How Product Standardization Affects Choice: Evidence from the Massachusetts Health Insurance Exchange*

DO NOT CITE OR CIRCULATE
PRELIMINARY DRAFT
For NBER Conference Submission

Keith M Marzilli Ericson and Amanda Starc
Boston University and University of Pennsylvania

January 25, 2013

Abstract

Standardization of complex products is widely touted as improving price competition and consumer satisfaction, but evidence on standardization's effect is limited. We examine a natural experiment: the standardization of health insurance plans on the Massachusetts Health Insurance Exchange. In the pre-period, firms had wide latitude to design contracts, which were then grouped into tiers of quality. A regulatory change forced firms to standardize their insurance plan design and offer seven defined plans. We examine the effect of standardization, and show that it altered the market equilibrium: consumers shifted to relatively more generous health insurance plans and to alternative brands. Using a discrete choice model, we estimate how standardization affected the weights consumers placed on plan attributes. We evaluate the welfare effects of standardization and conduct a number of counterfactuals. We use hypothetical choice experiments with different insurance menus to replicate the effect of standardization and conduct alternative counterfactuals.

*We thank (a lot of people) Raj Chetty, David Cutler, Jonathan Gruber, Larry Katz, Jon Kolstad, David Laibson, Ariel Pakes, Jim Rebitzer, Bob Town. We acknowledge funding from the Lab for Economic Applications and Policy (LEAP) at Harvard University, and the National Science Foundation.

1 Introduction

Health insurance exchanges (HIX) are a critical piece of the Patient Protection and Affordable Care Act (ACA) and are likely to become a critical channel through which consumers purchase their health care insurance. Approximately 20 million consumers will receive coverage via the exchanges, with subsidies totalling about $100 billion per year, tied to the premiums set by insurers. Therefore, understanding consumer demand for insurance is critical in implementing the ACA effectively.

Consumers are likely to face a wide range of choice in exchanges, and a variety of work from previous markets, such as Medicare Part D, indicates that consumers can become overwhelmed and make mistakes (Abaluck and Gruber 2010). However, a variety of behavioral nudges may encourage better choices (Kling et al 2011). In particular, providing clear, transparent information or simplifying a choice menu is often seen as a benign nudge that can improve market outcomes.

Insurance plans are multi-dimensional products (e.g. coinsurance, copayments, deductibles, maximum out-of-pocket limits) that are difficult for consumers to evaluate. We examine a natural experiment on the Massachusetts Health Insurance Exchange (HIX). Initially, the Massachusetts HIX gave firms wide latitude to design the terms of insurance plans. However, in 2010 the financial characteristics of all plans were standardized into seven distinct plan types.

We examine the impact of standardization on consumer choice, firm pricing, and equilibrium outcomes. These results are of general interest for economists examining how product regulation can 1) alter consumer choice and 2) affect equilibrium outcomes. They are also directly policy-relevant, as other states' HIX are considering whether to standardize plans or not.

To analyze the impact of standardization on choice we first show reduced form evidence that it had a substantial effect on brand and tier of plan chosen. We then estimate a discrete choice model in which that decision weights (a.k.a. salience or attention) on various insurance attributes (deductible, brand, etc.) are allowed to vary pre- and post-standardization. The differential decision weights can be interpreted as changes in attention (DellaVigna 2009), salience (Bordalo and Shleifer 2012), or contextual inference (Kamenica 2008).

In practice, standardization of products entails both an actual change in the choice menu, as well as some changes in the decision interface. While both types of changes are relevant for estimating the policy-relevant effect of standardization, it can be valuable to disentangle these two sources of standardization's effect in order to make predictions for other contexts. To separate these effects, we use hypothetical choice experiments in which participants are

presented with different menus and choice interfaces. We first replicate the Massachusetts change and show that when both products are standardized and the decision interface is changed, we see similar results in this experiment as we did in the actual HIX enrollment. This establishes both the validity of our field experiment, as well as the hypothetical choice approach. We further test an additional treatment, in which participants see the post-standardization menu using the pre-standardization choice interface. We find that the primary source of standardization's effect on choice is the menu itself, not the decision interface.

