University of Toronto
Department of Economics

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

By Victor Aguirregabiria and Arvind Magesan

May 24, 2016

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 10th, 2016

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 computation of a solution and to the evaluation of counterfactual experiments. We use a choice probabilities representation of the discrete decision problem to derive marginal conditions of optimality with the same features as the standard EEs in continuous decision problems. These EEs imply a fixed point mapping in the space of conditional choice values, that we denote the Euler equation-value (*EE-value*) operator. We show that, in contrast to Euler equation operators in continuous decision models, this operator is a contraction. We present numerical examples that illustrate how solving the model by iterating in the EE-value mapping implies substantial computational savings relative to iterating in the Bellman equation (that requires a much larger number of iterations) or in the policy function (that involves a costly valuation step). We define a sample version of the EE-value 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-value operator increases linearly with sample size, and provides an unbiased (in finite samples) and consistent estimator 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 several Monte Carlo experiments.

Keywords: Dynamic programming discrete choice models; Euler equations; Policy 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`

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 over alternative approaches 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, or at least an approximation to, the solution to the dynamic programming problem. In other words, even if the researcher can avoid full solution in estimating the structural parameters, full solution will be required when using the estimated model to study counterfactual policies in any case, 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 in the space of the policy function (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

---

[1] Another nice 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 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."*

state spaces and richer sources of individual specific heterogeneity. Nevertheless, in contrast to the Euler equations approach, the implementation of CCP or Hotz-Miller methods 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 main contribution of this paper is to extend the Euler Equation (EE) representation of dynamic decision problems to a general class of discrete choice models and to show that the advantages of this approach apply not only to the estimation of structural parameters but also to the computation of a solution and to the evaluation of counterfactual experiments.

First, we derive a representation of the discrete choice model as a continuous decision problem where the decision variables are choice probabilities. Using this equivalent representation, we derive marginal conditions of optimality similar in nature to the EEs in standard continuous decision problems. Second, we show that these EEs imply a fixed point mapping in the space of conditional choice values, that we denote as the *Euler Equation-value (EE-value) 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. Furthermore, in contrast to the standard policy-iteration mapping in DP problems, iterating in the EE mapping does not involve the computation of infinite-period-forward present values, but only one-period-forward expectations. 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 iterating in the Bellman equation (that requires a larger number of iterations because it is a weaker contraction) or in the policy function (that involves a costly infinite-periods-forward valuation step). These computational savings increase more than proportionally with the dimension of the state space.

Second, we define a sample version of the EE-value operator and use it to construct a sequence of consistent estimators of the structural parameters, and to evaluate counterfactual experiments. This sample-based EE-value operator is defined only at sample points of the exogenous state variables, and thus its dimensionality is relatively small and does not increase with the dimension of the state space. We show that this sample-based EE-value operator is also a contraction and the unique fixed

---

[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.

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. This is because the evaluation of the EE operator does not involve the computation of present values, only sample averages of next period's payoffs.

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 alternative methods for the exact solution of the model. For a moderately sized state space (relative to what is commonly found in applications), the standard method of policy function iterations takes over 200 times as long as the method of iterating in the *EE-value* mapping, and this difference increases with the dimensionality 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 the two methods. In particular, we show that although the standard policy iteration mapping needs fewer iterations to obtain a fixed point,[4] each iteration of the *EE mapping* is so relatively inexpensive that it ends up being considerably faster in total time to convergence.

In a second experiment, we study the finite sample properties and the computational cost of estimators of the structural parameters using standard methods in this literature and using the method based on the EE-value mapping. More specifically, we compare two-step Hotz-Miller (Hotz and Miller, 1993) and Maximum Likelihood Estimator (MLE) estimators with the two-step and K-step Pseudo Maximum Likelihood (PML) estimators based on our Euler equation representation. We find that the two-step Euler equation estimator has about 33% higher root mean squared error than the MLE. However, the K-step 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 two-step Hotz-Miller estimator takes almost 5000 times as long as the two-step Euler Equations estimator, and the K-step Euler equations estimator is over 2000 times faster than the MLE. 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

---

[4] Policy iterations are composite mappings which involve solution of infinite period forward expected values as an intermediate step. Each iteration is a larger step to convergence than is the case with value iterations.

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 that 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.

