DISCUSSION PAPER SERIES

DP17292

Dynamic demand for differentiated products with fixed effects unobserved heterogeneity

Victor Aguirregabiria

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

Dynamic demand for differentiated products with fixed effects unobserved heterogeneity

Victor Aguirregabiria

Discussion Paper DP17292
Published 09 May 2022
Submitted 08 May 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Industrial Organization

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Victor Aguirregabiria

Dynamic demand for differentiated products with fixed effects unobserved heterogeneity

Abstract

This paper studies identification and estimation of a dynamic discrete choice model of demand for differentiated product using consumer-level panel data with few purchase events per consumer (i.e., short panel). Consumers are forward-looking and their preferences incorporate two sources of dynamics: last choice dependence due to habits and switching costs, and duration dependence due to inventory, depreciation, or learning. A key distinguishing feature of the model is that consumer unobserved heterogeneity has a Fixed Effects (FE) structure -- that is, its probability distribution conditional on the initial values of endogenous state variables is unrestricted. I apply and extend recent results to establish the identification of all the structural parameters as long as the dataset includes four or more purchase events per household. The parameters can be estimated using a sufficient statistic - conditional maximum likelihood (CML) method. An attractive feature of CML in this model is that the sufficient statistic controls for the forward-looking value of the consumer's decision problem such that the method does not require solving dynamic programming problems or calculating expected present values.

JEL Classification: C23, C25, C51, D12

Keywords: Structural dynamic discrete choice models, Dynamic demand of differentiated products, Dynamic panel data models, fixed effects, Habits, Switching Costs, Storable products, Durable products

Victor Aguirregabiria - victor.aguirregabiria@utoronto.ca

University of Toronto and CEPR

Acknowledgements

I have benefited from helpful discussions with Jesus Carro, Jiaying Gu, Francis Guiton, Nail Kashaev, Lance Lochner, and Matthew Osborne. I would like to thank also comments from the audience in the Applied Economics and Econometrics seminar at Western University, and in the special session on Econometrics of Dynamic Discrete Choice at the 2021 conference of the Royal Economic Society.

Dynamic demand for differentiated products with fixed-effects unobserved heterogeneity*

Victor Aguirregabiria†
University of Toronto, CEPR

May 8, 2022

Abstract

This paper studies identification and estimation of a dynamic discrete choice model of demand for differentiated product using consumer-level panel data with few purchase events per consumer (i.e., short panel). Consumers are forward-looking and their preferences incorporate two sources of dynamics: *last choice dependence* due to habits and switching costs, and *duration dependence* due to inventory, depreciation, or learning. A key distinguishing feature of the model is that consumer unobserved heterogeneity has a *Fixed Effects (FE)* structure – that is, its probability distribution conditional on the initial values of endogenous state variables is unrestricted. I apply and extend recent results to establish the identification of all the structural parameters as long as the dataset includes four or more purchase events per household. The parameters can be estimated using a *sufficient statistic - conditional maximum likelihood* (CML) method. An attractive feature of CML in this model is that the sufficient statistic controls for the forward-looking value of the consumer's decision problem such that the method does not require solving dynamic programming problems or calculating expected present values.

Keywords: Structural dynamic discrete choice models; Dynamic demand of differentiated products; Panel data; Fixed effects; Habits; Switching costs; Storable products; Durable products.

JEL codes: C23, C25, C51, D12.

*I have benefited from helpful discussions with Jesús Carro, Jiaying Gu, Francis Guiton, Nail Kashaev, Lance Lochner, and Matthew Osborne. I would like to thank also comments from the audience in the Applied Economics and Econometrics seminar at Western University, and in the special session on Econometrics of Dynamic Discrete Choice at the 2021 conference of the Royal Economic Society.

†Department of Economics, University of Toronto. 150 St. George Street, Toronto, ON, M5S 3G7, Canada, victor.aguirregabiria@utoronto.ca.

1 Introduction

In many markets, consumer demand is dynamic in the sense that consumers' utility depends on past decisions. Sources of dynamics in demand include, among others, storable products,[1] durable products,[2] habit formation and switching costs,[3] adoption costs,[4] and learning.[5] These models exhibit two main forms of *state dependence* in consumers' purchasing decisions: *last-choice dependence* and *duration dependence*. We have *last-choice dependence* if a consumer's previous purchase of a product has a causal effect on her current probability of buying that product, for instance, as a result of habits, switching costs, or learning. We have *duration dependence* if the time elapsed since the last purchase has a causal effect on the current buying decision. Depletion of storable products and depreciation of durable products generate duration dependence in consumer demand. These forms of *state dependence* can induce substantial differences between short-run and long-run responses of demand to price changes. This has important economic implications in applications such as evaluating the effects of taxes and subsidies, measurement of firms' market power, or consumer welfare. The estimation of dynamic structural demand models using consumer panel data tries to measure these causal effects and use them for counterfactual analysis and welfare evaluation.

Unobserved heterogeneity plays a fundamental role in dynamic demand models using consumer panel data. Ignoring or incorrectly specifying the correlation between unobserved heterogeneity and pre-determined explanatory variables (e.g., previous purchasing decisions) can generate important biases in the estimation of the structural parameters that capture dynamic causal effects (Heckman, 1981b). Furthermore, the distribution of consumer taste heterogeneity has an important impact on demand price elasticities (Berry, Levinsohn, and Pakes, 1995). The empirical literature on dynamic demand of differentiated product has considered a *Random Effects (RE)* approach to model consumer unobserved heterogeneity. *RE* models impose parametric restrictions on the distribution of unobserved heterogeneity and on the correlation

[1] See Boizot, Robin, and Visser (2001), Pesendorfer (2002), Erdem, Imai, and Keane (2003), and Hendel and Nevo (2006).

[2] See Esteban and Shum (2007), Goettler and Gordon (2011), and Gowrisankaran and Rysman (2012).

[3] See Roy, Chintagunta, and Haldar (1996), Keane (1997), and Osborne (2011).

[4] See Ryan and Tucker (2012), and De Groote and Verboven (2019).

[5] See Ackerberg (2003), and Ching (2010).

between these unobservables and the initial values of the predetermined explanatory variables. Though this distribution is not identifiable in short panels, its misspecfication can generate important biases in the estimation of dynamic causal effects. This is the so called *initial conditions problem* (Heckman, 1981a). Therefore, *RE* models are not robust to misspecification of parametric restrictions on unobserved heterogeneity.

Fixed effects (FE) approaches impose no restriction on the distribution of consumers' unobserved heterogeneity such that the identification of parameters of interest is more robust than in *RE* models. However, several identification concerns have inclined researchers to avoid a *FE* approach in applications of dynamic discrete choice structural models, and more specifically, in dynamic demand models.

A first issue is related to the identification of structural parameters. All the existing (positive) identification results in *FE* dynamic discrete choice models impose the restriction that unobserved heterogeneity enters additively in the utility of each choice alternative. However, in dynamic programming models, unobserved heterogeneity enters not only in current utility but also in the continuation value of the forward-looking decision problem, and these continuation values depend non-additively (and in fact, without a closed-form expression) on both unobserved heterogeneity and observable state variables. The common wisdom was that *FE* models cannot deal with the non-additive unobserved heterogeneity that is inherent to discrete choice dynamic programming models.

A second important issue is related to the fact that *FE* methods cannot deliver identification of the distribution of unobserved heterogeneity with short-panels. Most empirical applications of dynamic structural models are interested in using the estimated model to obtain *Average Partial Effects (APE)* on endogenous variables of changes in explanatory variables or in structural parameters. Demand price elasticities are examples of these *APE*s. The common wisdom was that APEs are not identified in *FE* models, as they are expectations over the distribution of the unobserved heterogeneity (e.g., Abrevaya and Hsu, 2021; Honoré and DePaula, 2021).

Two recent studies provide positive identification results for the two issues discussed above. Aguirregabiria, Gu, and Luo (2021) establish the identification of structural parameters in a class of *FE* dynamic panel data logit models where agents are forward-looking. The class of models includes two types of endogenous state variables: the lagged choice variable, and the time duration in the last choice. Aguirregabiria and Carro (2021) prove the point identification of different AMEs in *FE* dynamic logit models. For instance, the average causal effect of changes in the lagged dependent variable or in the duration in last choice are identified.[6]

In this paper, I apply and extend identification results in Aguirregabiria, Gu, and Luo (2021) to a *FE* dynamic panel data model of consumer demand with differentiated product. The model can incorporate storable or durable products, habit formation, and brand switching costs. I present a new identification result in FE dynamic discrete choice with forward-looking agents. I show identification of utility parameters associated to state variables that follow exogenous stochastic processes. More specifically, in the context of demand models, I establish the identification of parameters that capture the effect of prices. In a FE forward-looking model, this identification result relies on a particular structure in the stochastic process of prices. There are two components in prices: a persistent component (i.e., regular price), and a transitory component (i.e., temporary promotions).

