# What Do Consumers Consider Before They Choose?

Abi Adams
Oxford, IFS, CEP

Jason Abaluck
Yale, NBER

October 2019

# Introduction

- When estimating consumer demand models we usually assume that consumers consider all the alternatives that we as the researcher see
- Lots of evidence that the assumption of full consideration is violated in reality for many applications of interest
- We should care about this for a number of reasons, e.g:
  - Cannot predict the impact/evaluate the benefits of making consumers aware of a wider set of alternatives
  - Biased estimates of preference parameters with implications for welfare analysis

# Introduction

- An exception: the literature on "consideration sets" — consumers might only consider an (unobserved) subset of alternatives
- Popular in marketing and a growing applied literature for providing a "simple" way to introduce unobserved choice sets
  - (Behavioural) decision theory provides a rich set of models: see, e.g. Masatlioglu et al (2012) and Cattaneo et al (2018)
  - **Default specific**: Ho, Hogan & Scott-Morton (2016); Heiss, McFadden, Winter, Wuppermann & Zhou (2016); Moshkin & Shachar (2002)
  - **Alternative specific**: Goeree (2008); Manzini and Mariotti (2012); Conlin and Mortimer (2013); Honka et al (2015); Gaynor, Propper & Seiler (2016)

# Challenge

- Wider application of these models has been held back by the difficulty of separately identifying “utility” and “consideration probability” parameters from observational data
- Two main strategies pursued to date:
  1. **Auxiliary data:** can we collect additional data on what options consumers considered?
  2. **Exclusion restrictions:** are there exogenous variables excluded from utility and from process generating consideration?

# This Paper

- In this paper we show that the restrictions from economic theory are sufficient for identification in many applied settings of interest
- Our approach relies on exploiting asymmetries in the "Slutsky" matrix
  - Changes in the characteristics of products impact the probability that you consider a good and not just utility
  - There is a particular pattern of cross-price asymmetries and violations of nominal illusion that are characteristic of a lack of consideration
  - Inspired by the theoretical work of Gabaix (2014) on inattention to characteristics although our focus is on inattention to goods

# This Paper

- Different strategy to that pursued in other current working papers on identification of consideration set models:
  - **Crawford, Griffith & Iaria**: results specific to Logit errors and rely on some assumptions about stability of choice sets over time
  - **Dardanoni, Manzini, Mariotti & Tyson**: limited allowance for preference heterogeneity
  - **Cattaneo, Ma, Masatlioglu & Suleymanov**: deterministic preferences but weaker assumptions on consideration
  - **Barseghyan, Coughlin, Molinari & Teitelbaum**: weaker assumptions on preference heterogeneity and consideration leading to set identification results

# This Paper

- Bring a parametric version of our framework to data to show that the variation at heart of our identification result is important for driving empirical results
  - Indirect inference estimator in which auxiliary model allows for cross derivative asymmetries
  - Structural parameters chosen to match the reduced form asymmetries
- **Lab validation**: can we recover the process generating consideration sets from choice data?
- **Medicare Part D**: to what extent is inertia driven by switching costs or lack of consideration?
  - Used to evaluate a proposed ‘active default’ policy

# Outline

# Basic Set-Up: Preferences

- Imagine that we are in a full-information environment
- Consumer $i$ selects the good $0, ..., J$ that gives her the highest utility
- Utility is a function of a good's characteristics, $\mathbf{x} \in \mathbb{R}^K$, plus a random error

$$\begin{aligned} u_{ij} &= v_j(\mathbf{x}_j) + \epsilon_{ij} \\ &= \beta p_j + w_j(\mathbf{z}_j) + \epsilon_{ij} \end{aligned}$$

- Here assume quasi-linearity but show can be (partially) relaxed within main paper
- Proof extends naturally to allow for individual heterogeneity through a random coefficient

$$u_{ij} = \beta_i p_j + w_j(\mathbf{z}_j) + \epsilon_{ij}$$

# Basic Set-Up: Preferences

- Imagine that we are in a full-information environment
- Consumer $i$ selects the good $0, ..., J$ that gives her the highest utility
- Utility is a function of a good's characteristics, $\mathbf{x} \in \mathbb{R}^K$, plus a random error

$$\begin{aligned} u_{ij} &= v_j(\mathbf{x}_j) + \epsilon_{ij} \\ &= \beta p_j + w_j(\mathbf{z}_j) + \epsilon_{ij} \end{aligned}$$

