Solution and Estimation of Dynamic Discrete Choice Structural Models Using Euler Equations

Victor Aguirregabiria*
University of Toronto and CEPR

Arvind Magesan*
University of Calgary

May 1st, 2018

Abstract

This paper extends the Euler Equation (EE) representation of dynamic decision problems to a general class of discrete choice models and shows that the advantages of this approach apply not only to the estimation of structural parameters but also to the solution of the model and the evaluation of counterfactual experiments. We use a choice probabilities representation of the discrete decision problem to derive marginal conditions of optimality with similar features as standard EEs in continuous decision problems. These EEs imply a fixed point mapping in the space of conditional choice values, that we denote the EE operator. We show that this operator is a stronger contraction than both the standard and the relative value function operators. Solving the dynamic programming problem by iterating in the EE operator implies substantial computational savings compared to value function and relative value function iterations (that require a much larger number of iterations) and to policy function iterations (that involves a costly valuation step at every iteration). We define a sample version of the EE operator and use it to construct a sequence of consistent estimators of the structural parameters, and to evaluate counterfactual experiments. The computational cost of evaluating this *sample-based EE* operator increases linearly with sample size, and provides a consistent estimator of the counterfactual. As such there is no curse of dimensionality in the consistent estimation of the model and in the evaluation of counterfactual experiments. We illustrate the computational gains of our methods using Monte Carlo experiments.

Keywords: Dynamic programming discrete choice models; Euler equations; Relative Value Function iteration; Estimation; Approximation bias.

JEL: C13; C35; C51; C61

Victor Aguirregabiria. Department of Economics. University of Toronto. 150 St. George Street. Toronto, Ontario victor.aguirregabiria@utoronto.ca

Arvind Magesan. Department of Economics. University of Calgary. 2500 University Drive, N.W. Calgary, Alberta arvindmagesan@gmail.com

*We would like to thank comments from Rob Bray, Thierry Magnac, Angelo Melino, Bob Miller, Pedro Mira, Jean Marc Robin, John Rust, Bertel Schjerning, Kunio Tsuyuhara, and from seminar participants at Calgary, North Carolina, Stony-Brook, Toronto, Western University, the Barcelona GSE Summer Forum (Structural Microeconometrics), the Canadian Economic Association conference, the Microeconometric Network Meeting in Copenhagen, the Society of Economic Dynamics conference and The Banff Empirical Microeconomics Conference 2015.

1 Introduction

The development of the Euler equation-GMM approach by Hansen and Singleton (1982) was a primary methodological contribution to the literature on estimation of dynamic structural models. One of the main advantages of this method is that it avoids the curse of dimensionality associated with the computation of present values. The computational cost of estimating structural parameters from Euler equations increases with sample size but not with the dimension of the state space.[1] However, the Euler equation-GMM approach also has some well-known limitations. First, the conventional wisdom in the literature is that this method cannot be applied to models of discrete choice because optimal decisions cannot be characterized in terms of marginal conditions in these models.[2] Second, while the Euler equation-GMM significantly reduces the computational burden associated with estimating structural parameters by avoiding full solution of the dynamic model, the end goal of structural work is typically to use an estimated model to study the effect of policies that have never occurred. The methods available for the estimation of the effect of such counterfactual experiments require the full solution of the dynamic programming problem, or at least an approximation. Even if the researcher can avoid full solution in estimating the structural parameters, solving the model is required when using the estimated model to study counterfactual policies and, in principle, Euler Equations do not help with this step. Though it is possible to use Euler equations to construct a fixed point operator (see Coleman, 1990, 1991), in general this operator is not a contraction such that convergence of this method is not guaranteed.

Given that the Hansen-Singleton method was believed to be inapplicable to the estimation of discrete choice models, the development of *Conditional Choice Probability* (CCP) methods for the estimation of these models, pioneered by Hotz and Miller (1993) and Hotz et al. (1994), represented a substantial methodological contribution in this literature. By avoiding the solution of the dynamic programming (DP) problem, these methods facilitate the estimation of specifications with larger state spaces and richer sources of individual specific heterogeneity. Nevertheless, in contrast to the

[1] Another attractive feature of the Euler equation-GMM approach when applied to panel data is that it can deal with different forms of non-stationarity of exogenous state variables without having to specify the stochastic process that governs the future out-of-sample evolution of these variables, e.g., evolution of future aggregate shocks, business cycle, regulatory changes, etc.

[2] For instance, this conventional wisdom is clearly ennunciated in Rust (1988, page 1008): *"In an important contribution, Hansen and Singleton (1982) have developed a practical technique for estimating structural parameters of a class of discrete-time, continuous control processes. Their method uses the generalized method of moments technique to estimate first-order necessary conditions of the agent's stochastic control problem (stochastic Euler equations), avoiding the need for an explicit solution for the optimal decision rule. The Hansen-Singleton method depends critically on the assumption that the agent's control variable is continuous in order to derive the first-order necessary conditions by the usual variational methods."*

Euler equations approach, the implementation of CCP or Hotz-Miller methods without a terminal state still requires the computation of present values defined as integrals or summations over the space of future state variables. In applications with continuous state variables or with very large state spaces, the exact solution of these present values is an intractable problem.[3] In this context, the work by Arcidiacono and Miller (2011, 2015) on models with a *finite dependence* property represents an important contribution.[4] They show that models that possess a *finite dependence* property permit a representation whereby the choice probabilities can be expressed as a function of expected payoffs at a finite number of states, meaning that a researcher interested in the estimation of structural parameters does not need to compute present discounted values of the stream of payoffs over infinite future periods. Nevertheless, the solution of the model and the estimation of counterfactual experiments is still subject to the curse of dimensionality.

This paper builds on results previously obtained in Aguirregabiria and Magesan (2013). In that previous paper, we derived two main results on dynamic discrete choice structural models: (i) we propose a representation of the discrete choice model as a continuous decision problem where the decision variables are choice probabilities, and we use this equivalent representation to derive marginal conditions of optimality similar in nature to the Euler equations in standard continuous decision problems; and (ii) we show that these marginal conditions imply moment conditions that can be used to estimate the structural parameters of the model by GMM, in the same spirit as Hansen and Singleton (1982) for continuous decision models. Building on these results, this paper presents two main contributions.

First, we show that these EEs imply a fixed point mapping in the space of conditional choice values, that we denote as the *Euler Equation (EE) operator*. We show that, in contrast to Euler equation operators in continuous decision models, this operator is a contraction, such that successive iterations in this mapping can be used to obtain the unique solution of the DP problem. We compare

[3] Applied researchers have used different approximation techniques such as discretization, Monte Carlo simulation, sieves, neural networks, etc. However, replacing true expected values with approximations introduces an approximation error, and this error typically induces a statistical bias in the estimation of the parameters of interest. In general, this bias does not go to zero as sample size increases and the level of approximation (e.g., number of Monte Carlo simulations) is constant.

[4] Consider a dynamic decision model with discrete time t, vector of state variables x_t and discrete decision variable a_t. Let j and k be two possible values of the decision variable at period t. Following Arcidiacono and Miller (2011), the model has the $\tau-$*periods finite dependence property* at (t, x_t, j, k) if there are two paths of feasible choices from period $t+1$ until period $t+\tau$, one path starting at $(t, x_t, a_t = j)$ and the other at $(t, x_t, a_t = k)$, such that the probability distribution of $x_{t+\tau+1}$ conditional on x_t is the same under the two paths of choices. Arcidiacono and Miller (2016) derive necessary and sufficient conditions for finite state dependence as rank conditions for matrices that depend on transition probabilities. We show that our necessary and sufficient condition for the existence of Euler equations, in Proposition 4 at Section 3, is closely related to Arcidiacono and Miller (2016) rank condition for one-period finite state dependence.

the computational properties of the EE iterations algorithm with those of the most commonly used solution methods: Value Function (VF) iterations, Relative Value Function (RVF) iterations, and Policy Function (PF) iterations. We show that the EE operator is a stronger contraction (i.e., it has a smaller Lipschitz constant) than VF and RVF, and has similar contraction properties as the PF operator. In terms of the time required to evaluate the operator a single time, the EE operator is also more efficient than VF, RVF, and PF algorithms, and this difference increases quadratically with the dimension of the state space. We present numerical examples that illustrate how solving the model by iterating in this EE fixed point mapping implies very substantial computational savings relative to these three alternative methods.

Second, we define a sample version of the EE operator and use it to construct a sequence of consistent estimators of the structural parameters, and to evaluate counterfactual experiments.[5] This sample-based EE operator is defined only at sample points of the exogenous state variables, and thus its dimensionality does not increase with the dimension of the state space. We show that this sample-based EE operator is also a contraction and the unique fixed point of this mapping is a root-N consistent estimator of the true solution. The sample operator can be used to define an estimator of the structural parameters. In contrast to most estimation methods of dynamic structural models, the computational cost to obtain this estimator does not depend on the dimension of the state space.

We illustrate the computational gains of our methods using several numerical experiments in the context of a dynamic model of market entry and exit. In a first experiment, we compare the computation time of EE, VF, RVF, and PF algorithms for the solution of the model. For moderately sized state spaces (up to 200,000 states), the EE algorithm is over 200 times faster than PF, between 50-70 times faster than VF, and between 5-40 times faster than RVF.[6] These differences increase with the dimension of the state space. This implies that many models that are computationally infeasible for all practical purposes using standard methods, are feasible using the method we propose. We also use this first set of experiments to study the source of the difference in total computation time across these algorithms. In particular, we show that the advantages of the EE algorithm relative to VF and RVF are both in terms of time-per-iteration and the number of iterations to convergence.

In a second experiment, we study the finite sample properties and the computational cost

[5] Our sample-based EE operator is related to the random grid method of Rust (1997), though Rust defines and applies this method to standard value function and policy function operators, and not to Euler equations.

[6] The EE algorithm is also faster that hybrid algorithms that combine PF and VF or RVF iterations.

of estimators of the structural parameters using standard methods in this literature and using the method based on the EE mapping. More specifically, we compare four estimation methods: two-step Hotz-Miller (Hotz and Miller, 1993); Maximum Likelihood Estimator (MLE) computed using the Nested Pseudo Likelihood (NPL) algorithm; two-step estimator using the EE mapping (2SEE); and NPL estimator using the EE mapping (NPL-EE). We find that the two-step Euler equation estimator has about 33% higher root mean squared error than the MLE. However, the NPL Euler equations estimator is statistically superior to the two-step Hotz-Miller estimator and statistically indistinguishable from the MLE. Very importantly, the main difference between the two approaches is their computational cost. The NPL Euler equations estimator is over 2000 times faster than the MLE-NPL.[7] Ultimately then, there is no trade-off (at least in the context of our dynamic entry/exit model), as researchers can obtain estimates close to the MLE with a small fraction of the computation time.

In our third and final set of experiments, we compare standard value function and policy function methods and Euler equations methods for the estimation of an equilibrium of the model associated to counterfactual policy. We study how these methods perform in predicting firm behavior in response to a counterfactual increase in the cost of entry, holding the computation time of the different methods fixed. We show that the finite sample properties of the Euler equation estimator are substantially better than those of the standard methods, i.e., mean absolute bias and squared error are between 35 and 60 percentage points smaller in the Euler equation method.

The rest of the paper is organized as follows. Section 2 presents the model and some fundamental results in the literature. Section 3 describes our derivation of Euler equations in discrete choice models, defines the *EE mapping*, and shows its contraction property. Section 4 defines the sample version of the EE mapping and uses this operator to define a family of estimators of structural parameters and a method to consistently estimate counterfactuals. We derive the statistical and computational properties of these methods, and compare them with those from previous methods in the literature. In section 5, we present results from Monte Carlo experiments where we illustrate the advantages of our proposed methods. We summarize and conclude in section 6. Proofs of Propositions are in the Appendix.

[7] The NPL method swaps the order of the two nested algorithms in the NFXP method. In NPL, the inner algorithm maximizes a pseudo likelihood with respect to the structural parameters, keeping fixed present values; and the outer algorithm updates present values. As shown in Aguirregabiria and Mira (2002) and Bray (2016), among others, computing the MLE using the NFXP algorithm takes several times longer than using NPL.

2 Model

2.1 Basic framework

This section presents a general class of dynamic programing (DP) models in discrete time with discrete actions and state variables. This framework follows Rust (1987, 1994) and it is standard in the structural microeconometrics literature. We describe some properties of the model that will be useful in the derivation of our main results.

Every period t, an agent takes a decision a_t to maximize his expected intertemporal payoff $\mathbb{E}_t[\sum_{j=0}^{T-t} \beta^j \ \Pi_t(a_{t+j}, \mathbf{s}_{t+j})]$, where $\beta \in (0,1)$ is the discount factor, T is the time horizon, which may be finite or infinite, $\Pi_t(.)$ is the one-period payoff function at period t, and $\mathbf{s}_t \in \mathcal{S}$ is the vector of state variables at period t, which we assume follows a controlled Markov process with transition probability function $f_t(\mathbf{s}_{t+1}|a_t, \mathbf{s}_t)$. The decision variable a_t belongs to the discrete and finite set $\mathcal{A} = \{0, 1, ..., J\}$. The sequence of value functions $\{V_t(.) : t \geq 1\}$ can be obtained recursively using the Bellman equation:

$$V_t(\mathbf{s}_t) = \max_{a_t \in \mathcal{A}} \left\{ \Pi_t(a_t, \mathbf{s}_t) + \beta \int V_{t+1}(\mathbf{s}_{t+1}) \ f_t(\mathbf{s}_{t+1}|a_t, \mathbf{s}_t) \ d\mathbf{s}_{t+1} \right\} \tag{1}$$

The optimal decision rule, $\alpha_t(.) : \mathcal{S} \to \mathcal{A}$, is obtained as the arg-max of the expression in brackets. This framework allows for both stationary and non-stationary models. In the stationary case, the time horizon T is infinite, and the payoff and transition probability functions are time-homogenous, which implies that the value function and the optimal decision rule are also invariant over time.

Following the standard model in this literature (Rust, 1994), we distinguish between two sets of state variables: $\mathbf{s}_t = (\mathbf{x}_t, \varepsilon_t)$, where $\mathbf{x}_t$ is the vector of state variables observable to the researcher, and ε_t represents the unobservables for the researcher. The vector of unobservables satisfies the standard assumptions of additive separability, conditional independence, and discrete support of the observable state variables. Specifically, the one-period payoff function is: $\Pi_t(a_t, \mathbf{s}_t) = \pi_t(a_t, \mathbf{x}_t) + \varepsilon_t(a_t)$, where $\varepsilon_t \equiv \{\varepsilon_t(a) : a \in \mathcal{A}\}$ is a vector of unobservable random variables, and the transition probability (density) function of the state variables factors as: $f_t(\mathbf{s}_{t+1}|a_t, \ \mathbf{s}_t) = f_t(\mathbf{x}_{t+1}|a_t, \mathbf{x}_t) \ dG(\varepsilon_{t+1})$, where $G(.)$ is the CDF of ε_t which is absolutely continuous with respect to Lebesgue measure, strictly increasing and continuously differentiable in all its arguments, and with finite means. The vector of state variables $\mathbf{x}_t$ belongs to a discrete set $\mathcal{X}$.

In this dynamic discrete choice problem, the value of choosing alternative a can be represented as $v_t(a, \mathbf{x}_t) + \varepsilon_t(a)$, where $v_t(a, \mathbf{x}_t)$ is the *conditional choice value function,*

$$v_t(a, \mathbf{x}_t) \equiv \pi_t(a, \mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1} \in \mathcal{X}} \int V_{t+1}(\mathbf{x}_{t+1}, \varepsilon_{t+1}) \ dG(\varepsilon_{t+1}) \ \ f_t(\mathbf{x}_{t+1}|a_t, \mathbf{x}_t) \tag{2}$$

Taking choice alternative $a_t = 0$ as a benchmark (without loss of generality), we can define the value differences $\widetilde{v}_t(a, \mathbf{x}_t) \equiv v_t(a, \mathbf{x}_t) - v_t(0, \mathbf{x}_t)$, and the optimal decision rule $\alpha_t(\mathbf{x}_t, \varepsilon_t)$ can be described as follows:

$$\{\alpha_t(\mathbf{x}_t, \varepsilon_t) = a\} \text{ if and only if } \{\widetilde{v}_t(a, \mathbf{x}_t) + \varepsilon_t(a) \geq \widetilde{v}_t(j, \mathbf{x}_t) + \varepsilon_t(j) \text{ for any } j\} \tag{3}$$

Let $\widetilde{\mathbf{v}}_t(\mathbf{x}_t)$ be the vector of J value differences at period t, i.e., $\widetilde{\mathbf{v}}_t(\mathbf{x}_t) = \{\widetilde{v}_t(a, \mathbf{x}_t) : a \in \mathcal{A} - \{0\}\}$. The *optimal choice probability (OCP) mapping*, $\mathbf{\Lambda}\left(\widetilde{\mathbf{v}}_t(\mathbf{x})\right) \equiv \{\Lambda\left(a, \widetilde{\mathbf{v}}_t(\mathbf{x})\right) : a \in \mathcal{A} - \{0\}\}$, is a mapping from the space of value differences in $\mathbb{R}^J$ into the space of choice probabilities in $[0,1]^J$ such that $\Lambda\left(a, \widetilde{\mathbf{v}}_t(\mathbf{x})\right)$ is the probability that given the vector of values $\widetilde{\mathbf{v}}_t(\mathbf{x})$ the optimal choice is alternative a. Given the form of the optimal decision rule in equation (3), the OCP function is:

$$\Lambda\left(a, \widetilde{\mathbf{v}}_t(\mathbf{x})\right) \equiv \int 1\left\{\widetilde{v}_t(a, \mathbf{x}_t) + \varepsilon_t(a) \geq \widetilde{v}_t(j, \mathbf{x}_t) + \varepsilon_t(j) \ \ \forall j\right\} \, dG\left(\varepsilon_t\right) \tag{4}$$

where $1\{.\}$ is the indicator function. Given a vector of choice probabilities $\mathbf{P}_t \equiv \{P_t(a) : a \in \mathcal{A} - \{0\}\}$, we say that this vector is optimal at period t given state $\mathbf{x}$ if and only if $\mathbf{P}_t = \mathbf{\Lambda}\left(\widetilde{\mathbf{v}}_t(\mathbf{x})\right)$ and $\widetilde{\mathbf{v}}_t(\mathbf{x})$ is the vector value differences at period t. Proposition 1 establishes that the OCP mapping is invertible.

PROPOSITION 1 [Hotz-Miller Inversion]. For any vector $\widetilde{\mathbf{v}}_t(\mathbf{x}) \in \mathbb{R}^J$ the OCP mapping $\mathbf{\Lambda}(.)$ is invertible such that there is a one-to-one relationship between the vector of value differences $\widetilde{\mathbf{v}}_t(\mathbf{x})$ and the vector of optimal choice probabilities $\mathbf{P}_t(\mathbf{x})$, i.e., $\widetilde{\mathbf{v}}_t(\mathbf{x}) = \mathbf{\Lambda}^{-1}(\mathbf{P}_t(\mathbf{x}))$. ■

Proof: Proposition 1 in Hotz and Miller (1993).

2.2 Dynamic decision problem in probability space

This dynamic discrete choice problem can be described in terms of the primitives or structural functions $\{\pi_t, f_t, G, \beta\}_{t=1}^T$. We now define a dynamic programming problem with the same primitives but where the agent does not choose a discrete action $a_t \in \{0, 1, ..., J\}$ but a probability distribution over the space of possible actions, i.e., a vector of choice probabilities $\mathbf{P}_t \equiv \{P_t(a) : a \in \mathcal{A} - \{0\}\}$. We denote this problem as the *probability-choice problem*. We show below that there is a one-to-one relationship between optimal decision rules (and value functions) in the original discrete choice problem and in the probability-choice problem.

Given an arbitrary vector of choice probabilities $\mathbf{P}_t$ in $[0,1]^J$, we define the following expected payoff function,

$$\Pi_t^P(\mathbf{P}_t, \mathbf{x}_t) \equiv \sum\nolimits_{a=0}^{J} P_t(a) \, [\pi_t\left(a, \mathbf{x}_t\right) + e_t(a, \mathbf{P}_t)], \tag{5}$$

where $e_t(a, \mathbf{P}_t)$ is the expected value of $\varepsilon_t(a)$ conditional on alternative a being the optimal choice given that the vector of value differences is $\widetilde{\mathbf{v}}_t = \mathbf{\Lambda}^{-1}(\mathbf{P}_t)$, i.e.,[8]

$$e_t(a, \mathbf{P}_t) = \mathbb{E}\left[\varepsilon_t(a) \mid \Lambda^{-1}(a, \mathbf{P}_t) + \varepsilon_t(a) \geq \Lambda^{-1}(j, \mathbf{P}_t) + \varepsilon_t(j) \ \ \forall j\right] \tag{6}$$

We also define the expected transition probability of the state variables,

$$f_t^P(\mathbf{x}_{t+1}|\mathbf{P}_t, \mathbf{x}_t) \equiv \sum\nolimits_{a=0}^{J} P_t(a) \ f_t(\mathbf{x}_{t+1}|a, \mathbf{x}_t). \tag{7}$$

Consider a dynamic programming problem where the decision at period t is the vector of choice probabilities $\mathbf{P}_t$, the current payoff function is $\Pi_t^P(\mathbf{P}_t, \mathbf{x}_t)$, and the transition probability of the state variables is $f_t^P(\mathbf{x}_{t+1}|\mathbf{P}_t, \mathbf{x}_t)$. This is the *probability-choice* problem associated to the original dynamic discrete choice model. By definition, the Bellman equation of this problem is:

$$V_t^P(\mathbf{x}_t) = \max_{\mathbf{P}_t \in [0,1]^J} \left\{ \Pi_t^P(\mathbf{P}_t, \mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1} \in \mathcal{X}} V_{t+1}^P(\mathbf{x}_{t+1}) \ f_t^P(\mathbf{x}_{t+1}|\mathbf{P}_t, \mathbf{x}_t) \right\} \tag{8}$$

The solution of this DP problem can be described in terms of a sequence of value functions $\{V_t^P(\mathbf{x}_t)\}_{t=1}^T$ and optimal decision rules $\{\mathbf{P}_t^*(\mathbf{x}_t)\}_{t=1}^T$. Proposition 2 presents several properties of this solution of the probability-choice problem. These properties play an important role in our derivation of Euler equations. These properties build on previous results but, as far as we know, they are new in the literature.

Let $W_t(\mathbf{P}_t, \mathbf{x}_t)$ be the intertemporal payoff function of the probability-choice problem, i.e., $W_t(\mathbf{P}_t, \mathbf{x}_t) \equiv \Pi_t^P(\mathbf{P}_t, \mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1}} V_{t+1}^P(\mathbf{x}_{t+1}) \ f^P(\mathbf{x}_{t+1}|\mathbf{P}_t, \mathbf{x}_t)$.

PROPOSITION 2. For any dynamic discrete choice problem that satisfies the assumptions of Additive Separability, and Conditional Independence, the associated probability-choice problem is such that:

(A) the intertemporal payoff function $W_t(\mathbf{P}_t, \mathbf{x}_t)$ is twice continuously differentiable and globally concave in $\mathbf{P}_t$ such that the optimal decision rule $\mathbf{P}_t^*(\mathbf{x}_t)$ is uniquely characterized by the first order condition $\partial W_t(\mathbf{P}_t^*, \mathbf{x}_t)/\partial \mathbf{P}_t = 0$;

(B) for any vector $\mathbf{P}_t$ in the J-dimension Simplex, the gradient vector $\partial W_t(\mathbf{P}_t, \mathbf{x}_t)/\partial \mathbf{P}_t$ is equal to $\widetilde{\mathbf{v}}_t(\mathbf{x}_t) - \mathbf{\Lambda}^{-1}(\mathbf{P}_t)$;

(C) the optimal decision rule in the probability-choice problem is equal to the optimal choice probability (OCP) function of the original discrete choice problem, i.e., $\mathbf{P}_t^*(\mathbf{x}_t) = \mathbf{\Lambda}(\widetilde{\mathbf{v}}_t(\mathbf{x}_t))$. ■

[8] For some distributions of the unobservables ε_t, this function has a simple closed form expression. For instance, if the unobservables are extreme value type 1, then $e_t(a, \mathbf{P}_t) = \gamma - \ln(P_t(a))$. And if the model is binary choice and the unobservables are independent standard normal, then $e_t(a, \mathbf{P}_t) = \phi(\Phi^{-1}[P_t(a)])/P_t(a)$.

Proof: In the Appendix.

Proposition 2 establishes two main results. First, it establishes a representation property of this class of discrete choice models. The dynamic discrete-choice model has a representation as a probability-choice problem with the particular definitions of expected payoff and expected transition probability functions presented in equations (5)-(7) above. Second, it shows that the optimal solution of the probability-choice problem can be described in terms of marginal conditions of optimality of a globally concave optimization problem. In section 3, we show that we can exploit this representation to derive Euler equations (EEs) for discrete choice DP models. The following Proposition 3 plays also a role in the particular structure of Euler equations in this model.