The structural parameters of the model are estimated using a *sufficient statistics - conditional maximum likelihood (CML)* method, in the spirit of Chamberlain (1985) and Honoré and Kyriazidou (2000). In the context of dynamic structural models, a very helpful implication of this CML method is that it is not subject to a *curse of dimensionality* associated with the computational cost of solving a dynamic programming problem, or calculating present values of future utilities. This is particularly relevant in applications of dynamic demand of differentiated product as the dimension of the state space increases exponentially with the number of products, which is typically large. The *sufficient statistics - CML* method "differences out" the continuation (forward-looking) value of the consumer's decision problem, as this value depends on the incidental parameters / unobserved heterogeneity. This implies that continuation values should not be computed to implement the CML estimator. Therefore, the CML estimation of this dynamic structural model is computationally as simple as in static or myopic models, and its cost does not depend on the dimension of the state space.[7]

[6]In related work, Chernozhukov, Fernandez-Val, Hahn, and Newey (2013), and more recently Davezies, D'Haultfoeuille, and Laage (2021), provide partial identification results for a general class of AMEs in binary choice panel data models.

[7]CML estimation implies a different type of curse of dimensionality. The evaluation of the conditional likeli-

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 presents identification results. Section 4 describes the estimation method. In section 5, I summarize and discuss some extensions for further research.

2 Model

I present a framework that includes both storable and durable differentiated products. Most of the features of the model are standard in the literature. The main distinguishing feature is the *Fixed Effects (FE)* nature of consumer unobserved heterogeneity.

2.1 Basics

There are J different brands of a differentiated product, and we index brands by $j \in \{1, 2, ..., J\}$. We index consumers by i. Time is discrete and indexed by t. There is a dichotomy in the definition of *time* in empirical applications of dynamic consumer demand. In most applications, t has the standard interpretation as calendar time: for instance, the unit of time can be a week. In these applications, the model and data account for time periods where a consumer does not purchase any variety of the differentiated product. In contrast, a good number of empirical applications in this literature consider t as the index for purchase events, such that $t = 1$ means a consumer's first purchase, $t = 2$ is her second purchase, and so on (see Keane, 1997; Osborne, 2011; among others). The different definition of time t has implications on the interpretation of dynamics in the model, such as duration dependence.[8] Here, I follow the standard definition of t as calendar time. However, all the identification results in this paper apply also to models where t indexes purchase events.

Every period t, consumer i decides whether to purchase or not one unit of the product, and which brand to buy. Variable $y_{it} \in \mathcal{Y} = \{0, 1, ..., J\}$ represents the decision of consumer i at

hood function involves adding up a function over all the possible values of the vector of sufficient statistics. The number of possible values increases exponentially with the number of time periods in the data, T. However, this dimensionality problem has an easy solution – at the cost of some efficiency lost in the CML estimator – which consists of splitting a T-periods history into shorter sub-histories.

[8] Also very importantly, the model where t indexes purchase events does not include the decision of "no purchase" as a choice alternative. That is, there is not an *outside alternative*.

period t, where $y_{it} = 0$ means "no purchase", and $y_{it} = j > 0$ means purchase of brand j. There are two endogenous state variables that capture dynamics in consumer demand: brand choice in the last purchase, $\ell_{it} \in \{1, 2, ..., J\}$; and time duration since the last purchase, $d_{it} \in \{1, 2, ..., D\}$. Last brand choice is related to habit formation and switching costs. Duration since last purchase captures the effect of inventory depletion (for storable products) or depreciation (for durable products). By definition, the transition rule of the vector of endogenous state variables $\mathbf{x}_{it} \equiv (\ell_{it}, d_{it})$ is:

$$\mathbf{x}_{i,t+1} \equiv (\ell_{i,t+1} \ , \ d_{i,t+1}) \ = \ f_x(y_{it}, \mathbf{x}_{it}) \ \equiv \ \begin{cases} (\ell_{it} \ , \ d_{it} + 1) & if \quad y_{it} = 0 \\ (j \ , \ 1) & if \quad y_{it} = j > 0 \end{cases} \tag{1}$$

In many applications, consumers are located in different geographic markets and face different prices. This is not a necessary condition for the identification results in this paper, but I allow for it. Let $p_{it}(j)$ be the price of product j at period t in the market where consumer i is located, and let $\mathbf{p}_{it}$ be the vector with the prices of all products, $\mathbf{p}_{it} \equiv (p_{it}(j) : j = 1, 2, ..., J)$.

2.2 Utility

Let U_{it} be the per-period utility that consumer i obtains at period t. It has four components:

$$U_{it} \ = \ b_{it} \ + \ \gamma_i \ m_{it} \ - \ sc_{it} \ + \ \varepsilon_{it} \tag{2}$$

The term b_{it} is the utility from consumption of the branded product; m_{it} represents consumption of the outside good or numeraire; γ_i is the marginal utility of the outside good, also referred as *marginal utility of income*; sc_{it} represents brand switching costs or habits; and ε_{it} represents consumer idiosyncratic taste shocks at the moment of purchase. I describe below each of these components.

(i) Heterogeneity in consumers' tastes. The model allows for product differentiation at the moment of consumption and not only at the moment of purchase. Parameter $\alpha_i(j)$ is the flow utility that consumer i receives from consuming product j when the product has not depreciated (or depleted) over time. It depends on product characteristics and consumer characteristics, some of them unobservable to the researcher. The vector $\boldsymbol{\alpha}_i \equiv (\alpha_i(1), \alpha_i(2), ..., \alpha_i(J))$ represents the fixed effects for consumer i. The fixed effect $\alpha_i(j)$ depends on product and consumer characteristics in an unrestricted way. For instance, it could have the structure $\alpha_i(j) = \boldsymbol{x}_j' \boldsymbol{\beta}_i + \boldsymbol{\xi}_j' \boldsymbol{\omega}_i$, where $\boldsymbol{x}_j$ and $\boldsymbol{\xi}_j$ are vectors of observable and unobservable product characteristics, respectively, and $\boldsymbol{\beta}_i$ and $\boldsymbol{\omega}_i$ are the vectors of marginal utilities of these characteristics for consumer i. Therefore, the specification of consumer heterogeneity in preferences is very flexible.

(ii) Depreciation / depletion. The flow utility from consumption of the branded product takes into account depletion or depreciation over time.[9] This flow utility has the following form:

$$b_{it} \;\equiv\; \begin{cases} \alpha_i(j) & if \quad y_{it} = j > 0 \\ \alpha_i(\ell_{it}) - \beta^{dep}(\ell_{it}, d_{it}) & if \quad y_{it} = 0\ . \end{cases} \tag{3}$$

If the consumer purchases a new product (if $y_{it} = j > 0$), there is no depreciation effect and the flow utility is $\alpha_i(j)$. Otherwise, if the consumer does not purchase a new product (if $y_{it} = 0$), she consumes product ℓ_{it} – her last purchase – and flow utility is $\alpha_i(\ell_{it}) - \beta^{dep}(\ell_{it}, d_{it})$. Parameter $\beta^d(j, d)$ measures the effect on utility of d periods of depreciation in product j. This depreciation/depletion effect can vary across products, and it is a nonparametric function of duration d.

Arguably, the depletion effect β^{dep} could vary across households. It is important to note that the model in this paper allows all the structural parameters, including this depletion effect, to depend on observable household characteristics, such as family size, income, education, or region. We impose the restriction that β^{dep} does not depend on consumer unobserved heterogeneity.

(iii) Utility from consumption of the composite good. The (dollar) amount of consumption of the composite good, m_{it}, is given by the consumer's budget constraint:

$$m_{it} \;=\; m_i - \sum_{j=1}^{J} p_{it}(j)\ 1\{y_{it} = j\}, \tag{4}$$

[9]The datasets used in this literature contain information on consumer purchase histories, but not on consumption and inventories of storable products. To deal with this missing information for storable products, researchers have imposed different restrictions. Aguirregabiria and Nevo (2013) discuss this issue and different restrictions in the literature. A common approach is using duration since last purchase, d_{it}, as a proxy of inventory. For storable products, this is the approach in Erdem, Imai, and Keane (2003).

where m_i is the consumer's (weekly) disposable income, and $\sum_{j=1}^{J} p_{it}(j)\ 1\{y_{it} = j\}$ represents expenditure in the branded product. Since this utility is linear in consumption of the composite good, income m_i does not enter explicitly in the difference between the utilities of two product choices. However, the parameter γ_i that represents the marginal utility of the composite good may depend on consumer income, or on other observable consumer characteristics.