- Here assume quasi-linearity but show can be (partially) relaxed within main paper
- Proof extends naturally to allow for individual heterogeneity through a random coefficient

$$u_{ij} = \beta_i p_j + w_j(\mathbf{z}_j) + \epsilon_{ij}$$

# Basic Set-Up: Preferences

- The probability that a consumer chooses some good $j$ is then:

$$
\begin{aligned}
Pr(i \text{ chooses } j) &= Pr\left(u_{ij} > u_{ij'} \quad \forall j' \neq j\right) \\
s_j^{\star} &= Pr\left(\epsilon_{ij'} < v_j + \epsilon_{ij} - v_{j'} \quad \forall j' \neq j\right)
\end{aligned}
$$

- Example: when $\epsilon_{ij}$ is distributed Type 1 Extreme Value, we get the popular logit model

$$
s_j^{\star} = \frac{exp\left(v_j\right)}{\sum_{j'=1}^{J} exp\left(v_{j'}\right)}
$$

- NB We allow for correlated unobservables in utility!

# Basic Set-Up: Consideration

- A consumer may not consider all goods in her choice set
- Good-0 represents an "inside" or "outside" default
- Let $\mathcal{P}(J)$ represent the power set of all goods, with any given element indexed by $C$
- Set of consideration sets containing good $j$ is given as:

$$\mathbb{P}(j) = \{C : C \in \mathcal{P}(J) \quad \& \quad j \in C \quad 0 \in C\}$$

# Basic Set-Up: Consideration

- Need some restrictions on consideration probabilities to achieve identification
- Two main classes of consideration set model found in the applied literature:
  - **Default specific**: with some probability $\mu(\mathbf{x}_0)$ you consider the full choice set, otherwise you only consider a (known) default option
  - **Alternative specific**: you consider good $j$ with probability $\phi_j(\mathbf{x}_j)$
- We consider a general framework that subsumes both of these classes of model
- NB throughout this presentation will be assuming independence of unobservables driving utility and consideration

# Basic Set-Up: Choice Probabilities

- In our model, observed choice probabilities take the form:

$$s_0 = (1 - \mu) +$$

# Basic Set-Up: Choice Probabilities

- In our model, observed choice probabilities take the form:

$$s_0 = (1 - \mu) + \mu$$

# Basic Set-Up: Choice Probabilities

- In our model, observed choice probabilities take the form:

$$s_0 = (1-\mu) + \mu \sum_{C \in \mathbb{P}(0)} \prod_{l \in C} \phi_l \prod_{l' \notin C} (1-\phi_{l'}) s_0^\star(C)$$

# Basic Set-Up: Choice Probabilities

- In our model, observed choice probabilities take the form:

$$s_0 = (1-\mu) + \mu \sum_{C \in \mathbb{P}(0)} \prod_{l \in C} \phi_l \prod_{l' \notin C} (1-\phi_{l'})\, s_0^{\star}(C)$$

# Basic Set-Up: Choice Probabilities

- In our model, observed choice probabilities take the form:

$$s_0 = (1-\mu) + \mu \sum_{C \in \mathbb{P}(0)} \prod_{l \in C} \phi_l \prod_{l' \notin C} (1-\phi_{l'})\, s_0^\star(C)$$

$$s_j = \mu \sum_{C \in \mathbb{P}(j)} \prod_{l \in C} \phi_l \prod_{l' \notin C} (1-\phi_{l'})\, s_j^\star(C)$$

for $j > 0$

# Extensions

- Dependence of $\phi_j$ on the characteristics of the default product
- Independence of unobservables influencing utility and attention implicit in the background
  - Consider case of finite set of “types”
  - Require exclusion restrictions for identification but fewer than if ignored results in this paper
- Asymmetries and nominal illusion results that we will now develop imply imperfect consideration in wider class of models

# Outline

# Symmetry

- With full consideration, choice probabilities will satisfy a symmetry restriction

$$\frac{\partial s_j^\star}{\partial p_k} = \frac{\partial s_k^\star}{\partial p_j}$$

- They will also satisfy absence of nominal illusion

$$s_j^\star(\mathbf{p}) = s_j^\star(\mathbf{p} + \delta)$$

- Given our assumptions on preferences, this result holds with correlation in unobserved tastes across products and in the mixed logit model More

# Symmetry

- With consideration sets, symmetry is violated and we suffer from nominal illusion

$$\frac{\partial s_j}{\partial p_k} \neq \frac{\partial s_k}{\partial p_j}$$
$$s_j^\star(\mathbf{p}) \neq s_j^\star(\mathbf{p} + \delta)$$

- Changes in characteristics do not just impact utility, but also the probability of paying attention to particular subsets of goods

- We can use these asymmetries to identify attention probabilities, $\mu(p_0)$ and $\phi_j(p_j)$

# Proof: Special Case

- For purposes of this presentation, will walk through the proof of a special case of our more general framework
- Default specific model: $\phi_j = 1$ for all $j$
- Choice probabilities take the form

$$\begin{aligned} s_0 &= (1-\mu) + \mu s_0^\star \\ s_j &= \mu s_j^\star \end{aligned}$$

where $s_j^\star \equiv s_j^\star (\mathbf{p} | \{0, ..., J\})$.