PROPOSITION 3. For any vector of choice probabilities $\mathbf{P}_t$ *in the J-dimensional simplex,*

$$\frac{\partial \Pi_t^P(\mathbf{P}_t, \mathbf{x}_t)}{\partial P_t(a)} = \widetilde{\pi}_t(a, \mathbf{x}_t) - \Lambda^{-1}(a, \mathbf{P}_t) \tag{9}$$

where $\widetilde{\pi}_t(a, \mathbf{x}_t) \equiv \pi_t(a, \mathbf{x}_t) - \pi_t(0, \mathbf{x}_t)$. ■

Proof: In the Appendix.

3 The Euler Equation fixed point mapping

3.1 Deriving Euler Equations

This subsection borrows from results in Aguirregabiria and Magesan (2013). In DP models where both decision and state variables are continuous, the standard approach to obtain EEs is based on the combination of marginal conditions of optimality at two consecutive periods together with an envelope condition for the value function. This standard approach, though very convenient for its simplicity, imposes strong restrictions on the model: the endogenous state variables should be continuous and follow transition rules where the stochastic component (innovation) is additively separable and independent of the decision, e.g., $\mathbf{x}_{t+1} = g_t(\mathbf{x}_t, \mathbf{P}_t) + \xi_{t+1}$. In general, the probability-choice model defined by Bellman equation (8) does not satisfy these conditions. However, these restrictions are far from being necessary for the existence of Euler equations. Here we apply a more general method to obtain these equations as proposed in Aguirregabiria and Magesan (2013).

The method is based on a *two-period deviation principle* that should be satisfied by the optimal solution of any DP problem. Consider a constrained optimization problem where the agent chooses the vector of probabilities at periods t and $t+1$, $\mathbf{P}_t$ and $\mathbf{P}_{t+1}$, to maximize the sum of expected and discounted payoffs at these two consecutive periods subject to the constraint that the distribution of the state variables at period $t+2$ stays the same as under the optimal solution of the DP problem.

This constrained optimization problem is formally given by:

$$\begin{aligned} \max_{\{\mathbf{P}_t, \mathbf{P}_{t+1}\}} \quad & \Pi_t^P(\mathbf{P}_t, \mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1} \in \mathcal{X}} \Pi_{t+1}^P(\mathbf{P}_{t+1}, \mathbf{x}_{t+1}) \; f_t^P(\mathbf{x}_{t+1} | \mathbf{P}_t, \mathbf{x}_t) \\ \text{subject to:} \quad & f_{(2)}^P(\mathbf{x}_{t+2} \mid \mathbf{P}_t, \mathbf{P}_{t+1}, \mathbf{x}_t) = f_{(2)}^P(\mathbf{x}_{t+2} \mid \mathbf{P}_t^*, \mathbf{P}_{t+1}^*, \mathbf{x}_t) \text{ for any } \mathbf{x}_{t+2} \end{aligned} \tag{10}$$

where $f_{(2)}^P$ represents the two-period-forward transition probability of the state variables, that by definition is a convolution of the one-period transitions at periods t and $t+1$:

$$f_{(2)}^P(\mathbf{x}_{t+2} \mid \mathbf{P}_t, \mathbf{P}_{t+1}, \mathbf{x}_t) \equiv \sum_{\mathbf{x}_{t+1} \in \mathcal{X}} f_{t+1}^P(\mathbf{x}_{t+2} \mid \mathbf{P}_{t+1}, \mathbf{x}_{t+1}) \; f_t^P(\mathbf{x}_{t+1} \mid \mathbf{P}_t, \mathbf{x}_t) \tag{11}$$

The two-period deviation principle establishes that the unique solution to this problem is given by the choice probability functions $\mathbf{P}_t^*(\mathbf{x}_t)$ and $\mathbf{P}_{t+1}^*(\mathbf{x}_{t+1})$ that solve the DP problem (8) at periods t and $t+1$.[9] Note that for each value of $\mathbf{x}_t$ there is a different constrained optimization problem, and therefore a different solution. We can solve this problem using Lagrange method. Under some additional conditions, we can operate in the Lagrange conditions of optimality to obtain equations that do not include Lagrange multipliers but only payoffs at periods t and $t+1$.

We have not distinguished so far between endogenous and exogenous state variables in the model. This distinction is important for the derivation of Euler equations. Let $\mathbf{x}_t$ be equal to $(\mathbf{y}_t, \mathbf{z}_t)$, where $\mathbf{y}_t$ is the vector of endogenous state variables with transition probability $f_t^y(\mathbf{y}_{t+1}|a_t, \mathbf{y}_t)$, and $\mathbf{z}_t$ is the vector of exogenous state variables with transition probability $f_t^z(\mathbf{z}_{t+1}|\mathbf{z}_t)$. By definition, the evolution of the exogenous state variables is not affected by the choice probabilities such that the set of constraints in problem (10) apply only to the distribution of the endogenous state variables at period $t+2$: i.e., $f_{(2)}^{P,y}(\mathbf{y}_{t+2} \mid \mathbf{P}_t, \mathbf{P}_{t+1}, \mathbf{x}_t) = f_{(2)}^{P,y}(\mathbf{y}_{t+2} \mid \mathbf{P}_t^*, \mathbf{P}_{t+1}^*, \mathbf{x}_t)$. Let $\mathcal{Y}$ be the (unconditional) support set of the vector of endogenous state variables $\mathbf{y}_t$. For a given vector $\mathbf{y}_t$, let $\mathcal{Y}_{(1)}(\mathbf{y}_t) \subseteq \mathcal{Y}$ be the support of $\mathbf{y}_{t+1}$ conditional on $\mathbf{y}_t$. That is, $\mathbf{y}_{t+1} \in \mathcal{Y}_{(1)}(\mathbf{y}_t)$ if and only if $f_t(\mathbf{y}_{t+1}|a_t, \mathbf{y}_t) > 0$ for some value of a_t. Similarly, for given $\mathbf{y}_t$, let $\mathcal{Y}_{(2)}(\mathbf{y}_t) \subseteq \mathcal{Y}$ be the set of all the vectors $\mathbf{y}_{t+2}$ with $\Pr(\mathbf{y}_{t+2}|a_t, a_{t+1}, \mathbf{y}_t) > 0$ for some value of a_t and a_{t+1}. Let $\widetilde{f}_t^y(\mathbf{y}_{t+1}|a_t, \mathbf{y}_t)$ be the *difference* transition probability $f_t^y(\mathbf{y}_{t+1}|a_t, \mathbf{y}_t) - f_t^y(\mathbf{y}_{t+1}|0, \mathbf{y}_t)$.[10] For notational simplicity we omit the state at period t, $\mathbf{x}_t$, as an argument in the expressions below, though we should keep in mind that there is a system of Euler equations for each value of $\mathbf{x}_t$.

There are two sets of Lagrange marginal conditions of optimality. The first set of Lagrange

[9] If the DP problem is stationary, then this solution will be such that the vectors $\mathbf{P}_t^*$ and $\mathbf{P}_{t+1}^*$ are the same.

[10] Without loss of generality, we use choice alternative 0 as the baseline.

conditions is a system of J equations, one for each probability $P_t(a)$ with $a \in \mathcal{A} - \{0\}$,

$$\frac{\partial \Pi_t^P}{\partial P_t(a)} + \sum_{\mathbf{y}_{t+1}} \left[\beta \ \Pi_{t+1}^P - \sum_{\mathbf{y}_{t+2}} \lambda_{t+2}(\mathbf{y}_{t+2}) \ f_{t+1}^{P,y}(\mathbf{y}_{t+2}|\mathbf{P}_{t+1}, \mathbf{x}_{t+1}) \right] \widetilde{f}_t^y(\mathbf{y}_{t+1}|a) = 0; \tag{12}$$

where $\lambda_{t+2}(\mathbf{y}_{t+2})$ is the Lagrange multiplier associated to the constraint in problem (10) for state $\mathbf{y}_{t+2}$.[11] The second set of Lagrange conditions is a system of $J * |\mathcal{Y}_{(1)}|$ equations, one for each probability $P_{t+1}(a|\mathbf{x}_{t+1})$ with $a \in \mathcal{A} - \{0\}$ and $\mathbf{y}_{t+1} \in \mathcal{Y}_{(1)}$,

$$\beta \frac{\partial \Pi_{t+1}^P}{\partial P_{t+1}(a|\mathbf{x}_{t+1})} - \sum_{\mathbf{y}_{t+2}} \lambda_{t+2}(\mathbf{y}_{t+2}) \ \widetilde{f}_{t+1}^y(\mathbf{y}_{t+2}|a, \mathbf{y}_{t+1}) = 0; \tag{13}$$

where we have taken into account that $\partial f_{t+1}^{P,y} / \partial P_{t+1}(a|\mathbf{x}_{t+1}) = \widetilde{f}_{t+1}^y(\mathbf{y}_{t+2}|a, \mathbf{y}_{t+1})$.

Our derivation of Euler equations consists of using the system of $J * |\mathcal{Y}_{(1)}|$ equations in (13) to solve for the vector of $|\mathcal{Y}_{(2)}| - 1$ Lagrange multipliers, and then plug-in this solution into the system of equations (12). A *necessary and sufficient* condition for the existence of Euler equations is that the system of equations (13) has a unique solution for the vector of Lagrange multipliers. Let $\widetilde{\mathbf{F}}_{t+1}$ be the matrix with elements $\widetilde{f}_{t+1}^y(\mathbf{y}_{t+2}|a_{t+1}, \mathbf{y}_{t+1})$ where the columns correspond to all the values $\mathbf{y}_{t+2} \in \mathcal{Y}_{(2)}(\mathbf{y}_t)$ leaving out one value, and the rows correspond to all the values $(a_{t+1}, \mathbf{y}_{t+1}) \in [\mathcal{A} - \{0\}] \times \mathcal{Y}_{(1)}(\mathbf{y}_t)$. Using this matrix, we can represent the system of equations (13) in vector form as:

$$\beta \frac{\partial \mathbf{\Pi}_{t+1}^P}{\partial \mathbf{P}_{t+1}} \quad = \quad \widetilde{\mathbf{F}}_{t+1} \ \lambda_{t+2} \tag{14}$$

where λ_{t+2} is the vector of Lagrange multipliers, and $\partial \mathbf{\Pi}_{t+1}^P / \partial \mathbf{P}_{t+1}$ is a vector with dimension $J * |\mathcal{Y}_{(1)}|$ that contains the marginal expected profits $\partial \Pi_{t+1}^P / \partial P_{t+1}(a|\mathbf{x}_{t+1})$. Proposition 4 is based on Proposition 3 in Aguirregabiria and Magesan (2013) and it establishes the conditions for existence of Euler equations and presents the general formula.[12]

PROPOSITION 4. Suppose that matrix $\widetilde{\mathbf{F}}_{t+1}$ is full column rank. Then, the marginal conditions of optimality for the constrained optimization problem (10) imply the following solution for the Lagrange multipliers, $\lambda_{t+2} = \lambda_{t+2}^* \left(\widetilde{\mathbf{F}}_{t+1}, \mathbf{\Pi}_{t+1}^P \right) \equiv \beta \ [\widetilde{\mathbf{F}}_{t+1}' \widetilde{\mathbf{F}}_{t+1}]^{-1} \ \widetilde{\mathbf{F}}_{t+1}' \frac{\partial \mathbf{\Pi}_{t+1}^P}{\partial \mathbf{P}_{t+1}}$, and the following system of Euler equations,

$$\frac{\partial \Pi_t^P}{\partial P_t(a)} + \beta \sum_{\mathbf{y}_{t+1}} \left[\pi_{t+1}(0, \mathbf{x}_{t+1}) + e(0, \mathbf{P}_{t+1}) - \sum_{\mathbf{y}_{t+2}} \lambda_{t+2}^*(\mathbf{y}_{t+2}) \ f_{t+1}^y(\mathbf{y}_{t+2}|0, \mathbf{y}_{t+1}) \right] \widetilde{f}_t^y(\mathbf{y}_{t+1}|a) \quad = \quad 0 \tag{15}$$

[11] For the derivation of this condition, note that $\partial f_t^{P,y} / \partial P_t(a) = f_t^y(\mathbf{y}_{t+1}|a) - f_t^y(\mathbf{y}_{t+1}|0) = \widetilde{f}_t^y(\mathbf{y}_{t+1}|a)$.

[12] The expression for the Euler equation in Proposition 4 is slightly different to the one in Aguirregabiria and Magesan (2013). The reason is that in our Proposition 4 we have applied several results that simplify the expression. We explain this in the Appendix.

where $\lambda^*_{t+2}(\mathbf{y})$ *is the element of solution vector* $\lambda^*_{t+2}\left(\widetilde{\mathbf{F}}_{t+1}, \mathbf{\Pi}^P_{t+1}\right)$ *associated to* $\mathbf{y}_{t+2} = \mathbf{y}$. ■

Proof: In the Appendix.

Remark 1. Note that the dimension of matrix $[\widetilde{\mathbf{F}}_{t+1}'\widetilde{\mathbf{F}}_{t+1}]$ is the number of values (minus one) that the endogenous state variables $\mathbf{y}_{t+2}$ can take two periods forward. In most applications, this is a small number. In particular, the dimension of this matrix does not depend on the dimension of the state space of the exogenous state variables. This is a feature of the Euler equation that is key for the substantial computational savings of the methods proposed in this paper.

Remark 2. The full column rank of matrix $\widetilde{\mathbf{F}}_{t+1}$ is a necessary and sufficient condition for the existence of this Euler equation. Arcidiacono and Miller (2016) present a similar condition for their one-period finite state dependence property. Therefore, our Euler equation is closely related to the representation they obtain with one-period finite state dependence.

Remark 3. Applying Proposition 3 to our Euler equation in (15), we can obtain a convenient expression. Remember that Proposition 3 establishes that the partial derivative $\partial \Pi^P_t / \partial P_t(a)$ is equal to $\widetilde{\pi}_t(a, \mathbf{x}_t) - \Lambda^{-1}(a, \mathbf{P}_t)$, where $\Lambda^{-1}(.)$ is the inverse of the Optimal Choice Probability mapping, $\mathbf{P}_t = \Lambda(\widetilde{\mathbf{v}}_t)$. Therefore, the solution for the vector of Lagrange multipliers, $\lambda_{t+2} = \beta\ [\widetilde{\mathbf{F}}_{t+1}'\widetilde{\mathbf{F}}_{t+1}]^{-1}\ \widetilde{\mathbf{F}}_{t+1}'\ [\partial \mathbf{\Pi}^P_{t+1} / \partial \mathbf{P}_{t+1}]$, is equal to a sum of elements $\widetilde{\pi}_{t+1}(a, \mathbf{x}_{t+1}) - \Lambda^{-1}(a, \mathbf{P}_{t+1})$ multiplied by transition probabilities.

EXAMPLE 1. Multi-armed bandit models. Dynamic discrete choice models of occupational choice, portfolio choice, or market entry-exit, among many other economic models, can be seen as examples of a general class of dynamic decision models called Multi-Armed Bandit problems (Gittins, 1979; Gittins, Glazebrook, and Weber, 2011). Every period t the agent chooses an occupation (or an asset; or a market to enter) among $J+1$ possible choices. There are costs of changing occupations such that the choice of occupation in the previous period is a state variable. The previous period's occupation is the only endogenous state variable of the model such that the endogenous state space is $\mathcal{Y} = \{0, 1, ..., J\}$, and the state transition is given by $x_{t+1} = a_t$ and $f^y_t(y_{t+1}|a_t, y_t) = 1\{y_{t+1} = a_t\}$. This implies that $f^{P,y}_t(y_{t+1} = y|\mathbf{P}_t, \mathbf{x}_t) = P_t(y|\mathbf{x}_t)$, and the two-periods forward transition is $f^P_{(2)}(y_{t+2} = y' \mid \mathbf{P}_t, \mathbf{P}_{t+1}, \mathbf{x}_t) = \sum_{a=0}^J P_t(a|\mathbf{x}_t)\ P_{t+1}(y'|a, \mathbf{z}_{t+1})$. In this model, the Lagrange condition with respect to $P_{t+1}(a'|a, \mathbf{z}_{t+1})$ is,

$$\beta \frac{\partial \Pi^P_{t+1}(\mathbf{P}_{t+1}, a)}{\partial P_{t+1}(a'|a, \mathbf{z}_{t+1})} = \lambda_{t+2}(a') \tag{16}$$

that provides a unique solution for the Lagrange multipliers. Substituting this solution into the first order condition with respect to $P_t(a|\mathbf{x}_t)$, and using Proposition 3, we can obtain the following

expression for the Euler equations in this class of models:

$$\begin{aligned} &\pi_t\left(a, \mathbf{x}_t\right) + e(a, \mathbf{P}_t) + \beta\left[\pi_{t+1}\left(0, a, \mathbf{z}_{t+1}\right) + e(0, \mathbf{P}_{t+1}(a, \mathbf{z}_{t+1}))\right] \quad = \\ &\pi_t\left(0, \mathbf{x}_t\right) + e(0, \mathbf{P}_t) + \beta\left[\pi_{t+1}\left(0, 0, \mathbf{z}_{t+1}\right) + e(0, \mathbf{P}_{t+1}(0, \mathbf{z}_{t+1}))\right] \end{aligned} \tag{17}$$

When the unobservables are i.i.d. extreme value distributed, we have that $e_t(a, \mathbf{P}_t) = \gamma - \ln P_t(a|\mathbf{x}_t)$. Euler equation (17) has a clear economic interpretation. There are multiple decision paths (a_t, a_{t+1}) than can reach a state at period $t+2$ (y_{t+2}) given a state at period t (y_t). For instance, for any value $a \in \mathcal{A}$ the decision path $(a_t = a, a_{t+1} = 0)$ reaches state $y_{t+2} = 0$. Euler equation (17) establishes that, at the optimal choice probabilities of the DP problem, any of these decision paths imply the same discounted expected payoff at periods t and $t+1$. This can be interpreted as an arbitrage condition. If this condition does not hold, then the agent can increase his expected intertemporal payoff by increasing the probability of the choice path with higher two-period expected payoff. The Euler equation in (17) is associated with the value $y_{t+2} = 0$. We have a similar expression for the Euler Equation associated with other state values $x_{t+2} = a > 0$. ■

EXAMPLE 2: Machine replacement model. Consider the bus engine replacement problem in Rust (1987). In this example, the endogenous state variable y_t is the mileage on the bus engine. The space of possible mileages is given by the discrete set $\mathcal{Y} = \{0, 1, 2, ...\}$, and mileage follows a transition rule $y_{t+1} = (1 - a_t)(y_t + 1)$, where $a_t \in \{0, 1\}$ represents the machine replacement decision. The set of possible states one period forward is $\mathcal{Y}_{(1)}(y_t) = \{0, y_t + 1\}$, and the set of possible states two periods forward is $\mathcal{Y}_{(2)}(y_t) = \{0, 1, y_t + 2\}$. The transition probability induced by the choice probability is $f_t^{P,y}(y_{t+1} = 0|\mathbf{P}_t, \mathbf{x}_t) = P_t(1|\mathbf{x}_t)$ and $f_t^{P,y}(y_{t+1} = y_t + 1|\mathbf{P}_t, \mathbf{x}_t) = 1 - P_t(1|\mathbf{x}_t)$. The two-periods forward transitions are $f_{(2)}^P(y_{t+2} = 1 \mid \mathbf{P}_t, \mathbf{P}_{t+1}, \mathbf{x}_t) = P_t(1|\mathbf{x}_t)\ P_{t+1}(0|0, \mathbf{z}_{t+1})$ and $f_{(2)}^P(y_{t+2} = y_t + 2 \mid \mathbf{P}_t, \mathbf{P}_{t+1}, \mathbf{x}_t) = P_t(0|\mathbf{x}_t)\ P_{t+1}(0|y_t + 1, \mathbf{z}_{t+1})$. The Lagrange conditions with respect to $P_{t+1}(1|0, \mathbf{z}_{t+1})$ and $P_{t+1}(1|y_t + 1, \mathbf{z}_{t+1})$ are,

$$\beta \frac{\partial \Pi_{t+1}^P(\mathbf{P}_{t+1}, 0)}{\partial P_{t+1}(1|0, \mathbf{z}_{t+1})} = \lambda_{t+2}(1) \quad \text{and} \quad \beta \frac{\partial \Pi_{t+1}^P(\mathbf{P}_{t+1}, x+1)}{\partial P_{t+1}(1|y_t + 1, \mathbf{z}_{t+1})} = \lambda_{t+2}(y_t + 2) \tag{18}$$

Solving these expressions into the marginal condition with respect to $P_t(1|\mathbf{x}_t)$, and using Proposition 3, we obtain that:

$$\begin{aligned} &\pi_t(1, \mathbf{x}_t) + e(1, \mathbf{P}_t) + \beta\left[\pi_{t+1}(1, 0, \mathbf{z}_{t+1}) + e(1, \mathbf{P}_{t+1}(0, \mathbf{z}_{t+1}))\right] \quad = \\ &\pi_t(0, \mathbf{x}_t) + e(0, \mathbf{P}_t) + \beta\left[\pi_{t+1}(1, y_t + 1, \mathbf{z}_{t+1}) + e(1, \mathbf{P}_{t+1}(y_t + 1, \mathbf{z}_{t+1}))\right] \end{aligned} \tag{19}$$

Again, there is a clear economic interpretation of this Euler equation. There are two possible decision paths from state y_t at period t to state $y_{t+2} = 0$ at period $t + 2$: decision path $(a_t =$

$1, a_{t+1} = 1)$, and decision path $(a_t = 0, a_{t+1} = 1)$. Euler equation (19) establishes that, at the optimal solution, the discounted expected payoff at periods t and $t+1$ should be the same under the two decision paths. ■

3.2 Euler Equation fixed point mapping

For the rest of the paper, we focus on stationary DP models: payoff and transition probability functions are constant over time, and time horizon T is infinite. Furthermore, matrix $\widetilde{\mathbf{F}}$ is full column rank, which is a necessary and sufficient condition for the existence of Euler equations, as established in Proposition 4. In this section, we show that the system of Euler equations implies a fixed point mapping in the space of value differences $\widetilde{\mathbf{v}}$ that we denote as the *Euler Equation (EE) mapping*. We show that the EE mapping is a contraction and its contraction (Lipschitz) constant is strictly smaller than the discount factor β. Therefore, this EE mapping is a stronger contraction than value function iterations and relative value function iterations.

The following Proposition 5 shows that we can use the system of Euler equations in (15) to define a fixed point mapping in the space of differential conditional choice value functions $\widetilde{v}_t(a, \mathbf{x}_t)$.

PROPOSITION 5 [Euler Equation mapping]. For the stationary probability-choice problem, the system of Euler equations in Proposition 4 implies the following fixed point mapping in the space of value differences $\widetilde{\mathbf{v}}$. *Given a vector* $\widetilde{\mathbf{v}} \in \mathcal{V} \subseteq \mathbb{R}^{J|\mathcal{X}|}$, *the Euler Equation (EE) mapping is defined as* $\mathbf{\Gamma}_{EE}(\widetilde{\mathbf{v}}) \equiv \{\Gamma_{EE}(a, \mathbf{x}_t, \widetilde{\mathbf{v}}) : (a, \mathbf{x}_t) \in (\mathcal{A} - \{0\}) \times \mathcal{X}\}$, *where*

$$\Gamma_{EE}(a, \mathbf{x}_t, \widetilde{\mathbf{v}}) \equiv \widetilde{\pi}\left(a, \mathbf{x}_t\right) + \beta \sum_{\mathbf{x}_{t+1}} \left[\pi_{t+1}\left(0, \mathbf{x}_{t+1}\right) + e(0, \Lambda\left(\widetilde{\mathbf{v}}\right)) - \overline{\lambda^*}(\mathbf{y}_{t+1}, \Lambda\left(\widetilde{\mathbf{v}}\right)) \right] \widetilde{f}(\mathbf{x}_{t+1}|a, \mathbf{x}_t), \quad (20)$$

$\overline{\lambda^*}(\mathbf{y}_{t+1}, \Lambda\left(\widetilde{\mathbf{v}}\right)) \equiv \sum_{\mathbf{y}_{t+2}} \lambda^*(\mathbf{y}_{t+2}, \Lambda\left(\widetilde{\mathbf{v}}\right)) \, f^y(\mathbf{y}_{t+2}|0, \mathbf{y}_{t+1})$, *and* $\widetilde{f}(\mathbf{x}_{t+1}|a, \mathbf{x}_t) \equiv \widetilde{f}^y(\mathbf{y}_{t+1}|a, \mathbf{y}_t) f^z(\mathbf{z}_{t+1}|\mathbf{z}_t)$.

■

Proof: It follows from the Euler equation in Proposition 4, the application of Proposition 3 that implies $\partial \Pi_t^P / \partial P_t(a) = \widetilde{\pi}_t\left(a, \mathbf{x}_t\right) - \Lambda^{-1}\left(a, \mathbf{P}_t\right)$, and the invertibility of the Optimal Choice Probability mapping Λ that implies that $\Lambda^{-1}\left(a, , \mathbf{x}_t, \mathbf{P}\right) = \widetilde{v}(a, \mathbf{x}_t)$.

