The Costs of Poor Health Plan Choice and Prescriptions for Reform

Saurabh Bhargava[1], George Loewenstein[1], and Shlomo Benartzi[2]
Carnegie Mellon University[1]
Anderson School of Management, University of California, Los Angeles[2]

This Version: March 2016

Summary

We investigate the costs and consequences of plan choice in the ACA. In a first study, we assess the potential financial consequences of inefficient plan choices using pricing data for the several thousand plans listed on federal exchanges. In a second study, we investigate the efficiency with which consumers make plan decisions by observing hypothetical plan choices from experimentally varying menus. We show that the metal labels of the ACA lead to less efficient choices than alternative labels that emphasize considerations of health risk, or even generic plan labels. We argue that a new choice architecture, or structurally redesigned insurance, informed by a deeper understanding of how consumers make complicated health decisions, could substantially improve consumer welfare.

Introduction

Amidst the many polarizing narratives surrounding the Affordable Care Act (ACA), scarce attention has been paid towards understanding whether the millions who enrolled through the exchanges actually made sensible plan choices. Enrollees, many of whom had no prior experience with health insurance, confronted an average of 47 plans which varied by insurance provider, medical coverage, and cost.[1] Indeed, a key selling point of the ACA, beyond reducing the ranks of the uninsured, was its promise to increase the options available to consumers and encourage beneficial insurer competition through the creation of health plan exchanges. Recent research, however, has challenged the theoretical rationale for this expansion of choice by raising concerns about the ability of consumers to make informed decisions between the choices they encounter.

Evaluating the rationality of consumer plan choices, and the financial consequences of such choices, is difficult. The traditional approach of economics has been to evaluate an insurance choice based on whether it minimizes expected health costs, adjusting for plausible levels of tolerance for financial risk. Such analyses tend to ignore, however, other considerations that matter to people, such as the doctors included in-network, the breadth of medical services covered, and the insurer's reputation for handling claims expediently. Researchers have recently attempted to circumvent this problem with a variety of different strategies.

---

[1] As reported by ASPE in a June 2014 report: https://aspe.hhs.gov/sites/default/files/pdf/76896/2014MktPlacePremBrf.pdf

Prior Research Evaluating Health Plan Choices

One can organize existing research evaluating the efficiency of consumer insurance decisions as broadly reflecting one of three approaches: (i) analyses of prescription drug plan choices from Medicare Part D, (ii) experiments involving hypothetical plan choices, and (iii) analyses of administrative data on spending and choices of employees.

A first set of studies (1-7) reveals that many seniors chose prescription drug plans from Medicare Part D that led to higher spending than available alternatives, even after adjusting for health and financial risk tolerance. In one influential example, Abaluck and Gruber (2011) found that elderly enrollees made financially sub-optimal plan choices that over-weighted plan premiums relative to estimated out-of-pocket expenses. In another study, Kling et al. (2012) found that providing enrollees with a letter highlighting the personalized cost-savings from switching plans led 28% of recipients to switch, relative to only 17% in a comparison group who received a control mailing.

In a second set of studies, hypothetical plan choices among representative samples show that sub-optimal plan decisions are not limited to the elderly and the well-chronicled complexity of plan choices in Medicare Part D (8-10). The experimental paradigm involves asking subjects to select plans from stylized settings in which the content and presentation of the choice interface is experimentally manipulated to understand how consumers make plan decisions. Johnson et al. (2013), for example, presented subjects, all of whom passed an initial insurance comprehension test, with information about future health needs and a small menu of plan options. Apart from a sample of MBA students, a majority of subjects, even those financially incentivized to choose efficiently, chose financially sub-optimal plans unless aided by cost-calculators or customized plan defaults.

Skeptics might continue to insist that working-age consumers, when confronted with significant financial stakes and reasonably standardized plan menus, should fare better than the elderly or experimental subjects. At the least, such consumers should be expected to improve their plan choices over time. A final set of studies addresses these claims through analyses of administrative data on plan choices and health spending of employees from firms with employer-sponsored health coverage (11-14). One particularly clear litmus test for assessing the quality of plan choices was described in a recent study which examined employee decisions from an unusually large and highly standardized plan menu (11, hereafter BLS). The firm permitted employees to "build" their own health plan through the choice of four "cost-sharing" features: deductible, copayment, out-of-pocket spending limit, and coinsurance. Beyond these differences in cost-sharing and price, the 48 available plans were identical – e.g., they were offered by the same insurer and covered the same network of providers. However, because of how these plans were priced, nearly every one of the 36 lower deductible plan options was more costly to the employee than the otherwise identical high deductible plan, regardless of the employee's medical use, (e.g., employees paid $528 to reduce their deductible from $1,000 to $750, all else equal, despite a maximum potential savings of $250).[2]

---

[2] Of the 36 plans available to employees with deductibles less than $1,000, 35 were financially dominated in this manner, while the remaining plan fell $20 short of domination. After adjusting for the differential tax treatment of premiums and out-of-pocket spending, the authors estimate that 30 of the 36 low-deductible plans were dominated.

