NBER WORKING PAPER SERIES

ESTIMATING THE HETEROGENEOUS WELFARE EFFECTS OF CHOICE ARCHITECTURE: AN APPLICATION TO THE MEDICARE PRESCRIPTION DRUG INSURANCE MARKET

Jonathan D. Ketcham
Nicolai V. Kuminoff
Christopher A. Powers

Working Paper 22732
http://www.nber.org/papers/w22732

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2016

Ketcham and Kuminoff's research was supported by a grant from the National Institute for Health Care Management (NIHCM) Research and Educational Foundation. The findings do not necessarily represent the views of the NIHCM Research and Education Foundation or the National Bureau of Economic Research. We are grateful for insights and suggestions from Gautam Gowrisankaran, Kate Ho, Sebastien Houde, Mike Keane, Christos Makridis, Alvin Murphy, Sean Nicholson, Jaren Pope, Dan Silverman, Meghan Skira, V. Kerry Smith, and seminar audiences at the AEA/ASSA Annual Meeting, the Congressional Budget Office, Health and Human Services Office of the Assistant Secretary for Planning and Evaluation, the ASU Health Economics Conference, the Annual Health Economics Conference, the Quantitative Marketing and Economics Conference, the Health Econometrics Conference, Brigham Young University, Cornell University, Iowa State University, Michigan State University, Northern Arizona University, Stanford University, University of Arizona, UC Santa Barbara, University of Calgary, University of Chicago, University of Maryland, University of Miami, University of Southern California, Vanderbilt University, and Yale University.

Estimating the Heterogeneous Welfare Effects of Choice Architecture: An Application to the Medicare Prescription Drug Insurance Market
Jonathan D. Ketcham, Nicolai V. Kuminoff, and Christopher A. Powers
NBER Working Paper No. 22732
October 2016
JEL No. D02,D61,D81,I11

ABSTRACT

We develop a structural model for bounding welfare effects of policies that alter the design of differentiated product markets when some consumers may be misinformed about product characteristics and inertia in consumer behavior reflects a mixture of latent preferences, information costs, switching costs and psychological biases. We use the model to analyze three proposals to redesign markets for Medicare prescription drug insurance: (1) reducing the number of plans, (2) providing personalized information, and (3) defaulting consumers to cheap plans. First we combine administrative and survey data to determine which consumers make informed enrollment decisions. Then we analyze the welfare effects of each proposal, using revealed preferences of informed consumers to proxy for concealed preferences of misinformed consumers. Results suggest that each policy produces large gains and losses for some consumers, but the menu reduction would unambiguously harm most consumers whereas personalized information would unambiguously benefit most consumers.

Jonathan D. Ketcham
Earl G. and Gladys C. Davis Distinguished
Research Professor in Business
Department of Marketing, Box 4106
W.P. Carey School of Business
Arizona State University
300 E. Lemon Street
Tempe, AZ 85287-4106
ketcham@asu.edu

Nicolai V. Kuminoff
Department of Economics
Arizona State University
P.O. Box 879801
Tempe, AZ 85287
and NBER
kuminoff@asu.edu

Christopher A. Powers
U.S. Department of Health and Human Services
Centers for Medicare and Medicaid Services
7500 Security Boulevard
Mailstop B2-29-04
Baltimore, MD 21244
Christopher.Powers@cms.hhs.gov

One of the frontiers in empirical microeconomics is to assess the equity and efficiency of polices that alter a market's design to nudge consumers toward making certain decisions. Thaler and Sunstein (2008) dubbed this approach to policy "choice architecture". Examples of choice architecture include restricting the number of differentiated products in a market, providing consumers with personalized information about their options, and making default choices for consumers but letting them opt out. Understanding how such polices affect consumer welfare is increasingly important for program evaluation. The United Kingdom, the United States, the World Bank, and other government organizations have begun using choice architecture to nudge program beneficiaries.[1]

A stated goal of choice architecture is to benefit consumers who do not make fully informed decisions. Such paternalistic policies may also harm some consumers by eliminating their preferred products, by making it harder to buy those products, and by causing prices to increase (Camerer et al. 2003). Yet little work has attempted to predict the distribution of gains and losses of prospective choice architecture policies. To do so requires addressing at least two fundamental challenges with revealed preference analysis. First, an analyst must identify which consumer decisions are misinformed and hence potentially misleading about the consumer's preferences. Second, the preferences of both informed and misinformed consumers must be inferred. In this paper we develop a revealed preference framework to address these challenges and use it to predict the welfare effects of three recent proposals to redesign Medicare markets for prescription drug insurance.

Prescription drug insurance is an ideal setting for studying choice architecture. In 2006, Medicare Part D created government-designed, taxpayer-subsidized geographic markets for standalone prescription drug insurance plans. By 2014, these markets annually enrolled 24 million seniors with federal outlays of $65 billion (US Department of Health and Human Services 2015). When obtaining coverage, the typical enrollee chooses among 50 plans that differ in cost, risk protection, and quality. A new enrollee's choice becomes her future default; she will be passively reassigned to that same plan the follow-

[1] For example, in 2015 President Obama issued an executive order directing the newly created US Social and Behavioral Sciences Team to help federal agencies "identify programs that offer choices and carefully consider how the presentation and structure of those choices, including the order, number, and arrangement of options, can most effectively promote public welfare, as appropriate, giving particular consideration to the selection and setting of default options." (U.S. Executive Order #13707, Section 1.b.iii).

ing year unless she actively switches to a different one during the open enrollment window. Due to concerns about market complexity and consumer inertia, researchers and federal agencies have proposed several reforms (McFadden 2006, Thaler and Sunstein 2008, Federal Register 2014). These include reducing the number of plans, providing consumers with personalized information about their options, and auto-assigning people to default plans that are expected to minimize cost. We assess the welfare effects of these proposals using a novel combination of administrative records and survey data on a national panel of enrollees from 2006-2010. For the first time in academic research, we link the Medicare Current Beneficiary Survey (MCBS) to administrative records of the respondents' annual enrollment decisions, drug claims, and chronic medical conditions. Linking these data sets allows us to identify enrollees who are misinformed and analyze their decisions. The longitudinal MCBS tracks enrollees' effort to learn about the market, tests their knowledge of how the market works, observes whether they self-enrolled in plans or had help from advisors, and reports a rich set of demographics that do not exist in administrative data.

We model annual plan choices as a static repeated-choice process in which people may incur costs to learn about their options or to switch plans.[2] We first identify a subset of choices that will not necessarily reveal the person's preferences for plan attributes because the person appears to be misinformed. We characterize a choice as misinformed if the MCBS knowledge test reveals that the decision maker misunderstood a critical feature of the market, or if her choice can only be rationalized under full information by preference orderings that violate weak risk aversion or basic axioms of consumer theory. Based on these criteria, we find that 44% of 2006-2010 plan choices appear misinformed. The probability of being in this group increases as enrollees age, as they develop Alzheimer's disease and other cognitive impairments, and as their drug expenditures increase. The probability decreases with education and with their effort to learn about the market.

[2] A static model seems appropriate here because it is difficult for consumers to forecast their own future prescription drug needs, let alone the drug needs and enrollment decisions of other consumers together with the implications for plan prices and offerings. Our static approach is similar to other health insurance applications such as Handel (2013) and Handel and Kolstad (2015).

We then estimate and validate separate multinomial logit models for informed and misinformed choices. We find that informed enrollees are sensitive to price and risk averse at levels consistent with prior evidence (Cohen and Einav 2007, Handel 2013, Handel and Kolstad 2015). In contrast, the decisions made by misinformed enrollees seemingly imply that they are risk-loving, less price sensitive, and more averse to switching plans. We infer the preferences of these misinformed enrollees from the behavior of observationally identical enrollees in the informed group. The underlying assumption is that being informed is uncorrelated with preferences after conditioning on measures for health, prescription drug use and demographics. Using this assumption, we derive welfare measures which recognize that misinformed consumers may be made better off or worse off by policies that alter choice architecture. Our framework nests as special cases the welfare measures derived by Small and Rosen (1981) for the case of full information and by Leggett (2002) for the case of misinformation.

We use our estimates to simulate three prospective policies under a range of assumptions about consumer foresight, about the causes of inertia, and about how the policies will affect consumers' decisions.[3] Specifically, we report the share of consumers who benefit from each policy, consumer surplus, and other outcomes as bounds on ranges that we obtain by repeating our analyses under extreme assumptions about the efficacy of choice architecture. In our "most effective" scenario we assume that each policy causes misinformed consumers to behave like their analogs in the informed group. This scenario also assumes that inertia is caused entirely by misinformation. At the opposite extreme, our "least effective" scenario assumes the policies would not change consumer behavior and that inertia by informed consumers reflects their hassle cost of switching insurance plans and/or their utility from latent welfare-relevant features of their preferred plans.

The first policy we simulate is the government's proposal to limit each insurer to sell no more than two plans per market (Federal Register 2014). Second, we calibrate our model to replicate treatment effects from a field experiment by Kling et al. (2012) in

[3] We employ a partial-equilibrium approach and hold constant insurers' behavior and plan design. A comprehensive general equilibrium approach would require modeling of insurer interactions with consumers, other insurers, pharmaceutical manufacturers, and government regulators. We consider how each policy would affect insurers' net revenues holding premiums constant, as insurers' responses are predicated on such changes.

which enrollees were told which plan would be cheapest for them and how much money they could expect to save by switching. In the third experiment, we simulate the government's proposal to automatically reassign people to their lowest cost plans (Health and Human Services 2014). Our framework formalizes ways in which each policy has potential to create winners and losers.[4]

We find that reducing the number of plans makes at least two thirds of consumers worse off because people are heterogeneous and no plans are universally poor matches for consumers. This policy also embeds strong incentive for regulatory capture as insurers can increase their rents by influencing which plans are retained. In contrast, we find that personalized information benefits 48 to 92 percent of consumers, with average welfare gains of 2 to 11 percent of consumers' out-of-pocket spending. Similarly, defaults benefit over 80 percent of consumers if they can costlessly opt out. However, average opt out costs of $65 to $198 entirely eliminate these gains. All of these findings persist under a range of more inclusive or more exclusive rules for identifying misinformed choices.

This article advances on prior work that adapts Bernheim and Rangel's (2009) welfare framework to evaluate policies that target consumer inertia and misinformation. Prior studies have sought to recover preferences in such environments by leveraging experiments and surveys to distinguish between active and passive choices made by consumers who are assumed to differ in their knowledge of market institutions. Most identify preferences by assuming that information treatments make consumers fully informed or that consumers making active decisions are fully informed (Handel 2013, Allcott and Kessler 2015, Allcott and Taubinsky 2015, Taubinsky and Rees-Jones 2015, Ho, Hogan and Scott Morton 2015, Polyakova 2015). We relax both assumptions. Like Ambuehl, Bernheim, and Lusardi (2014), Handel and Kolstad (2015), Handel, Kolstad and Spinnewijn (2015) and Wiswall and Zafar (2015) we assess whether active decision makers are informed by testing their knowledge directly. We sharpen our test by leveraging novel features of our data to identify who actually makes enrollment decisions (beneficiaries or their advisors)

[4] For example, the menu restrictions may benefit misinformed consumers by reducing their ability to choose low utility plans. The information treatment and default assignment policies could create losers due to asymmetric information because the government would only use prior drug claims, and by creating incentives for consumers to choose plans that are cheaper but potentially lower utility due to lower quality or risk protection

and whether those decisions violate axioms of consumer theory.[5] Like Handel (2013) and Bernheim, Fradkin, and Popov (2015) we estimate bounds on welfare that recognize consumer inertia may arise from a mixture of latent preferences, information costs, switching costs, and psychological biases. We extend this partial identification logic to consider alternative hypotheses for how consumers will respond to choice architecture policies and the implications of those responses for consumer welfare and firm revenue. From a policy perspective, we believe our study is the first to use Bernheim and Rangel's framework to evaluate federal proposals to simplify a high-stakes differentiated product market that is both subsidized and regulated by the federal government.

More broadly, this article adds to research that aims to disentangle heterogeneity in information from heterogeneity in preferences, and to explore the welfare implications of counterfactual polices, especially when consumers make financially important decisions in complex or novel environments (Harris and Keane 1999, Keane and Wasi 2013, Choi et al. 2014, Keane and Thorp 2016). Specifically, we test decision makers' knowledge both directly and indirectly, we account for the role of advisors, and we recognize that consumer responses to survey-based knowledge questions may or may not be consistent with the way they make financial decisions. Our findings also advance knowledge on decision making among seniors, who are of particular interest because they control a large share of wealth and are relatively vulnerable to declines in cognitive function (Fang, Silverman and Keane 2008, Agarwal et al. 2009, Keane and Thorp 2016).

I. Medicare Part D

US citizens typically become eligible for Medicare benefits when they turn 65. In 2006, Medicare Part D extended these benefits to include prescription drug insurance. A novel and controversial feature of Part D is that it created quasi-private markets for delivering insurance.[6] Part D created 34 state or multistate markets within which the average enrollee chose among 50 standalone prescription drug insurance plans (PDPs) sold by 20

[5] In cases where advisors made the enrollment decisions, the MCBS tests the advisors' knowledge.
[6] Prior to the ACA, Part D was the largest expansion of public insurance programs since the start of Medicare.

private insurers.[7] The default for new beneficiaries is to be uninsured.[8] After an enrollee chooses a plan she is automatically reassigned to the same plan the following year unless she switches to a different one during open enrollment. Enrollees pay monthly premiums as well as out of pocket (OOP) costs for the drugs they purchase and taxpayers subsidize the total costs of non-poor enrollees by an average of 75.5%.

PDPs differ in terms of premiums, OOP costs of specific drugs, and quality measures such as customer service, access to pharmacy networks, the ability to obtain drugs by mail order, and the prevalence and stringency of prior authorization requirements.[9] The novelty of the market together with the complexity of the product led many analysts to speculate that consumers would struggle to navigate the market. Liebman and Zeckhauser (2008) summarize this view when they write: "Health insurance is too complicated a product for most consumers to purchase intelligently and it is unlikely that most individuals will make sensible decisions when confronted with these choices." Some analysts flagged Part D as a candidate for libertarian paternalism (McFadden 2006, Thaler and Sunstein 2008). Moreover, the government has expressed a desire to simplify health insurance markets and nudge enrollees toward cheaper plans. In 2014, CMS proposed limiting insurers to selling no more than two plans per region, which would reduce the average consumer's choice set by about 20% (Federal Register 2014). The US Department of Health and Human Services also announced that it is considering redesigning federal health insurance exchanges to automatically reassign people to low-cost plans unless they opt out (Health and Human Services 2014). The welfare effects of these types of policies will depend on consumers' preferences for PDP attributes, the cost of switching plans, and how the policies affect consumers' decision processes.

Several prior studies have investigated the role of information and consumer behavior in Medicare Part D. Over the first five years of the program, the average enrollee could have reduced annual expenditures (premium + out of pocket) by 25% (or $341) by switching to their cheapest available plan (Ketcham, Lucarelli and Powers 2015). Yet, the

[7] Subject to CMS approval, insurers can sell multiple PDPs in each market and make annual changes to existing plans.
[8] Enrollees who qualify for low-income subsidies are autoenrolled to certain plans, but we exclude them from our analysis.
[9] Many insurers require consumers to have prior authorization from a doctor in order to obtain certain drugs, but the stringency of these requirements differs from insurer to insurer.

implications for consumer welfare remain ambiguous. When enrollees are surveyed about their experiences in Part D most report being satisfied with the plans they chose (Heiss, McFadden and Winter 2010, Kling et al. 2012). Furthermore, Ketcham, Kuminoff and Powers (2015) demonstrate that most of the people who could have saved money by switching chose plans that were either superior in some measure of quality or provided greater protection from negative health shocks. These consumers could be making informed decisions to pay for quality and risk protection. On the other hand, when Kling et al. (2012) asked 406 Wisconsin enrollees how much they thought they could save by switching plans, most respondents underestimated the true figure. Kling et al. also found that sending enrollees a letter with personalized information about their potential savings increased the rate at which enrollees switched plans by 11.5 percentage points. Overall, the existing evidence suggests that some consumers are misinformed, but others may be choosing to pay more for plans with higher quality and/or greater risk protection.

II. Linking Administrative Records to Enrollee Surveys

For the first time in academic research, we have linked the Medicare Current Beneficiary Survey (MCBS) to the respondents' administrative records at the US Centers for Medicare and Medicaid Services (CMS). The MCBS is a national rotating panel questionnaire that began in 1991 and is administered to approximately 16,000 people annually.[10] It collects information about Medicare beneficiaries and their use of health care services. Each participant is interviewed up to three times per year for four consecutive years, regardless of whether they stay at the same address or move into and out of long term care facilities. Importantly for our purposes, participants are tested on their knowledge of the PDP market. The MCBS also asks participants if and how they

[10] A potential limitation of working with the MCBS sample is that it is not designed to be nationally representative without weighting, and selecting the appropriate weights is complicated by panel rotation and by our exclusive focus on respondents who participated in the standalone PDP market. Respondents who do not purchase a standalone PDP can instead obtain prescription drug insurance through an employer sponsored plan or a Medicare Advantage plan. Further, the MCBS does not sample individuals from 3 PDP regions: 1(Maine and New Hampshire), 20 (Mississippi), and 31 (Idaho and Utah). To assess whether using unweighted MCBS data might compromise the external validity of our results, we compared the unweighted demographics of the average enrollee in our linked sample with a random 20% sample of all Part D enrollees from CMS's administrative files. Table A2 shows that the average enrollee in our linked sample is 1 to 2 years older. Otherwise, the two samples are virtually identical in terms of race, gender, rates of dementia and depression, number of PDP brands and plans available, expenditures on plan premiums and OOP costs, and the maximum amount of money that the average enrollee could have been saved by enrolling in their cheapest available plan. Given the strong similarity between the two samples, we expect that our findings from the linked MCBS-administrative sample can be generalized to the broader population of non-poor Part D enrollees.

searched for information about Medicare services and it provides rich demographic data. Also of particular value for our study, the MCBS indicates whether a proxy responded to the survey, and whether the beneficiary makes health insurance decisions on her own, with help from someone else, or whether the proxy makes decisions for her.