(iv) Product switching costs / habits. Following the standard approach, brand switching costs take place at the moment of purchase, and not through the consumption of the product. The specification of brand switching costs is:

$$sc_{it} \ = \ \beta^{sc}(\ell_{it}, y_{it}), \tag{5}$$

Parameter $\beta^{sc}(k, j)$ represents the cost of switching from brand k to brand j. There is no switching cost without switching or without a new purchase, such that such that $\beta^{sc}(k, k) = 0$ and $\beta^{sc}(k, 0) = 0$.

This component of the utility can be also interpreted in terms of *habit formation*. In that case, we have that $sc_{it} = -\beta^{hab}(\ell_{it}, y_{it})$, where parameter $\beta^{hab}(j, j) \geq 0$ for $j > 0$ is the increase in utility from purchasing the same brand as in last purchase, and for any $k \neq j$, $\beta^{hab}(k, j)$ is restricted to be zero. This *habit formation* model is equivalent to the *brand switching cost* model under the restrictions $\beta^{sc}(k, j) = \beta^{sc}(k', j)$ for any $k, k' \neq j$. This more parsimonious specification is quite common in this literature.

Regardless of consumer unobserved heterogeneity or/and forward-looking behavior, it is not possible to identify all the switching cost parameters $\beta^{sc}(k, j)$ using data of consumers' product choices. In the best case, for any pair of products (k, j), it is possible to identify the following linear combination of these parameters:

$$\widetilde{\beta}^{sc}(k, j) \ = \ \beta^{sc}(k, j) + \beta^{sc}(j, k) - \beta^{sc}(k, k) - \beta^{sc}(j, j) \tag{6}$$

Therefore, we study the identification of these $\widetilde{\beta}^{sc}(k, j)$ parameters.

(v) Extreme value distributed shocks in preferences. Finally, variable $\varepsilon_{it}(j)$ captures other consumer idiosyncratic factors affecting the utility from purchasing product j at period t. Variables $\boldsymbol{\varepsilon}_{it} \equiv (\varepsilon_{it}(0), ..., \varepsilon_{it}(J))$ are i.i.d. over (i,t,j) with a type I extreme value distribution. Both $\boldsymbol{\alpha}_i$ and $\boldsymbol{\varepsilon}_{it}$ are unobservable to the econometrician.

Putting together the different components of the utility function, we have:

$$U_{it} = u_{\boldsymbol{\alpha}_i}(y_{it}, \mathbf{x}_{it}, \mathbf{p}_{it}) + \varepsilon_{it}(y_{it}) = \begin{cases} \alpha_i(\ell_{it}) - \beta^{dep}(\ell_{it}, d_{it}) + \varepsilon_{it}(0) & if \quad y_{it} = 0 \\ \\ \alpha_i(j) - \gamma_i \; p_{it}(j) - \beta^{sc}(\ell_{it}, j) + \varepsilon_{it}(j) & if \quad y_{it} = j > 0, \end{cases} \tag{7}$$

where $u_{\boldsymbol{\alpha}_i}(y_{it}, \mathbf{x}_{it}, \mathbf{p}_{it})$ represents utility excluding unobservable logit shocks.

2.3 Stochastic process for prices

The structure of the stochastic process of prices that I describe here is not necessary for the identification of the parameters β^{sc} and β^{dep}, but it plays an important role in the identification of the marginal utility of income parameter, γ, in this forward-looking model. The assumption is that the price of any product j has two components, one that is persistent over time $(z_{it}(j))$, and other that is transitory $(e_{it}(j))$. That is:

$$p_{it}(j) \;\; = \;\; \rho\left(z_{it}(j), \;\; e_{it}(j)\right) \tag{8}$$

where $\rho(.)$ is a known function. Some examples for the $\rho(.)$ function are the linear function $p_{it}(j) = z_{it}(j) + e_{it}(j)$, or the linear in logs function, $p_{it}(j) = \exp\{z_{it}(j) e_{it}(j)\}$. The vector of permanent components $\mathbf{z}_{it} \equiv (z_{it}(j) : j = 1, 2, ..., J)$ follows a Markov process of order n. The vector of transitory components $\mathbf{e}_{it} \equiv (e_{it}(j) : j = 1, 2, ..., J)$ is i.i.d. over time, and for any periods t and s, $\mathbf{z}_{it}$ and $\mathbf{e}_{is}$ are independently distributed. The stochastic relationship between the prices of the different products is unrestricted.

This structure can be interpreted in terms of "regular prices", represented by $\mathbf{z}_{it}$, and transitory promotions, represented by $\mathbf{e}_{it}$. For notational simplicty, for the rest of the paper, I consider that $\mathbf{z}_{it}$ follows a first-order Markov process with transition density function $f_z(\mathbf{z}_{i,t+1}|\mathbf{z}_{it})$.

This structure has an important implication on a consumer's dynamic decision model. The whole vector of prices $\mathbf{p}_t$ (both $\mathbf{z}_t$ and $\mathbf{e}_t$) affects a consumer's current utility, but the expected and discounted value of future utilities (the continuation value) depends on $\mathbf{z}_t$ but not on $\mathbf{e}_t$. This exclusion restriction plays an important role in the identification of parameter γ in the Fixed Effects dynamic forward-looking model in this paper.

It is relevant to note that, given time series data on prices and a specification of the $\rho(.)$ function, it is possible to identify the parameters in the two stochastic processes that characterize the evolution of prices, and based on these parameters, it is possible to identify the two components $\mathbf{z}_{it}$ and $\mathbf{e}_{it}$. For the rest of the paper, I assume that these two vectors, $\mathbf{z}_{it}$ and $\mathbf{e}_{it}$, are observable to the researcher.

2.4 Consumer decision problem

Every period t, the consumer observes the vectors of state variables $\mathbf{x}_{it}$, $\mathbf{z}_{it}$, $\mathbf{e}_{it}$, and $\boldsymbol{\varepsilon}_{it}$, and makes a purchasing decision y_{it} to maximize her expected and discounted intertemporal utility $\mathbb{E}_t\left[\sum_{s=0}^{\infty} \delta_i^s \; U_{i,t+s}\right]$, where $\delta_i \in [0,1)$ is consumer i's time discount factor. This consumer's problem is a stationary *Markov Decision Process (MDP)*, and Blackwell's Theorem establishes that the value function and the optimal decision rule are time-invariant (Blackwell, 1965).

The decision problem of consumer i at period t is:

$$y_{it} \;=\; \arg\max_{j\in\mathcal{Y}} \left\{ \; u_{\boldsymbol{\alpha}_i}(j, \mathbf{x}_{it}, \mathbf{p}_{it}) \;+\; \varepsilon_{it}(j) \;+\; v_{\boldsymbol{\alpha}_i}(f_x(j,\mathbf{x}_{it}), \mathbf{z}_{it}) \; \right\} \tag{9}$$

where, as defined in equation (1), $f_x(j, \mathbf{x}_{it})$ represents the value of $\mathbf{x}_{i,t+1}$ given state $\mathbf{x}_{it}$ and decision $y_{it} = j$; and $v_{\boldsymbol{\alpha}_i}(f_x(j,\mathbf{x}_{it}), \mathbf{z}_{it})$ is the *continuation value function*, i.e., the expected and discounted value of future utility given current state is $(\mathbf{x}_{it}, \mathbf{z}_{it})$ and current choice j.[10] Let $\sigma_{\boldsymbol{\alpha}_i}(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it})$ be the integrated (or smoothed) value function, that is defined as the expectation of the value function over the distribution of the i.i.d. unobservable state variables $\boldsymbol{\varepsilon}_{it}$ (Rust, 1994).[11] This value function is the unique solution of the integrated Bellman equation. For the

[10] In the literature of structural dynamic discrete models, letter v is commonly used to denote the *conditional choice value function*. Instead, here I use letter v to denote the continuation value, such that the *conditional choice value function* is $u_{\boldsymbol{\alpha}_i}(j, \mathbf{x}_{it}, \mathbf{p}_{it}) + v_{\boldsymbol{\alpha}_i}(f_x(j,\mathbf{x}_{it}), \mathbf{z}_{it})$.