We then run counterfactuals that examine the equilibrium policy implications of standardization. First, we show that choices are affected by both the salience of product characteristics and firm decisions, examining how many more consumers would have chosen a HDHP in the absence of a firm price response. We discuss a variety of methods for welfare evaluation (c.f. Bernheim and Rangel 2011).

The paper is organized as follows. Section 2 describes the Massachusetts Connector and the policy change. Section 3 describes the data and provides reduced form evidence of the impact of standardization. Section 4 outlines the empirical model and presents the structural estimates. Section 5 conducts counterfactual analyses and describes the hypothetical choice experiments. Section 7 concludes.

2 The Massachusetts HIX

2.1 History and Existing Literature

The Massachusetts' HIX was established by the 2006 Massachusetts health reform. The HIX we examine is an unsubsidized health insurance exchange (termed "Commonwealth Choice") for individuals and families making over 300% of the poverty line who were not offered employer-sponsored insurance; a separate, subsidized program serves individuals under 300% of the poverty line. In the time frame we analyze (2009-2010), the exchange had been operating for 2 years, and was regulated by the state. Eventually, it will transition to being an ACA exchange, with slightly different regulation. See Ericson and Starch (2013) for more detail on the exchange. Previous work on the Massachusetts HIX has modeled consumer demand (Ericson and Starc 2012a), and has found that accounting for imperfect competition is important (Ericson and Starc 2012b).

2.2 Plan Standardization

States have a great deal of latitude in designing exchanges, including plan design. However, throughout its existence, the Massachusetts HIX has taken an active approach. Initially, a

number of tiers were defined (bronze, silver, gold) by actuarial value, in a model that was subsequently duplicated by the Affordable Care Act. The Connector required carriers to offer a minimum number of products (six, distributed across tiers) and awarded a seal of approval only to selected providers (cite: Toolkit Series). This system evolved over time to a situation in late 2009 in which 25 distinct plans were offered by one of 6 carriers: Blue Cross Blue Shield, Neighborhood Health Plan, Tufts, Health New England, Harvard Pilgrim, and Fallon.

Interest in standardizing the plans grew out consumer feedback, as consumers were confused by the existing choice architecture. Board members further saw this as an opportunity to improve choice, stating that "consumers don't have to worry that there's some sort of 'gotcha' in the health insurance purchase. They can know that they are comparing equivalent products and so make better informed decisions based on premium and provider differences".[1] However, this left board members with a difficult choice of *how* to standardize the products; rather than relying on competitive pressure forcing insurers to offer the products demand by consumers, the decision would suddenly be top-down. Furthermore, the Connector had little research to guide their decision, and relied heavily on common sense.[2]

The initial standardization led to the creation of seven product categories: Gold, Silver-High, Medium and Low, and Bronze-High, Medium, and Low. The plans were initially offered by the same set of six insurers; in addition, Fallon began offering plans with tiered/limited networks. As a result, standardization lowered the number of contract designs (financial parameters) used in the market, it actually increased the number of plans, in the sense of contract design-carrier combinations. Crucially, standardization also unbundled the decision making process into a decision about a contract design, followed by a decision about a carrier, as seen below in Figure 1. The standardization process is an ongoing one; the silver medium plan has been eliminated due to low demand and additional carriers have been added.

Our paper is most closely related to two ideas within the empirical industrial organization literature. The first concerns costly consumer search, which can lead to equilibrium dispersion in prices for even homogeneous goods. Various studies, including Cebul et al (2011) and Hortascu and Syverson (2004) use this equilibrium dispersion in prices to trace out a distribution of search costs. This literature has primarily focused on homogenous goods markets, though the search problem is almost certainly exacerbated by product differentiation, which implies that consumers must search over a variety of product characteristics, in addition to price. By contrast, the marketing literature has considered the possibility

[1] Nancy Turnbull, Toolkit Series.

[2] Toolkit.

that not all products enter the choice set, and instead allows for consideration sets (cite). Finally, within the health context, Sorenson (2000) considers the role of price posting regulation in driving the equilibrium distribution of prices of prescription drugs. This last paper illustrates and import point: information disclosure can lead to increased price competition.

The idea that information provision affects market outcomes is the second stream of literature with which we hope to connect our paper. While Sorenson considers the role of information in affecting prices, Jin and Leslie (2003) consider the impact on quality, finding that increased disclosure can improve quality. Again, within the health setting, Dafny and Dranove (2005) show that health plan report cards do "tell consumers something they don't know" and increase enrollment beyond the role of market-based information.