How did employees fare when confronted with this standardized but financially consequential menu? The researchers found that a majority of employees chose financially dominated plans, which resulted in excess spending equivalent to 42% of annual premiums. Moreover, lower income employees were especially likely to choose dominated plans, and employees who enrolled in dominated plans in one year were unlikely to switch into alternative plans in the following year. This behavior cannot be rationalized by any set of expectations regarding one's health or willingness to take on financial risk.

Why would employees pay significantly more in premiums, which, with complete certainty, could not be offset by reductions in out-of-pocket expenses? Through a series of follow-up experiments, the authors investigated whether the choices reflected the potentially debilitating search costs associated with an expansive menu, explicit preferences for low-deductibles, however expensive (e.g., due to the convenience of predictable budgets or a psychological distaste for unpredictable out-of-pocket spending), or confusion about how to compare plan values. The experiments suggest that plan choices emerged from a lack of understanding about how to translate basic cost-sharing concepts, such as a deductible, into a projection of total health costs. When subjects were shown a simple dominated menu featuring four plans that differed just in their deductible and price, a majority of subjects still chose a dominated plan. Yet when subjects were shown menus in which cost-sharing definitions, plan trade-offs, both financial and non-financial, were made explicit, the demand for dominated plans virtually disappeared.

Plan Choice in the ACA
Recognizing the difficulty consumers have had in navigating health plan choice, the ACA instituted a number of measures to guide consumers. Such measures included a thoughtfully designed exchange interface, instructional videos and online decision-aides, customer service representatives available by phone, and, perhaps most importantly a requirement that plans must conform to one of four standardized cost-sharing "tiers." Plans within each tier were, on average, expected to cover a pre-specified share of health expenses ranging from 60 percent for the lowest tier to 90 percent for the highest tier. Tiers were also given metal labels—bronze, silver, gold, platinum—following the model of the Massachusetts health exchange.[3] Finally, all plans were required to cover a set of basic medical services so that all enrollees, regardless of their plan choice, could avail themselves of services deemed essential.

Against this backdrop—consumers challenged by the complexity of insurance choice and a choice architecture carefully crafted to shepherd consumers through such complexity—we sought to understand the financial consequences of plan choice for ACA enrollees and the likelihood that such enrollees chose judiciously. In the first study presented below, we analyze pricing data from several thousand plans listed on the first year of the federal ACA exchange to assess the approximate spending consequences of choosing one plan tier over another, given different levels of medical need. The exercise, which abstracts away from the many other factors that shape choice (e.g., differences in insurer reputation, or network configurations), aspires to capture the average sensitivity of medical

---

[3] A fifth tier of "catastrophic" plans, with very high deductibles and low premiums, were available in some regions to individuals under 30 or those who qualified for an insurance exemption.

spending to plan tier choice, for a fixed set of health risks, using actual plan prices across the numerous pricing regions of the exchanges.

Given the absence of administrative data on actual choices and medical claims, in a second study we sought to experimentally test whether experimental subjects differing in their medical spending needs are likely to make efficient plan choices when confronted with a stylized environment resembling that of the exchange. In a baseline condition, subjects chose from a set of composite health plans constructed, and labeled, to resemble the cost-sharing tiers of the exchange. In additional conditions, we present subjects with more transparent plan labels or even plan recommendations. Our primary measure of choice efficiency is whether plan choices minimize expected health costs estimated from elicited beliefs regarding medical needs. Given that other considerations, such as aversion to financial risk or liquidity constraints, might shape plan choice, we additionally assess efficiency by observing the consistency of choices under the metal labels and more informative labeling regimes. We close by discussing the implications of research on plan choices, and findings from Behavioral Economics more generally, for the success of health reform.

Study 1: The Financial Consequences of Plan Choice in the Exchange

Overview
To assess the financial consequences of plan tier choice on the exchange, for given levels of medical need, we examined average differences in projected spending associated with each cost-sharing tier for consumers based on their location and age. Using data on actual plans from year-one of the exchange, we estimated the age by region specific health care spending associated with plans representative of each tier choice across three categories of medical care (low, medium, and high). We then calculated the average "overspending" associated with each representative plan relative to the cheapest plan in expectation. For simplicity, we focused the pricing analysis on couples with no children.

Data Sources
From the original 78,522 plans listed on the federal and state sponsored exchanges, we generated our sample after excluding data from 6,282 Catastrophic plans, 1,792 plans for which we were unable to determine premiums, and 30,563 plans from counties which failed to offer at least one plan in each of the four tiers of interest—Bronze, Silver, Gold, and Platinum.[4] We analyzed the remaining 39,885 plans across 220 distinct insurance pricing regions.[5] For each plan, we recorded plan prices and cost-sharing features for four each of the four age categories (30 years, 40 years, 50 years, 60 years) for which data are provided. While our health spending calculations only require knowledge of premiums and cost-sharing features for the "low" and "high" utilization categories, to calculate the spending associated

---

[4] Additional analysis of a sample not restricted by the requirement that each region include plans reflecting all four tiers produced qualitatively similar results to those reported here and is available upon request from the authors.
[5] We extracted the data from the QHP landscape files at healthcare.gov (accessed in July 2014). The data represents plans from 220 insurance rating areas – geographic regions within which insurance firms must price plans equivalently, conditioned on age, and smoking status.

with the "middle" category of utilization, we estimated average, age-specific, utilization rates (i.e., primary care and specialist visits) from employee data reported in BLS.

