Online Appendix to Accompany "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program: Comment"

By JONATHAN D. KETCHAM, NICOLAI V. KUMINOFF, AND CHRISTOPHER A. POWERS

I. DERIVING ABALUCK AND GRUBER'S WELFARE MEASURES

AG begin by assuming that consumers' decisions are guided by a logit model that is linear and additively separable in PDP characteristics, as shown in (4) and repeated here for convenience.

(A1) $$\hat{u}_{ij} = \hat{\omega}_{ij} + \hat{\varepsilon}_{ij} = p_j\hat{\alpha} + \mu_{ij}\hat{\beta}_1 + \sigma_{ij}^2\hat{\beta}_2 + \mathbf{c_j}\hat{\beta}_3 + \mathbf{q_j}\hat{\gamma} + \hat{\varepsilon}_{ij}.$$

In contrast, the actual utility a consumer experiences from her selected PDP is instead defined by the following hedonic utility function that satisfies AG's three normative restrictions,

(A2) $$u_{ij} = \omega_{ij} + \varepsilon_{ij} = p_j\hat{\alpha} + \mu_{ij}\hat{\alpha} + \sigma_{ij}^2\tilde{\beta}_2 + \mathbf{q_j}\hat{\gamma} + \varepsilon_{ij}, \text{ where } \tilde{\beta}_2 < 0.$$

Because AG assume the marginal utility of income is a constant revealed by $\hat{\alpha}$, a consumer's expected welfare from choosing plan *j* can be expressed as

(A3) $$E[CS_i] = \frac{1}{\hat{\alpha}} E\left[u_{ij} \middle| \hat{u}_{ij} > \hat{u}_{ik}\ \forall\ k\right].$$

PDP choice and welfare are both deterministic from the consumer's perspective. The expectation in (A3) simply reflects the analyst's inability to observe $\hat{\varepsilon}_{ij}$.

AG aim to calculate the partial equilibrium welfare gain from a hypothetical intervention "that would make individuals fully informed and fully rational" (p 1208). In other words, they want to calculate the welfare gain from a policy that would induce the consumer to choose the PDP that maximizes AG's normative

utility function (A2) instead of (A1). Assuming the policy has no effect on the marginal utility of income, the welfare gain can be expressed in general terms as

$$\text{(A4)} \qquad \Delta E[CS_i] = \frac{1}{\hat{\alpha}} E\left[\max_k\{u_{ik}\} - \left(u_{ij} \middle| \hat{u}_{ij} > \hat{u}_{ik}\ \forall\ k\right)\right].$$

The analytical formula depends on the interpretation of the residual utility terms, $\hat{\varepsilon}_{ij}$ and ε_{ij}.

In appendix D of the NBER (2009) version of their paper, AG outline two different approaches to interpreting residual utility. The more conventional approach, laid out in earlier papers such as Leggett (2002), is to interpret $\hat{\varepsilon}_{ij}$ as the idiosyncratic utility from PDP characteristics that consumers observe but the analyst does not. Examples include proximity to in-network pharmacies, availability of mail-order pharmacies, individual-specific experience with the insurers, coordination with spouses, disutility from prior authorization requirements, uncertainty about whether other plans will approve prior authorization requests, and so on. In this case the policy intervention has no effect on the utility residual because the same unobserved PDP attributes enter hedonic utility.[1] Thus $\varepsilon_{ij} = \hat{\varepsilon}_{ij}$.

In contrast, the approach that Abaluck and Gruber (2011, 2013) use for their published empirical analyses is to assume that the policy intervention also eliminates the utility residual: $\varepsilon_{ij} = 0$. That is, ε_{ij} itself is treated as an optimization mistake in addition to violations in the three parametric restrictions that they explicitly mention as reducing welfare (p.1208). This approach embeds at least three important assumptions. First, it assumes there are no omitted variables. The analyst must have data on every PDP attribute that affects consumers' hedonic utility. Second, it assumes (A1) and (A2) are correctly specified. The analyst must know the true parametric forms of decision utility and hedonic utility. Third, it assumes the policy intervention has no direct effect on utility. For example, the two poli-

[1] As we point out in section IV, $\hat{\varepsilon}_{ij}$ may also reflect misspecification of the true parametric form of decision utility. In this case ε_{ij} may differ from $\hat{\varepsilon}_{ij}$ if the policy affects the marginal decision utility of one or more PDP attributes included in $\hat{\varepsilon}_{ij}$.

cies suggested in AG may affect welfare due to distaste for being nudged or distaste for sacrificing control over plan choices to a surrogate decider. Together, these three assumptions are required for AG to treat $\hat{\varepsilon}_{ij}$ as an idiosyncratic optimization mistake that is eliminated by their hypothetical policy.

In the remainder of this section we derive analytical formulas for consumer welfare under each of the two approaches to interpreting residual utility. Whereas Abaluck and Gruber (2009) derive measures of baseline consumer surplus prior to any policy intervention, we derive the key statistic used in their welfare calculations (and ours)—the *change* in consumer welfare caused by the hypothetical policy "that would make individuals fully informed and fully rational".

Case 1. Residual Utility is an Optimization Mistake: $\boldsymbol{\varepsilon_{ij} \equiv 0\ \forall ij}$

In this case the analyst can calculate baseline consumer surplus for each individual by using the marginal utility of income to translate utils into dollars:

$$E[CS_i] = CS_i = \frac{\omega_{ij}}{\hat{\alpha}}. \tag{A5}$$

After consumers are made to choose the plans that maximize AG's normative utility function the post-policy consumer surplus becomes

$$E[CS_i^*] = CS_i^* = \frac{1}{\hat{\alpha}} \max_k \{\omega_{ik}\}. \tag{A6}$$

Hence the change in welfare generated by the hypothetical policy is

$$\Delta E[CS_i] = CS_i^* - CS_i = \frac{1}{\hat{\alpha}} \left[\max_k \{\omega_{ik}\} - \omega_{ij} \right]. \tag{A7}$$

Because $\varepsilon_{ij} \equiv 0$ the analyst can calculate actual consumer surplus instead of expected consumer surplus.

Case 2: Residual Utility Reflects Omitted Attributes $\hat{\varepsilon}_{ij} = \varepsilon_{ij}\ \forall ij$

In this case the analyst must integrate over the assumed Type I EV distribution for $\hat{\varepsilon}_{ij}$ to calculate expected consumer surplus prior to the policy. The resulting expression in (A8) depends on the standard log sum rule as well as the difference between decision utility and hedonic utility weighted by the probability of selecting each PDP (e.g. Small and Rosen 1981, Leggett 2002, Abaluck and Gruber 2009).

$$E[CS_i] = \frac{1}{\hat{\alpha}}\left[ln\sum_k e^{\hat{\omega}_{ik}} + \sum_j\left(\omega_{ij} - \hat{\omega}_{ij}\right)\frac{e^{\hat{\omega}_{ij}}}{\sum_k e^{\hat{\omega}_{ik}}}\right] + \hat{C}. \tag{A8}$$

In the equation, $\hat{C}$ represents the constant of integration divided by $\hat{\alpha}$. It arises from the assumed Type I EV distribution for $\hat{\varepsilon}_{ij}$ and the fact that the level of utility is unknown.

