Birkbeck Working Papers in Economics & Finance

ISSN 1745-8587

Birkbeck
UNIVERSITY OF LONDON

Department of Economics, Mathematics and Statistics

BWPEF 1503

Choice in the Presence of Experts

Walter Beckert
Birkbeck, University of London

April 2015

▪ Birkbeck, University of London ▪ Malet Street ▪ London ▪ WC1E 7HX ▪

Choice in the Presence of Experts

Walter Beckert

April 28, 2015

Abstract

This paper considers the micro-econometric analysis of discrete choice problems in which the choice set is strategically pre-selected by a third-party advisor or expert. It delineates measures of efficiency loss arising from (i) the sets of relevant choice attributes being imperfectly aligned between ultimate beneficiary of the choice outcome and third-party expert, and (ii) pre-selected choice-sets being uniform across ultimate beneficiaries who differ in their subjective evaluations of relevant choice attributes. And it identifies inconsistencies in estimation when strategically pre-selected choice sets are treated as exogenous. Some applications to choice in healthcare and financial services markets are sketched.

Preliminary and Incomplete. Comments welcome.

Keywords: Discrete choice, expert, endogenous choice sets.
JEL classification: D120, C510, I110, G110.

1 Introduction

As a by-product to some work I have done in the recent past on competition issues in healthcare (Beckert et al. (2012)), and in particular on patients' choice of hospital for elective medical procedures, I have become interested more generally in the econometric analysis of discrete choice outcomes when choice sets are endogenous. By this, I am thinking of situations in which an "expert" (advisor or broker) with arguably superior information strategically presents a set of pre-selected choice alternatives to a decision maker (e.g. a general practitioner (GP) presents sets of hospitals, out of the universe of all hospitals, for a patient to choose from for an elective medical procedure). This proposed research is concerned with biases in estimation when the potential endogeneity of choice sets is ignored in the econometric model that forms the basis of analysis. In the area of healthcare provision, the role of GPs within the topology of National Health Service (NHS) funded services has been enhanced by the Health and Social Care Act (2012) which gives GPs significant responsibilities for purchasing healthcare services in England. Consequently, there is also increased substantive interest in understanding the impact of GPs on patient choice and medical outcomes.

Endogeneity of choice sets is also an issue in the area of financial decision making. Here, a financial advisor or broker may offer sets of financial contracts to a retail client (e.g. different investment funds or assets, out of all traded assets; or different insurance products). This is also an area of regulatory interest. The then Financial Services Authority[1], for example, in its recent Retail Distribution Review (RDR) proposed various changes to the remuneration, capital and independence requirements for financial advisors. Some real estate decisions have similar characteristics, as do certain types of art purchases. Chamley (2004) summarizes the growing theoretical microeconomic literature on the role of experts in consumer and investor choice decisions.

Finally, strategic composition of choice sets emerges as a feature of online markets. Social media platforms are at the point of becoming gateways to online service providers. For example, Facebook in the future may host contents of selected online

[1]Now, Financial Conduct Authority.

news media[2] and already now acts as platform for app-install ads[3]. Furthermore, antitrust authorities have focussed on Googles competition with so-called vertical, or specialised, search services, such as comparison shopping sites, travel search engines and search sites aimed at local services, out of concern that rivals are disadvantaged because Google's search platform allegedly gives preferential treatment to results from its own services; this concern has culminated in the launch of a formal inquiry by the European Commission's Directorate for Competition into Google's shopping searches[4]. In these instances, the design of the online platform, acting as a gateway to services relevant to their ultimate users, is likely governed by revenue considerations of the platform operator - such as revenue from advertisement or proprietary services - that are not aligned with those relevant to the service users.

Exogeneity of the set of choice alternatives is an implicit assumption in conventional discrete choice analysis (McFadden (1974))[5]. To my knowledge, the implications for econometric analysis of endogenous choice sets have not been investigated in detail, and no general econometric choice model exists that allows for endogenous choice sets. McFadden (1977) offers two conditions - positive and uniform conditioning, characterizing an exogenous selection mechanism - that are sufficient to yield consistent estimates in the presence of exogenously limited choice sets; Santos et al. (2013) refer to this result as justification for the consistency of their maximum likelihood estimator with imposed choice sets that are subsets of the true choice sets. From an econometric perspective, the endogeneity of the set of choice alternatives constitutes a potential sample selection problem that tends to bias estimation results. This is similar to the well-known issue of incidental truncation (Heckman (1976)) whereby decision outcomes of interest are only observed for a selected subsample and where failure to properly model the sample selection mechanism induces the estimates of the outcome relationship to be biased and inconsistent. Similar is-

[2]See New York Times, 24 March 2015;
http://www.nytimes.com/2015/03/24/business/media/facebook-may-host-news-sites-content.html

[3]See New York Times, 26 March 2015;
http://www.nytimes.com/2015/03/26/technology/debunking-the-latest-predictions-of-facebooks-demise.h

[4]See, for example, Financial Times, 02 and 15 April 2015;
http://www.ft.com/cms/s/0/97a4dc62-e360-11e4-9a82-00144feab7de.html?siteedition=uk#axzz3XIZ3NHfN
http://www.ft.com/cms/s/0/0c2b2840-d8d3-11e4-8a23-00144feab7de.html?siteedition=intl#axzz3W8LdSMDi

[5]See Maddala (1983) for a general introduction and overview.

sues also arise in the analysis of endogenous sample attrition (Hausman and Wise (1979)).