This paper is related to a literature that exploits properties of dynamic discrete decision problems to obtain a representation of the model that does not involve the calculation of present discounted values of the stream of payoffs over infinite future periods. Key contributions in this literature are Hotz and Miller (1993) and Arcidiacono and Miller (2011, 2015) who 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 does not need to compute present values to estimate the structural parameters of the model.[5] Our sample-based EE operator is also 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.

The rest of the paper is organized as follows. Section 2 presents the model and fundamental results in the literature. Section 3 describes our derivation of Euler equations in discrete choice models, defines the *EE-value mapping*, and shows that it is a contraction. Section 4 defines the sample version of the EE-value 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.

---

[5]Following Arcidiacono and Miller (2011), a dynamic decision problem has the finite dependence property if, beginning at some state X, two different decisions a and a' arrive at the same state X' in finite time with probability one.

4

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 $\mathbf{x}_t$ itself is comprised by two types of state variables, exogenous variables $\mathbf{z}_t$ and endogenous variables $\mathbf{y}_t$. They are distinguished by the fact that the transition probability of the endogenous variables depends on the action a_t, while the transition probability of the exogenous variables does not depend on a_t. The vector of unobservables satisfies the standard assumptions of additive separability (AS), conditional independence (CI), and discrete support of the observable state variables. Specifically, the one-period payoff function is additively separable in the unobservables: $\Pi_t(a_t, \mathbf{s}_t) = \pi_t(a_t, \mathbf{x}_t) + \varepsilon_t(a_t)$, where $\varepsilon_t \equiv \{\varepsilon_t(a) : a \in A\}$ is a vector of unobservable random variables, and the transition probability (density) function of the state variables factors as: $f(\mathbf{s}_{t+1}|a_t, \mathbf{s}_t) = f_x(\mathbf{x}_{t+1}|a_t, \mathbf{x}_t) \ dG(\varepsilon_{t+1})$, 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 $\mathcal{X}$. For notational convenience, unless necessary, we omit the exogenous state variables $\mathbf{z}_t$ and treat the whole vector $\mathbf{x}_t$ as endogenous.

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*,

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

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., $\widetilde{\mathbf{v}}_t(\mathbf{x}_t) = \{\widetilde{v}_t(a, \mathbf{x}_t) : a \neq 0\}$. The *optimal choice probability (OCP) mapping*, $\mathbf{\Lambda}\left(\widetilde{\mathbf{v}}_t(\mathbf{x})\right) \equiv \{\Lambda\left(a, \widetilde{\mathbf{v}}_t(\mathbf{x})\right) : a \neq 0\}$, is defined as a mapping from $\mathbb{R}^J$ into $[0, 1]^J$. It is the probability that given the observable state $\mathbf{x}$ the optimal choice at period t is alternative a. Given the form of the optimal decision rule in equation (3), the OCP function is:

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

where $1\{.\}$ is the indicator function. In vector form, the OCP mapping is defined as $\mathbf{\Lambda}\left(\widetilde{\mathbf{v}}_t(\mathbf{x})\right) \equiv \{\Lambda\left(a, \widetilde{\mathbf{v}}_t(\mathbf{x})\right) : a \neq 0\}$. Given a vector of choice probabilities $\mathbf{P}_t \equiv \{P_t(a) : a = 0, 1, ..., J\}$, we say that this vector is optimal at period t given state $\mathbf{x}$ if and only if $\mathbf{P}_t = \mathbf{\Lambda}\left(\widetilde{\mathbf{v}}_t(\mathbf{x})\right)$ where $\widetilde{\mathbf{v}}_t(\mathbf{x})$ is the vector value differences as defined in (2), and the value function solves the Bellman equation in (1).

Proposition 1 establishes that the OCP mapping is invertible.