Research Design

To estimate the financial consequences of plan tier choice, we first constructed a "composite" plan that reflected the pricing and cost-sharing features of actual plans available within a tier for each of the 4 age categories x 220 insurance pricing regions in our data. We constructed the composite plans by taking the un-weighted average premium (assuming no tobacco surcharge), out-of-pocket maximum, and deductible for a married couple without children for every age x region for all plans within a tier (prices in the federal exchange are permitted to vary by consumer age, in decades, and across insurance rating regions). For analytic tractability, we assume that each composite plan features the modal copayment and coinsurance for all plans within that tier (reported in Table 1 of the Appendix).

Having constructed a set of composite plan features corresponding to each tier, we then calculated the total health care spending implied by each composite tier-plan under three scenarios of medical need. For low care, we assumed that the enrolled couple required no medical visits, while for medium care, we assumed the each enrollee required the same amount of care (i.e. primary care and specialist visits) as the typical employee analyzed in BLS.[6] For this middle category, we further assumed a cost of $200 for each primary care visit and $350 for each specialist visit. For the high care scenario, we assumed medical costs that exceeded each composite plan's out-of-pocket maximum.

Finally, for the three categories of medical need and every age-by-region cell, we ranked each composite plan with respect to total projected costs, and calculated the overspending associated with the non-first-best tiers. We then averaged these spending differences by age, region, and plan tier rank, and report them in Figure 1.[7] The figure presents average rates of ex-post overspending by rank, in both dollar terms and as a percent share of the average premium of the plan associated with the lowest amount of spending. The standard errors reflect heterogeneity across age and region and are clustered at the state level to account for proximal counties which feature identical plan offerings.

Results

Figure 1 indicates that a couple, requiring little medical care, who chose the second best of the four plan tiers, would spend $1,662 (95% confidence interval [CI], $1494 to $1831) in excess of the spending associated with the cheapest plan tier. If the couple had instead chosen the most expensive of the four tiers, the couple would have spent an excess of $4,708 (95% CI, $4,125 to $5,291)—equivalent to 61% of the annual premium of the cheapest tier. For the typical couple requiring a high amount of care, the most inefficient tier choice would lead to excess spending equivalent to 59% of the annual premium of the least expensive tier. While the figure averages across enrollee ages, these estimates of potential overspending are highly stable across age groups.

---

[6] Specifically, for the middle care category we assume that the average number of visits per enrollee (PCP, Specialist) is (5.7, 5.8) for 30 year olds, (5.9, 7.5) for 40 year olds, (6.3, 9.5) for 50 year olds, and (7.1, 11.7) for 60 year olds. These utilization figures were doubled since our exercise involves couples.

[7] The use of un-weighted averages produces approximate estimates of spending differences for typical age-by-region cells.

Figure 1. Potential Health Overspending by Choice of Plan Tier on the Exchange
Estimated for Couples (averaged across age x county) by Medical Care Category

