare (restricted, normative plan characteristics)	196	128	184	362
Welfare (restricted, positive brand random effects)	150	125	159	240
Welfare (restricted, normative brand random effects)	160	154	206	261
Welfare (restricted, normative nonparametric brand random effects)	64	49	100	61
Welfare (restricted, no inertial errors)	135	14	11	17
Number of beneficiaries: restricted	9,127	9,127	9,127	9,127

Notes: Table shows various measures of choice quality from 2006 through 2009 in conditional logit and random coefficients models. The first row replicates row 3 of Table 7 in the restricted sample; this reports foregone welfare in our conditional logit model, computed as total costs plus the dollar equivalent value of risk protection and quality ratings. Welfare (restricted, normative plan characteristics) uses an alternative metric in which financial plan characteristics are included in the welfare computation. Welfare (restricted, positive brand random effects) uses the same metric as the first row, but in the model with brand random effects. Welfare (restricted, normative brand random effects) allows these random effects to enter welfare (and computes welfare using simulated choices). Welfare (restricted, normative nonparametric brand random effects) computes foregone welfare relative to the best plan in the same brand as the chosen plan. Welfare (restricted, no inertial errors computes foregone welfare using only the choices of switchers and averaging in zeros for all inertial beneficiaries (assuming that all inertial choices are rational).

(ii) The case in which brand preferences are heterogeneous and normally distributed with the correlation structure specified in Table 6A above, but are not welfare relevant (row 3 of Table 8);

(iii) The case in which brand preferences have the same parametric structure as in (ii) but are welfare relevant (row 4 of Table 8); and

(iv) The case in which brand preferences are unrestricted and fully welfare relevant (row 5 of Table 8).

In case (iv), foregone welfare is mechanically much lower because the only possible mistakes are choosing the wrong plan within the chosen brand (foregone welfare falls from an average of $168 to $69). We nonetheless find comparable welfare losses in cases (i)–(iii). Case (iii) would mechanically approach case (iv) as the variance of the preference heterogeneity becomes arbitrarily large. The fact that case (iii) is much closer to case (ii) than case (iv) reflects the fact that the choice inconsistencies we estimate are not well explained by (parametric) heterogeneous brand preferences.

We can think of this exercise as follows. We observe variation in choices not explained by observables. This variation could reflect either heterogeneous brand preferences or noise (the epsilon term). The nonparametric test in (iv) assumes that all of the across brand noise reflects heterogeneous brand preferences. The parametric test attempts to separately identify brand preferences from noise. The identifying assumption is that—so long as brand preferences are additively separable—more heterogeneity in brand preferences will mean less sensitivity to plan characteristics (such as premiums) in across brand choices than within brand choices. If choices are fairly sensitive to premiums within brands but not sensitive to premiums across brands, this would suggest that the across brand choices are driven by preferences

for particular brands; on the other hand, if choices are as sensitive to premiums across brand as within brand, this suggests little heterogeneity.

While we estimate statistically significant heterogeneity, this heterogeneity explains only a small fraction of the residual variation in choices in our model. Note that this is not merely a consequence of our parametric assumption: as the variance of the random coefficients grows larger in the parametric model, the results converge to the nonparametric model in which only within brand errors are possible. Our results suggest, however, that there is only a small amount of heterogeneity in brand preferences. Therefore, foregone welfare in the fourth row of Table 8 is very close to the first row.

Moreover, the fourth row of Table 8 shows that allowing brand preference heterogeneity to count normatively increases foregone welfare. This is a counterintuitive result: one might expect that allowing brand preferences to matter for normative welfare would reduce foregone welfare because it would help to rationalize the observed patterns of choices across brands. To gain some intuition, it is helpful to imagine a world where there is a single "best" plan and all consumers have a 10 percent chance of choosing the best plan and a 90 percent chance of choosing randomly from amongst 40 or so other plans in their choice set. Adding brand fixed effects makes the best plan look even better from a normative standpoint (a disproportionate share choose it given the fact that people otherwise choose randomly) but this increases foregone welfare since most people do not choose the best plan. This contrasts with the world in which preferences are highly heterogeneous, nearly all choices are driven by brand preferences, and thus all across brand choices are welfare maximizing. Note that whether choices reflect idiosyncratic noise or heterogeneous brand preferences, there is a separate question as to whether these factors should enter the model normatively. Our finding here is that, to the extent that the two can be distinguished, the former factors explain a much larger share of choices.

A second important normative driver of our results is the assumption that inertia is welfare irrelevant—that is, there are no true adjustment costs to switching plans and no additional unobserved features of plans which make chosen plans desirable. Fully relaxing this assumption would imply that only switchers can err. Row 6 of Table 8 reports foregone welfare in this scenario; it is less than one tenth as large as all years after 2006 because it is now mechanically zero for all non-switchers (note that this is distinct from the exercise above where we simulated a world with no inertia but still allowed inertial consumers to have welfare loss as they chose only as well as switchers). Once again, we cannot settle the question of whether inertia represents inattention or true adjustment costs given the data available, but this is an important question for future work with large normative implications. Ho, Hogan, and Scott Morton (2015) present some evidence for the inattention hypothesis in the Part D setting, arguing that those who switch plans do so in response to shocks to their existing plan as opposed to just the possibility of savings from switching. This supports the normative assumption in our baseline specification.