The policy intervention eliminates the wedge between decision utility and hedonic utility, simplifying calculation of post-policy consumer surplus:

$$E[CS_i^*] = \frac{1}{\hat{\alpha}}[ln\sum_k e^{\omega_{ik}}] + C, \text{ where} = \hat{C} + \frac{\rho}{\hat{\alpha}}. \tag{A9}$$

If the policy intervention has a direct effect on utility, defined here by ρ, then the post-policy constant of integration, C, differs from the pre-policy constant of integration.[2] On the other hand, if we follow AG in assuming that the policy has no direct effect on utility then $\rho = 0$ and $C = \hat{C}$. In this case, the change in expected consumer surplus is

$$\Delta E[CS_i] = E[CS_i^*] - E[CS_i] = \frac{1}{\hat{\alpha}}\left[ln\frac{\sum_k e^{\omega_{ik}}}{\sum_k e^{\hat{\omega}_{ik}}} - \sum_j\left(\omega_{ij} - \hat{\omega}_{ij}\right)\frac{e^{\hat{\omega}_{ij}}}{\sum_k e^{\hat{\omega}_{ik}}}\right]. \tag{A10}$$

[2] We assume that any direct effect of the policy on utility is additive and invariant to PDP choice so that $E[CS_i^*] = \frac{1}{\hat{\alpha}}\max_k\{\omega_{ik} + \varepsilon_{ij}\} = \frac{1}{\hat{\alpha}}[ln\sum_k e^{\omega_{ik}+\rho}] + \hat{C} = \frac{1}{\hat{\alpha}}[ln\sum_k e^{\rho}\, e^{\omega_{ik}}] + \hat{C} = \frac{1}{\hat{\alpha}}[ln(e^{\rho}\sum_k e^{\omega_{ik}})] + \hat{C} = \frac{1}{\hat{\alpha}}[ln(e^{\rho})] + \frac{1}{\hat{\alpha}}[ln\sum_k e^{\omega_{ik}}] + \hat{C} = \frac{1}{\hat{\alpha}}[ln\sum_k e^{\omega_{ik}}] + C$.

Equation (A10) isolates the combined welfare effect of imposing the three normative restrictions on utility that AG emphasize. In contrast, the 27% welfare gain that AG report in their conclusion is based on the calculation in (A7) that embeds their normative restrictions along with the added assumption that residual utility consists entirely of optimization mistakes. Therefore, comparing empirical results for (A7) and (A10) will reveal the extent to which AG's reported 27% potential welfare gain is driven by the particular optimization mistakes they emphasize relative to their novel interpretation of the Type I EV logit error term.

Leggett, Christopher G. 2002. "Environmental Valuation with Imperfect Information". *Environmental and Resource Economics*. 23: 343-355.

Small, Kenneth A. and Harvey S. Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica*. 49(1): 105-130.

II. Additional Results

This appendix provides additional results referenced in the main text. Table A1 provides an example of the difference between AG's definition for brand dummies that relies on CMS contract ID codes that are unobserved by consumers and our definition that relies on company and plan names observed by consumers. We define AARP and UnitedHealth as two distinct brands, whereas AG group one AARP plan and one UnitedHealth plan into one brand, and two AARP plans and one UnitedHealth plan into a separate brand.

Table A1—Example of the Difference between Contract ID and Brand name Dummy Variables

Plan Name	Brands #1 and #2 using:	
	contract ID	brand name
AARP MedicareRx Plan	1	1
AARP MedicareRx Plan - Enhanced	2	1
AARP MedicareRx Plan - Saver	2	1
UnitedHealth Rx Basic	2	2
UnitedHealth Rx Extended	1	2

Note: Example is from the Region 2 (CT, MA, RI and VT) in 2007.

Figure A1 reports the gap premium and gap enrollment rates for various alternative samples. Panel A shows that the divergence between AG's results and results from the CMS data widens when part-year enrollees are included in the CMS sample as they likely were in AG's sample. The remaining panels provide further evidence that people responded to how gap coverage mattered for themselves. Panel B depicts CMS region 25 which was the region with the largest number of (non-poor) PDP enrollees. It is comprised of Iowa, Minnesota, Montana, Nebraska, North Dakota, South Dakota, and Wyoming. People in these states had exclusive access to a plan with especially generous gap coverage, as seen from comparing the cost premia in panel B with that in Figure 1B. They re-

sponded by enrolling at much higher rates—up to 75% at the 98th expenditure quantile. Thus, enrollment in gap plans varied dramatically across regions with the regional rate of enrollment increasing in the generosity of coverage.

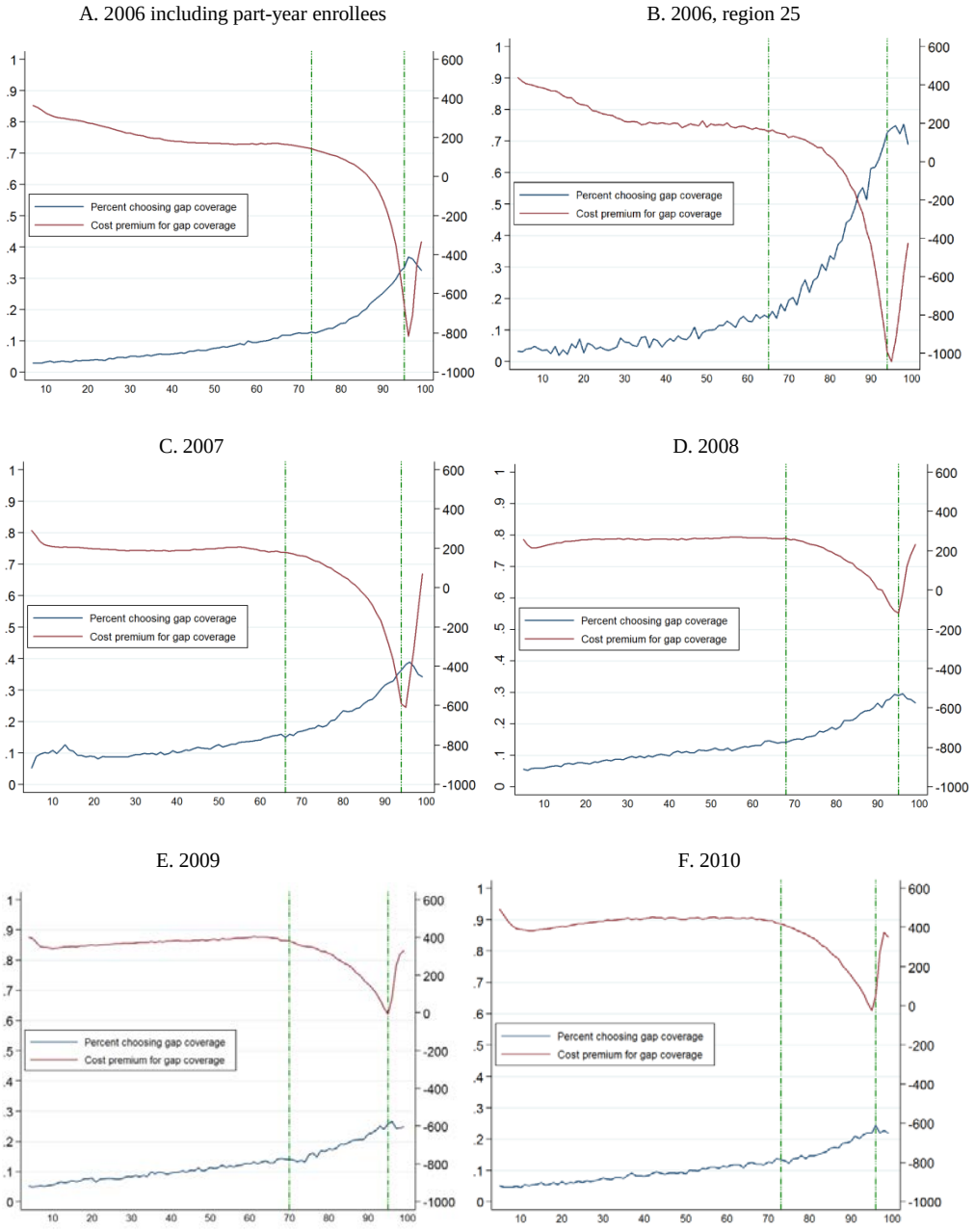

FIGURE A1: PERCENT CHOOSING GAP COVERAGE AND ADDED COST BY EXPENDITURE QUANTILE

Table A2 summarizes how the average consumer's chosen plan differs from other plans the consumer could have chosen. Each cell reports the difference between an attribute of the consumer's chosen plan and the mean value of that same attribute calculated over all of the plans that the consumer could have chosen but did not. For example, in 2006 the average consumer paid $112 less in out of pocket costs for prescription drugs under her chosen plan then she would have paid, on average, if she had enrolled in a different plan than was available to her.