For each MCBS respondent who purchased a standalone PDP between 2006 and 2010 we obtained administrative records on their prescription drug claims, the set of PDPs available to them, and their annual enrollment decisions. Then we calculated what each enrollee would have spent had they purchased the same bundle of drugs under each alternative PDP in their choice set. This was done by combining their actual claims with the cost calculator developed in Ketcham, Lucarelli and Powers (2015).[11] Next we used administrative data from CMS's Chronic Condition Data Warehouse to determine if and when each individual had depression or dementia, which are associated with diminished cognitive performance (Agarwal et al. 2009). Like prior studies of PDP choice we limit our analysis to enrollees who did not receive a low-income subsidy.[12]

Our linked sample includes 3,547 individuals who made 10,867 annual enrollment decisions between 2006 and 2010.[13] Table A1 reports annual means of the key variables. The typical enrollee is a retired high school graduate with living children. Approximately 22% are college graduates, 55% are married, and 55% have annual pre-tax household incomes over $25,000. Only 35% report that they ever personally use the internet to get information of any kind. However, among those who do use the internet most have used it to search for information on Medicare programs (27%). Another 17% report having called 1-800-Medicare for information. The average beneficiary's total expenditures on premiums and out of pocket costs increased from $1,203 in 2007 to $1,400 in 2010.[14]

[11] There is a correlation of .94-.98 each year between the out of pocket costs predicted for the actual plan and the realized cost observed in the administrative data. Differences between the calculator's predictions and realized costs are due to changes in plan design or drug pricing that occur after open enrollment and are not observable to consumers at the time they make enrollment decisions.

[12] We exclude those receiving low-income subsidies because they are autoenrolled into plans, they receive larger premium subsidies, and their copayments are much more uniform across plans. Hence, they are less relevant for our evaluation of prospective policies designed to alter choice architecture. Despite excluding them, our sample has similar income levels to the national average of people age 65 and above. In our sample 54% of households have annual income over $25,000 (weighted 2006-2010 dollars), compared with 63% (constant 2010 dollars) based on all householders 65 and older in the 2010 Census American Community Survey.

[13] This excludes observations on beneficiaries who reenrolled in plans they had originally chosen prior to joining the MCBS. We drop these observations because we cannot observe the beneficiaries' knowledge at the time they first selected their current plans.

[14] The figure for 2006 is $1,013. It is smaller because during the inaugural year of the program open enrollment extended through May. Less than half the enrollees in our sample were enrolled for all of 2006. If we limit the sample to full-year enrollees, the 2006 mean annual consumer expenditure is $1,366.

This is a significant share of income given that 45% of beneficiaries have household incomes below $25,000. The data also reveal that by the end of our study period significant fractions of enrollees had been diagnosed with dementia (12%) and depression (11%).

TABLE 1—CHARACTERISTICS OF PEOPLE WHO MAKE THEIR OWN DECISIONS OR GET HELP

	Who makes health insurance decisions?		
	Beneficiary	Beneficiary gets help	Proxy
number of enrollment decisions	6,790	2,906	1,171
high school graduate (%)	83	75	61
college graduate (%)	25	19	14
income>$25k (%)	57	53	48
uses the internet (%)	39	33	18
mean age	77	78	80
dementia including Alzheimer's (%)	5	11	31
depression (%)	9	11	14
mean number of drug claims	32	36	40
mean premium ($)	416	411	426
mean out-of-pocket costs ($)	885	1,030	1,285
mean potential savings ($)	325	325	357

Note: The table reports means for key variables for the sample of Medicare Part D enrollees found in both the MCBS and cost calculator samples from 2006-2010. See the text for details.

Given the relatively large amount of money at stake, the age range of the eligible population and the prevalence of cognitive illnesses it is unsurprising to find that 38% of enrollees did not make health insurance decisions on their own: 27% had help and 11% relied on a proxy to make the decision for them. Table 1 shows that beneficiaries who get help are likely to be older, sicker, lower income, less educated, and less internet savvy than beneficiaries who made decisions on their own. Those getting help are also more likely to have been diagnosed with depression or dementia. All of these differences are amplified when we compare beneficiaries who make their own health insurance decisions to those who rely on proxies to make decisions for them.

III. Identifying Enrollment Decisions Suspected to be Misinformed

Only 8% of the enrollment decisions in our data minimize ex post expenditures. In

2006 the average enrollee could have saved \$460 by choosing their cheapest available plan.[15] This is equivalent to reducing total expenditures by 45%. Potential savings declined to \$349 in 2007 (or 29% of expenditures) and remained similar thereafter. Why are people leaving money on the table? We hypothesize that the answers differ from person to person. Some may be making informed decisions to pay more for plans that provide better risk protection and higher quality. Others may misunderstand how the market works or underestimate their potential savings. We must distinguish between these groups to evaluate the welfare effects of prospective choice architecture policies.

For the group we identify as informed, we apply standard revealed preference logic to infer their preferences for cost reduction, risk protection, and quality. But revealed preference logic cannot be applied when consumers have latent beliefs about products that contradict the information we observe. With this in mind, we adapt two features of Bernheim and Rangel's (2009) proposed approach to revealed preference analysis in the presence of latent heterogeneity in beliefs.[16] First, we use theory and data to identify enrollment decisions that we suspect may fail to reveal preferences. We label these choices as *suspect*, using Bernheim and Rangel's terminology. Second, for the consumers making suspect choices, we calibrate their preference relations using proxy measures derived from the behavior of observationally identical consumers who we observe making *non-suspect* choices. Thus, we implement Bernheim and Rangel's proposal to respect consumer sovereignty and apply standard revealed preference methods unless theory and data suggest the standard approach may fail to reveal consumers' preferences.

A. Defining Suspect Choices

Like prior Part D studies, we assume that consumer *i*'s utility from drug plan *j* in year *t* depends on the mean and variance of her potential expenditures in that plan under all possible health states. Expenditures equal the plan premium, p_{jt}, plus out of pocket costs,

[15] This figure sums over premiums and out of pocket costs. See Table A1 for details. This average falls below the \$520 figure reported by Ketcham, Lucarelli and Powers (2015) based on CMS's 20% sample of 2006 full year enrollees because our average also includes people who only enrolled for part of the year. The primary reason for part-year enrollment in 2006 was the fact that the initial open enrollment period was extended through May (Heiss, McFadden, and Winter 2010).

[16] Latent heterogeneity in beliefs is one case of what Bernheim and Rangel refer to as "ancillary conditions" on decision making.

$oop_{jt}(x_{it})$, of an exogenously given vector of drug quantities, x_{it}. Utility also depends on a vector of measures of plan quality, q_{ijt}, that reflect the time and effort required for an individual to obtain her eligible benefits under the plan.

Our first indicator of suspect choices is derived by applying Ketcham, Kuminoff, and Powers' (2016) test for whether consumers are actively choosing plans that cannot be rationalized as maximizing a well behaved utility function under full information.[17] To simplify notation we denote total costs as $c_{ijt} = p_{jt} + oop_{ijt}$. We assume that consumers are weakly risk averse and have preference orderings that are complete, transitive, and strongly monotonic over expected cost savings, risk protection, and quality. Under this assumption, a fully informed consumer will never actively enroll in a plan, *j*, that is dominated by another, *k*, in the sense that the following four conditions hold simultaneously.

$(1.a)\ E(c_{ikt}) \leq E\left(c_{ijt}\right).$
$(1.b)\ var(c_{ikt}) \leq var\left(c_{ijt}\right).$
$(1.c)\ q_{ijt} \leq q_{ikt}.$
$(1.d)\ At\ least\ one\ of\ the\ inequalities\ is\ strict.$

An informed consumer will never choose a plan that has higher cost, higher variance, and lower quality than some feasible alternative. We refer to choices that satisfy (1.a)-(1.d) as being *dominated*. In theory, a consumer may choose a dominated plan if she is risk loving, if she dislikes quality, if she has a negative marginal utility of income, or, more likely, if she is misinformed about her options. Hence, if we observe a consumer actively choosing a dominated plan then we label her choice as "suspect". We suspect that the consumer is misinformed and, therefore, that her enrollment decision may not reveal her preferences.[18]

To test whether enrollees chose dominated plans we define cost, variance, and quality using methods from the literature on PDP choice (Abaluck and Gruber 2011, Ketcham,

[17] Similar to Chetty et al. (2015) we define an enrollment choice as *active* if either of the following statements is true: (1) the person is new to the market and must select a plan to become insured or (2) the person switched to a new plan during open enrollment. If neither statement is true, then the enrollee took no action during open enrollment and was automatically reenrolled in the plan she chose last year—her default—in which case we define her choice as *passive*. After the inaugural enrollment cycle in 2006 between 20% and 23% of enrollees made active choices each year.

[18] Consumers who violate at least one condition are choosing plans on what Lancaster (1966) called the "efficiency frontier" in attribute space. Every plan on the frontier can be rationalized as maximizing some utility function that satisfies the preference axioms and weak risk aversion under full information. For example, an informed risk averse consumer may optimally choose a more expensive and lower quality plan that better insures her against negative health shocks.

Kuminoff, and Powers 2016). First we assume that informed consumers have unbiased expectations of their drug needs for the upcoming year: $E(c_{ijt}) = c_{ijt}$.[19] Next, we use a cohort approach to calculate variance. We calculate $var(c_{ijt})$ from the distribution of expenditures under plan *j* for the drugs used in year *t* by people in consumer *i*'s cohort in terms of year *t-1* drug claims. Specifically, we use CMS's random 20% sample of all PDP enrollees to assign each individual in the MCBS sample to 1 of 1000 cells defined by the deciles to which she belonged in the national distributions of the prior year's total drug spending, days' supply of branded drugs, and days' supply of generic drugs.[20] Then we calculate $var(c_{ijt})$ for the distribution of drugs used by everyone in consumer *i*'s cell. Finally, we allow utility to depend on indicators for insurance companies. These indicators reflect all aspects of PDP quality that vary across insurers, such as customer service, pharmacy networks, mail order options, and prior authorization requirements.[21]

Because we allow utility to depend on insurer dummies, a chosen plan will be dominated if and only if the enrollee could have chosen a different plan offered by the same insurer that would have lowered the mean and variance of her drug expenditures, or lowered one holding the other constant. The first row of Table 2 shows that 19% of beneficiaries actively enrolled in dominated plans in the first year of the program, when everyone had to actively enroll. The share ranged from 4% to 7% in subsequent years. The decline after 2006 is mostly due to a decline in active decision making as most returning enrollees reenrolled in plans they had chosen previously. That said, the probability of choosing a dominated plan conditional on making an active choice also declined by three percentage points between 2006 and 2010.

To hedge against potential Type II error in using active choices of dominated plans to identify misinformation, we use an MCBS knowledge question to develop a second sus-

[19] Our econometric estimates and policy conclusions are robust to assuming that consumers are myopic: $E(c_{ijt}) = c_{ijt-1}$. This is unsurprising since individual prescription drug use is strongly persistent over time.

[20] In cases where CMS did not have the person's drug claims from the prior year, such as 2006, we predicted their deciles based on current and future drug claims and past, current and future health.

[21] For example, stringent prior authorization requirements for certain drugs may be unattractive to consumers who believe they have a high likelihood of purchasing those drugs and irrelevant to consumers who do not. Likewise, consumers differ in their proximity to in-network pharmacies. These factors vary across insurance brands and consumers but not across plans within a brand.

pect choice indicator. Each year, respondents were asked to state whether the following sentence is true or false.

$$(1.e)\ Your\ OOP\ costs\ are\ the\ same\ in\ all\ Medicare\ prescription\ drug\ plans.$$

For people with no drug claims, the statement is true. For people with any claims the statement is false due to variation in formularies, deductibles, and coinsurance. Understanding that drug costs vary across plans is the central to understanding how the market works.[22] Moreover, this variation is financially important: the average beneficiary's OOP costs for her purchased drugs vary by over $1,100 across her available plans.

TABLE 2—POTENTIAL INDICATORS OF SUSPECT CHOICES

	Percent of enrollees					
	2006	2007	2008	2009	2010	2007-2010
Actively enrolling in a plan:						
that is dominated	19	6	6	4	5	5
while not answering knowledge question correctly	44	6	8	6	9	7
Passively reenrolling in a plan that was:						
dominated when actively chosen		12	12	12	10	11
actively chosen while not answering knowledge question correctly		31	26	23	19	24
Suspect choices (union of the first four rows)	54	48	45	40	38	42

Note: The table reports the share of choices triggering each indicator, by year. The MCBS knowledge question asks whether the enrollee's out of pocket costs are the same under every available drug plan. The correct answer is coded as yes for enrollees who filed drug claims in both the prior and current years if their out of pocket costs did in fact vary across plans in both years. The last row reports the share of enrollees satisfying the criteria in either of the first two rows. See the text for additional details.

We use each person's drug claims to determine their correct answer to the MCBS question. Because respondents may be unsure about which enrollment year the question is referring to, we code a person's answer for year *t* as correct if it is correct for either year *t* or year *t-1*. Table 2 shows that 44% of respondents failed to give the correct answer in 2006 and 6% to 9% answered incorrectly when making active choices in subsequent years.[23] On average, respondents who answered incorrectly could have saved 16% more by switching to a different plan than those who answered correctly.[24]

[22] The MCBS asks five other questions that test knowledge of Part D, but they are less relevant for forecasting individual drug expenditures. Howell, Wolff and Herring (2012) provide further analysis of the MCBS knowledge questions.

[23] Aggregating over active and passive choices, the total share of respondents answering incorrectly was 29% in 2007, 31% in 2008,

Finally, when beneficiaries are passively reenrolled in their default plans we defer to their preceding active choices of those plans when coding their passive reenrollment decisions as suspect or non-suspect. Table 2 shows that from 2007 to 2010, 11% of beneficiaries were passively reenrolled in plans that were dominated when they were actively chosen and 24% were reenrolled in plans that were actively chosen during enrollment cycles in which they answered the knowledge question incorrectly.

In summary, we define a choice as suspect if the decision maker (i) actively enrolled in a dominated plan; (ii) actively enrolled in a plan while answering the knowledge question incorrectly; or (iii) passively reenrolled in a plan that satisfied (i) and/or (ii) at the time it was first chosen. The last row of Table 2 shows that 54% of all enrollment decisions satisfied at least one of these criteria in 2006 and 42% between 2007 and 2010. In the supplemental appendix (Table A10) we demonstrate that our policy conclusions are robust to several alternative ways of defining suspect choices. This includes focusing exclusively on dominated plan choices; focusing exclusively on the knowledge test; using a more inclusive definition that adds enrollees who could have reduced their expenditures by more than 50%; including choices for 2006; excluding mid-year enrollment decisions; excluding beneficiaries who get help choosing plans; and assuming that enrollees are myopic in the sense that they expect their drug use in the upcoming year to be identical to the prior year.

B. Who is More Likely to Make Suspect Choices?

To develop intuition for potential mechanisms driving suspect choices, we estimate linear probability models in which the dependent variable, S_{irt}, is an indicator for whether person *i* in CMS region *r* made a suspect choice in the year *t* enrollment cycle,

$$(2)\ S_{irt} = \kappa + \lambda d_{irt} + \phi_r + \rho_t + e_{it}.$$

---

and 28% in 2009 and 2010. The 15 percentage point reduction between 2006 and 2007 is consistent with prior evidence on learning in PDP markets (Ketcham, Lucarelli, and Powers 2015, Ketcham et al. 2012).

[24] Table A3 shows that when we focus on active enrollment decisions, failing to answer the knowledge question correctly is associated with a 1.3 percentage point increase in the probability of choosing a dominated plan and a $68 increase in the amount of money that could be saved by switching to the cheapest available plan, even when conditioning on education, income, employment status, presence of living children, internet use, effort to search for information about CMS programs online or by calling 1-800-Medicare, getting help making enrollment decisions, the number of available plans, gender, race, age, dementia, depression, number of drug claims, and dummies for year and CMS region. For 11% of our sample the person who responds to the survey and makes the enrollment decision is a proxy for the beneficiary, such as a spouse or child (Table A1).