PROPOSITION 1 [Hotz-Miller Inversion]. The mapping $\mathbf{\Lambda}\left(\widetilde{\mathbf{v}}\right)$ 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 $\mathbf{P}_t(\mathbf{x})$ for a given value of $\mathbf{x}$, 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 primitive or structural functions $\{\pi_t, f_x, G, \beta\}_{t=1}^T$. 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 $\mathbf{P}_t \equiv \{P_t(a) : a_t = 0, 1, ..., J\}$. We denote this problem as the *dynamic probability-choice problem*. We show below that there is a close 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, $\mathbf{P}_t$, we define the following expected payoff function,

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

where $e_t(a, \mathbf{P}_t)$ is the expected value of $\varepsilon_t(a)$ conditional on: (i) alternative a being the optimal choice; and (ii) $\mathbf{P}_t$ being the vector of optimal choice probabilities. Condition (i) implies that $e_t(a, \mathbf{P}_t)$ is the expectation $\mathbb{E}_t[\varepsilon_t(a) \mid \widetilde{v}_t(a) + \varepsilon_t(a) \geq \widetilde{v}_t(j) + \varepsilon_t(j)$ for any $j]$. By *Hotz-Miller Inversion Theorem*, condition (ii) implies that $\widetilde{v}_t(a) = \Lambda^{-1}(a, \mathbf{P}_t)$. Therefore,[6]

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

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

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

Now, we define a dynamic programming problem where the decision at period t is the vector of choice probabilities $\mathbf{P}_t$, the current payoff function is $\Pi_t^P(\mathbf{P}_t, x_t)$, and the transition probability of the state variables is $f^P(x_{t+1} \mid \mathbf{P}_t, x_t)$. By definition, the Bellman equation of this problem is:

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

This is the *dynamic probability-choice* problem associated to the original dynamic discrete choice model. The solution of this DP problem can be described in terms of the 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 dynamic 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 dynamic probability-choice problem, i.e., $W_t(\mathbf{P}_t, \mathbf{x}_t) \equiv \Pi_t^P(\mathbf{P}_t, \mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1}} V_{t+1}^P(\mathbf{x}_{t+1}) \ f^P(\mathbf{x}_{t+1}|\mathbf{P}_t, \mathbf{x}_t)$.

---

[6] 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 $e_t(a, \mathbf{P}_t) = \phi(\Phi^{-1}[P_t(a)])/P_t(a)$.

PROPOSITION 2. For any dynamic discrete choice problem that satisfies the assumptions of Additive Separability (AS) and Conditional Independence (CI), the associated dynamic 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., $\mathbf{P}_t^*(\mathbf{x}_t) = \mathbf{\Lambda}(\widetilde{\mathbf{v}}_t(\mathbf{x}_t))$.* ∎

Proof: In the Appendix.

Proposition 2 establishes a representation property of this class of discrete choice models. The dynamic discrete-choice model has a representation as a dynamic probability-choice problem with the particular definitions of expected payoff and expected transition probability functions presented in equations (5)-(7) above. In section 3, we show that we can exploit this representation to derive Euler equations (EEs) for discrete choice DP models. Proposition 3 plays also an important role in our derivation of Euler equations.

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 Euler Equations and fixed point mappings

3.1 Deriving Euler Equations

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, e.g., $\mathbf{x}_{t+1} = f(\mathbf{x}_t, a_t) + \xi_{t+1}$. The dynamic probability-choice model defined by Bellman equation (8) does not necessarily 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.

We follow an approach that builds on and extends the one in Pakes (1994). 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:

$$\max_{\{\mathbf{P}_t, \mathbf{P}_{t+1}\}} \quad \Pi_t^P(\mathbf{P}_t, \mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1} \in \mathcal{X}} \Pi_{t+1}^P(\mathbf{P}_{t+1}, \mathbf{x}_{t+1}) \, f^P(\mathbf{x}_{t+1} | \mathbf{P}_t, \mathbf{x}_t)$$

$$\text{subject to:} \quad f_{(2)}^P(\mathbf{x}_{t+2} \mid \mathbf{P}_t, \mathbf{P}_{t+1}, \mathbf{x}_t) = f_{(2)}^P(\mathbf{x}_{t+2} \mid \mathbf{P}_t^*, \mathbf{P}_{t+1}^*, \mathbf{x}_t) \text{ for any } \mathbf{x}_{t+2}$$

(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 $t + 1$:

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

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 $t + 1$.[7] 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 marginal payoffs at periods t and $t + 1$.