TABLE A2—DIFFERENCE BETWEEN THE CHOSEN PLAN AND THE MEAN ALTERNATIVE

	2006	2007	2008	2009	2010
sample size	464,543	566,962	602,992	614,714	629,225
premium (difference in $)	-89	-73	-65	-62	-53
out of pocket costs (difference in $)	-112	-109	-164	-140	-187
variance of OOP costs (difference in percentage points)	-16	27	36	-5	-7
count of top 100 drugs covered (difference in number of drugs)	2	1	1	1	1
CMS quality index (difference in percentage points)	6	5	2	7	0

Note: Each row is calculated as the average over all people of the difference between the attribute of their chosen plan and the average of that same attribute calculated over all others plans in the individual's choice set. The unit of analysis is the individual person.

Table A3 provides the share of people in 2006 and 2007 that could reduce their spending by certain amounts by moving from their plan without gap coverage into the cheapest plan with gap coverage, or by moving from their plan with gap coverage into the cheapest plan without gap coverage.

TABLE A3—POTENTIAL SAVINGS FROM MOVING INTO OR OUT OF A GAP PLAN, 2006-2007

	2006		2007	
Percent who could save more than $X by moving	Into a gap plan	Out of a gap plan	Into a gap plan	Out of a gap plan
$100	9.9	4.2	6.6	8.6
$300	8.0	2.3	3.7	1.3
$500	4.2	0.1	2.2	0.2
$750	2.5	0.0	1.2	0.1
$1,000	1.4	0.0	0.6	0.0

Table A4 repeats the nonparametric analysis in Table 2 after replacing our brand dummies (based on company name) with AG's brand dummies (based on contract IDs).

TABLE A4— NONPARAMETRIC TEST OF CHOICE INCONSISTENCY WITH BRAND DUMMY VARIABLES DEFINED USING CONTRACT ID

	Plan attributes affecting utility	Assumption on expected drug expenditures in year *t*	% Consumers choosing frontier plans				
			2006	2007	2008	2009	2010
(1)	E[cost]	year t drug consumption	7	7	10	6	8
(2)	E[cost], var(cost)	year t drug consumption	25	24	24	26	36
(3)	E[cost], var(cost), CMS quality	year t drug consumption	35	33	46	42	45
(4)	E[cost], var(cost), brand	year t drug consumption	74	77	80	87	87
(5)	E[cost], var(cost), brand	year t or t-1 drug consumption		81	86	91	90

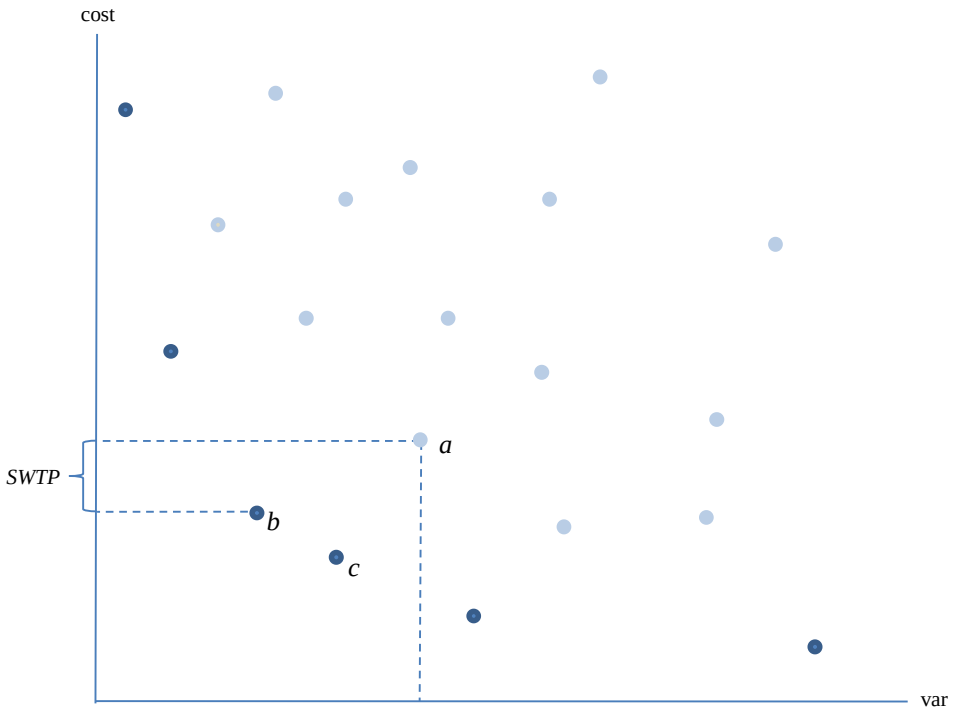

FIGURE A2—ILLUSTRATION OF THE SUFFICIENT WILLINGNESS TO PAY FOR BRAND

Figure A2 illustrates how we calculate the sufficient willingness to pay (*SWTP*) for the bundle of unobserved PDP attributes that vary from brand to

brand. To begin, consider a plan, *a*, that lies on the efficiency frontier in cost-variance-brand space, where cost means the total cost (premiums plus ex post OOP drug costs) to the individual. Figure A2 is projected in cost-variance space. The dots represent other available plans. Plans on the efficiency frontier in cost-variance space have dark shading; plans off the frontier have light shading. The area inside the rectangle defined by the dashed lines that intersect at point *a* defines the portion of the efficiency frontier where other plans dominate *a* in cost-variance space. In the figure there are two such plans, *b* and *c*. We define *SWTP* as the amount of income the consumer gives up by choosing to purchase plan *a* instead of the most expensive plan on the portion of the cost-variance frontier that dominates plan *a*. Hence, $SWTP = cost_a - cost_b$.

SWTP can be interpreted as an arbitrarily close approximation to the willingness to pay for latent attributes of the consumer's preferred brand for a consumer with preferences satisfying basic axioms of consumer preference theory. To see why, suppose that plan *a* is sold by brand *A* whereas plans *b* and *c* are sold by brand *B*, and the two brands differ in a vector of latent quality attributes, *q*. Consider a consumer who prefers plan *b* to plan *c* and is indifferent between plans *b* and *a* such that

$$\begin{aligned} & U(y - cost_b, var_b, q_B) \\ & = U(y - cost_a, var_a, q_A) \\ & = U(y - cost_b - SWTP, var_a, q_A), \end{aligned} \tag{A11}$$

where the last line follows from the definition of *SWTP*. The consumer's exact willingness to pay (WTP) to switch from q_B to q_A, evaluated at the best available point on the efficiency frontier in cost-variance space, is implicitly defined by the following equation

$$U(y - cost_b, var_b, q_B) = U(y - cost_b - WTP, var_b, q_A). \tag{A12}$$

Combining (A11) and (A12) yields the following expression

(A13) $$U(y - cost_b - SWTP, var_a, q_A) = U(y - cost_b - WTP, var_b, q_A).$$

Assuming the consumer's preferences satisfy global risk aversion and strong monotonicity it must be the case that $WTP > SWTP$. This follows from (A13) because quality is held constant at q_A. That is, in order to hold utility constant when the variance decreases from var_a to var_b, the risk averse consumer's income must be reduced. Thus, $WTP = SWTP + \varepsilon$, where ε is a positive constant that reflects the willingness to pay to reduce the variance from var_a to var_b at q_A. Finally, notice that ε can be made arbitrarily close to zero (e.g. one tenth of one cent) without violating completeness, transitivity, strong monotonicity or risk aversion. It follows that $SWTP$ provides an arbitrarily close approximation to the willingness to pay for latent attributes of the consumer's preferred brand, conditional on cost and variance, that is sufficient to rationalize the consumer's observed choice. Also notice that SWTP equals 0 for any plan on the efficiency frontier in cost-variance space, whereas no value for SWTP can rationalize the choice of a plan that lies off the efficiency frontier in cost-variance-brand space.