Panel A. Annual Overspending by Plan Tier Rank

Panel B. Annual Overspending as a Percent of Best Tier Premium
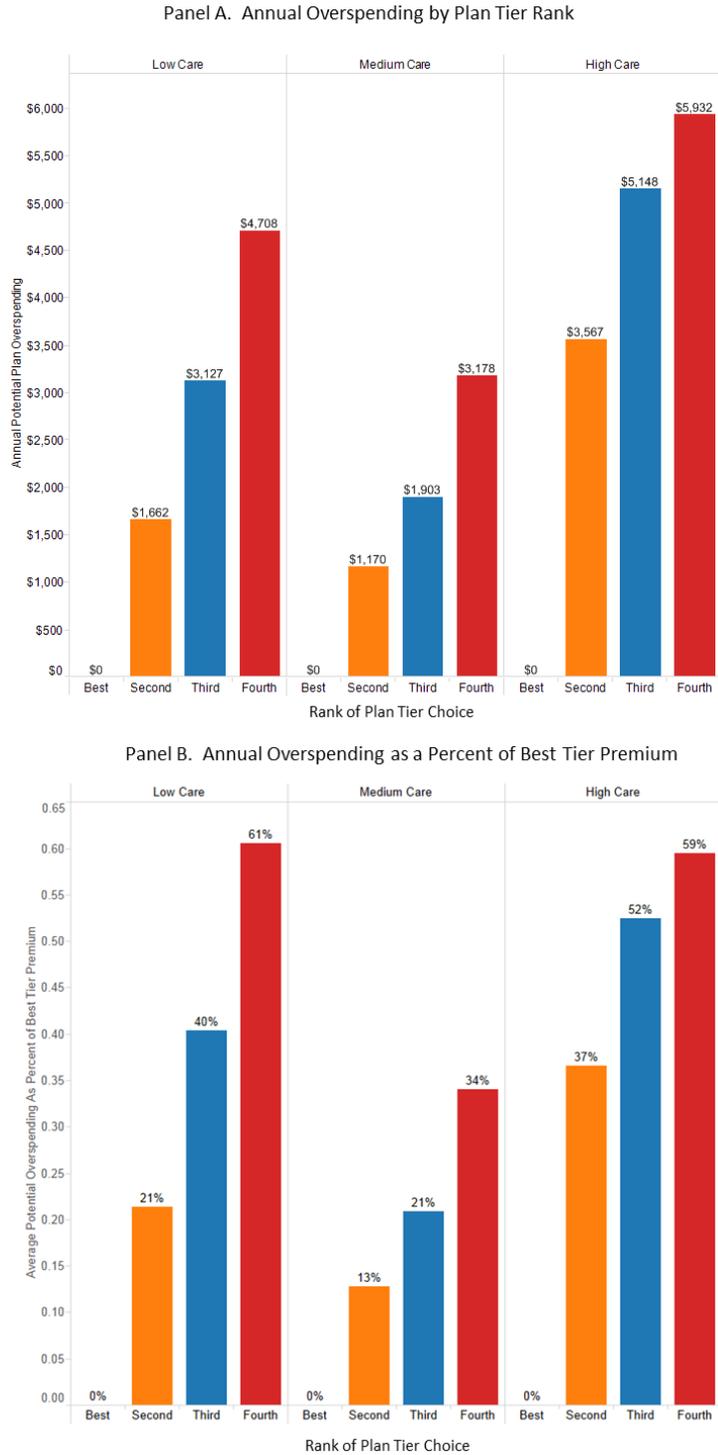
Notes: Figure reports the estimated average difference in absolute (Panel A) and relative (Panel B) total health spending for childless couples by choice of ACA plan tier and level of realized care. Composite plans representing each tier were calculated by averaging plan features for all plans within a tier and age group by rating area. Standard errors reported in the paper are clustered by state to account for identical plan offerings across many proximal regions.

An important feature of exchange choice is the premium tax credit for which most enrollees (i.e., those with income between 100 to 400% of FPL) are eligible. The presence of such refundable credits do not influence our estimates of overspending since these subsidies, while calculated from a baseline plan, can be applied to any plan chosen (such subsidies would significantly raise the share of premium equivalences depicted in Panel B of the figure). However, for the smaller sub-set of individuals additionally eligible for cost-sharing subsidies, (i.e., eligible enrollees earning at or less than 250% of the Federal Poverty Level), potential overspending may differ from the reported figures, since such subsidies are contingent on the selection of Silver Plans.

Study 2: Plan Choice in an Exchange-Like Environment

Overview
Given the high potential cost of choosing a plan that does not match one's medical need, Study 2 explored the likelihood of efficient choice on the exchange through experiments in which subjects were asked to choose plans from hypothetical menus resembling those in the exchange. We gauge the efficiency of decisions by assessing whether the observed choices minimize total health spending, based on estimates inferred from reported medical need, and, given the potential importance of factors such as risk or liquidity, by evaluating the consistency of choice across menus in which we vary the transparency of plan labels or include explicit plan recommendations.

Research Design
Subjects (n=304) were U.S. adults over the age of 25, recruited from the Qualtrics Survey Panel. The less than 10 minute survey was hosted on Qualtrics. Subjects were diverse in gender (male: 38%, female: 62%), age (26 to 35: 18%, 36 to 45: 23%, 46 to 55: 25%, > 55: 33%), yearly income (<= $30,000: 34%, $30,001 to $50,000: 24%, $50,001 to $80,000: 25%, $80,001 to $120,000: 10%, > $120,000: 7%) and education (college: 40%, some college: 39%, high school: 19%, less than high school: 2%).

Subjects were first asked to report background demographics and then asked a series of questions about their overall health and expected medical utilization over the subsequent year.[8] Finally, we asked subjects to make a hypothetical choice of an insurance plan, covering just themselves for next year, from a menu reflecting plans whose features and prices represented age-specific averages of plans on the exchange. We restricted each menu to three plans whose prices and plan features were constructed, following the procedure discussed in Study 1, to represent the Bronze, Silver, and Platinum tiers. For simplicity, we averaged plan prices only by age group, rather than age x region such that participants were presented with one of the plan menus displayed in Appendix Table 1. Subjects were told that all plans covered the same essential benefits (corresponding to the basic services covered in the ACA) which were displayed.

---

[8] We elicited expectations about health utilization in two ways. First we asked subjects to indicate their anticipated use of medical care from a set of three options corresponding to "little or no" care, "a moderate amount of" care, or "a great deal of" care. Second, we asked subjects to project how many times they anticipated visiting a primary care doctor, specialist, and hospital.