# Default Specific Consideration

- Changes in the characteristics of the default have two impacts on non-default goods:

$$\frac{\partial s_j}{\partial p_0} = \mu \frac{\partial s_j^\star}{\partial p_0}$$

# Default Specific Consideration

- Changes in the characteristics of the default have two impacts on non-default goods:

$$\frac{\partial s_j}{\partial p_0} = \mu \frac{\partial s_j^\star}{\partial p_0} + s_j^\star \frac{\partial \mu}{\partial p_0}$$

# Default Specific Consideration

- Cross derivative differences take the form:

$$
\begin{aligned}
\frac{\partial s_j}{\partial p_0} - \frac{\partial s_0}{\partial p_j} &= s_j^\star \frac{\partial \mu}{\partial p_0} + \mu \left( \frac{\partial s_j^\star}{\partial p_0} - \frac{\partial s_0^\star}{\partial p_j} \right) \\
&= s_j^\star \frac{\partial \mu}{\partial p_0} \\
&= s_j^\star \frac{\partial \mu}{\partial p_0} \frac{\mu}{\mu} \\
&= \frac{s_j^\star}{\mu} \frac{\partial \log(\mu)}{\partial p_0} \\
&= s_j \frac{\partial \log(\mu)}{\partial p_0}
\end{aligned}
$$

# Default Specific Consideration

- Cross derivative differences take the form:

$$
\begin{aligned}
\frac{\partial s_j}{\partial p_0} - \frac{\partial s_0}{\partial p_j} &= s_j^\star \frac{\partial \mu}{\partial p_0} + \mu \left( \frac{\partial s_j^\star}{\partial p_0} - \frac{\partial s_0^\star}{\partial p_j} \right) \\
&= s_j^\star \frac{\partial \mu}{\partial p_0} \\
&= s_j^\star \frac{\partial \mu}{\partial p_0} \frac{\mu}{\mu} \\
&= \frac{s_j^\star}{\mu} \frac{\partial \log(\mu)}{\partial p_0} \\
&= s_j \frac{\partial \log(\mu)}{\partial p_0}
\end{aligned}
$$

# Default Specific Consideration

- Cross derivative differences take the form:

$$
\begin{aligned}
\frac{\partial s_j}{\partial p_0} - \frac{\partial s_0}{\partial p_j} &= s_j^\star \frac{\partial \mu}{\partial p_0} + \mu \left( \frac{\partial s_j^\star}{\partial p_0} - \frac{\partial s_0^\star}{\partial p_j} \right) \\
&= s_j^\star \frac{\partial \mu}{\partial p_0} \\
&= s_j^\star \frac{\partial \mu}{\partial p_0} \frac{\mu}{\mu} \\
&= \frac{s_j^\star}{\mu} \frac{\partial \log(\mu)}{\partial p_0} \\
&= s_j \frac{\partial \log(\mu)}{\partial p_0}
\end{aligned}
$$

# Default Specific Consideration

- Cross derivative differences take the form:

$$\begin{aligned}
\frac{\partial s_j}{\partial p_0} - \frac{\partial s_0}{\partial p_j} &= s_j^\star \frac{\partial \mu}{\partial p_0} + \mu \left( \frac{\partial s_j^\star}{\partial p_0} - \frac{\partial s_0^\star}{\partial p_j} \right) \\
&= s_j^\star \frac{\partial \mu}{\partial p_0} \\
&= s_j^\star \frac{\partial \mu}{\partial p_0} \frac{\mu}{\mu} \\
&= \frac{s_j^\star}{\mu} \frac{\partial \log(\mu)}{\partial p_0} \\
&= s_j \frac{\partial \log(\mu)}{\partial p_0}
\end{aligned}$$