[11] That is, if $V_{\boldsymbol{\alpha}_i}(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}, \boldsymbol{\varepsilon}_{it})$ is the value function, then the integrated value function is defined as $\sigma_{\boldsymbol{\alpha}_i}(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}) \equiv \int V_{\boldsymbol{\alpha}_i}(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}, \boldsymbol{\varepsilon}_{it}) \; f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}_{it}) \; d\boldsymbol{\varepsilon}_{it}$, where $f_{\boldsymbol{\varepsilon}}$ is the density function of $\boldsymbol{\varepsilon}_{it}$.

type I extreme value distribution, the integrated Bellman equation is:

$$\sigma_{\boldsymbol{\alpha}_i}(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}) \;=\; \log\left(\sum_{j=0}^{J} \exp\left\{u_{\boldsymbol{\alpha}_i}(j, \mathbf{x}_{it}, \mathbf{p}_{it}) \;+\; v_{\boldsymbol{\alpha}_i}(f_x(j, \mathbf{x}_{it}), \mathbf{z}_{it})\right\}\right) \tag{10}$$

and the *continuation value function* is:

$$v_{\boldsymbol{\alpha}_i}(f_x(j, \mathbf{x}_{it}), \mathbf{z}_{it}) \;\equiv\; \delta_i \int_{\mathbf{z}_{i,t+1}} \int_{\mathbf{e}_{i,t+1}} \sigma_{\boldsymbol{\alpha}_i}(f_x(j, \mathbf{x}_{it}), \mathbf{z}_{i,t+1}, \mathbf{e}_{i,t+1}) \; f_z(d\mathbf{z}_{i,t+1}|\mathbf{z}_{it}) \; f_e(d\mathbf{e}_{i,t+1}) \tag{11}$$

Define the *conditional choice probability* (CCP) function as:

$$P(j|\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}, \boldsymbol{\alpha}_i) \equiv Pr\left(j = \arg\max_{k \in \mathcal{Y}} \left\{u_{\boldsymbol{\alpha}_i}(k, \mathbf{x}_{it}, \mathbf{p}_{it}) + \varepsilon_{it}(k) + v_{\boldsymbol{\alpha}_i}(f_x(k, \mathbf{x}_{it}), \mathbf{z}_{it})\right\} |\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}, \boldsymbol{\alpha}_i\right) \tag{12}$$

Importantly, this CCP function is conditional on the observable state variables $(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it})$ and on the unobservable fixed effects $\boldsymbol{\alpha}_i$. The extreme value type I distribution of the unobservables $\boldsymbol{\varepsilon}$ implies that the logarithm of the CCP function has the following form. For any $j \in \mathcal{Y}$:

$$\log P\left(j|\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}, \boldsymbol{\alpha}_i\right) \;=\; u_{\boldsymbol{\alpha}_i}(j, \mathbf{x}_{it}, \mathbf{p}_{it}) \;+\; v_{\boldsymbol{\alpha}_i}(f_x(j, \mathbf{x}_{it}), \mathbf{z}_{it}) \;-\; \sigma_{\boldsymbol{\alpha}_i}(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}) \tag{13}$$

Note that the integrated value $\sigma_{\boldsymbol{\alpha}_i}(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it})$ is also the logarithm of the denominator of the logit CCP function.

3 Identification

The researcher observes a panel dataset of N households over T periods with information on households' purchasing decisions and prices. The time length of the panel, T, is short in the sense that it contains only a few purchases per household. The identification results in this paper consider that T is fixed and – as it is common in proofs of identification – that N is infinite such that we have an infinite population of households.

Assumption 1 summarizes restrictions on the model for the identification of structural parameters $\widetilde{\beta}^{sc}$, β^{dep}, and γ.

ASSUMPTION 1. *(A) (i.i.d. Logit shocks)* $\varepsilon_{it}(j)$ *is i.i.d. over* (i, t, j) *with type I extreme value distribution, and is independent of* $\boldsymbol{\alpha}_i$*. (B) (Strict exogeneity of prices with respect to shocks* $\boldsymbol{\varepsilon}_{it}$*) For any two periods,* t *and* s*, the variables* $\varepsilon_{it}(j)$ *and prices* $\mathbf{p}_{is}$ *are independently distributed. (C) (Sticky prices) The vectors* $\mathbf{z}_t - \mathbf{z}_{t-1}$ *and* $\mathbf{e}_t - \mathbf{e}_{t-1}$ *have supports that include a neighborhood around zero.* ■

Assumptions 1(A) and 1(B) rule out aggregate market-level shocks correlated with prices. This restriction is very common in the literature of structural dynamic demand models, especially in applications using consumer level data.[12] This may seem a strong assumption in an econometric demand model, and at odds with the literature on estimation of static demand models using the BLP framework (Berry, Levinsohn, and Pakes, 1995). However, as I explain in the next two paragraphs, the FE approach in this paper controls for two important sources of endogeneity in prices.

First, note that the vector of fixed effects $\boldsymbol{\alpha}_i = (\alpha_i(1), \alpha_i(2), ..., \alpha_i(J))$ includes product fixed effects: in fact, it accounts for any interaction of product effects and consumer effects. Therefore, these incidental parameters account for time-invariant differences in product quality that can be correlated with the cross-sectional variation in prices across products. That is, the model accounts – in a very general way – for endogeneity of price levels. This means that the potential concern with Assumption 1(B) is because endogeneity of changes in prices over time.

Second, all the identification results in this paper control for variation over time in the persistent component of prices, $\mathbf{z}_{it}$ and exploit only variation in the transitory component, $\mathbf{e}_{it}$. Therefore, the method controls for endogeneity of *regular prices*.

Importantly, we consider a fixed effects (FE) model, in the sense that both the unconditional distribution $F_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}_i)$ and the sequence of conditional distributions $F_{\boldsymbol{\alpha}|\mathbf{x},t}(\boldsymbol{\alpha}_i|\ell_{it}, d_{it}, \mathbf{z}_{it}, \mathbf{e}_{it})$ for $t = 1, 2, ..., T$ are completely unrestricted functions.

[12] For instance, Erdem, Imai, and Keane (2003), Hendel and Nevo (2006), Pavlidis and Ellickson, (2017), Griffith, Nesheim, and O'Connell, (2018), or Mysliwski, Sanches, Junior, and Srisuma, (2020), among many other papers in this literature, assume that prices are not correlated with transitory shocks. In contrast, some dynamic demand models using market level data, such as Gowrisankaran and Rysman (2012), allow for unobserved market-level demand shocks.

3.1 Preliminaries

Let $\boldsymbol{\theta}$ be the vector with the structural parameters β^{sc}, β^{d}, and γ. To establish the identification of $\boldsymbol{\theta}$, I follow Aguirregabiria, Gu, and Luo (2021) who consider a *sufficient statistic - conditional likelihood* approach in the spirit of Chamberlain (1985), Magnac (2000), and Honoré and Kyriazidou (2000). I start describing some general features of this approach in the context of the dynamic demand model.

Let $\mathbf{y}_i \equiv (\ell_{i1}, d_{i1}, y_{i1}, y_{i2}, ..., y_{iT})$ be the vector with consumer i's choice history, including the initial condition (ℓ_{i1}, d_{i1}), and let $\widetilde{\mathbf{z}}_i \equiv (\mathbf{z}_{i1}, \mathbf{z}_{i2}, ..., \mathbf{z}_{iT})$ and $\widetilde{\mathbf{e}}_i \equiv (\mathbf{e}_{i1}, \mathbf{e}_{i2}, ..., \mathbf{e}_{iT})$ be the vectors with the histories of the two components in prices. Taking into account equation (13) for the log-CCP function, we have that the log-probability of a choice history $\mathbf{y}_i$ conditional on the history of prices and parameters $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}_i$ is:

$$\log \mathbb{P}\left(\mathbf{y}_i|\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, \boldsymbol{\alpha}_i, \boldsymbol{\theta}\right) = \log p_1(\mathbf{x}_{i1}|\boldsymbol{\alpha}_i) + \sum_{t=1}^{T} u_{\boldsymbol{\alpha}_i}(j, \mathbf{x}_{it}, \mathbf{p}_{it}) + v_{\boldsymbol{\alpha}_i}(f_x(j, \mathbf{x}_{it}), \mathbf{z}_{it}) - \sigma_{\boldsymbol{\alpha}_i}\left(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}\right) \tag{14}$$

where $p_1(.)$ is the probability of the endogenous state variables at period $t = 1$. In this model, the log-probability of a choice history has the following structure:

$$\log \mathbb{P}\left(\mathbf{y}_i|\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, \boldsymbol{\alpha}_i, \boldsymbol{\theta}\right) \;=\; \mathbf{s}(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)' \, \mathbf{g}(\boldsymbol{\alpha}_i) + \mathbf{c}(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)' \, \boldsymbol{\theta} \tag{15}$$

where $\mathbf{s}(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)$ and $\mathbf{c}(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)$ are vectors of statistics which are functions of $(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)$, such as, for instance, $\sum_{t=1}^{T} 1\{y_{it} = j\}$ or $\sum_{t=1}^{T} 1\{y_{it} = j\} p_{it}(j)$, and $\mathbf{g}(\boldsymbol{\alpha}_i)$ is a vector of functions which include the fixed effects $\boldsymbol{\alpha}_i$.