On the right of the equality d_{irt} is a vector of demographics, some of which change over time, and ρ_t and ϕ_r are indicators for enrollment year and region.[25]

TABLE 3—ASSOCIATION BETWEEN SUSPECT CHOICES AND DEMOGRAPHICS

	Suspect choice 2006 - 2010		Suspect choice 2007 - 2010	
college graduate	-0.058	[0.021]***	-0.058	[0.021]***
income>$25k	-0.012	[0.018]	-0.012	[0.019]
currently working	0.011	[0.025]	0.009	[0.026]
married	0.012	[0.020]	0.011	[0.020]
has living children	-0.057	[0.033]*	-0.064	[0.034]*
uses the internet	-0.020	[0.021]	-0.015	[0.022]
searched for CMS info: internet	-0.090	[0.021]***	-0.083	[0.021]***
searched for CMS info: 1-800-Medicare	-0.058	[0.019]***	-0.066	[0.020]***
has help making insurance decisions	0.025	[0.017]	0.016	[0.018]
number of available plans (standardized)	-0.005	[0.014]	-0.003	[0.016]
female	0.024	[0.019]	0.028	[0.019]
nonwhite	0.118	[0.035]***	0.114	[0.036]***
age: 70-74	0.050	[0.021]**	0.047	[0.023]**
age: 75-79	0.066	[0.025]***	0.065	[0.027]**
age: 80-84	0.072	[0.027]***	0.071	[0.028]**
age: over 84	0.120	[0.029]***	0.118	[0.030]***
dementia including Alzheimer's	0.048	[0.026]*	0.040	[0.027]
depression	0.012	[0.022]	0.011	[0.023]
number of drug claims (standardized)	0.027	[0.008]***	0.033	[0.008]***
number of plan choices	10,867		9,119	
number of enrollees	3,547		3,444	
mean of the dependent variable	0.44		0.42	
R-squared	0.064		0.059	

Note: The table reports coefficients and standard errors from linear probability models of individual's plan choices. The dependent variable equals one if we suspect the choice was misinformed. See the text for a formal definition. All explanatory variables are binary expect the number of available plans and the number of drug claims, both of which are standardized. The omitted indicators define the baseline enrollee as a 65 to 69 year old white male who did not finish high school, has income below $25k, does not get help making insurance decisions, has not searched for CMS information using the internet or 1-800-Medicare, has the mean number of drug claims, and has not been diagnosed with dementia or depression. All regressions include indicators for enrollment year and region. Robust standard errors are clustered by enrollee. *,**, and *** indicate the p-value is less than 0.1, 0.05, and 0.01.

The first column of Table 3 reports results for enrollment decisions from 2006-2010.

[25] These indicators capture variation in the complexity of choice sets across space and time. For example, in the first year of the program the number of available plans per region ranged from 27 to 52. The number of plans also changed over time, increasing noticeably between 2006 and 2007. This variation allows us to test the choice overload hypothesis that consumers are less likely to make informed decisions as the number of options grows. Ketcham, Lucarelli and Powers (2015) test choice overload in Part D more extensively, capitalizing on individual-specific variation in the number of plans available by the person's relative cost of those plans.

The omitted indicators define the reference person as a 65 to 69 year old unmarried and retired white male with no high school diploma who has not searched for information on CMS programs and makes his own enrollment decisions. The coefficients imply that obtaining a college degree is associated with a 5.8 percentage point reduction in the probability of making a suspect choice. The probability is higher for nonwhites (+11.8) which might proxy for unobserved differences in wealth or education. The probability is lower for enrollees who searched for information about CMS programs using the internet (-9.0) or by called 1-800-Medicare (-5.8), but it is not any lower for beneficiaries who had help making enrollment decisions.[26]

Looking at the administrative variables, the probability of making a suspect choice is increasing in age, consistent with prior evidence on the decline in cognitive performance for individuals over 65 (Agarwal et al. 2009, Tymula et al. 2013). The predicted probability is approximately 7 percentage points higher for enrollees in their late 70's and 12 percentage points higher for enrollees in their late 80's. This is after controlling separately for diagnosed cognitive illnesses normally associated with aging, namely dementia (+4.8), and conditioning on the increased complexity of decisionmaking associated with greater drug needs via a measure of total drug claims (+2.7 for a one standard deviation increase in claims). Having living children, even conditional on receiving help choosing, is associated with a nearly 6 percent reduction in the probability of making a suspect choice. In comparison we find that income, gender, and marital status have small and statistically insignificant effects. We also obtain a precisely estimated zero on the number of available plans, providing evidence against the hypothesis that choice overload causes suspect choices (Ketcham, Lucarelli and Powers 2015).

The last column of Table 3 shows that the results are largely unchanged if we drop 2006. We exclude 2006 enrollment decisions from our main analysis because of the improvement in knowledge question responses in 2007. Because consumers appear to have learned during the inaugural year of the program, their choices in that first year may be

[26] The lower probability for those calling 1-800-Medicare is consistent with Kling et al.'s (2012) audit of the Medicare help line in which actors calling the number for information found that customer service representatives consistently identified low-cost plans based on the actors' fictional drug needs. The positive (but insignificant) coefficient for those getting help could be driven by principal agent problems, the helpers' opportunity costs of time, and/or added complexity in the decision process since those getting help tend to use more drugs and are more likely to be diagnosed with dementia and depression (Table 1).

less informative for analyzing prospective policies. That said, we show in the appendix that our main findings are invariant to whether we include or exclude 2006 choices.

IV. A Parametric Model of Decision Making with Heterogeneity in Beliefs

To evaluate the welfare effects of prospective polices we must select a parametric approximation to utility. The novelty of our approach to is to allow for heterogeneity in beliefs about plan attributes. We focus on identifying parameters that describe how plan attributes affect plan choice and then use our indicators for *suspect* and *non-suspect* choices to guide how we interpret those parameters and use them for policy evaluation.

A. Initial Enrollment Decision

When a beneficiary first enters the market in year 0 she must actively choose a plan to obtain insurance. She will choose the plan that maximizes her utility, conditional on her beliefs about plan attributes. We approximate this process with a model similar to the ones used by Abaluck and Gruber (2011), Kling et al. (2012) and Ketcham, Kuminoff, and Powers (2016),

$$(3)\ U_{ij0} = \alpha_{it}\acute{c}_{ij0} + \beta_{it}\acute{\sigma}_{ij0}^{2} + \gamma_{it}\acute{q}_{ij0} + \epsilon_{ij0}.$$

$\acute{c}_{ij0}$ denotes the amount that person *i* expects to spend under plan *j* in terms of the premium plus out of pocket costs for prescription drugs, $\acute{\sigma}_{ij0}^{2}$ is the variance of out of pocket costs, $\acute{q}_{ij0}$ is a vector of quality attributes, and ϵ_{ij0} is an idiosyncratic person-plan specific taste shock. The accents indicate that the variables reflect person *i*'s beliefs about plan attributes. Heterogeneity in beliefs is discussed below. Beneficiaries may also have heterogeneous marginal rates of substitution between expected cost, variance, and quality. We model this heterogeneity as a linear function of demographics, some of which may evolve over time: $\alpha_{it} = \alpha_0 + \alpha_1 d_{it}$, and similarly for β_{it} and γ_{it}. Finally, people may lose utility from the time and effort required to learn about a plan and enroll in it. We assume that this cost is constant across plans so that it cancels out of between-plan comparisons and can therefore be suppressed in (3).

B. Subsequent Enrollment Decisions

After an enrollee chooses a plan in year 0 she is automatically reassigned to that plan in year 1 unless she actively switches to a different plan during open enrollment.[27] As before, making an active decision may be costly. In contrast, no effort is required to reenroll in the default plan:

(4) $U_{ij1} = \alpha_{it}\acute{c}_{ij1} + \beta_{it}\acute{\sigma}^2_{ij1} + \gamma_{it}\acute{q}_{ij1} + \eta_{it}\Delta\acute{B}_{ij1} + \delta_{it}\Delta\acute{P}_{ij1} + \epsilon_{ij1}.$

Two terms capture the utility loss from actively switching plans: $\Delta\acute{P}_{ijt}$ is an indicator for whether plan j is a non-default plan sold by the same insurer as the default plan, and $\Delta\acute{B}_{ijt}$ is an indicator for whether plan *j* is a non-default plan sold by a different insurer. The disutility of switching plans is captured by the parameters $\eta_{it} = \eta_0 + \eta_1 d_{it}$ and $\delta_{it} = \delta_0 + \delta_1 d_{it}$, which summarize how inertia varies with demographics. We consider how to interpret inertia when we discuss welfare measurement in Section V. After a consumer chooses a plan in year *1*, the decision process is the same in years *2*,…,*T*.

C. Heterogeneity in Information

We model heterogeneity in information by allowing suspect and non-suspect choices to be driven by different beliefs. Non-suspect choices are assumed to be informed in the sense that decision makers' beliefs about plan attributes coincide with the measures we collected. Put differently, we respect consumer sovereignty and invoke the standard assumption of full information in the absence of evidence to the contrary. In contrast, we do not observe the beliefs about plan attributes that led to suspect choices. While the non-suspect (*n*) and suspect (*s*) groups may have different beliefs about plans, we assume that they share the same underlying preference parameters conditional on demographics.

(5) $U^n_{ijt} = \alpha_{it}c_{ijt} + \beta_{it}\sigma^2_{ijt} + \gamma_{it}q_{jt} + \eta_{it}\Delta B_{ijt} + \delta_{it}\Delta P_{ijt} + \epsilon_{ijt}.$

(6) $U^s_{ijt} = \alpha_{it}\acute{c}_{ijt} + \beta_{it}\acute{\sigma}^2_{ijt} + \gamma_{it}\acute{q}_{ijt} + \eta_{it}\Delta\acute{B}_{ijt} + \delta_{it}\Delta\acute{P}_{ijt} + \epsilon_{ijt}.$

[27] Plans are occasionally discontinued, which can force people to make an active choice. In such case, we can revert to equation (3) to model the new enrollment decision.

We dropped the accents in (5) to indicate that we are using our empirical measures of plan attributes for the non-suspect group.[28]

Because we do not observe the beliefs of people making suspect choices, we do not necessarily identify their preferences from their observed behavior. To see this notice that if we replace the subjective beliefs in (6) with empirical measures of plan attributes then, in general, we must also allow the values of the preference parameters and the error term to change in order to maintain their utility ranking of plans:

(7) $U_{ijt}^{s} = \acute{\alpha}_{it} c_{ijt} + \acute{\beta}_{it} \sigma_{ijt}^{2} + \acute{\gamma}_{it} q_{ijt} + \acute{\eta}_{it} \Delta B_{ijt} + \acute{\delta}_{it} \Delta P_{ijt} + \acute{\epsilon}_{ijt}$.

For example, if people make suspect choices because they have downward biased expectations about their drug needs at the time they choose a plan (i.e. $c_{ijt} > \acute{c}_{ijt}$) then we would expect $\alpha_{it} < \acute{\alpha}_{it}$. Likewise, if they have downward biased expectations about their potential savings from switching plans, then we would expect $\eta_{it} < \acute{\eta}_{it}$ and $\delta_{it} < \acute{\delta}_{it}$.

To facilitate estimation we assume that the person-plan specific taste shocks in (5) and (7) are *iid* draws from type I extreme value distributions. The variances may differ between the suspect and non-suspect groups because the idiosyncratic shocks in (7) will absorb any residual utility differences needed to maintain the preference ordering over plans when we move from (6) to (7). Therefore, when we normalize the model variances to $\pi^2/6$, the coefficients estimated for the suspect group will be scaled by the ratio of the group-specific variances (Train 2009). After making this normalization, we can rewrite the estimating equation for the suspect group (s) as

(8) $U_{ijt}^{s} = \alpha_{it}^{s} c_{ijt} + \beta_{it}^{s} \sigma_{ijt}^{2} + \gamma_{it}^{s} q_{ijt} + \eta_{it}^{s} \Delta B_{ijt} + \delta_{it}^{s} \Delta P_{ijt} + \epsilon_{ijt}$,

where $\alpha_{it}^{s} = \acute{\alpha}_{it} \sqrt{var(\epsilon_{ijt})/var(\acute{\epsilon}_{ijt})}$ and similarly for β_{it}^{s}, γ_{it}^{s}, η_{it}^{s}, and δ_{it}^{s}. Our econometric model identifies the parameters of (5) and (8).

[28] Their expected PDP costs are defined as $c_{ijt} = p_{jt} + E[oop_{ijt}]$, their type-specific variance is defined as $\sigma_{ijt}^2 = var(oop_{ijt})$, and q_{jt} is a vector containing indicators for insurance companies and an index of overall plan quality developed by CMS. All variables are calculated using the techniques developed in prior studies of PDP choice as described in III.A.

D. Identification

Equations (3)-(4) illustrate how the model parameters can be identified from data on suspect and non-suspect enrollment decisions. Our ability to observe each individual's plan choices when they first enter the market allows us to overcome the initial conditions problem. Consider the non-suspect group. Given the parametric form for utility and the distributional assumption on ϵ_{ijt}, we can use a multinomial logit model of initial plan choices (3) to identify the parameters that describe how marginal rates of substitution between cost, variance, and quality vary with demographics, $\alpha_0, \alpha_1, \beta_0, \beta_1, \gamma_0, \gamma_1$. Then we can use a model of their subsequent plan choices (4) to identify the inertia parameters, $\eta_0, \eta_1, \delta_0, \delta_1$, via the rates at which individuals actively switched out of their initial plans. In practice, we pool data from all plan choices and estimate the parameters simultaneously using (5). The same arguments can be made to identify the parameters of (8) for the suspect group. Prior studies have analyzed the properties of this model and underlying identification arguments in detail (Ketcham, Kuminoff, and Powers 2016, Polyakova 2015). The novelty of our identification strategy is to estimate separate parameters for suspect and non-suspect groups. The ability to differentiate their decision processes is critical to assessing welfare effects of prospective policies.

V. Welfare Effects of Choice Architecture Policies

When some decisions are misinformed, reforms that reduce information costs and/or simplify the choice process can, in principle, increase some consumers' welfare. Consider a policy implemented between periods 0 and 1 that changes the set of available plans from J to K. Consumer welfare may be affected through multiple channels. First, the policy may change the menu of options by adding choices, removing choices, and regulating their costs or quality. Second, the policy may change how consumers make decisions, e.g. by lowering the cost of switching plans or changing default assignment rules.[29]

A. Non-Suspect Group

[29] In general equilibrium, if the policy induces consumers and firms to adjust their behavior then those adjustments may feed back into the levels of endogenous attributes such as premiums.

The expected change in welfare for people in the non-suspect group (*n*) is derived by integrating over ϵ_{ijt} in the standard expression for consumer surplus to generate the log sum ratio from Small and Rosen (1981).

$$(9)\ \Delta E[CV_i^n] = \frac{1}{\alpha_{it}^n}\left\{ln\frac{\sum_{k\in K}[exp(V_{ik}^{n1})]}{\sum_{j\in J}\left[exp\left(V_{ij}^{n0}\right)\right]}\right\},$$

where V_{ij}^{n0} and V_{ik}^{n1} denote the observed part of utility in (5) evaluated for PDPs j and k before and after the policy. The temporal subscript is suppressed for brevity such that $V_{ij}^{n0} = V_{ijt}^{n0}(\theta^n, d_{it}) = U_{ijt}^{n0} - \epsilon_{ijt}$, where $\theta^n = [\alpha^n, \beta^n, \gamma^n, \eta^n, \delta^n]$ and each letter is a vector of parameters describing how preferences vary with demographics.

B. Suspect Group

Welfare calculation is more involved for the suspect group. The observed part of (8) determines how PDP attributes affect their enrollment decisions, but their ex post realized utility from those decisions is determined by (5). This follows from our assumption that, conditional on prescription drug use and demographics, the suspect and non-suspect groups share the same underlying preference parameters. Therefore, a single plan's contribution to expected utility is defined by integrating over the product of (5) and the probability of choosing that plan based on (8). Aggregating over the PDP menu prior to the policy yields the following general expression

$$(10)\ E[U_i^{s0}] = \sum_{j\in J}\int_{-\infty}^{\infty}\left(V_{ij}^{n0} + \epsilon_{ij}\right)F_j\left(V_{ij}^{s0} - V_{i1}^{s0} + \epsilon_{ij}, \dots, V_{ij}^{s0} - V_{iK}^{s0} + \epsilon_{ij}\right)d\epsilon_{ij},$$

where $F_j(\cdot)$ is the derivative of the joint CDF of the idiosyncratic taste shocks with respect to ϵ_{ij}. Subtracting this expression from the post-policy measure of expected utility, dividing by the marginal utility of income, and integrating over the idiosyncratic taste shocks yields an expression for welfare that was first derived by Leggett (2002) as a way to describe decision making under misinformation in a static environment without inertia.

$$(11)\ \Delta E[CV_i^s] = \frac{1}{\alpha_{it}^n}\left\{ln\frac{\sum_{k\in K}[exp(V_{ik}^{s1})]}{\sum_{j\in J}\left[exp\left(V_{ij}^{s0}\right)\right]} + \sum_{k\in K}\left[\psi_{ik}^{s1}\left(V_{ik}^{n1} - V_{ik}^{s1}\right)\right] - \sum_{j\in J}\left[\psi_{ij}^{s0}\left(V_{ij}^{n0} - V_{ij}^{s0}\right)\right]\right\},$$

where $V_{ij}^{s0} = V_{ijt}^{s0}(\theta^s) = U_{ijt}^{s0} - \epsilon_{ijt}$, $\theta^s = [\alpha^s, \beta^s, \gamma^s, \eta^s, \delta^s]$, and ψ_{ij} is the logit probability of choosing plan *j* so that $\psi_{ij}^{s0} = exp\left(V_{ij}^{s0}\right)/\sum_{m\in J}[exp(V_{im}^{s0})]$. The first term inside braces in (11) is the standard log sum ratio evaluated at θ^s. The second and third terms adjust the log sum ratio to account for the welfare implications of the difference between θ^s and θ for each choice, weighted by the predicted probability of making that choice before and after the policy. In the special case where $\theta^s = \theta$, equation (11) reduces to the standard welfare measure in (9).