Plan choice proceeded in three steps which, for the baseline condition, we depict in Appendix Figure 1. After the choice paradigm was introduced (Panel A), subjects were given the option of filtering plans by tier, or viewing all plans (Panel B). Following this decision, subjects were shown the prices and cost-sharing features associated with the indicated plan(s) (Panel C). Subjects subsequently had the option of choosing a plan, or, in the case all plans hadn't been displayed, viewing all available plans. To test the sensitivity of plan choice to alternative choice environments, subjects were randomized into one of four treatment arms, each associated with a different set of plan labels: (1) the metal labels of the exchange; (2) generic plan labels ("Plan A," "Plan B," and "Plan C"); (3) medical use labels ("High Use," "Medium Use," and "Low Use"); and (4) generic labels accompanied by a recommendation indicating the least expensive plan estimated from self-reported medical need. Because the opportunity to filter plans upfront wasn't meaningful for the generic label conditions, this step was bypassed so that subjects in these conditions proceeded straight to a menu featuring all plan options.

We assess the efficiency of plan choice in two ways. First, we document whether the selected plan minimized total spending based on the two measures of anticipated utilization collected from the survey (i.e., the self-reported estimates of medical visits, and the categorical expectation of medical care). To calculate the health spending associated with plan choices, we assumed the same per-visit costs as in the first study and additionally assumed a cost of $2,000 for each hospital visit. Second, to evaluate choice efficiency associated with the metal labels, we compare choices prompted by the metal labels with those produced from more informative labels, under the presumption that a property of optimal choice is its consistency across settings of high and low clarity.[9]

Results

Figure 2 depicts the likelihood that subjects minimized expected health spending with their choice of plan under different labeling regimes. Based on expenses estimated from anticipated medical visits, as depicted in Panel A, only 33% of respondents chose the plan which would have been cheapest for them, while 43% chose plans providing coverage exceeding anticipated need ("Over-insured"), while 24% chose plans providing too little coverage ("Under-insured"). Focusing simply on expected costs, the chosen plans implied average overspending of approximately $888 (95% CI, $681 to $1095), or 16% of a typical premium across all subjects. Those not choosing to minimize their total spending instead chose plans resulting in average overspending equivalent to 24% of the typical premium.[10]

Panel B of the figure shows that metal labels did not significantly improve choice efficiency relative to generic plan labels ($888 compared to $794 in average overspending across all respondents, p = .52). However, when choosing plans with labels emphasizing considerations of medical use, ($559 average overspending; 95% CI, $378 to $740), and plan recommendations ($591 average overspending; 95% CI, $401 to $781) participant choices were more financially efficient than those yielded from metal labels

---

[9] See BLS for a detailed discussion of using this type of criterion for evaluating choice efficiency.

[10] If we instead evaluated choice by whether respondents sorted themselves into the plan tier corresponding to their self-reported category of expected medical need (e.g., those anticipating "little to no care" sorted into Bronze Plans), 44% of respondents chose a plan commensurate with their health risk.

(both, p < 0.05).[11] When we calculated health expenses only for those respondents who chose inefficiently, the metal labels resulted in an average of $1324 (95% CI, $1080 to $1568) in excess spending which represents 24% of the typical respondent's annual premium.

Figure 2. Demand for Insurance and Overspending in Marketplace Experiment
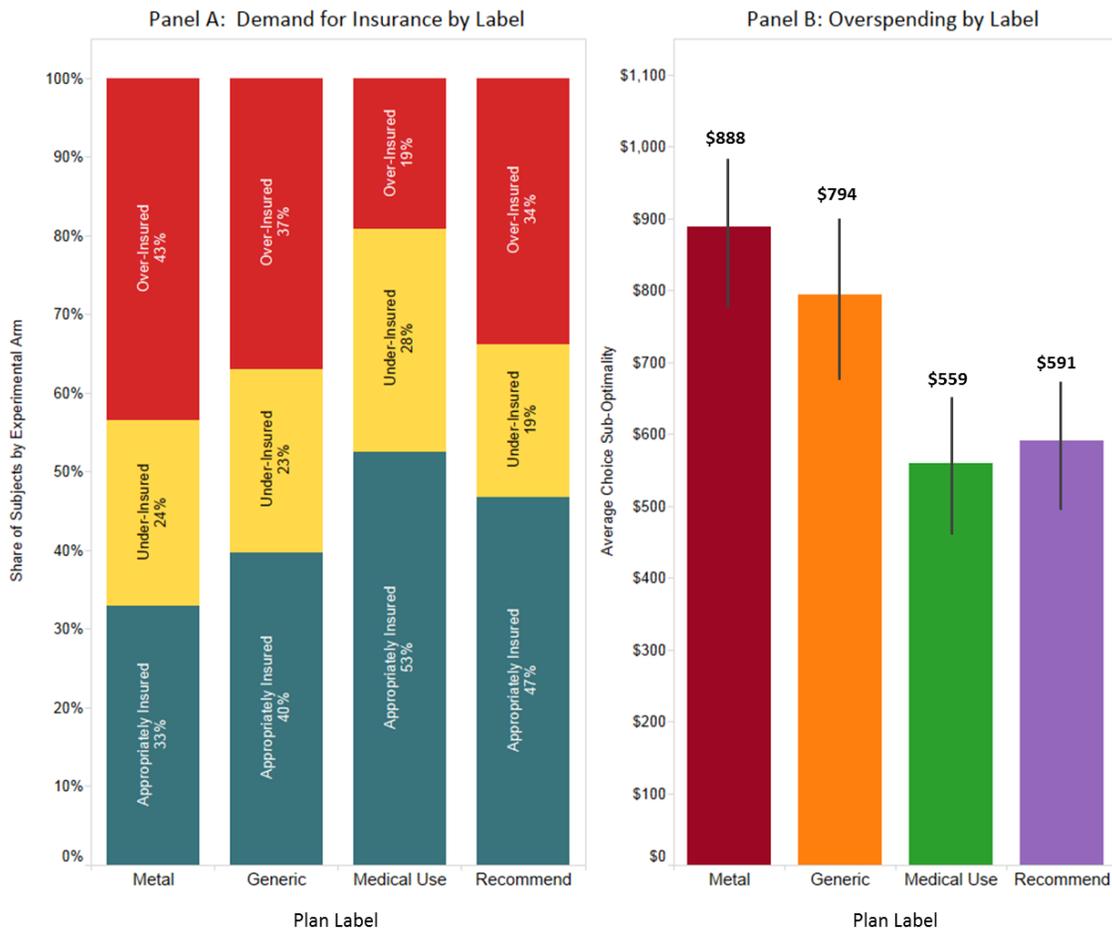
Notes: Figure reports results of a between-subject experiment (N = 304) in which subjects were presented with a hypothetical choice between insurance plans representative of each cost-sharing tier in the exchange. Panel A reports subject demand for insurance relative to the least expensive plan tier implied by health expenses estimated from self-reported expectations of medical utilization. Each column represents a different labeling condition in the experiment. Panel B reports average overspending relative to each individual's cost-minimizing plan, by plan label condition. Bars indicate an interval of +/- 1 standard error.