This note works out some preliminary results that demonstrate the source of potential inconsistency of estimators when the endogeneity of choice sets is ignored.

2 Micro-level Econometric Model

2.1 Basic Setup and Unconstrained Pre-Selection

Consider a decision maker i who is modelled to make a discrete choice out of a set of alternatives $\mathcal{J}$. Suppose this set is presented to him by an advisor who arguably possesses superior information, say on the hard-to-assess quality of all possible choice alternatives, collected in the set $\mathcal{H}$. The role of the advisor is to pre-select $\mathcal{J} \subseteq \mathcal{H}$ for the benefit of the ultimate decision maker i.

Consider the expert's selection mechanism. Denote the expert's latent assessment of alternative j's net benefit by $v_j^\star$; this could incorporate anticipated benefits accruing to i, any benefits accruing to the expert as a result of incentivization schemes put in place by the producer of j; or any benefits accruing to the expert's reputation from promoting alternative j. Suppose that the expert includes j in $\mathcal{J}$ if, and only if, $v_j^\star > 0$:

$$\begin{aligned} v_j^\star &= \alpha_j - \xi_j, \\ v_j &= 1_{\{v_j^\star > 0\}}, \quad j \in \mathcal{H}, \end{aligned}$$

where α_j denotes the measurable component of $v_j^\star$, ξ_j is unobserved by the econometrician, and v_j is a binary inclusion indicator, taking value one when the expert's net benefit assessment is positive so that j is included in $\mathcal{J}$, and zero otherwise. Here, ξ_j might capture, in particular, the unquantifiable quality assessment of alternative j by the expert, e.g. to the extent that it affects the expert's prospective reputation or other subjective or "soft" attributes of alternative j. In this framework, the expert has all the information relevant to him, each choice alternative is assessed by the expert individually and independently, and $\mathcal{J} = \{j \in \mathcal{H} : v_j = 1\}$. An alternative pre-selection mechanism is outlined below.

Now consider the ultimate decision maker i. Suppose with any conceivable choice alternative i associates an indirect conditional utility $u^\star_{ij}$,

$$u^\star_{ij} = \delta_{ij} + \zeta_j + \epsilon_{ij},$$

that comprises a measurable component δ_{ij}, next to unobserved components ζ_j and ϵ_{ij}. Here, δ_{ij} might capture observable attributes of j that relate directly to i, e.g. geographic distance, coverage of specific idiosyncratic risks, etc. The unobservable ζ_j might reflect quality aspects of alternative j that are unobserved by the econometrician, as is i's idiosyncratic taste or preference for j, modelled by ϵ_{ij}. The indirect utility that decision maker i associates with alternative j is latent, but inference about δ_{ij} is possible to the extent that j is included in $\mathcal{J}$, in that it can be observed whether or not j is chosen by i. This is akin to the well-known incidental truncation and non-random sample selection issues first addressed in a regression framework by Heckman (1976).

Consider the case when ξ_j and ζ_j are allowed to be correlated. This may arise when unobserved quality aspects of alternative j are at least partly relevant to both, the decision maker and the expert. This is plausibly so when the expert's reputation hinges on matching up consumers, like i, with beneficial choice outcomes, like j. It can also arise from subjective assessments of "soft" (i.e. not easily quantifiable or measurable) attributes of the choice alternative.[6] Then, given $j \in \mathcal{J}$,

$$\begin{aligned} \tilde{u}^\star_{ij} &:= \mathbb{E}\left[u^\star_{ij} | j \in \mathcal{J}\right] \\ &= \delta_{ij} + \mathbb{E}[\zeta_j | \xi_j < \alpha_j] + \epsilon_{ij} \\ &= \delta_{ij} + \phi(\alpha_j) + \epsilon_{ij}, \end{aligned}$$

where $\phi(\alpha_j) = \mathbb{E}[\zeta_j | \xi_j < \alpha_j]$ accounts for the effect of the expert's inclusion of j in $\mathcal{J}$; for example, if ξ_j and ζ_j are bivariate standard normal with correlation $\rho \in (-1, 1)$,

[6] In the medical context, for example, patient and GP may differ in terms of what they consider relevant aspects of the perioperative care and environment: The GP may focus on strictly medical aspects (e.g. availability of specialist expertise for treating any comorbidities), while a patient may focus also on psychosocial aspects (e.g. psychological support to mitigate anxiety) which may affect somatic recovery.

then this terms is the well-known Mills ratio, evaluated at α_j and pre-multiplied by ρ. The observed choice outcome is an indicator Y_{ij} taking value one when i choose j out of $\mathcal{J}$, i.e.

$$Y_{ij} = 1_{\left\{\tilde{u}^{\star}_{ij}=\max\{\tilde{u}^{\star}_{ik},k\in\mathcal{J}\}\right\}},$$

and the probability distribution of Y_{ij} is induced by distributional assumptions on ϵ_{ij}.