This logic generalizes to any number of plans on the portion of the efficiency frontier that dominates plan *a* in cost-variance space. Regardless of the thickness or sparseness of plans in attribute space, we can always set ε to be less than *e*, where *e* is an arbitrarily small positive constant. Likewise, this logic can be generalized to any assignment of plans to brands by restricting the consumer to have identical tastes for the vector of latent attributes associated with brands *B, C, D,* and so on.

Table A5 reports results from our replication of the first three columns of Table 1 in AG. The columns of the two tables are directly comparable, and both rely on AG's definition of the brand dummy variables. Models with AG's brand-state dummies (AG's column (4)) do not converge.

TABLE A5— REPLICATION OF AG TABLE 1 USING CMS DATA

	(1)	(2)	(3)
Premium (hundreds)	-0.352***	-0.430***	-0.559***
	(0.00110)	(0.00143)	(0.00249)
OOP costs (hundreds)	-0.161***	-0.102***	-0.102***
	(0.000478)	(0.000562)	(0.000578)
Variance (millions)	-0.000136***	-0.000124**	-4.86e-05
	(5.28e-05)	(5.17e-05)	(6.54e-05)
Deductible (hundreds)		-0.101***	-0.0201***
		(0.00169)	(0.00281)
full gap coverage		0.818***	1.897***
		(0.00887)	(0.0146)
generic gap coverage		0.216***	0.529***
		(0.00685)	(0.00907)
Cost sharing		-1.205***	-0.313***
		(0.0156)	(0.0248)
Number of top 100 drugs on formulary		0.184***	0.190***
		(0.00107)	(0.00153)
CMS quality index	4.217***	3.322***	
	(0.00996)	(0.0112)	
Brand dummies	No	No	Yes
Number of consumers	464,543	464,543	464,543
Number of plans	1,348	1,348	1,348
Number of states	48	48	48
Number of brands	73	73	73
Pseudo R^2	0.17	0.20	0.32

Table A6 reports results from estimating the model in column (3) of Table 4 for each year from 2006 through 2010.

TABLE A6— SENSITIVITY OF MAIN RESULTS FROM AG'S FULL MODEL TO THE STUDY YEAR

	2006	2007	2008	2009	2010
Premium (hundreds)	-0.402***	-0.144***	-0.355***	-0.505***	-0.565***
	(0.00246)	(0.00180)	(0.00161)	(0.00183)	(0.00201)
OOP costs (hundreds)	-0.108***	-0.108***	-0.214***	-0.250***	-0.194***
	(0.000573)	(0.000580)	(0.000836)	(0.000849)	(0.000734)
Variance (millions)	-6.48e-05	1.51e-06***	0.000162***	-0.124***	1.46e-05
	(6.42e-05)	(3.39e-07)	(4.70e-05)	(0.0114)	(6.02e-05)
Deductible (hundreds)	0.0510***	-0.341***	-0.233***	-0.709***	-0.713***
	(0.00279)	(0.00175)	(0.00243)	(0.00351)	(0.00247)
Full gap coverage	1.162***	0.326***	-0.136	1.503***	-1.269***
	(0.0146)	(0.0225)	(8,047)	(0.120)	(0.0422)
Generic gap coverage	0.356***	-1.065***	-0.184***	0.281***	0.328***
	(0.00893)	(0.00654)	(0.00749)	(0.00860)	(0.00934)
Cost sharing	0.683***	5.198***	1.067***	-4.636***	-5.149***
	(0.0244)	(0.0208)	(0.0378)	(0.0432)	(0.0361)
Number of top 100 drugs on formulary	0.175***	0.275***	0.181***	0.150***	0.334***
	(0.00144)	(0.00264)	(0.00245)	(0.00268)	(0.00191)
Brand dummies	Yes	Yes	Yes	Yes	Yes
Number of people	464,543	566,962	602,992	614,714	629,225

Table A7 reports results from estimating the model in column (1) of AG's Table 3 for each year from 2006 through 2010. As shown, the premium coefficient is slightly below the OOP coefficient for 2008, 2009 and 2010, and the variance coefficient has a negative sign for both 2009 and 2010.

TABLE A7— SENSITIVITY OF AG'S BASE RESULTS TO THE STUDY YEAR

	2006	2007	2008	2009	2010
Premium (hundreds)	-0.352***	-0.441***	-0.342***	-0.287***	-0.236***
	(0.00110)	(0.00111)	(0.000832)	(0.000793)	(0.000719)
OOP costs (hundreds)	-0.161***	-0.176***	-0.382***	-0.297***	-0.296***
	(0.000477)	(0.000516)	(0.000710)	(0.000680)	(0.000599)
Variance (millions)	-0.000136***	4.76e-07*	0.000252***	-1.269***	-0.000307***
	(5.28e-05)	(2.74e-07)	(3.79e-05)	(0.0128)	(5.09e-05)
CMS quality index	4.208***	5.064***	1.040***	1.332***	-0.0115***
	(0.00994)	(0.00925)	(0.00425)	(0.00318)	(0.00260)
Brand dummies	no	no	no	no	no
Number of people	464,543	566,962	602,992	614,714	629,225

Table A8 reports the correlation coefficients between placebo plan characteristics and real plan characteristics calculated across all consumer-plan observations.

TABLE A8— CORRELATIONS BETWEEN PLACEBO AND REAL PLAN CHARACTERISTICS

	premium	OOP costs	variance	deductible	full gap coverage	generic gap coverage	cost sharing	top 100 count
premium	1.00							
OOP costs	-0.10	1.00						
variance	0.00	0.01	1.00					
deductible	-0.30	0.13	0.00	1.00				
full gap coverage	0.33	-0.05	0.00	-0.13	1.00			
generic gap coverage	0.30	-0.08	0.00	-0.30	-0.06	1.00		
cost sharing	-0.33	0.18	0.00	-0.04	0.04	-0.08	1.00	
top 100 count	0.21	-0.10	0.00	-0.09	0.08	0.09	-0.43	1.00
count 8	-0.01	-0.02	0.00	-0.08	-0.03	-0.02	-0.06	0.05
count 9	0.06	-0.03	0.00	-0.07	0.08	0.06	-0.03	0.07
count D	0.06	0.00	0.00	-0.03	0.09	-0.01	0.01	-0.05
count d	-0.01	0.00	0.00	0.04	-0.05	0.11	-0.05	0.06
count e	0.01	0.03	0.00	0.08	-0.06	0.07	0.02	0.02
count k	0.07	-0.02	0.00	-0.07	-0.11	0.14	-0.13	0.07
count l	-0.06	0.01	0.00	0.04	-0.01	-0.03	0.06	-0.03
count o	0.03	0.00	0.00	0.01	-0.03	0.11	0.01	0.04
count r	0.06	-0.01	0.00	-0.03	0.05	-0.01	0.02	-0.03
count x	-0.15	0.04	0.00	0.10	0.07	-0.29	0.15	-0.14

TABLE A9— RESULTS FROM MODELS WITH PLACEBO PLAN CHARACTERISTICS

Variable	Coefficient
Count of 8's	-0.0243***
	(0.00568)
Count of 9's	-0.0523***
	(0.00604)
Count of D's	-0.0154**
	(0.00653)
Count of d's	0.158***
	(0.00701)
Count of e's	0.0929***
	(0.00737)
Count of k's	-0.215***
	(0.00450)
Count of l's	0.0301***
	(0.00631)
Count of o's	-0.0522***
	(0.00782)
Count of r's	-0.228***
	(0.00716)
Premium (hundreds)	-0.429***
	(0.00260)
OOP costs (hundreds)	-0.108***
	(0.000571)
Variance (millions)	-5.11e-05
	(6.38e-05)
Deductible (hundreds)	-0.0250***
	(0.00315)
Full gap coverage	1.314***
	(0.0151)
Generic gap coverage	0.383***
	(0.00913)
Cost sharing	1.019***
	(0.0254)
Number of top 100 drugs on formulary	0.172***
	(0.00141)
Brand dummies	Yes
Number of people	464,543