EXAMPLE 3: Consider the Multi-armed bandit problem in Example 1 with extreme value distribution of the unobservables and the system of Euler equations in (**??**). For this model, the OCP mapping $\Lambda\left(\widetilde{\mathbf{v}}\right)$ has the form of multinomial logit probabilities, and the inverse of this mapping is $\widetilde{v}(a, \mathbf{x}_t) = \ln P(a|\mathbf{x}_t) - \ln P(0|\mathbf{x}_t)$. The form of the $e(a, \mathbf{P})$ function is $\gamma - \ln P(a)$, and therefore $e(a, \Lambda\left(\widetilde{\mathbf{v}}\right))$ is equal to $\gamma - \widetilde{v}(a) + \ln(1 + \sum_{j=1}^{J} \exp\{\widetilde{v}(j)\})$. Taking into account these expressions,

we have that EE mapping in this model is:

$$\begin{aligned}\Gamma_{EE}(a,\mathbf{x}_t,\widetilde{\mathbf{v}}) &= \widetilde{\pi}(a,\mathbf{x}_t)+\beta\,\mathbb{E}_{\mathbf{z}_{t+1}|\mathbf{z}_t}\left[\pi\left(0,a,\mathbf{z}_{t+1}\right)-\pi\left(0,0,\mathbf{z}_{t+1}\right)\right] \\ &+ \beta\,\mathbb{E}_{\mathbf{z}_{t+1}|\mathbf{z}_t}\left[\ln\left(1+\textstyle\sum_{j=1}^{J}\exp\{\widetilde{v}(j,0,\mathbf{z}_{t+1})\}\right)-\ln\left(1+\textstyle\sum_{j=1}^{J}\exp\{\widetilde{v}(j,a,\mathbf{z}_{t+1})\}\right)\right]\end{aligned} \tag{21}$$

Proposition 6 establishes that the Euler Equation mapping is a contraction. A property of the EE mapping that plays a key role in the proof of contraction (and in the property that its Lipschitz constant is smaller than β) is that $\mathbf{\Gamma}_{EE}(a,y,\mathbf{z};\widetilde{\mathbf{v}})-\mathbf{\Gamma}_{EE}(a,0,\mathbf{z};\widetilde{\mathbf{v}})=\Delta(a,y,\mathbf{z})$, where $\Delta(a,y,\mathbf{z})$ is a primitive of the model that is known to the researcher before solving the model and does not depend on the vector of values $\widetilde{\mathbf{v}}$ where we are evaluating the mapping. This property implies that all the vectors in the sequence $\{\widetilde{\mathbf{v}}^k : k\geq 1\}$, obtained by applying fixed-point iterations from some initial $\widetilde{\mathbf{v}}^0$ (i.e., $\widetilde{\mathbf{v}}^{k+1}=\Gamma(\widetilde{\mathbf{v}}^k)$), satisfy the property $\widetilde{v}(a,y,\mathbf{z})-\widetilde{v}(a,0,\mathbf{z})=\Delta(a,y,\mathbf{z})$. Therefore, without loss of generality, we can restrict our analysis of the fixed point mapping Γ to the subspace of values $\widetilde{\mathbf{v}}$ that satisfy this restriction. That is, we consider $\Gamma(\widetilde{\mathbf{v}})$ on $\mathcal{V}^R$ where,

$$\mathcal{V}^R \equiv \left\{\widetilde{\mathbf{v}}\in\mathbb{R}^{J(J+1)|\mathcal{Z}|} : \widetilde{v}(a,y,\mathbf{z})=\widetilde{v}(a,0,\mathbf{z})+\Delta(a,y,\mathbf{z}) \text{ for any } (a,y,\mathbf{z})\right\} \tag{22}$$

PROPOSITION 6. The Euler Equation - value mapping Γ_{EE} is a contraction in the complete metric space $(\mathcal{V}^R,\|.\|_\infty)$, where $\|.\|_\infty$ is the infinity norm, i.e., there is a constant $\delta\in(0,1)$ such that for any pair $\widetilde{\mathbf{v}}$ and $\widetilde{\mathbf{w}}$ in $\mathcal{V}^R$, we have that $\|\mathbf{\Gamma}_{EE}(\widetilde{\mathbf{v}})-\mathbf{\Gamma}_{EE}(\widetilde{\mathbf{w}})\|_\infty\leq\delta\|\widetilde{\mathbf{v}}-\widetilde{\mathbf{w}}\|_\infty$. Furthermore, $\delta<\beta$ such that the Euler equation - value mapping is a stronger contraction than the value function iterations mapping. ■

Proof: In the Appendix.

A corollary of Proposition 6 is that successive iterations in the EE operator is a method to solve this discrete choice dynamic programming problem. Below we compare this method to the most commonly used methods for solving DP problems: value iterations (i.e., iterations in the Bellman equation), relative value iterations, and policy iterations (or Newton-Kantorovich iterations).

We also define the Euler equation - probability mapping as $\mathbf{\Gamma}_{EE-p}(\mathbf{P})\equiv\{\Gamma_{EE-p}(a,\mathbf{x}_t,\mathbf{P}) : (a,\mathbf{x}_t)\in(\mathcal{A}-\{0\})\times\mathcal{X}\}$, where $\Gamma_{EE-p}(a,\mathbf{x}_t,\mathbf{P})=\Lambda(a,\mathbf{x}_t,\Gamma_{EE}(\Lambda^{-1}(\mathbf{P})))$. We show in the Appendix that the Euler equations - probability mapping is not always a contraction.

3.3 Alternative solution methods

In this section, we describe the standard algorithms and our new algorithm for the solution of dynamic programming Markov discrete choice models, and explain their computational complexity.

For the sake of concreteness, we illustrate these methods for the multi-armed bandit dynamic logit model in Example 1. This is a canonical model in the literature and has been used in a large number of empirical applications. We describe these algorithms using vector-form notation. Let $|\mathcal{Z}|$ be the dimension of the space of exogenous state variables. All of the following vectors have dimension $|\mathcal{Z}| \times 1$: $\mathbf{V}(y)$ is the vector of values when the endogenous state is y; and $\pi(a,y)$, $\widetilde{\mathbf{v}}(a,y)$, $\mathbf{P}(a,y)$ are the vectors of one-period payoffs, differential values, and conditional choice probabilities, respectively, when current action is a and the endogenous state is y. Let $\mathbf{F}_z$ be the $|\mathcal{Z}| \times |\mathcal{Z}|$ matrix of transition probabilities of the exogenous state variables.

Value function (VF) iterations. The value function operator is a fixed point mapping in the space of the vector of values $\mathbf{V} = \{\mathbf{V}(y) : y \in \mathcal{A}\}$. It is defined as $\mathbf{\Gamma}_{VF}(\mathbf{V}) = \{\mathbf{\Gamma}_{VF}(y,\mathbf{V}) : y \in \mathcal{A}\}$ with

$$\mathbf{\Gamma}_{VF}(y,\mathbf{V}) = \ln\left(\sum_{a=0}^{J} \exp\left\{\pi(a,y) + \beta\ \mathbf{F}_z\ \mathbf{V}(a)\right\}\right) \tag{23}$$

The algorithm starts with an initial vector of values $\mathbf{V}_0$. At every iteration $n \geq 1$ it updates the vector using $\mathbf{V}_n = \mathbf{\Gamma}_{VF}(\mathbf{V}_{n-1})$. The computational complexity per iteration derives from the number of multiplications involved in matrix products $\mathbf{F}_z\ \mathbf{V}_{n-1}(a)$ for the $J+1$ choice alternatives, and is thus of the order $(J+1)\ |\mathcal{Z}|^2$. The degree of contraction of this mapping, as measured by the Lipschitz constant, is equal to the discount factor β (Puterman, 1994, and Rust, 1996).

Relative Value function (RVF) iterations. Let $\mathbf{x}_0 \in \mathcal{X}$ be an arbitrary value of the vector of state variables. The *relative value function operator* is a fixed point mapping in the space of the vector of values $\mathbf{V} = \{\mathbf{V}(y) : y \in \mathcal{A}\}$, that is defined as $\mathbf{\Gamma}_{RVF}(\mathbf{V}) = \{\mathbf{\Gamma}_{RVF}(y,\mathbf{V}) : y \in \mathcal{A}\}$ with

$$\mathbf{\Gamma}_{RVF}(y,\mathbf{V}) = \ln\left(\sum_{a=0}^{J} \exp\left\{\pi(a,y) + \beta\ \mathbf{F}_z\ \left[\mathbf{V}(a) - V(\mathbf{x}_0)\ \mathbf{1}_{|\mathcal{Z}|}\right]\right\}\right) \tag{24}$$

and $\mathbf{1}_{|\mathcal{Z}|}$ is a column vector of ones. Given an initial $\mathbf{V}_0$, this vector is updated at every iteration $n \geq 1$ using $\mathbf{V}_n = \mathbf{\Gamma}_{RVF}(\mathbf{V}_{n-1})$. The computational complexity of one iteration in this algorithm is the same as one value function iteration. The difference between the two algorithms is in their contraction properties. The Lipschitz constant of the RVF operator is $\beta\ \rho(\mathbf{F}_z)$, where $\rho(\mathbf{F}_z) \in (0,1)$ is the spectral radius of the transition matrix $\mathbf{F}_z$ (Puterman, 1994, section 6.6). When the stochastic process of the exogenous state variables has strong time persistence, the spectral radius $\rho(\mathbf{F}_z)$ is close to one and the degree of contraction of RVF is very similar to VF. However, when the exogenous state variables are not so persistent, the value of $\rho(\mathbf{F}_z)$ can be substantially smaller than one and the RVF algorithm can converge to a solution much faster than VF iterations.[13]

[13] In a recent paper, Bray (2016) exploits this property of the RVF algorithm to propose improved versions of the Nested Fixed Point (Rust, 1987), and the Nested Pseudo Likelihood (Aguirregabiria and Mira, 2002) estimation methods.

Policy function (PF) iterations (or Newton-Kantorovich method, Puterman and Brumelle, 1979, and Puterman, 1994). The policy function operator is a fixed point mapping in the space of the vector of conditional choice probabilities $\mathbf{P} \in \{\mathbf{P}(a,y) : (a,y) \in [\mathcal{A} - \{0\}] \times \mathcal{A}\}$. It is defined as $\mathbf{\Gamma}_{PF}(\mathbf{P}) = \{\mathbf{\Gamma}_{PF}(a,y,\mathbf{P}) : (a,y) \in [\mathcal{A} - \{0\}] \times \mathcal{A}\}$ with

$$\mathbf{\Gamma}_{PF}(a,y,\mathbf{P}) = \exp\left\{\pi(a,y) + \beta\ \mathbf{F}_z\ \mathbf{W}^{\mathbf{P}}(a)\right\} ./ \left(\sum_{j=0}^{J} \exp\left\{\pi(j,y) + \beta\ \mathbf{F}_z\ \mathbf{W}^{\mathbf{P}}(j)\right\}\right), \tag{25}$$

where ./ represents the element-by-element division; and $\mathbf{W}^{\mathbf{P}}(a)$ is a $|\mathcal{Z}| \times 1$ vector that contains the present discounted values of future payoffs conditional on every possible value of $\mathbf{z}_t$ and on $y_t = a$, and conditional on the behavior of the agent at future periods follows the vector of choice probabilities $\mathbf{P}$. The vectors of present values $\{\mathbf{W}^{\mathbf{P}}(a) : a \in \mathcal{A}\}$ are obtained solving the system of linear equations: for any $y \in \mathcal{A}$,

$$\mathbf{W}^{\mathbf{P}}(y) = \sum_{a=0}^{J} \mathbf{P}(a,y) * \left[\pi(a,y) + \beta\ \mathbf{F}_z\ \mathbf{W}^{\mathbf{P}}(a)\right]; \tag{26}$$

where $*$ is the element-by-element product. The linear operator described in the system (26), that delivers the vectors $\mathbf{W}^{\mathbf{P}}(a)$ for a given vector of CCPs $\mathbf{P}$, is denoted the *valuation operator*. The probability operator that returns a new vector of CCPs given the vectors of valuations $\mathbf{W}^{\mathbf{P}}$ is denoted the *policy improvement operator*. Therefore, the policy function operator in (25) is the composition of the valuation and the policy improvement operators. The PF algorithm starts with an initial $\mathbf{P}_0$, and at every iteration $n \geq 1$ updates the vector using $\mathbf{P}_n = \mathbf{\Gamma}_{PF}(\mathbf{P}_{n-1})$. The computational complexity of one iteration in the PF operator is given by the complexity of solving the system of linear equations (26). This complexity is of the order $O(|\mathcal{X}|^3)$. Therefore, one PF iteration is more costly than one VF or RVF iteration, and the difference increases with the dimension of the state space. However, the PF operator is a stronger contraction that VF and RVF operators, and therefore it requires a smaller number of iterations to achieve convergence.

Euler Equation (EE) iterations. The EE operator is a fixed point mapping in the space of the vector of differential values $\widetilde{\mathbf{v}} = \{\widetilde{\mathbf{v}}(a,y) : (a,y) \in [\mathcal{A} - \{0\}] \times \mathcal{A}\}$. It is defined as $\mathbf{\Gamma}_{EE}(\widetilde{\mathbf{v}}) = \{\mathbf{\Gamma}_{EE}(a,y,\widetilde{\mathbf{v}}) : (a,y) \in [\mathcal{A} - \{0\}] \times \mathcal{A}\}$ with

$$\begin{aligned} \mathbf{\Gamma}_{EE}(a,y,\widetilde{\mathbf{v}}) &= [\pi(a,y) - \pi(a,0)] + \beta\ \mathbf{F}_z\ [\pi(0,a) - \pi(0,0)] \\ &+ \beta\ \mathbf{F}_z \left[\ln\left(1 + \sum_{j=1}^{J} \exp\{\widetilde{\mathbf{v}}(j,0)\}\right) - \ln\left(1 + \sum_{j=1}^{J} \exp\{\widetilde{\mathbf{v}}(j,a)\}\right)\right] \end{aligned} \tag{27}$$

The algorithm starts with an initial vector $\widetilde{\mathbf{v}}_0$ and at every iteration $n \geq 1$ updates the vector using $\widetilde{\mathbf{v}}_n = \mathbf{\Gamma}_{EE}(\widetilde{\mathbf{v}}_{n-1})$. The computational complexity of each iteration is of the order $J\ |\mathcal{Z}|^2$, and it comes from the calculation of the matrix products $\mathbf{F}_z\ \mathbf{w}_{n-1}(a)$ where $\mathbf{w}_{n-1}(a)$ is the term

$\ln(1 + \sum_{j=1}^{J} \exp\{\widetilde{\mathbf{v}}_{n-1}(j,0)\}) - \ln(1 + \sum_{j=1}^{J} \exp\{\widetilde{\mathbf{v}}_{n-1}(j,a)\})$. This algorithm has two computational advantages with respect to VF and RVF methods. First and most important, as shown in Proposition 6, the EE mapping is a stronger contraction than compared to VF and RVF such that it converges to the solution using a smaller number of iterations. Second, its cost per iteration is smaller: it is of the order $J\ |\mathcal{Z}|^2$ for EE, and $(J+1)\ |\mathcal{Z}|^2$ for VF and RVF. This is because the EE operator needs to calculate expectations of next period values not for every current choice alternative (as VF and RVF) but for all alternatives except one. This difference increases with the dimension of the state space. We illustrate these advantages in section 5.

4 Estimation

Suppose that the researcher's dataset consists of panel data of N agents, indexed by i, over T periods of time with information on agents' actions and state variables, $\{a_{it}\ , \mathbf{x}_{it} : i = 1, 2, ..., N; t = 1, 2, ..., T\}$. Here we consider a sample where the number of agents N is large and the number of time periods T is small, i.e., asymptotic results are for $N \to \infty$ with T fixed. The researcher is interested in using this sample to estimate the structural parameters in the payoff function. We assume that the payoff function is known to the researcher up to a finite vector of structural parameters θ. The researcher is also interested in using the estimated model to make predictions about how a change in some structural parameters affects agents' behavior. This prediction exercise is described in the literature as a *counterfactual experiment*. In this section, we present estimation methods for structural parameters and counterfactual experiments that use the Euler Equation operator.

4.1 Empirical Euler-Equation mapping

Given a sample, the researcher can construct an empirical counterpart of the EE mappings defined in Section 3. We use this empirical EE mapping for the estimation of structural parameters and counterfactual experiments. Here we define the empirical EE mappings and prove some important properties.

Let $\{\mathbf{z}_{it} : i = 1, 2, ..., N;\ t = 1, 2, ..., T\}$ be the sample observations of the vector of exogenous state variables. Define the empirical set $\mathcal{Z}_N = \{\mathbf{z} \in \mathcal{Z} : \text{there is an observation } (i,t) \text{ with } \mathbf{z}_{it} = \mathbf{z}\}$, and the empirical transition probability function $f_{(N)}(\mathbf{z}'|\mathbf{z})$ defined on $\mathcal{Z}_N \times \mathcal{Z}_N$ into $[0,1]$, such that for any $\mathbf{z} \in \mathcal{Z}_N$, $f_{(N)}(\mathbf{z}'|\mathbf{z}) = \sum_{i=1}^{N} 1\{\mathbf{z}_{it+1} = \mathbf{z}' \text{ and } \mathbf{z}_{it} = \mathbf{z}\} / \sum_{i=1}^{N} 1\{\mathbf{z}_{it} = \mathbf{z}\}$. Stationarity of the transition probability $f_z(\mathbf{z}_{t+1}|\mathbf{z}_t)$ implies that: the set $\mathcal{Z}_N$ is a random sample from the ergodic

set $\mathcal{Z}$; $\mathcal{Z}_N$ converges to $\mathcal{Z}$ and $f_{(N)}(\mathbf{z}'|\mathbf{z})$ converges uniformly to $f_z(\mathbf{z}'|\mathbf{z})$ as N goes to infinity. Let $\mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}\}}[.]$ be a sample conditional mean operator from $\mathbb{R}$ into $\mathbb{R}$ such that for any real-valued function $h(\mathbf{z}')$ the operator is defined as:

$$\mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}\}}\left[h(\mathbf{z}')\right] \equiv \sum_{\mathbf{z}' \in \mathcal{Z}_N} f_{(N)}(\mathbf{z}'|\mathbf{z})\ h(\mathbf{z}') \tag{28}$$

The *Empirical EE mapping* $\Gamma^{(N)}_{EE,v}(\widetilde{\mathbf{v}})$ is defined as the sample counterpart of the EE mapping in equation (20) where we replace the conditional expectation at the population level with its empirical counterpart $\mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}\}}$. For instance, for the dynamic logit model, we have that

$$\begin{aligned}\Gamma^{(N)}_{EE}(a,y,\mathbf{z};\widetilde{\mathbf{v}}) &= [\pi(a,y,\mathbf{z}) - \pi(0,y,\mathbf{z})] + \beta\ \mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}\}}\left[\pi\left(0,a,\mathbf{z}'\right) - \pi\left(0,0,\mathbf{z}'\right)\ \right] \\ &+ \beta\ \mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}\}}\left[\ln\left(1+\sum_{j=1}^{J}\exp\{\widetilde{v}_{t+1}(j,0,\mathbf{z}')\}\right) - \ln\left(1+\sum_{j=1}^{J}\exp\{\widetilde{v}_{t+1}(j,a,\mathbf{z}')\}\right)\right]\end{aligned} \tag{29}$$

and $\Gamma^{(N)}_{EE}(\widetilde{\mathbf{v}})$ is $\{\Gamma^{(N)}_{EE}(a,y,\mathbf{z};\widetilde{\mathbf{v}}) : (a,y,\mathbf{z}) \in \mathcal{A}\times\mathcal{Y}\times\mathcal{Z}_N\}$. This is a fixed point mapping in the space of value differences such that we can obtain a sample-based solution to the DP problem by solving the fixed point problem $\widetilde{\mathbf{v}} = \mathbf{\Gamma}^{(N)}_{EE}(\widetilde{\mathbf{v}})$. Importantly, the dimension of this fixed point mapping is $J * |\mathcal{Y}| * \mathcal{Z}_N$, which can be many orders of magnitude smaller than the dimension of Γ_{EE} when the dimension of $\mathcal{Z}$ is large relative to sample size.

Proposition 7 establishes that the Empirical EE mapping is a contraction and it converges uniformly in probability to the true (population) EE mapping. We now include explicitly the vector of structural parameters θ as an argument in this mapping.

PROPOSITION 7. The Empirical EE mapping $\Gamma^{(N)}_{EE}(\widetilde{\mathbf{v}},\theta)$ is a contraction mapping and it converges uniformly in probability to the population EE mapping $\Gamma_{EE}(\widetilde{\mathbf{v}},\theta)$.

Proof: In the Appendix.

We also define the *Empirical EE-probability mapping* as $\Gamma^{(N)}_{EE-p}(\mathbf{P})$ by combining the EE mapping and the Optimal Choice Probability mapping, $\Lambda(\widetilde{\mathbf{v}})$, and its inverse, $\Lambda^{-1}(\mathbf{P})$. Let $\widetilde{\mathbf{v}}$ be the fixed point of the empirical EE mapping such that $\widetilde{\mathbf{v}} = \Gamma^{(N)}_{EE}(\widetilde{\mathbf{v}},\theta)$, and let $\mathbf{P} = \Lambda(\widetilde{\mathbf{v}})$. Then, applying the inverse mapping we have that $\mathbf{P}$ should solve the following fixed point mapping: $\mathbf{P} = \Lambda(\Gamma^{(N)}_{EE}(\Lambda^{-1}(\mathbf{P}),\theta))$. The EE-probability mapping is defined as $\Gamma^{(N)}_{EE-p}(\mathbf{P}) \equiv \Lambda(\Gamma^{(N)}_{EE}(\Lambda^{-1}(\mathbf{P}),\theta))$.

For the dynamic logit model,

$$\Gamma^{(N)}_{EE-p}(a,\mathbf{x};\mathbf{P}) = \frac{\exp\left\{\widetilde{\pi}\left(a,\mathbf{x}\right)+\beta\ \mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}\}}\left[\pi\left(0,a,\mathbf{z}'\right)-\pi\left(0,0,\mathbf{z}'\right)-\ln\left(\frac{P(0|a,\mathbf{z}')}{P(0|0,\mathbf{z}')}\right)\right]\right\}}{\sum_{j=0}^{J}\exp\left\{\widetilde{\pi}\left(j,\mathbf{x}\right)+\beta\ \mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}\}}\left[\pi\left(0,j,\mathbf{z}'\right)-\pi\left(0,0,\mathbf{z}'\right)-\ln\left(\frac{P(0|j,\mathbf{z}')}{P(0|0,\mathbf{z}')}\right)\right]\right\}} \tag{30}$$

Using the same argument as in Proposition 7, it is possible to show that $\Gamma^{(N)}_{EE-p}(\mathbf{P},\theta)$ converges uniformly in probability to the population EE-prob mapping $\Gamma_{EE,p}(\mathbf{P},\theta)$. However, the population EE-prob operator is not necessarily a contraction, and this is also the case for its sample counterpart.

4.2 Estimation of structural parameters using Euler equations

Given the empirical EE operator and the , define the Pseudo Likelihood function:

$$Q_N\left(\theta,\mathbf{P}\right)=\sum_{i=1}^{N}\sum_{t=1}^{T}\ln\Gamma^{(N)}_{EE-p}(a_{it},\mathbf{x}_{it};\ \theta,\mathbf{P}) \tag{31}$$

We can construct a root-N consistent and asymptotically normal estimator of θ using a two-step Pseudo Maximum Likelihood (PML) estimator. The first step consists of the nonparametric estimation of the conditional choice probabilities: $\widehat{\mathbf{P}}_N \equiv \{\widehat{P}_t(a|\mathbf{x}) : t=1,2,...,T\}$. For instance, $\widehat{P}_t(a|\mathbf{x})$ can be a kernel (Nadaraya-Watson) estimator of the regression of $1\{a_i=a\}$ on $\mathbf{x}_{it}$. Importantly, and in contrast to other so called CCP estimators, we do not need to estimate conditional choice probabilities at states which are not observed in the sample. In the second step, the PML estimator of θ is:[14]

$$\widehat{\theta}_N = \arg\max_{\theta\in\Theta}\ Q_N\left(\theta,\widehat{\mathbf{P}}_N\right) \tag{32}$$

This two-step semiparametric estimator is root-N consistent and asymptotically normal under mild regularity conditions (see Theorems 8.1 and 8.2 in Newey and McFadden, 1994). The variance matrix of this estimator can be estimated using the semiparametric method in Newey (1994), or as shown by Ackerberg, Chen, and Hahn (2012) using a computationally simpler parametric-like method as in Newey (1984). We can also define a sequence of $K-step$ estimators in the same spirit as the Nested Pseudo Likelihood estimator in Aguirregabiria and Mira (2002). For $K\geq 2$, the K-step PML-EE estimators of θ and $\mathbf{P}$ are defined sequentially as:

$$\begin{aligned}\widehat{\theta}^{(K)}_N &= \arg\max_{\theta\in\Theta}\ Q_N\left(\theta,\widehat{\mathbf{P}}^{(K-1)}_N\right)\\ \widehat{\mathbf{P}}^{(K)}_N &= \Gamma^{(N)}_{EE-p}\left(\widehat{\theta}^{(K)}_N,\widehat{\mathbf{P}}^{(K-1)}_N\right)\end{aligned} \tag{33}$$

[14] We can also use the Empirical operators $\Gamma^{(N)}_{EE-p}$ or $\Gamma^{(N)}_{EE-v}$ to define GMM estimators of the structural parameters.