# Default Specific Consideration

- Cross derivative differences take the form:

$$\begin{aligned}
\frac{\partial s_j}{\partial p_0} - \frac{\partial s_0}{\partial p_j} &= s_j^\star \frac{\partial \mu}{\partial p_0} + \mu \left( \frac{\partial s_j^\star}{\partial p_0} - \frac{\partial s_0^\star}{\partial p_j} \right) \\
&= s_j^\star \frac{\partial \mu}{\partial p_0} \\
&= s_j^\star \frac{\partial \mu}{\partial p_0} \frac{\mu}{\mu} \\
&= \frac{s_j^\star}{\mu} \frac{\partial \log(\mu)}{\partial p_0} \\
&= s_j \frac{\partial \log(\mu)}{\partial p_0}
\end{aligned}$$

# Default Specific Consideration

- Changes in consideration probabilities constructively identified by cross-derivative differences:

$$\frac{\partial \log(\mu)}{\partial p_0} = \frac{1}{s_j}\left[\frac{\partial s_j}{\partial p_0} - \frac{\partial s_0}{\partial p_j}\right]$$

- Get the level of attention by integrating over the support of characteristics and pinning down the constant at point of symmetry

$$\mu = exp\left(-\int \frac{1}{s_j}\left[\frac{\partial s_j}{\partial p_0} - \frac{\partial s_0}{\partial p_j}\right] \mathrm{d}p_0\right)$$

# Full Proof: Sketch

- Choice probabilities take the form

$$s_0 = (1-\mu) + \mu \sum_{C \in \mathbb{P}(0)} \prod_{l \in C} \phi_l \prod_{l' \notin C} (1-\phi_{l'})\, s_0^\star(C)$$

$$s_j = \mu \sum_{C \in \mathbb{P}(j)} \prod_{l \in C} \phi_l \prod_{l' \notin C} (1-\phi_{l'})\, s_j^\star(C)$$

  for $j > 0$

- Need further source of variation in this model, with slight abuse of notation:
  - $s_j(\mathcal{J}/j')$: market share of $j$ when $j'$ not available
  - NB Similar to Kawaguchi et al (MS, 2016) but without additional exclusion restriction
  - Can also express in terms of a full support assumption — required for nonparametric identification of RUM

# Full Proof: Sketch

- Choice probabilities take the form

$$s_0 = (1-\mu) + \mu \sum_{C \in \mathbb{P}(0)} \prod_{l \in C} \phi_l \prod_{l' \notin C} (1-\phi_{l'})\, s_0^\star(C)$$

$$s_j = \mu \sum_{C \in \mathbb{P}(j)} \prod_{l \in C} \phi_l \prod_{l' \notin C} (1-\phi_{l'})\, s_j^\star(C)$$

  for $j > 0$

- Need further source of variation in this model, with slight abuse of notation:
  - $s_j(\mathcal{J}/j')$: market share of $j$ when $j'$ not available
  - NB Similar to Kawaguchi et al (MS, 2016) but without additional exclusion restriction
  - Can also express in terms of a full support assumption — required for nonparametric identification of RUM

# Full Proof: Sketch

- Changes in consideration probabilities are the unique solution to a system of linear equations:

$$\frac{\partial s_j}{\partial p_0} - \frac{\partial s_0}{\partial p_j} = \frac{\partial \log(\mu)}{\partial p_j} s_j + \frac{\partial \log(\phi_j)}{\partial p_j} \left( s_0(\mathcal{J}/j) - s_0 \right)$$

  where $s_j = s_j(\mathbf{x}|\mathcal{J})$.

- Final piece of puzzle: use nominal illusion to identify latent market shares

# Full Proof: Sketch

- Changes in consideration probabilities are the unique solution to a system of linear equations:

$$\frac{\partial s_j}{\partial p_0} - \frac{\partial s_0}{\partial p_j} \quad = \quad \frac{\partial \log(\mu)}{\partial p_j} s_j + \frac{\partial \log(\phi_j)}{\partial p_j} \left( s_0(\mathcal{J}/j) - s_0 \right)$$

  where $s_j = s_j\left(\mathbf{x}|\mathcal{J}\right)$.