D. *Structural Decomposition*

In this section, we attempt to understand the factors driving changes in welfare over time. To make the changes over time more transparent, we eliminate the

TABLE 9—STRUCTURAL DECOMPOSITION

	2007	2008	2009
Year $t - 1$ foregone welfare	161	170	211
Change due to choice set + choice function	26	68	39
Choice function	21	21	0
Individual learning	3	9	8
Inertia	12	20	0
Cohort learning	7	−8	−8
Choice set	4	47	39
Premiums	−1	106	37
OOP	128	14	74
Plans (exit)	−133	−73	−76
Plans (entry)	10	1	4

Notes: Table reports the results of the welfare decomposition exercise described in equations (3)–(8). As noted in the text, the model is estimated on the full sample but the results are reported in a randomly chosen subset of the full panel with 75,829 beneficiaries per year (the remaining full panel beneficiaries after the model is estimated on a 10 percent sample of the full sample). The first row is identical to the third row of Table 7 and reports foregone welfare from the previous year. The next row reports the sum of the changes due to choice set and choice function factors. The first row of the Choice function and Choice set panels gives the sum of the changes due to each of their respective components. The Inertia term represents the change in welfare due to the *change* in inertia between years which differs from the exercise in Table 7 where we report the impact of inertia relative to a world with no inertia.

normative terms and the claims term in the decomposition above and evaluate welfare in all years using 2006 claims as well as the risk preferences and quality ratings estimated in 2006. This restriction has little impact on the results but it makes more transparent the degree to which choice set and choice function factors are impacting welfare. Overall, we find that choice set factors predominate, and that increasing premiums combined with consumer inertia and the exit of the most generous plans over time lead foregone welfare to increase and had an even more dramatic effect on absolute welfare once we control for changes in claims over time.

The first panel in Table 9 shows the level of foregone welfare in the previous year, and the second row shows the change due to choice set and choice function factors holding fixed 2006 claims. The remainder of the table shows the contribution of each choice set and choice function side factor to the overall change in welfare.

Our results suggest that choice function measures were relatively unimportant. Individual learning, inertia, and cohort learning all led welfare to change relatively little over time. The one exception is a small increase in foregone welfare due to beneficiaries becoming more inertial in 2008 (while roughly the same fraction of consumers switched, the potential savings from switching—and thus the magnitude of the inertia term in our model—were larger in 2008). But as highlighted above, the problem here is not simply inertia which accounts for only a small fraction of foregone welfare; it is that even those who switch do not remove choice inconsistencies when they do so.

Table 9 also shows that choice set changes substantially increased foregone welfare over time. We see the same pattern across all years. Foregone welfare increases both due to changes in premiums and changes in plan generosity. This was partly but not fully offset by reductions in foregone welfare due to plan exit.

TABLE 10—STRUCTURAL DECOMPOSITION: ABSOLUTE WELFARE

	2007	2008	2009
Year $t-1$ absolute welfare	−1,318	−1,309	−1,310
Change due to choice set + choice function	−169	−87	−72
Choice function	−21	−21	0
Individual learning	−3	−9	−8
Inertia	−12	−20	0
Cohort learning	−7	8	8
Choice set	−148	−67	−72
Premiums	−27	−41	−76
OOP	-31	81	144
Plans (exit)	−93	−109	−139
Plans (entry)	3	2	−2

Notes: Table reports the results of the welfare decomposition exercise described in equations (3)–(6) using absolute welfare rather than foregone welfare. As noted in the text, the model is estimated on the full sample but the results are reported in a randomly chosen subset of the full panel. The first row reports absolute welfare from the previous year. The next row reports the sum of the changes due to choice set and choice function. The first row of the Choice function and Choice set panels gives the sum of the changes due to each of their respective components. The Inertia term represents the change in welfare due to the *change* in inertia between years which differs from the exercise in Table 7 where we report the impact of inertia relative to a world with no inertia.

We are unable to distinguish from Table 9 whether the effects are due to changes in the chosen plan or changes in the best foregone plan. To address this, in Table 10 we present the same decomposition of choice set factors but for absolute, not foregone, welfare. Absolute welfare is normalized to 0 for a plan with zero premiums, zero out-of-pocket costs, mean risk aversion, and mean quality rating; dollar equivalent welfare is then computed for each plan relative to this normalized plan. In this case more negative numbers mean a larger reduction in welfare. While absolute welfare is roughly constant over time, this is because claims are increasing; when we isolate choice set and choice function factors, we see that absolute welfare is falling over time—consumers are being made worse off for a given set of claims. The choice function impacts are identical to those for foregone welfare because changes in choice functions impact only the chosen plan, not the best available plan. The choice set effects are different because choice set changes impact both the chosen plan (which would impact both absolute and foregone welfare) and the best available plan (which only impacts foregone welfare).

The absolute welfare results add some nuance to the foregone welfare results. In all years, we can see that consumers are being made worse off by premium increases as well as the exit of desirable plans from the market. They are made better off by increases in plan generosity among existing plans; the fact that plan generosity changes tended to increase foregone welfare suggests that generosity increased even more for the best unchosen plan.