C. Bounding the Welfare Implications of Inertia

Equations (9) and (11) treat the non-suspect group's inertia parameters as being directly relevant for welfare. This is consistent with interpreting inertia as a mixture of latent preferences and hassle costs of switching plans. However, Kling et al. (2002) argue that inertia is more likely to reflect downward biased expectations for the savings from switching plans along with other psychological factors such as status quo bias, procrastination, and limited attention. These mechanisms have no direct effect on consumer welfare; they affect welfare indirectly by lowering the rate at which consumers switch plans. Our data do not allow us to distinguish the importance of psychological bias relative to latent preferences and switching costs. One can separate them, in principle, by adding assumptions on the form of statistical distributions for unobserved preference heterogeneity and switching costs (e.g. Heckman 1981, Dube et al. 2010, Polyakova 2015). We prefer to avoid such assumptions by instead taking a partial identification approach similar to Handel (2013) and Bernheim, Fradkin, and Popov (2015). We calculate welfare for two extreme cases that provide bounds on the share of inertia that is welfare relevant. In the first case, inertia is assumed to be entirely welfare relevant (as in (9) and (11)) and in the second case it is assumed to be entirely irrelevant, e.g. due to psychological bias.

To calculate the change in expected welfare when inertia reflects psychological biases we replace equations (9) and (11) with (9') and (11').

$$(9') \ \Delta E[CV_i^n] = \frac{1}{\alpha_{it}^n}\left\{ln\frac{\sum_{k\in K}[exp(V_{ik}^{n1})]}{\sum_{j\in J}\left[exp\left(V_{ij}^{n0}\right)\right]} + \sum_{k\in K}\left[\psi_{ik}^{n1}\left(V_{ik}^{n^*1} - V_{ik}^{n1}\right)\right] - \sum_{j\in J}\left[\psi_{ij}^{n0}\left(V_{ij}^{n^*0} - V_{ij}^{n0}\right)\right]\right\}.$$

$$(11') \ \Delta E[CV_i^s] = \frac{1}{\alpha_{it}^n}\left\{ln\frac{\sum_{k\in K}[exp(V_{ik}^{s1})]}{\sum_{j\in J}\left[exp\left(V_{ij}^{s0}\right)\right]} + \sum_{k\in K}\left[\psi_{ik}^{s1}\left(V_{ik}^{n^*1} - V_{ik}^{s1}\right)\right] - \sum_{j\in J}\left[\psi_{ij}^{s0}\left(V_{ij}^{n^*0} - V_{ij}^{s0}\right)\right]\right\}.$$

These equations differ from (9) and (11) in that $V_{ik}^{n^*1} = V_{ik}^{n1} - \eta_{it}^n \Delta B_{ijt} - \delta_{it}^n \Delta P_{ijt}$ and $V_{ij}^{n^*0} = V_{ik}^{n0} - \eta_{it}^n \Delta B_{ijt} - \delta_{it}^n \Delta P_{ijt}$. Hence, in this case inertia has no direct effect on consumer welfare; it only affects welfare indirectly via consumers' enrollment decisions.

D. Bounding the Policy's Effect on Consumer Behavior

Prospective welfare analysis also requires us to take a stance on whether a counterfactual choice architecture policy would induce consumers to behave differently. In principle, a policy designed to simplify the choice process could induce decision makers in the suspect group to update their beliefs about the market and behave like decision makers in the non-suspect group. Or it could have no effect at all. In the absence of empirical evidence, we again take a partial identification approach and consider two extreme scenarios. One scenario assumes that the policy has no effect on behavior; the other assumes that the policy induces consumers in the suspect group to behave like those in the non-suspect group, conditional on demographics and prescription drug utilization. The second case involves replacing V_{ik}^{s1} with V_{ik}^{n1} and ψ_{ik}^{s1} with ψ_{ik}^{n1} in equations (11) and (11').

E. Discussion

Our welfare framework is consistent with divergent theories of consumer decision making. When it is costly for consumers to acquire information, to make a decision, or to negotiate a transaction they may choose not to become fully informed (Stigler and Becker 1977). Misinformation may also stem from psychological biases (Kahneman, Wakker, and Sarin 1997).[30] Our framework requires observing which decisions are affected by some combination of these mechanisms, but it avoids the need to model them or take a

[30] To use the terminology from Kahneman, Wakker, and Sarin (1997), one can think of $V_{ij}^n(\theta)$ as approximating the "hedonic utility" derived by consuming a good and $V_{ij}^s(\theta^s)$ as approximating the "decision utility" function maximized by people who are misinformed.

stance on their relative importance. The disadvantage of being unable to disentangle these mechanisms is that we only recover bounds on welfare. Whether the bounds are informative is an empirical question.

The bounds that we derive extend Small and Rosen (1981) to recognize that consumers differ in the information they use to make decisions. Our adjustment for misinformation implements Bernheim and Rangel's (2009) proposal for how to measure welfare when the analyst suspects that some choices will not reveal preferences. This allows us to recognize that choice architecture may create winners and losers. For example, consider the partial equilibrium welfare effects of a policy that automatically assigns each consumer to a plan, but allows them to opt out and choose a different plan if they prefer. Nobody can be made better off from such a policy within a model that assumes all consumers are fully informed and freely mobile (e.g. Lucarelli, Prince, and Simon 2012). At the opposite extreme, nobody can be made worse off within a model that assumes the policy is implemented by a benevolent regulator who knows consumers' preferences better than they know their own preferences (e.g. Abaluck and Gruber 2011). Our approach provides a middle ground between these extremes. Equation (9) and its analogs recognize that informed consumers can be made worse off from restrictions on choice. Equation (11) and its analogs introduce flexibility so that misinformed consumers may gain or lose from restrictions on choice. Aggregating the gains and losses can yield criteria for policy evaluation consistent with the concept of asymmetric paternalism (Camerer et al. 2003).

Our framework also highlights the information needed to evaluate a prospective policy. First we must estimate parameters describing how suspect and non-suspect choice probabilities vary with plan attributes, θ^n and θ^s, in order to calibrate ψ_{ij}^{s0}, V_{ij}^{n0}, V_{ij}^{s0}, and $V_{ik}^{n^*0}$. Then we must map the policy onto plan attributes and utility to calibrate ψ_{ij}^{s1}, V_{ij}^{n1}, V_{ij}^{s1}, and $V_{ik}^{n^*1}$ and calculate bounds on welfare.

VI. Multinomial Logit Estimation

A. Main Results

Table 4 presents the estimates that we use as the basis for policy experiments.[31] The first column reports results for a naïve model that ignores heterogeneity in consumers' decision making processes by pooling data on suspect and non-suspect choices. The main effects have the expected signs and are precisely estimated, with the exception of variance. Its insignificant coefficient mirrors the finding from Abaluck and Gruber (2011) and Ketcham, Kuminoff and Powers (2016) that if we ignore heterogeneity in decision making, then the representative enrollee appears to ignore risk protection.

Columns 2 and 3 repeat the estimation for non-suspect and suspect choices separately. Comparing main effects across the three columns reveals that the insignificant coefficient on variance in the pooled model is driven by aggregating over heterogeneous decision making processes for the suspect and non-suspect groups. Taken literally, the coefficient on variance for the suspect group implies they are risk loving. In contrast, the non-suspect group is risk averse at levels consistent with findings from prior studies (Cohen and Einav 2007, Handel 2013, Handel and Kolstad 2015). For example, our results imply that enrollees in the non-suspect group would be indifferent between a 50-50 bet of winning \$1,000 and losing between \$854.7 and \$937.3.[32] Further, the non-suspect group is more sensitive to price with the implication that the monetary value of inertia—defined by dividing the switching indicators by the expected cost coefficient—is nearly three times larger for the suspect group.

Focusing on non-suspect choices in column 2, the interaction coefficients are consistent with intuition. Interactions between cost and indicators for whether the beneficiary is in the top or bottom terciles of the claims distribution imply that the marginal utility of income declines as people become sicker. People who have previously taken the time to search for information about Medicare programs on the internet or by calling 1-800-Medicare tend to be more sensitive to price and to have stronger preferences for CMS's index of overall plan quality which is based, in part, on customer satisfaction. Preferences

[31] We also estimated more flexible models that interacted PDP attributes with more comprehensive sets of demographic variables. However the additional interactions tend to have small and statistically insignificant effects (Table A4), which led us to use the more parsimonious specification in Table 4. A notable result from the more comprehensive model is that enrollees who do and do not get help making health insurance decisions make choices that imply virtually identical marginal rates of substitution between cost, variance, and quality. The main difference between the two groups is that those who get help exhibit less inertia, as shown in Table 4.

[32] These calculations are based on the fact that our specification for utility provides a 1st order approximation to a CARA model. Our calculations are additional discussion are provided in Table A5 and associated discussion in the supplemental appendix.

for plan quality are also higher among higher income enrollees. One explanation is that the opportunity cost of time is increasing in income and that choosing a higher quality plan reduces the time and effort required to interact with the insurer.

TABLE 4—LOGIT MODELS OF PRESCRIPTION DRUG PLAN CHOICE

	All Choices		Non-Suspect choices		Suspect choices	
expected cost	-0.283	[0.017]***	-0.377	[0.029]***	-0.197	[0.021]***
variance	0.076	[0.085]	-0.433	[0.118]***	0.621	[0.126]***
quality (CMS index)	0.035	[0.078]	0.056	[0.104]	-0.012	[0.124]
within-brand switch	-3.307	[0.109]***	-3.239	[0.152]***	-3.396	[0.155]***
between-brand switch	-5.181	[0.095]***	-4.923	[0.128]***	-5.591	[0.141]***
cost x 1{ bottom tercile of claims }	-0.172	[0.034]***	-0.194	[0.039]***	-0.089	[0.053]*
cost x 1{ top tercile of claims }	0.082	[0.021]***	0.128	[0.035]***	0.027	[0.024]
cost x 1{ sought CMS info }	-0.043	[0.022]*	-0.074	[0.032]**	0.037	[0.030]
quality x 1{ income > $25k }	0.170	[0.091]*	0.202	[0.118]*	0.095	[0.147]
quality x 1{ sought CMS info }	0.283	[0.096]***	0.241	[0.122]**	0.326	[0.165]**
switch within brand x standardized age	-0.162	[0.069]**	-0.138	[0.093]	-0.179	[0.103]*
switch within brand x 1{ income > $25k }	-0.383	[0.126]***	-0.364	[0.169]**	-0.373	[0.183]**
switch within brand x 1{ help }	0.335	[0.122]***	0.271	[0.170]	0.474	[0.181]***
switch within brand x 1{ sought CMS info }	0.126	[0.131]	0.262	[0.167]	-0.200	[0.208]
switch within brand x 1{ nonwhite }	-0.812	[0.297]***	-1.211	[0.450]***	-0.587	[0.396]
switch brand x standardized age	-0.122	[0.055]**	-0.167	[0.073]**	0.025	[0.081]
switch brand x 1{ income > $25k }	-0.390	[0.106]***	-0.411	[0.139]***	-0.429	[0.163]***
switch brand x 1{ help }	0.263	[0.105]**	0.233	[0.141]*	0.383	[0.160]**
switch brand x 1{ sought CMS info }	0.285	[0.102]***	0.178	[0.133]	0.263	[0.165]
switch brand x 1{ nonwhite }	-0.794	[0.239]***	-1.371	[0.348]***	-0.107	[0.341]
pseudo R^2	0.66		0.64		0.71	
number of enrollment decisions	9,119		5,248		3,871	
number of enrollees	3,442		2,175		1,560	

Note: The table summarizes logit models estimated from data on all choices; non-suspect choices only; and suspect choices only. All models include indicators for insurers. Excluded demographic interactions define the reference person as white and 78 years old with no college degree and annual income below $25,000. This person is in the middle tercile of the distribution of total drug claims, did not get help making an enrollment decision, and did not use the internet or 1-800-Medicare to search for information. Robust standard errors are clustered by enrollee. *,**, and *** indicate that the p-value is less than 0.1, 0.05, and 0.01 respectively.

Inertia tends to be lower for people who get help choosing a plan and who searched for information about CMS programs, whereas it tends to be higher for people who are older, nonwhite and who have higher incomes, though some of these effects are not precisely estimated. The income effect could again be due to heterogeneity in the opportunity cost of time. The directions of these effects are mostly consistent across the suspect and non-suspect groups, but the monetary implications are larger for the suspect group. The average non-suspect enrollee would have to be paid $846 to hold their utility constant if they were randomly reassigned to a different plan offered by the same insurer or $1,292 if they were reassigned to a plan offered by a different insurer. Comparable figures for the suspect group are $1,888 and $2,958. The fact that we see greater inertia for between-insurer switches compared to within-insurer switches is consistent with the inertia parameters reflecting latent preferences and hassle costs. Between-insurer switches are likely to require more time and effort than within-insurer switches as different plans offered by the same insurer tend to have the same formularies, pharmacy networks, customer service, and so on. In contrast, insurers typically differ along these dimensions, so that switching insurers may require new prior authorization requests, transferring prescriptions to new pharmacies, and becoming familiar with new formulary and customer service systems. Psychological biases might also be greater for between-brand switches.

B. Validation Tests

A potential concern with our approach to modeling heterogeneity in consumer decision making is that it could be overfitting the data and consequently yielding less accurate predictions for how consumers will respond to prospective policies. We assess the model's predictive power by using validation tests similar to Keane and Wolpin (2007) and Galiani, Murphy, and Pantano (2015). The idea is to compare the out of sample predictions from our model with the standard pooled model that assumes a homogeneous decision process. Our validation test is powered by the largest year-to-year change in the PDP choice set that occurred during our study period. Between 2008 and 2009 the number of plans fell by 10%. We use data from 2008 to estimate the standard and refined models and then use each set of estimates to predict how consumers would adapt to their new choice sets in 2009.[33] Table A6 shows that among suspect choosers the refined model

[33] We exclude indicators for insurance brand because some new insurers joined the market in 2009 so we are unable to estimate indicators for them in 2008.

more accurately predicts the share that chose dominated plans; the share that chose the least expensive plans offered by their insurers; mean expenditures; the average amount that consumers who chose dominated plans could save by switching; and the share who chose to switch plans and the share who chose plans with gap coverage. The refined model likewise outperforms the pooled model in making out-of-sample predictions for the choices of non-suspect choosers for all but two of these measures. Overall, this exercise suggests that distinguishing between suspect and non-suspect choice processes improves the model's predictive power out of sample.

As an indirect test of our maintained assumption that people in the suspect and non-suspect groups share the same underlying utility parameters, conditional on demographics and prescription drug use, we leverage the panel structure of our data to repeat the estimation for four mutually exclusive sets of enrollment decisions: (1) choices made by enrollees who always make suspect choices (n=3,311); (2) suspect choices made by enrollees who sometimes make non-suspect choices (n=560); (3) non-suspect choices made by enrollees who sometimes make suspect choices (n=634); and (4) choices made by enrollees who always make non-suspect choices (n=4,616). The results, shown in Tables A7-A8, reveal that the estimated marginal rates of substitution between cost, variance, and quality are similar between groups 1 and 2, and between groups 3 and 4, despite some reduction in statistical precision. In other words, when people who switch between the suspect and non-suspect groups make non-suspect choices they behave in similar ways to the people who always make non-suspect choices. This supports the assumptions underlying our approach of using non-suspect preference parameters to predict welfare effects for people in the suspect group.

VII. Evaluating Prospective Choice Architecture Policies

A. Preliminaries

i. Calculating Changes in Insurer Revenue

After adjusting consumers' choice sets to depict each prospective policy, we use the resulting changes in choice probabilities to calculate changes in insurer revenue per enrollee, holding premiums fixed. The purpose of this exercise is to develop insight on the strength of insurers' incentives to respond to prospective government policies by adjusting premiums and other PDP attributes, without having to assume a parametric form for the PDP production function or having to model how it arises from interactions between competing insurers, drug companies and the government.[34] Developing a comprehensive supply side model of prescription drug insurance markets is a challenging and potentially important task for future research. Our partial equilibrium approach in this paper is similar to prior studies that have assumed premiums adjust in proportion to changes in insurer revenue per enrollee (Handel 2013, Ho, Hogan, and Scott-Morton 2015).

Equation (12) defines the change in insurer revenue per enrollee:

$$(12)\ \Delta\pi = \frac{1}{N}\sum_i \sum_{k\in K} \psi_{ik}^1\, \pi_{ik}^1 - \frac{1}{N}\sum_i \sum_{j\in J} \psi_{ij}^0\, \pi_{ij}^0,$$

where π_{ij}^0 and π_{ik}^1 measure insurer revenue per enrollee before and after the policy.[35] The change in revenue per enrollee is determined by whether the policy mitigates or exacerbates adverse selection based on predicted changes to choice probabilities (Handel 2013)

ii. Bounding Estimated Outcomes

Section V explained our approach to bounding the welfare effect of inertia and the policy's effect on consumer behavior. We use these bounds to report results for two extreme cases. At one extreme is the case where the policy is "most effective" as a nudge in the sense that it causes the suspect group to start behaving like the non-suspect group ***and*** the inertia parameters estimated for the non-suspect group reflect psychological bias and hence have no direct effect on welfare, i.e. using equation 9' and 11' with V_{ik}^{n1} and ψ_{ik}^{n1}. At the other extreme is the case where the policy is "least effective" as a nudge in that it does not change the suspect group's behavior ***and*** the inertia parameters for the non-

[34] In general, prospective policies may also have important implications for government spending due to the large subsidies for PDP purchases. We omit them here because changes in these subsidies are almost entirely due to changes in premiums. In our partial-equilibrium counterfactuals, the implied government spending never changes by more than a few dollars per enrollee.