In applied work, it is advantageous to allow for an outside good, in a nested model. Suppose the indirect utility of the outside option is normalized to zero.[7]. Under extreme value type 1 assumptions on the ϵ_{ij}s, this yields the well known nested logit choice probabilities (McFadden (1978), Cardell (1991)),

$$\begin{aligned}\Pr\left(Y_{ij}=1\right) &= \frac{\exp((\delta_{ij}+\phi(\alpha_j))/(1-\sigma))}{\sum_{k\in\mathcal{J}}\exp((\delta_{ik}+\phi(\alpha_k))/(1-\sigma))}\frac{\left(\exp((\delta_{ij}+\phi(\alpha_j))/(1-\sigma))\right)^{1-\sigma}}{\left(1+\sum_{k\in\mathcal{J}}\exp((\delta_{ik}+\phi(\alpha_k))/(1-\sigma))\right)^{1-\sigma}} \\ &= \frac{\exp((\delta_{ij}+\phi(\alpha_j))/(1-\sigma))}{\exp(I_{i\mathcal{J}})}\frac{\exp((1-\sigma)I_{i\mathcal{J}})}{1+\exp((1-\sigma)I_{i\mathcal{J}})},\end{aligned}$$

where $\sigma \in [0,1)$ is a measure of the correlation of the indirect utilities of the inside goods, and $I_{i\mathcal{J}} = \ln\left(\sum_{k\in\mathcal{J}}\exp((\delta_{ik}+\phi(\alpha_k))/(1-\sigma))\right)$ is the inclusive value of the nest comprising the inside goods. The expression for $\Pr(Y_{ij}=1)$ demonstrates that the selection terms $\phi(\alpha_j)$, $j \in \mathcal{J}$, constitute regressors that are omitted in analyses that ignore strategic choice set pre-selection by an expert or advisor, provided correlation between ξ_j and ζ_j cannot be ruled out and the selection terms vary across $j \in \mathcal{J}$. Such omission will yield inconsistent maximum likelihood estimates, as a consequence of model mis-specification.

Sovinsky Goeree (2008) presents a related model of random choice or consideration sets at the level of the decision maker in which the probability of the decision maker being informed about a choice alternative j takes the place of the inclusion probability $\Pr(v_j = 1)$. In her model of the US personal computer industry, these probabilities are exogenously driven by product level advertising and consumer level media exposure.[8] Her model can be viewed as a special case of the present model in which ζ_j and ξ_j are independent, conditional on observed attributes.

[7]This is the approach taken in Besanko et al (1990). An alternative is Anderson and de Palma (1992) who specify the conditional utility of the outside option as an independent EV(0,1) variable, i.e. as varying across decision makers

[8]See also Dinerstein et al. (2014) for an application to consumer search in internet commerce.

This type of model is less compelling in situations when information about choice alternatives is costly to acquire and disseminate. On the one hand, this renders decision making complex for the uninformed layman. And on the other hand, it creates a role for the informed expert, namely to reduce the complexity of the decision process for the layman. The following subsection describes an alternative model that captures these ideas.

2.2 Constrained Pre-Selection and Efficiency

Suppose that α_j and δ_{ij} share some, but not all attributes relevant to decision maker i and the expert advisor. For example, in the hospital choice context, let $\mathbf{x}_{ij}$ denote hospital j's attributes that are taken into account by both, $\mathbf{w}_{ij}$ those that only matter to the patient, and $\mathbf{z}_j$ those that only matter to the GP, in the role of the expert. For simplicity, suppose that patient and GP attach the same weights (coefficients) δ to $\mathbf{x}_{ij}$, so that

$$\begin{aligned} \delta_{ij} &= \mathbf{x}'_{ij}\delta + \mathbf{w}'_{ij}\beta, \\ \alpha_{ij} &= \mathbf{x}'_{ij}\delta + \mathbf{z}'_{j}\alpha \end{aligned}$$

where α and β are parameter vectors and α_{ij}, taking the role of α_j above, reflects the variation of $\mathbf{x}$ across i, in addition to j. The indirect utility of the ultimate, layman decision maker is then

$$\begin{aligned} u^{\star}_{ij} &= \delta_{ij} + \zeta_{ij} + \epsilon_{ij} \\ &= \mathbf{x}'_{ij}\delta + \mathbf{w}'_{ij}\beta + \zeta_{ij} + \epsilon_{ij}. \end{aligned}$$

Condition on the event that the outside option is not chosen, i.e. focus the choice of an alternative from the set of inside goods $\mathcal{J}_i$ pre-selected by the expert.[9] Under the i.i.d. EV(0,1) assumption on ϵ_{ij} and assuming that decision maker i takes the

Gaynor et al. (2014) emphasize the promise this approach holds in healthcare industrial organization research.