This PML estimator based on the Euler equation Pseudo Likelihood function implies an efficiency loss relative to the PML estimator based on a pseudo likelihood where the mapping Γ is the standard policy-iterations operator (or Newton-Kantorovich operator). As shown in Aguirregabiria and Mira (2002, Proposition 4), the two-step pseudo maximum likelihood estimator based on the policy-iteration operator is asymptotically equivalent to the maximum likelihood estimator. This efficiency property is not shared by any of the other Hotz-Miller type of two-step estimator. However, there is a trade-off in the choice between the PML estimator based on Euler equations and the one based on the policy-iteration operator. While the later is asymptotically efficient, its computational cost can be orders of magnitude larger than the computational cost for the estimator based on Euler equations (see Section 5). In models with large state spaces the implementation of the asymptotically optimal PML estimator may require approximation methods. In that case, the EE-based estimator can provide more precise estimates because it avoids approximation biases. We illustrate these trade-offs in our Monte Carlo experiments in Section 5.

4.3 Estimation of counterfactuals

Given a sample and an estimate of the structural parameters, $\widehat{\theta}$, the researcher is interested in estimating the behavioral effects of a change in the structural parameters from the estimate $\widehat{\theta}$ to an alternative vector θ^*. To estimate the effects of this counterfactual experiment on agents' behavior and payoffs, the researcher needs to solve the DP problem given the structural parameters θ^*. We can represent this solution either in terms of the vector of conditional choice probabilities $\mathbf{P}^*$ or in terms of the vector of value differences $\widetilde{\mathbf{v}}^*$. The vector $\widetilde{\mathbf{v}}^*$ is defined as the unique fixed point of the contraction mapping $\Gamma_{EE}(.,\theta^*)$, i.e., $\widetilde{\mathbf{v}}^* = \Gamma_{EE}(\widetilde{\mathbf{v}}^*, \theta^*)$.

In most empirical applications, the dimension of the state space, and in particular the dimension of $\mathcal{Z}$, is very large such that the exact computation of $\widetilde{\mathbf{v}}^*$ is computationally unfeasible. Here we propose an approximation to the solution using the Empirical EE mapping. We approximate $\widetilde{\mathbf{v}}^*$ using $\widetilde{\mathbf{v}}_N^*$. This approximate solution is defined as the unique fixed point of the Empirical EE mapping, $\widetilde{\mathbf{v}}_N^* = \mathbf{\Gamma}_{EE}^{(N)}(\widetilde{\mathbf{v}}_N^*, \theta^*) \equiv \{\Gamma_{EE}^{(N)}(a, y, \mathbf{z}; \widetilde{\mathbf{v}}) : (a, y, \mathbf{z}) \in \mathcal{A} \times \mathcal{Y} \times \mathcal{Z}_N\}$. And the corresponding vector of conditional choice probabilities is $\mathbf{P}_N^* = \Lambda(\widetilde{\mathbf{v}}_N^*)$. This approximate solution has the following interesting properties.

(a) Low computational cost and no curse of dimensionality. The vector $\widetilde{\mathbf{v}}_N^*$ and the mapping $\mathbf{\Gamma}_{EE}^{(N)}$ have the same dimension as the sample size. In most empirical applications, this dimension is many orders of magnitude smaller than the dimension of the state space. This reduction in the

dimension of the fixed point problem, together with the other computational properties of the EE operator, imply substantial computational savings. The Empirical EE mapping is an Euler equation version of the random operators defined in Rust (1997). Rust shows that these operators succeed in breaking the curse of dimensionality for Markov dynamic decision models with discrete choices and continuous state variables. This property also applies to our dynamic decision model when the endogenous state variables are discrete and exogenous state variables are continuous.

(b) Consistency. $\widetilde{\mathbf{v}}_N^*$ and $\mathbf{P}_N^*$ are consistent estimators of the true counterfactuals $\widetilde{\mathbf{v}}^*$ and $\mathbf{P}^*$.

PROPOSITION 8. The vector of value differences $\widetilde{\mathbf{v}}_N^*$ *that is defined as the fixed point* $\widetilde{\mathbf{v}}_N^* = \mathbf{\Gamma}_{EE}^{(N)}(\widetilde{\mathbf{v}}_N^*, \theta^*)$ *is a root-N consistent and asymptotically normal estimator of* $\widetilde{\mathbf{v}}^*$.
Proof: In the Appendix.

5 Monte Carlo experiments

In this section we present Monte Carlo experiments to illustrate the performance of the *Euler equation methods* in terms of computational savings and statistical precision in three problems: the exact solution of the DP problem; the estimation of structural parameters; and the estimation of counterfactual experiments. We evaluate our solution and estimation methods in the context of a dynamic model of market entry and exit.

First we examine the differences in the computational burdens of the four solution algorithms presented in section 3.3 above. Generally speaking, the total time required to obtain a model solution is comprised by two factors, the amount of time per iteration and the number of iterations. We use the experiments to compare the time per iteration and the number of iterations each method takes to convergence to obtain a better understanding of the computational costs.

Second, we present Monte Carlo experiments to evaluate the finite sample properties and computational costs of four estimators of the structural parameters of the model: two-step PML-EE estimator; sequential PML-EE estimator; two-step PML-policy function estimator (a variant of the Hotz-Miller CCP estimator); and the Maximum Likelihood estimator computed using the Nested Pseudo Likelihood algorithm in Aguirregabiria and Mira (2002).

Third, given an estimated model and a counterfactual experiment that consists of an increase in the parameter that represents the cost of market entry, we present Monte Carlo experiments to evaluate the finite sample properties of these methods to estimate counterfactual choice probabilities. These methods consist of finding a fixed point of the corresponding empirical operator: EE mapping, EE-prob mapping, and empirical versions of value function and policy function operators.

5.1 Design of the experiments

We consider a dynamic model of firm entry and exit decisions in a market. The decision variable a_t is the indicator of being active in a market, such that the action space is $\mathcal{A} = \{0, 1\}$. The endogenous state variable y_t is the lagged value of the decision variable, $y_t = a_{t-1}$, and it represents whether the firm has to pay an entry cost or not. The vector $\mathbf{z}_t$ of exogenous state variables includes firm productivity, and market and firm characteristics that affect variable profit, fixed cost, and entry cost.[15] We specify each of these components in turn.

An active firm earns a profit $\pi(1, \mathbf{x}_t) + \varepsilon_t(1)$ where $\pi(1, \mathbf{x}_t)$ is equal to the variable profit (VP_t) minus fixed cost (FC_t), and minus entry cost (EC_t). The payoff to being inactive is $\pi(0, \mathbf{x}_t) + \varepsilon_t(0)$, where we assume $\pi(0, \mathbf{x}_t) = 0$ for all possible values of $\mathbf{x}_t$. We assume that $\varepsilon_t(0)$ and $\varepsilon_t(1)$ are extreme value type 1 distributed with dispersion parameter $\sigma_\varepsilon = 1$. The variable profit function is $VP_t = [\theta_0^{VP} + \theta_1^{VP} z_{1t} + \theta_2^{VP} z_{2t}] \exp(\omega_t)$ where ω_t is the firm's productivity shock, z_{1t} and z_{2t} are exogenous state variables that affect the firm's price-cost margin in the market; and θ_0^{VP}, θ_1^{VP}, and θ_2^{VP} are parameters. The fixed cost is, $FC_t = \theta_0^{FC} + \theta_1^{FC} z_{3t}$, and the entry cost is, $EC_t = (1 - y_t)$ $[\theta_0^{EC} + \theta_1^{EC} z_{4t}]$, where the term $(1 - y_t)$ indicates that the entry cost is paid only if the firm was not active in the market at previous period, z_{3t} and z_{4t} are exogenous state variables, and and θ's are parameters. The vector of structural parameters in the payoff function is $\theta = (\theta_0^{VP}, \theta_1^{VP}, \theta_2^{VP}, \theta_0^{FC}, \theta_1^{FC}, \theta_0^{EC}, \theta_1^{EC})'$. The vector of exogenous state variables $\mathbf{z} = (z_1, z_2, z_3, z_4, \omega)$ has discrete and finite support. Each of the exogenous state variables takes K values. The dimension of the state space $|\mathcal{X}|$ is then $2 * K^5$. Each exogenous state variable follows a discrete-AR(1) process, and we use Tauchen's method to construct the transition probabilities of these discrete state variables (Tauchen, 1986).[16]

The DGP used in our numerical and Monte Carlo experiments is summarized in table 1.

[15] We treat productivity as observable. For instance, using data on firms' output and inputs the researcher can estimate production function parameters and productivity taking into account the selection problem due to endogenous entry and exit decisions, e.g., Olley and Pakes (1996), and Ackerberg, Caves, and Frazer (2015).

[16] Let $\{z_j^{(k)} : k = 1, 2, ..., K\}$ be the support of the state variable z_j, and define the width values $w_j^{(k)} \equiv z_j^{(k+1)} - z_j^{(k)}$. Let $\tilde{z}_{jt}$ be a continuous latent variable that follows the AR(1) process $\tilde{z}_{jt} = \gamma_0^j + \gamma_1^j \tilde{z}_{jt-1} + e_{jt}$, with $e_{jt} \sim$ i.i.d. $N(0, \sigma_j^2)$. Then, the transition probability for the discrete state variable z_{jt} is given by: $\Phi([z_j^{(1)} + (w_j^{(1)}/2) - \gamma_0^j - \gamma_1^j z]/\sigma_j)$ for $z' = z_j^{(1)}$; $\Phi([z_j^{(k)} + (w_j^{(k)}/2) - \gamma_0^j - \gamma_1^j z]/\sigma_j) - \Phi([z_j^{(k-1)} + (w_j^{(k-1)}/2) - \gamma_0^j - \gamma_1^j z]/\sigma_j)$ for $z' = z_j^{(k)}$ with $2 \leq k \leq K-1$; and $1 - \Phi([z_j^{(K-1)} + (w_j^{(K-1)}/2) - \gamma_0^j - \gamma_1^j z]/\sigma_j)$ for $z' = z_j^{(K)}$.

Table 1
Parameters in the DGP

Payoff Parameters:	$\theta_0^{VP} = 0.5; \quad \theta_1^{VP} = 1.0; \quad \theta_2^{VP} = -1.0$
	$\theta_0^{FC} = 0.5; \quad \theta_1^{FC} = 1.0$
	$\theta_0^{EC} = 1.0; \quad \theta_1^{EC} = 1.0$
Each z_j state variable:	z_{jt} is AR(1): $z_{jt} = \gamma_0^j + \gamma_1^j \; z_{jt-1} + \sigma_e \; e_{jt}$
	$\gamma_0^j = 0.0; \quad \gamma_1^j = 0.6$
Productivity :	ω_t is AR(1): $\omega_t = \gamma_0^\omega + \gamma_1^\omega \; \omega_{t-1} + \sigma_e^\omega \; e_t$
	$\gamma_0^\omega = 0.2; \;\; \gamma_1^\omega = 0.9$
Low persistence model:	$\sigma_e^\omega = \sigma_e = 1$
High persistence model:	$\sigma_e^\omega = \sigma_e = 0.01$
Discount factor	$\beta = 0.95$

5.2 Comparing solution methods

We compare the computing times of five methods for the exact solution of the model: (a) value function iterations; (b) relative value function iterations; (c) policy function iterations; (d) iterations in the EE operator; and (e) iterations in the EE-prob operator. Methods (a) to (c) are the most standard algorithms for the solution of DP problems. Methods (d) and (e) are new algorithms that we propose in this paper. We compare these methods for six different dimensions of the state space $|\mathcal{X}|$: 64, 486, 2048, 6250, 15552, and 200,000 that correspond to values 2, 3, 4, 5, 6, and 10, respectively, for the number of points in the support of each exogenous state variable. Despite we use the same starting values to initialize the different algorithms, it might be the case that the relative performance of these methods depends on the initial value. To check for this possibility, we have implemented this experiment using ten different initial values, the same for all the algorithms. We find very small differences in the relative performance of the algorithms across the different initial values. The numbers in tables 3 and 4 are the averages over these initial values.

First, we analyze the degree of contraction of these algorithms. The number of iterations that a fixed point algorithm needs to achieve convergence is closely related to the degree of contraction of the mapping. To compare the degree of contraction, we calculate an approximation to the Lipschitz constants of the mappings. The Lipschitz constant of a mapping Γ in a complete metric space $(\mathcal{V}, \|.\|)$ is defined as the smallest constant c that satisfies the condition $\|\Gamma(\mathbf{V}) - \Gamma(\mathbf{W})\|/\|\mathbf{V} - \mathbf{W}\| \leq c$ for any two points $\mathbf{V}$ and $\mathbf{W}$ in the domain $\mathcal{V}$. For instance, for the value function mapping Γ_{VF} and using the sup-norm (or uniform norm), the Lipschitz constant

is defined as:

$$L(\Gamma_{VF}) \equiv \sup_{\mathbf{V},\mathbf{W}\in\mathbb{R}^{|\mathcal{X}|}} \left[\frac{\sup_{\mathbf{x}\in\mathcal{X}} |\Gamma_{VF}(\mathbf{x},\mathbf{V}) - \Gamma_{VF}(\mathbf{x},\mathbf{W})|}{\sup_{\mathbf{x}\in\mathcal{X}} |V(\mathbf{x}) - W(\mathbf{x})|} \right] \tag{34}$$

Calculating the exact value of the Lipschitz constant for any of the mappings we consider is not a practical option because the dominion of all these mappings is infinite. As such we obtain the following approximation. Let $\{\mathbf{V}_k : k = 0, 1, ..., \mathcal{I}_{\mathbf{V}_0}\}$ be the sequence of values that we obtain by applying successive iterations in the mapping Γ given an initial value $\mathbf{V}_0$, where $\mathcal{I}_{\mathbf{V}_0}$ is the number of iterations to reach convergence. Then, we obtain an approximation (i.e., a lower bound) to the Lipschitz constant of this mapping by considering the ratios $\|\Gamma(\mathbf{V}) - \Gamma(\mathbf{W})\|/\|\mathbf{V} - \mathbf{W}\|$ at the pair of values $\mathbf{V}_k$ and $\mathbf{V}_{k+1}$ generated in the sequence. That is, the approximation to the Lipschitz constant of mapping Γ is:

$$\widetilde{L}_{\mathbf{V}_0}(\Gamma) \equiv \max_{k\in\{0,1,...,\mathcal{I}_{\mathbf{V}_0}\}} \frac{\|\Gamma(\mathbf{V}_{k+1}) - \Gamma(\mathbf{V}_k)\|}{\|\mathbf{V}_{k+1} - \mathbf{V}_k\|} \tag{35}$$

To obtain a better approximation, we generate sequences from many initial guesses, and take as our approximation to the Lipschitz constant the maximum over all these sequences.

Table 2 reports the Lipschitz constants of the five operators for the different dimensions of the state space, and for two different versions of the model according to the time persistence of the exogenous state variables: a model with *low persistence* where $\sigma_e^\omega = \sigma_e = 1$ that implies $\rho(\mathbf{F}_z) = 0.56$; and a model with *high persistence* where $\sigma_e^\omega = \sigma_e = 0.001$ that implies a $\rho(\mathbf{F}_z)$ close to one. The Lipschitz constants are very stable across the different dimensions of the state space. For the low persistence model, there is a very substantial difference in the degree of contraction of the VF operator and the other operators. The PF and the EE operators are always the strongest contractions. In the model with low persistence, the RVF operator is a stronger contraction than VF, but it is still far away from the EE and the PF operators. The good contraction properties of the RVF and the EE-prob operators completely disappear in the model with highly persistent state variables. Importantly, that is not the case for the EE operator: the Lipschitz constant of this operator increases when $\rho(\mathbf{F}_z)$ gets close to one but it is still substantially smaller than the Lipschitz constant of VF and RVF operators. This is the main result that we want to emphasize in table 2. The EE operator has strong contraction properties, similar to the ones of the PF operator, regardless of the degree of persistence in the exogenous state variables.

Table 2
Degree of Contraction (Lipschitz Constants) of the mappings

	Low Persistence Model					High Persistence Model				
$\|\mathcal{X}\|$	**EE**	**EE-p**	**PF**	**VF**	**RVF**	**EE**	**EE-p**	**PF**	**VF**	**RVF**
64	0.20	0.30	0.14	0.95	0.59	0.34	0.90	0.56	0.95	0.95
486	0.18	0.21	0.13	0.95	0.54	0.34	0.82	0.47	0.95	0.95
2,032	0.18	0.18	0.11	0.95	0.53	0.31	0.88	0.86	0.95	0.95
6,250	0.18	0.18	0.10	0.95	0.53	0.32	0.93	0.69	0.95	0.95
15,552	0.18	0.18	0.09	0.95	0.53	0.28	0.96	0.62	0.95	0.95
200,000	0.18	0.18	-	0.95	0.53	0.28	0.96	-	0.95	0.95

Table 3 presents time per iteration, number of iterations, and total computation time for the different algorithms and dimensions of the state space for the model with low persistence of the state variables, which is a very optimistic scenario for the relative value iterations algorithm. For every dimension of the state space, iterating in the EE mapping is always the most efficient algorithm, followed closely by iterating in the EE-prob mapping. The computational gains relative to the standard methods are very substantial, as shown in the columns at the bottom-right of the table that report the ratio of computing times relative to the EE method. Furthermore, these gains increase with the dimension of the state space.

The PF algorithm is the least efficient algorithm despite the fact that this operator is the stronger contraction and converges to the solution after only 5 iterations. The inefficiency of the PF algorithm comes from the cost of its valuation step that increases cubically with the dimension of the state space. EE is more efficient than VF and RVF both because its time-per-iteration is lower and because it is a stronger contraction such that it requires a smaller number of iterations. These advantages are quite substantial. The time-per-iteration of VF and RVF is more than twice that of the EE. For the largest dimension of the state space (still a modest dimension relative to most empirical applications), the total times of VF and RVF are 85 and 7 times, respectively, more than the EE method.

Table 3
Comparison of Solution Methods
(Model with low persistence)

Number of states	**Number of Iterations**					**Time per Iteration (in seconds)**				
$\lvert\mathcal{X}\rvert$	EE	EE-p	PF	VF	RVF	EE	EE-p	PF	VF	RVF
64 (= 2*2^5)	13	12	5	351	36	<0.001	<0.001	0.01	<0.001	<0.001
486 (= 2*3^5)	13	12	5	246	31	<0.001	<0.001	0.60	<0.001	<0.001
2048 (= 2*4^5)	13	12	5	345	30	0.005	0.005	16.29	0.01	0.01
6250 (= 2*5^5)	13	12	5	344	30	0.04	0.04	150.0	0.08	0.08
15552 (= 2*6^5)	13	12	5	344	30	0.17	0.21	916.2	0.46	0.46
200,000 (= 2*10^5)	13	12	5*	344	30	21.05	23.37	198,978	67.64	64.47

Number of states	**Total Time (in seconds)**					**Time Ratios**				
$\lvert\mathcal{X}\rvert$	EE	EE-p	PF	VF	RVF	$\frac{\text{EE}}{\text{EE}}$	$\frac{\text{EE-p}}{\text{EE}}$	$\frac{\text{PF}}{\text{EE}}$	$\frac{\text{VF}}{\text{EE}}$	$\frac{\text{RVF}}{\text{EE}}$
64 (= 2*2^5)	<0.001	<0.001	0.05	0.03	<0.001	1	1	60	37.0	4.0
486 (= 2*3^5)	0.01	0.01	3.02	0.35	0.03	1	1	300	26.6	3.4
2048 (= 2*4^5)	0.05	0.06	81.44	2.76	0.27	1	1	1320	53.1	5.2
6250 (= 2*5^5)	0.47	0.47	750.0	26.14	2.37	1	1	1,595	55.9	5.1
15552 (= 2*6^5)	2.25	2.53	4,581	158.9	13.89	1	1	2,040	70.7	6.2
200,000 (= 2*10^5)	273.6	280.4	$\simeq$ 1M*	23,270	1,934	1	1	3,630	85.0	7.1

Note: EE = EE iterations; EE-p = EE-prob iterations; PF = Policy iterations; VF = Value iterations; RVF = Relative value iterations.
Note (*): Estimated number of iterations, and estimated total time given the observed time-per-iteration.

Table 4 compares the EE and RVF algorithms in the model with high persistence of the state variables. As we have shown in table 3 above, in this model the RVF does not have strong contraction properties and its computational properties are practically the same as VF iterations. In contrast, the EE still has strong contraction properties. As a result, table 4 shows that for this model the performance of the EE algorithm is strongly superior to RVF.

Rust (1987, 1988), Powell (2007), and Bertsekas (2011) advocate using a hybrid value-policy

iteration method. The algorithm starts with value function iterations until a loose convergence criterion is reached. Then, the algorithm switches to policy function iterations. When the switching point is appropriately tuned, this algorithm can be faster than both value function and policy function iterations. We have not reported results from this hybrid method in this experiment. The main reason is that, even for the moderate dimensions of the state space in tables 2-4, one single policy iteration takes almost one hundred times longer than the whole time to convergence of the EE method. In other words, even for this moderate dimension, the optimal 'hybrid' algorithm is the pure value function iteration method.

Table 4
Comparison of EE and RVF Solution Methods
(Model with high persistence)

Number of states	**EE iterations**			**Relative value iterations**			**Ratio total time**
$\lvert\mathcal{X}\rvert$	# iter.	Time-per-iter.	Total time	# iter.	Time-per-iter.	Total time	**RVF / EE**
64 (= 2*2^5)	24	<0.001	0.002	378	<0.001	0.05	24.5
486 (= 2*3^5)	24	<0.001	0.03	350	<0.001	0.79	30.4
2048 (= 2*4^5)	17	0.01	0.16	341	0.02	6.28	38.8
6250 (= 2*5^5)	18	0.06	0.99	322	0.12	37.94	38.0
15552 (= 2*6^5)	17	0.43	7.15	333	0.87	289.13	40.5
200,000 (= 2*10^5)	16	46.2	721.1	319	98.5	18,818	42.5

5.3 Estimation of structural parameters

For the Monte Carlo experiments that deal with the estimation of the structural parameters, we consider the DGP described in table 1 with a state space with size $|\mathcal{X}| = 6250$ and low persistence of the state variables. We generate $1,000$ samples with a sample size of $N = 1,000$ firms and $T = 2$ time periods. For the first sample period, the value of the vector of state variables is drawn from its ergodic or steady-state distribution.

Table 5
Monte Carlo Experiments. Estimation of Structural Parameters
Sample size: $N = 1,000$ & $T = 2$; Monte Carlo samples = $1,000$

	Mean Absolute Bias				**Root Mean Squared Error**			
Parameter (True value)	**2-step PF**	**2-step EE**	**MLE (NPL)**	**K-step EE**	**2-step PF**	**2-step EE**	**MLE (NPL)**	**K-step EE**
$\theta_0^{VP}(0.5)$	0.061 (12.3 %)	0.073 (14.6 %)	0.057 (11.3 %)	0.057 (11.4 %)	0.078 (15.6 %)	0.095 (18.9 %)	0.073 (14.6 %)	0.073 (14.6 %)
$\theta_1^{VP}(1)$	0.062 (6.2 %)	0.078 (7.8 %)	0.062 (6.2 %)	0.062 (6.2 %)	0.077 (7.7 %)	0.100 (10.0 %)	0.077 (7.7 %)	0.078 (7.8 %)
θ_2^{VP}(-1)	0.057 (5.7 %)	0.074 (7.4 %)	0.058 (5.8 %)	0.058 (5.8 %)	0.073 (7.3 %)	0.093 (9.3 %)	0.073 (7.3 %)	0.074 (7.4 %)
$\theta_0^{FC}(1.5)$	0.108 (7.2 %)	0.206 (13.7 %)	0.096 (6.4 %)	0.097 (6.5 %)	0.140 (9.3 %)	0.247 (16.5 %)	0.125 (8.3 %)	0.127 (8.5 %)
$\theta_1^{FC}(1)$	0.058 (5.8 %)	0.093 (9.3 %)	0.060 (6.0 %)	0.059 (5.9 %)	0.072 (7.2 %)	0.118 (11.8 %)	0.075 (7.5 %)	0.074 (7.4 %)
$\theta_0^{EC}(1)$	0.096 (9.6 %)	0.096 (9.6 %)	0.095 (9.5 %)	0.096 (9.6 %)	0.121 (12.1 %)	0.122 (12.2 %)	0.121 (12.1 %)	0.121 (12.1 %)
$\theta_1^{EC}(1)$	0.114 (11.4 %)	0.130 (13.0 %)	0.109 (10.9 %)	0.111 (11.1 %)	0.142 (14.2 %)	0.160 (16.0 %)	0.136 (13.6 %)	0.138 (13.8 %)
Total MAB	0.557	0.751	0.535	0.540				
Ratio $\frac{MAB\ \text{PF}}{MAB\ \text{EE}}$	0.74		0.99					
Total RMSE					0.703	0.935	0.680	0.684
Ratio $\frac{RMSE\ \text{PF}}{RMSE\ \text{EE}}$					0.75		0.99	
Time (in secs)	1067.80	0.218	7345.79	3.261	1067.80	0.218	7345.79	3.261
Ratio $\frac{Time\ \text{PF}}{Time\ \text{EE}}$	4898		2252		4898		2252	

Note: In parentheses, the statistic as percentage of the true parameter value.