Our paper is also related to the behavioral economics of limited attention and salience. See for instance, DellaVigna (2009)'s simple model of attention.

3 Data and Reduced-Form Evidence

Our dataset is transaction-level data (purchase, cancellation, and payments) from the unsubsidized market (Commonwealth Choice) from the beginning of the Connector's existence in July 2007 until July 2010 (with additional data to be added). We observe approximately 50,000 transactions. There are large spikes in initial enrollment during the first month of the Connector's existence as well as just before the individual mandate took effect in December 2007, with a steady-state enrollment of approximately one thousand households per month.

Our choice analyses focuses on a subset of the data: Nov. 2009-Feb. 2010. Because we observe transaction-level data, we do not observe all the plan prices that individuals face. However, for the subsample, we collected an extensive set of price quotes from the Connector website using a Perl script. The choice of sample period does not have a strong effect on the results, and we show the robustness of our results to various sample selections.

First, we provide basic summary statistics in 1. Bronze plans are popular during both time periods, though they are slightly more popular in the earlier period. The shift in enrollment from bronze plans is absorbed by silver plans. Interestingly, there is a big decrease in the number of consumers who choose high deductible plans (defined as in the tax code as plans with at least a $1200 individual deductible[3]): over half of all enrollees choose high-deductible plans in 2009, while just over one-fourth choose high-deductible plans in 2010. Furthermore, premiums do not decrease, which, while surprising, could be due to a change in composition of enrollees. However, we note that, while the population is slightly

[3] $1150 in 2009.

older in 2010 (which speaks to the composition of enrollees), the information provision does not seem to have a huge impact on prices.

To examine the effect of the standardization on prices, we present a series of hedonic regressions in 2. Unadjusted for generosity, plans are slightly cheaper in 2010. However, column 2 shows that this is largely because more lower generosity (bronze and silver) plans were available in 2010. If anything, plans became slightly more expensive in 2010. However, columns 3 and 4 show that this effect is not uniform across plan types. Less generous tiers become *relatively* more expensive, while the cost savings associated with choosing a HDHP are lower in 2010. Therefore, we should expect consumers to choose more generous plans in 2010 given the price change.

4 Standardization and Decision Weights

We estimate a discrete choice model of demand for products in which 1) we model the utility of insurance plans as a function of their characteristics and 2) allow the weight on these characteristics to differ pre- and post- standardization. Estimating separate decision weights pre- and post-standardization allow us to model how demand differs post-standardization; these demand estimates can then be used to make inference about optimal firm response.

In this section, we remain agnostic about why the decision weights change post-standardization, but discuss and test among the alternatives in subsequent sections. Decision weights can result from context-dependent inference (Kamenica 2009), context-dependent salience (Bordalo, Gennaioli, and Shleifer 2012), or from attention in the way information is presented (DellaVigna 2009).

4.1 Model of Consumer Choice

In the simplest discrete choice model, consumers face discrete choice from a set of plans, and value various plan characteristics; the mean value of a plan will be denoted by δ_j and the individual-specific component will be denoted by μ_{ij}. Initially, we assume that consumer i's utility of plan j is given by:

$$u_{ij} = \delta_j + \mu_{ij} + \varepsilon_{ij}$$

where ε_{ij} is an error term that i.i.d extreme value.

Additionally, we will allow consumer utility to deviate from the decision index to accommodate a complex decision process (see also Ericson and Starc 2012). We model both utility and salience of product characteristics and use a conditional logit specification to estimate consumer demand for insurance plans. Define the decision index to be $v_{ij} = f(u_{ij})$, where

$f(u_{ij})$ is an affine transformation of the utility. We assume the consumer chooses the plan with the largest decision index. Given the assumption that ε_{ij} is i.i.d. extreme value, this implies that a plan's enrollment share can be written as:

$$s_{ij} = \frac{\exp(v_{ij})}{1 + \Sigma_j \exp(v_{ij})}$$

where s_{ij} represents the probability that consumer i purchases product j.[4]

The decision index depends not only on the underlying utility of the product, but on the salience of various product characteristics. That is, under the standard model, $\delta_j = X_j^* \beta$, while in the decision index model, $\delta_j = X_j^* (\beta\sigma)$, where the vector β represents the primitive preference and the vector σ represents the salience of the product characteristics. This σ may be a function of the way information is presented, the products position in the choice set, or any "ancillary condition" in the language of Bernheim and Rangel (2011).