[9]In the setting of this subsection, $\mathcal{J}$ may also depend on i, to the extent that the expert espouses the attributes that i values and that these vary with i, e.g. distance.

pre-selected choice set $\mathcal{J}_i$ as given[10],

$$\begin{aligned}\Pr(Y_{ij} = 1|\mathcal{J}_i) &= \frac{\exp(\delta_{ij} + \zeta_{ij})}{\sum_{k\in\mathcal{J}} \exp(\delta_{ik} + \zeta_{ik})}, \quad j \in \mathcal{J}_i \\ &= 0 \qquad\qquad \text{o.w., i.e. } j \notin \mathcal{J}_i,\end{aligned}$$

while, absent the pre-selection,

$$\Pr(Y_{ij} = 1|\mathcal{H}) = \frac{\exp(\delta_{ij} + \zeta_{ij})}{\sum_{k\in\mathcal{H}} \exp(\delta_{ik} + \zeta_{im})} \quad j \in \mathcal{H}.$$

This implies that the divergence of the distribution of patient level choice outcomes under pre-selection relative to their distribution absent pre-selection, in terms of the Kullback-Leibler measure (relative entropy), is

$$\begin{aligned}D(\mathcal{J}_i||\mathcal{H}; \mathbf{x}_i, \mathbf{w}_i) &= \sum_{j\in\mathcal{J}_i} \Pr(Y_{ij} = 1|\mathcal{J}_i) \ln\left(\frac{\Pr(Y_{ij} = 1|\mathcal{J}_i)}{\Pr(Y_{ij} = 1|\mathcal{H})}\right) \\ &= \ln\left(\sum_{k\in\mathcal{H}} \exp(\delta_{ik})\right) - \ln\left(\sum_{m\in\mathcal{J}_i} \exp(\delta_{im})\right) \\ &= \ln\left(\sum_{k\in\mathcal{H}} \exp(\mathbf{x}_{ik}\delta + \mathbf{w}'_{ij}\beta + \zeta_{ij})\right) - \ln\left(\sum_{m\in\mathcal{J}_i} \exp(\mathbf{x}'_{im}\delta + \mathbf{w}'_{im}\beta + \zeta_{im})\right) \\ &= I_{\mathcal{H}}(\mathbf{x}_i, \mathbf{w}_i) - I_{\mathcal{J}_i}(\mathbf{x}_i, \mathbf{w}_i).\end{aligned}$$

This divergence can be viewed as a loss in efficiency that arises from reducing the complexity of the choice problem, limiting it to evaluating $J_i = \#\mathcal{J}_i$ alternatives, instead of $H = \#\mathcal{H} \geq J_i$. This can be seen as a benefit by uninformed laymen decision makers if there are information acquisition and decision costs for them.

Comment: The Kullback - Leibler divergence for two measures P and Q is $D(P||Q) = \mathbb{E}_P[\ln(P/Q)] = \sum_j P(j)\ln(P(j)/Q(j))$ and not symmetric. It requires that $Q(j) = 0$ implies $P(j) = 0$, i.e. that $\Pr(Y_{ij} = 1|\mathcal{H}) = 0$ implies $\Pr(Y_{ij} = 1|\mathcal{J}_i) = 0$. In the present model, this is plausible, and the reverse is neither plausible nor true. This dictates how the two probability distributions enter the measure, i.e. as $D(\mathcal{J}_i||\mathcal{H}; \mathbf{x}_i, \mathbf{w}_i)$, not as $D(\mathcal{H}||\mathcal{J}_i; \mathbf{x}_i, \mathbf{w}_i)$. Typically, however, the divergence is interpreted as Q approximating P, i.e. as $\{\Pr(Y_{ij} = 1|\mathcal{H}), j \in \mathcal{H}\}$ approximating $\{\Pr(Y_{ij} = 1|\mathcal{J}_i), j \in \mathcal{H}\}$. Here, the interpretation is the reverse. White (1994)

[10] This amounts to assuming that the decision maker behaves non-strategically and does not question how the expert arrived at the pre-selection outcome $\mathcal{J}_i$.

offers an interpretation[11] that, adapted to this model, implies that the divergence measures the "surprise" from learning that decision outcomes are in fact governed by $\{\Pr(Y_{ij} = 1|\mathcal{J}_i), j \in \mathcal{H}\}$, rather than by $\{\Pr(Y_{ij} = 1|\mathcal{H}), j \in \mathcal{H}\}$.