Thus, we find that over time in the Part D program there was an exit of the most generous plans, with a shift toward higher premiums and more generous coverage among plans that remain in the program. More generous coverage tended to increase foregone welfare because the best unchosen plans increased their coverage even more, and the exit of generous plans tended to reduce foregone welfare since unchosen plans were even more likely to exit. In 2006, increases in coverage generosity failed to offset premium increases; in 2007 and 2008 they more than offset premium

TABLE 11—STRUCTURAL DECOMPOSITION: NORMATIVE BRAND DUMMIES

	2007	2008	2009
Year $t-1$ foregone welfare	224	178	303
Change due to choice set + choice function	−25	78	22
Normative	−15	79	11
Choice function	−26	2	6
Individual learning	5	6	10
Inertia	−37	11	4
Cohort learning	6	−15	−7
Choice set	1	76	16
Premiums	1	147	60
OOP	95	−37	57
Plans (exit)	−113	−36	−105
Plans (entry)	18	2	3

Notes: Table reports the results of the welfare decomposition exercise described in equations (3)–(6). Table 11 differs from Table 9 in that all welfare evaluations use a normative model in which the dollar-equivalent of brand fixed effects is included. As noted in the text, the model is estimated on the full sample but the results are reported in the full panel. The first row reports foregone welfare from the previous year. The next row reports the sum of the changes due to choice set and choice function. The first row of the Choice function and Choice set panels gives the sum of the changes due to each of their respective components. The Inertia term represents the change in welfare due to the *change* in inertia between years which differs from the exercise in Table 7 where we report the impact of inertia relative to a world with no inertia.

increases but this effect was outweighed by the exit of desirable plans. Thus, taken together, these choice set changes reduced absolute welfare and increased foregone welfare.

Table 11 replicates the results in Table 9 with normative brand dummies. The stylized facts noted on the choice function side persist. Likewise, in both cases, choice set changes are predominantly responsible for the increase in foregone welfare over time.

VI. Conclusion

The bold experiment with consumer choice across health insurance plans embodied in the Medicare Part D program provides an excellent opportunity to assess how consumers perform in choosing insurance plans. We find that, using the best available data, consumers are very inconsistent in their choices, overweighting premiums relative to out-of-pocket costs, weighting plan characteristics above and beyond the effect on that consumer, and ignoring variance in coverage across plans. Moreover, we find that these choice inconsistencies persist over time, and that the foregone welfare from choosing inconsistently rises during the first four years of the Part D program.

This increase in foregone welfare occurs primarily through choice set changes. The set of plans available to consumers were changing over time in a way which heightened the disadvantages of choosing poorly. And this was not offset by factors on the choice function side. Most consumers remain inertial, passively allowing the changes in the choice set to impact them. Most strikingly, even those consumers who are not inertial continue to choose inconsistently, so even the limited amount of switching that takes place does not do much to offset the welfare losses. Ultimately,

we conclude that there was little learning at both the individual and cohort levels over time within the Part D program.

Of course, the choice inconsistencies we identify do not necessarily mean that consumers are behaving in a sub-optimal way given the information and computational tools available to them. It would be consistent with our results for consumers to rationally be choosing not to exert the time and effort to choose the very best plan available, but this still means that consumers could be made better off were they enrolled in different plans. This suggests the value of continued improvement in decision support tools available to enrollees in the Part D program. A recent GAO report highlighted the value of continued CMS oversight of the Planfinder tool on Medicare.gov (on which our calculator is based) (US GAO 2014).

A shortcoming of our analysis is that we do not endogenize the choice set: we document how choice sets changed over time but we do not examine the factors driving these changes in equilibrium. If the model developed here were supplemented by a model of competition between firms for behavioral consumers, we could simulate how policy changes such as providing additional information to consumers or new rules governing the types of benefits firms may offer might impact the premium setting and entry and exit of firms and thereby consumer welfare. Ericson (2014) shows that some of the changes in premiums reported here could arise in a model with firms pricing to inertial consumers. Analyzing the choice set evolution is especially complicated due to the rules governing assignment of low-income subsidy enrollees to plans which seem to play an important role in determining premiums (Decarolis 2015). A full model should take into account these rules, as well as two additional supply-side effects: informing consumers could exacerbate adverse selection (Handel 2013) but more information could intensify price competition by eliminating the artificial differentiation which arises from consumer confusion.

APPENDIX A: CALCULATOR CONSTRUCTION

Our calculator tool determines for a fixed set of claims what out-of-pocket costs would have been for that set of claims for each of the plans in an individual's choice set. In Abaluck and Gruber (2009), we considered several different versions of the calculator tool allowing varying degrees of substitution across drugs if individuals choose different plans. We found that this made little difference for our final results—here, we make the assumption that individuals substitute toward the cheapest (lowest tiered) alternative within each branded/generic class and Generic Code Number (GCN), a classification which groups together all products with identical ingredients, strength, dosage, and route of administration.

The first major task is to reconstruct the appropriate formularies for each plan from the observed claims—the formulary should tell us, for each (plan, branded/generic ID, GCN) the lowest tier in which a drug with that branded/generic ID and GCN is available and whether or not the drug is included in the plan formulary at all. This file is constructed empirically based on the observed claims of enrollees in each plan. The CMS Part D Events file contains drugs classified by National Drug Code (NDC) number; we use a crosswalk provided by First Data Bank to determine the associated GCN for each of these drugs.

CMS provides beneficiary cost files linked to encrypted plan IDs which determine the copays or coinsurances associated with a given day's supply of a drug in a given tier for each coverage range. Given our formulary files, we can determine the appropriate tier (including no tier) for each drug an individual consumes and given the beneficiary cost files we can determine the appropriate copays and coinsurances. We run each beneficiary's claims through our calculator in chronological order keeping track of cumulative expenditures to date (and thus coverage range) in order to compute each beneficiary's total out-of-pocket costs. There are a large number of subtleties omitted in this explanation (e.g., how we determine copays/coinsurances when days supply is not an exact multiple of 30, how we order claims if they appear chronologically on the same day among many others). The source code for our calculator which details all of these assumptions is available upon request.