4.2 Identification

In a model in which certain product characteristics receive more weight than others, identification is likely to be difficult. In the simplest models, salience of product characteristic for a good is measured by function σ which serves as a multiple on the utility derived from that product characteristic. It is easy to see that such a measure is not non-parametrically identified. Is a product popular because it contains a bundle of popular characteristics or because it contains characteristics particularly salient to consumers? The literature suggests a number of solutions to this problem. While some papers rely of documenting dominated decisions or modelling things that directly affect utility, like switching costs (Handel 2012), other papers rely on restrictions from theory (Abaluck and Gruber 2012).

However, policy variation of the kind we observe can identify *changes* in salience. This requires a number of additional assumptions, including stable preferences, a lack of heteroskedasticity, and no measurement error. More formally, let δ_{10} be the average quality of a plan with characteristics X in 2010 and δ_{09} be the average quality of a plan with the same characteristics in 2009. Consider the simples model of salience in which σ_k is simply a multiple on the utility derived from a plan characteristics. That is,

$$\begin{aligned} \delta_{10} &= X\beta\sigma_{10} + \xi \\ \delta_{09} &= X\beta\sigma_{09} + \xi. \end{aligned}$$

[4] In the absence of individual heterogeneity μ_{ij} and heuristics h_{ij}, the δ_j parameters simply represent an inversion of the observed market shares for each plan.

Then $\Delta\delta = X\beta\left(\sigma_{10} - \sigma_{09}\right)$ and we can normalize $\sigma_{09} = 0$.

We need a number of additional assumptions to identify σ_{10}. First, there cannot be differential measurement error across the two years of data. The characteristics we measure (such as brand and metal dummies) cannot have increased or decreased in value: bronze plans must be equally generous across the two regimes. This is likely to be satisfied, as the 2010 plans were often modelled after 2009 plans. Second, underlying preferences must be stable across the two policy regimes. While this is obviously impossible to prove, we do note that the consumer characteristics that are correlated with preferences are similar before and after the policy change. Finally, there can be no heteroskedasticity that requires rescaling of the coefficients (see Train 2003). In our example, these assumptions are likely to be satisfied. However, our welfare calculations (below) will consider robustness checks that loosen these assumptions (at the cost of separate identification).

Of course, the researchers does not always have access to policy variation of the type we observe. In such cases, the research may be willing to place additional structure on the problem in order to separately identify preferences from the effect of salience. As with any modelling choice, the results will depend on the function used. A sensible approach is to use the theories of salience developed in the behavior literature with describe how local thinkers are likely to behave. We can then discuss whether such a model is identified and, furthermore, what patterns of consumer heterogeneity (in either preferences or choice process) can be identified.

4.3 Estimates

We estimate a consumer utility function that allows for changes in decision-weights in 2010:

$$u_{ij} = X_{ij}\beta\left(\sigma^{*}1\left[2010\right]\right) - \alpha p_{ij}\left(\sigma^{*}1\left[2010\right]\right) + \varepsilon_{ij},$$

allowing for separate changes in weightson premiums, the HDHP and the tiers, and insurer brand. The results are in Table 3. The second specification shows that, holding the weight on brand constant, the premium is slightly less salient to consumers, while financial characteristics, including plan tier and the HDHP dummy, are much more salient to consumers. The full model shows information disclosure had a much stronger effect on the salience of premiums: the σ coefficient on premiums is much lower than one. Again, financial characteristics become more important. Finally, consumers are less concerned with insurer name and network post-standardization.