Table A10 shows results provided to us by AG regarding the placebo test. It also compares the implied WTP for actual plan financial attributes from AG's 2011 article, our replication of them, their new results and our replication of them. The results on the financial attributes show that their new results diverge from their old ones by at least $90 for 4 of the 5 attributes. In contrast, all of ours are within $65. The lower half of the Table reports the results from their placebo attributes and our replication of their placebo model. For several reasons these results are not directly comparable to the results we report, yet they yield similar qualitative insights: first, they replaced our count variables for each alphanumeric with indicator variables for any positive count of the alphanumeric. Although this makes it impossible to isolate the marginal effects comparable to the financial attributes, to facilitate comparison we replicate their approach here. Second, they stated that they normalized these placebo attributes to zero, relative to whether a "9" is present, but they did not implement any similar normalization for the financial attributes. Third, two separate values are reported for the presence of "k", and no values are reported for the presence of an "e". Nonetheless, as with our results the test implies that these imaginary attributes influence people's PDP choices in economically meaningful ways. For example, AG's results imply that people would be willing to pay $117 more for a plan with a "d", "o" and "l" in the encrypted plan ID than for a plan with three "9s", whereas they would pay $124 for a plan with three "9s" to avoid a plan with an "8", "D" and "x". Both of these, as well as a number of other combinations, exceed the magnitude estimated for of all of the real plan attributes in AG 2011 other than full gap coverage.

TABLE A10— COMPARING ESTIMATED WILLINGNESS TO PAY FOR REAL AND PLACEBO PLAN CHARACTERISTICS FROM AG 2011 BASELINE MODEL OUR REPLICATION OF AG, AND NEW RESULTS PROVIDED BY AG

	AG 2011	Our replication		AG placebo specification		Our replication	
	WTP ($)	WTP ($)	Difference from AG 2011 ($)	WTP ($)	Difference from AG 2011	WTP ($)	Difference from AG 2011
Decreasing the deductible from $250 to $0	80	15	-65	293	213	20	-60
Covering one additional "top 100" drug	50	40	-10	9	-41	39	-11
Adding generic gap coverage	50	89	39	142	92	87	37
Increasing cost sharing from 25% to 65%	80	95	15	541	461	92	12
Adding full gap coverage	300	306	6	434	134	297	-3
Encrypted plan ID includes at least one							
“d”				60		-97	
“o”				40		-32	
“k”--result 1				29		11	
“k”--result 2				-27		--	
“l”				17		-40	
“r”				16		14	
"9"				Reference		-30	
"e"				Not provided		-60	
“8”				-7		-37	
“D”				-34		-38	
“x”				-83		-77	

TABLE A11— RESULTS FROM FIVE LARGEST REGIONS DEFINING BRAND DUMMIES BASED ON CONTRACT ID

	region 25	region 17	region 11	region 22	region 4
Estimated parameter ratios					
premium / OOP	1.3	7.3	9.7	5.7	78.9
	(0.0)	(0.2)	(0.3)	(0.2)	(5.4)
variance / premium	0.0	0.0	0.0	-0.3	0.0
	(0.0)	(0.0)	(0.0)	(0.1)	(0.0)
deductible / premium	-3.0	0.5	0.3	-0.2	0.0
	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)
full gap / premium	-2.9	-3.3	-4.3	-3.2	-4.4
	(0.1)	(0.1)	(0.0)	(0.1)	(0.0)
generic gap / premium	0.6	1.8	-0.9	-2.3	-1.8
	(0.5)	(0.1)	(0.0)	(0.1)	(0.0)
cost share / premium	-2.9	2.4	-2.8	5.0	1.8
	(0.8)	(0.1)	(0.2)	(0.2)	(0.1)
top 100 / premium	-0.3	-0.3	-0.2	-0.3	-0.1
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Welfare loss (% of costs)					
$\varepsilon \equiv 0$	101	19	25	25	19
ε is unrestricted	19	9	17	8	20
PDP Menu					
# plans	41	42	43	47	44
# brands	17	17	19	21	19
# plans w/ gap coverage	7	6	8	6	6
# plans w/ no deductible	23	25	25	27	25
Consumers					
mean age	76	78	76	76	78
% with dementia	6.2	8.2	7.8	8.6	9.4
% off cost-var frontier	74	79	76	72	82
% off cost-var-brand frontier	36	27	26	25	16
mean potential savings	621	469	543	517	606
number	46,997	37,939	30,138	29,387	24,162

Table A12 replicates the results in Table 5 after limiting the sample to white females under 80 who have not been diagnosed with Alzheimer's, dementia, or depression. See the discussion of Table 5 for additional details.

TABLE A12: RESULTS FROM MODELS IN TABLE 5 BUT WITH THE SAMPLE RESTRICTED TO WHITE FEMALES AGE<80 WITHOUT ALZHEIMER'S DISEASE OR DEMENTIA OR DEPRESSION

	United States	region 25	region 17	region 11	region 22	region 4
Decision utility parameters						
premium / OOP	3.3	1.0	5.6	4.2	1.7	-5.0
	(0.0)	(0.1)	(0.2)	(0.3)	(0.2)	(2.1)
variance / premium	0.0	0.0	0.0	-0.1	-0.2	-0.8
	(0.0)	(0.0)	(0.0)	(0.0)	(0.2)	(0.6)
deductible / premium	-0.1	-3.6	0.1	0.3	-0.7	4.9
	(0.0)	(0.3)	(0.0)	(0.1)	(0.2)	(1.8)
full gap / premium	-2.5	-2.6	-3.0	-3.3	1.6	-8.3
	(0.0)	(0.2)	(0.1)	(0.2)	(0.9)	(1.6)
generic gap / premium	-0.8	2.0	1.9	-0.1	-3.6	3.2
	(0.0)	(1.0)	(0.2)	(0.1)	(0.6)	(2.0)
cost share / premium	-2.8	-4.2	2.7	-9.1	5.7	-30.5
	(0.1)	(1.7)	(0.3)	(0.6)	(1.2)	(11.9)
top 100 / premium	-0.4	-0.3	-0.3	-0.4	-0.6	-0.1
	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)
Welfare loss (% of costs)						
$\varepsilon \equiv 0$	41	94	18	31	63	--
ε is unrestricted	7	27	8	11	4	--
Number of consumers	155,115	17,196	11,448	9,869	9,174	6,916

Table A13 reports summary statistics of the distribution of region-level results, restricted to the 24 regions with statistically significant positive estimates for the marginal utility of income. The last row reports the premium-to-OOP ratio that is predicted from an extended version of AG's DU model from equation (4) that allows the premium-to-OOP ratio to vary with the number of plans in the choice set, the number of brands, the number of plans with gap coverage, the number of plans with zero deductible, the consumer's age, and an indicator for whether the consumer is diagnosed with dementia including Alzheimer's disease. Coefficient estimates are reported in Table A16.