For the description of these Euler equations, it is convenient to incorporate some additional notation. Let $\mathcal{X}$ be the (unconditional) support set of the vector of endogenous state variables $\mathbf{x}_t$. For a given vector $\mathbf{x}_t$, let $\mathcal{X}_{(1)}(\mathbf{x}_t) \subseteq \mathcal{X}$ be the support of $\mathbf{x}_{t+1}$ conditional on $\mathbf{x}_t$. That is, $\mathbf{x}_{t+1} \in \mathcal{X}_{(1)}(\mathbf{x}_t)$ if and only if $f(\mathbf{x}_{t+1} | a_t, \mathbf{x}_t) > 0$ for some value of a_t. Similarly, for given $\mathbf{x}_t$, let $\mathcal{X}_{(2)}(\mathbf{x}_t) \subseteq \mathcal{X}$ be the set with all the vectors $\mathbf{x}_{t+2}$ with $\Pr(\mathbf{x}_{t+2} | a_t, a_{t+1}, \mathbf{x}_t) > 0$ for some value of a_t and a_{t+1}. Let $\widetilde{f}(\mathbf{x}_{t+1} | a_t, \mathbf{x}_t)$ be the *difference* transition probability $f(\mathbf{x}_{t+1} | a_t, \mathbf{x}_t) - f(\mathbf{x}_{t+1} | 0, \mathbf{x}_t)$, where using

---

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

choice alternative 0 as the baseline is without loss of generality. Let $\widetilde{\mathbf{F}}_{t+1}(\mathbf{x}_t)$ be a matrix with elements $\widetilde{f}(\mathbf{x}_{t+2}|a_{t+1}, \mathbf{x}_{t+1})$ where the columns correspond to all the values $\mathbf{x}_{t+2} \in \mathcal{X}_{(2)}(\mathbf{x}_t)$ leaving out one value, and the rows correspond to all the values $(a_{t+1}, \mathbf{x}_{t+1}) \in [A - \{0\}] \times \mathcal{X}_{(1)}(\mathbf{x}_t)$. 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 conditions is a system of J equations, one for each probability $P_t(a)$ with $a > 0$,

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

where $\lambda(\mathbf{x}_{t+2})$ is the Lagrange multiplier associated to the constraint for state $\mathbf{x}_{t+2}$.[8] The second set of Lagrange conditions is a system of $J * |\mathcal{X}_{(1)}|$ equations, one for each probability $P_{t+1}(a|\mathbf{x}_{t+1})$ with $a > 0$ and $\mathbf{x}_{t+1} \in \mathcal{X}_{(1)}$,

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

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

Our derivation of Euler equations consists of using the system of $J * |\mathcal{X}_{(1)}|$ equations in (13) to solve for the vector of $|\mathcal{X}_{(2)}| - 1$ Lagrange multipliers, and then plug-in this solution into the system of equations (12). The key condition for the existence of Euler equations comes from the existence of a unique solution for the Lagrange multipliers in the system of equations (13). Using the definition of the matrix $\widetilde{\mathbf{F}}_{t+1}$ above, we can represent the system of equations (13) in vector form as:

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

where λ is the vector of Lagrange multipliers, and $\partial \mathbf{\Pi}_{t+1}^P/\partial \mathbf{P}_{t+1}$ is a vector with dimension $J * |\mathcal{X}_{(1)}|$ that contains the marginal expected profits $\partial \Pi_{t+1}^P/\partial P_{t+1}(a|\mathbf{x}_{t+1})$. Proposition 4 establishes the conditions for existence of Euler equations and presents the general formula.

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, $\lambda = \beta \ \mathbf{M}_{t+1} \ \dfrac{\partial \mathbf{\Pi}_{t+1}^P}{\partial \mathbf{P}_{t+1}}$, where $\mathbf{M}_{t+1}$ is the matrix $[\widetilde{\mathbf{F}}_{t+1}'\widetilde{\mathbf{F}}_{t+1}]^{-1} \ \widetilde{\mathbf{F}}_{t+1}'$, and the

---

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

following system of Euler equations,

$$\frac{\partial \Pi_t^P}{\partial P_t(a)} + \beta \sum_{\mathbf{x}_{t+1}} \left[\Pi_t^P(\mathbf{P}_{t+1}, \mathbf{x}_{t+1}) - \beta \sum_{\mathbf{x}_{t+2}} \left(\mathbf{M}_{t+1}(\mathbf{x}_{t+2}) \frac{\partial \Pi_{t+1}^P}{\partial \mathbf{P}_{t+1}} \right) f^P(\mathbf{x}_{t+2} | \mathbf{P}_{t+1}, \mathbf{x}_{t+1}) \right] \widetilde{f}(\mathbf{x}_{t+1} | a) = 0 \tag{15}$$

where $\mathbf{M}_{t+1}(\mathbf{x})$ is the row vector in matrix $\mathbf{M}_{t+1}$ associated to $\mathbf{x}_{t+2} = \mathbf{x}$. ∎

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 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 that show in this paper.