We compute four estimators: (a) two-step PML-EE, as defined in equations (31) and (32); (b) sequential or K-step PML-EE, as defined in equation (33); (c) two-step PML - Policy Function; and (d) and Maximum Likelihood estimator (MLE) computed using the Nested Pseudo Likelihood (NPL) algorithm.[17] Aguirregabiria and Mira (2002) show that the two-step PML - Policy Function estimator is asymptotically equivalent to the MLE. First-step choice probabilities were estimated using a simple raw frequency estimator. For the number of steps of the recursive estimators (i.e., K-step PML-EE and the NPL implementation of MLE), we set $K = 15$, but we achieved convergence of the estimation algorithms in less that 15 iterations for almost all the $1,000$ Monte Carlo simulations. By definition of the PML-EE estimators, the expectation of $\mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}$ is replaced by its sample counterpart at the observed states in the sample. In contrast, the implementation of the two-step PML-PF and MLE require the transition probabilities at every possible value of the state variables. For these two estimators, we have used the true values of the transition probabilities. Note that this aspect of our experiment contributes to increase the estimation error of the EE estimator relative to the policy iteration estimator, as the second uses the true transition probabilities.

Table 5 presents the results of these Monte Carlo experiments. We report the mean (over simulations) absolute bias of the parameter estimates, and the root-mean squared error, and these statistics as percentages of the true parameter value. We also report the average (over simulations) computing time to implement each estimator. These experiments provide several interesting results. First, the two-step PML-PF estimator performs extremely well. As shown in Aguirregabiria and Mira (2002), this two-step estimator is asymptotically efficient but it has finite sample bias due to the fact that it relies on imprecise first step estimates of the conditional choice probabilities. However, for this Monte Carlo experiment, there are not large improvements from recursively updating the conditional choice probabilities with new estimates of the parameters. Most of the improvement from this iterative procedure concentrates on the estimate of the entry cost parameter θ_1^{EC}: as percentage of the true parameter, the Mean Absolute Bias goes from 11.4% to 10.9%, and the Root Mean Square Error declines from 14.2% to 13.6%. The average computing times of the two-step PML-PF and the MLE (using NPL) estimators are 1067 and 7345 seconds, respectively. In this experiment, given that the two-step PML-PF estimator is already quite precise, the additional computational cost of the MLE (seven times the cost of the two-step estimator) has low returns in terms of statistical properties, i.e., a reduction in the RMSE of 0.6 percentage points at a cost of

[17] For the computation of the MLE in single-agent models, the NPL algorithm is several times faster than the nested fixed point algorithm (see Aguirregabiria and Mira, 2002, or Bray, 2016).

6378 CPU seconds.

Compared to the (asymptotically efficient) two-step PML-PF estimator, the two-step Euler Equations estimator has more Mean Absolute Bias (MAB) (i.e., 26% greater) and Root Mean Square Error (RMSE) (i.e., 25% greater). Then, if the two-step PML-PF estimator is so close the MLE and also works so well relative to the two-step EE estimator, then why bother with Euler Equation estimators? The answer lies in: (a) the relative CPU time of the two methods; and (b) the fact that when we consider the K-step estimators, the statistical differences between the EE estimators and the PF estimator all but vanish. First, note that the two-step PML-EE estimator requires less than a quarter of a second (0.218 seconds) while the two-step PML-PF estimator requires more than 17 minutes (1067 seconds), almost 5000 times longer. This is the trade-off for the two-step estimator - 25% better statistical properties in exchange for about 5000 times longer computation time. This trade-off becomes even more stark when we compare the K-step PML - EE estimator with either the two-step PML-PF or the MLE (NPL) estimators. In terms of MAB and RMSE, the K-step PML-EE estimator is superior to the two-step PML-PF and it is essentially indistinguishable from the MLE. However the K-step PML-EE estimator is thousands of times faster than either the two-step PML-PF or MLE (NPL) estimators, and there is no real trade-off to be made. The K-step PML-EE estimator is statistically as good as the PF estimators, but more than 2000 times faster.

Overall, the experiment illustrates that the K-step EE estimator performs very well relative to the best possible estimator, the MLE. There is no substantial difference in MAB or RMSE between them, and K-step EE estimator has better statistical properties than the two-step estimators. That is, iterations (or sequential estimation) tend to improve the EE estimators more than the PF estimators. This is intuitive. The two-step PML-PF estimator is asymptotically equivalent to the MLE, so that all the improvements are in terms of finite sample bias, while for the EE estimators the improvements are both finite sample and asymptotic (i.e., the zero Jacobian property does not hold for the EE and therefore the sequence of EE K-step are not asymptotically equivalent). This experiment indicates that the massive computational savings associated with using the EE K-step estimator seems to justify its application, given there is little statistical cost relative to the MLE.

In a recent paper, Bray (2016) proposes a modified version of the NPL method that replaces the policy function iterations in outer algorithm with relative value iterations. For models with low persistence of the state variables, Bray shows that this modified NPL method (*Strong NPL*, as defined by the author) is 10 times faster than the standard NPL method. Though this a substantial

improvement in the amount of time to compute the MLE, it is of a very different order of magnitude from the ratio of 2252 between NPL and K-stage PML-EE that we report in table 5.

5.4 Estimation of counterfactuals

We now study how standard methods and Euler Equation-based methods perform in counterfactual policy experiments. Given the large dimension of the state space in actual applications, the exact computation of a solution of the model (under the factual and counterfactual scenarios) is computationally unfeasible, at least using standard solution methods of value function or policy function iterations. The purpose of this section is twofold. First, we show that the Euler Equation solution method makes it feasible solving the model *exactly* when the dimension of the state space is relatively large, i.e., of the order of a few million values. We show that for this dimension, and using standard computing equipment, the Euler Equation method solves the model in a few hours while standard methods require computing times between a few weeks (for value function iterations) and more than one year (for policy function iterations). A second purpose of this section is to study the relative performance of the Euler Equation and the standard methods in empirical applications where the model cannot be solved exactly and we need to use approximation methods such as the Empirical EE operator.

5.4.1 Counterfactual question

The counterfactual policy we consider is an increase in the cost of entry in the context of the dynamic entry-exit model presented above. The presence of entry costs can generate misallocation in an industry. There may be potential entrants that are more productive than incumbent firms but are not willing to enter in the market and replace the less efficient firms because the entry cost makes this unprofitable. Furthermore, the presence of entry costs makes exit less attractive to incumbent firms, because re-entry is more expensive, and in this way higher entry costs may discourage low productivity incumbents from exiting. We are interested in the quantification of the net effect of entry cost on total industry productivity.

Suppose that the industry consists of N potential entrants, indexed by i. Competition in this industry is characterized by monopolistic competition, i.e., a single-agent model. The expected value of the total output produced by firms active in the industry, is:

$$Q^* = \mathbb{E}\left(\sum_{i=1}^{N} a_{it} \ \exp\left(\omega_{it}\right)\right) = N \ \sum_{\mathbf{z},\omega} p(\mathbf{z},\omega) \ \exp\left(\omega\right) \ f^*(\mathbf{z},\omega) \tag{36}$$

where $f^*(\mathbf{z},\omega)$ represents the steady-state or ergodic distribution of the exogenous variables, and

$p(\mathbf{z},\omega)$ is the probability that a firm is active when the exogenous state variables take the values $(\mathbf{z},\omega)$, i.e., $p(\mathbf{z},\omega) \equiv \Pr(a_{it}=1|\mathbf{z}_t=\mathbf{z},\omega_{it}=\omega)$. Note that $p(\mathbf{z},\omega)$ is different from the conditional choice probability function because the probability $p(\mathbf{z},\omega)$ is not conditional on the firm's incumbent status at previous period. However, by definition, the steady-state condition implies the following relationship between $p(\mathbf{z},\omega)$ and the conditional choice probabilities $P(0,\mathbf{z},\omega)$ and $P(1,\mathbf{z},\omega)$:

$$p(\mathbf{z},\omega) = (1-p(\mathbf{z},\omega))\ P(0,\mathbf{z},\omega) + p(\mathbf{z},\omega)\ P(1,\mathbf{z},\omega) \tag{37}$$

Rearranging we get, $p(\mathbf{z},\omega) = P(0,\mathbf{z},\omega)/[1-P(1,\mathbf{z},\omega)+P(0,\mathbf{z},\omega)]$. Using this expression to obtain the partial derivative $\partial p(\mathbf{z},\omega)/\partial EC$, where EC represents the entry cost, it is possible to show that the effect of entry cost on $p(\mathbf{z},\omega)$ and Q^* is ambiguous and it depends on the values of the state variables and the structural parameters. An increase in the entry cost has a negative effect on the probabilities of entry $P(0,\mathbf{z},\omega)$ for potential entrants, but it has a positive effect on the probability of staying in the industry $P(1,\mathbf{z},\omega)$ for incumbent firms. Since these effects are of opposite sign, the entry cost has an ambiguous net effect on the steady-state probability that a firm is active, $p(\mathbf{z},\omega)$, and on total output Q^*.

5.4.2 Exact solution of factual and counterfactual models

Suppose that a researcher has estimated this model and is interested in the effect of the entry cost on expected total industry output: i.e., the effect of a change in EC on Q^*. More specifically, the counterfactual experiment we consider is an increase in the entry cost parameter θ_0^{EC} from 1 to 2.5. First, we show that when the dimension of the state space is relatively large (i.e., more than 1 million points), the Euler Equation solution method can be used to solve the model *exactly* under the factual and counterfactual scenarios, while standard solution methods are infeasible.

In this experiment, we consider a DGP with a state space with $|\mathcal{Z}_1| = |\mathcal{Z}_2| = |\mathcal{Z}_3| = |\mathcal{Z}_4| = |\Omega| = 14$, such that the number of points in the complete state space is $|\mathcal{X}| = 2 * 14^5 = 1,075,648$. The values of the structural parameters are the ones given in table 1 above. We have solved the model, under the factual and counterfactual scenarios, by iterating in the EE mapping. Using standard computing equipment, the time-per-iteration was 546 seconds, and the total time (with 13 iterations) was $7,048$ seconds, i.e., less than two hours. In contrast, the VF algorithm have a time-per-iteration of 1528 seconds and its total time is 146 hours, more than 6 days.[18]

[18] This total time for the VF method is estimated using the condition that the number of iterations to convergence is 344, which is consistent with the results in table 3. The estimated total time for the PF algorithm is more than two years.

Table 6
Properties of Factual and Counterfactual Scenarios in Steady-State

	(1) Probability Being Active	(2) Entry Probability	(3) Exit Probability	(4) State Persistence	(5) Output
(A) Factual DGP	0.323	0.274	0.580	0.768	0.529
(B) Counterfactual DGP	0.258	0.157	0.513	0.884	0.423
Policy Effect: (B) - (A)	-0.065	-0.117	-0.068	+0.116	-0.106
(Percentage change)	(-20.1%)	(-42.7%)	(-11.7%)	(15.1%)	(-20.0%)

Table 6 presents predictions from the exact solution of the model under the factual and counterfactual scenarios. We calculate the following average outcomes: (1) the probability of being active: $p^* = \sum_{\mathbf{z},\omega} p(\mathbf{z},\omega)\ f^*(\mathbf{z},\omega)$; (2) the probability of entry: $P_0^* = \sum_{\mathbf{z},\omega} P(0,\mathbf{z},\omega)\ f^*(\mathbf{z},\omega)$; (3) the probability of exit: $1 - P_1^* = \sum_{\mathbf{z},\omega}(1 - P(1,\mathbf{z},\omega))\ f^*(\mathbf{z},\omega)$; (4) state persistence: $\Pr(a_t = y_t) = \sum_{\mathbf{z},\omega}[p(z,\omega)\ P(1,\mathbf{z},\omega) + (1-p(z,\omega))\ (1-P(0,\mathbf{z},\omega))]\ f^*(\mathbf{z},\omega)$; and (5) output per potential entrant: $q^* = \sum_{\mathbf{z},\omega} p(\mathbf{z},\omega)\ \exp(\omega)\ f^*(\mathbf{z},\omega)$. Column (1) shows that, in this experiment, the effect of an increase in θ_0^{EC} from 1 to 2.5 is a reduction in the probability of being active by 6.5 percentage points. The increase in entry cost reduces the average probability of entry (by 11.7 percentage points), and the average probability of exit (by 6.8 percentage points). As non-incumbents are more likely to remain outside the market when the cost of entry rises, and incumbents are less likely to exit, we expect an increase in the persistence of the activity decision, which we see in column (4). Column (5) shows that the increase in the entry cost implies an 20% reduction in industry output.

5.4.3 Estimation of counterfactuals: Monte Carlo experiment

The dimension of the state space in this experiment, with more than 1 million points, is still small relative to the dimensions that we find in actual applications. For instance, in an empirical application of this model the five exogenous state variables can be continuous variables with substantial variability across firms and over time. Even if the researcher is willing to discretize each of these continuous variables, an accurate representation of the distribution of these variables and their variation over time may require around one hundred grid points per variable. In our model, this implies a state space with approximately 20 billion points. We would like to compare the performance of the different approximation methods in this type of realistic scenario. However, that

model cannot be solved exactly, and therefore, we cannot calculate the true approximation errors of the different methods. Instead, we consider that the true DGP is the model in previous section with $2 * 14^5$ possible states and compare the performance of different approximation methods by fixing the computational time that the researcher is willing to pay in order to obtain an approximate solution of the model.

Table 7
Monte Carlo Results: Average Computation Times

	VF	PF	EE-prob	EE
# Points in spaces $\mathcal{Z}_{N_{method}}$	100	100	1,735	1,735
Number of Iterations	350	5	13.4	12.7
Time Per Iteration	0.001	0.007	0.002	0.002
Total Time	0.062	0.036	0.024	0.026

To make a fair comparison, we keep the time required to solve the fixed points equal for the different solution/approximation methods. As defined above, $\mathcal{Z}$ and $\mathcal{Z}_N$ denote the true set and the sample set of values of the exogenous state variables. Let N_{EE}, N_{PF}, and N_{VF} be three integers smaller than the sample size (number of firms) N. Define $\mathcal{Z}_{N_{EE}}$, $\mathcal{Z}_{N_{PF}}$, and $\mathcal{Z}_{N_{VF}}$ as the sets of sample values of the exogenous state variables for the first N_{EE}, N_{PF}, and N_{VF} firms in the sample, respectively, and over the T sample periods. In the Monte Carlo experiments we solve the model by iterating in the empirical EE mapping defined on the space $\mathcal{Z}_{N_{EE}}$, and by iterating in the PF and VF mappings defined on the spaces $\mathcal{Z}_{N_{PF}}$ and $\mathcal{Z}_{N_{VF}}$. In order to keep the computation time (roughly) the same across the three methods, we choose $N_{PF} < N_{VF} < N_{EE}$ appropriately. That is, we use a smaller space of the exogenous state variables when we solve the model using PF and VF iterations than when we use EE iterations. This is a practical approach for the PF and VF algorithms, as the researcher would likely need to use some type of approximation method with a large state space. To solve the mappings on these reduced spaces, we must define the transition probabilities of the exogenous variables in each case. As in the case of the parameter estimates in the previous subsection, we use the true transition probabilities, normalized to the reduced space. Table 7 presents the number of points in the spaces $\mathcal{Z}_{N_{EE}}$, $\mathcal{Z}_{N_{VF}}$, and $\mathcal{Z}_{N_{PF}}$ and the average (across the Monte Carlo simulations) computing times of the different methods.

Table 8
Monte Carlo Results: Counterfactual Estimates
Effect of Entry Cost on Number of firms, Turnover rates, and Output

	(1) Probability Being Active	(2) Entry Probability	(3) Exit Probability	(4) State Persistence	(5) Total Output
True Policy Effect	-0.065	-0.117	-0.068	+0.116	-0.106
Mean Absolute Bias					
VF iterations	0.036 (55.9%)	0.016 (13.7%)	0.016 (24.2%)	0.011 (9.4%)	0.070 (66.3%)
PF iterations	0.036 (55.9%)	0.016 (13.7%)	0.016 (24.2%)	0.011 (9.4%)	0.070 (66.3%)
EE iterations	0.024 (36.8%)	0.007 (6.2%)	0.010 (14.4%)	0.008 (7.1%)	0.040 (38.1%)
EE-prob iterations	0.024 (36.8%)	0.007 (6.2%)	0.010 (14.4%)	0.008 (7.1%)	0.040 (38.1%)
Root Mean Square Error					
VF iterations	0.044 (67.5%)	0.020 (17.1%)	0.020 (30.1%)	0.014 (11.8%)	0.088 (82.8%)
PF iterations	0.044 (67.5%)	0.020 (17.1%)	0.020 (30.1%)	0.014 (11.8%)	0.088 (82.8%)
EE iterations	0.024 (37.1%)	0.007 (6.3%)	0.010 (14.8%)	0.008 (7.2%)	0.041 (38.6%)
EE-prob iterations	0.024 (37.1%)	0.007 (6.3%)	0.010 (14.8%)	0.008 (7.2%)	0.041 (38.6%)

Note: In parenthesis, the statistic in percentage over the true value of the parameter.

Table 8 presents the results of this Monte Carlo experiment. We estimate the policy effects (1) to (5) that we describe in table 6. We implement four different iterative methods: EE-prob, EE, Value function iterations, and Policy function iterations. We use 500 Monte Carlo simulations from the DGP, and calculate Root Mean Squared Error (RMSE) and the Mean Absolute Bias (MAB) based on these simulations. The statistical properties of the approximation methods based on the EE-prob and EE mappings are essentially equivalent. The results clearly indicate that the EE-methods have considerably lower bias and RMSE than both the VF and the PF approximation methods for all the five statistics. Systematically over the different parameters, the RMSE of the VF and PF methods are twice as large than these statistics for the EE methods.

6 Conclusion

Aguirregabiria and Magesan (2013) use a representation of the discrete choice model as a continuous decision problem where the decision variables are choice probabilities to derive marginal conditions of optimality similar in nature to the Euler equations in standard continuous decision problems. We build on this previous result to make two contributions.

First, we show that these Euler equations imply a fixed point mapping in the space of conditional

choice values (EE operator), and that this mapping is a contraction with a contraction constant that it is strictly smaller than the discount factor. We show that solving the dynamic programming problem by iterating in this operator provides very substantial computational gains relative to the standard solution methods of value function, relative value function, and policy function iterations. In particular, the Euler equations operator is a stronger contraction than value function and relative value function iterations, and its time-per-iteration is lower.

Second, we define a sample version of the EE operator and use it to construct consistent estimators of the structural parameters and counterfactual experiments. This sample-based EE operator is also a contraction but it is defined only at sample points of the exogenous state variables, and thus its dimensionality does not increase with the dimension of the state space.

We use Monte Carlo experiments to illustrate the computational gains associated with the new solution and estimation methods. In the context of a dynamic model of entry and exit, computing the exact solution on a moderately sized state space using the standard policy iteration can take thousands of times as long as computing the exact solution using the EE mapping, implying that models that are computationally infeasible for all practical purposes using standard methods, are feasible using the method we propose. We then illustrate using Monte Carlo experiments that the finite sample properties of (pseudo) likelihood estimates using the Euler equations representation are statistically indistinguishable from the Maximum Likelihood estimates, but at significant computational savings - the maximum likelihood estimator requires thousands of times as much CPU time as the EE-based pseudo likelihood estimator. Finally, we study the methods' relative ability to estimate the effect of counterfactual increase in the cost of entry. We show that, for a fixed computation time the finite sample properties of the EE mapping are better than those of the estimator associated with policy iterations in terms of both root mean squared error and mean absolute bias.

APPENDIX. PROOFS OF PROPOSITIONS

Our proofs of Propositions 2 to 4 apply Lemmas 1, 2, and 3 below. Therefore, we first present and prove these Lemmas. For the sake of completeness, we also present a proof of Williams-Daly-Zachary (WDZ) Theorem.

For Lemmas 1 to 3 and WDZ, we consider the following general Random Utility Model that includes our framework as a particular case. Given a vector of values $\mathbf{v} \equiv (v_0, v_1, ..., v_J)$, and a vector of random variables $\varepsilon \equiv (\varepsilon_0, \varepsilon_1, ..., \varepsilon_J)$ with joint CDF $G(\varepsilon)$ that is continuously differentiable, consider the discrete choice problem $\max_{a\in\{0,1,...,J\}} [v_a + \varepsilon_a]$. Without loss of generality, we can represent this discrete choice problem as $\max_{a\in\{0,1,...,J\}} [\widetilde{v}_a + \widetilde{\varepsilon}_a]$ where $\widetilde{v}_a \equiv v_a - v_0$ and $\widetilde{\varepsilon}_a \equiv \varepsilon_a - \varepsilon_0$. Let $\widetilde{\mathbf{v}}$ be the vector of value differences $(\widetilde{v}_1, \widetilde{v}_2, ..., \widetilde{v}_J)$, and let $\mathbf{P} = (P_1, P_2, ..., P_J)'$ be a vector of probabilities in the J dimensional simplex. Given this problem, we define the following functions:

(i) the Optimal Choice Probability (OCP) mapping, $\Lambda(\widetilde{\mathbf{v}}) \equiv \{\Lambda(a, \widetilde{\mathbf{v}}) : a = 1, 2, ..., J\}$ with $\Lambda(a, \widetilde{\mathbf{v}}) = \int 1\{\widetilde{v}_a + \widetilde{\varepsilon}_a \geq \widetilde{v}_j + \widetilde{\varepsilon}_j \text{ for any } j\}\, dG(\varepsilon)$;

(ii) the inverse of the OCP mapping, $\widetilde{\mathbf{v}} = \Lambda^{-1}(\mathbf{P})$;

(iii) *McFadden's Social Surplus function,* $S(\widetilde{\mathbf{v}}) \equiv \int \max_{j\in\{0,1,...,J\}} [\widetilde{v}_j + \widetilde{\varepsilon}_j]\, dG(\varepsilon)$;

(iv) the expectation function, $e^{(v)}(a, \widetilde{\mathbf{v}}) \equiv \mathbb{E}[\widetilde{\varepsilon}_a \mid \widetilde{v}_a + \widetilde{\varepsilon}_a \geq \widetilde{v}_j + \widetilde{\varepsilon}_j \text{ for any } j]$;

(v) the expectation function $e(a, \mathbf{P}) \equiv e^{(v)}(a, \Lambda^{-1}(\mathbf{P})) = \mathbb{E}[\widetilde{\varepsilon}_a \mid \Lambda^{-1}(a, \mathbf{P}) + \widetilde{\varepsilon}_a \geq \Lambda^{-1}(j, \mathbf{P}) + \widetilde{\varepsilon}_j \text{ for any } j]$.

(vi) the expected payoff function, $W(\mathbf{P}) \equiv \sum_{j=1}^{J} P_j\, [\widetilde{v}_j + e(j, \mathbf{P})] = \mathbf{P}'[\widetilde{\mathbf{v}} + \mathbf{e}(\mathbf{P})]$.