TABLE A13—SUMMARY STATISTICS OF THE DISTRIBUTION OF REGION-SPECIFIC ESTIMATED PARAMETER RATIOS, PDP MENU ATTRIBUTES, CONSUMER ATTRIBUTES AND NONPARAMETRIC OUTCOMES, 2006

	Mean	Standard deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
Estimated parameter ratios							
premium / OOP	4.8	2.8	1.1	2.7	4.2	6.6	12.3
variance / premium	-0.2	0.3	-0.8	-0.4	-0.2	0.0	0.6
deductible / premium	-0.4	1.8	-7.5	-0.1	0.1	0.2	1.1
full gap / premium	-3.0	1.7	-4.4	-4.2	-3.3	-2.5	0.6
generic gap / premium	-1.6	2.7	-9.5	-1.7	-1.1	-0.4	1.6
cost share / premium	-4.5	13.0	-25.8	-10.7	-6.0	-2.0	48.1
top 100 / premium	-0.4	0.2	-0.9	-0.6	-0.4	-0.3	-0.2
Welfare loss (% of costs)							
$\varepsilon \equiv 0$	42.1	24.1	17.7	25.1	33.6	54.5	109.9
ε is unrestricted	14.9	13.1	3.3	8.8	11.3	15.9	68.4
PDP Menu							
# plans	42	9	38	41	42	44	47
# brands	19	4	17	18	19	20	23
# plans w/ gap coverage	7	1	6	6	7	7	9
Consumers							
number	16,638	11,036	3,710	7,035	14,712	23,659	46,997
mean age	76	15	75	75	76	76	78
% with Alzheimer's	7	2	6	7	8	8	9
% off cost-var frontier	75	15	67	72	75	78	82
% off cost-var-brand frontier	19	6	8	16	19	23	34
mean potential savings	506	103	355	485	508	535	621
premium / oop ratio predicted by interaction model	4.1	0.6	3.3	3.8	4.0	4.2	5.9

TABLE A14—REGION-SPECIFIC ESTIMATED PARAMETER RATIOS, PDP MENU ATTRIBUTES, CONSUMER ATTRIBUTES AND NONPARAMETRIC OUTCOMES, 2006

	region 1	region 2	region 3	region 4	region 5	region 6	region 7	region 8
Estimated parameter ratios								
premium / OOP	3.2	12.3	-0.7	3.1	3.9	0.3	0.3	6.7
	(0.3)	(0.5)	(0.4)	(1.1)	(0.6)	(0.3)	(0.3)	(0.3)
variance / premium	-0.2	-0.2	5.7	0.6	-0.4	-0.3	-6.4	0.0
	(0.1)	(0.1)	(3.8)	(0.5)	(0.2)	(0.5)	(7.2)	(0.0)
deductible / premium		0.3	1.7	-7.5	0.3	-10.9	6.3	0.2
		(0.0)	(0.7)	(2.9)	(0.1)	(11.5)	(6.3)	(0.0)
full gap / premium		-4.3	-18.4	0.5	-3.0	23.1	29.8	-4.3
		(0.1)	(8.9)	(1.9)	(0.3)	(28.4)	(37.4)	(0.1)
generic gap / premium	-2.9	-1.6	1.1	-8.8	0.2	-10.6	15.0	-1.5
	(0.1)	(0.0)	(2.3)	(2.9)	(0.4)	(9.1)	(18.0)	(0.1)
cost share / premium	-25.8	-5.6	22.2	48.1	-9.7	-9.5	-107.7	-3.9
	(1.7)	(0.2)	(16.8)	(18.0)	(1.8)	(9.4)	(112.9)	(0.2)
top 100 / premium	-0.5	-0.2	0.3	-0.5	-0.6	-1.5	-4.1	-0.3
	(0.0)	(0.0)	(0.4)	(0.2)	(0.1)	(1.4)	(4.1)	(0.0)
Welfare loss (% of costs)								
$\varepsilon \equiv 0$	29	31	-193	110	34	421	403	24
ε is unrestricted	11	22	-12	68	7	23	42	13
PDP Menu								
# plans	41	44	46	44	47	52	41	38
# brands	19	22	22	20	22	23	19	17
# plans w/ gap coverage	6	7	6	6	6	8	6	7
# plans w/ zero deductible	24	28	25	25	26	30	23	22
Consumers								
number	5,729	18,248	10,661	24,162	10,570	19,417	11,148	19,447
mean age	76	77	76	78	76	76	75	75
% with Alzheimer's	7.0	7.7	7.6	9.4	7.8	8.0	7.1	6.6
% off cost-var frontier	67	76	71	82	77	74	73	68
% off cost-var-brand frontier	10	19	9	8	14	19	17	16
mean potential savings	355	491	463	606	493	475	511	533
premium / oop ratio predicted by interaction model	3.9	5.7	3.3	3.6	3.3	3.5	3.7	4.2

Note: the last row reports the premium-to-oop ratio predicted from a generalized version of AG's model that allows the ratio to vary with the proxy measures for menu complexity and cognitive ability. For more details see the explanation of Tables A13 and A16.

TABLE A14 (CONTINUED)—REGION-SPECIFIC ESTIMATED PARAMETER RATIOS, PDP MENU ATTRIBUTES, CONSUMER ATTRIBUTES AND NONPARAMETRIC OUTCOMES, 2006

	region 9	region 10	region 11	region 12	region 13	region 14	region 15	region 16
Estimated parameter ratios								
premium / OOP	0.3	7.8	3.3	6.2	5.5	2.1	3.6	2.6
	(0.3)	(0.4)	(0.2)	(0.3)	(0.4)	(0.2)	(0.2)	(0.2)
variance / premium	-11.1	-0.2	0.0	-0.4	-0.2	-0.4	-0.7	0.0
	(12.0)	(0.1)	(0.0)	(0.1)	(0.1)	(0.2)	(0.1)	(0.2)
deductible / premium	-17.3	-0.1	0.1	-0.1	0.6	0.2	-0.1	0.0
	(18.9)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.0)	(0.1)
full gap / premium	13.5	-4.4	-3.3	-4.4	-4.2	-1.2	-4.2	-2.5
	(20.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.5)	(0.1)	(0.3)
generic gap / premium	21.8	-1.1	-0.1	-1.4	-1.0	0.5	-1.2	-0.6
	(25.5)	(0.1)	(0.1)	(0.1)	(0.2)	(0.3)	(0.1)	(0.1)
cost share / premium	40.0	-2.0	-10.9	-3.5	-12.7	-14.7	-11.0	-7.6
	(44.9)	(0.2)	(0.6)	(0.2)	(0.6)	(1.5)	(0.6)	(0.6)
top 100 / premium	-1.4	-0.2	-0.5	-0.4	-0.4	-0.9	-0.5	-0.4
	(1.3)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.0)	(0.0)
Welfare loss (% of costs)								
$\varepsilon \equiv 0$	398	24	38	27	34	66	40	42
ε is unrestricted	61	12	9	16	21	9	11	7
PDP Menu								
# plans	45	42	43	41	40	43	42	45
# brands	21	19	20	19	19	20	19	19
# plans w/ gap coverage	6	7	8	6	6	7	7	9
# plans w/ zero deductible	24	24	25	23	23	25	25	29
Consumers								
number	7,650	17,268	30,138	16,928	10,389	15,932	23,832	9,340
mean age	76	75	76	75	76	76	76	75
% with Alzheimer's	7.4	8.0	7.8	7.9	7.5	7.7	7.0	6.1
% off cost-var frontier	79	79	76	71	74	72	76	78
% off cost-var-brand frontier	20	20	15	23	22	16	24	21
mean potential savings	558	495	543	522	499	495	540	427
premium / oop ratio predicted by interaction model	3.3	3.9	4.0	3.7	3.9	3.9	4.0	4.6

Note: the last row reports the premium-to-oop ratio predicted from a generalized version of AG's model that allows the ratio to vary with the proxy measures for menu complexity and cognitive ability. For more details see the explanation of Tables A13 and A16.