5 Counterfactuals

5.1 Impact of Salience on Choice

We first run a counterfactual experiment that attempts to disentangle the supply- and demand- side forces that explain changes in market shares across plans. We focus on the number of consumers choosing bronze policies. We simulate the percentage of consumers choosing bronze plans given salience change inclusive and salience change exclusive preference under both the old and new price vectors. Table 4 shows the results. Choice change very little given the changes in premiums. However, the salience change inclusive preferences make a large difference. In the first column, we see that under either set of prices, consumers who have salience change inclusive preferences purchase bronze plans approximately 64% of the time. The second column shows that, under either set of prices, consumers who have salience change exclusive preference purchase bronze plans only 53% of the time. The roughly 10% fall in the market share of bronze plans is due to demand-side factors, rather than firm pricing or the choice set.

5.2 Welfare

A natural question is whether standardization made consumers better off and, if so, what the source of the gains in consumer surplus were. Because we use a logit specification to estimate demand, we face the problem that consumer utility grows as the size of the choice set increases. Therefore, we also compare the 2009 choice set to a limited 2010 choice set that included only the three most popular carriers from 2010. Therefore, we hold the number of plans fixed, choosing the 25 most popular plans in 2010 to compare to the 25 plans in 2009. We measure equivalent variation using the standard formula of Nevo (2001) and McFadden (1999), which is given by:

$$EV = (1/\alpha) \left[\log \left(\Sigma_j \delta_{j2010} \right) - \log \left(\Sigma_j \delta_{j2010} \right) \right].$$

We evaluate welfare separately, in one colulmn assuming the 2009 decision weights are the "real preferences", and in the other assuming the 2010 weights are the real preferences. In each row, we compare welfare pre and post standardization, where standardization changes the choice menu. In row 1, we estimate the total effect of standardization, which includes a shift in menu, premiums, and consumer decision weights/salience. In row 2, we hold decision weights constant at their 2009 levels, and so standardization's effect includes both the menu change and the change in firm's premiums. In row 3, we hold premiums constant,

but allow standardization to change decision weights and the menu. Finally, in row 4, we hold premiums constant as well, and allow standardization to affect the menu only.

The results are in Table **??**. The overall improvement in consumer welfare is around $25 per member per month, which is about 7% of the average monthly premium paid. The effect of the changed menu alone, nearly $30 per member/per month, is quite large. If for no other reason, the standardization improved consumer outcomes by providing a set of products consumers liked. Allowing salience to vary between the two regimes also increases welfare, as seen in the third estimate. If one is willing to say that 2010 preferences are the "real", primitive preferences, standardization is an even better policy for consumers. The second estimate is smaller because it no longer holds premiums constant; firms reoptimize and capture some of the surplus generated by the standardization policy. This represents the lower bound of the impact of standardization on consumer welfare. At $15 per member per month, this still represents approximately 4% of premiums.

Table **??** also presents the marginal effect of each aspect of the policy change: changed consumer salience, premiums, and menus of premiums. Firms capture part of the surplus: the effect of changing premiums is a negative $10 per member per month. The changed menu created the bulk of the benefit, around $30, and the change in consumer surplus captured the rest, around $10.

5.3 Pilot Experiment

Table 5 shows the results of a pilot experiment in which participants were randomly to assigned to one of three choice scenarios (between subject), and asked to choose an insurance plan. Participants were recruited from an online labor market.

The "Pre-Standardization" choice menu was nearly identical to the interface and HIX's menu for insurance plans December 2009 (for someone in the Boston area, at the prices of a 30 year-old. The "Post-Standardization" choice menu replicated the January 2010 situation: it contained the post-standardization menu, and had participants first choose a tier of insurance generosity and then choose a brand. The "Post Standardization, Pre-Standardization Inferface" menu has the exactly same plans as in the "Post-Standardization" condition, but uses the pre-standardization decision interface.

While the experimental results differ in some ways from the field results, they generally support the pattern we find: moving from the "Pre-Standardization" to "Post-Standardization" condition raises the market share of Neighborhood, and lowers that of Fallon. The fraction choosing HDHP plans declines. (Though, in contrast to what we saw in the field, the fraction choosing silver plans declines rather than increases.)

The counterfactual condition "Post Standardization, Pre-Standardization Inferface" allows us to test whether the observed shifts in choice or due to the menu or the interface. There are only small differences in the brands chosen, comparing this condition to the "Post-Standardization" condition. This indicates that the effect of standardization on choice came through the choice menu, rather than the interface. (Again, however, tier chosen differed, with these participants being most likely to choose gold plans).