Williams-Daly-Zachary (WDZ) Theorem. *For any choice alternative a,*

$$\frac{\partial S(\widetilde{\mathbf{v}})}{\partial \widetilde{v}_a} = \Lambda(a, \widetilde{\mathbf{v}}) \tag{A.1}$$

Proof. Define the function $m(\widetilde{\mathbf{v}} + \widetilde{\varepsilon}) = \max_{j\in\{0,1,...,J\}} [\widetilde{v}_j + \widetilde{\varepsilon}_j]$, such that $S(\mathbf{v}) = \int m(\widetilde{\mathbf{v}} + \widetilde{\varepsilon})\, dG(\varepsilon)$. It is clear that $\partial m(\widetilde{\mathbf{v}} + \widetilde{\varepsilon})/\partial \widetilde{v}_a = 1\{\widetilde{v}_a + \widetilde{\varepsilon}_a \geq \widetilde{v}_j + \widetilde{\varepsilon}_j \text{ for any } j\}$. Note that $\partial S(\widetilde{\mathbf{v}})/\partial \widetilde{v}_a$ is equal to $\int \partial m(\widetilde{\mathbf{v}} + \widetilde{\varepsilon})/\partial \widetilde{v}_a\, dG(\varepsilon)$. Therefore,

$$\frac{\partial S(\widetilde{\mathbf{v}})}{\partial \widetilde{v}_a} = \int 1\{\widetilde{v}_a + \widetilde{\varepsilon}_a \geq \widetilde{v}_j + \widetilde{\varepsilon}_j \text{ for any } j\}\, dG(\varepsilon) = \Lambda(a, \widetilde{\mathbf{v}}) \tag{A.2}$$

Lemma 1. *For any vector of value differences $\widetilde{\mathbf{v}} = (\widetilde{v}_1, \widetilde{v}_2, ..., \widetilde{v}_J)'$,*

$$\left[\frac{\partial \mathbf{\Lambda}(\widetilde{\mathbf{v}})'}{\partial \widetilde{\mathbf{v}}}\right]^{-1} \frac{\partial \mathbf{e}^{(v)}(\widetilde{\mathbf{v}})'}{\partial \widetilde{\mathbf{v}}} \mathbf{\Lambda}(\widetilde{\mathbf{v}}) = -\widetilde{\mathbf{v}} - \mathbf{e}^{(v)}(\widetilde{\mathbf{v}}) \tag{A.3}$$

Proof. Using the definitions of Social Surplus and optimal choice probability functions, we have that:

$$S(\widetilde{\mathbf{v}}) = \sum_{j=1}^{J} \Lambda(j, \widetilde{\mathbf{v}}) \left[\widetilde{v}_j + e^{(v)}(j, \widetilde{\mathbf{v}})\right] \tag{A.4}$$

Therefore,

$$\frac{\partial S(\widetilde{\mathbf{v}})}{\partial \widetilde{v}_a} = \sum_{j=1}^{J} \frac{\partial \Lambda(j, \widetilde{\mathbf{v}})}{\partial \widetilde{v}_a} \left[\widetilde{v}_j + e^{(v)}(j, \widetilde{\mathbf{v}})\right] + \Lambda(a, \widetilde{\mathbf{v}}) + \sum_{j=1}^{J} \Lambda(j, \widetilde{\mathbf{v}}) \frac{\partial e^{(v)}(j, \widetilde{\mathbf{v}})}{\partial \widetilde{v}_a} \tag{A.5}$$

In vector form,

$$\frac{\partial S\left(\widetilde{\mathbf{v}}\right)}{\partial \widetilde{\mathbf{v}}} = \frac{\partial \Lambda(\widetilde{\mathbf{v}})'}{\partial \widetilde{\mathbf{v}}}\left[\widetilde{\mathbf{v}} + \mathbf{e}^{(v)}(\widetilde{\mathbf{v}})\right] + \Lambda(\widetilde{\mathbf{v}}) + \frac{\partial \mathbf{e}^{(v)}(\widetilde{\mathbf{v}})'}{\partial \widetilde{\mathbf{v}}}\Lambda(\widetilde{\mathbf{v}}) \tag{A.6}$$

By WDZ Theorem, we have that $\partial S\left(\widetilde{\mathbf{v}}\right)/\partial \widetilde{\mathbf{v}} = \Lambda(\widetilde{\mathbf{v}})$. Plugging this result into the expression for $\partial S\left(\widetilde{\mathbf{v}}\right)/\partial \widetilde{\mathbf{v}}$ above, we have that:

$$\frac{\partial \Lambda(\widetilde{\mathbf{v}})'}{\partial \widetilde{\mathbf{v}}}\left[\widetilde{\mathbf{v}} + \mathbf{e}^{(v)}(\widetilde{\mathbf{v}})\right] + \frac{\partial \mathbf{e}^{(v)}(\widetilde{\mathbf{v}})'}{\partial \widetilde{\mathbf{v}}}\Lambda(\widetilde{\mathbf{v}}) = 0 \tag{A.7}$$

Hotz-Miller inversion theorem, together with the inverse function theorem, implies that for any value of the vector $\widetilde{\mathbf{v}}$ the matrix $\partial\Lambda(\widetilde{\mathbf{v}})'/\partial\widetilde{\mathbf{v}}$ is invertible. Therefore,

$$\left[\frac{\partial \mathbf{\Lambda}\left(\widetilde{\mathbf{v}}\right)'}{\partial \widetilde{\mathbf{v}}}\right]^{-1}\frac{\partial \mathbf{e}^{(v)}(\widetilde{\mathbf{v}})'}{\partial \widetilde{\mathbf{v}}}\mathbf{\Lambda}\left(\widetilde{\mathbf{v}}\right) = -\widetilde{\mathbf{v}} - \mathbf{e}^{(v)}(\widetilde{\mathbf{v}}) \qquad \blacksquare \tag{A.8}$$

Lemma 2. *For any vector of probabilities in the J-dimension Simplex,* $\mathbf{P} = (P_1, P_2, ..., P_J)'$, *and any choice alternative* a,

$$\frac{\partial W(\mathbf{P})}{\partial P_a} = \widetilde{v}_a - \Lambda^{-1}(a, \mathbf{P}) \tag{A.9}$$

Proof. By definition of $\Pi^P(\mathbf{P})$ as $\sum_{j=1}^J P_j\ [\widetilde{v}_j + e(j, \mathbf{P})]$ and taking into account that $e(j, \mathbf{P}) = e^{(v)}(j, \Lambda^{-1}(\mathbf{P}))$, we have that:

$$\begin{aligned}
\frac{\partial W(\mathbf{P})}{\partial P_a} &= \widetilde{v}_a + e(a, \mathbf{P}) + \sum_{j=1}^J P_j \frac{\partial e(j, \mathbf{P})}{\partial P_a} \\
&= \widetilde{v}_a + e(a, \mathbf{P}) + \sum_{j=1}^J P_j \left[\left.\frac{\partial e^{(v)}(j, \widetilde{\mathbf{v}})}{\partial \widetilde{\mathbf{v}}'}\right|_{\widetilde{\mathbf{v}}=\mathbf{\Lambda}^{-1}(\mathbf{P})}\right]\frac{\partial \Lambda^{-1}(\mathbf{P})}{\partial P_a} \qquad \text{(A.10)} \\
&= \widetilde{v}_a + e(a, \mathbf{P}) + \mathbf{P}'\left[\left.\frac{\partial \mathbf{e}^{(v)}(\widetilde{\mathbf{v}})}{\partial \widetilde{\mathbf{v}}'}\right|_{\widetilde{\mathbf{v}}=\mathbf{\Lambda}^{-1}(\mathbf{P})}\right]\frac{\partial \Lambda^{-1}(\mathbf{P})}{\partial P_a}
\end{aligned}$$

where $\mathbf{e}^{(v)}(\widetilde{\mathbf{v}}) \equiv (e^{(v)}(1, \widetilde{\mathbf{v}}),\ e^{(v)}(2, \widetilde{\mathbf{v}}),\ ...,\ e^{(v)}(J, \widetilde{\mathbf{v}}))'$. In vector form,

$$\frac{\partial W(\mathbf{P})}{\partial \mathbf{P}} = \widetilde{\mathbf{v}} + \mathbf{e}(\mathbf{P}) + \frac{\partial \Lambda^{-1}(\mathbf{P})'}{\partial \mathbf{P}}\left[\left.\frac{\partial \mathbf{e}^{(v)}(\widetilde{\mathbf{v}})'}{\partial \widetilde{\mathbf{v}}}\right|_{\widetilde{\mathbf{v}}=\mathbf{\Lambda}^{-1}(\mathbf{P})}\right]\mathbf{P} \tag{A.11}$$

Given that $\mathbf{\Lambda}^{-1}\left(\mathbf{P}\right)$ is the inverse of the mapping $\mathbf{\Lambda}\left(\widetilde{\mathbf{v}}\right)$, the inverse function theorem implies that $\partial\mathbf{\Lambda}^{-1}\left(\mathbf{P}\right)'/\partial\mathbf{P} = \left[\partial\mathbf{\Lambda}\left(\widetilde{\mathbf{v}}\right)/\partial\widetilde{\mathbf{v}}'|_{\widetilde{\mathbf{v}}=\mathbf{\Lambda}^{-1}(\mathbf{P})}\right]^{-1}$, such that:

$$\frac{\partial W(\mathbf{P})}{\partial \mathbf{P}} = \widetilde{\mathbf{v}} + \mathbf{e}(\mathbf{P}) + \left[\left.\frac{\partial \mathbf{\Lambda}\left(\widetilde{\mathbf{v}}\right)}{\partial \widetilde{\mathbf{v}}'}\right|_{\widetilde{\mathbf{v}}=\mathbf{\Lambda}^{-1}(\mathbf{P})}\right]^{-1}\left[\left.\frac{\partial \widetilde{\mathbf{e}}^{(v)}(\widetilde{\mathbf{v}})'}{\partial \widetilde{\mathbf{v}}}\right|_{\widetilde{\mathbf{v}}=\mathbf{\Lambda}^{-1}(\mathbf{P})}\right]\mathbf{P} \tag{A.12}$$

Lemma 1 implies that for $\widetilde{\mathbf{v}} = \mathbf{\Lambda}^{-1}\left(\mathbf{P}\right)$,

$$\left[\left.\frac{\partial \mathbf{\Lambda}\left(\widetilde{\mathbf{v}}\right)}{\partial \widetilde{\mathbf{v}}'}\right|_{\widetilde{\mathbf{v}}=\mathbf{\Lambda}^{-1}(\mathbf{P})}\right]^{-1}\left[\left.\frac{\partial \widetilde{\mathbf{e}}^{(v)}(\widetilde{\mathbf{v}})'}{\partial \widetilde{\mathbf{v}}}\right|_{\widetilde{\mathbf{v}}=\mathbf{\Lambda}^{-1}(\mathbf{P})}\right]\mathbf{P} = -\mathbf{\Lambda}^{-1}\left(\mathbf{P}\right) - \mathbf{e}(\mathbf{P}) \tag{A.13}$$

And solving this expression into the equation for $\partial W(\mathbf{P})/\partial \mathbf{P}$ above, we have that:

$$\frac{\partial W(\mathbf{P})}{\partial \mathbf{P}} = \widetilde{\mathbf{v}} - \mathbf{\Lambda}^{-1}(\mathbf{P}) \tag{A.14}$$

or $\partial W(\mathbf{P})/\partial P_a = \widetilde{v}_a - \Lambda^{-1}(a, \mathbf{P})$. ■

Lemma 3. *The Social Surplus function $S(\mathbf{v})$ is a weak contraction (i.e., non-expansive mapping) in the complete metric space $(\mathcal{V}, \|.\|_\infty)$. That is, for any $\mathbf{v}$, $\mathbf{w} \in \mathcal{V}$, we have that:*

$$|S(\mathbf{v}) - S(\mathbf{w})| \leq \|\mathbf{v} - \mathbf{w}\|_\infty \tag{A.15}$$

Proof: For any $\mathbf{v}$, $\mathbf{w} \in \mathcal{V}$ we have that:

$$\begin{aligned} |S(\mathbf{v}) - S(\mathbf{w})| &= \left| \int \max_j [v_j + \varepsilon_j] \, dG(\varepsilon) - \int \max_j [w_j + \varepsilon_j] \, dG(\varepsilon) \right| \\ &\leq \left| \int \max_j |v_j + \varepsilon_j - w_j - \varepsilon_j| \, dG(\varepsilon) \right| \\ &= \max_j |v_j - w_j| = \|\mathbf{v} - \mathbf{w}\|_\infty \end{aligned} \tag{A.16}$$

■

Proof of Proposition 2. Given the model assumptions, it is clear that the function $W_t(\mathbf{P}_t, \mathbf{x}_t)$ is twice continuously differentiable in $\mathbf{P}_t$. By definition, the intertemporal payoff function $W_t(\mathbf{P}_t, \mathbf{x}_t)$ is equal to $v_t(0, \mathbf{x}_t) + \sum_{j=1}^J P_t(j) \, [\widetilde{v}_t(j, \mathbf{x}_t) + e(j, \mathbf{P}_t)]$. Therefore, $W_t(\mathbf{P}_t, \mathbf{x}_t)$ belongs to the class of expected payoff functions in Lemma 2. Applying Lemma 2 to function $W_t(\mathbf{P}_t, \mathbf{x}_t)$, we have that:

$$\frac{\partial W_t(\mathbf{P}_t, \mathbf{x}_t)}{\partial \mathbf{P}_t} = \widetilde{\mathbf{v}}_t(\mathbf{x}_t) - \mathbf{\Lambda}^{-1}(\mathbf{P}_t) \tag{A.17}$$

Suppose for the moment that $W_t(\mathbf{P}_t, \mathbf{x}_t)$ is globally concave in $\mathbf{P}_t$. Under this condition, the optimal decision rule $\mathbf{P}_t^*(\mathbf{x}_t)$ is uniquely characterized by the first order condition $\partial W_t(\mathbf{P}_t^*, \mathbf{x}_t)/\partial \mathbf{P}_t = \widetilde{\mathbf{v}}_t(\mathbf{x}_t) - \mathbf{\Lambda}^{-1}(\mathbf{P}_t^*) = 0$. By definition of the inverse mapping $\mathbf{\Lambda}^{-1}(\mathbf{P}_t)$, we have that $\mathbf{P}_t^*(\mathbf{x}_t) = \mathbf{\Lambda}(\widetilde{\mathbf{v}}_t(\mathbf{x}_t))$.

Now, we prove that $W_t(\mathbf{P}_t, \mathbf{x}_t)$ is globally concave in $\mathbf{P}_t$. Using the expression above for the first derivative $\partial W_t(\mathbf{P}_t, \mathbf{x}_t)/\partial \mathbf{P}_t$, and applying the inverse function theorem, we have that:

$$\frac{\partial^2 W_t(\mathbf{P}_t, \mathbf{x}_t)}{\partial \mathbf{P}_t \partial \mathbf{P}_t'} = -\frac{\partial \mathbf{\Lambda}^{-1}(\mathbf{P}_t)}{\partial \mathbf{P}_t'} = -\left[\frac{\partial \mathbf{\Lambda}(\widetilde{\mathbf{v}})}{\partial \widetilde{\mathbf{v}}'} \bigg|_{\widetilde{\mathbf{v}} = \mathbf{\Lambda}^{-1}(\mathbf{P})} \right]^{-1} \tag{A.18}$$

The Jacobian matrix $\partial \mathbf{\Lambda}(\widetilde{\mathbf{v}})/\widetilde{\mathbf{v}}'$ is symmetric and has all principal minors positive. Theorem 29 in Chapter 1 of Magnus and Neudecker (1988) establishes that this is equivalent to being positive definite. This implies that the Hessian matrix $\partial^2 W_t(\mathbf{P}_t, \mathbf{x}_t)/\partial \mathbf{P}_t \partial \mathbf{P}_t'$ is negative definite and therefore $W_t(\mathbf{P}_t, \mathbf{x}_t)$ is globally concave in $\mathbf{P}_t$. ■

Proof of Proposition 3. The intertemporal payoff function of the probability-choice model is defined as $W_t(\mathbf{P}_t, \mathbf{x}_t) \equiv \Pi_t^P(\mathbf{P}_t, \mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1}} V_{t+1}^P(\mathbf{x}_{t+1}) \, f^P(\mathbf{x}_{t+1} | \mathbf{P}_t, \mathbf{x}_t)$. Then, for any arbitrary

vector $\mathbf{P}_t$ in the Simplex, we have that:

$$\frac{\partial W_t(\mathbf{P}_t, \mathbf{x}_t)}{\partial P_t(a)} = \frac{\partial \Pi_t^P(\mathbf{P}_t, \mathbf{x}_t)}{\partial P_t(a)} + \beta \sum_{\mathbf{x}_{t+1}} V_{t+1}^P(\mathbf{x}_{t+1}) \ \widetilde{f}(\mathbf{x}_{t+1}|a, \mathbf{x}_t) \tag{A.19}$$

where $\widetilde{f}(\mathbf{x}_{t+1}|a, \mathbf{x}_t) \equiv f_x(\mathbf{x}_{t+1}|a, \mathbf{x}_t) - f_x(\mathbf{x}_{t+1}|0, \mathbf{x}_t)$. From Proposition 2, we have that:

$$\frac{\partial W_t(\mathbf{P}_t, \mathbf{x}_t)}{\partial P_t(a)} = \widetilde{v}_t(a, \mathbf{x}_t) - \Lambda^{-1}(a, \mathbf{P}_t) \tag{A.20}$$

Combining equations (A.19) and (A.20), we have that:

$$\frac{\partial \Pi_t^P(\mathbf{P}_t, \mathbf{x}_t)}{\partial P_t(a)} = \widetilde{v}_t(a, \mathbf{x}_t) - \Lambda^{-1}(a, \mathbf{P}_t) - \beta \sum_{\mathbf{x}_{t+1}} V_{t+1}^P(\mathbf{x}_{t+1}) \ \widetilde{f}(\mathbf{x}_{t+1}|a, \mathbf{x}_t) \tag{A.21}$$

Now, by definition of the value difference $\widetilde{v}_t(a, \mathbf{x}_t) = v_t(a, \mathbf{x}_t) - v_t(0, \mathbf{x}_t)$, we have that:

$$\widetilde{v}_t(a, \mathbf{x}_t) = \pi_t(a, \mathbf{x}_t) - \pi_t(0, \mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1}} V_{t+1}^P(\mathbf{x}_{t+1}) \ \widetilde{f}(\mathbf{x}_{t+1}|a, \mathbf{x}_t) \tag{A.22}$$

Then, combining equations (A.21) and (A.22), we obtain that:

$$\frac{\partial \Pi_t^P(\mathbf{P}_t, \mathbf{x}_t)}{\partial P_t(a)} = \pi_t(a, \mathbf{x}_t) - \pi_t(0, \mathbf{x}_t) - \Lambda^{-1}(a, \mathbf{P}_t) \qquad \blacksquare \tag{A.23}$$

Proof of Proposition 4. Multiplying by $P_{t+1}(a|\mathbf{x}_{t+1})$ the Lagrange condition in (13) and summing over all the choice alternatives $a \neq 0$, we have that:

$$\beta \left[\sum_{a \neq 0} P_{t+1}(a|\mathbf{x}_{t+1}) \frac{\partial \Pi_{t+1}^P}{\partial P_{t+1}(a|\mathbf{x}_{t+1})} \right] = \sum_{\mathbf{y}_{t+2}} \lambda_{t+2}(\mathbf{y}_{t+2}) \left[\sum_{a \neq 0} P_{t+1}(a|\mathbf{x}_{t+1}) \widetilde{f}_{t+1}^y(\mathbf{y}_{t+2}|a, \mathbf{y}_{t+1}) \right] \tag{A.24}$$

Taking into account the definitions of Π_{t+1}^P and $f_{t+1}^{P,y}$ and Proposition 3, we have that equation (A.24) implies that:

$$\beta \left[\Pi_{t+1}^P - \pi_{t+1}(0, \mathbf{x}_{t+1}) + e(0, \mathbf{P}_{t+1}) \right] = \sum_{\mathbf{y}_{t+2}} \lambda_{t+2}(\mathbf{y}_{t+2}) \left[f_{t+1}^{P,y}(\mathbf{y}_{t+2}|\mathbf{P}_{t+1}, \mathbf{x}_{t+1}) - f_{t+1}^y(\mathbf{y}_{t+2}|0, \mathbf{y}_{t+1}) \right] \tag{A.25}$$

Solving equation (A.25) into the Lagrangian condition (12), we obtain:

$$\frac{\partial \Pi_t^P}{\partial P_t(a)} + \beta \sum_{\mathbf{y}_{t+1}} \left[\pi_{t+1}(0, \mathbf{x}_{t+1}) + e(0, \mathbf{P}_{t+1}) - \sum_{\mathbf{y}_{t+2}} \lambda_{t+2}(\mathbf{y}_{t+2}) \ f_{t+1}^y(\mathbf{y}_{t+2}|0, \mathbf{y}_{t+1}) \right] \widetilde{f}_t^y(\mathbf{y}_{t+1}|a) \ = \ 0 \tag{A.26}$$

Finally, given the system of Lagrange conditions $\beta \dfrac{\partial \mathbf{\Pi}_{t+1}^P}{\partial \mathbf{P}_{t+1}} = \widetilde{\mathbf{F}}_{t+1} \ \lambda_{t+2}$ and the condition of full column rank on matrix $\widetilde{\mathbf{F}}_{t+1}$, we have that $\lambda_{t+2} = \beta \ [\widetilde{\mathbf{F}}_{t+1}' \widetilde{\mathbf{F}}_{t+1}]^{-1} \ \widetilde{\mathbf{F}}_{t+1}' \dfrac{\partial \mathbf{\Pi}_{t+1}^P}{\partial \mathbf{P}_{t+1}}$. $\blacksquare$

Proof of Proposition 6. Let $\mathbf{v} = (v_0, v_1, ..., v_J)$ be a vector of choice specific values, and let $\widetilde{\mathbf{v}} = (\widetilde{v}_1, ..., \widetilde{v}_J)$ be the corresponding vector of value differences. McFadden's Social Surplus function is defined as:

$$S(\widetilde{\mathbf{v}}) \equiv \mathbb{E}_\varepsilon \left[\max_{j \in \mathcal{A}} \{\widetilde{v}_j + \varepsilon_j\} \right] \tag{A.27}$$

with $\widetilde{v}_0 = 0$. For instance, when ε's are i.i.d. extreme value type I, we have that $S(\widetilde{\mathbf{v}}) = \ln\left(\sum_{j=0}^{J} \exp\{\widetilde{v}_j\}\right)$. Let $\widetilde{\mathbf{v}} = \{\widetilde{v}(a, y, \mathbf{z}) : (a, y, \mathbf{z}) \in \mathcal{A} - \{0\} \times \mathcal{A} \times \mathcal{Z}\}$ be a vector of value differences, such that $\widetilde{v}(a, y, \mathbf{z}) \equiv v(a, y, \mathbf{z}) - v(0, y, \mathbf{z})$ and v is the conditional choice value function. The Euler Equation mapping is the fixed point mapping in the space of the vector of value differences $\widetilde{\mathbf{v}}$. For the dynamic multi-armed bandit model, this mapping is:

$$\begin{aligned} \Gamma(a, y, \mathbf{z}; \widetilde{\mathbf{v}}) &= \pi(a, y, \mathbf{z}) - \pi(0, y, \mathbf{z}) + \beta\, \mathbb{E}_{\mathbf{z}'|\mathbf{z}} [\pi(0, a, \mathbf{z}') - \pi(0, 0, \mathbf{z}')] \\ &+ \beta\, \mathbb{E}_{\mathbf{z}'|\mathbf{z}} [S(\widetilde{v}(j, a, \mathbf{z}')) - S(\widetilde{v}(j, 0, \mathbf{z}'))] \end{aligned} \tag{A.28}$$

This EE mapping has the following form:

$$\Gamma(a, y, \mathbf{z}; \widetilde{\mathbf{v}}) = c(a, y, \mathbf{z}) + \beta \sum_{\mathbf{z}' \in \mathcal{Z}} \left[S\left(\widetilde{\mathbf{v}}(a, \mathbf{z}')\right) - S\left(\widetilde{\mathbf{v}}(0, \mathbf{z}')\right) \right] \; f_z(\mathbf{z}'|\mathbf{z}) \tag{A.29}$$

where $c(a, y, \mathbf{z}) \equiv \pi(a, y, \mathbf{z}) - \pi(0, y, \mathbf{z}) + \beta \sum_{\mathbf{z}' \in \mathcal{Z}} [\pi(0, a, \mathbf{z}') - \pi(0, 0, \mathbf{z}')]\; f_z(\mathbf{z}'|\mathbf{z})$.

First, we note a property of the mapping $\Gamma(a, y, \mathbf{z}; \widetilde{\mathbf{v}})$. The term $\beta\, \mathbb{E}_{\mathbf{z}'|\mathbf{z}} [S(\widetilde{\mathbf{v}}(a, \mathbf{z}')) - S(\widetilde{\mathbf{v}}(0, \mathbf{z}'))]$ does not depend on y. This implies that $\Gamma(a, y, \mathbf{z}; \widetilde{\mathbf{v}}) - \Gamma(a, 0, \mathbf{z}; \widetilde{\mathbf{v}}) = c(a, y, \mathbf{z}) - c(a, 0, \mathbf{z}) \equiv \Delta(a, y, \mathbf{z})$, where $\Delta(a, y, \mathbf{z})$ is a primitive of the model that is known to the researcher before solving the model and does not depend on the vector of values $\widetilde{\mathbf{v}}$ where we are evaluating the mapping Γ. Consider the sequence of values $\{\widetilde{\mathbf{v}}^k : k \geq 0\}$ obtained by applying fixed-point iterations from some initial $\widetilde{\mathbf{v}}^0$, i.e., $\widetilde{\mathbf{v}}^{k+1} = \Gamma(\widetilde{\mathbf{v}}^k)$. It is clear that for any $k \geq 1$, the vector of values $\widetilde{\mathbf{v}}^k$ satisfies the property $\widetilde{v}(a, y, \mathbf{z}) - \widetilde{v}(a, 0, \mathbf{z}) = \Delta(a, y, \mathbf{z})$. Therefore, without loss of generality, we can restrict our analysis of fixed point mapping Γ to the subspace of values $\widetilde{\mathbf{v}}$ that satisfy this restriction. That is, we consider $\Gamma(\widetilde{\mathbf{v}})$ on $\mathcal{V}^R$ where,

$$\mathcal{V}^R \equiv \left\{ \widetilde{\mathbf{v}} \in \mathbb{R}^{J(J+1)|\mathcal{Z}|} : \widetilde{v}(a, y, \mathbf{z}) = \widetilde{v}(a, 0, \mathbf{z}) + \Delta(a, y, \mathbf{z}) \text{ for any } (a, y, \mathbf{z}) \right\} \tag{A.30}$$

Binary choice model. For the moment, consider the binary choice version of this model. For given state variables $(y, \mathbf{z})$, there is only one value difference, $\widetilde{v}(1, y, \mathbf{z}) \equiv v(1, y, \mathbf{z}) - v(0, y, \mathbf{z})$. For notational simplicity, we omit the argument 1 in function $\widetilde{v}(1, y, \mathbf{z})$ and represent this value difference as $\widetilde{v}(y, \mathbf{z})$. Similarly, we omit the argument $a = 1$ from the EE mapping $\Gamma(a, y, \mathbf{z}; \widetilde{\mathbf{v}})$ and represent this mapping as $\Gamma(y, \mathbf{z}; \widetilde{\mathbf{v}})$. Applying equation (A.29) to the binary choice model, we can represent the EE mapping using the following expression:

$$\Gamma(y, \mathbf{z}; \widetilde{\mathbf{v}}) = c(y, \mathbf{z}) + \beta \sum_{\mathbf{z}' \in \mathcal{Z}} \left[S\left(\widetilde{v}(1, \mathbf{z}')\right) - S\left(\widetilde{v}(0, \mathbf{z}')\right) \right] \; f_z(\mathbf{z}'|\mathbf{z}) \tag{A.31}$$

[1] McFadden's social surplus function $S(v) = \mathbb{E}_\varepsilon [\max\{\varepsilon_0 \; ; \; v + \varepsilon_1\}]$ has several properties that we exploit in our proof. (i) It is twice continuously differentiable. (ii) Its first derivative is the choice

probability for alternative $a = 1$, i.e., $S'(v) \in (0,1)$. (iii) By the Mean Value Theorem, for any pair of finite values w and v, we have that $S(w) - S(v) = S'(v^*)\ (w - v)$ where $v^* = \lambda w + (1-\lambda)v$ for a scalar $\lambda \in [0,1]$. (iv) By the Mean Value Theorem, for any pair of finite values w and v and any scalar $\Delta \in \mathbb{R}$, we have that $[S(w+\Delta) - S(v+\Delta)] - [S(w) - S(v)] = [S'(v^*_\Delta) - S'(v^*)]\ (w - v)$ where $v^*_\Delta = \lambda_\Delta(w+\Delta) + (1-\lambda_\Delta)(v+\Delta)$ and $v^* = \lambda w + (1-\lambda)v$ for scalars $\lambda_\Delta \in [0,1]$ and $\lambda \in [0,1]$. Since $S'(v^*_\Delta)$ and $S'(v^*)$ belong to $(0,1)$, it is clear that $|S'(v^*_\Delta) - S'(v^*)| < 1$.