We now provide some examples to illustrate the derivation of these Euler equations and to present the simple form of these equations in some models that have received substantial attention in empirical applications.

3.1.1   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, $a_t \in \{0, 1, ..., J\}$. There are costs of changing occupations such that the choice of occupation in the previous period is a state variable. Suppose that the previous period's occupation is the only endogenous state variable of the model. Then, the state space is $\mathcal{X} = \{0, 1, ..., J\}$, and the transition function is given by $x_{t+1} = a_t$ such that $f(x_{t+1} | a_t, x_t) = 1\{x_{t+1} = a_t\}$. This implies that $f^P(x_{t+1} = a | \mathbf{P}_t, x_t) = P_t(a | x_t)$, and the two-periods forward transition is $f_{(2)}^P(x_{t+2} = a' | \mathbf{P}_t, \mathbf{P}_{t+1}, x_t) = \sum_{a=0}^{J} P_t(a | x_t) \, P_{t+1}(a' | a)$. The constrained optimization problem is:

$$\max_{\{\mathbf{P}_t, \mathbf{P}_{t+1}\}} \quad \Pi_t^P(\mathbf{P}_t, x_t) + \beta \sum_{a=0}^{J} P_t(a | x_t) \, \Pi_{t+1}^P(\mathbf{P}_{t+1}, a)$$

$$\text{subject to:} \quad \sum_{a=0}^{J} P_t(a | x_t) \, P_{t+1}(a' | a) = \text{constant} \quad \text{for any } a' \tag{16}$$

Taking into account that there are only J free probabilities such that $P_t(0 | x_t) = 1 - \sum_{a=1}^{J} P_t(a | x_t)$, the Lagrange condition with respect to $P_t(a | x_t)$ is:

$$\frac{\partial \Pi_t^P}{\partial P_t(a | x_t)} + \beta \left[\Pi_{t+1}^P(\mathbf{P}_{t+1}, a) - \Pi_{t+1}^P(\mathbf{P}_{t+1}, 0) \right] - \sum_{a'=1}^{J} \lambda(a') \left[P_{t+1}(a' | a) - P_{t+1}(a' | 0) \right] = 0 \tag{17}$$

11

The Lagrange condition with respect to $P_{t+1}(a'|a)$ is,

$$\beta \frac{\partial \Pi_{t+1}^P(\mathbf{P}_{t+1}, a)}{\partial P_{t+1}(a'|a)} = \lambda(a') \tag{18}$$

Combining these two sets of conditions we get the following system of Euler equations: for any $a > 1$,

$$\frac{\partial \Pi_t^P}{\partial P_t(a|x_t)} + \beta \left[\Pi_{t+1}^P(\mathbf{P}_{t+1}, a) - \Pi_{t+1}^P(\mathbf{P}_{t+1}, 0) \right]$$

$$\tag{19}$$

$$-\beta \sum_{a'=1}^{J} \left[\frac{\partial \Pi_{t+1}^P(\mathbf{P}_{t+1}, a)}{\partial P_{t+1}(a'|a)} P_{t+1}(a'|a) - \frac{\partial \Pi_{t+1}^P(\mathbf{P}_{t+1}, 0)}{\partial P_{t+1}(a'|0)} P_{t+1}(a'|0) \right] = 0$$

We can simplify further this expression. Using Proposition 3 and after some operations we can obtain:

$$\pi_t(a, x_t) + e(a, \mathbf{P}_t(x_t)) + \beta \left[\pi_{t+1}(0, a) + e(0, \mathbf{P}_{t+1}(a)) \right] =$$

$$\tag{20}$$

$$\pi_t(0, x_t) + e(0, \mathbf{P}_t(x_t)) + \beta \left[\pi_{t+1}(0, 0) + e(0, \mathbf{P}_{t+1}(0)) \right]$$