[35] Empirically, we define insurer revenue per enrollee as the total premium (paid partly by enrollees and partly by the government) less residual drug expenditures, defined as total expenditures less the sum of consumers' OOP costs and government payments for consumers who exceed the threshold for catastrophic spending. We assume the average cost of plan management and operations per enrollee is unchanged by the policy so that it cancels out of the difference in (12).

suspect group reflect the hassle cost of switching plans and/or preferences for latent plan attributes and hence are welfare relevant (i.e. using equations 9 and 11).[36] To provide statistical bounds on our estimates, we report the 2.5th percentile from a 100 replication bootstrap for the least effective scenario and the 97.5th percentile for the most effective scenario.

B. Distributional Effects of a Menu Restriction

In early 2014, CMS proposed a series of changes to Medicare Part D that included a provision to limit each parent organization to offering no more than one basic and one enhanced plan per region (Department of Health and Human Services 2014).[37,38] This would have forced some current enrollees to switch plans. While the proposal was controversial and has yet to be implemented, it provides an opportunity to investigate the effects of a realistic menu restriction.

We use the sets of enrollees and plans in 2010—the last year of our sample—as the baseline for simulating the welfare effects of CMS's proposed menu restriction. Our data for that year describe 2,611 individuals, both new enrollees and those with experience. CMS must approve each PDP that an insurer offers, but the proposed regulation did not specify how, exactly, CMS would determine which plans to retain. Therefore we start by assuming that CMS would require each sponsor to continue to offer their most popular plans; i.e. the single basic plan and the single enhanced plan with the highest enrollments.[39] Then we consider alternative rules as robustness checks. The menu restriction reduces the number of plans on the average enrollee's menu from 47 to 31.

The menu restriction affects consumer welfare in several ways. First, people will be made worse off if their utility maximizing plans are eliminated. Second, individuals who switch plans may incur utility costs of switching. Third, individuals in the suspect group

[36] Alternatively, one could solve jointly for a continuous faction of inertia that is welfare relevant and a continuous faction of suspect group consumers who start behaving like their analogs in the non-suspect group in order to minimize and maximize particular moments of the distribution of welfare effects.
[37] "Parent organizations" or "sponsors" are entities that contract with CMS to sell PDPs. They may include multiple brand names. Basic plans may differ in design but must be deemed actuarially equivalent to the standard benefits package for some representative enrollee(s). Enhanced plans offer supplemental benefits.
[38] The proposal included the rationale to "…ensure that beneficiaries can choose from a less confusing number of plans that represent the best value each sponsor can offer" (Federal Register 2014).
[39] This is consistent with our interpretation of CMS' impact analysis (Federal Register 2014).

may be made better off if the policy forces them to switch out of a dominated plan or if contracting their choice sets reduces their inertia and nudges them to switch to plans that are cheaper, higher quality, and provide better insurance against health shocks. The magnitude of each of these gains or losses depends on which plans are eliminated and the relative benefits of switching.

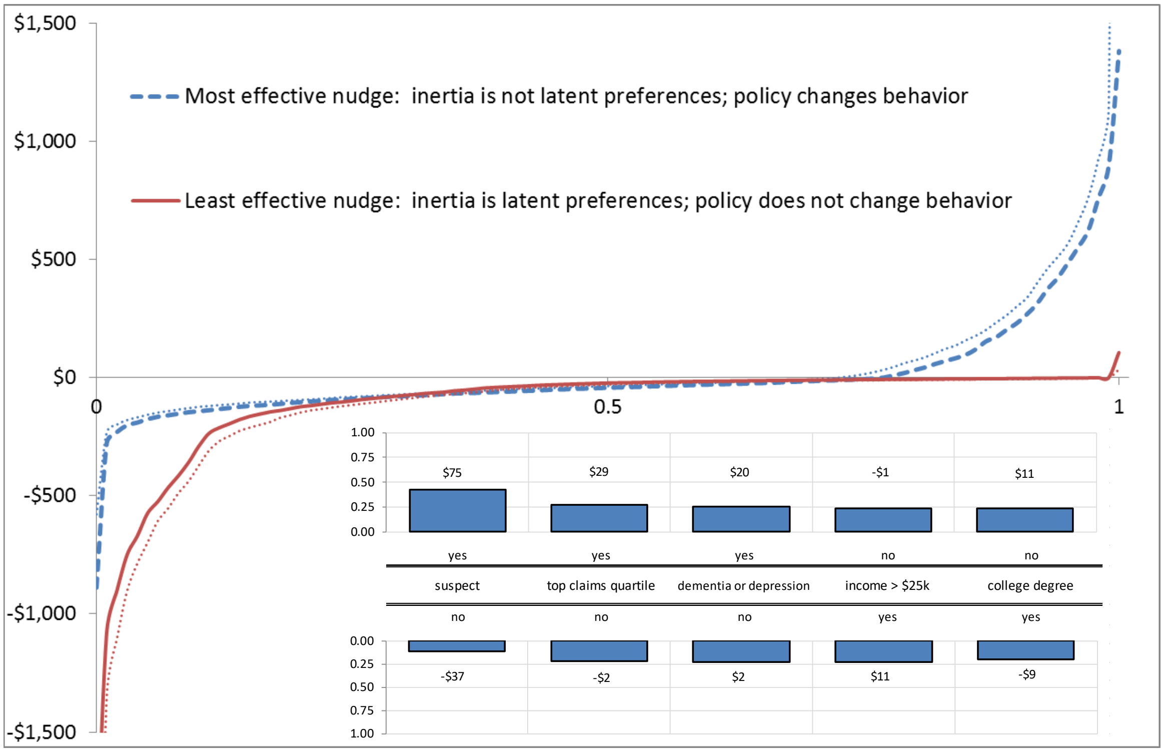
To summarize results we start by focusing on the case in which CMS requires each insurer to retain their basic and enhanced plans with the highest numbers of enrollees. Figure 1 summarizes the distributional effects on the beneficiary population. It shows CDFs of the expected consumer surplus under the "most effective" and "least effective" scenarios for the efficacy of the policy in nudging consumers (henceforth ME and LE). The bar charts in the bottom half of the figure summarize the average changes in expected consumer surplus and the shares of consumers with expected welfare gains under the ME scenario for several types of people who might be of interest to policymakers: (i) those making suspect choices, or not, (ii) those in the top quartile of the distribution of total drug claims, or not, (iii) those with dementia or depression, or not, (iv) those with income over $25,000, or not, and (v) those with a college degree, or not.[40] In both the ME and LE scenarios fewer than 25% of consumers are made better off by the menu restriction. Further, the median consumer in every one of the 10 demographic groups is made worse off. While those in the suspect group have larger average gains and a higher probability of gains than those in the non-suspect group (the bootstrap confidence intervals show these are significantly different at 1%) even the median consumer in the suspect group is expected to lose from menu restrictions.

Figure 2 summarizes the mechanisms that drive welfare effects in the ME and LE scenarios. It reports the shares of winners and losers who are forced to switch because the policy eliminates their default plans, followed by the expected reductions in their premiums and OOP expenditures, the expected reductions in their expenditure variances, and the expected increases in plan quality (both the CMS quality index and the index of latent quality defined by the insurer dummy variables). Changes in variance and quality are

[40] This comparison is less interesting in the LE scenario because in that scenario virtually all consumers have welfare losses.

converted to dollar equivalents using the non-suspect group's marginal utility of income.

FIGURE 1: DISTRIBUTION OF THE WELFARE EFFECTS FROM A MENU RESTRICTION



Note: The figure shows CDFs of the expected change in welfare from limiting each insurer to selling one basic plan and one enhanced plan, assuming that CMS requires insurers to keep the plans with the highest current enrollment. The small dotted lines represent the nonparametric 95% upper bound on the most effective nudge and the 95% lower bound on the least effective nudge based on a 100 replication bootstrap. The bar charts show the fractions of consumers with welfare gains by demographic group and the numbers above or below each bar report average consumer surplus within the groups.

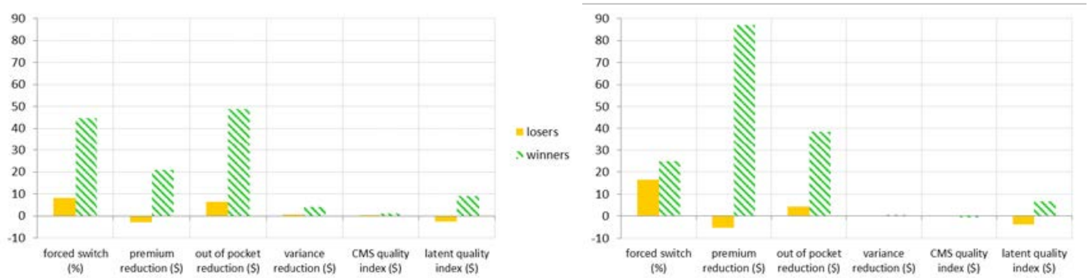
In the ME scenario just under 25% of consumers are made better off.[41] Nearly half of those winners are forced to switch plans. Many of the people who are forced to switch, particularly those in the suspect group, are better off from switching because their new plans provide more generous coverage and there is no utility cost of switching in the ME scenario. Furthermore, by assumption, people in the suspect group now place more emphasis on cost and risk protection when selecting plans. As a result, the average winner pays $21 less in expected premiums and $41 less in expected out of pocket costs after the policy. Their expected risk exposure declines by an amount equivalent to a certain pay-

[41] We code anyone with changes in welfare of |$0.01| or less as having no change. In tables 5, 6, and 7 the percentages of consumers with no change in welfare equal 100 minus the reported percentages with welfare gains and losses.

ment of $4 and they have an expected improvement in plan quality worth just over $10 (summing the effects of the CMS index and insurer dummies). Nevertheless, most people experience welfare losses because they become enrolled in plans with higher costs and they have more desirable plans eliminated. A small number of consumers, particularly those in the non-suspect group, experience relatively large losses because the policy eliminates their chosen plans, resulting in substantially higher expected premiums and lower expected quality.

FIGURE 2: MECHANISMS UNDERLYING THE WELFARE EFFECTS OF A MENU RESTRICTION

A. MOST EFFECTIVE NUDGE B. LEAST EFFECTIVE NUDGE



Note: The first column reports the shares of consumers with expected welfare gains (winners) and expected welfare losses (losers) who are forced to switch because their chosen plans are eliminated. The next two columns report expected reductions in premiums and out of pocket expenditures. The last three columns use the marginal utility of income for the non-suspect group to report the expected reduction in variance and expected increases in plan quality in monetary equivalents.

TABLE 5: SUMMARY OF OUTCOMES FOR ALTERNATIVE MENU RESTRICTION RULES

	Max enrollment		Max frontier		Min expenditures		Max profit	
	most effective	least effective	most effective	least effective	most effective	least effective	most effective	least effective
% enrollees with default plan eliminated	16.7 (0.0)	16.7 (0.0)	24.7 (0.0)	24.7 (0.0)	19.4 (0.0)	19.4 (0.0)	37.0 (0.0)	37.0 (0.0)
Δ expected welfare / enrollee ($)	6.0 (9.0)	-106.8 (8.3)	26.0 (9.2)	-162.5 (10.3)	32.1 (9.0)	-125.4 (9.0)	22.2 (8.9)	-218.8 (11.3)
% enrollees with expected welfare gain	23.1 (1.6)	1.2 (0.6)	29.0 (1.6)	1.2 (0.5)	27.8 (1.6)	1.4 (0.6)	31.9 (1.2)	0.6 (0.4)
% enrollees with expected welfare loss	76.9 (1.6)	98.8 (0.6)	71.0 (1.6)	98.8 (0.5)	72.2 (1.6)	98.6 (0.6)	68.1 (1.2)	99.3 (0.4)
Δ insurer revenue / enrollee ($)	-8.0 (4.0)	9.7 (3.3)	-14.4 (4.2)	8.0 (3.0)	-15.5 (4.5)	2.4 (3.6)	21.2 (5.4)	45.9 (3.9)

Note: The table shows the sensitivity of outcomes to the menu restriction rule. Max enrollment is the baseline that corresponds to figures 1 and 2. Max frontier retains the basic and enhanced plans with the highest shares of enrollees on the efficiency frontier. Min expenditure retains plans with the lowest average expenditures. Max profit allows insurers to retain the plans with the highest average profit per enrollee. Standard errors from a 100 replication bootstrap are in parentheses.

In the LE scenario, only 2% of consumers are made better off. For most people, the utility loss from being forced to switch plans more than offsets the cost savings, risk reduction, and improvements in plan quality experienced by switchers. The small fraction of winners is comprised entirely of individuals in the suspect group who have large reductions in expected premiums and expected OOP costs. Hence, if we think that inertia primarily reflects hassle costs and consumer preferences, then the menu restriction significantly harms the vast majority of consumers in exchange for small benefits for a small share of people in the suspect group who become less able to choose inferior plans.

The first two columns of Table 5 summarize the shares of people who have their default plans eliminated by the policy, the average changes in expected welfare per enrollee, the shares of winners and losers and the changes in insurer revenue per enrollee. The ME scenario predicts a net effect on consumer welfare that is statistically indistinguishable from zero, as large gains for a small fraction of consumers offset smaller losses for the majority. The LE scenario predicts a statistically significant mean welfare reduction of -$107, as 99 percent of consumers are made worse off. The last six columns show comparable results for three other hypothetical rules for how CMS could determine which plans to keep on the menu: the plans that are on the efficiency frontier for the greatest number of people; the plans with the minimum average cost to the enrollee; and the plans with the highest net revenue per enrollee.[42] Our results on consumer welfare are qualitatively robust across these scenarios. The most striking differences are the reductions in consumer welfare and increases in insurer revenue that occurs when insurers are allowed to retain their highest profit plans. Under the LE scenario, welfare is expected to fall by $219, amounting to 15.6% of enrollees' average spending. Insurer net revenue per enrollee has expected increases even under the ME scenario ($21) nearly as large as the gains in average consumer welfare ($22). The more profitable plans tend to be the higher-premium ones that provide more risk reduction and have higher quality ratings. Hence, insurers would have strong incentives to persuade regulators to allow them to retain their more comprehensive plans. With approximately 7.7 million people participating in the standalone Medicare prescription drug markets, this partial equilibrium change in insurer

[42] For profitability, we assume that there is sufficiently little variation in costs of plan operations and management per enrollee within the set of plans offered by each insurer that it does not affect the ranking of plans by revenue per enrollee. .

net revenue under the maximum profit rule increases insurer revenues by $163 to $353 million per year.

C. Distributional Effects of Personalized Decision Support

Our second policy experiment is a hypothetical decision support tool modeled on a randomized field experiment conducted by Kling, Mullainathan, Shafir, Vermeulen, and Wrobel (2012) [henceforth KMSVW]. Their study is motivated by the observation that while Medicare enrollees can learn about their PDP options and potential savings by calling 1-800-Medicare or using various online cost calculators, a minority of enrollees report doing so, as seen in Table A1. KMSVW attribute this to "comparison friction" which they define as the wedge between available information and consumers' use of it. KMSVW tested an intervention in which several hundred treatment group enrollees were sent a decision support letter containing personalized information about their potential personal cost savings from switching to their lowest cost available plan. The letter also identified the name of the low cost insurer and contact information to initiate a switch. KMSVW observed a 7 percentage point increase in the rate at which the treatment group switched to the plan identified in the letter relative to a control group that received a general letter with no personalized decision support, and an 11.5 percentage point increase in the overall switching rate for the treatment group.