Suppose that, from the expert's perspective, there is a constant unit cost $C > 0$ of including an alternative into $\mathcal{J}_i$. This cost may be specific to the expert. For example, in the context of hospital choice in the UK where a GP (practice) plays the role of the expert, this cost might be expected to be a function of practice list size, the number of GPs in the practice, their work experience and whether they obtained their qualification in the UK or abroad. It imposes a constraint that can be thought of as the effort the expert needs to exert in order to explain the features, pros and cons of the alternative to the decision maker. Let $\mathcal{P}$ denote the set of all partitions of $\mathcal{H}$, i.e. $\mathcal{P} = \{\mathcal{G} \subset \mathcal{H} : \#\mathcal{G} \leq \#\mathcal{H}\}$. Furthermore, suppose the expert's objective in selecting $\mathcal{J}_i$ is to minimize the efficiency loss, or "surprise" in the sense of White's interpretation, on the basis of his own assessments of the choice alternatives in $\mathcal{H}$, taking account the cost constraint of including alternatives. Let the expert's assessment of i's valuation of alternative j be $v^{\star}_{ij} = \alpha_{ij} + \sigma^2 \xi_{ij}$, where ξ_{ij} captures the expert's incomplete information about i's preferences, i.e. σ^2 reflects his uncertainty about decision maker i's valuation of alternative j.[12] Just as the cost parameter C, the variance parameter σ^2 might also be specific to the expert; for example, in the case of GP (practices) in the role of the expert, this might be a function of the heterogeneity of patients within the practice list. More specifically, suppose that the econometric error term ζ_{ij} in $u^{\star}_{ij}$ can be decomposed into uncertainty $\mu^x_{ij} + \xi^x_{ij}$ with regard to the attributes taken into account by both, decision maker and expert, and uncertainty $\mu^w_{ij} + \xi^w_{ij}$ with regard to attributes that only matter to the decision maker,

$$\zeta_{ij} = \mu^x_{ij} + \mu^w_{ij} + \xi^x_{ij} + \xi^w_{ij},$$

where μ^x_{ij} and μ^w_{ij} are those parts of the econometrician's uncertainty about the two parts of δ_{ij} that are known to the expert, while ξ^x_{ij} and ξ^w_{ij} are unknown to both. From

[11]See p.31; his Assumption 3.4 is satisfied because $\{\Pr(Y_{ij} = 1|\mathcal{H}), j \in \mathcal{H}\}$ is a probability distribution.

[12]It is worth emphasizing that it is necessary to allow for uncertainty about the laymen decision maker's valuations on the part of the expert because absent such uncertainty the expert would simply choose on behalf of the laymen.

the perspective of the expert who cares only about the utility contribution related to $\mathbf{x}$, only the former matters. So, $\xi_{ij} = \xi^x_{ij}$. Consequently, from the perspective of the econometrician, in the model for the expert μ^x_{ij} matters in addition to ξ_{ij}.

Assuming the ξ_{ij} are i.i.d. extreme value type 1, the distribution of choice outcomes from the expert's perspective is given by logit choice probabilities based on attributes $\mathbf{x}$ and $\mathbf{z}$. Denote the econometrician's incomplete information about the expert specific relevant attributes $\mathbf{z}$ by μ^z_j. Once the $\{\xi_{ij}\}_{i\in\mathcal{H}}$ are integrated out, the econometrician's remaining uncertainty with regard to the expert's assessment of alternative j is therefore $\mu_{ij} = \mu^x_{ij} + \mu^z_j$. This leads to

$$\mathcal{J}_i = \arg\min_{\mathcal{G}\in\mathcal{P}} D(\mathcal{G}||\mathcal{H};\mathbf{x}_i,\mathbf{z}) + C\#\mathcal{G}.$$

The solution to this problem is to order the alternatives in $\mathcal{H}$ according to their indirect utilities,

$$\begin{aligned}
\exp\left(\frac{\alpha_{i(1:H)}+\mu_{i(1:H)}}{\sigma}\right) &= \exp\left(\frac{\mathbf{x}'_{i(1:H)}\delta+\mathbf{z}'_{(1:H)}\alpha+\mu_{i(1:H)}}{\sigma}\right)\\
&\geq \cdots\\
&\geq \exp\left(\frac{\alpha_{i(H:H)}+\mu_{i(H:H)}}{\sigma}\right)\\
&= \exp\left(\frac{\mathbf{x}'_{i(H:H)}\delta+\mathbf{z}'_{(H:H)}\alpha+\mu_{i(H:H)}}{\sigma}\right) \qquad (\star)
\end{aligned}$$

and to include the ones up to the point that

$$\begin{aligned}
J_i &= \arg\max_{h\in\{1,\cdots,H\}}\left\{\ln\left(\sum_{k=1}^{h}\exp\left(\frac{\alpha_{i(k:H)}+\mu_{i(k:H)}}{\sigma}\right)\right)-\ln\left(\sum_{m=1}^{h-1}\exp\left(\frac{\alpha_{i(m:H)}+\mu_{i(m:H)}}{\sigma}\right)\right)\geq C\right\}\\
&= \arg\max_{h}\left\{-\ln\left(1-\frac{\exp\left(\frac{\alpha_{i(h:H)}+\mu_{i(h:H)}}{\sigma}\right)}{\sum_{m=1}^{h}\exp\left(\frac{\alpha_{i(m:H)}+\mu_{i(m:H)}}{\sigma}\right)}\right)\geq C\right\}
\end{aligned}$$

This also implies that

$$-\ln\left(1-\frac{\exp\left(\frac{\alpha_{i(k:H)}+\mu_{i(k:H)}}{\sigma}\right)}{\sum_{m=1}^{h}\exp\left(\frac{\alpha_{i(m:H)}+\mu_{i(m:H)}}{\sigma}\right)}\right)\geq C \quad \text{for } k=1,\cdots,J_i.$$

It is at this stage of pre-selection that the distinction between the expert and the layman decision maker emerges and can be defined: The expert has sufficient information and expertise to establish a ranking ($\star$) of the alternatives in $\mathcal{H}$ without cost, while the layman decision maker does not. This distinction is an implicit assumption in the present setup. The distinction creates a role for the expert, namely to pre-select and thereby narrow down the set of choice alternatives in order to render the layman's choice problem less complex and more tractable.