This pilot experiment establishes the general validity of the hypothetical choice approach. A large scale experiment will be deployed with a more representative age-sample of New England residents in the appropriate income range. This sample should be more representative in their familiarity with regional brand differences, and in willingness-to-pay for insurance.

5.4 Discussion

Given that, in many scenarios, standardization improves outcomes for consumers and firms can capture some of the newly generated surplus, there is an open question of why they did not offer such an assortment of plans initially. We believe that this is due to the central role of information provision in shaping consumer choice. However, other explanations for insurer behavior are equally plausible. The first, firm learning, argues that the market simply may not be in equilibrium yet. The market is relatively new (approximately 4 years old at the end of our sample period) and that may not have been enough time for firms to learn about both costs and demand. Demand will obviously affect pricing (see Ericson and Starc 2012) as well as product offerings. Furthermore, as we explore below, firms may also need to know about the nature of selection. Finally, given that this may be a relatively small proportion of a carrier's book of business, they may not have a huge incentive to perfect their offerings in this particular market.

Second, regulation has a large impact on the offerings of firms in many insurance markets, and especially in this insurance market. In some sense, this is a paper about improved regulation via information provision. Initially, firms were told to submit just two bids for bronze plans. They were to be of the same basic design, with one having prescription drug coverage and one omitting prescription drug coverage (the latter were later abandoned). Underlying this logic was the idea that consumers, contrary to economists, may actually prefer less choice, so that they don't have to compare "apples to lava" (cite: Toolkit Series). Given this view, the standardization may have improved welfare by introducing additional choice while providing additional decision support tools that allowed them to express their preferences. Therefore, it may be the interaction of the choice set and the information provision that led to gains in consumer surplus.

A third possibility is that selection led to the product assortment in the pre-standardization period. It is possible that a low-deductible Neighborhood bronze plan would have attracted a very high cost subset of the population and would not have been profitable. However, it then seems interesting that BCBS would offer such a plan, given that sicker consumers would be drawn to their more extensive network and stronger brand name. Regardless, this highlights the fact that one-shot deviations may or may not be profitable. As a case-in-point, we estimate that the Neighborhood Bronze High plan, which captures a 16% market share in 2010, would capture roughly half that number of consumers, or an 8% market share, under the 2009 preferences.

Finally, we note that the policy change was not without risk. In the bottom panel of Table 5, we see that if the insurers were forced to offer only one plan of each medal - bronze, silver, and gold - the consumers are made worse off under the 2010 prices regardless of their preferences. Therefore, this type of policy intervention requires caution. Without expanding the choice set and providing a high-value option to consumers, the policy change could have reduced welfare.

6 Conclusion

We analyze a real-world change in "choice architecture" and examine its effect on consumers. In 2010, the Massachusetts Connector standardized the policies available to consumers and changed the way information was presented. We argue that the change altered the salience of multiple product characteristics to consumers. In this market, consumers trade-off financial and network generosity off against premiums. The change in "choice architecture" made financial generosity more salient and network generosity less salient to consumers. As a result, consumers tended to choose more financial generous plans and less HDHP, while simultaneously opting for insurers with narrower networks. We show that this is due to demand-side, rather than supply-side, factors.

Consistent with the behavioral economics literature, we find that well-designed "choice architecture" generates benefits for consumers. Depending on the exact specification, consumer welfare increases 4-10% due to the policy change. However, firms are able to capture some of the surplus associated with the change. As a result, we argue that potential choice architects take firm behavior into account when setting policy. Finally, our results can inform the design of federal and state exchanges under the Affordable Care Act.

Table 1: Basic Summary Statistics

	2009	2010
Bronze	0.6135	0.5547
Silver	0.3012	0.3397
Gold	0.0853	0.1056
HDHP	0.4493	0.2858
BCBS	0.1513	0.1573
Fallon	0.2166	0.0775
Harvard Pilgrim	0.1892	0.1344
HNE	0.0471	0.0237
Neighborhood	0.3408	0.4911
Tufts	0.0550	0.1167