---

[11] While 47% of respondents in the plan recommendation condition chose optimally as determined by health expenses calculated from the respondent's expected number of medical visits, if optimal choice is instead evaluated by whether respondents sorted themselves in accordance with their self-reported categorical need (the dimension along which plan recommendations were made in the study), 64% of respondents appear to have chosen efficiently.

Discussion and Implications for Policy

Across the two studies, our findings suggest that, despite the sensitivity of total health spending to the choice of plan tier, individuals are not likely to choose cost-minimizing plans, even when aided by the metal labels of the exchange. The economic magnitude of decisions is significant. Study 1 indicates that an individual's choice of the second best available tier— an efficacy of decision-making that exceeds that observed in other settings where individuals make choices that appear no better than chance— would cause enrollees to overspend an equivalent of 13% to 37% of plan premiums, while the worst choice of tier could lead to overspending amounting to over 60% of plan premiums (Figure 1, Panel B). While the pricing analysis involves a highly simplified representation of choice, the estimates imply the financial significance of plan tier choice for consumers, of all ages, across the hundreds of pricing regions serviced by the federal exchanges.

The two-thirds of subjects in Study 2 who chose a plan incommensurate with their medical expectations overspent by $1324, equivalent to 24% of the typical annual premium.[12] Although the hypothetical decisions examined made were not financially consequential, unlike consumers making real choices, there is reason to believe that this paradigm does capture authentic behavior. BLS reported a striking similarity in the distribution of plan choices across the experimental setting and the actual choices of several thousand firm employees. Moreover, Johnson et al. (2013) find no significant difference between hypothetical plan choices in conditions with and without incentives tied to choice efficiency.

One might contend that the behavior we observe might reflect informed considerations of financial risk for those who appear to "over-insure," or the simple inability to pay higher premiums ("liquidity constraints," in economic parlance) on the part of those who appear to "under-insure." There are several reasons, however, why the choices we document are not likely to be rationalized by such factors. First, one mark of whether plan choices reflect a coherent and informed set of preferences is the consistency of such choice under menus in which the presentation of choices is varied. That subjects are significantly more likely to minimize spending under more transparent labeling regimes (e.g., medical usage), or even plan recommendations, suggests that the behavior in the baseline condition does not reflect clear preferences. Second, to understand if plan choices can be explained by a preference for risk-taking, we asked respondents about their general willingness to take on financial risks (on a scale of 1 to 10). After controlling for self-reported health and income, our elicited measure of financial risk-taking does not predict either demand for "over-insurance" (b = -0.02, p = 0.33) or "under-insurance" (b = -0.005, p = 0.81).[13] Finally, the recent research discussed above implies that the behavior of subjects in our studies is more likely reflective of deficits in health insurance literacy than informed preferences for avoiding risk, liquidity constraints, or other factors.

---

[12] This reflects a weighted average of subjects demanding too much insurance ($1102 excess spending) and too little insurance ($1730 excess spending).