The set $\mathcal{J}_i$ resulting from the expert's pre-selection may differ, however, from the one that would be chosen if the assessment were based on δ_{ij} (encompassing $\mathbf{x}_i$ and $\mathbf{w}_i$), instead of α_{ij} (encompassing $\mathbf{x}_i$ and $\mathbf{z}$), i.e. if the decision maker's and expert's assessment criteria were perfectly aligned, in the sense that they were to consider the same set of attributes of the choice alternatives as decision relevant. Denote the choice set based on $\{\alpha_{ij}\}$ by $\mathcal{J}_{i\alpha}$, and the one that would have been pre-selected on the basis of $\{\delta_{ij}\}$ by $\mathcal{J}_{i\delta}$. The efficiency loss due to pre-selection by the expert can then be cast as

$$\begin{aligned}
\Delta_i &= D(\mathcal{J}_{i\alpha}||\mathcal{H}; \mathbf{x}_i, \mathbf{w}_i) \\
&= I_{\mathcal{H}}(\mathbf{x}_i, \mathbf{w}_i) - I_{\mathcal{J}_{i\alpha}}(\mathbf{x}_i, \mathbf{w}_i) \\
&= I_{\mathcal{H}}(\mathbf{x}_i, \mathbf{w}_i) - I_{\mathcal{J}_{i\delta}}(\mathbf{x}_i, \mathbf{w}_i) + I_{\mathcal{J}_{i\delta}}(\mathbf{x}_i, \mathbf{w}_i) - I_{\mathcal{J}_{i\alpha}}(\mathbf{x}_i, \mathbf{w}_i) \\
&= D(\mathcal{J}_{i\delta}||\mathcal{H}; \mathbf{x}_i, \mathbf{w}_i) + D(\mathcal{J}_{i\alpha}||\mathcal{J}_{i\delta}; \mathbf{x}_i, \mathbf{w}_i).
\end{aligned}$$

The first term captures the efficiency loss due to the reduction in complexity of the choice problem, while the second term captures the additional efficiency loss arising from a misalignment of assessment criteria between decision maker and expert which results in a choice set $\mathcal{J}_{i\alpha}$ which is suboptimal when evaluated on the basis of the attributes $\mathbf{x}$ and $\mathbf{w}$ relevant to the decision maker and ultimate beneficiary of the choice outcome.

The econometric implications of this model of constrained pre-selection for the econometric model are similar to the ones of unconstrained pre-selection. The econometrician cannot observe the ranking of the alternatives included in $\mathcal{J}_i$. From the inequalities ($\star$) above, the set $\{\mu_{ij}\}_{j\in\mathcal{J}_i}$ must satisfy the necessary condition for inclusion of the jth alternative, so that

$$\begin{aligned}
G(\mathcal{J}_i; \alpha_i, C) &= \left\{ \{\mu_{ij}\}_{j\in\mathcal{J}_i} : -\ln\left(1 - \frac{\exp\left(\frac{\alpha_{ij}+\mu_{ij}}{\sigma}\right)}{\sum_{m\in\mathcal{J}_i}\exp\left(\frac{\alpha_{im}+\mu_{im}}{\sigma}\right)}\right) \geq C \right\} \\
\Pr(\mathcal{J}_i; C) &= \Pr\left(G(\mathcal{J}_i; \alpha_i, C)\right).
\end{aligned}$$

To the extent that μ_{ij} is correlated with ζ_{ij}, i.e. to the extent that μ_{ij}^x is non-zero with positive probability, observing $\mathcal{J}_i$ is informative about ζ_{ij}, so that $\Phi(\alpha_i, C) = \mathbb{E}[\zeta_{ij}|G(\mathcal{J}_i; \alpha_i, C)]$ accounts for pre-selection in this model, analogous to $\phi(\alpha_i)$ in the model with unconstrained pre-selection. Unlike in the model of unconstrained pre-selection, the selection term here does not permit a closed-form solution and needs to be simulated.