- Final piece of puzzle: use nominal illusion to identify latent market shares

# Outline

I Basic Set-Up

II Asymmetry-Based Identification

III **Estimation**

IV Experimental Validation

V Field Application

# Estimation

- Identification results constructive and so, in theory, consistent nonparametric estimators could be based on them
- However, in practice, nonparametric estimation is infeasible given the dimensionality problems
- Place a set of functional form assumptions on utility and the process driving consideration that are consistent with our framework
- We estimate special cases of our general framework in two scenarios, showing that asymmetries important for driving the ultimate results

# Estimation: Application Assumptions

- Functional form assumptions a simple version of those followed in marketing literature and Goeree (2008):

$$s_j^\star(C) = \frac{exp\left(\alpha_j + x_j\beta\right)}{\sum_{j'\in C} exp\left(\alpha_{j'} + x_{j'}\beta\right)}$$

$$\phi_j = \frac{exp\left(\delta_j + x_j\gamma\right)}{1 + exp\left(\delta_j + x_j\gamma\right)}$$

$$\mu = \frac{exp\left(\delta_0 + x_0\omega\right)}{1 + exp\left(\delta_0 + x_0\omega\right)}$$

- Typical to estimate the parameters of the parametric model by maximum (simulated) likelihood (e.g. Goeree 2008)

# Indirect Inference

- We instead pursue an estimation strategy that is grounded in the identifying variation at the heart of our identification proof
- Estimate the model by indirect inference
  - Match the parameters of a flexible auxiliary model that is able to capture cross-derivative asymmetries in the data
  - Intuitively, if estimate the auxiliary model on data simulated from the 'true' DGP, should get the same parameters as when estimating the auxiliary model on simulated data

# Applications

1. **Lab: Choice Experiment**
   - Alternative to a simulation exercise: we know that the model is misspecified
   - We set the process generating which subset of 10 goods a respondent considers in a choice experiment
   - Can we recover the parameters of this process without using information on what options a respondent considered?
2. **Field: Medicare Part D Choice**
   - Recent set of papers looking to disentangle switching costs and inattention in insurance choices
   - Are the exclusion restrictions employed valid?
   - Evaluation on an "active default" policy

# Applications

1. **Lab: Choice Experiment**
   - Alternative to a simulation exercise: we know that the model is misspecified
   - We set the process generating which subset of 10 goods a respondent considers in a choice experiment
   - Can we recover the parameters of this process without using information on what options a respondent considered?
2. **Field: Medicare Part D Choice**
   - Recent set of papers looking to disentangle switching costs and inattention in insurance choices
   - Are the exclusion restrictions employed valid?
   - Evaluation on an “active default” policy

# Choice Experiment

- Endowed respondents with $25 and asked them to select their most preferred option from a set of goods that appeared on their screen
- 10 goods in full choice set chosen from Yale Bookstore with the price randomly drawn
- We set the probability that a particular good showed up on a respondent *i*'s screen in round *r* as:

$$\phi_j(p_{ijr}) = \frac{exp\left(\delta_j + \gamma p_{ijr}\right)}{1 + exp\left(\delta_j + \gamma p_{ijr}\right)}$$

- Can we recover the (known) $\delta_j$ and $\gamma$?

# Choice Experiment



**Collegiate Pacific Banner ("Yale University Lux et Veritas")**
**$8.00**

**Embroidered Towel From Team Golf**
**$20.00**

**Mug w/ Thumb Piece**
**$11.00**


**LXG Power Bank**
**$12.00**

**Moleskin Large Notebook with Debossed Wordmark, Unruled**
**$23.00**

(You must wait 10 seconds before clicking next to make sure you consider all options)

Next

# Choice Experiment

- Auxiliary model specified as a flexible logit with good specific parameters:

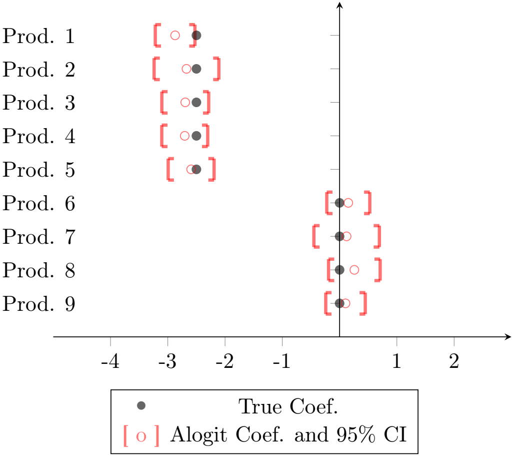
$$v_{ijr} = \omega_j + \theta_p p_{ijr} + \sum_{j'=0}^{J} \theta_{jj'} p_{ijr} p_{ij'r}$$

$$\widetilde{s}_{ijr} = \frac{\exp\left(v_{ijr}\right)}{\sum_{j'} \exp\left(v_{ij'r}\right)}$$