Appendix B: Mismeasured Out-of-Pocket Costs

An important concern with the restrictions imposed by our model is that our out-of-pocket cost variable may be mismeasured. If this is the case, then what our model attributes to beneficiaries overweighting salient plan features may actually reflect the true financial value of those characteristics not adequately captured by our out-of-pocket cost variables.

Consider three alternative types of measurement error:

(i) Measurement error due to calculator error;

(ii) Measurement error due to private information; and

(iii) Measurement error due to moral hazard.

In Section II, we note that our calculator is extremely accurate in replicating observed out-of-pocket costs and that errors in our calculator do not appear to be correlated with choice quality. This suggests that the first type of error should be small.

Measurement error due to private information might arise if beneficiaries have information about their future expenditures beyond what can be forecast given our predicted cost model. We can test for this directly (and control for it) by including in our model the difference between perfect foresight out-of-pocket cost and predicted out-of-pocket costs and checking whether this term explains choices after conditioning on out-of-pocket costs. Abaluck and Gruber (2009) formally derive such a model. If we partial out the other covariates from realized out-of-pocket costs, we can also interpret the model as instrumenting for realized out-of-pocket costs using our predicted costs measure. The rationale is that realized costs C_{ij} can be written as $C_{ij} = \mu_{ij} + e_{ij} + \eta_{ij}$ where μ_{ij} is our predicted costs measure, e_{ij} represents a component of costs observables to beneficiaries at the time when they choose but not to the econometrician, and η_{ij} represents a component of realized costs unknown even to beneficiaries at the time when they choose. We assume that μ_{ij} and e_{ij} are independent of η_{ij}. The coefficient on realized costs in a choice model would be biased toward zero due to the measurement error from η_{ij}, but this problem can be rectified by instrumenting for C_{ij} with μ_{ij}. This allows us to take into account the

possibility that beneficiaries have some private information and allow for it in our model. Any remaining weight attached to plan characteristics after conditioning on predicted costs and the control function constructed by regressing realized costs on predicted costs and covariates should not reflect private information. Appendix Table A7 shows the estimated coefficients in our model when this realized cost term is added. While we find evidence of private information, including this term has little impact on the estimated plan characteristics.

A third potential source of mismeasurement in our out-of-pocket cost variable is moral hazard. We can consider the impact of moral hazard in a stylized model. Suppose that total drug spending S is a linear function of the coinsurance rate so that we can write $S = a - b \cdot CI$ and out-of-pocket costs are a linear function of total spending, $OOP = S \cdot CI$. Let CI_0 denote the coinsurance rate in the plan in which the individual was actually enrolled and let CI_j denote the coinsurance rate of an alternative plan j. Appendix A of Abaluck and Gruber (2009) shows that with these assumptions, there is an omitted term in our choice model equal to $\frac{1}{2} b CI_j (CI_j - CI_0)^2$ which captures the welfare gain from adjusting one's consumption if one chooses a plan different from the one which was actually chosen. A first point to note is that our assumption of no moral hazard should bias downward our estimates of foregone welfare because we are ignoring behavioral adjustments which make alternative plans more valuable. Second, we find that the absolute magnitude of this omitted term given prescription drug elasticities estimated in the literature is small—the additional value due to moral hazard of adjusting ones claims after moving from a plan with 25 percent coinsurances to 75 percent coinsurances is just \$21 (Abaluck and Gruber 2009 gives explicit calculations). For these reasons, we do not believe that moral hazard accounts for the choice inconsistencies we estimate.

A final concern with our model is whether it is empirically identified. For example, does our model yield different predictions relative to a model in which plan characteristics are omitted but premiums and OOP costs are roughly equally weighted? To see whether it does, we perform a simulation exercise. We estimate our model with and without plan characteristics included. Assuming identical i.i.d. error terms across the two simulations (so that none of the differences are due to noise alone), we then simulate choices in the models with and without plan characteristics included.

The results of this exercise strongly suggest that plan characteristics are identified based on empirical features of choices. In both models, the average probability that a plan is chosen is 0.025; the average absolute change in the probability a plan is chosen is 0.0049, or about a 20 percent difference. 7,719 of 64,231 choices are predicted to change in our simulation sample (itself a 10 percent sample of the full sample used in the paper). In other words, our model suggests that 12 percent of beneficiaries are predicted to choose different plans in the model fit to the data with plan characteristics omitted (a number greater than the percentage who are actually observed to switch plans in the data).

Appendix C: Random Coefficients Model

The random coefficients model with heterogeneous brand preferences is given by equation (2), but, as discussed in the text, we assume that the brand fixed effects

in equation (2) can be replaced by brand random effects. These are heterogeneous and normally distributed and can be decomposed into brand and time components: $\xi_{i,b,t} = \xi_{ib}^B + \xi_{it}^T$. where $\text{cov}\left(\xi_{ib}^B, \xi_{ib'}^B\right) \equiv \sigma_{bb'}^B$ and $\text{cov}\left(\xi_{it}^T, \xi_{it'}^T\right) \equiv \sigma_{tt'}^T$ are unrestricted but $\text{cov}\left(\xi_{ib}^B, \xi_{it}^T\right) \equiv 0$ for all b and t. Let B denote the number of brands included in estimation (excluding the most popular brand which is omitted) and T the number of time periods.