For instance, when the unobservables are i.i.d. extreme value distributed, we have that $e_t(a, \mathbf{P}_t(x_t)) = \gamma - \ln P_t(a|x_t)$ and this Euler equation becomes:

$$\pi_t(a, x_t) - \ln P_t(a|x_t) + \beta \left[\pi_{t+1}(0, a) - \ln P_{t+1}(0|a) \right] =$$

$$\tag{21}$$

$$\pi_t(0, x_t) - \ln P_t(0|x_t) + \beta \left[\pi_{t+1}(0, 0) - \ln P_{t+1}(0|0) \right]$$

Euler equation (20) has a clear economic interpretation. There are two possible decision paths from state x_t at period t to state $x_{t+2} = 0$ at period $t + 2$: the decision path $(a_t = 1, a_{t+1} = 0)$, and the decision path $(a_t = 0, a_{t+1} = 0)$. Euler equation (20) establishes that, at the optimal solution the discounted expected payoff at periods t and $t + 1$ should be the same for the two decision paths. This is 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 implied by equation (21) is associated with the value $x_{t+2} = 0$. We note here as well that one can obtain a different expression using the Euler Equation associated with the state value $x_{t+2} = a > 0$, as in general, there are as many Euler Equations as there are possible values of x_{t+2}.

3.1.2 Example 2: Machine replacement model

Consider the bus engine replacement problem in Rust (1987). In this example, the endogenous state variable x_t is the mileage on the bus engine. Suppose that the space of possible mileages is given by the discrete set $\mathcal{X} = \{0, 1, 2, ...\}$, and that mileage follows a transition rule $x_{t+1} = (1 - a_t)(x_t + 1)$,

where $a_t \in \{0, 1\}$ represents the machine replacement decision. Given that the state at period t is $x_t = x$, the set of possible states one period forward is $\mathcal{X}_{(1)}(x) = \{0, x+1\}$, and the set of possible states two periods forward is $\mathcal{X}_{(2)}(x) = \{0, 1, x+2\}$. The transition probability induced by the choice probability is $f^P(x_{t+1} = 0|\mathbf{P}_t, x_t) = P_t(1|x_t)$ and $f^P(x_{t+1} = x_t + 1|\mathbf{P}_t, x_t) = 1 - P_t(1|x_t)$, and the two-periods forward transitions are $f^P_{(2)}(x_{t+2} = 1 \mid \mathbf{P}_t, \mathbf{P}_{t+1}, x_t) = P_t(1|x_t) \ P_{t+1}(0|0)$ and $f^P_{(2)}(x_{t+2} = x + 2 \mid \mathbf{P}_t, \mathbf{P}_{t+1}, x) = P_t(0|x) \ P_{t+1}(0|x+1)$. The restrictions in the constraint optimization problem are:

$$P_t(1|x_t) \ P_{t+1}(0|0) \quad = \quad \text{constant}$$

$$P_t(0|x_t) \ P_{t+1}(0|x_t + 1) \quad = \quad \text{constant}$$

(22)

The Lagrange condition with respect to $P_t(1|x)$ is:

$$\frac{\partial \Pi^P_t}{\partial P_t(1|x)} + \beta \left[\Pi^P_{t+1}(\mathbf{P}_{t+1}, 0) - \Pi^P_{t+1}(\mathbf{P}_{t+1}, x+1) \right] + \lambda(1) \ P_{t+1}(0|0) + \lambda(x+2) \ P_{t+1}(0|x+1) = 0 \quad (23)$$

The Lagrange conditions with respect to $P_{t+1}(1|0)$ and $P_{t+1}(1|x+1)$ are,

$$\beta \frac{\partial \Pi^P_{t+1}(\mathbf{P}_{t+1}, 0)}{\partial P_{t+1}(1|0)} = \lambda(1) \quad \text{and} \quad \beta \frac{\partial \Pi^P_{t+1}(\mathbf{P}_{t+1}, x+1)}{\partial P_{t+1}(1|x+1)} = \lambda(x+2) \quad (24)$$