- Estimator of structural utility and consideration parameters, $\psi = [\delta, \gamma, \alpha, \beta]$, defined as:

$$\widehat{\psi} = \arg\min_{\psi} \left(\widehat{\theta}^t - \widehat{\theta}^s(\psi)\right)' W \left(\widehat{\theta}^t - \widehat{\theta}^s(\psi)\right)$$

# Results: Attention Fixed Effects

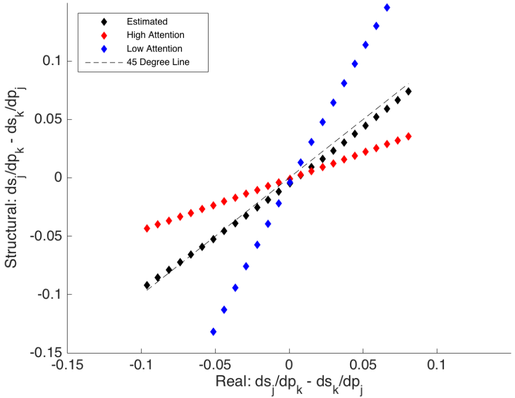


# Results: Price Coefficients

Table: Price Coefficients

| | Conditional Logit | ALogit MLE | ALogit II | 'Truth' |
|---|---|---|---|---|
| *Utility* | -0.054*** | -0.1644*** | -0.1284** | -0.173*** |
| | (0.003) | (0.037) | (0.048) | (0.004) |
| *Attention* | | 0.137*** | 0.141*** | 0.15 |
| | | (0.017) | (0.025) | |

# Asymmetries



# Field Application: Health Insurance

- Apply the Default Specific Model to Medicare Part D data on:
  - 20% sample of Part D beneficiaries from 2008-2009
  - Low Income Subsidy (LIS) beneficiaries “with stakes”
- DSC model applied by Heiss et al (2017) and Ho, Hogan & Scott-Morton although both rely on additional exclusion restrictions for identification
- Key question: how to explain inertia in choices over time?
- Often get implausibly large estimates of switching costs ($> \$1,000$)

# Health Insurance

- Important for welfare evaluation of a smart default policy
  - Low switching because of high inattention?
  - Low switching because of utility relevant switching costs?
- Two sources of switching costs:
  - Paperwork costs, $\rho$: hassle and time to enrol in new scheme
  - Acclimation costs, $\alpha$: cost of rescheduling deliveries and switching to new drugs
- Identification strategy:
  - Asymmetries: disentangle inattention from switching costs
  - Random reassignment of LIS beneficiaries: separately identify $\rho$ and $\alpha$

# Utility

- Choice probabilities in the DSC model given by:

$$s_{ijt} \equiv s_{jt}(\mathbf{x}_{it}) = (1 - \mu_t(\mathbf{x}_{idt}))\, Default_{ijt} + \mu_t(\mathbf{x}_{idt}) s^{\star}_{jt}(\mathbf{x}_{it})$$

- Conditional on being awake, the utility of individual *i* from choosing plan *j* at time *t* is given by:

$$u_{ijt} = \mathbf{x}_{ijt}\beta + (\alpha + \rho) Default_{ijt} + \epsilon_{ijt}$$

- When LIS beneficiaries no longer qualify for full premium subsidies, utility is given by:

$$u_{ijt} = x_{ijt}\beta + (\alpha + \rho) Default_{ijt} + \alpha \left( Default_{ij,t-1} \times Reassigned_{ijt} \right) + \epsilon_{ijt}$$

# Utility

- Choice probabilities in the DSC model given by:

$$s_{ijt} \equiv s_{jt}(\mathbf{x}_{it}) = (1 - \mu_t(\mathbf{x}_{idt}))\, Default_{ijt} + \mu_t(\mathbf{x}_{idt}) s^{\star}_{jt}(\mathbf{x}_{it})$$

- Conditional on being awake, the utility of individual *i* from choosing plan *j* at time *t* is given by:

$$u_{ijt} = \mathbf{x}_{ijt}\beta + (\alpha + \rho) Default_{ijt} + \epsilon_{ijt}$$

- When LIS beneficiaries no longer qualify for full premium subsidies, utility is given by:

$$\begin{aligned} u_{ijt} &= x_{ijt}\beta + (\alpha + \rho) Default_{ijt} \\ &\quad + \alpha \left( Default_{ij,t-1} \times Reassigned_{ijt} \right) + \epsilon_{ijt} \end{aligned}$$