Because the model is much more computationally demanding than our baseline model, we use a 2.5 percent random sample and restrict to beneficiaries who chose amongst the 11 most popular brands; this yields 9,127 beneficiaries. In addition to our usual plan-level inertial terms, we add inertial terms at the brand level to aid in separate identification of brand preferences.

We estimate this model by maximum simulated likelihood. To improve the computational efficiency, we use scrambled Halton sequences (following Train 2003). This approach also means that the likelihood function is deterministic so that we can analytically compute the gradient to aid in estimation. To impose appropriate normalizations, we construct Σ, the covariance matrix of the vector of random effects given by: $\left\{\xi_{i2}^T - \xi_{i1}^T, \ldots, \xi_{iT}^T - \xi_{i1}^T, \xi_{i1}^B + \xi_{i1}^T, \ldots, \xi_{iB}^B + \xi_{i1}^T\right\}$. Note that we can write $\xi_{i,b,t}$ in any year as a linear combination of elements of this vector (the sum of the corresponding brand and time terms). Denote by Σ_B the covariance matrix of random effects in year 1 and Σ_T the matrix of time effects relative to year 1. Σ_B is the bottom-right $B \times B$ elements of Σ while Σ_T is the top-right $T - 1 \times T - 1$ elements. These are the brand and time covariance matrices reported in the text.

To ensure that all covariance matrices are symmetric and positive definite, we cannot directly estimate parameters of these matrices. Instead, we write $\Sigma = \begin{bmatrix} \Sigma_T & \Sigma_{\text{cov}} \\ \Sigma_{\text{cov}'} & \Sigma_B \end{bmatrix}$ where every row of Σ_{cov} is identical and is given by $\left\{\sigma_{2,1}^T, \ldots, \sigma_{T,1}^T\right\}$. We can Cholesky decompose Σ into $\mathbf{LL}'$ where $= \begin{bmatrix} \mathbf{L}_T & \mathbf{0} \\ \tilde{\mathbf{L}}_{\text{cov}} & \tilde{\mathbf{L}}_B \end{bmatrix}$, $\mathbf{L}_T \mathbf{L}_T' = \Sigma_T$ is the Cholesky decomposition of Σ_T, $\tilde{\mathbf{L}}_{\text{cov}} = (\mathbf{L}_T^{-1} \Sigma_{\text{cov}})'$ where and $\tilde{\mathbf{L}}_B \tilde{\mathbf{L}}_B' = \Sigma_B - \tilde{\mathbf{L}}_{\text{cov}} \tilde{\mathbf{L}}_{\text{cov}'}$. The parameters we estimate are the elements of the lower triangular matrix L_T, the $T - 1$ elements of each row of Σ_{cov}, and the lower triangular matrix $\tilde{\mathbf{L}}_B$.

To compute standard errors, we use the negative of the Hessian of the likelihood function to recover the standard errors of the estimated parameters and then use the delta method to compute the standard errors of the elements of the covariance matrices.

TABLE A1—1,000 CELL MODEL VERSUS JUST JANUARY MODEL

	January (2007)	1,000 cell (2007)
Foregone savings PF ($)	306	306
Foregone savings PF (%)	25	25
Foregone savings pred ($)	193	195
Foregone savings pred (%)	16	16
Efficient frontier pred ($)	152	153
Efficient frontier pred (%)	12	12
Number of beneficiaries	5,275	5,275

Notes: Table replicates the corresponding rows of Table 1 in the text in the final sample for which reports are resulted in the structural model. In this sample, we compare the predicted savings measures computed by conditioning on decile of January expenditures with a model in which we assign each individual to one of 1,000 cells based on deciles of prior year expenditures and days supply of branded and generic drugs (as in AG). The latter model is not feasible in 2006 since we do not observe prior year data but it is feasible in 2007.

TABLE A2—REALIZED OVERSPENDING, PREDICTED OVERSPENDING, AND EFFICIENT FRONTIER OVERSPENDING USING ALL ENROLLEES REGARDLESS OF MONTH OF ENROLLMENT

	2006	2007	2008	2009
Foregone savings PF ($)	241	268	299	361
Foregone savings PF (%)	26	25	29	33
Foregone savings pred ($)	172	171	231	302
Foregone savings pred (%)	1	15	21	24
Efficient frontier pred ($)	137	127	177	242
Efficient frontier pred (%)	13	11	16	20
Number of beneficiaries	1,497,250	1,433,225	1,382,194	1,238,630

Notes: Table shows various measures of choice quality from 2006 through 2009, both in absolute terms and as a percentage of total costs (computed as the sum of premiums paid and out-of-pocket costs) in the full sample described in the text using all enrollees from January–June. Foregone savings PF gives our perfect foresight measure: realized total costs relative to the plan which minimizes realized total costs (ex post). Foregone savings pred compares predicted costs in the chosen plan to predicted costs in the costs minimizing plan, where predicted costs are computed as average costs among all individuals in the same decile of costs in the first month of enrollment in that year. Efficient frontier pred gives the same measure, but compares the chosen plan only with plans which have weakly lower variance (computed as the variance in simulated out-of-pocket costs for that plan among 200 beneficiaries in the same decile of expenditure during the first month of enrollment).