Combining these conditions, we get the following system of Euler equations: for any $a > 1$,

$$\frac{\partial \Pi^P_t}{\partial P_t(a|x_t)} + \beta \ \left[\Pi^P_{t+1}(\mathbf{P}_{t+1}, 0) - \Pi^P_{t+1}(\mathbf{P}_{t+1}, x+1) \right]$$

$$-\beta \left[\frac{\partial \Pi^P_{t+1}(\mathbf{P}_{t+1}, 0)}{\partial P_{t+1}(1|0)} \ P_{t+1}(0|0) - \frac{\partial \Pi^P_{t+1}(\mathbf{P}_{t+1}, x+1)}{\partial P_{t+1}(1|x+1)} \ P_{t+1}(0|x+1) \right] \quad = \quad 0$$

(25)

We can simplify further this expression. Using Proposition 3, we obtain that:

$$\pi_t(1, x_t) + e(1, \mathbf{P}_t(x_t)) + \beta \left[\pi_{t+1}(1, 0) + e(1, \mathbf{P}_{t+1}(0)) \right] \quad =$$

$$\pi_t(0, x_t) + e(0, \mathbf{P}_t(x_t)) + \beta \left[\pi_{t+1}(1, x_t + 1) + e(1, \mathbf{P}_{t+1}(x_t + 1)) \right]$$

(26)

Again, there is a clear economic interpretation of this Euler equation. There are two possible decision paths from state x_t at period t to state $x_{t+2} = 0$ 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 (26) establishes that the discounted expected payoff at periods t and $t + 1$ should be the same for the two decision paths.

3.2 Euler Equation fixed point mappings

In this section, we show that the system of Euler equations derived above imply two different fixed point mappings: (i) a mapping in the space of conditional choice probabilities, that we denote as the *Euler Equation - Probability (EE-p) mapping*; and (ii) a mapping in the space of conditional choice values, that we denote as the *Euler Equation - Value (EE-v) mapping*. We show that the Euler Equation-Value mapping is a contraction.

13

3.2.1 Euler Equation-Probability mapping

Consider the system of Euler equations in (21). Solving for $P_t(a|\mathbf{x}_t)$ we have that:

$$P_t(a|\mathbf{x}_t) = \frac{\exp\left\{\pi(a,\mathbf{x}_t) - \pi(0,\mathbf{x}_t) + \beta\,\mathbb{E}_t\left[\pi(0,a,z_{t+1}) - \pi(0,0,z_{t+1}) - \ln\left(\frac{P_{t+1}(0|a,z_{t+1})}{P_{t+1}(0|0,z_{t+1})}\right)\right]\right\}}{\displaystyle\sum_{j=0}^{J}\exp\left\{\pi(j,\mathbf{x}_t) - \pi(0,\mathbf{x}_t) + \beta\,\mathbb{E}_t\left[\pi(0,j,z_{t+1}) - \pi(0,0,z_{t+1}) - \ln\left(\frac{P_{t+1}(0|j,z_{t+1})}{P_{t+1}(0|0,z_{t+1})}\right)\right]\right\}} \tag{27}$$

The right-hand-side of this equation describes a mapping $\Gamma_{EE,p}(a,\mathbf{x}_t,\mathbf{P}_{t+1})$ from the vector of choice probabilities at period $t+1$, $\mathbf{P}_{t+1}$, into the probability space, such that $\Gamma_{EE,p}(a,\mathbf{x}_t,.)$: $[0,1]^{|\mathcal{X}|J} \to [0,1]$. Let $\Gamma_{EE,p}(\mathbf{P})$ be the vector-valued function that consists of the collection of the functions $\Gamma_{EE,p}(a,\mathbf{x}_t,.)$ for every value of $(a,\mathbf{x}_t)$ in the space $(\mathcal{A} - \{0\}) \times \mathcal{X}$:

$$\mathbf{\Gamma}_{EE,p}(\mathbf{P}) \equiv \{\Gamma_{EE,p}(a,\mathbf{x}_t,\mathbf{P}) : \text{for any } (a,\mathbf{x}_t) \in (\mathcal{A} - \{0\}) \times \mathcal{X}\} \tag{28}$$