The structure in equation (15) has two key implications for the identification of $\boldsymbol{\theta}$. First, equation (15) implies that $\mathbf{s}_i \equiv \mathbf{s}(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)$ is a *sufficient statistic* for $\boldsymbol{\alpha}_i$. That is, the probability $\mathbb{P}(\mathbf{y}_i|\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, \boldsymbol{\alpha}_i, \boldsymbol{\theta}, \mathbf{s}_i)$ does not depend on $\boldsymbol{\alpha}_i$. To show this, note that:

$$\begin{aligned} \mathbb{P}(\mathbf{y}_i|\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, \boldsymbol{\alpha}_i, \boldsymbol{\theta}, \mathbf{s}_i) \;&=\; \frac{\mathbb{P}(\mathbf{y}_i|\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, \boldsymbol{\alpha}_i, \boldsymbol{\theta})}{\mathbb{P}(\mathbf{s}_i|\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, \boldsymbol{\alpha}_i, \boldsymbol{\theta})} = \frac{\exp\left\{\mathbf{s}_i' \, \mathbf{g}(\boldsymbol{\alpha}_i) + \mathbf{c}_i' \, \boldsymbol{\theta}\right\}}{\sum_{\mathbf{y}:\, \mathbf{s}(\mathbf{y}) = \mathbf{s}_i} \exp\left\{\mathbf{s}_i' \, \mathbf{g}(\boldsymbol{\alpha}_i) + \mathbf{c}(\mathbf{y})' \, \boldsymbol{\theta}\right\}} \\ &=\; \frac{\exp\left\{\mathbf{c}_i' \, \boldsymbol{\theta}\right\}}{\sum_{\mathbf{y}:\, \mathbf{s}(\mathbf{y}) = \mathbf{s}_i} \exp\left\{\mathbf{c}(\mathbf{y})' \, \boldsymbol{\theta}\right\}} = \mathbb{P}(\mathbf{y}_i|\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, \boldsymbol{\theta}, \mathbf{s}_i) \end{aligned} \tag{16}$$

where $\sum_{\mathbf{y}:\ \mathbf{s}(\mathbf{y})=\mathbf{s}_i}$ is the sum over all the possible choice histories $\mathbf{y}$ with $\mathbf{s}(\mathbf{y})$ equal to $\mathbf{s}_i$.

Second, using the expression at the bottom line of equation (16), we have the following conditional log-likelihood function (at the population level):

$$\mathbb{E}\left[\log \mathbb{P}(\mathbf{y}_i|\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, \boldsymbol{\theta}, \mathbf{s}_i)\right] \;=\; \mathbb{E}\left[\mathbf{c}_i'\,\boldsymbol{\theta} \;-\; \log\left(\sum_{\mathbf{y}:\ \mathbf{s}(\mathbf{y})=\mathbf{s}_i} \exp\left\{\mathbf{c}(\mathbf{y})'\,\boldsymbol{\theta}\right\}\right)\right] \tag{17}$$

The first order conditions for the maximization of this likelihood function with respect to $\boldsymbol{\theta}$ imply the following moment conditions (i.e., likelihood equations):

$$\boldsymbol{m}(\boldsymbol{\theta}) \;\equiv\; \mathbb{E}\left[\mathbf{c}_i \;-\; \sum_{\mathbf{y}:\ \mathbf{s}(\mathbf{y})=\mathbf{s}_i} \mathbf{c}(\mathbf{y})\; \mathbb{P}(\mathbf{y}|\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, \boldsymbol{\theta}, \mathbf{s}_i)\right] \;=\; \mathbf{0} \tag{18}$$

The Jacobian matrix for this vector of moment conditions – or equivalently, the Hessian of the log-likelihood function, or the negative of Fisher's information matrix – is:

$$\frac{\partial \boldsymbol{m}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \;=\; -\,\mathbb{E}\left(\left[\mathbf{c}_i - \mathbb{E}(\mathbf{c}_i|\mathbf{s}_i)\right]\;\left[\mathbf{c}_i - \mathbb{E}(\mathbf{c}_i|\mathbf{s}_i)\right]'\right) \tag{19}$$

This is the negative of a variance-covariance matrix, and therefore, it is negative semidefinite for any value of $\boldsymbol{\theta}$. In other words, the likelihood function is globally concave in $\boldsymbol{\theta}$. Furthermore, based on these moment conditions, a necessary and sufficient condition for (point) identification of $\boldsymbol{\theta}$ is that this Jacobian matrix $\partial\boldsymbol{m}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}'$ is nonsingular (i.e., rank identification condition). The right-hand-side in equation (19) shows that this is the case if and only if the statistics in vector $\mathbf{c}_i$ are linearly independent conditional on $\mathbf{s}_i$.

A more intuitive description of the identification of the structural parameters consists in showing that, for every parameter in the vector $\boldsymbol{\theta}$, say θ_k (the k-th element of vector $\boldsymbol{\theta}$), there exist two choice histories, say A and B, such that $\mathbf{s}(A) = \mathbf{s}(B)$ and $\mathbf{c}(A) - \mathbf{c}(B)$ is a vector where all the elements are zero except element k that is one. Under these conditions, equation (15) implies that:

$$\theta_k \;=\; \log \mathbb{P}(A) - \log \mathbb{P}(B), \tag{20}$$

which shows that parameter θ_k is identified from the log odds ratio of histories A and B.[13] This intuitive description of the identification of parameters in FE discrete choice models has been used by Chamberlain (1985) and Honoré and Kyriazidou (2000), among others. In this paper, I follow this approach.

It is important to note that for some FE nonlinear panel data models the identification conditions based on this conditional likelihood approach are sufficient but not necessary. Bonhomme (2012) provides a systematic approach (the *functional differencing* method) to construct moment restrictions for a general class of FE models. Recently, Honoré and Weidner (2020), Dobronyi, Gu, and Kim (2021), and Honoré, Muris, and Weidner (2021) have used Bonhomme's *functional differencing* approach to establish new point identification results of parameters in FE models where the conditional likelihood approach, at least apparently, does not provide identification.

For the rest of this subsection, I present results on the identification of structural parameters in two versions of the demand model: with and without duration dependence. Here, for simplicity, I do not present the expression for the statistics $\mathbf{s}_i$ and $\mathbf{c}_i$ and instead provide examples of pairs of histories A and B that identify the different parameters in $\boldsymbol{\theta}$.

3.2 Dynamic demand without duration dependence

Suppose that there is not depreciation or depletion such that $\beta^d(\ell, d) = 0$, but there is habit formation or/and brand switching costs, such that $\beta^{sc}(\ell, y) \neq 0$. This is the demand model in empirical applications like Roy, Chintagunta, and Haldar (1996), Keane (1997), Osborne (2011), or Mysliwski, Sanches, Junior, and Srisuma, (2020), among others.