The pre-selected choice sets $\mathcal{J}_i$ vary across laymen i, to the extent that the attributes considered by both, expert and layman, $\mathbf{x}_{ij}$ vary with i; e.g. distance between i and hospital j. In practice, the expert may pre-select a uniform choice set $\mathcal{J}_\alpha$ at the outset on the basis of $\mathbf{z}$ and choice attributes $\mathbf{x}$ as they relate to the average layman, and then offer this to all laymen who consult him. This wedge between the pre-selected choice set based on average attributes, rather than those specific to i, introduces yet another layer of potential inefficiency into the choice mechanism, so that the total inefficiency measured by the KL divergence is

$$\begin{aligned}\Delta &= \sum_i \left[D(\mathcal{J}_\alpha||\mathcal{J}_{i\alpha}; \mathbf{x}_i, \mathbf{w}_i) + D(\mathcal{J}_{i\delta}||\mathcal{H}; \mathbf{x}_i, \mathbf{w}_i) + D(\mathcal{J}_{i\alpha}||\mathcal{J}_{i\delta}; \mathbf{x}_i, \mathbf{w}_i)\right] \\ &= \sum_i \left[D(\mathcal{J}_\alpha||\mathcal{J}_{i\alpha}; \mathbf{x}_i, \mathbf{w}_i) + \Delta_i\right].\end{aligned}$$

Uniformity of the pre-selected choice set across i adds, for each layman i, an additional potential efficiency to Δ_i.

3 Market-level Model

The same considerations apply when the micro-level econometric choice model is aggregated to market levels, as in Berry (1994) and Berry, Levinsohn and Pakes (1995). For ease of exposition, suppose $\delta_{ij} = \delta_j$ for all decision makers i advised by the expert. Then, the share of alternative j of the market defined or procured by the expert, s_j, is given by

$$s_j = \frac{\exp((\delta_j + \phi(\alpha_j))/(1-\sigma))}{\exp(I_\mathcal{J})} \frac{\exp((1-\sigma)I_\mathcal{J})}{1 + \exp((1-\sigma)I_\mathcal{J})}.$$

For s_0 the analogously defined share of the outside good and $s_{j|\mathcal{J}}$ the market share of j within the nest of inside goods, this yields

$$\ln(s_j) - \ln(s_0) = \delta_j + \phi(\alpha_j) + \sigma \ln(s_{j|\mathcal{J}}) + \nu_j,$$

where ν_j is any part of ζ_j that is uncorrelated with ξ_j. Again, the selection term $\phi(\alpha_j)$ persists as a regressor that is omitted in the conventional estimating equation that does not account for sample selection.

4 Data Requirements

An empirical illustration could either be based on micro-level choice data, within a framework as in Section 2, or on market-level data, within a framework as in Section 3 of this note.

With regard to micro-level data, my initial plan was to use British Health Episode Statistics (HES), as in earlier work. As already suggested, investigating the role of general practitioners (GPs), as experts, is rather topical in the UK because the Health and Social Care Act (2012) gives GPs significant responsibilities for purchasing healthcare services in England. In the HES data, recording National Health Service funded medical procedures in England, patients can be linked to GPs, and assuming that patients only choose from what their GP offers them and GPs only offer alternatives that are taken up with positive probability, the sets $\mathcal{J}$ and $\mathcal{H}$ can be constructed for each GP. Attributes of choice alternatives are also available (e.g. from Dr Foster, Lang Buisson etc.), and some hypothesized GP incentivization schemes can be identified (e.g. market-forces-factor based uplifts on reimbursements accruing to certain foundation trust hospitals with residual profit motives). Unfortunately, as a consequence of recent increased parliamentary scrutiny of health data releases, the Department of Health is currently reviewing its data access procedures, so that there is a risk that work on current and forthcoming HES data may be held up.

An alternative I am investigating is similar data for Denmark, but it is my current understanding that the Danish healthcare system is not yet subject to the same market forces and competitive pressures that would be the ultimate substantive motivation for this kind of research.

With regard to market level data, the choice data on employer-sponsored health insurance in the US, used in Dafny, Ho and Varela (2013), might be ideal. Since both employer plan offerings and employee plan selections are observed, this identifies $\mathcal{J}$ and the respective market shares, next to $\mathcal{H}$. It is a bit unclear to me whether the data contain attributes that are in α_j and not in δ_j, and vice versa. I understand that this is proprietary data, however.

5 Example

Consider a financial advisor facing a client or advisee i, and contemplate this situation pre and post RDR, indexed by $t = \{0, 1\}$. Here, the client can be thought of as a set of relevant characteristics (investment volume[13], attitude towards risk, socio-demographics); cast in this way, it is not necessary that the same individual is seen to make decisions in both time periods, only that individuals with broadly similar relevant characteristics are advised by the financial advisor in both periods.

- $\mathcal{H}$: The set of all financial products the advisor has access to and hence could conceivably suggest to i, both pre and post RDR; the two points of observation around the policy implementation should be close enough so that one could argue that no materially new products appeared on the market and that no relevant products were discontinued. Post RDR, in principle this ought to be the entire market, but for all practical purposes this is presumable a substantially smaller set for any given advisor. Empirically, it is presumably not (much) wider than the set of products observed as being chosen post RDR by the advisor's clients.
- $\mathcal{J}_{it}$: The set of choice alternatives presented by the advisor to client i at time t. This ought to be captured in the advisor's records and minutes of consultation

[13]This investment volume is assumed to be exogenous. This may be considered a strong assumption, to the extent that some funds require minimum amounts to be invested, etc..

meetings with the client.