# Welfare

- Smart Default Policy: individuals are reassigned to an alternative plan and given the option of immediately switching back if they desire without enrolling in the new plan for a year.
- Assume that normative utility depends only on total cost and other observable factors
- Change in welfare associated with the policy can be expressed as:

$$\begin{aligned} \triangle W_i &= W_i^1 - W_i^0 \\ &= \rho\left(s_{id}^1 - s_{io}^0\right) + \alpha \triangle s_{io} + \sum_j \triangle s_{ij} v_{ij} \end{aligned} \tag{5.1}$$

# Preference Parameters: Medicare Part D

| | Conditional Logit | | DSC Model | |
|---|---|---|---|---|
| *Utility*: | | | | |
| Annual Premium (hundreds) | -0.505*** | (0.005) | -1.034*** | (0.010) |
| Annual Out of Pocket Costs (hundreds) | -0.214*** | (0.007) | -0.297*** | (0.012) |
| Variance of Costs (millions) | 2.246*** | (0.089) | 2.579*** | (0.165) |
| Deductible (hundreds) | -0.516*** | (0.009) | -0.724*** | (0.013) |
| Donut Hole Coverage | 0.691*** | (0.027) | 0.335*** | (0.051) |
| Average Consumer Cost Sharing % | -1.181*** | 0.107 | -4.128*** | 0.163 |
| # of Top 100 Drugs in Formulary | 0.038*** | (0.004) | 0.172*** | (0.006) |
| Normalized Quality Rating | 0.438*** | (0.010) | 0.515*** | (0.015) |
| Original Plan | 0.988*** | (0.238) | 1.314*** | (0.257) |
| Assigned Plan | 6.428*** | (0.012) | 4.240*** | (0.078) |
| Acclimation Costs | $196 | | $127 | |
| Paperwork Costs | $1078 | | $283 | |
| Attention Probability | | | 19.7% | |

# Attention Parameters: Medicare Part D

| | Conditional Logit | | DSC Model | |
|---|---|---|---|---|
| *Attention*: | | | | |
| Annual Premium (hundreds) | - | - | $0.062^{***}$ | (0.014) |
| Annual Out of Pocket Costs (hundreds) | - | - | $0.030^{*}$ | (0.012) |
| Variance of Costs (millions) | - | - | $-0.627^{***}$ | (0.159) |
| Deductible (hundreds) | - | - | $0.069^{***}$ | (0.020) |
| Donut Hole Coverage | - | - | $-0.761^{***}$ | (0.052) |
| Average Consumer Cost Sharing % | - | - | $-1.447^{***}$ | (0.219) |
| # of Top 100 Drugs in Formulary | - | - | -0.002 | (0.010) |
| Normalized Quality Rating | - | - | $-0.511^{***}$ | (0.019) |
| Acclimation Costs | $196 | | $127 | |
| Paperwork Costs | $1078 | | $283 | |
| Attention Probability | | | 19.7% | |

# Welfare Simulations: Smart Default Policy 1

| | Attention Cost | | | | |
|---|---|---|---|---|---|
| | \$0 | \$50 | \$100 | \$200 | \$300 |
| Conditional Logit Parameters | \$31 | \$31 | \$31 | \$31 | \$31 |
| DSC Parameters | \$177 | \$177 | \$177 | \$177 | \$177 |
| *Direct Effect on Attention Probability* | | | | | |
| 25% | | | | | |
| 50% | | | | | |
| 75% | | | | | |
| 100% | | | | | |

# Welfare Simulations: Smart Default Policy 1

| | Attention Cost | | | | |
|---|---|---|---|---|---|
| | $0 | $50 | $100 | $200 | $300 |
| Conditional Logit Parameters | $31 | $31 | $31 | $31 | $31 |
| DSC Parameters | $177 | $177 | $177 | $177 | $177 |
| *Direct Effect on Attention Probability* | | | | | |
| 25% | $172 | $170 | $168 | $164 | $160 |
| 50% | $144 | $129 | $115 | $86 | $57 |
| 75% | $112 | $85 | $58 | $4 | -$50 |
| 100% | $77 | $37 | -$2 | -$81 | -$161 |

# Welfare Simulations: Smart Default Policy 2

| | Attention Cost | | | | |
|---|---|---|---|---|---|
| | $0 | $50 | $100 | $200 | $300 |
| DSC Parameters | $222 | $222 | $222 | $222 | $222 |
| *Direct Effect on Attention Probability* | | | | | |
| 25% | $215 | $213 | $210 | $204 | $199 |
| 50% | $184 | $168 | $153 | $122 | $91 |
| 75% | $150 | $122 | $95 | $39 | -$17 |
| 100% | $114 | $74 | $35 | -$45 | -$124 |