We estimate the welfare effects of a prospective national rollout of the decision support tool in which the government mails letters to all existing enrollees that would be worded similarly to the one sent to KMSVW's treatment group. Because the information relies on prior drug claims, the policy would not affect new enrollees. Such a policy may affect welfare via several pathways. First, providing enrollees with personalized information may make them better off by mitigating psychological biases and/or reducing information costs. In our model, this would be realized as increases in the switch rate and cost savings. Because KMSVW's decision support tool does not embed information about risk protection and quality, however, the net effect on welfare is ambiguous. Second, an important feature of the information campaign—if it were implemented by the government—is that it would necessarily be based on incomplete information about enrollees' drug needs. While CMS has full information about existing enrollees' individual claims over their prior years in the PDP market, individuals may have private information about their own drug needs over the upcoming year. If enrollees with private information about changes in their drug needs choose to switch plans based on outdated information provided by CMS then these misinformed individuals could experience welfare losses.[43]

We use KMSVW's estimated treatment effects as moments to calibrate V_{ij}^{n1} and V_{ij}^{s1}. Specifically, in the ME scenario we multiply the estimated inertia parameters by $\omega_1(1+\omega_2 1\{j=j^*\})$ as shown in (13.a) and (13.b), where $1\{j=j^*\}$ is an indicator for whether plan *j* is the individual's minimum cost plan that would be featured in the letter. We calibrate ω_1 to generate a 7 percentage point increase in the rate at which the treatment group switches to their lowest cost plan relative to the baseline we observe in the data, and we calibrate ω_2 to simultaneously generate an 11.5 percentage point increase in the overall switch rate subject to the constraints that $0 \leq \omega_1, \omega_2, \omega_1+\omega_2 \leq 1$.

$$(13.a)\ \ V_{ijt}^{n1} = \hat{\alpha}_{it}^{n} c_{ijt} + \hat{\beta}_{it}^{n} \sigma_{ijt}^{2} + \hat{\gamma}_{it}^{n} q_{jt} + \omega_1(1+\omega_2 1\{j=j^*\})\left(\hat{\eta}_{it}^{n} \Delta B_{ijt} + \hat{\delta}_{it}^{n} \Delta P_{ijt}\right).$$

$$(13.b)\ \ V_{ijt}^{s1} = \hat{\alpha}_{it}^{s} c_{ijt} + \hat{\beta}_{it}^{s} \sigma_{ijt}^{2} + \hat{\gamma}_{it}^{s} q_{jt} + \omega_1(1+\omega_2 1\{j=j^*\})\left(\hat{\eta}_{it}^{s} \Delta B_{ijt} + \hat{\delta}_{it}^{s} \Delta P_{ijt}\right).$$

$$(13.c)\ \ V_{ijt}^{s1} = \hat{\alpha}_{it}^{s} c_{ijt} + \hat{\beta}_{it}^{s} \sigma_{ijt}^{2} + \hat{\gamma}_{it}^{s} q_{jt} + \hat{\eta}_{it}^{s} \Delta B_{ijt} + \hat{\delta}_{it}^{s} \Delta P_{ijt}.$$

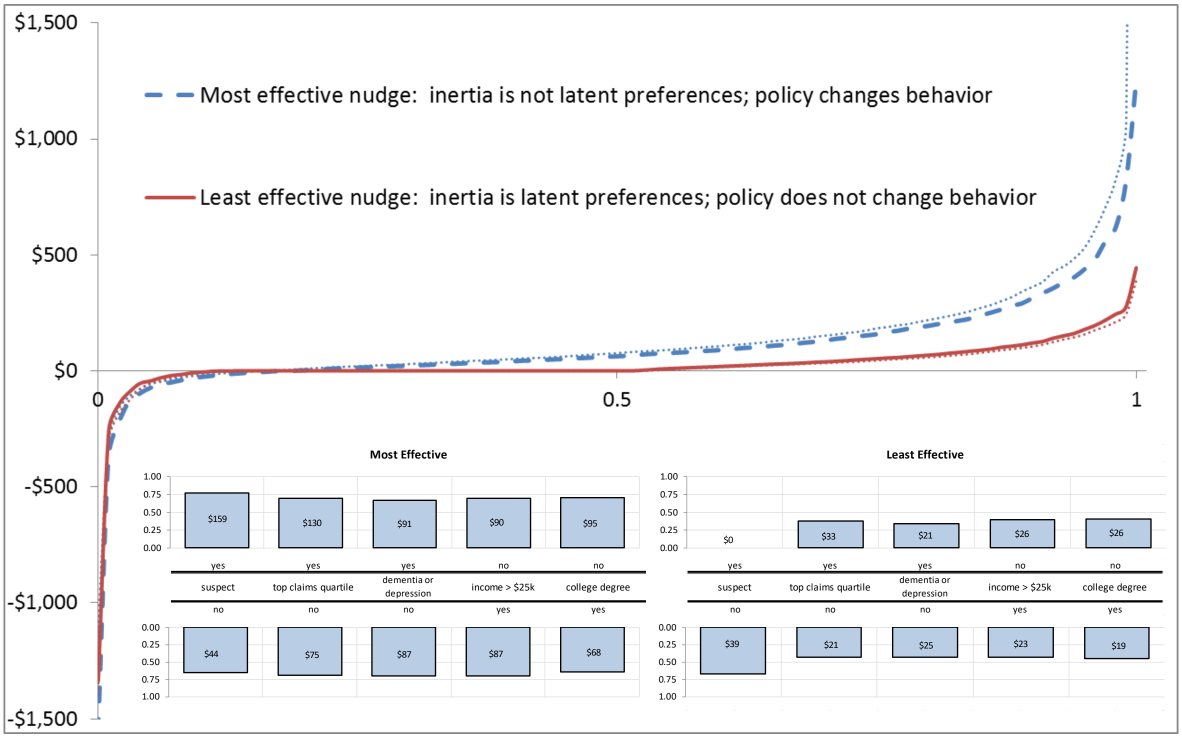
In the LE scenario, there is assumed to be no change in the behavior of the suspect group so we use (13.a) and (13.c), in which case ω_1 and ω_2 will have to be larger than in the ME scenario in order to induce sufficient switching among the non-suspect group to replicate the treatment effects estimated by KSMVW.

Figure 3 summarizes the distributional effects of the decision support tool. In the ME scenario 81 percent of consumers are made better off by the policy. Those who made suspect choices under the status quo policy are more likely to win and experience larger gains than those who did not (significant at 1%) and those with the highest number of drug claims are expected to have larger average gains than those with fewer claims (significant at 1%), but we do not find any other notable differences across demographic

[43] In principle such a phenomenon could occur if the free but imperfect information from CMS reduces individuals' efforts to acquire private information about their own future drug needs. Carlin, Gervais, and Manso (2013) explore these ideas more generally.

groups. In the LE scenario, the share of consumers with welfare gains declines to 48 percent because the suspect group is assumed to ignore the information treatment. Thus, they are unaffected by the policy.

FIGURE 3: DISTRIBUTION OF WELFARE EFFECTS FROM PERSONALIZED DECISION SUPPORT



<u>Note</u>: The figure shows CDFs of the expected change in welfare from a personalized decision support tool that is based on the field experiments of Kling et al. (2012). The model is calibrated to reproduce their estimated treatment effects on the rates at which people switch plans. The small dotted lines represent the nonparametric 95% upper bound on the most effective nudge and the 95% lower bound on the least effective nudge based on a 100 replication bootstrap. The bar charts show the fractions of consumers with welfare gains by demographic group and the numbers above or below each bar report average consumer surplus within the groups.

To reveal the mechanisms underlying the welfare losses, Table A9 shows that under both scenarios, losers had much larger changes in actual OOP drug spending between 2009 and 2010. This is because the low cost plan featured in the information treatment is the one that minimizes their expenditures based on their 2009 drug claims. Some of the people who experience significant health shocks would have spent substantially more in their government recommended minimum cost plans than in the plans that they actually chose for themselves in 2010. These individuals are more likely to have made non-suspect choices. This illustrates the potential welfare losses that can arise from a nudge based on incomplete information. More broadly, this suggests a tradeoff between the potential benefits of simplifying the presentation of information and the potential costs of deemphasizing important details about the assumptions underlying that information.

TABLE 6: SUMMARY OF OUTCOMES FROM THE PERSONALIZED DECISION SUPPORT TOOL AND SENSITIVITY TO DECISION MAKERS' EXPECTATIONS

	Unbiased Expectations		Myopia	
	most effective	least effective	most effective	least effective
Δ expected welfare / enrollee ($)	102.9 (9.2)	28.1 (3.7)	158.2 (12.2)	62.4 (3.1)
% enrollees with expected welfare gain	81.1 (1.0)	48.4 (1.0)	91.8 (0.8)	54.4 (1.1)
% enrollees with expected welfare loss	18.8 (1.0)	12.4 (0.7)	2.1 (0.7)	0.0 (0.0)
% enrollees switching to the advertised plan	8.0 (0.1)	8.0 (0.1)	8.0 (0.1)	8.0 (0.1)
Δ insurer revenue / enrollee ($)	-10.8 (3.8)	0.4 (2.8)	-34.2 (3.5)	-16.6 (1.8)

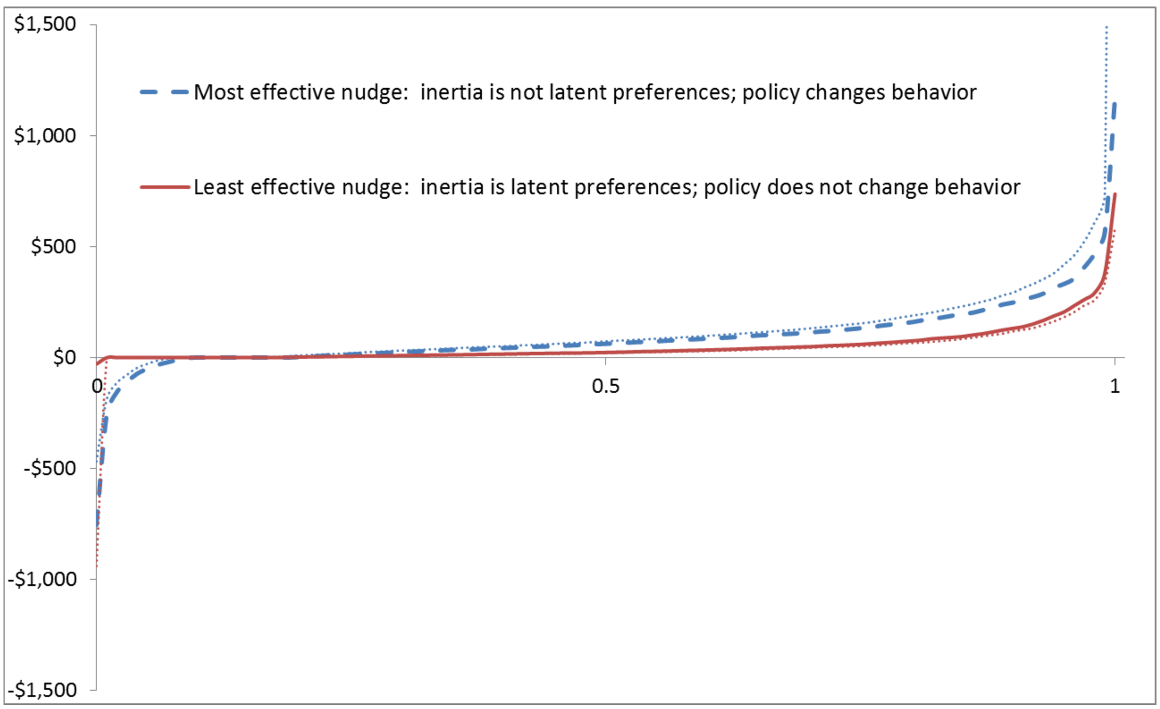
Note: The table shows the sensitivity of outcomes to the assumed form of decision makers' expectations for their own drug needs in the upcoming year. The baseline scenario that corresponds to figures 3 and 4 (perfect foresight) assumes that decision makers accurately forecast changes in their drug needs. The myopia scenario assumes that decision makers expect their future drug needs to be identical to the prior year. Standard errors from a 100 replication bootstrap are in parentheses.

The first two columns of Table 6 provides summary statistics for the outcomes under the ME and LE scenarios while maintaining our model's assumption that consumers have unbiased expectations of their actual drug use in the upcoming year. The average welfare gains range from $28 to $103. The unbiased expectations assumption could cause us to understate the policy's benefits. If consumers are myopic in the sense that they expect their drug use to be the same as the prior year then the information treatment has less scope to reduce some consumers' welfare. The last two columns of Table 6 demonstrate this and show that when we repeat the estimation and simulation based on the assumption that consumers are myopic when they enroll in insurance plans, then between 54% and 92% of consumers benefit from the policy and the average change in welfare is an increase of between $62 and $158.

D. Distributional Effects of Default Assignment to a Low Cost Plan

Our final policy experiment replaces CMS's current approach to defining each person's default plan for reenrollment as the plan they previously chose for themselves with a policy that would set the default plan to be the one that minimizes CMS's expectation for each enrollee's costs. We envision the policy being implemented as a stronger version of the decision support tool. Instead of informing enrollees of their minimum cost options, the enrollees would be automatically assigned to those options unless they chose to opt out by overriding the reassignment and choosing a different plan. As before, we assume CMS would predict each enrollee's minimum cost plan using their drug claims from the prior year. Consistent with CMS's current approach, first-time enrollees would still be required to make active decisions.

FIGURE 4: DISTRIBUTION OF WELFARE EFFECTS FROM ASSIGNMENT TO A DEFAULT PLAN



Note: The figure shows CDFs of the expected change in welfare from automatically assigning people to default plans, assuming it is costless to opt out. People are automatically assigned to the plan that would minimize their expenditures based on their prior year of drug use. The small dotted lines represent the nonparametric 95% upper bound on the most effective nudge and the 95% lower bound on the least effective nudge based on a 100 replication bootstrap.

In the ME scenario, the policy completely erases inertia for enrollment in the new low-cost default. Nevertheless, some consumers may still prefer their original plans if those plans provide greater quality or variance reduction. Assuming it is costless for enrollees to opt out and continue in their old plans, under ME assumptions the policy could reduce consumer welfare from (mis)assignment to plans requiring higher expenditures due to changes in drug needs or by reducing the probability of switching to higher cost plans that also provide higher utility due to superior risk protection and/or quality. Figure 4 shows that for a large share of consumers the net change is dominated by the aggregate effect of lower expenditures and the elimination of inertia. Overall, just over 80% of consumers have gains in expected welfare in both scenarios, and this is accompanied by reductions in insurer revenue of $42 to $128 as shown in Table 7.[44]

In the LE scenario, being assigned to a default plan does not eliminate the hassle cost of learning to navigate a plan offered by a different insurer (e.g. prior authorization paperwork, new pharmacy networks, new customer service protocols). To account for this we recalibrate the model so that the policy reduces the cost of switching to the low-cost default from $\hat{\eta}_{it}\Delta B_{ijt} + \hat{\delta}_{it}\Delta P_{ijt}$ to $(\hat{\eta}_{it} - \hat{\delta}_{it})\Delta B_{ijt}$. Under this interpretation, the welfare-relevant hassle costs are the difference in the estimated cost of switching between brands relative to switching within brands. The continued presence of navigation costs reduces the share of enrollees choosing their assigned default to 14%.[45] The right half of Table 7 shows that the share of consumers who benefit, their average welfare gain, and the implications for government spending and insurer revenue are virtually unchanged if we repeat the estimation and simulation under the assumption that consumers have myopic expectations of their own drug needs for the upcoming year.

Table 7 also illustrates the importance of the design of the opt-out feature. People may incur some disutility from the time and effort required to pay attention to the new policy, learn how the opt-out feature works, determine whether they prefer their newly assigned default to their old plan and, if not, to exercise their opt out option. Under the

---

[44] We do not find any differences in average gains or the probability of gain across observed consumer attributes, so we suppress the complementary bar chart for brevity.

[45] This approach may still overstate benefits to the extent that $\hat{\eta}$ and $\hat{\delta}$ represent latent preferences. As we increase the post-policy cost of switching to the new default option to $\hat{\eta}\Delta B_{ijt} + \hat{\delta}\Delta P_{ijt}$ the benefits to consumers approach zero. The extreme case in which $\hat{\eta}$ and $\hat{\delta}$ are entirely latent preferences is equivalent to reverting to the pre-policy equilibrium in which case the policy has no effect on consumer welfare.

assumption that everyone faces the same disutility parameter from opting out we solve for the mean opt-out cost needed to set the average change in expected welfare to zero. It ranges from a low of \$65 in the LE scenario with unbiased expectations to a high of \$198 in the ME scenario under myopia. When people incur such utility losses from opting out, some of them choose the newly assigned default even when it is welfare reducing relative to their prior plan in the absence of opt out costs.

TABLE 7: SUMMARY OF OUTCOMES FROM THE DEFAULT ASSIGNMENT RULE AND SENSITIVITY TO DECISION MAKERS' EXPECTATIONS

	Unbiased Expectations		Myopia	
	most effective	less effective	most effective	less effective
Δ expected welfare / enrollee ($)	88.7	50.4	116.7	66.1
	(9.2)	(16.6)	(13.2)	(26.6)
% enrollees with expected welfare gain	81.0	82.8	82.5	83.2
	(1.1)	(1.0)	(1.1)	(1.0)
% enrollees with expected welfare loss	9.3	0.5	7.9	0.0
	(1.0)	(0.4)	(0.9)	(0.3)
opt out cost needed to set average Δ in expected welfare to zero ($)	134.7	64.6	197.9	88.0
	(16.2)	(27.3)	(28.9)	(46.9)
% enrollees switching to the default plan	40.0	14.0	44.0	15.3
	(1.1)	(0.8)	(1.4)	(0.9)
Δ insurer revenue / enrollee ($)	-128.2	-41.9	-129.8	-41.8
	(6.7)	(3.2)	(7.6)	(3.5)
Suspect group only				
Δ expected welfare / enrollee ($)	157.0	48.6	198.5	63.0
	(21.2)	(6.6)	(30.1)	(11.5)
Δ insurer revenue / enrollee ($)	-132.2	-25.3	-135.5	-23.7
	(22.1)	(10.3)	(26.0)	(9.1)
Non-suspect group only				
Δ expected welfare / enrollee ($)	45.0	51.6	63.4	68.2
	(5.2)	(24.0)	(6.7)	(37.6)
Δ insurer revenue / enrollee ($)	-97.4	-40.8	-103.4	-42.4
	(13.3)	(5.9)	(14.9)	(6.1)

Note: The table shows the sensitivity of outcomes to the assumed form of decision makers' expectations for their own drug needs in the upcoming year. The first four rows assume no opt out cost. See the text for additional details and definitions. Standard errors from a 100 replication bootstrap are in parentheses.

Finally, Table 7 illustrates that default assignment can substantially reduce insurer revenue. Under the ME scenario, 40% to 44% of enrollees remain in their new default plans. These plans transfer enough of consumers' OOP costs to the insurance companies that expected revenue per enrollee declines by more than the increase in expected consumer welfare. Hence, the policy exacerbates adverse selection as in Handel (2013). The bottom half of the table shows that expected welfare gains for enrollees in the suspect group exceed the expected reductions in revenue for their chosen insurers, whereas the reverse is true for enrollees in the non-suspect group. Thus, if insurers were to pass on revenue reductions to consumers via higher premiums, less risk protection, or lower quality then consumers in the non-suspect group seem more likely to lose from the policy.