Table 2: Hedonic Regressions, List Prices

	(1)	(2)	(3)	(4)
Silver		-141.0***	-146.6***	-146.6***
		(0.426)	(0.535)	(0.534)
Bronze		-267.2***	-275.0***	-272.3***
		(0.450)	(0.597)	(0.831)
2010 Dummy	-14.71***	21.98***	7.268***	7.266***
	(0.735)	(0.320)	(0.759)	(0.757)
Silver*2010			16.37***	16.40***
			(0.887)	(0.885)
Bronze*2010			19.66***	27.62***
			(0.923)	(1.185)
HDHP				-4.093***
				(0.893)
HDHP*2010				-11.75***
				(1.141)
Constant	338.7***	538.0***	542.8***	542.5***
	(0.830)	(0.560)	(0.604)	(0.607)
Fixed Effect		age cat, carrier	age cat, carrier	age cat, carrier
Observations	106,940	106,940	106,940	106,940
R-squared	0.514	0.912	0.912	0.912

Table 3: Discrete Choice Model: Decision Weights Vary Pre- and Post-Standardization

	(1)	(2)	(3)	(4)
	Conditional Logit		Mixed Logit	
Premium*2009	-0.0259***	-0.0274***		
	(0.00126)	(0.00131)		
Premium*2010	-0.00122**	-0.00119**		
	(0.000554)	(0.000580)		
Premium*2009, Mean			-3.484***	-3.298***
			(0.0490)	(0.0489)
Premium*2010, Mean			-3.442***	-3.326***
			(0.0473)	(0.0466)
Premium*2009, SD			0.189***	0.255***
			(0.0207)	(0.0211)
Premium*2010, SD			0.277***	0.269***
			(0.0246)	(0.0229)
Premium*Age	0.000287***	0.000303***	0.000316***	0.000359***
	(1.79e-05)	(1.84e-05)	(2.13e-05)	(2.45e-05)
Silver*2009	0.362***	0.395***	0.778***	0.856***
	(0.0850)	(0.0927)	(0.101)	(0.102)
Gold*2009	1.350***	1.469***	1.885***	2.151***
	(0.155)	(0.158)	(0.170)	(0.175)
Silver*2010	1.119***	0.689***	1.565***	1.226***
	(0.0907)	(0.0947)	(0.104)	(0.109)
Gold*2010	2.531***	2.168***	2.942***	2.744***
	(0.155)	(0.156)	(0.172)	(0.176)
HDHP*2009		-0.0850		-0.549***
		(0.0690)		(0.0902)
HDHP*2010		-0.832***		-0.900***
		(0.0736)		(0.0740)
Observations	106,940	106,940	106,940	106,940

Table 4: Counterfactuals

Percent Choosing Bronze		
	Pre-Std DW	Post-Std DW
Price		
2009	0.64	0.54
2010	0.63	0.52

Table 5: Hypothetical Choice Experiment: Pilot Results

	Pre-Std	Post-Std	Post-Std Menu, Pre-Std Interface
BRAND			
Blue Cross	7.84	4.17	4.08
Fallon	11.76	0	2.04
Harvard	7.84	0	2.04
Neighborhood	23.53	62.5	75.51
Tufts	49.02	33.33	16.33
TIER			
Bronze	33.93	30.19	26.53
Silver	58.93	50.97	30.61
Gold	7.14	18.87	42.86
HDHP	32.14	20.75	18.36
N = 158			

Note: Pilot Experimental Data

Web Appendix

A.1 Appendix Tables and Figures

Table A.1: Enrollment-Weighted Hedonics

	(1)	(2)	(3)	(4)
Silver		-108.8***	-115.0***	-115.8***
		(2.373)	(2.988)	(2.956)
Bronze		-226.0***	-233.4***	-248.7***
		(2.256)	(2.832)	(3.405)
2010 Dummy	15.01***	15.52***	-1.225	-2.064
	(3.277)	(1.384)	(4.262)	(4.215)
Silver*2010			16.33***	17.43***
			(4.872)	(4.820)
Bronze*2010			20.15***	45.12***
			(4.645)	(5.253)
HDHP				21.76***
				(2.797)
HDHP*2010				-39.70***
				(4.090)
Constant	264.3***	524.9***	532.0***	538.9***
	(3.274)	(2.809)	(3.267)	(3.345)
Fixed Effect		age cat, carrier	age cat, carrier	age cat, carrier
Observations	4,156	4,156	4,156	4,156
R-squared	0.542	0.923	0.923	0.925