TABLE A3—ENTRANTS BY YEAR

	By year			
	2006	2007	2008	2009
Sample	538,807	1,345,328	1,316,396	1,179,665
Number who entered in year *t*		913,791	174,721	203,535
Entered because not Jan. in prior year		859,242	50,963	48,803
Turned 65		6,116	7,294	6,503
Geographic location change		10	18	14
Claims from multiple plans		24,643	4,209	4,038
Dual or LIS in prior year		7,325	2,476	1,949
No PDP in previous year		16,433	109,757	142,226
No data in previous year		22	4	2

TABLE A4—EXITS BY YEAR

	By year			
	2006	2007	2008	2009
Sample	538,807	1,345,328	1,316,396	1,179,665
Number who exited following year t	107,270	203,653	340,266	
Died	21,466	55,462	46,383	
Geographic location change	37	747	76	
Claims from multiple plans	18,482	52,715	43,338	
Dual or LIS in prior year	3,854	8,664	9,618	
Switched to non-PDP plan	61,401	80,779	93,355	
No data in following year	2,030	5,286	147,496	

Notes: Appendix Table A3 decomposes all reasons for entry into our main sample (which includes only beneficiaries enrolled for the full year). A beneficiary enters the sample in year t if they were present in year t but not in $t - 1$. Beneficiaries enter either because they were excluded from the prior year sample for enrolling after January, because they turned 65, because their geographic location changed (some territories such as the US Virgin Islands are excluded from the sample), because they were previously excluded for having claims from multiple plans, because they were previously excluded for receiving low-income subsidies, because they were previously not enrolled in a PDP plan or because there is simply no record of them having existed in the previous year. Appendix Table A4 details the reasons for exit from the sample. Beneficiaries exit because of death, because of geographic location changes, because they had claims from multiple plans in a subsequent year, because they became eligible for low-income subsidies, because they switched to a non-PDP plan or because they disappeared entirely from the data.

TABLE A5—CRRA VERSUS CARA ESTIMATES

	CRRA (wealth = 17,000)			CARA		
	1	3	10	0	0.0003	0.0005
Premium (hundreds)	−5.39	−5.1	−3.4	−5.3	−4.77	−3.88
	(0.04)	(0.00)	(0.00)	(0.00)	(0.03)	(0.02)
OOP cost (hundreds)	−5.36	−4.9	−2.8	−5.3	−4.52	−3.38
	(0.04)	(0.00)	(0.00)	(0.00)	(0.03)	(0.03)
Variance ($\times 10^6$)	−1.90	−6.30	−20.00	−2.60	−9.24	−16.28
	(0.15)	(0.10)	(0.20)	(0.20)	(0.14)	(0.19)
Deductible (hundreds $)	0.01	0.03	0.08	0.01	0.04	0.07
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Donut hole (hundreds $)	0.01	0.05	−0.1	0.0	0.0	0.04
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Generic donut hole (hundreds $)	0.01	0.03	0.05	0.01	0.02	0.05
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)

Notes: Table excerpted from Appendix Table 1 in Abaluck and Gruber (2009). The table shows the results of estimating the model using simulated choices assuming the utility function and risk parameter listed in a given column. Each column shows coefficients from a single regression. The coefficients reported are the parameters of the utility function, not marginal effects. Standard errors are in parentheses. The sample differs slightly from that in Table 1 because individuals with greater than 17,000 in total costs for any plan are dropped. All simulated choices are based on the cost distribution generated from the realized costs of 200 individuals in the same decile of 2005 total costs, decile of 2005 total days supply of branded drugs, and decile of 2005 days supply of generic drugs. The first three columns compute expected utility using a CRRA utility function with wealth of 17,000 and the indicated coefficient of relative risk aversion, assuming that individuals select the choice which maximizes expected utility. The final three columns compute expected utility using a CARA utility function with the indicated coefficient of absolute risk aversion.

TABLE A6—CONDITIONAL LOGIT RESTRICTED SAMPLE COEFFICIENTS

Brand dummies	2006	2007	2008	2009
Premium (hundreds)	−0.91	−0.89	−0.81	−1.00
	(0.04)	(0.06)	(0.04)	(0.05)
OOP (hundreds)	−0.19	−0.51	−0.25	−0.63
	(0.03)	(0.06)	(0.07)	(0.09)
Variance ($\times 10^6$)	0.10	−0.12	−0.48	−2.34
	(0.41)	(0.61)	(0.75)	(0.89)
Inertia (plan)	X	1.67	3.84	5.66
		(1.56)	(0.30)	(0.34)
Inertia (brand)	X	1.90	2.53	1.60
		(0.09)	(0.10)	(0.10)
Deductible (hundreds)	−1.26	−0.36	−0.73	−0.60
	(0.08)	(0.07)	(0.06)	(0.07)
Full donut hole coverage	3.43	5.77	X	X
	(0.28)	(0.72)		
Generic coverage	1.21	1.50	1.30	2.97
	(0.17)	(0.15)	(0.17)	(0.27)
Cost sharing	−10.24	0.64	0.18	−4.08
	(0.73)	(0.88)	(0.77)	(0.83)
Number of top 100 on form	0.09	0.38	0.23	−0.15
	(0.01)	(0.10)	(0.08)	(0.06)

Notes: Table shows conditional logit results from estimating equation (2) by maximum likelihood using our restricted sample. Standard errors are in parentheses. In addition to the coefficients reported here, all specifications include brand fixed effects, separate coefficients for active and forced switchers (only those for active switchers are reported), interactions between all the reported coefficients and experience variables, and interactions between the inertia coefficient and (demeaned) values of the plan characteristics for the minimum cost plan, deciles of expenditure in the previous year, and characteristics of the chosen plan in the previous year. In later years, separate inertia dummies are included for the plan one was enrolled in during each preceding year.