TABLE A14 (CONTINUED)—REGION-SPECIFIC ESTIMATED PARAMETER RATIOS, PDP MENU ATTRIBUTES, CONSUMER ATTRIBUTES AND NONPARAMETRIC OUTCOMES, 2006

	region 17	region 18	region 19	region 20	region 21	region 22	region 23	region 24
Estimated parameter ratios								
premium / OOP	7.0	4.4	-0.2	4.7	2.2	1.9	8.1	9.9
	(0.2)	(0.3)	(0.3)	(0.4)	(0.6)	(0.1)	(0.6)	(0.6)
variance / premium	0.0	-0.6	0.0	-0.3	0.0	-0.8	-0.2	-0.4
	(0.0)	(0.1)	(0.0)	(0.1)	(0.3)	(0.2)	(0.1)	(0.1)
deductible / premium	0.0	0.2	15.0	0.0	-1.6	-0.5	0.2	0.1
	(0.0)	(0.0)	(25.6)	(0.1)	(0.7)	(0.1)	(0.0)	(0.0)
full gap / premium	-3.5	-3.4	-46.6	-3.5	-3.3	0.6	-3.7	-4.0
	(0.1)	(0.1)	(74.3)	(0.2)	(0.7)	(0.5)	(0.1)	(0.1)
generic gap / premium	1.6	-1.1		-1.8	-9.5	-2.8	-0.3	-0.8
	(0.1)	(0.1)		(0.2)	(2.9)	(0.3)	(0.1)	(0.1)
cost share / premium	3.8	-6.8	32.5	-1.3	3.2	5.0	-2.5	-2.2
	(0.2)	(0.4)	(58.0)	(0.5)	(2.4)	(0.6)	(0.3)	(0.2)
top 100 / premium	-0.4	-0.4	3.5	-0.3	-0.6	-0.7	-0.4	-0.2
	(0.0)	(0.0)	(6.4)	(0.0)	(0.1)	(0.0)	(0.0)	(0.0)
Welfare loss (% of costs)								
$\varepsilon \equiv 0$	18	31	-295	24	40	62	20	23
ε is unrestricted	9	11	-50	10	8	3	10	16
PDP Menu								
# plans	42	41	40	38	39	47	42	40
# brands	18	18	17	17	18	22	18	17
# plans w/ gap coverage	6	6	6	6	6	6	7	7
# plans w/ zero deductible	25	25	24	23	24	27	25	25
Consumers								
number	37,939	13,492	7,317	6,785	4,265	29,387	6,880	7,499
mean age	78	76	75	75	75	76	76	77
% with Alzheimer's	8.2	7.9	8.5	6.9	7.8	8.6	7.6	7.5
% off cost-var frontier	79	73	79	74	75	72	79	78
% off cost-var-brand frontier	15	21	30	26	19	18	24	24
mean potential savings	469	491	591	520	547	517	520	536
premium / oop ratio predicted by interaction model	4.0	4.1	4.1	4.3	4.2	3.5	4.1	4.5

Note: the last row reports the premium-to-oop ratio predicted from a generalized version of AG's model that allows the ratio to vary with the proxy measures for menu complexity and cognitive ability. For more details see the explanation of Tables A13 and A16.

TABLE A14 (CONTINUED)—REGION-SPECIFIC ESTIMATED PARAMETER RATIOS, PDP MENU ATTRIBUTES, CONSUMER ATTRIBUTES AND NONPARAMETRIC OUTCOMES, 2006

	region 25	region 26	region 27	region 28	region 29	region 30	region 31	region 32
Estimated parameter ratios								
premium / OOP	1.1 (0.0)	-0.2 (0.9)	4.5 (0.6)	1.5 (0.7)	1.8 (2.0)	3.9 (0.4)	-2.2 (0.7)	5.6 (0.4)
variance / premium	0.0 (0.0)	0.9 (7.1)	-0.8 (0.2)	0.1 (0.3)	-3.1 (3.6)	0.0 (0.0)	2.7 (1.0)	0.0 (0.0)
deductible / premium	-3.6 (0.2)	6.5 (22.6)	0.2 (0.1)	-0.7 (0.6)	0.8 (0.5)	1.1 (0.1)	1.5 (0.4)	0.5 (0.0)
full gap / premium	-2.9 (0.1)	-45.5 (147.7)	-3.1 (0.3)	-0.3 (2.3)	1.8 (7.3)	-2.3 (0.3)	-9.1 (1.5)	-3.1 (0.1)
generic gap / premium	0.8 (0.6)		-0.7 (0.2)	-2.9 (1.3)	0.3 (1.4)	-0.8 (0.3)		-1.1 (0.1)
cost share / premium	-3.2 (0.9)	-29.3 (114.5)	-6.4 (0.8)	-10.1 (5.0)	-25.8 (25.8)	-22.1 (1.9)	0.3 (2.1)	-6.4 (0.4)
top 100 / premium	-0.4 (0.0)	62.6 (225.2)	-0.3 (0.0)	-0.7 (0.3)	-0.7 (0.7)	-0.6 (0.0)	0.3 (0.2)	-0.4 (0.0)
Welfare loss (% of costs)								
$\varepsilon \equiv 0$	92	-324	33	79	109	59	-86	31
ε is unrestricted	24	-118	10	4	9	32	-29	15
PDP Menu								
# plans	41	43	43	43	44	45	44	47
# brands	23	19	19	20	20	22	20	20
# plans w/ gap coverage	7	6	7	6	7	6	6	7
# plans w/ zero deductible	23	26	26	25	25	25	23	28
Consumers								
number	46,997	1,587	3,710	4,926	1,703	12,314	5,594	23,141
mean age	76	75	76	75	75	76	76	76
% with Alzheimer's	6.2	6.9	7.4	5.7	6.4	6.6	6.3	6.6
% off cost-var frontier	74	74	73	69	78	74	76	78
% off cost-var-brand frontier	34	13	18	13	17	18	20	17
mean potential savings	621	414	468	444	510	483	532	521
premium / oop ratio predicted by interaction model	5.9	3.9	4.1	3.7	3.7	3.4	3.4	3.8

Note: the last row reports the premium-to-oop ratio predicted from a generalized version of AG's model that allows the ratio to vary with the proxy measures for menu complexity and cognitive ability. For more details see the explanation of Tables A13 and A16.

FIGURE A3—RATIO OF PREMIUM-TO-OOP COEFFICIENTS IN 2006, BY CMS REGION USING AG'S DEFINITION OF BRAND DUMMIES BASED ON CONTRACT ID

Note: The figure reports the premium-to-OOP coefficient ratio obtained by estimating region-specific models with contract id dummies. The econometric specification is the same as the national model in column 2 of Table 4. In regions with light-shaded numbers, we fail to reject the null hypothesis that the marginal utility of income is negative at the 5% level. All estimates are statistically indistinguishable from 1 at the 5% level.

Table A15 provides the coefficients and standard errors from meta-regressions of the conditional relationship between the premium-to-OOP ratio and proxy measures for menu complexity and cognitive ability. The models are limited to the 24 regions with statistically significant positive estimates for the marginal utility of income. The main text provides additional details.

TABLE A15: RESULTS FROM MODELS OF THE REGION-LEVEL ESTIMATES FOR AG'S PARAMETRIC MEASURES OF CHOICE QUALITY ON PROXY MEASURES FOR MENU COMPLEXITY AND COGNITIVE ABILITY

	(1)	(2)	(3)	(4)	(5)
				% welfare loss	
	premium-to-OOP coef ratio			$\varepsilon \equiv 0$	$\varepsilon \neq 0$
Number of plans	-0.508		-0.559	1.100	0.707
	(0.553)		(0.577)	(4.420)	(2.386)
Number of brands	-0.234		-0.214	7.613*	1.082
	(0.521)		(0.553)	(4.241)	(2.289)
Number of plans w/ gap coverage	-0.284		0.0865	0.837	0.703
	(0.896)		(0.943)	(7.227)	(3.901)
Number of plans w/ zero deductible	0.858		0.802	-4.213	-2.574
	(0.710)		(0.725)	(5.555)	(2.998)
mean age		0.706	0.811	8.567	10.43**
		(0.960)	(1.004)	(7.697)	(4.154)
% with Alzheimer's		0.277	0.445	-3.410	0.129
		(0.861)	(0.947)	(7.261)	(3.919)
Constant	11.33	-50.83	-52.80	-678.1	-769.2**
	(10.30)	(69.39)	(71.95)	(551.5)	(297.7)
Observations	24	24	24	24	24
R^2	0.151	0.061	0.239	0.404	0.422
Adjusted R^2	-0.028	-0.028	-0.029	0.193	0.218
P-value of model Wald Chi-Square	0.514	0.515	0.523	0.136	0.111

Note: *** p<0.01, ** p<0.05, * p<0.1

To check robustness of the results from the meta-regression in equation (9) we estimate AG's DU model (4) after adding interactions between the OOP ratio and the proxy measures for menu complexity and cognitive ability. This logit model accounts for variation in menu complexity across CMS regions and for variation in cognitive ability within and across CMS regions. Each of the interaction coefficients is statistically significant at the 1% level. To evaluate their economic magnitudes we use the estimates to predict how the premium-to-oop ratio would change as we move from the lowest value of each variable observed in our data to the highest value, while evaluating all other variables at their means. The resulting ranges are reported in the last two columns of Table A16. For example, the results in the first row of the table imply that *increasing* the number of plans in a consumer's choice set from 38 plans to 52 plans would *decrease* the premium-to-oop ratio from 4.6 to 2.9, contrary to the hypothesis of choice overload.