[13] We separately model demand for over/under insurance as defined in Figure 2 with a linear probability model in which demand is a function of a participant's self-reported risk, after controlling flexibly for self-reported health and income. We note that this estimate is not very precise given the small experimental sample in the metal condition. A similar model indicates that risk does not linearly predict ln(overspending), b = -0.015, p = 0.75.

As a rough estimate of how our findings might translate to consumer welfare in the ACA, if two-thirds of the 8 million enrollees in the first-year of the ACA made choices resulting in average overspending of $1,324, the result would be nearly $7.1 billion of excess spending each year. The policy implications of consumers engaging in sub-optimal decisions extend beyond direct welfare consequences. Recently, some economists have argued that in markets with a significant share of not fully informed consumers, firms may be subject to less competitive pressures to reduce prices and improve quality, and may even compete over confusing consumers into purchasing lucrative suboptimal plans (15-18). In the same way that certain welfare-reducing product attributes such as credit card fees or mortgage closing costs may persist in competitive equilibrium given some share of naïve consumers, insurers may not be driven to eliminate the complicated, profit-generating, features of health insurance products.[14]

Why are consumers prone to making sub-optimal plan choices despite metal labels designed to facilitate decisions? Astute choice of plan tiers requires careful consideration of one's expected medical expenses. Those anticipating a modest need for medical care would spend less, on average, by selecting a plan in a low cost-sharing tier, while those anticipating substantial care should benefit, on average, by choosing a plan with greater cost-sharing. An explanation consistent with our findings is that our experimental subjects failed to interpret the metal labels as signaling a gradation in the degree of cost-sharing associated with each tier, and instead interpreted them as implying differences in the *quality* of medical care or access to such care. This explanation is supported by the second study in that labels designed to encourage consumers to choose plans based on expected use led to choices closer to the cost-minimization criterion. The possibility that consumers rely on metal labels as a global measure of quality is also suggested by a survey which found that, among survey respondents deemed to be below the median in mathematical ability, gold plans were preferred to other plans regardless of the underlying plan features (20).

Collectively, our evidence, and the research upon which it builds, suggests that the psychological features of the implementation of health reform warrant as much attention as the economic structure of the reform. Practically, the second study suggests that the adoption of labels which imply gradations in expected medical use—a more logical dimension for plan comparison than the gradation in the quality of medical care implied by the metal labels—would improve the quality of enrollee choices, as would specific plan recommendations.

However, since replacing metal labels is impractical, at least in the near-term, Behavioral Economics offers several alternative strategies through which to encourage more efficient choices. These strategies include the use of customized defaults or highly personalized calculators, education through real-time, scenario-based, examples, or the simplified presentation of the trade-offs most relevant for plan comparison (see 8 and 21 for review). To the credit of policy-makers, the interface of the ACA has evolved since its inaugural year. This past year, a "beta" version of a health cost calculator, using

---

[14] Other work supports this thesis in highlighting how the growing complexity of certain financial products is correlated with higher bank profits and lower consumer welfare suggesting potentially strategic motives (19).

consumer inputs of expected utilization, was introduced on the federal exchange. While the move towards projections of estimated yearly costs is commendable, we suspect that presenting such information as the *fourth* most prominently displayed figure in a plan description (after monthly premium, deductible, out- of-pocket maximum) has as much potential to confuse enrollees as it does help them. Ultimately, while a smarter choice interface may improve plan choice, a superior path towards reform may be to move beyond redesigning the choice environment to fundamentally simplifying the insurance incentives whose complexity continues to overwhelm enrollees (22,23).

REFERENCES

(1) Abaluck, J., and Gruber, J. (2011). "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program," *American Economic Review*, Vol. 101, No. 4, pp. 1180-1210.

(2) Ho, K., Hogan, J., and Morton, FS. (2015). "The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program," NBER Working Paper No. 21028.

(3) Kling, J., Mullainathan, S., Shafir, E., Vermeulen, L., and Wrobel, M. (2012). "Comparison Friction: Experimental Evidence from Medicare Drug Plans," *Quarterly Journal of Economics*, Vol. 127, No. 1, pp. 199-235.

(4) Ericson, Keith M. (2014). "Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange," *American Economic Journal: Economic Policy,* Vol. 6 No. 1, pp. 38-64.

(5) Heiss, Florian, McFadden, Daniel, and Winter, Joachim (2010). "Mind the Gap! Consumer Perceptions and Choices of Medicare Part D Prescription Drug Plans." In *Research Findings in the Economics of Aging*, edited David A. Wise, pp. 413-481, Chicago: University of Chicago Press.

(6) Ketcham, Jonathan D., Claudio Lucarelli, and Powers, Christopher A. (2012). "Sinking, Swimming, or Learning to Swim in Medicare Part D." *American Economic Review*, Vol. 102, No. 6, pp. 2639-2673.

(7) Ketcham, Jonathan D., Claudio Lucarelli, and Powers, Christopher A. (2015). "Paying Attention or Paying Too Much in Medicare Part D." *American Economic Review*, Vol. 105, No. 1, pp. 204–33.