TABLE A7—CONDITIONAL LOGIT MODEL COEFFICIENTS WITH BRAND FIXED EFFECTS AND REALIZED COST

Brand dummies	2006	2007	2008	2009
Premium (hundreds)	−0.68	−0.71	−0.96	−0.82
	(0.03)	(0.06)	(0.07)	(0.06)
OOP (hundreds)	−0.02	−0.22	−0.22	−0.37
	(0.03)	(0.04)	(0.04)	(0.10)
Realized OOP (hundreds)	−0.16	−0.19	−0.16	−0.23
	(0.01)	(0.02)	(0.01)	(0.02)
Variance ($\times 10^6$)	0.61	0.53	−0.32	−2.96
	(0.56)	(0.29)	(0.51)	(0.76)
Inertia	X	3.82	6.42	7.13
		(0.43)	(0.07)	(0.12)
Deductible (hundreds)	−0.96	−0.25	−0.80	−0.41
	(0.09)	(0.04)	(0.03)	(0.05)
Full donut hole coverage	2.30	4.84	X	X
	(0.17)	(0.36)		
Generic coverage	0.70	0.69	1.26	1.59
	(0.04)	(0.06)	(0.08)	(0.10)
Cost sharing	−7.96	0.99	−0.44	−1.26
	(0.93)	(0.36)	(0.32)	(0.45)
Number of top 100 on form	0.08	0.33	0.45	−0.03
	(0.01)	(0.02)	(0.02)	(0.01)
Quality rating	0.64	0.27	0.77	0.61
	(0.02)	(0.01)	(0.02)	(0.03)

Notes: Table is identical to Table 2, but with realized OOP included in addition to predicted OOP. The rationale for this, and the interpretation as a test of private information, is described in Appendix B.

REFERENCES

Abaluck, Jason T., and Jonathan Gruber. 2009. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." NBER Working Paper 14759.

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

Abaluck, Jason, and Jonathan Gruber. 2016. "Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance: Dataset." *American Economic Review*. http://dx.doi.org/10.1257/aer.20130778.

Agarwal, Sumit, John Driscoll, Xavier Gabaix, and David Laibson. 2009. "The Age of Reason: Financial Decisions over the Life-Cycle with Implications for Regulation." *Brookings Papers on Economic Activity* (Fall): 51–117.

Decarolis, Francesco. 2015. "Medicare Part D: Are Insurers Gaming the Low Income Subsidy Design?" *American Economic Review* 105 (4): 1547–80.

Ericson, Keith M. 2014. "Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange." *American Economic Journal: Economic Policy* 6 (1): 38–64.

Handel, Benjamin R. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review* 103 (7): 2643–82.

Heiss, Florian, Adam Leive, Daniel McFadden, and Joachim Winter. 2013. "Plan Selection in Medicare Part D: Evidence from Administrative Data." *Journal of Health Economics* 32 (6): 1325–44.

Heiss, Florian, Daniel McFadden, and Joachim Winter. 2006. "Who Failed to Enroll in Medicare Part D, and Why? Early Results." *Health Affairs* 25 (5): w344–54.

Ho, Kate, Joseph Hogan, and Fiona Scott Morton. 2015. "The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program." NBER Working Paper 21028.

Hoadley, Jack, Elizabeth Hargrave, Juliette Cubanski, and Tricia Neuman. 2006. "An In-Depth Examination of Formularies and Other Features of Medicare Drug Plans." Menlo Park, CA: Henry J. Kaiser Family Foundation.

Kaiser Family Foundation and the Harvard School of Public Health. 2006. "Survey Finds Few Seniors Expect To Switch Medicare Drug Plans For 2007." http://benefitslink.com/pr/detail.php?id=40246 (accessed April 28, 2011).

Ketcham, Jonathan D., Claudio Lucarelli, Eugenio J. Miravete, and M. Christopher Roebuck. 2012. "Sinking, Swimming, or Learning to Swim in Medicare Part D." *American Economic Review* 102 (6): 2639–73.

Ketcham, Jonathan D., Claudio Lucarelli, and Christopher A. Powers. 2014. "Paying Attention or Paying Too Much in Medicare Part D." *American Economic Review* 105 (1): 204–33.

Kling, Jeffrey R., Sendhil Mullainathan, Eldar Shafir, Lee C. Vermeulen, and Marian V. Wrobel. 2012. "Comparison Friction: Experimental Evidence from Medicare Drug Plans." *Quarterly Journal of Economics* 127 (1): 199–235.

McClellan, Mark B. 2006. Implementation of the Medicare Prescription Drug Benefit. Testimony in the Hearing before the Subcommittee on Health of the Committee on Ways and Means, US House of Representatives, One Hundred Ninth Congress, Second Session.

Neipp, Joachim, and Richard Zeckhauser. 1985. "Persistence in the Choice of Health Plans." *Advances in Health Economics and Health Services Research* 6: 47–72.

US Government Accountability Office. 2014. "Medicare Part D: CMS Has Implemented Processes to Oversee Plan Finder Pricing Accuracy and Improve Website Usability." GAO 14-143.