- α_{ijt}: This is a function of attributes of product j that are stipulated to govern the advisor's decision on whether or not to include product j in $\mathcal{J}_{it}$ at time t. Such attributes are likely to be the riskiness of product j, relative to client i's risk tolerance; the commission accruing to the advisor in the event i purchases product j, possibly expressed as a percentage of funds invested. Collect these attributes as vector $\mathbf{z}_{ijt}$. Typically, α_{ijt} would be specified as a linear function in $\mathbf{z}_{ijt}$, say $\alpha_{ijt} = \mathbf{z}'_{ijt}\mu_t$, where μ_t is a vector of coefficients of the same dimension as $\mathbf{z}_{ijt}$; its elements capture the impact that the corresponding product attribute has on the advisor's decision at time t on whether or not to include product j in the set of choice alternatives presented to i in the consultation.

- δ_{ijt}: This is a function of attributes of product $j \in \mathcal{J}_{it}$ that are stipulated to govern client i's product choice at time t. This function may or may not be different pre and post RDR; this may depend on whether the RDR induced some change in consumer awareness and literacy. One would expect that the set of products $\mathcal{J}_{it}$ presented to client i is chosen by the advisor such that all its elements entail risk commensurate with i's risk tolerance; so there ought not be any (or much) variation in product level risk across $\mathcal{J}_{it}$. Product attributes that are likely to matter are the product specific acquisition and exit costs, requirements on minimum holding period, liquidity. Collect these attributes as vector $\mathbf{x}_{ijt}$, and adopt again a linear specification for δ_{ijt}, say $\delta_{ijt} = \mathbf{x}'_{ijt}\theta_t$; here, it would also be possible to allow for heterogeneity across clients, i.e. to let the coefficient vector θ_{it} vary across clients, depending on socio-demographic characteristics (random coefficient model).

Estimation of this model would yield estimates of a number of parameters of interest for an empirical assessment of the effect the RDR had on the way financial advice is given and, more generally, on consumer level financial decision making:

(i) It would permit assessment of the extent to which advisors place different weight on product level risk attributes and remuneration implications before and after the policy initiative; such differences would be expected to manifest themselves in differences in the estimates $\hat{\mu}_t$ of μ_t for the two periods.

(ii) It would also permit assessment of whether advisees themselves change their evaluation of product attributes over time, as a result of enhanced awareness and financial literacy, possibly attributable to the nature of advice given on the presented choice options; such differences would be expected to manifest themselves in differences of the estimates $\hat{\theta}_t$ of θ_t for the two periods (or, in a random coefficient model, in the estimates of the distribution of θ_{it} across customers i for the two periods).

References

[1] Anderson, S.P. and A. de Palma (1992): "The Logit as a Model of product Differentiation", *Oxford Economic Paper*, **44**, 51-67

[2] Beckert, W., Christensen, M. and K. Collyer (2012): "Choice of NHS-Funded Hospital Services in England", *The Economic Journal*, Vol. 122, Issue 560, 400-417

[3] Besanko, D., Perry, M.K. and R.H. Spady (1990): "The Logit Model of Monopolistic Competition: Brand Diversity", *The Journal of Industrial Economics*, **38(4)**, 397-415

[4] Chamley, C.P. (2004): *Rational Herds: Economic Models of Social Learning*, Cambridge: Cambridge University Press

[5] Dafny, L., Ho, K. and M. Varela (2013): "Let Them Have Choice: Grains from Shifting Away From Employer-Sponsored Health Insurance and Toward and Individual Exchange", *American Economic Journal: Economic Policy*, **5(1)**, 32-58

[6] Dinerstein, M., Einav, L., Levin, J. and N. Sunderesan (20014): "Consumer Price Search and Platform Design in Internet Commerce", mimeo, Stanford University

[7] Gaynor, M., Ho, K. and R. Town (2014): "The Industrial Organization of Healthcare Markets", NBER working paper 19800

[8] Hausman, J.AS. and D.A. Wise (1979): "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment", *Econometrica*, **47(2)**, 455-473

[9] Heckman, J. (1976): "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models", *Annals of Economic and Social Measurement*, **5**, 475-492

[10] Maddala, G.S. (1983): *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

[11] McFadden, D.L. (1974): "Conditional logit analysis of qualitative choice behaviour", in P. Zarembka, ed.: *Frontiers in Economics*, 105-142, New York: Academic Press

[12] McFadden, D.L. (1977): "Modelling the Choice of Residential Location", *Cowles Foundation Discussion Paper* No. 477

[13] Santos, R., Gravelle, H. and C. Propper (2013): "Does quality affect patients' choice of doctor? Evidence from the UK", University of York, Centre for Health Economics, working paper No. 88

[14] Sovinsky Goeree, M. (2008): "Limited Information and Advertising in the US Personal Computer Industry", *Econometrica*, **76(5)**, 1017-74

[15] White, H. (1994): *Estimation, Inference and Specification Analysis*, Econometric Society Monographs No.22, Cambridge: Cambridge University Press