In this model, the only endogenous state variables is ℓ_{it}, and the transition function is:

$$\ell_{i,t+1} \;=\; f_x(y_{it}, \ell_{it}) \;=\; \begin{cases} \ell_{it} & if \quad y_{it} = 0 \\ j & if \quad y_{it} = j > 0 \end{cases} \tag{21}$$

Then, combining equation (13) with this transition rule, we have the following expression for

[13] Note that, in this class of models, every choice history has a strictly positive probability such that $\log \mathbb{P}(A)$ and $\log \mathbb{P}(B)$ are finite real numbers.

the log-CCP function:

$$\log P_{it}\left(y_{it}\right) = \begin{cases} \alpha_i(\ell_{it}) + v_{\boldsymbol{\alpha}_i}(\ell_{it}, \mathbf{z}_{it}) - \sigma_{\boldsymbol{\alpha}_i}\left(\ell_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}\right) & if \quad y_{it} = 0 \\ \alpha_i(j) + v_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_{it}) - \sigma_{\boldsymbol{\alpha}_i}\left(\ell_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}\right) - \beta^{sc}(\ell_{it}, j) - \gamma\ p_{it}(j) & if \quad y_{it} = j > 0 \end{cases} \tag{22}$$

Suppose that $T = 4$. Let k and j be two different products, and consider the following pair of choice histories:

$$A \ = \ (k,\ j,\ k,\ j) \quad ; \quad B \ = \ (k,\ k,\ j,\ j) \tag{23}$$

Taking into account the structure of the log-CCP function in equation (22), we have the following expression for the log-probabilities of the choice histories:

$$\begin{array}{rcl} \log \mathbb{P}(A) & = & \log p_1(k, \boldsymbol{\alpha}_i) + \alpha_i(j) + \alpha_i(k) + \alpha_i(j) \\ & + & v_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_2) + v_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_3) + v_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_4) \\ & - & \sigma_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_2, \mathbf{e}_2) - \sigma_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_3, \mathbf{e}_3) - \sigma_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_4, \mathbf{e}_4) \\ & - & \beta^{sc}(k, j) - \beta^{sc}(j, k) - \beta^{sc}(k, j) - \gamma\ \left(p_2(j) + p_3(k) + p_4(j)\right), \end{array} \tag{24}$$

and

$$\begin{array}{rcl} \log \mathbb{P}(B) & = & \log p_1(k, \boldsymbol{\alpha}_i) + \alpha_i(k) + \alpha_i(j) + \alpha_i(j) \\ & + & v_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_2) + v_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_3) + v_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_4) \\ & - & \sigma_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_2, \mathbf{e}_2) - \sigma_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_3, \mathbf{e}_3) - \sigma_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_4, \mathbf{e}_4) \\ & - & \beta^{sc}(k, j) - \gamma\ \left(p_2(k) + p_3(j) + p_4(j)\right), \end{array} \tag{25}$$

Therefore, the difference between log-probabilities of the two histories is:

$$\begin{array}{rcl} \log \mathbb{P}(A) - \log \mathbb{P}(B) & = & v_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_2) + v_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_3) - v_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_2) + v_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_3) \\ & - & \sigma_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_3, \mathbf{e}_3) - \sigma_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_4, \mathbf{e}_4) + \sigma_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_3, \mathbf{e}_3) + \sigma_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_4, \mathbf{e}_4) \\ & - & \widetilde{\beta}^{sc}(k, j) - \gamma\ \left(p_2(j) - p_3(j) - p_2(k) + p_3(k)\right), \end{array} \tag{26}$$

Given equation (26), we can establish the identification of parameters $\widetilde{\beta}^{sc}(k, j)$ and γ. To control for the unobserved heterogeneity in the functions $v(., \boldsymbol{\alpha}_i)$ and $\sigma(., \boldsymbol{\alpha}_i)$, we need to impose the restrictions $\mathbf{z}_2 = \mathbf{z}_3 = \mathbf{z}_4$ and $\mathbf{e}_3 = \mathbf{e}_4$. Given these restrictions, we have that:

$$\log \mathbb{P}(A) - \log \mathbb{P}(B) = -\widetilde{\beta}^{sc}(k,j) - \gamma \ \left(p_2(j) - p_3(j) - p_2(k) + p_3(k)\right), \tag{27}$$

Therefore, with $p_2(j) - p_3(j) = 0$ and $p_2(k) - p_3(k) = 0$ we have identification of the switching cost parameter $\widetilde{\beta}^{sc}(k,j)$. And with $p_2(j) - p_3(j) - p_2(k) + p_3(k) \neq 0$ we have identification of the price-sensitivity parameter γ. That is, keeping constant the permanent component of prices between periods 2 and 4, and the transitory component between periods 3 and 4, we can identify γ using the variation between periods 2 and 3 in the transitory component of the prices in products j and k.

The pair of histories A and B in equation (23) is only one of the many history pairs with identification power for the structural parameters. For instance, we can extend this example by including periods of no purchase. Let $\mathbf{0}_n$ represent a vector of n zeros. For $k, j \geq 1$ with $k \neq j$, and any two natural numbers n_1 and n_2, consider the following choice histories:

$$A \ = \ (k,\ \mathbf{0}_{n_1},\ j,\ \mathbf{0}_{n_2},\ k,\ \mathbf{0}_{n_2},\ j) \quad ; \quad B \ = \ (k,\ \mathbf{0}_{n_1},\ k,\ \mathbf{0}_{n_2},\ j,\ \mathbf{0}_{n_2},\ j) \tag{28}$$

It is straightforward to show that the difference between the log-probabilities of these two histories has the following expression:

$$\begin{aligned} \log \mathbb{P}(A) - \log \mathbb{P}(B) \ &= \ \sum_{t=n_1+2}^{n_1+n_2+2} v_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_t) - v_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_t) + \sum_{t=n_1+n_2+3}^{n_1+2n_2+3} v_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_t) - v_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_t) \\ &- \ \sum_{t=n_1+3}^{n_1+n_2+3} \sigma_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_t, \mathbf{e}_t) - \sigma_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_t, \mathbf{e}_t) + \sum_{t=n_1+n_2+4}^{n_1+2n_2+4} \sigma_{\boldsymbol{\alpha}_i}(k, \mathbf{z}_t, \mathbf{e}_t) - \sigma_{\boldsymbol{\alpha}_i}(j, \mathbf{z}_t, \mathbf{e}_t) \\ &- \ \widetilde{\beta}^{sc}(k,j) - \gamma \ \left(p_{n_1+2}(j) - p_{n_1+n_2+3}(j) - p_{n_1+2}(k) + p_{n_1+n_2+3}(k)\right). \end{aligned} \tag{29}$$

To eliminate the unobserved heterogeneity $\boldsymbol{\alpha}_i$ from this difference we need the following conditions on prices: (i) the permanent component $\mathbf{z}_t$ is constant from period $n_1 + 2$ to $n_1 + 2n_2 + 4$; and (ii) the transitory component $\mathbf{e}_t$ is constant between periods $n_1 + 3$ and $n_1 + 2n_2 + 4$. Under these conditions, we have that:

$$\log \mathbb{P}(A) - \log \mathbb{P}(B) = -\widetilde{\beta}^{sc}(k,j) - \gamma \ \left(p_{n_1+2}(j) - p_{n_1+3}(j) - p_{n_1+2}(k) + p_{n_1+3}(k)\right). \tag{30}$$

This equation shows that a change between periods n_1+2 and n_1+3 in the transitory component of the price of product j or k identifies parameter γ. The switching cost parameter $\widetilde{\beta}^{sc}(k,j)$ is identified from histories where this transitory component is constant.

3.3 Dynamic demand with duration dependence

Consider now the demand model with duration dependence. The expression for the log-CCP function is:

$$\log P_{it} = \begin{cases} \alpha_i(\ell_{it}) - \beta^{dep}(\ell_{it}, d_{it}) + v_{\boldsymbol{\alpha}_i}(\ell_{it}, d_{it}+1, \mathbf{z}_t) - \sigma_{\boldsymbol{\alpha}_i}\left(\ell_{it}, d_{it}, \mathbf{z}_t, \mathbf{e}_t\right) & if \quad y_{it} = 0 \\ \alpha_i(j) - \gamma \ p_{it}(j) - \beta^{sc}(\ell_{it}, j) + v_{\boldsymbol{\alpha}_i}(j, 1, \mathbf{z}_t) - \sigma_{\boldsymbol{\alpha}_i}\left(\ell_{it}, d_{it}, \mathbf{z}_t, \mathbf{e}_t\right) & if \quad y_{it} = j > 0 \end{cases} \tag{31}$$

First, it is straightforward to verify that the pair of histories in equation (23) still identifies the switching cost parameters $\widetilde{\beta}^{sc}_{kj}$ and the price sensitivity parameter γ. Therefore, I focus here on the identification of the depretiation parameters $\beta^{dep}(j,d)$. Furthermore, for notational simplicity, I consider here that the two price components are constant over the considered choice histories and omit $\mathbf{z}_t$ and $\mathbf{e}_t$ as arguments.