[2] Taking into account the properties in point [1], now we can prove that mapping Γ is a contraction on the metric space $(\mathcal{V}^R, \|.\|_\infty)$. For any pair of vectors $\widetilde{\mathbf{v}}$ and $\widetilde{\mathbf{w}}$ on the space $\mathcal{V}^R$, we have that:

$$
\begin{aligned}
&\Gamma(y, \mathbf{z}; \widetilde{\mathbf{w}}) - \Gamma(y, \mathbf{z}; \widetilde{\mathbf{v}}) \\
&= \beta \sum_{\mathbf{z}' \in \mathcal{Z}} \left([S\left(\widetilde{w}(1, \mathbf{z}')\right) - S\left(\widetilde{w}(0, \mathbf{z}')\right)] - [S\left(\widetilde{v}(1, \mathbf{z}')\right) - S\left(\widetilde{v}(0, \mathbf{z}')\right)] \right)\ f_z(\mathbf{z}'|\mathbf{z}) \\
&= \beta \sum_{\mathbf{z}' \in \mathcal{Z}} \left([S\left(\widetilde{w}(1, \mathbf{z}')\right) - S\left(\widetilde{v}(1, \mathbf{z}')\right)] - [S\left(\widetilde{w}(0, \mathbf{z}')\right) - S\left(\widetilde{v}(0, \mathbf{z}')\right)] \right)\ f_z(\mathbf{z}'|\mathbf{z}) \\
&= \beta \sum_{\mathbf{z}' \in \mathcal{Z}} \left([S\left(\widetilde{w}(0, \mathbf{z}') + \Delta(\mathbf{z}')\right) - S\left(\widetilde{v}(0, \mathbf{z}') + \Delta(\mathbf{z}')\right)] - [S\left(\widetilde{w}(0, \mathbf{z}')\right) - S\left(\widetilde{v}(0, \mathbf{z}')\right)] \right)\ f_z(\mathbf{z}'|\mathbf{z}) \\
&= \beta \sum_{\mathbf{z}' \in \mathcal{Z}} \left([S'(v^*_\Delta(\mathbf{z}')) - S'(v^*(\mathbf{z}'))]\ \ [\widetilde{w}(0, \mathbf{z}') - \widetilde{v}(0, \mathbf{z}')] \right)\ f_z(\mathbf{z}'|\mathbf{z})
\end{aligned}
\tag{A.32}
$$

The third equality comes from the property of the subspace $\mathcal{V}^R$. The fourth equality comes from the application of property (iv) of function $S(.)$.

Define $\delta(\mathbf{z}') \equiv S'(v^*_\Delta(\mathbf{z}')) - S'(v^*(\mathbf{z}'))$. From properties (ii) and (iv) of function $S(.)$, it is clear that $|\delta(\mathbf{z})| < 1$ for any $\mathbf{z} \in \mathcal{Z}$. Define $\delta^* \equiv \sup_{\mathbf{z} \in \mathcal{Z}} |\delta(\mathbf{z})|$. It is clear that $0 < \delta^* < 1$. Applying these definitions to equation (A.32), we have that:

$$
\begin{aligned}
\Gamma(y, \mathbf{z}; \widetilde{\mathbf{w}}) - \Gamma(y, \mathbf{z}; \widetilde{\mathbf{v}}) &= \beta \sum_{\mathbf{z}' \in \mathcal{Z}} \delta(\mathbf{z}')\ \ [\widetilde{w}(0, \mathbf{z}') - \widetilde{v}(0, \mathbf{z}')]\ \ f_z(\mathbf{z}'|\mathbf{z}) \\
&\leq \beta \sum_{\mathbf{z}' \in \mathcal{Z}} |\delta(\mathbf{z}')\ \ [\widetilde{w}(0, \mathbf{z}') - \widetilde{v}(0, \mathbf{z}')]|\ \ f_z(\mathbf{z}'|\mathbf{z}) \\
&\leq \beta \sum_{\mathbf{z}' \in \mathcal{Z}} |\delta(\mathbf{z}')|\ \ |\widetilde{w}(0, \mathbf{z}') - \widetilde{v}(0, \mathbf{z}')|\ \ f_z(\mathbf{z}'|\mathbf{z}) \\
&\leq \beta\delta^* \sum_{\mathbf{z}' \in \mathcal{Z}} |\widetilde{w}(0, \mathbf{z}') - \widetilde{v}(0, \mathbf{z}')|\ \ f_z(\mathbf{z}'|\mathbf{z}) \\
&\leq \beta\delta^* \sup_{\mathbf{z}' \in \mathcal{Z}} |\widetilde{w}(0, \mathbf{z}') - \widetilde{v}(0, \mathbf{z}')|
\end{aligned}
\tag{A.33}
$$

The first and the second inequalities come from the application of the triangle inequality, i.e., $|A + B| \leq |A| + |B|$. The third inequality comes from $\delta^* \geq \delta(\mathbf{z})$ for any $\mathbf{z} \in \mathcal{Z}$. And the fourth inequality comes from the fact that $\sum_{\mathbf{z}' \in \mathcal{Z}} |\widetilde{w}(0, \mathbf{z}') - \widetilde{v}(0, \mathbf{z}')|\ f_z(\mathbf{z}'|\mathbf{z})$ is an expectation for the random variable $|\widetilde{w}(0, \mathbf{z}') - \widetilde{v}(0, \mathbf{z}')|$, and for any random variable its expected value is smaller or equal than the maximum value in the support of the random variable.

For any of vectors $\widetilde{\mathbf{v}}$ and $\widetilde{\mathbf{w}}$ on the space $\mathcal{V}^R$, we have that $|\widetilde{w}(1,\mathbf{z}') - \widetilde{v}(1,\mathbf{z}')| = |\widetilde{w}(0,\mathbf{z}') - \widetilde{v}(0,\mathbf{z}')|$. Therefore, we have that:

$$\|\widetilde{\mathbf{w}} - \widetilde{\mathbf{v}}\|_\infty = \sup_{y,\mathbf{z}'} \left|\widetilde{w}(y,\mathbf{z}') - \widetilde{v}(y,\mathbf{z}')\right| = \sup_{\mathbf{z}'} \left|\widetilde{w}(0,\mathbf{z}') - \widetilde{v}(0,\mathbf{z}')\right| \tag{A.34}$$

Equations (A.33) and (A.34) imply that, for any $(y,\mathbf{z})$, $\Gamma(y,\mathbf{z};\widetilde{\mathbf{w}}) - \Gamma(y,\mathbf{z};\widetilde{\mathbf{v}}) \leq \beta\delta^* \|\widetilde{\mathbf{w}} - \widetilde{\mathbf{v}}\|_\infty$. Therefore,

$$\|\Gamma(\widetilde{\mathbf{w}}) - \Gamma(\widetilde{\mathbf{v}})\|_\infty \leq \beta\delta^* \|\widetilde{\mathbf{w}} - \widetilde{\mathbf{v}}\|_\infty \tag{A.35}$$

with $\beta\delta^* \in (0,1)$. This establishes that the mapping Γ is a contraction. As shown above, $\delta^* \in (0,1)$. Therefore, the contraction constant of the mapping Γ is strictly smaller than β. Q.E.D.

Multinomial choice model. We now extend the previous proof to the multinomial model.

[1] The surplus function $S(\mathbf{v})$ has several properties that we exploit in our proof. (i) It is continuously differentiable. (ii) Its partial derivative with respect to $\widetilde{v}_j$ is equal to the choice probability of alternative j, i.e., $\partial S(\widetilde{\mathbf{v}})/\partial v_j = P_j(\widetilde{\mathbf{v}}) \in (0,1)$. (iii) By the Mean Value Theorem, for any pair of finite vectors $\mathbf{w}$ and $\mathbf{v}$, we have that $S(\mathbf{w}) - S(\mathbf{v}) = \sum_{j=1}^J P_j(\mathbf{v}^*)(w_j - v_j)$ where $\mathbf{v}^* = \lambda\mathbf{w} + (1-\lambda)\mathbf{v}$ for a scalar $\lambda \in [0,1]$. (iv) By the Mean Value Theorem, for any pair of finite vectors $\mathbf{w}$ and $\mathbf{v}$ and vector $\Delta \in \mathbb{R}^J$, we have that $[S(\mathbf{w}+\Delta) - S(\mathbf{v}+\Delta)] - [S(\mathbf{w}) - S(\mathbf{v})] = \sum_{j=1}^J [P_j(\mathbf{v}_\Delta^*) - P_j(\mathbf{v}^*)](w_j - v_j)$ where $\mathbf{v}_\Delta^* = \lambda_\Delta(\mathbf{w}+\boldsymbol{\Delta}) + (1-\lambda_\Delta)(\mathbf{v}+\boldsymbol{\Delta})$ and $\mathbf{v}^* = \lambda\mathbf{w} + (1-\lambda)\mathbf{v}$ for scalars $\lambda_\Delta \in [0,1]$ and $\lambda \in [0,1]$. Furthermore, if $\delta_j \equiv P_j(\mathbf{v}_\Delta^*) - P_j(\mathbf{v}^*)$, them for any j, $|\delta_j| < 1$. Also, for any subset $\mathcal{J} \subseteq \{1,2,...,J\}$, $\sum_{j\in\mathcal{J}} P_j(\mathbf{v}) = P_\mathcal{J}(\mathbf{v}) \in (0,1)$ such that $\left|\sum_{j\in\mathcal{J}}\delta_j\right| = |P_\mathcal{J}(\mathbf{v}_\Delta^*) - P_\mathcal{J}(\mathbf{v}^*)| < 1$.

[2] The following Lemma is applied in our proof below.

LEMMA. Let $D = \sum_{j=1}^J \delta_j\ x_j$ with: (a) $|\delta_j| < 1$ for any j; (b) for any subset $\mathcal{J} \subseteq \{1,2,...,J\}$, $\left|\sum_{j\in\mathcal{J}}\delta_j\right| < 1$; and (c) $x_j \geq 0$ for any j. Then, $|D| \leq \delta^* \sup_j |x_j|$ with $\delta^* \in (0,1)$.
Proof of Lemma: Define the sets of indexes: $\mathcal{J}_+ = \{j : \delta_j \geq 0\}$ and $\mathcal{J}_- = \{j : \delta_j < 0\}$. Given that $x_j \geq 0$ for any j, we have that:

$$|D| = \left|\sum_{j\in\mathcal{J}_+} \delta_j x_j + \sum_{j\in\mathcal{J}_-} \delta_j x_j\right| \leq \sup\left\{\left|\sum_{j\in\mathcal{J}_+} \delta_j x_j\right| \ , \ \left|\sum_{j\in\mathcal{J}_-} \delta_j x_j\right|\right\}$$

Suppose that $\left|\sum_{j\in\mathcal{J}_+}\delta_j x_j\right| \geq \left|\sum_{j\in\mathcal{J}_-}\delta_j x_j\right|$. Define $\delta_+^* \equiv \sum_{j\in\mathcal{J}_+}\delta_j$ and $p_j \equiv \delta_j/\delta_+^*$ such that $\delta_+^* \in (0,1)$, $p_j \in (0,1)$, and $\sum_{j\in\mathcal{J}_+} p_j = 1$. Then, we have:

$$|D| \leq \left|\sum_{j\in\mathcal{J}_+} \delta_j x_j\right| = \left|\delta_+^* \sum_{j\in\mathcal{J}_+} p_j\ x_j\right| \leq \delta_+^* \sum_{j\in\mathcal{J}_+} p_j\ |x_j| \leq \delta_+^* \sup_{j\in\mathcal{J}_+} |x_j| \leq \delta_+^* \sup_j |x_j|$$

Alternatively, suppose that $\left|\sum_{j\in\mathcal{J}_+}\delta_j x_j\right| < \left|\sum_{j\in\mathcal{J}_-}\delta_j x_j\right|$. Define $\delta_-^* \equiv -\sum_{j\in\mathcal{J}_-}\delta_j$ and $p_j \equiv -\delta_j/\delta_-^*$ such that $\delta_-^* \in (0,1)$, $p_j \in (0,1)$, and $\sum_{j\in\mathcal{J}_-} p_j = 1$. Then, we have:

$$|D| \leq \left|\sum_{j\in\mathcal{J}_-} \delta_j x_j\right| = \left|\delta_-^* \sum_{j\in\mathcal{J}_-} p_j\ x_j\right| \leq \delta_-^* \sum_{j\in\mathcal{J}_-} p_j\ |x_j| \leq \delta_-^* \sup_{j\in\mathcal{J}_-} |x_j| \leq \delta_-^* \sup_j |x_j| \quad \blacksquare$$

[3] Taking into account the properties in points [1] and [2] above, now we can prove that mapping Γ is a contraction on the metric space $(\mathcal{V}^R, \|.\|_\infty)$. For any pair of vectors $\widetilde{\mathbf{v}}$ and $\widetilde{\mathbf{w}}$ on the space $\mathcal{V}^R$, we have that:

$$\begin{aligned}
&\Gamma(a,y,\mathbf{z};\widetilde{\mathbf{w}}) - \Gamma(a,y,\mathbf{z};\widetilde{\mathbf{v}}) \\
&= \beta \sum_{\mathbf{z}'\in\mathcal{Z}} \left([S\left(\widetilde{\mathbf{w}}(a,\mathbf{z}')\right) - S\left(\widetilde{\mathbf{w}}(0,\mathbf{z}')\right)] - [S\left(\widetilde{\mathbf{v}}(a,\mathbf{z}')\right) - S\left(\widetilde{\mathbf{v}}(0,\mathbf{z}')\right)] \right) \; f_z(\mathbf{z}'|\mathbf{z}) \\
&= \beta \sum_{\mathbf{z}'\in\mathcal{Z}} \left([S\left(\widetilde{\mathbf{w}}(a,\mathbf{z}')\right) - S\left(\widetilde{\mathbf{v}}(a,\mathbf{z}')\right)] - [S\left(\widetilde{\mathbf{w}}(0,\mathbf{z}')\right) - S\left(\widetilde{\mathbf{v}}(0,\mathbf{z}')\right)] \right) \; f_z(\mathbf{z}'|\mathbf{z}) \\
&= \beta \sum_{\mathbf{z}'\in\mathcal{Z}} \left([S\left(\widetilde{\mathbf{w}}(a,\mathbf{z}') + \mathbf{\Delta}(a,\mathbf{z}')\right) - S\left(\widetilde{\mathbf{v}}(a,\mathbf{z}') + \mathbf{\Delta}(a,\mathbf{z}')\right)] - [S\left(\widetilde{\mathbf{w}}(0,\mathbf{z}')\right) - S\left(\widetilde{\mathbf{v}}(0,\mathbf{z}')\right)] \right) f_z(\mathbf{z}'|\mathbf{z}) \\
&= \beta \sum_{\mathbf{z}'\in\mathcal{Z}} \sum_{j=1}^{J} [P_j(\mathbf{v}^*_\Delta(\mathbf{z}')) - P_j(\mathbf{v}^*(\mathbf{z}'))] \;\; [\widetilde{w}(j,0,\mathbf{z}') - \widetilde{v}(j,0,\mathbf{z}')] \;\; f_z(\mathbf{z}'|\mathbf{z})
\end{aligned} \tag{A.36}$$

The third equality comes from the property of the subspace $\mathcal{V}^R$. The fourth equality comes from the application of property (iv) of function $S(.)$, where $\mathbf{v}^*_\Delta(\mathbf{z}) = \lambda_\Delta(\mathbf{z}) \; (\widetilde{\mathbf{w}}(0,\mathbf{z}) + \mathbf{\Delta}(\mathbf{z})) + (1 - \lambda_\Delta(\mathbf{z}))(\widetilde{\mathbf{v}}(0,\mathbf{z}) + \mathbf{\Delta}(\mathbf{z}))$ and $\mathbf{v}^*(\mathbf{z}) = \lambda(\mathbf{z}) \; \widetilde{\mathbf{w}}(0,\mathbf{z}) + (1 - \lambda(\mathbf{z})) \; \widetilde{\mathbf{v}}(0,\mathbf{z})$ for scalars $\lambda_\Delta(\mathbf{z}) \in [0,1]$ and $\lambda(\mathbf{z}) \in [0,1]$.

Define $\delta_j(\mathbf{z}) \equiv P_j(\mathbf{v}^*_\Delta(\mathbf{z})) - P_j(\mathbf{v}^*(\mathbf{z}))$ and $D(\mathbf{z}) \equiv \sum_{j=1}^{J} \delta_j(\mathbf{z}) \; [\widetilde{w}(j,0,\mathbf{z}) - \widetilde{v}(j,0,\mathbf{z})]$ such that

$$\Gamma(a,y,\mathbf{z};\widetilde{\mathbf{w}}) - \Gamma(a,y,\mathbf{z};\widetilde{\mathbf{v}}) \;\; = \;\; \beta \sum_{\mathbf{z}'\in\mathcal{Z}} D(\mathbf{z}') \; f_z(\mathbf{z}'|\mathbf{z}) \tag{A.37}$$

It is clear that: (a) $|\delta_j(\mathbf{z})| < 1$ for any j; (b) for any subset $\mathcal{J} \subseteq \{1,2,...,J\}$, $\left|\sum_{j\in\mathcal{J}} \delta_j(\mathbf{z})\right| < 1$; and suppose that (c) $\widetilde{w}(j,0,\mathbf{z}) - \widetilde{v}(j,0,\mathbf{z}) \geq 0$ for any j. Then, the Lemma in point [3] implies that $|D(\mathbf{z})| \leq \delta^*(\mathbf{z}) \; \sup_j |\widetilde{w}(j,0,\mathbf{z}) - \widetilde{v}(j,0,\mathbf{z})|$ with $\delta^*(\mathbf{z}) \in (0,1)$. Therefore, if $\delta^* \equiv \sup_{\mathbf{z}\in\mathcal{Z}} \delta^*(\mathbf{z})$, we have that:

$$\begin{aligned}
\Gamma(a,y,\mathbf{z};\widetilde{\mathbf{w}}) - \Gamma(a,y,\mathbf{z};\widetilde{\mathbf{v}}) &\leq \beta \sum_{\mathbf{z}'\in\mathcal{Z}} \delta^*(\mathbf{z}') \; \sup_j |\widetilde{w}(j,0,\mathbf{z}') - \widetilde{v}(j,0,\mathbf{z}')| \; f_z(\mathbf{z}'|\mathbf{z}) \\
&\leq \beta\delta^* \sum_{\mathbf{z}'\in\mathcal{Z}} \sup_j |\widetilde{w}(j,0,\mathbf{z}') - \widetilde{v}(j,0,\mathbf{z}')| \; f_z(\mathbf{z}'|\mathbf{z}) \\
&\leq \beta\delta^* \sup_{j,\mathbf{z}} |\widetilde{w}(j,0,\mathbf{z}') - \widetilde{v}(j,0,\mathbf{z}')|
\end{aligned} \tag{A.38}$$

For any of vectors $\widetilde{\mathbf{v}}$ and $\widetilde{\mathbf{w}}$ on the space $\mathcal{V}^R$, we have that $|\widetilde{w}(j,y,\mathbf{z}') - \widetilde{v}(j,y,\mathbf{z}')| = |\widetilde{w}(j,0,\mathbf{z}') - \widetilde{v}(j,0,\mathbf{z}')|$. Therefore, we have that:

$$\|\widetilde{\mathbf{w}} - \widetilde{\mathbf{v}}\|_\infty = \sup_{j,y,\mathbf{z}'} \left|\widetilde{w}(j,y,\mathbf{z}') - \widetilde{v}(j,y,\mathbf{z}')\right| = \sup_{j,\mathbf{z}'} \left|\widetilde{w}(j,0,\mathbf{z}') - \widetilde{v}(j,0,\mathbf{z}')\right| \tag{A.39}$$

Equations (A.38) and (A.39) imply that, for any $(a,y,\mathbf{z})$, $\Gamma(a,y,\mathbf{z};\widetilde{\mathbf{w}}) - \Gamma(a,y,\mathbf{z};\widetilde{\mathbf{v}}) \leq \beta\delta^* \|\widetilde{\mathbf{w}} - \widetilde{\mathbf{v}}\|_\infty$. Therefore,

$$\|\Gamma(\widetilde{\mathbf{w}}) - \Gamma(\widetilde{\mathbf{v}})\|_\infty \leq \beta\delta^* \|\widetilde{\mathbf{w}} - \widetilde{\mathbf{v}}\|_\infty \tag{A.40}$$

with $\beta\delta^* \in (0,1)$. This establishes that the mapping Γ is a contraction. As shown above, $\delta^* \in (0,1)$. Therefore, the contraction constant of the mapping Γ is strictly smaller than β. Q.E.D.

Proof of Proposition 7.