# Overview of Additional Analysis

- Reduced form evidence of asymmetries: differential sensitivity of switching to changes in the characteristics of the default and rival plans
- Overidentification tests: test whether the exclusion restrictions used in the literature are valid

# Conclusion

- Show identification of a class of consideration set models that are likely to be useful to applied researchers
- Exploit violations in symmetry of cross derivatives
- Assumptions already made by researchers in specifications with full-consideration typically sufficient for identification with limited consideration
- Demonstrate model utility/tractability through applications to choice in a variety of different settings including welfare evaluation of Smart Default Policy

# Thank you!

# Symmetry Proof: Example Nested Logit

- To see in a very simple case, consider the nested logit in which cross-price effects take the form:

$$\frac{\partial s_{jm}}{\partial p_{km}} = \begin{cases} \beta s_{km} \left( \frac{\sigma}{1-\sigma} s_{jm|g} + s_{jm} \right) & \text{if } j \text{ and } k \text{ in the same nest} \\ \beta s_{jm} s_{km} & \text{otherwise} \end{cases}$$

  - $s_{jm|g}$: gives the within-nest market share of good $j$
  - $\sigma$: how different substitution patterns are within and across nests.

- Clear that these are symmetric in products in different nests, but what about those in the same nest?

# Symmetry Proof: Example Nested Logit

- For products in the same nest we have:

$$\frac{\partial s_{jm}}{\partial p_{km}} - \frac{\partial s_{km}}{\partial p_{jm}} = \beta \frac{\sigma}{1-\sigma} \left( s_{km} s_{jm|g} - s_{jm} s_{km|g} \right)$$

- Given that

$$s_{jm} = s_{jm|g} g_m$$

where $g_m$ is the probability of buying a good from nest $g$. We have:

$$\begin{aligned} \frac{\partial s_{jm}}{\partial p_{km}} - \frac{\partial s_{km}}{\partial p_{jm}} &= \beta \frac{\sigma}{1-\sigma} \left( s_{km|g} g_m s_{jm|g} - s_{jm|g} g_m s_{km|g} \right) \\ &= 0 \end{aligned}$$

Back

# Symmetry Proof: General

- With [] denoting exclusion, the probability that option $j$ is chosen under full consideration is given by:

$$s^{\star}_{jm} = Pr\left(v_{jm} + \epsilon_{ijm} = \max_{j'} v_{j'm} + \epsilon_{ij'm}\right)$$
$$= \int \int_{-\infty}^{v_{jm}+e-v_{0m}} ... \int_{-\infty}^{v_{jm}+e-v_{Jm}} f(z_0, ..., e, ..., z_J)\mathrm{d}z_J...[\mathrm{d}z_j]...\mathrm{d}z_0\mathrm{d}e$$

- This allows for an arbitrary correlation structure in the random utility errors.

# Symmetry Proof: General

- Then:

$$\frac{\partial s_{jm}}{\partial p_{j'm}} = -\beta \int \int_{-\infty}^{v_{jm}+e-v_{0m}} ... \left[ \int_{-\infty}^{v_{jm}+e-v_{jm}} \right] ... \left[ \int_{-\infty}^{v_{jm}+e-v_{j'm}} \right] ... \int_{-\infty}^{v_{jm}+e-v_{Jm}}$$

$$f(z_0, ..., e, ..., v_{jm} + e - v_{j'm}, ..., z_J) \mathrm{d}z_J ... [\mathrm{d}z_j'] ... [\mathrm{d}z_j] ... \mathrm{d}z_0 \mathrm{d}e$$

# Symmetry Proof: General

- Using the change of variables $t = v_{jm} + e - v_{j'm}$, one obtains:

$$\frac{\partial s_{jm}}{\partial p_{j'm}} = -\beta \int \int_{-\infty}^{v_{j'm}+t-v_{0m}} .. \left[ \int_{-\infty}^{v_{j'm}+t-v_{j'm}} \right] .. \left[ \int_{-\infty}^{v_{j'm}+t-v_{jm}} \right] .. \int_{-\infty}^{v_{j'm}+t-v_{iJ}}$$

$$f(z_0, ..., v_{j'm} + t - v_{jm}, ..., t, ..., z_J) \mathrm{d}z_J ... [\mathrm{d}z'_j] ... [\mathrm{d}z_j] ... \mathrm{d}z_0 \mathrm{d}t$$

$$= \frac{\partial s_{j'm}}{\partial p_{jm}}$$

Back