VIII. Summary

We developed a structural model for conducting partial equilibrium evaluations of the equity and efficiency of choice architecture reforms in a differentiated product market where some consumers' choices may not reveal their preferences. Specifically we used administrative and survey data to first identify which consumers appear to make informed and informative decisions. We then estimated separate models of decision making for the informed and misinformed groups and used parameters from the former to assess welfare implications of prospective policies for the latter, implementing Bernheim and Rangel's (2009) proposal for adapting revealed preference analysis to situations of nonstandard decision making. Finally, we reported bounds on welfare that are robust to extreme assumptions about the latent mechanisms underlying consumer inertia and the effects of counterfactual polices on consumer behavior.

The results from any empirical implementation of this approach may depend on the specific rules used to define the subset of choices as misinformed, and any such rules could be controversial. In our context of the Medicare prescription drug markets, however, our main qualitative and quantitative results are robust to using a variety of more inclusive or exclusive rules. As shown in the supplemental appendix (Table A10), this includes (i) focusing exclusively on violations of preference axioms; (ii) focusing exclusively on the MCBS knowledge test; (iii) classifying enrollees as misinformed if they could have reduced expenditures by more than 50%; (iv) extending the sample to include choices for 2006; (v) contracting the sample to exclude mid-year enrollment decisions; (vi) contracting the sample to exclude beneficiaries who get help choosing plans; and (vii) assuming consumers are myopic about their future drug needs when selecting plans.

The results from our policy experiments suggest that CMS's recent proposal to simplify the choice process by reducing the number of drug plans would reduce welfare for the median consumer by up to 16% of consumer expenditures and potentially increase transfers to insurers. In contrast, our results suggest that providing personalized information about the potential savings from switching plans or assigning people to low-cost default plans would benefit the median enrollee. Under the most optimistic scenario, these gains are 11% of consumer expenditures. Comparing the decision support and default assignment policies suggests that defaults have higher downside risk for consumers due to opt-out costs and larger losses in insurer revenue. Both have the potential to erode the consumer welfare gains observed in our partial equilibrium approach. More generally, because both of these policies emphasize cost minimization, insurers may respond by simultaneously lowering plans' costs, quality and risk protection in ways that have ambiguous effects on consumer welfare.

A key challenge for future research is to understand how suppliers adjust to choice architecture policies. Polyakova (2016) takes a step in this direction, finding that eliminating switching costs would lower adverse selection and increase consumer surplus primarily through lower premiums.[46] Similarly Ho, Hogan and Scott Morton (2015) conclude that eliminating consumer inattention would lead to lower levels and lower growth of prescription drug insurance premiums. Another challenge is to determine the conditional probabilities of responding to information treatments for the consumer groups who are informed and misinformed in the baseline. Such information would help to reduce the economic uncertainty in predictions for which consumers would win and lose from information based policies. Finally, our analysis does not embed consumers' decisions about whether to participate in the PDP market. Given the large taxpayer subsidies to all PDP enrollees, such decisions could have significant effects on expected consumer sur-

[46] Decarolis, Polyakova and Ryan (2015) also model plans responses to prospective policy changes. Their focus is on the subsidy structure rather than choice architecture.

plus and total insurer revenues. Similarly, our study holds constant the drugs consumed across plans and under alternative policies, again excluding some potentially welfare-relevant changes from the policies. We consider each of these issues as important avenues for further research.

References

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.

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

Allcott, Hunt and Dmitry Taubinsky. 2015. "Evaluating Behaviorally-Motivated Policy: Experimental Evidence from the Lightbulb Market." *American Economic Review*, 105(8): 2501-2538.

Allcott, Hunt and Judd B. Kessler. 2015. "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons." *NBER Working Paper #21671*.

Ambuehl, Sandro, B. Douglas Bernheim and Annamaria Lusardi. 2014. "The Effect of Financial Education on the Quality of Decision Making". *NBER Working Paper #20618*.

Bernheim, B. Douglas and Antonio Rangel. 2009. "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." *The Quarterly Journal of Economics*, 124(1): 51-104.

Bernheim, B. Douglas, Andrey Fradkin, and Igor Popov. 2015. "The Welfare Economics of Default Options in 401(k) Plans." *American Economic Review,* 105(9): 2798-2837.

Camerer, Colin, Samuel Issacharoff, George Lowenstein, Ted O'Donoghue, and Matthew Rabin. 2003. "Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism"". *University of Pennsylvania Law Review*. 151(3): 1211-1254.

Carlin, Bruce, Simon Gervais, and Gustavo Manso. 2013. "Libertarian Paternalism, Information Production, and Financial Decision Making." *Review of Financial Studies*, 26(9): 2204-28.

Chetty, Raj, John N. Friedman, Soren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen. 2015. "Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark". *Quarterly Journal of Economics*, 129(3): 1141-1219.

Choi, Syngjoo, Shachar Kariv, Wieland Müller and Dan Silverman. 2014. "Who is (More) Rational?" *American Economic Review*, 104(6): 1518-50.

Cohen and Liran Einav. 2007. "Estimating Risk Preferences from Deductible

Choice.”.*American Economic Review*, 97(3): 745-788.

Dube, Jean-Pierre, Gunter J. Hitsch, and Peter E. Rossi. 2010. “State Dependence and Alternative Explanations for Consumer Inertia.” *RAND Journal of Economics*, 41(3): 417–45.

Fang, Hanming, Michael P. Keane, and Dan Silverman. 2008. “Sources of Advantageous Selection: Evidence from the Medigap Insurance Market.” *Journal of Political Economy*, 116(2): 303-350.

Federal Register. 2014. “Policy and Technical Changes to the Medicare Advantage and the Medicare Prescription Drug Benefit Programs for Contract Year 2015 (CMS-4159-P).” January 10: 1917-2073. Accessed online September 1, 2015, http://www.gpo.gov/fdsys/pkg/FR-2014-01-10/html/2013-31497.htm.

Galiani Sebastian, Alvin Murphy, and Juan Pantano. 2015. “Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment”. *American Economic Review*, 105(11): 3385-3415.

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

Handel, Benjamin R. and Jonathan T. Kolstad. 2015. “Health Insurance for “Humans”: Information Frictions, Plan Choice, and Consumer Welfare.” *American Economic Review*, 105(8): 2449-2500.

Handel, Benjamin R., Jonathan T. Kolstad, and Johannes Spinnewijn. 2015. “Information Frictions and Adverse Selection: Policy Interventions in Health Insurance Markets.” *NBER Working Paper #21759.*

Health and Human Services. 2014. “Proposed HHS Notice of Benefit and Payment Parameters for 2016 Fact Sheet.” November 20, 2014. Accessed online November 27, 2015, https://www.cms.gov/cciio/resources/fact-sheets-and-faqs/downloads/fact-sheet-11-20-14.pdf.

Heiss, Florian, Daniel McFadden, and Joachim Winter. 2010. “Mind the Gap! Consumer Perceptions and Choices of Medicare Part D Prescription Drug Plans.” In *Research Findings in the Economics of Aging*, edited by David A. Wise. Chicago, IL: University of Chicago Press.

Ho, Katherine, 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.*

Howell, Benjamin L., Jennifer Wolff, and Bradley Herring. 2012. “Medicare Beneficiary Knowledge of the Part D Program and Its Relationship with Voluntary Enrollment.” *Medicare Research & Review*. 2(4): E1-E21.

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

Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 1997. “Back to Bentham? Explorations of Experienced Utility.” *Quarterly Journal of Economics* 112(2): 375-405.

Keane, Michael P. and Susan Thorp. 2016. “Complex Decision Making: The Roles of

Cognitive Limitations, Cognitive Decline and Aging." In *The Handbook of Population Aging*, edited by J. Piggott and A. Woodland. Elsevier.
Keane, Michael P. and Nada Wasi. 2013. "Comparing alternative models of heterogeneity in consumer choice behavior." *Journal of Applied Econometrics* 28: 1018-1045.
Keane, Michael P. and Kenneth I. Wolpin. 2007. "Exploring the usefulness of a non-random holdout sample for model validation: Welfare effects on female behavior." International Economic Review 48(4): 1351-1378.
Ketcham, Jonathan D., Claudio Lucarelli, Eugenio Miravete and Christopher Roebuck. 2012. "Sinking, Swimming or Learning to Swim in Medicare Part D." *American Economic Review*. 102(6): 2639-2673.
Ketcham, Jonathan D., Claudio Lucarelli, and Christopher Powers. 2015. "Paying Attention or Paying Too Much in Medicare Part D." *American Economic Review*. 105(1): 204-233.
Ketcham, Jonathan D., Nicolai Kuminoff, and Christopher Powers. 2016. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program: Comment." *Forthcoming, American Economic Review*.
Kling, Jeffrey, Sendhil Mullainathan, Eldar Shafir, Lee Vermeulen, and Marian V. Wrobel. 2012. "Comparison Friction: Experimental Evidence from Medicare Drug Plans." *Quarterly Journal of Economics*, 127(1): 199-235.
Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory." *Journal of Political Economy*, 74(2): 132-57.
Leggett, Christopher G. 2002. "Environmental Valuation with Imperfect Information". *Environmental and Resource Economics*. 23: 343-355.
Liebman, Jeffrey and Richard Zeckhauser. 2008. "Simple Humans, Complex Insurance, Subtle Subsidies. *NBER Working Paper 14330*.
Lucarelli, Claudio, Jeffrey Prince, and Kosali Simon. 2012. "The Welfare Impact of Reducing Choice in Medicare Part D: A Comparison of Two Regulation Strategies." *International Economic Review*, 53(4): 1155-1177.
MedPAC, 2006. "Chapter 8. How beneficiaries learned about the drug benefit and made plan choices." In *Report to the Congress, Increasing the Value of Medicare*.
McFadden, Daniel. 2006. "Free Markets and Fettered Consumers." *American Economic Review*, 96 (1): 5–29.
Polyakova, Maria. 2016. "Regulation of Insurance with Adverse Selection and Switching Costs: Evidence from Medicare Part D." *American Economic Journal: Applied Economics*, 8(3): 165-95.
Decarolis, Francesco, Maria Polyakova and Stephen P. Ryan. 2015. "The Welfare Effects of Supply-Side Regulations in Medicare Part D." *NBER Working Paper 21298*.
Small, Kenneth A. and Harvey S. Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica*. 49(1): 105-130.
Stigler, George J. and Gary S. Becker. 1977. "De Gustibus Non Est Disputandum". *American Economic Review*. 67(2): 76-90.
Taubinsky, Dmitry and Alexander Robert Rees-Jones. 2015. "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment." *Working Paper*.

Thaler, Richard H. and Cass R. Sunstein. 2008. *Nudge*. New York: Penguin Books.

Train, Kenneth. 2009. *Discrete Choice Methods with Simulation.* Cambridge University Press.

Tymula, Agnieszka, Lior A. Rosenberg Belmaker, Lital Ruderman, Paul W. Glimcher, and Ifat Levy. 2013. "Like Cognitive Function, Decision Making Across the Life Span Shows Profound Age-Related Changes. *Proceedings of the National Academy of Sciences*, 110(42): 17143-17148.

US Department of Health and Human Services. 2015. "2015 CMS Statistics." Centers for Medicare & Medicaid Services Office of Information Products and Data Analytics CMS Pub. No. 03512, December. Accessed online on August 30, 2016 at https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMS-Statistics-Reference-Booklet/Downloads/2015CMSStatistics.pdf#page=13.

Wiswall, Matthew and Basit Zafar. 2015. "Determinants of College Major Choice: Identification using an Information Experiment". *Review of Economic Studies.* 82(2): 791-824.

Supplemental Appendix: For Online Publication

Table A1—Summary Statistics for the MCBS-Administrative Sample

	Overall	2006	2007	2008	2009	2010
number of enrollees	10,867	1,748	1,975	2,167	2,366	2,611
Medicare Current Beneficiary Survey						
high school graduate (%)	79	77	77	78	80	80
college graduate (%)	22	21	21	22	23	25
income>$25k (%)	55	52	53	53	56	57
currently working (%)	13	14	12	13	12	13
married (%)	55	57	55	54	56	56
has living children (%)	93	93	93	93	93	93
uses the internet (%)	35	33	32	34	37	38
searched for CMS info: internet (%)	27	22	24	27	30	30
searched for CMS info: 1-800-Medicare (%)	17	29	23	17	12	8
makes own health insurance decisions (%)	62	63	62	63	63	62
gets help making insurance decisions (%)	27	27	26	26	26	28
insurance decisions made by proxy (%)	11	10	12	11	11	10
CMS Administrative Data						
mean age	78	77	77	78	78	79
female (%)	63	62	63	63	63	63
white (%)	93	93	92	93	93	94
dementia including Alzheimer's (%)	9	6	8	9	11	12
depression (%)	10	8	9	10	11	11
mean number of drug claims	34	28	34	36	35	35
mean number of available plans	51	43	56	55	51	47
mean number of available brands	22	19	24	23	23	21
has a default plan (%)	65	0	80	83	83	77
switches out of the default plan (%)	11	0	11	16	15	13
active enrollment decisions (%)	46	100	31	33	32	36
mean premium ($)	407	330	355	398	459	493
mean out-of-pocket costs ($)	851	683	847	883	936	907
mean potential savings, ex post ($)	333	435	326	277	316	313

Note: The table reports means for key variables for the sample of Medicare Part D enrollees found in both the MCBS and cost calculator samples in the given year. See the text for details.

TABLE A2— COMPARING MCBS SAMPLE MEANS WITH ADMINISTRATIVE DATA

	Overall	2006	2007	2008	2009	2010
number of enrollees	10,867	1,748	1,975	2,167	2,366	2,611
Medicare Current Beneficiary Survey						
high school graduate (%)	79	77	77	78	80	80
college graduate (%)	22	21	21	22	23	25
income>$25k (%)	55	52	53	53	56	57
currently working (%)	13	14	12	13	12	13
married (%)	55	57	55	54	56	56
has living children (%)	93	93	93	93	93	93
uses the internet (%)	35	33	32	34	37	38
searched for CMS info: internet (%)	27	22	24	27	30	30
searched for CMS info: 1-800-Medicare (%)	17	29	23	17	12	8
makes own health insurance decisions (%)	62	63	62	63	63	62
gets help making insurance decisions (%)	27	27	26	26	26	28
insurance decisions made by proxy (%)	11	10	12	11	11	10
CMS Administrative Data						
mean age	78	77	77	78	78	79
female (%)	63	62	63	63	63	63
white (%)	93	93	92	93	93	94
dementia including Alzheimer's (%)	9	6	8	9	11	12
depression (%)	10	8	9	10	11	11
mean number of drug claims	34	28	34	36	35	35
mean number of available plans	51	43	56	55	51	47
mean number of available brands	22	19	24	23	23	21
has a default plan (%)	65	0	80	83	83	77
switches out of the default plan (%)	11	0	11	16	15	13
active enrollment decisions (%)	46	100	31	33	32	36
mean premium ($)	407	330	355	398	459	493
mean out-of-pocket costs ($)	851	683	847	883	936	907
mean potential savings, ex post ($)	333	435	326	277	316	313

Note: The top half of the table reports means based on enrollees in the merged administrative-MCBS sample that we use for estimation. The bottom half of the table reports means based on a random 20% sample of all individuals who enrolled in Medicare Part D for the entire year. The two data sets differ in that our merged sample includes individuals who enrolled during the middle of the year. We drop these individuals before calculating sample means in order to ensure comparability between the two data sets.

TABLE A3—ASSOCIATION BETWEEN MCBS KNOWLEDGE QUESTION AND MARKET OUTCOMES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Pass Knowledge Test			Pass Knowledge Test			Pass Knowledge Test			Pass Knowledge Test		
	yes	no	p-val: equal means	yes	no	p-val: equal means	yes	no	p-val: equal means	yes	no	p-val: equal means
Conditional on demographics	no	no		yes	yes		no	no		yes	yes	
Active choices only	no	no		no	no		yes	yes		yes	yes	
Number of plan choices	7,560	3,307		7,560	3,307		3,330	1,433		3,330	1,433	
Percent choosing dominated plans	18.5	18.3	0.598	16.5	17.7	0.000	16.3	18.9	0.016	16.7	18	0.000
Mean potential savings ($)	314	363	0.020	282	330	0.000	296	393	0.036	305	373	0.000

The table reports the percentages of enrollees in dominated plans and their mean potential savings, conditional on the accuracy of answers to the MCBS knowledge question. The first six columns report results for all choices. Columns 1-3 show that potential savings is $49 higher for the average enrollee who answers the knowledge question incorrectly ($363 compared to $314) and that this difference is statistically significant at the 2% level. In contrast, there is virtually no difference in the probability of choosing a dominated plan. To isolate the association between knowledge and decision making separately from demographics, we repeat the comparison using residuals from regressions of the percent choosing dominated plans and mean potential savings on indicators for high school degree, college degree, income over $25,000, current working, married, living children, has used the internet to get information on Medicare programs, has used 1-800-Medicare to get information, gets help making health insurance decisions, the number of plans available, female, $70 \leq age \leq 74$, $75 \leq age \leq 79$, $80 \leq age \leq 84$, $85 \leq age$, has dementia, has depression, number of claims, year dummies and region dummies. Columns 4-6 show that after removing the variation in outcomes associated with a linear function of demographics, the percent choosing dominated plans is 1.2 percentage points higher for those answering the knowledge question incorrectly, potential savings is $48 higher, and both differences are statistically significant at the 0.1% level. Columns 7-12 show that the association between knowledge and decision making is stronger if we focus exclusively on active choices. Conditioning on demographics, the probability of actively choosing a dominated plan is 1.3 percentage points higher for the uninformed group and potential savings is $68 higher.