This article has been cited by:

1. Adam Atherly, Roger D. Feldman, Bryan Dowd, Eline van den Broek-Altenburg. 2020. Switching Costs in Medicare Advantage. *Forum for Health Economics and Policy*, ahead of print. [Crossref]

2. Amanda Starc, Robert J Town. 2019. Externalities and Benefit Design in Health Insurance. *The Review of Economic Studies* **101**. . [Crossref]

3. Christian Bünnings, Hendrik Schmitz, Harald Tauchmann, Nicolas R. Ziebarth. 2019. The Role of Prices Relative to Supplemental Benefits and Service Quality in Health Plan Choice. *Journal of Risk and Insurance* **86**:2, 415-449. [Crossref]

4. Andrew J. Barnes, Michael Karpman, Sharon K. Long, Yaniv Hanoch, Thomas Rice. 2019. More intelligent designs: Comparing the effectiveness of choice architectures in US health insurance marketplaces. *Organizational Behavior and Human Decision Processes* . [Crossref]

5. Michael Sinkinson, Amanda Starc. 2019. Ask Your Doctor? Direct-to-Consumer Advertising of Pharmaceuticals. *The Review of Economic Studies* **86**:2, 836-881. [Crossref]

6. Amitabh Chandra, Benjamin Handel, Joshua Schwartzstein. Behavioral economics and health-care markets 459-502. [Crossref]

7. Rebecca Myerson. 2019. Information Gaps and Health Insurance Enrollment: Evidence from Affordable Care Act Assister Programs. *SSRN Electronic Journal* . [Crossref]

8. Caitlin N. McKillop, Teresa M. Waters, Cameron M. Kaplan, Erin K. Kaplan, Michael P. Thompson, Ilana Graetz. 2018. Three years in – changing plan features in the U.S. health insurance marketplace. *BMC Health Services Research* **18**:1. . [Crossref]

9. Marcus Dillender. 2018. What happens when the insurer can say no? Assessing prior authorization as a tool to prevent high-risk prescriptions and to lower costs. *Journal of Public Economics* **165**, 170-200. [Crossref]

10. Liran Einav, Amy Finkelstein, Maria Polyakova. 2018. Private Provision of Social Insurance: Drug-Specific Price Elasticities and Cost Sharing in Medicare Part D. *American Economic Journal: Economic Policy* **10**:3, 122-153. [Abstract] [View PDF article] [PDF with links]

11. Brian E. McGarry, Nicole Maestas, David C. Grabowski. 2018. Simplifying The Medicare Plan Finder Tool Could Help Older Adults Choose Lower-Cost Part D Plans. *Health Affairs* **37**:8, 1290-1297. [Crossref]

12. Jason Abaluck, Jonathan Gruber, Ashley Swanson. 2018. Prescription drug use under Medicare Part D: A linear model of nonlinear budget sets. *Journal of Public Economics* **164**, 106-138. [Crossref]

13. Konan Hara, Jun Tomio, Thomas Svensson, Rika Ohkuma, Akiko Kishi Svensson, Tsutomu Yamazaki. 2018. Association measures of claims-based algorithms for common chronic conditions were assessed using regularly collected data in Japan. *Journal of Clinical Epidemiology* **99**, 84-95. [Crossref]

14. Cornel Kaufmann, Tobias Müller, Andreas Hefti, Stefan Boes. 2018. Does personalized information improve health plan choices when individuals are distracted?. *Journal of Economic Behavior & Organization* **149**, 197-214. [Crossref]

15. Benjamin Handel, Joshua Schwartzstein. 2018. Frictions or Mental Gaps: What's Behind the Information We (Don't) Use and When Do We Care?. *Journal of Economic Perspectives* **32**:1, 155-178. [Abstract] [View PDF article] [PDF with links]

16. Francesco Decarolis, Andrea Guglielmo. 2017. Insurers' response to selection risk: Evidence from Medicare enrollment reforms. *Journal of Health Economics* **56**, 383-396. [Crossref]

17. Adam Pilny, Ansgar Wübker, Nicolas R. Ziebarth. 2017. Introducing risk adjustment and free health plan choice in employer-based health insurance: Evidence from Germany. *Journal of Health Economics* **56**, 330-351. [Crossref]

18. James Marton, Aaron Yelowitz, Jeffery C. Talbert. 2017. Medicaid program choice, inertia and adverse selection. *Journal of Health Economics* **56**, 292-316. [Crossref]

19. Katherine A. Desmond, Thomas H. Rice, Arleen A. Leibowitz. 2017. Do Medicare Beneficiaries Living With HIV/AIDS Choose Prescription Drug Plans That Minimize Their Total Spending?. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* **54**, 004695801773403. [Crossref]

20. Jonathan Gruber. 2017. Delivering Public Health Insurance Through Private Plan Choice in the United States. *Journal of Economic Perspectives* **31**:4, 3-22. [Abstract] [View PDF article] [PDF with links]

21. Adam Pilny, Ansgar WWbker, Nicolas R. Ziebarth. 2017. Introducing Risk Adjustment and Free Health Plan Choice in Employer-Based Health Insurance: Evidence from Germany. *SSRN Electronic Journal* . [Crossref]

22. Jason Abaluck, Jonathan Gruber. 2016. Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program: Reply. *American Economic Review* **106**:12, 3962-3987. [Abstract] [View PDF article] [PDF with links]

23. Jonathan Ketcham, Nicolai Kuminoff, Christopher A. Powers. 2016. Rejoinder to Abaluck and Gruber. *SSRN Electronic Journal* . [Crossref]