(A) $\Gamma^{(N)}_{EE,v}(\widetilde{\mathbf{v}},\theta)$ is a contraction. This proof is a straightforward extension of the proof of proposition 6. For a given sample with N cross sectional observations, denote the space in which the value differences live to be $\mathcal{V}_N$. Additionally, let $\widehat{f}_N(\mathbf{x}'|a,\mathbf{x})$ be the estimated transition probabilities for a given action a. For any $\tilde{v}$ and $\tilde{w}$ in $\mathcal{V}_N$, using the same inequalities as in (A.24) but replacing $\mathbf{\Gamma}_{EE,v}$ with $\mathbf{\Gamma}^{(N)}_{EE,v}$, and f with f_N, we have that:

$$\begin{aligned}
\left\|\mathbf{\Gamma}^{(N)}_{EE,v}(\widetilde{\mathbf{v}},\theta) - \mathbf{\Gamma}^{(N)}_{EE,v}(\widetilde{\mathbf{w}},\theta)\right\|_\infty &= \max_{a,\mathbf{x}} \left|\Gamma^{(N)}_{EE,v}(a,\mathbf{x},\widetilde{\mathbf{v}}) - \Gamma^{(N)}_{EE,v}(a,\mathbf{x},\widetilde{\mathbf{w}})\right| \\
&= \beta \max_{a,\mathbf{x}} \left|\sum_{\mathbf{x}'} \left[S\left(\widetilde{\mathbf{w}}(\mathbf{x}')\right) - S\left(\widetilde{\mathbf{v}}(\mathbf{x}')\right)\right] \left[f_N(\mathbf{x}'|a,\mathbf{x}) - f_N(\mathbf{x}'|0,\mathbf{x})\right]\right| \\
&\leq \beta \left\|\widetilde{\mathbf{v}} - \widetilde{\mathbf{w}}\right\|_\infty \qquad \blacksquare
\end{aligned} \tag{A.42}$$

(B) $\Gamma^{(N)}_{EE,v}(\widetilde{\mathbf{v}},\theta)$ converges uniformly to $\mathbf{\Gamma}_{EE,v}(\widetilde{\mathbf{v}},\theta)$. We need to show that, for any value of $(a,\mathbf{x})$, $\sup_{\widetilde{\mathbf{v}},\theta} \left|\mathbf{\Gamma}^{(N)}_{EE,v}(a,\mathbf{x},\widetilde{\mathbf{v}},\theta) - \mathbf{\Gamma}_{EE,v}(a,\mathbf{x},\widetilde{\mathbf{v}},\theta)\right| \longrightarrow_p 0$. Or equivalently, for any $\varepsilon > 0$ and any $\delta > 0$, there is an integer $N_0(\varepsilon,\delta)$ such that for every value of $(\widetilde{\mathbf{v}},\theta)$ and any $N > N_0(\varepsilon,\delta)$ we have that,

$$\Pr\left(\left|\mathbf{\Gamma}^{(N)}_{EE,v}(a,\mathbf{x},\widetilde{\mathbf{v}},\theta) - \mathbf{\Gamma}_{EE,v}(a,\mathbf{x},\widetilde{\mathbf{v}},\theta)\right| > \varepsilon\right) < \delta \tag{A.43}$$

For the rest of this proof and for the sake of notational simplicity, we omit the arguments $(a,\mathbf{x})$ and use γ to represent the vector of parameters $(\widetilde{\mathbf{v}},\theta)$. By definition, we have that $\mathbf{\Gamma}_{EE,v}(\gamma) = \sum_{\mathbf{z}'\in\mathcal{Z}} h(\mathbf{z}',\gamma)\ f(\mathbf{z}')$ and $\mathbf{\Gamma}^{(N)}_{EE,v}(\gamma) = \sum_{\mathbf{z}'\in\mathcal{Z}_N} h(\mathbf{z}',\gamma)\ f_{(N)}(\mathbf{z}')$, where: (a) $h(\mathbf{z}',\gamma)$ is a function of payoffs and choice probabilities that comes from the Euler equation; (b) $h(\mathbf{z}',\gamma)$ is a bounded function such that $h^* \equiv \sup_{\mathbf{z}',\gamma} |h(\mathbf{z}',\gamma)| < \infty$; (c) f is the true population distribution (transition probability) of $\mathbf{z}'$, and $f_{(N)}$ is the empirical distribution based on the sample; (d) for convenience, and without loss of generality, we consider that $f_{(N)}(\mathbf{z}') = 0$ for values $\mathbf{z}'$ outside the sample set $\mathcal{Z}_N$, such that we can write $\mathbf{\Gamma}^{(N)}_{EE,v}(\gamma) = \sum_{\mathbf{z}'\in\mathcal{Z}} h(\mathbf{z}',\gamma)\ f_{(N)}(\mathbf{z}')$; and (e) $f_{(N)}$ is a uniformly consistent estimator of f and this implies that, for any $\varepsilon_f > 0$ and any $\delta_f > 0$, there is an integer $N_f(\varepsilon_f,\delta_f)$ such that for any $N > N_f(\varepsilon_f,\delta_f)$ we have that,

$$\Pr\left(\sup_{\mathbf{z}'\in\mathcal{Z}} \left|f_{(N)}(\mathbf{z}') - f(\mathbf{z}')\right| > \varepsilon_f\right) < \delta_f \tag{A.44}$$

Given points (a) to (e), we now prove uniform convergence of $\mathbf{\Gamma}^{(N)}_{EE,v}$ to $\mathbf{\Gamma}_{EE,v}$. Note that for any

N and γ,

$$\begin{aligned}\left|\mathbf{\Gamma}^{(N)}_{EE,v}(\gamma) - \mathbf{\Gamma}_{EE,v}(\gamma)\right| &= \left|\sum_{\mathbf{z}'\in\mathcal{Z}} \left[f_{(N)}(\mathbf{z}') - f(\mathbf{z}')\right] h(\mathbf{z}',\gamma)\right| \\ &\leq h^* \sum_{\mathbf{z}'\in\mathcal{Z}} \left|f_{(N)}(\mathbf{z}') - f(\mathbf{z}')\right| \\ &\leq h^* \sup_{\mathbf{z}'\in\mathcal{Z}} \left|f_{(N)}(\mathbf{z}') - f(\mathbf{z}')\right|\end{aligned} \tag{A.45}$$

This implies that, for any $\varepsilon > 0$,

$$\begin{aligned}\Pr\left(\left|\mathbf{\Gamma}^{(N)}_{EE,v}(\gamma) - \mathbf{\Gamma}_{EE,v}(\gamma)\right| > \varepsilon\right) &\leq \Pr\left(h^* \sup_{\mathbf{z}'\in\mathcal{Z}} \left|f_{(N)}(\mathbf{z}') - f(\mathbf{z}')\right| > \varepsilon\right) \\ &= \Pr\left(\sup_{\mathbf{z}'\in\mathcal{Z}} \left|f_{(N)}(\mathbf{z}') - f(\mathbf{z}')\right| > \frac{\varepsilon}{h^*}\right)\end{aligned} \tag{A.46}$$

Therefore, for any $\varepsilon > 0$ and $\delta > 0$, we can fix $\varepsilon_f = \varepsilon/h^*$, $\delta_f = \delta$, and $N_0(\varepsilon,\delta) = N_f(\varepsilon_f,\delta_f)$, such that for any $N > N_0(\varepsilon,\delta)$ we have that $\Pr\left(\left|\mathbf{\Gamma}^{(N)}_{EE,v}(\gamma) - \mathbf{\Gamma}_{EE,v}(\gamma)\right| > \varepsilon\right) < \delta$. ■

Proof of Proposition 8.

(A) Consistency. For notational simplicity, we omit now θ^* as an argument in functions $\mathbf{\Gamma}_{EE,v}(\widetilde{v},\theta^*)$ and $\mathbf{\Gamma}^{(N)}_{EE,v}(\widetilde{v},\theta^*)$. The true $\widetilde{v}^*$ is defined as the unique fixed point $\widetilde{v}^* = \mathbf{\Gamma}_{EE,v}(\widetilde{v}^*)$, and $\widetilde{v}^*_N$ is defined as the unique fixed point $\widetilde{v}^*_N = \mathbf{\Gamma}^{(N)}_{EE,v}(\widetilde{v}^*_N)$. Given that, (a) $\mathbf{\Gamma}_{EE,v}(.)$ is a continuous function, and (b) $\mathbf{\Gamma}^{(N)}_{EE,v}$ converges uniformly in probability to $\mathbf{\Gamma}_{EE,v}$, we have by Slutsky's that $\widetilde{v}^*_N$ converges in probability to $\widetilde{v}^*$.

(B) Asymptotic Normality. By the fixed point conditions that define the value vectors $\widetilde{v}^*_N$ and $\widetilde{v}^*$, we have that:

$$\begin{aligned}\widetilde{v}^*_N - \widetilde{v}^* &= \mathbf{\Gamma}^{(N)}_{EE,v}(\widetilde{v}^*_N) - \mathbf{\Gamma}_{EE,v}(\widetilde{v}^*) \\ &= \sum_{\mathbf{z}\in\mathcal{Z}} h(\mathbf{z},\widetilde{v}^*_N)\ f_{(N)}(\mathbf{z}) - \sum_{\mathbf{z}\in\mathcal{Z}} h(\mathbf{z},\widetilde{v}^*)\ f(\mathbf{z})\end{aligned} \tag{A.47}$$

Applying the Mean Value Theorem to the vector function $h(\mathbf{z},\widetilde{v}^*_N)$ at $\widetilde{v} = \widetilde{v}^*$, and using the consistency of $\widetilde{v}^*_N$ (i.e., $\widetilde{v}^*_N = \widetilde{v}^* + o_p(1)$), we have that:

$$h(\mathbf{z},\widetilde{v}^*_N) = h(\mathbf{z},\widetilde{v}^*) + \frac{\partial h(\mathbf{z},\widetilde{v}^*)}{\partial \widetilde{v}'}\left(\widetilde{v}^*_N - \widetilde{v}^*\right) + o_p(1) \tag{A.48}$$

Solving this expression into (A.47) and using the consistency of $f_{(N)}$ (i.e., $f_{(N)}(\mathbf{z}) = f_{(N)}(\mathbf{z}) + o_p(1)$), we obtain:

$$\widetilde{v}^*_N - \widetilde{v}^* = \sum_{\mathbf{z}\in\mathcal{Z}} h(\mathbf{z},\widetilde{v}^*)\ \left(f_{(N)}(\mathbf{z}) - f(\mathbf{z})\right) + \left[\sum_{\mathbf{z}\in\mathcal{Z}} f(\mathbf{z})\ \frac{\partial h(\mathbf{z},\widetilde{v}^*)}{\partial \widetilde{v}'}\right]\left(\widetilde{v}^*_N - \widetilde{v}^*\right) + o_p(1) \tag{A.49}$$

Solving for $\widetilde{v}^*_N - \widetilde{v}^*$,

$$\widetilde{v}^*_N - \widetilde{v}^* = \left[\mathbf{I} - \mathbf{D}\right]^{-1}\mathbf{H}\ \left(\mathbf{f}_{(N)} - \mathbf{f}\right) + o_p(1) \tag{A.50}$$

where: $\mathbf{I}$ is the identity matrix; $\mathbf{D}$ is the matrix $\sum_{\mathbf{z}\in\mathcal{Z}} f(\mathbf{z})\ \partial h(\mathbf{z},\widetilde{v}^*)/\partial\widetilde{v}'$; $\mathbf{f}_{(N)}$ and $\mathbf{f}$ are the column vectors that contain the probabilities $f_{(N)}(\mathbf{z})$ and $f(\mathbf{z})$, respectively, for every value of $\mathbf{z}$; and $\mathbf{H}$ is the matrix with columns $[h(\mathbf{z}^{(1)},\widetilde{v}^*),\ h(\mathbf{z}^{(2)},\widetilde{v}^*),\ ...,\ h(\mathbf{z}^{(|\mathcal{Z}|)},\widetilde{v}^*)]$. Under mild regularity conditions, a standard Central Limit Theorem implies that the frequency estimator $\mathbf{f}_{(N)}$ is such that $\sqrt{N}\left(\mathbf{f}_{(N)}-\mathbf{f}\right)$ converges in distribution to $N(0,\mathbf{V}_f)$. Then, applying the Mann-Wald Theorem we have that:

$$\sqrt{N}\left(\mathbf{f}_{(N)}-\mathbf{f}\right) \to_d N\left(0,\ \left[\mathbf{I}-\mathbf{D}\right]^{-1}\mathbf{H}\ \mathbf{V}_f\ \mathbf{H}'\ \left[\mathbf{I}-\mathbf{D}'\right]^{-1}\right) \qquad \blacksquare \tag{A.51}$$

Euler Equation-Probability mapping. The following Proposition 9 shows that the system of Euler equations defines a fixed point mapping in the space of conditional choice probabilities. We also show below with an example that this fixed point mapping is not always a contraction.

PROPOSITION 9 [Euler Equation-Probability mapping]. For the stationary probability-choice problem, the system of Euler equations in Proposition 4 implies the following fixed point mapping in the space of conditional choice probabilities. Given a vector $\mathbf{P}\in[0,1]^{J|\mathcal{X}|}$*, the Euler Equation-Probability (EE-p) mapping is defined as* $\mathbf{\Gamma}_{EE-p}(\mathbf{P})\equiv\{\Gamma_{EE-p}(a,\mathbf{x}_t,\mathbf{P}):(a,\mathbf{x}_t)\in(\mathcal{A}-\{0\})\times\mathcal{X}\}$*, where:* $\Gamma_{EE-p}(a,\mathbf{x}_t,\mathbf{P})=\Lambda\left(a,\widetilde{\mathbf{v}}^{EE}(\mathbf{x}_t,\mathbf{P})\right)$*; $\Lambda(.)$ is the Optimal Choice Probability mapping;* $\widetilde{\mathbf{v}}^{EE}(\mathbf{x}_t,\mathbf{P})\equiv\left\{\widetilde{v}^{EE}(a,\mathbf{x}_t,\mathbf{P}):a\in(\mathcal{A}-\{0\})\right\}$*, with*

$$\widetilde{v}^{EE}(a,\mathbf{x}_t,\mathbf{P})\equiv\widetilde{\pi}\left(a,\mathbf{x}_t\right)+\beta\sum_{\mathbf{x}_{t+1}}\left[\pi_{t+1}\left(0,\mathbf{x}_{t+1}\right)+e(0,\mathbf{P})-\overline{\lambda^*}(\mathbf{y}_{t+1},\mathbf{P})\ \right]\widetilde{f}(\mathbf{x}_{t+1}|a,\mathbf{x}_t), \tag{A.52}$$

$\overline{\lambda^*}(\mathbf{y}_{t+1},\mathbf{P})\equiv\sum_{\mathbf{y}_{t+2}}\lambda^*(\mathbf{y}_{t+2},\mathbf{P})f^y(\mathbf{y}_{t+2}|0,\mathbf{y}_{t+1})$*, and* $\widetilde{f}(\mathbf{x}_{t+1}|a,\mathbf{x}_t)\equiv\widetilde{f}^y(\mathbf{y}_{t+1}|a,\mathbf{y}_t)f^z(\mathbf{z}_{t+1}|\mathbf{z}_t)$. ■

Proof: It follows from the Euler equation in Proposition 4, the application of Proposition 3 that implies $\partial\Pi_t^P/\partial P_t(a)=\widetilde{\pi}_t\left(a,\mathbf{x}_t\right)-\Lambda^{-1}\left(a,\mathbf{P}_t\right)$, and the invertibility of the Optimal Choice Probability mapping Λ.

EXAMPLE: Consider the Multi-armed bandit problem in Example 1 with extreme value distribution of the unobservables. For this model, the OCP mapping $\Lambda\left(\widetilde{\mathbf{v}}_t\right)$ has the form of multinomial logit probabilities. Then, we have that:

$$\Gamma_{EE-p}(a,\mathbf{x}_t,\mathbf{P})=\frac{\exp\left\{\widetilde{\pi}\left(a,\mathbf{x}_t\right)+\beta\ \mathbb{E}_{\mathbf{z}_{t+1}|\mathbf{z}_t}\left[\pi\left(0,a,\mathbf{z}_{t+1}\right)-\pi\left(0,0,\mathbf{z}_{t+1}\right)-\ln\left(\frac{P(0|a,\mathbf{z}_{t+1})}{P(0|0,\mathbf{z}_{t+1})}\right)\right]\right\}}{\sum_{j=0}^{J}\exp\left\{\widetilde{\pi}\left(j,\mathbf{x}_t\right)+\beta\ \mathbb{E}_{\mathbf{z}_{t+1}|\mathbf{z}_t}\left[\pi\left(0,j,\mathbf{z}_{t+1}\right)-\pi\left(0,0,\mathbf{z}_{t+1}\right)-\ln\left(\frac{P(0|j,\mathbf{z}_{t+1})}{P(0|0,\mathbf{z}_{t+1})}\right)\right]\right\}} \tag{A.53}$$

where $\mathbb{E}_{\mathbf{z}_{t+1}|\mathbf{z}_t}[.]$ represents the expectation over the distribution $f^z(\mathbf{z}_{t+1}|\mathbf{z}_t)$. ■

By definition, $\Gamma_{EE,p}(\mathbf{P})$ is a fixed point mapping in the probability space $[0,1]^{|\mathcal{X}|J}$. In the stationary model, the vector of optimal choice probabilities can be described as a fixed point of this mapping: $\mathbf{P}^*=\Gamma_{EE,p}(\mathbf{P}^*)$. We can use this mapping to solve the DP problem. Unfortunately, the EE-P mapping is not a contraction such that an algorithm that applies fixed-point iterations in this mapping may not converge to the solution of the DP problem. The following example illustrates this issue.

EXAMPLE (Entry-Exit model). Consider a binary-choice version of the multi-armed bandit model. The model is stationary and so we omit the time subindex. There are only two free conditional choice probabilities: $\mathbf{P} = (P(1|0), P(1|1))$, where $P(1|0)$ is the probability of moving from state 0 to state 1 (i.e., market entry), and $P(1|1)$ is the probability of staying in state 1 (i.e., staying active in the market). In this example we use the more compact notation $P_0 \equiv P(1|0)$ and $P_1 \equiv P(1|1)$. Euler equation (**??**) takes the following form: $\pi(1,x) - \ln P_x + \beta\left[\pi(0,1) - \ln(1-P_1)\right] = \pi(0,x) - \ln(1-P_x) + \beta\left[\pi(0,0) - \ln(1-P_0)\right]$. The EE-p mapping is $\Gamma_{EE,p}(\mathbf{P}) = \{\Gamma_{EE,p}(x,\mathbf{P}) : x = 0,1\}$ with

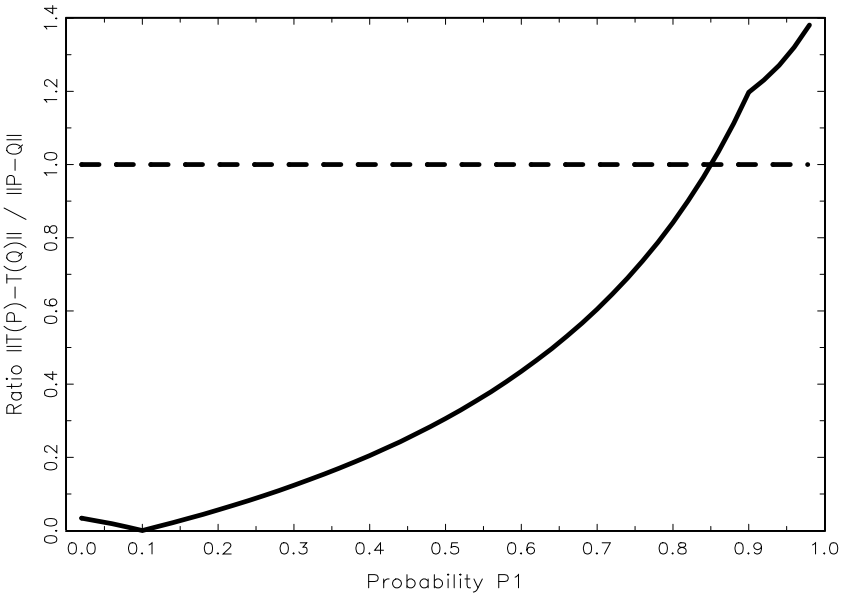
$$\Gamma_{EE,p}(x,\mathbf{P}) = \frac{\exp\left\{c\left(x\right) + \beta\left[-\ln(1-P_1) + \ln(1-P_0)\right]\right\}}{1 + \exp\left\{c\left(x\right) + \beta\left[-\ln(1-P_1) + \ln(1-P_0)\right]\right\}} \tag{A.54}$$

and $c\left(x\right) \equiv \left[\pi(1,x) - \pi(0,x)\right] + \beta[\pi(0,1) - \pi(0,0)]$. Given two vectors of CCPs in $[0,1]^2$, say $\mathbf{P}$ and $\mathbf{Q}$, consider the ratio $r(\mathbf{P},\mathbf{Q}) = \left\|\mathbf{\Gamma}_{EE,p}(\mathbf{P}) - \mathbf{\Gamma}_{EE,p}(\mathbf{Q})\right\| / \left\|\mathbf{P} - \mathbf{Q}\right\|$. The mapping $\Gamma_{EE,p}$ is a contraction if and only if the ratio $r(\mathbf{P},\mathbf{Q})$ is strictly smaller than one for any $\mathbf{P}$ and $\mathbf{Q}$ in $[0,1]^2$. Figure 1 presents an example where this condition is not satisfied. This figure presents the ratio $r(\mathbf{P},\mathbf{Q})$ as a function of P_1 for fixed $P_0 = 0.1$ and $\mathbf{Q} = (0.5, 0.5)$, in a model with $c\left(0\right) = -1$, $c\left(1\right) = 1$, and $\beta = 0.95$. We see that the ratio becomes strictly greater than one for values of P_1 close to 1. Therefore, in this example the mapping $\Gamma_{EE,p}(\mathbf{P})$ is not a contraction. Note also that,

$$\frac{\partial \Gamma_{EE,p}(1,\mathbf{P})}{\partial P_1} = \beta \frac{\Gamma_{EE,p}(1,\mathbf{P})\ \left[1 - \Gamma_{EE,p}(1,\mathbf{P})\right]}{1 - P_1} \tag{A.55}$$

The value of this derivative is greater than one for values of P_1 that are close enough to one. It is also possible to show that the spectral radius of the Jacobian matrix $\partial \Gamma_{EE,p}(\mathbf{P}) / \partial \mathbf{P}'$ is equal to $\beta \left| \frac{\Gamma_{EE,p}(1,\mathbf{P})\left[1-\Gamma_{EE,p}(1,\mathbf{P})\right]}{1-P_1} - \frac{\Gamma_{EE,p}(0,\mathbf{P})\left[1-\Gamma_{EE,p}(0,\mathbf{P})\right]}{1-P_0} \right|$, and it is greater than one for values of P_1 close enough to one. ■

Figure 1. Ratio $\|\mathbf{\Gamma}_{EE,p}(\mathbf{P}) - \mathbf{\Gamma}_{EE,p}(\mathbf{Q})\| \,/\, \|\mathbf{P} - \mathbf{Q}\|$ as a function of P_1



References

[1] Ackerberg, D., K. Caves, and G. Frazer (2015): "Identification properties of recent production function estimators," *Econometrica*, 83(6), 2411-2451.

[2] Ackerberg, D., X. Chen, and J. Hahn (2012): "A Practical Asymptotic Variance Estimator for Two-Step Semiparametric Estimators," *The Review of Economics and Statistics*, 94(2): 481-498.

[3] Aguirregabiria, V. (1997): "Estimation of Dynamic Programming Models with Censored Decision Variables," *Investigaciones Economicas*, 21, 167-208.

[4] Aguirregabiria, V., and A. Magesan (2013): "Euler Equations for the Estimation of Dynamic Discrete Choice Structural Models," *Advances in Econometrics*, 31, 3-44.

[5] Aguirregabiria, V. and P. Mira (2002): "Swapping the nested fixed point algorithm: A class of estimators for discrete Markov decision models," *Econometrica*, 70, 1519-1543.

[6] Arcidiacono, P. and R. Miller (2011): "CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," *Econometrica*, 79, 1823-1867.

[7] Arcidiacono, P. and R. Miller (2015): "Identifying Dynamic Discrete Choice Models off Short Panels," manuscript. Carnegie Mellon University.

[8] Arcidiacono, P. and R. Miller (2016): "Nonstationary Dynamic Models with Finite Dependence," manuscript. Carnegie Mellon University.

[9] Bertsekas, D. (2011): "Approximate policy iteration: A survey and some new methods," *Journal of Control Theory and Applications*, 9(3), 310-335.

[10] Bray, R. (2016): "Ergodicity and the Estimation of Markov Decision Processes," manuscript available at SSRN.

[11] Coleman, W. (1990): "Solving the Stochastic Growth Model by Policy Function Iteration," *Journal of Business and Economic Statistics*, 8, 27–29.

[12] Coleman, W. (1991): "Equilibrium in a Production Economy With Income Tax," *Econometrica*, 59, 1091-1104.

[13] Gittins, J. (1979): "Bandit Processes and Dynamic Allocation Indices," *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2), 148-177.

[14] Gittins, J., K. Glazebrook, and R. Weber (2011): "Multi-armed bandit allocation indices." John Wiley & Sons.

[15] Hansen, L. P., and K. J. Singleton (1982): "Generalized instrumental variables estimation of nonlinear rational expectations models," *Econometrica*, 50, 1269-1286.

[16] Hotz, J., and R.A. Miller (1993): "Conditional choice probabilities and the estimation of dynamic models," *Review of Economic Studies*, 60, 497-529.

[17] Hotz, J., R.A. Miller, S. Sanders, and J. Smith (1994): "A simulation estimator for dynamic models of discrete choice," *Review of Economic Studies*, 61, 265-89.

[18] Magnus, J. and H. Neudecker (1988): "Matrix Differential Calculus with Applications in Statistics and Econometrics," New York: Wiley.

[19] McFadden, D. (1981): "Econometric Models of Probabilistic Choice," in C. Manski and D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge, MA.

[20] Newey, W.K. (1984): "A Method of Moments Interpretation of Sequential Estimators," Economics Letters 14, pp. 201 –206.

[21] Newey, W.K. (1994): "The Asymptotic Variance of Semiparametric Estimators," Econometrica, 62, 1349 –1382.

[22] Newey, W.K. and D. F. McFadden (1994): "Large sample estimation and hypothesis testing", in R.F. Engle III and D.F. McFadden (eds.), The Handbook of Econometrics, vol. 4. North-Holland, Amsterdam.

[23] Olley, S. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263-1297.

[24] Powell, W. (2007): "Approximate Dynamic Programming: Solving the curses of dimensionality," Vol. 703. John Wiley & Sons.

[25] Puterman, M. and S. Brumelle (1979): "On the convergence of policy iteration in stationary dynamic programming," *Mathematics of Operations Research*, 4(1), 60-69.

[26] Puterman, M. (1994): "Markov decision processes: discrete stochastic dynamic programming." John Wiley & Sons. New Jersey.

[27] Rust, J. (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica*, 55, 999-1033.

[28] Rust, J. (1988): "Maximum Likelihood Estimation of Discrete Control Processes," *SIAM Journal of Control and Optimization*, 26(5), 1006-1024.

[29] Rust, J. (1992): "Do People Behave According to Bellman's Principle of Optimality?" Working Paper E-92-10. The Hoover Institute. Stanford University.

[30] Rust, J. (1994): "Structural estimation of Markov decision processes," in R. E. Engle and McFadden (eds.) *Handbook of Econometrics Volume 4*, North-Holland.

[31] Rust, J. (1996): "Numerical dynamic programming in economics," Handbook of computational economics, 1, 619-729.

[32] Rust, J. (1997): "Using Randomization to Break the Curse of Dimensionality," *Econometrica*, 65 (3), 487-516.

[33] Tauchen, G. (1986): "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20(2), 177-181.