TABLE A4—LOGIT MODELS WITH ADDITIONAL DEMOGRAPHIC INTERACTIONS

	All Choices		Non-Suspect choices		Suspect choices	
expected cost	-0.288	[0.021]***	-0.391	[0.035]***	-0.196	[0.025]***
variance	0.066	[0.176]	-0.389	[0.274]	0.445	[0.172]***
quality (CMS index)	0.053	[0.087]	0.097	[0.114]	-0.051	[0.140]
within-brand switch	-3.306	[0.108]***	-3.246	[0.151]***	-3.397	[0.154]***
between-brand switch	-5.183	[0.093]***	-4.937	[0.126]***	-5.601	[0.139]***
cost x 1{ income > $25k }	0.018	[0.021]	0.033	[0.034]	0.014	[0.025]
cost x 1{ bottom tercile of claims }	-0.173	[0.034]***	-0.196	[0.039]***	-0.089	[0.053]*
cost x 1{ top tercile of claims }	0.084	[0.021]***	0.130	[0.035]***	0.030	[0.024]
cost x 1{ help }	-0.012	[0.022]	-0.011	[0.036]	-0.024	[0.026]
cost x 1{ sought CMS info }	-0.046	[0.023]**	-0.078	[0.033]**	0.035	[0.030]
variance x 1{ college graduate }	0.001	[0.186]	-0.135	[0.236]	0.928	[0.295]***
variance x standardized age	-0.004	[0.086]	-0.046	[0.113]	-0.032	[0.118]
variance x 1{ female }	0.146	[0.174]	-0.111	[0.232]	0.519	[0.233]**
variance x 1{ help }	0.014	[0.176]	0.088	[0.240]	-0.249	[0.274]
variance x 1{ sought CMS info }	-0.226	[0.178]	0.068	[0.241]	-0.604	[0.262]**
quality x 1{ income > $25k }	0.160	[0.092]*	0.181	[0.120]	0.097	[0.148]
quality x 1{ help }	-0.034	[0.094]	-0.102	[0.122]	0.108	[0.152]
quality x 1{ sought CMS info }	0.278	[0.096]***	0.248	[0.123]**	0.302	[0.164]*
switch within brand x standardized age	-0.162	[0.069]**	-0.138	[0.092]	-0.172	[0.103]*
switch within brand x 1{ income > $25k }	-0.368	[0.125]***	-0.335	[0.165]**	-0.363	[0.183]**
switch within brand x 1{ help }	0.321	[0.122]***	0.257	[0.169]	0.462	[0.181]**
switch within brand x 1{ sought CMS info }	0.122	[0.131]	0.258	[0.167]	-0.200	[0.208]
switch within brand x 1{ nonwhite }	-0.811	[0.297]***	-1.214	[0.450]***	-0.578	[0.397]
switch brand x standardized age	-0.121	[0.055]**	-0.168	[0.073]**	0.029	[0.081]
switch brand x 1{ income > $25k }	-0.368	[0.103]***	-0.368	[0.137]***	-0.409	[0.160]**
switch brand x 1{ help }	0.247	[0.103]**	0.222	[0.139]	0.360	[0.157]**
switch brand x 1{ sought CMS info }	0.280	[0.102]***	0.170	[0.134]	0.270	[0.164]*
switch brand x 1{ nonwhite }	-0.794	[0.240]***	-1.369	[0.351]***	-0.108	[0.341]
pseudo R^2	0.66		0.64		0.71	
number of enrollment decisions	9,831		5,465		4,366	
number of enrollees	3,511		2,166		1,675	

Note: The table reports parameter estimates from logit models estimated from data on all choices; from non-suspect choices only; and from suspect choices only. All models include indicators for insurers. Robust standard errors are clustered by enrollee. *,**, and *** indicate that the p-value is less than 0.1, 0.05, and 0.01 respectively.

TABLE A5—RISK PREMIUMS FOR 50-50 BETS FOR NON-SUSPECT CHOICES

Risk premium as a fraction of the bet	Size of Bet
0.01	100
0.11	1,000
0.21	2,000
0.31	3,000
0.39	4,000
0.46	5,000
0.52	6,000
0.58	7,000
0.62	8,000
0.66	9,000
0.69	10,000

To assess the estimates from the logit model for non-suspect choices, we compare its implied risk premiums in a manner comparable with prior literature. Specifically, deriving the risk premium from the logit model as a 1st order approximation to a CARA model yields the following expression for the risk aversion coefficient:

$$\rho_{it} = \frac{-2\beta_{it}/1{,}000{,}000}{\alpha_{it}/100}, \text{ where } U_{ijt} = \alpha_{it}\acute{c}_{ijt} + \beta_{it}\acute{\sigma}^2_{ijt} + \gamma_{it}\acute{q}_{ijt} + \eta_{it}\Delta\acute{B}_{ijt} + \delta_{it}\Delta\acute{P}_{ijt} + \epsilon_{ij1}.$$

The estimates in Table 4 for the reference individual in the non-suspect group yields $\rho = .000217$. Table A4 translates this into a risk premium for various 50-50 bets. These results are broadly consistent with the range of prior results, e.g. as reported in Table 5 of Cohen and Einav (2007). Cohen and Einav find the mean_consumer would be indifferent between a 50-50 bet of winning $100 and losing $76.5, whereas the median consumer is virtually risk neutral. In contrast, our results imply the mean non-suspect consumer is indifferent between a 50-50 bet of winning $100 and losing $98.9 although Cohen and Einav argue that preferences likely differ between their automobile insurance context other contexts like drug insurance. In the health insurance context, Handel (2013) finds that the median individual is indifferent between a 50-50 bet of winning $100 and losing $94.6. In the model preferred by Handel and Kolstad (2015), the mean consumer is indifferent between a 50-50 bet of winning $1,000 and losing $913. This controls for friction and inertia. In comparison, our results imply indifference between winning $1,000 and losing $892.

TABLE A6—VALIDATION OF LOGIT MODELS STRATIFIED BY SUSPECT VS NON-SUSPECT AGAINST ANALOG POOLED MODEL

	In-sample fit (2008)						Out-of-sample fit (2009)								Weighted absolute errors			
	suspect			non-suspect			suspect				non-suspect				in-sample		out-of-sample	
	data	|model error|		data	|model error|		data	|model error|			data	|model error|			|model error|		|model error|	
		s=ns	s		s=ns	ns		s=ns	s	ns		s=ns	s	ns	s=ns	s≠ns	s=ns	s≠ns
Percent of consumers choosing:																		
gap coverage	14	1	0	10	2	2	15	4	3	5	7	1	2	0	2	1	2	1
dominated plan	33	9	8	14	8	7	37	9	8	10	24	1	2	0	9	8	5	4
min cost plan within brand	46	7	5	64	9	12	42	9	4	6	58	3	9	6	9	9	6	5
Mean consumer expenditures ($)																		
premium + OOP	1,385	14	0	1,266	12	0	1,578	29	13	41	1,374	17	35	4	14	0	23	9
overspending on dominated plans	49	17	14	28	13	14	54	26	23	29	17	7	5	7	16	15	16	15
Percent of consumer switching plans	15	4	0	23	3	0	13	6	2	10	22	4	8	1	4	0	5	2

Table A6 reports results from a logit model validation exercise. The purpose is to determine whether the models estimated separately by suspect and non-suspect choices outperform the pooled model, and whether the suspect model better predicts suspect choices than the non-suspect model does and vice versa. For this exercise the estimation sample is 2008 while the prediction sample is 2009. We chose these two years because they incorporate the largest year-to-year change in the choice set in our data—a central aspect to out-of-sample validation methods (Keane and Wolpin 2007). In particular, the number of plans available fell by 10%, although three new brands entered the market, precluding our use of brand indicators in the models. The results show that both in-sample and out-of-sample predictions are closer to the data along a number of policy-relevant outcomes when we base the predictions on separate models for the given type of choice. Blue shading is used to indicate the moments where our preferred model that distinguishes between suspect and non-suspect choices outperforms the pooled model. Red shading indicates moments where the pooled model performs better. We summarize the results in the main text.

TABLE A7—CHARACTERISTICS OF PEOPLE WHO ALWAYS, SOMETIMES, OR NEVER MAKE SUSPECT CHOICES

	Always suspect	Sometimes suspect	Never suspect
number of enrollees	3,311	1,194	4,616
Medicare Current Beneficiary Survey			
high school graduate (%)	77	79	80
college graduate (%)	18	23	26
income>$25k (%)	51	53	59
currently working (%)	12	9	14
married (%)	52	52	58
has living children (%)	92	93	94
uses the internet (%)	28	37	41
searched for CMS info: internet (%)	21	29	33
searched for CMS info: 1-800-Medicare (%)	11	18	16
makes own health insurance decisions (%)	60	61	65
gets help making insurance decisions (%)	27	29	26
insurance decisions made by proxy (%)	13	10	10
CMS Administrative Data			
mean age	79	78	77
female (%)	64	71	59
white (%)	91	96	94
dementia including Alzheimer's (%)	13	10	8
depression (%)	11	13	9
mean number of drug claims	37	39	32
mean number of available plans	52	53	52
mean number of available brands	23	23	23
has a default plan (%)	85	79	78
switches out of the default plan (%)	9	33	12
active enrollment decisions (%)	24	54	34
mean premium ($)	454	406	422
mean out-of-pocket costs ($)	946	1,032	825
mean potential savings, ex post ($)	339	325	282

TABLE A8—LOGIT ESTIMATES FOR PEOPLE WHO ALWAYS, SOMETIMES, AND NEVER MAKE SUSPECT CHOICES

	Always suspect		Sometimes suspect: suspect choice		Sometimes suspect: non-suspect choice		Never suspect	
expected cost	-0.218	[0.024]***	-0.103	[0.041]**	-0.393	[0.068]***	-0.381	[0.033]***
variance	0.491	[0.116]***	1.125	[0.344]***	-1.100	[0.296]***	-0.338	[0.136]**
quality (CMS index)	-0.280	[0.138]**	1.101	[0.306]***	-0.033	[0.233]	0.088	[0.121]
within-brand switch	-3.623	[0.194]***	-2.673	[0.284]***	-2.051	[0.357]***	-3.475	[0.173]***
between-brand switch	-6.101	[0.180]***	-4.283	[0.267]***	-3.353	[0.254]***	-5.253	[0.153]***
cost x 1{ bottom tercile of claims }	-0.130	[0.044]***	-0.054	[0.088]	-0.170	[0.107]	-0.209	[0.043]***
cost x 1{ top tercile of claims }	0.031	[0.027]	-0.023	[0.051]	0.062	[0.081]	0.153	[0.040]***
cost x 1{ sought CMS info }	0.015	[0.028]	0.030	[0.054]	-0.075	[0.065]	-0.064	[0.037]*
quality x 1{ income > $25k }	0.161	[0.168]	-0.166	[0.336]	-0.206	[0.289]	0.262	[0.134]**
quality x 1{ sought CMS info }	0.207	[0.193]	0.337	[0.346]	0.372	[0.328]	0.218	[0.135]
switch within brand x standardized age	-0.070	[0.123]	-0.413	[0.185]**	0.034	[0.178]	-0.185	[0.114]
switch within brand x 1{ income > $25k }	-0.519	[0.225]**	0.149	[0.330]	-0.061	[0.392]	-0.536	[0.200]***
switch within brand x 1{ help }	0.538	[0.216]**	0.431	[0.355]	0.565	[0.336]*	0.159	[0.204]
switch within brand x 1{ sought CMS info }	-0.453	[0.268]*	-0.057	[0.345]	0.117	[0.345]	0.380	[0.201]*
switch within brand x 1{ nonwhite }	-0.351	[0.445]	-0.893	[0.779]	0.473	[1.174]	-1.103	[0.514]**
switch brand x standardized age	0.092	[0.104]	-0.133	[0.140]	0.206	[0.129]	-0.325	[0.086]***
switch brand x 1{ income > $25k }	-0.244	[0.210]	-0.664	[0.279]**	-0.388	[0.309]	-0.444	[0.158]***
switch brand x 1{ help }	0.563	[0.195]***	0.283	[0.311]	0.482	[0.274]*	0.167	[0.166]
switch brand x 1{ sought CMS info }	0.046	[0.222]	0.248	[0.277]	-0.300	[0.287]	0.290	[0.154]*
switch brand x 1{ nonwhite }	0.177	[0.370]	0.106	[0.681]	0.419	[1.377]	-1.291	[0.376]***
pseudo R^2	0.75		0.54		0.46		0.68	
number of enrollment decisions	3,311		560		634		4,614	

Note: The table reports parameter estimates from logit models estimated from data on all choices; from non-suspect choices only; and from suspect choices only. All models include indicators for insurers. Robust standard errors are clustered by enrollee. *,**, and *** indicate that the p-value is less than 0.1, 0.05, and 0.01 respectively.

TABLE A9—CHARACTERISTICS OF WINNERS AND LOSERS FROM THE DECISION SUPPORT TOOL

	Most effective nudge		Least effective nudge	
	enrollees with welfare gains	enrollees with welfare losses	Enrollees with welfare gains	Enrollees with welfare losses
% making suspect choices	42	25	0	0
$\mid oop_{2010} - oop_{2009} \mid$	356	600	324	648

Table A9 shows that enrollees with welfare losses are more likely to come from the non-suspect group and to have larger changes in OOP drug spending between the policy year and the prior year used to determine the minimum cost plan. The text accompanying Figure 3 provides additional details.

TABLE A10—ADDITIONAL ROBUSTNESS CHECKS ON OUR MAIN RESULTS

	Menu Restriction		Decision Support		Default Assignment	
	most effective	least effective	most effective	least effective	most effective	least effective
	A. Baseline results					
Δ expected welfare / enrollee ($)	6	-107	103	28	76	50
% enrollees with expected welfare gain	23	1	81	48	81	83
Δ insurer revenue / enrollee ($)	-8	10	-11	0	-128	-42
	B. Enrollees expect their drug needs to be the same as last year					
Δ expected welfare / enrollee ($)	10	-118	158	62	117	66
% enrollees with expected welfare gain	24	1	92	54	82	83
Δ insurer revenue / enrollee ($)	-7	9	-34	-17	-130	-42
	C. Suspect choices based on dominated plans only					
Δ expected welfare / enrollee ($)	-4	-109	104	21	68	49
% enrollees with expected welfare gain	20	2	82	66	77	83
Δ insurer revenue / enrollee ($)	-5	9	-9	3	-126	-42
	D. Suspect choices based on knowledge test only					
Δ expected welfare / enrollee ($)	-15	-133	126	36	81	55
% enrollees with expected welfare gain	19	0	84	58	79	83
Δ insurer revenue / enrollee ($)	9	11	0	-2	-111	-43
	E. Suspect choices expanded to include potential savings > 50%					
Δ expected welfare / enrollee ($)	26	-91	92	22	89	48
% enrollees with expected welfare gain	30	3	78	43	82	83
Δ insurer revenue / enrollee ($)	-23	9	-20	2	-144	-40
	F. Exclude mid-year enrollment decisions					
Δ expected welfare / enrollee ($)	-24	-107	84	33	36	43
% enrollees with expected welfare gain	21	1	76	50	69	79
Δ insurer revenue / enrollee ($)	-6	13	-8	3	-120	-37
	G. Exclude beneficiaries who get help choosing plans					
Δ expected welfare / enrollee ($)	-2	-97	100	30	70	44
% enrollees with expected welfare gain	23	2	81	50	79	80
Δ insurer revenue / enrollee ($)	-7	11	-12	-2	-125	-41
	H. Include choices for 2006					
Δ expected welfare / enrollee ($)	5	-115	114	33	77	49
% enrollees with expected welfare gain	23	1	82	49	80	83
Δ insurer revenue / enrollee ($)	-12	11	-17	0	-130	-38

Table A10 reports the sensitivity of our main estimates for consumer welfare and insurer revenue to several alternative specifications of our model. The columns match the main policy scenarios summarized in the tables and figures of the previous section. Panels A and B repeat the results from those scenarios for convenience. Panels C, D, and E report the sensitivity of our main results alternative approaches to defining suspect choices under the baseline approach using ex post drug claims to determine plan costs and choice of dominated plan. Panel C ignores the MCBS knowledge question and defines choices as suspect based solely on dominated plans whereas Panel D defines choices as suspect based solely on the MCBS knowledge question. Panel E uses a more inclusive definition of suspect choices based on the union of dominated plan choices, the knowledge question, and being able to reduce expenditures by more than 50%. Moving from C to E increases the set of choices labeled as suspect from 17% to 48%, with the base results in Panel A fitting logically between these figures. Altering how suspect choices are defined has little effect on our main qualitative results. The main reason is that when we classify a greater share of choices as suspect, the difference between θ^s and θ^n declines. More people benefit from certain simplifications to choice architecture, but the average gain among those who benefit is smaller. These effects offset each other in a way that leads to small increases in expected welfare in some scenarios and small decreases in expected welfare in others.

As a next set of robustness checks, we refine the sample in multiple ways. In Panel F we exclude 3,358 choices made by enrollees who first entered the market mid-year. A potential concern in that they may have been forward looking with respect to the following year's drug needs at the time they made their enrollment decisions, especially as they neared or entered the open enrollment period for the following year. Dropping them has little effect on our results. In Panel G we drop 4,044 choices made by enrollees (44% of our sample) who had help choosing a plan or relied on a proxy to choose a plan for them. The logit estimates and subsequent policy implications are similar to the full sample. This suggests that while the research value of having access to better information on how family, friends, and advisors influence decision making is self-evident, in our context of Medicare Part D it does not alter the predicted effects of policy reforms. Finally, in Panel H we include data from the inaugural year of the Medicare Part D program. Again, this only produces minimal changes in our estimates relative to the baseline results in Panel A.