(8) Johnson E., Hassin R., Baker T., Bajger A., and Treuer G. (2013). "Can Consumers Make Affordable Care Affordable? The Value of Choice Architecture," *PLoS ONE*, Vol. 8, No. 12.

(9) Kairies, Nadja, Kokot, Johanna, Vomhof, Markus, and Wessling, Jens (2014). "How Do Consumers Choose Health Insurance? An Experiment on Heterogeneity in Attribute Tastes and Risk Preferences," Ruhr Economic Paper No. 537. http://ssrn.com/abstract=2579733.

(10) Schram, A., & Sonnemans, J. (2011). "How individuals choose health insurance: An experimental analysis," *European Economic Review*, Vol. 55, No. 6, pp. 799-819.

(11) Bhargava, S., Loewenstein, G., and Sydnor, J. (2015). "Do Individuals Make Sensible Health Insurance Decisions?" NBER Working Paper No. 21160.

(12) Handel, B., (2013). "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *American Economic Review*, Vol. 103, No. 7, pp. 2643-2682.

(13) Strombom, Bruce A., Thomas C. Buchmueller, and Paul J. Feldstein, (2002). "Switching Costs, Price Sensitivity and Health Plan Choice." *Journal of Health Economics*, Vol. 21, No. 1, pp. 89–116.

(14) Handel, Ben R., and Kolstad, Jonathan T. (2015). "Health Insurance for 'Humans': Information Frictions, Plan Choice, and Consumer Welfare." *American Economic Review*, Vol. 105, No. 8, pp. 2449–2500.

(15) Sandroni, A., and Squintani, F. (2007). "Overconfidence, Insurance, and Paternalism," *American Economic Review*, Vol. 97, No. 5, pp. 1994-2004.

(16) Gabaix, X., and Laibson, D. (2006). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *The Quarterly Journal of Economics*, Vol. 121, No. 2, pp. 505-540

(17) DellaVigna, S., and Malmendier, U. (2004). "Contract Design and Self-Control: Theory and Evidence," *The Quarterly Journal of Economics*, Vol. 119, No. 2, pp. 353-402.

(18) Heidhues, P., and Kőszegi, B. (2014). "Using Information about Naivete to Price Discriminate," mimeo, ESMT.

(19) Celerier, Claire, and Boris Vallee (2015). "Catering to Investors through Product Complexity," Harvard Business School, Working Paper 16-050.

(20) Ubel, Peter A., Comerford, David A., and Johnson, Eric (2015). "Healthcare.gov 3.0 — Behavioral Economics and Insurance Exchanges," *New England Journal Medicine*, Vol. 372, No. 8, pp. 695-698.

(21) Handel, B., and Kolstad, J. (2015). "Getting the Most from Marketplaces: Smart Policies on Health Insurance Choice," The Hamilton Project: Discussion Paper 2015-08.

(22) Loewenstein, G., et al. (2013). "Consumers' Misunderstanding of Health Insurance." *Journal of Health Economics*, Vol. 32, pp. 850-862.

(23) Bhargava, S., and Loewenstein, G., (2015). "Behavioral Economics and Public Policy 102: Beyond Nudging." *American Economic Review*, *Papers & Proceedings*, Vol. 105, No. 5, pp. 396-401.

APPENDIX – TABLES AND FIGURES

Appendix Table 1.
Plan Features by Tier and Enrollee Age for Study 2

Plan Tier	Deductible	Premium by Enrollee Age				MOOP	Copayments or Coinsurance		
		25 to 35	36 to 45	46 to 55	56 to 65		PCP	Specialist	Hospital
Bronze	$5,094	$238	$267	$374	$567	$6,268	0%	0%	0%
Silver	$2,911	$276	$311	$434	$659	$5,723	$30	$50	0%
Platinum	$344	$345	$389	$543	$825	$1,866	$10	$20	0%

Appendix Figure 1
Study 2 - Experimental Interface for Metal Label Condition

Panel A – Choice Introduction

Now imagine that you must enroll in a health insurance plan for next year.

In a moment we will ask you to choose a health plan from a set of plan options based on prices and features from actual health plans.

For this hypothetical decision, please imagine that the plans cover just yourself and not a spouse and/or dependents.

Note that all health plans must offer the same essential benefits including:

- Doctor visits
- Prescription drugs
- Hospitalization
- Maternity and newborn care
- Preventive care

Please click START to begin.

Panel B – Plan Menu

Based on your age and region, 3 health plans are available to you.

Which plan(s) would you like to see to make your choice?

- ALL PLANS (3 plans)
- Bronze Plan (1 plan)
- Silver Plan (1 plan)
- Platinum Plan (1 plan)

Panel C – Plan Detail

The following health plan is available to you:

For all plans, there is a modest copay for prescriptions, and no charge for ER visits once the deductible is met.

SILVER

Monthly Premium: **$311 / month**

Deductible: **$2911 / yr**

Out-of-Pocket Maximum: **$5750 / year**

Copayments/Coinsurance:
Primary Doctor: $30/visit
Specialist Doctor: $50/visit

Do you want to choose this plan or see all available plans?

- Choose Silver Plan
- See All Plans