Let n be a natural number such that $2 \leq n \leq (T-2)/2$. Consider the following choice histories, both with initial duration $d_1 = 1$:

$$A \ = \ (j, \ \mathbf{0}_{n-1}, \ j, \ \mathbf{0}_{n+1}) \quad ; \quad B \ = \ (j, \ \mathbf{0}_{n}, \ j, \ \mathbf{0}_{n}, \ ) \tag{32}$$

Note that these histories do not contain any product switching event such that switching cost parameters do not appear in the probabilities of these choice histories. Taking into account the structure of the log-CCP in equation (31), we have the following expressions for log-probabilities of these choice histories:

$$\begin{aligned}\log \mathbb{P}(A) &= \log p_1(j, d_1|\boldsymbol{\alpha}_i) + 2(n+1)\alpha_i(j) - \sum_{d=1}^{n-1}\beta^{dep}(j,d) - \sum_{d=1}^{n+1}\beta^{dep}(j,d) \\ &+ \sum_{d=2}^{n} v_{\boldsymbol{\alpha}_i}(j,d) + \sum_{d=1}^{n+2} v_{\boldsymbol{\alpha}_i}(j,d) - \sum_{d=1}^{n}\sigma_{\boldsymbol{\alpha}_i}(j,d) - \sum_{d=1}^{n+1}\sigma_{\boldsymbol{\alpha}_i}(j,d).\end{aligned} \tag{33}$$

$$\begin{aligned}\log \mathbb{P}(B) &= \log p_1(j, d_1|\boldsymbol{\alpha}_i) + 2(n+1)\alpha_i(j) - \sum_{d=1}^{n}\beta^{dep}(j,d) - \sum_{d=1}^{n-1}\beta^{dep}(j,d) \\ &+ \sum_{d=2}^{n+1} v_{\boldsymbol{\alpha}_i}(j,d) + \sum_{d=1}^{n+1} v_{\boldsymbol{\alpha}_i}(j,d) - \sum_{d=1}^{n+1}\sigma_{\boldsymbol{\alpha}_i}(j,d) - \sum_{d=1}^{n}\sigma_{\boldsymbol{\alpha}_i}(j,d).\end{aligned} \tag{34}$$

They imply the following expression for the difference between the log-probabilities:

$$\log \mathbb{P}(A) - \log \mathbb{P}(B) = -\beta^{dep}(j, n+1) + \beta^{dep}(j,n) + v_{\boldsymbol{\alpha}_i}(j, n+2) - v_{\boldsymbol{\alpha}_i}(j, n+1) \tag{35}$$

The expression in the right-hand-side still depends on the incidental parameters $\boldsymbol{\alpha}_i$ such that, without further restrictions, this pair of histories does not identify β^{dep} parameters.

In general, without further restrictions, the duration dependence structural parameters $\beta^{dep}(j,d)$ are not identified in the forward-looking model.[14] To obtain identification of these parameters, I follow the same approach as Aguirregabiria, Gu, and Luo (2021) and impose the following restriction.

ASSUMPTION 2. *For any product j, there is a value of duration d_j^* – which can vary across products – such that $\beta^{dep}(j,n) = \beta^{dep}(j, d_j^*)$ for any duration $n \geq d_j^*$.* ■

Importantly, as established in Proposition 6 in Aguirregabiria, Gu, and Luo (2021), the value of d_j^* is identified from the data as long as it is not larger than $(T-1)/2$. I reproduce here this result from Aguirregabiria, Gu, and Luo (2021).

PROPOSITION. *Let d_j^* be the value of duration defined in Assumption 2 above. For any product j and any duration n with $2n+1 \leq T$, define the pair of histories $A_{j,n} = (j, \mathbf{0}_{n-1}, j, \mathbf{0}_{n+1})$*

[14]Note that in a myopic model (i.e., $\delta_i = 0$) all the continuation values $v_{\boldsymbol{\alpha}_i}(j,d)$ are zero, such that equation (35) implies the identification of $\beta^{dep}(j, n+1) - \beta^{dep}(j,n)$ from $\log \mathbb{P}(A) - \log \mathbb{P}(B)$. Here, I do not impose this restriction and consider that consumers can be forward-looking.

and $B_{j,n} = (j, \mathbf{0}_n, j, \mathbf{0}_n)$. *If* $d_j^* \leq (T-1)/2$, *then the value* d_j^* *is point identified from the following expression:*

$$d_j^* \;=\; max\{n \;:\; \log \mathbb{P}(A_{j,n}) - \log \mathbb{P}(B_{j,n}) \neq 0\} \qquad \blacksquare \tag{36}$$

An important implication of Assumption 2 is that the continuation value function is such that $v_{\boldsymbol{\alpha}_i}(j,n) = v_{\boldsymbol{\alpha}_i}(j,d_j^*)$ for any duration $n \geq d_j^*$. Combining this property with equation (35), we have that for $n = d_j^* - 1$:

$$\log \mathbb{P}(A) - \log \mathbb{P}(B) \;=\; -\,\beta^{dep}(j,d_j^*) + \beta^{dep}(j,d_j^* - 1) \tag{37}$$

such that the (local) depreciation rate $\beta^{dep}(j,d_j^*) - \beta^{dep}(j,d_j^* - 1)$ is identified. If $\beta^{dep}(j,d)$ is a linear function, i.e., $\beta^{dep}(j,d) = \overline{\beta}_j^{dep}\, d$, then equaton (37) implies the identification of the product-specific depreciation rate $\overline{\beta}_j^{dep}$.

4 Estimation

Let $\boldsymbol{\theta}$ be the vector of parameters $(\gamma, \widetilde{\beta}_{kj}^{sc}, \overline{\beta}_j^{dep}, k, j \in \{1,2,...,J\}$ with $k > j)$. The dataset is $\{y_{it}, \mathbf{z}_{it}, \mathbf{e}_{it} : i = 1,2,...,N; t = 1,2,...,T\}$. Let $\widetilde{\mathbf{z}}_i$ and $\widetilde{\mathbf{e}}_i$ the vectors with the time series of prices $\{\mathbf{z}_{it} : t = 1,2,...,T\}$ and $\{\mathbf{e}_{it} : t = 1,2,...,T\}$, respectively.

In the identification results in section 3, a sufficient statistic for $\boldsymbol{\alpha}_i$ was represented as a binary indicator that combines the condition $\mathbf{y}_i \in \{A \cup B\}$, where A and B are two choice histories, and restrictions on prices, that we can represent as $r(\widetilde{\mathbf{z}}_\mathbf{i}, \widetilde{\mathbf{e}}_\mathbf{i}) = \mathbf{0}$. That is:

$$s_i \equiv s\left(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i\right) = 1\{\mathbf{y}_i \in A \cup B \text{ and } r(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i) = \mathbf{0}\} \tag{38}$$

We have shown that:

$$\mathbb{P}(\mathbf{y}_i | \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, s_i = 1) = \frac{\exp\{c(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\}}{\exp\{c(A, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\} + \exp\{c(B, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\}} \tag{39}$$

where $c(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)$ is a vector of known statistics.

There are many pairs of histories A and B that provide sufficient statistics for $\boldsymbol{\alpha}_i$ and have identification power for $\boldsymbol{\theta}$. Let index these sufficient statistics by $m \in \{1, 2, ..., M\}$. We also index by m the different elements that define sufficient statistic s_i^m, that is: the corresponding pair of choice histories, (A^m, B^m); the restrictions on prices, $r^m(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i) = \mathbf{0}$; and the identifying statistics, $c^m(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)$. Note that we can use m also to index different sub-periods in the panel dataset.[15] Given these M sufficient statistics, we can define the following *conditional log-likelihood function*:

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{m=1}^{M}\sum_{i=1}^{N} 1\{\mathbf{y}_i \in A^m \cup B^m\}\ 1\{r^m(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i) = \mathbf{0}\}\ \log\left(\frac{\exp\{c(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\}}{\exp\{c(A, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\} + \exp\{c(B, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\}}\right) \tag{40}$$

The CML estimator is the value of $\boldsymbol{\theta}$ that maximizes $\mathcal{L}(\boldsymbol{\theta})$, which is a globally concave function.

In many possible applications, such as those using weekly or daily supermarket scanner data, price stickiness implies such that the restrictions $r^m(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i) = \mathbf{0}$ hold for a non-negligible fraction of observations. Nevertheless, imposing exactly these restrictions typically implies loosing a substantial amount of observations. This is exactly the issue that motivates the *Kernel Weighted CML* in Honoré and Kyriazidou (2000). In the log-likelihood function, we replace the indicator $1\{r^m(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i) = \mathbf{0}\}$ with a weight that depends inversely on the magnitude of vector $|r^m(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)|$ such that we put more weight in the log-likelihood of observations for which $|r^m(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)|$ is close to zero. More specifically, we consider the following Kernel Weighted conditional log-likelihood function:

$$\mathcal{L}^{KW}(\boldsymbol{\theta}) = \sum_{m=1}^{M}\sum_{i=1}^{N} 1\{\mathbf{y}_i \in A^m \cup B^m\}\ K\left(\frac{r^m(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)}{b_N}\right)\ \log\left(\frac{\exp\{c(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\}}{\exp\{c(A, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\} + \exp\{c(B, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\}}\right) \tag{41}$$

where $K(.)$ is a kernel density function that satisfies the regularity condition $K(v) \to 0$ as $||v|| \to \infty$, and b_N is a bandwidth parameter such that $b_N \to 0$ and $Nb_N \to \infty$ and $N \to \infty$. As shown by Honoré and Kyriazidou (2000), this estimator is consistent and asymptotically normal. If prices $\mathbf{z}_{it}$ and $\mathbf{e}_{it}$ have discrete support, the rate of convergence of the estimator is

[15] For instance, we can divide a panel dataset with $T = 12$ periods into three different sub-panels with four time periods each, and construct sufficient statistics separately for each sub-panel.