TABLE A16: ESTIMATED EFFECTS OF PROXY MEASURES FOR MENU COMPLEXITY AND COGNITIVE ABILITY ON THE PREMIUM-TO-OOP RATIO

	Summary statistics			Econometric estimates		Premium-to-OOP ratio	
	mean	min	max	interaction with OOP	standard error	predicted at the Min	predicted at the Max
Number of plans in choice set	43.15	38	52	-0.0035	(0.0004)	4.6	2.9
Number of brands in choice set	20.05	17	23	-0.0014	(0.0003)	4.0	3.7
Number of plans with gap coverage	6.67	6	9	0.0033	(0.0007)	3.7	4.1
Number of plans with zero deductible	25.08	22	30	0.0043	(0.0005)	3.4	4.7
Age	76.04	66	108	0.0002	(0.0001)	3.7	4.0
Dementia including Alzeheimer's	0.08	0	1	-0.0102	(0.0014)	3.8	3.5

Note: The estimated coefficients on premium and OOP are -0.406 and -0.066 respectively. Both have p-values of zero out to four decimal places.

Table A17 provides results from validation tests for the cases where the set of brands in an estimation region spans the set of brands in the prediction region. Two pairs of regions meet this criterion in 2006. As a result, we estimate the AG's two competing models for region 14 and then use the resulting coefficients

to predict outcomes in region 15, and we use estimates for region 30 to predict outcomes in region 28. Both region pairs are similar in their consumer populations and PDP menu complexity. AG's EUM model yields closer out-of-sample predictions than their DU model in every case but one. The shading indicates which prediction is closer to the data.

TABLE A17: RESULTS FROM BETWEEN-REGION VALIDATION TESTS FOR THE ONLY TWO PAIRS OF REGIONS IN 2006 FOR WHICH ONE REGION'S BRANDS ARE NESTED WITHIN THE OTHER'S

	region 14 → 15			region 30 → 28		
	data	AG's DU	AG's EUM	data	AG's DU	AG's EUM
	In-sample data and predictions					
	region 14			region 30		
Percent of consumers choosing:						
gap coverage	11	11	13	6	6	9
dominated plan	16	18	17	18	19	17
min cost plan within brand	47	40	41	41	35	40
Median consumer expenditures ($)						
premium + OOP	1,261	1,262	1,267	1,074	1,093	1,095
overspending on dominated plans	0	65	58	0	63	58
Market concentration						
Hirfindahl-Hirschman index	25	25	25	25	25	25
market share of top brand	44	44	44	39	39	39
	Out-of-sample data and predictions					
	region 15			region 28		
Percent of consumers choosing:						
gap coverage	9	6	10	18	11	13
dominated plan	12	25	19	24	18	17
min cost plan within brand	50	32	40	42	40	41
Median consumer expenditures ($)						
premium + OOP	1,096	1,205	1,178	1,418	1,352	1,355
overspending on dominated plans	0	102	70	0	67	57
Market concentration						
Hirfindahl-Hirschman index	44	30	31	0 21	27	26
market share of top brand	62	42	46	0 31	45	44

Table A18 provides results from the national validation test shown in Table 6 except using the root mean square error in predictions across regions in place of the mean absolute error.

TABLE A18— NONRANDOM HOLDOUT SAMPLE TESTS OF MODEL VALIDATION, 2006

		In-sample fit		Out-of-sample fit	
		AG's DU	AG's EUM	AG's DU	AG's EUM
	data	RMSE		RMSE	
		Using CMS Star Ratings for Quality			
Percent of consumers choosing:					
gap coverage	13	0	7	9	7
dominated plan	20	5	6	8	7
min cost plan within brand	52	8	9	11	9
Median consumer expenditures ($)					
premium + OOP	1,255	16	46	113	88
overspending on dominated plans	0	72	55	70	51
Market concentration					
Hirfindahl-Hirschman index	25	11	15	11	15
market share of top brand	37	12	18	14	18
		Using Brand Indicators for Quality			
Percent of consumers choosing:					
gap coverage	13	0	4	9	8
dominated plan	20	2	4	9	9
min cost plan within brand	52	6	9	14	13
Median consumer expenditures ($)					
premium + OOP	1,256	16	21	102	101
overspending on dominated plans	0	82	71	88	70
Market concentration					
Hirfindahl-Hirschman index	25	0	0	14	13
market share of top brand	37	0	0	18	17

Note: RMSE refers to the root mean square error between the regional-level model predictions and data, weighted across regions by the number of people in the sample in the region. The results are based on every possible pairwise combination of regions in 2006 except that they exclude regions 33 and 34 (HI and AK), and the lower half also excludes region 26 (NM). Thus the values in the top half are based on the results from all 992 of the possible regional out-of-sample predictions while those in the lower half are based on 930 of them.

Table A19 provides results from the national validation test suggested to us by Abaluck and Gruber. Specifically, we estimate the models using the 2006 data from 31 regions and use it to predict a single out-of-sample region, repeated using each of the 32 regions as the holdout region (excluding Alaska and Hawaii). This is very similar to an in-sample validation test as the set of plans and plan attributes in the single out-of-sample region is typically very close to being nested within the in-sample set (see Keane and Wolpin 2007). As before the measures of market concentration are defined at the region level as that is the policy-relevant market definition. Hence while the models with brand indicators match the market concentration perfectly across the 31 in-sample regions in each of the 32 separate tests (yielding a mean absolute deviation of 0), they do not perfectly predict the region-level market concentration for any single given in-region sample.

TABLE A19— RESULTS FROM THE NATIONAL MODEL VALIDATION TESTS SUGGESTED BY ABALUCK AND GRUBER

		In-sample fit		Out-of-sample fit	
		AG's DU	AG's EUM	AG's DU	AG's EUM
	data	|model error|		|model error|	
		Using CMS Star Ratings for Quality			
Percent of consumers choosing:					
gap coverage	13	0	2	6	6
dominated plan	20	2	2	6	6
min cost plan within brand	52	3	4	6	5
Median consumer expenditures ($)					
premium + OOP	1,255	4	15	61	60
overspending on dominated plans	0	63	51	63	51
Market concentration					
Hirfindahl-Hirschman index	25	11	14	11	14
market share of top brand	37	9	17	11	17
		Using Brand Indicators for Quality			
Percent of consumers choosing:					
gap coverage	13	0	2	6	6
dominated plan	20	1	1	5	5
min cost plan within brand	52	7	8	8	9
Median consumer expenditures ($)					
premium + OOP	1,255	10	14	61	57
overspending on dominated plans	0	64	61	71	63
Market concentration					
Hirfindahl-Hirschman index	25	4	4	5	4
market share of top brand	37	6	6	8	7

Note: | Model error | refers to the mean absolute deviation between the model predictions and data.