$N^{-1/2}$. Otherwise, with continuous prices, the rate of convergence is slower than $N^{-1/2}$.

5 Conclusions

This paper presents a Fixed Effects dynamic panel data model of demand for different products where consumers are forward looking. I apply and extend recent results from Aguirregabiria, Gu, and Luo (2021) to establish the identification of all structural parameters in this model. Several extensions of the results in this paper are interesting topics for further research.

First, an important motivation for the estimation of structural models is using them for counterfactual experiments that consist in evaluating the effects on agents' behavior of hypothetical changes in structural parameters or/and exogenous variables. Recent research by Aguirregabiria and Carro (2021), Davezies, D'Haultfoeuille, and Laage (2021), and Pakel and Weidner (2021) present new positive results on the identification of average marginal effects and counterfactuals in Fixed Effects discrete choice models. It would be interesting applying these results to this dynamic demand model.

Second, the specification of the sources of demand dynamics in the model of this paper, though relatively flexible, is restrictive. For instance, the model does not accommodate some forms of consumer learning for experienced goods used in the literature with Random Effects models (e.g., Ching, 2010).

Third, the model assumes that consumers buy at most one unit of the product per period. However, it is well-known that forward-looking consumers can buy for inventory (Hendel and Nevo, 2006; Hendel and Nevo, 2013). It would be interesting and useful to extend the model to incorporate the possibility of consumers purchasing multiple units/

Fourth, in the same spirit as the the recent work by Mysliwski, Sanches, Junior, and Srisuma, (2020), this dynamic demand can be combined with a dynamic game of price competition. The assumption on the two price components – persistent and transitory – is consistent with an equilibrium of a dynamic pricing game with i.i.d. firms' private information.

Last but not least, eventually, empirical applications using the model and identification results in this paper will be key judges of the relative merits and limitations of this framework.

References

Abrevaya, J., and Y. Hsu (2021): "Partial effects in non-linear panel data models with correlated random effects," *The Econometrics Journal*.

Ackerberg, D. A. (2003): "Advertising, learning, and consumer choice in experience good markets: an empirical examination*," *International Economic Review*, 44(3), 1007–1040.

Aguirregabiria, V., and J. Carro (2021): "Identification of average marginal effects in fixed effects dynamic discrete choice models," *CEPR Discussion Paper No. DP16354*.

Aguirregabiria, V., J. Gu, and Y. Luo (2021): "Sufficient statistics for unobserved heterogeneity in dynamic structural logit models," *Journal of Econometrics*, 223(2), 280–311.

Aguirregabiria, V., and A. Nevo (2013): "Recent developments in empirical IO: Dynamic demand and dynamic games," *Advances in Economics and Econometrics*, 3, 53–122.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica*, 63(4), 841–890.

Blackwell, D. (1965): "Discounted dynamic programming," *The Annals of Mathematical Statistics*, 36(1), 226–235.

Boizot, C., J.-M. Robin, and M. Visser (2001): "The demand for food products: an analysis of interpurchase times and purchased quantities," *The Economic Journal*, 111(470), 391–419.

Bonhomme, S. (2012): "Functional differencing," *Econometrica*, 80(4), 1337–1385.

Chamberlain, G. (1985): "Heterogeneity, omitted variable bias, and duration dependence," in *Longitudinal Analysis of Labor Market Data*. Cambridge University Press.

Chernozhukov, V., I. Fernandez-Val, J. Hahn, and W. Newey (2013): "Average and quantile effects in nonseparable panel models," *Econometrica*, 81(2), 535–580.

Ching, A. T. (2010): "Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration," *International Journal of Industrial Organization*, 28, 619–638.

Davezies, L., X. D'Haultfoeuille, and L. Laage (2021): "Identification and estimation of average marginal effects in fixed effect logit models," *arXiv preprint arXiv:2105.00879*.

De Groote, O., and F. Verboven (2019): "Subsidies and time discounting in new technology adoption: Evidence from solar photovoltaic systems," *American Economic Review*, 109(6), 2137–72.

Dobronyi, C., J. Gu, and K. Kim (2021): "Identification of Dynamic Panel Logit Models with Fixed Effects," *arXiv:2104.04590*.

Erdem, T., S. Imai, and M. Keane (2003): "Brand and quantity choice dynamics under price uncertainty," *Quantitative Marketing and Economics*, 1(1), 5–64.

Esteban, S., and M. Shum (2007): "Durable-goods oligopoly with secondary markets: the case of automobiles," *The RAND Journal of Economics*, 38(2), 332–354.

Goettler, R. L., and B. R. Gordon (2011): "Does AMD spur Intel to innovate more?," *Journal of Political Economy*, 119(6), 1141–1200.

Gowrisankaran, G., and M. Rysman (2012): "Dynamics of consumer demand for new durable goods," *Journal of Political Economy*, 120(6), 1173–1219.

Griffith, R., L. Nesheim, and M. O'Connell (2018): "Income effects and the welfare consequences of tax in differentiated product oligopoly," *Quantitative Economics*, 9(1), 305–341.

Heckman, J. (1981a): "The incidental parameters problem and the problem of initial conditionsin estimating a discrete time - discrete data stochastic process," in *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press.

——— (1981b): "Statistical models for discrete panel data," in *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press.

Hendel, I., and A. Nevo (2006): "Measuring the implications of sales and consumer inventory behavior," *Econometrica*, 74(6), 1637–1673.

——— (2013): "Intertemporal price discrimination in storable goods markets," *American Economic Review*, 103(7), 2722–51.

Honoré, B., and A. DePaula (2021): "Identification in simple binary outcome panel data models," *The Econometrics Journal*.

Honoré, B., and E. Kyriazidou (2000): "Panel data discrete choice models with lagged dependent variables," *Econometrica*, 68(4), 839–874.

Honoré, B., and M. Weidner (2020): "Dynamic Panel Logit Models with Fixed Effects," *arXiv preprint arXiv:2005.05942*.

Honoré, B. E., C. Muris, and M. Weidner (2021): "Dynamic Ordered Panel Logit Models," *arXiv preprint arXiv:2107.03253*.

Keane, M. P. (1997): "Modeling heterogeneity and state dependence in consumer choice behavior," *Journal of Business & Economic Statistics*, 15(3), 310–327.

Magnac, T. (2000): "Subsidised training and youth employment: distinguishing unobserved heterogeneity from state dependence in labour market histories," *The Economic Journal*, 110(466), 805–837.

Mysliwski, M., F. M. Sanches, D. S. Junior, and S. Srisuma (2020): "The Welfare Effects of Promotional Fees," Discussion paper, CeMMAP Working Paper, CWP35/20.

Osborne, M. (2011): "Consumer learning, switching costs, and heterogeneity: A structural examination," *Quantitative Marketing and Economics*, 9(1), 25–70.

Pakel, C., and M. Weidner (2021): "Bounds on Average Effects in Discrete Choice Panel Data Models," Discussion paper, Working paper.

Pavlidis, P., and P. B. Ellickson (2017): "Implications of parent brand inertia for multiproduct pricing," *Quantitative Marketing and Economics*, 15(4), 369–407.

Pesendorfer, M. (2002): "Retail sales: A study of pricing behavior in supermarkets," *The Journal of Business*, 75(1), 33–66.

Roy, R., P. Chintagunta, and S. Haldar (1996): "A framework for investigating habits,"The Hand of the Past," and heterogeneity in dynamic brand choice," *Marketing Science*, 15(3), 280–299.

Rust, J. (1994): "Structural estimation of Markov decision processes," *Handbook of Econometrics*, 4, 3081–3143.

Ryan, S. P., and C. Tucker (2012): "Heterogeneity and the dynamics of technology adoption," *Quantitative Marketing and Economics*, 10(1), 63–109.