By definition, $\mathbf{\Gamma}_{EE,p}(\mathbf{P})$ is a fixed point mapping in the probability space $[0,1]^{|\mathcal{X}|J}$. Using this mapping we can represent the relationship between choice probabilities at periods t and $t+1$ implied by the Euler Equations as follows:

$$\mathbf{P}_t = \mathbf{\Gamma}_{EE,p}(\mathbf{P}_{t+1}) \tag{29}$$

In the stationary version of the model (i.e., infinite horizon and time-homogeneous payoff function), the optimal choice probabilities are time invariant: $\mathbf{P}_t = \mathbf{P}_{t+1}$. Therefore, expression (29) describes the optimal choice probabilities as a fixed point of the mapping $\Gamma_{EE,p}$. We denote $\Gamma_{EE,p}$ as the *Euler Equation - Probability (EE-p) mapping.*

We can use this mapping to solve the DP problem. A computational advantage of solving the model using fixed-point iterations in the EE-p mapping is that each iteration does not involve calculating present values, as in the standard policy iteration or Newton-Kantorovich method (Puterman and Brumelle, 1979, and Puterman, 1994). 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). Here we use the more compact notation $P_0 \equiv P(1|0)$ and $P_1 \equiv P(1|1)$. The Euler

14

equation (21) 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

$$\Gamma_{EE,p}(x,\mathbf{P}) = \frac{\exp\left\{c\left(x\right) + \beta\left[-\ln(1-P_1) + \ln(1-P_0)\right]\right\}}{1 + \exp\left\{c\left(x\right) + \beta\left[-\ln(1-P_1) + \ln(1-P_0)\right]\right\}} \tag{30}$$

and $c\left(x\right) \equiv [\pi(1,x) - \pi(0,x)] + \beta[\pi(0,1) - \pi(0,0)]$. For vectors $\mathbf{P}$ and $\mathbf{Q}$ with the set $[0,1]^2$, consider the ratio $r(\mathbf{P},\mathbf{Q}) = \|\mathbf{\Gamma}_{EE,p}(\mathbf{P}) - \mathbf{\Gamma}_{EE,p}(\mathbf{Q})\| / \|\mathbf{P} - \mathbf{Q}\|$. 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 the set $[0,1]^2$. 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,

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

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. $\blacksquare$

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
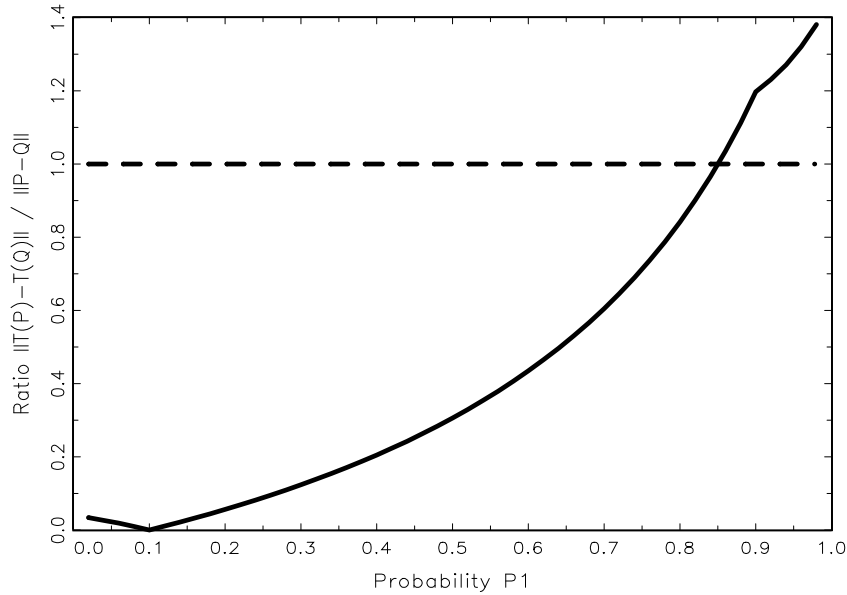
3.2.2 Euler Equation-Value mapping

Let $v_t(a, \mathbf{x}_t)$ be the conditional choice value function, and let $\widetilde{v}_t(a, x_t)$ be the differential conditional choice value function, i.e., $\widetilde{v}_t(a, \mathbf{x}_t) \equiv v_t(a, \mathbf{x}_t) - v_t(0, \mathbf{x}_t)$. Hotz-Miller (1993) inversion theorem (Proposition 1 above) establishes that there is a one-to-one mapping between value differences and choice probabilities. In the logit model, this mapping is such that $P_t(a|\mathbf{x}_t) = \exp\{\widetilde{v}_t(a, \mathbf{x}_t)\}/[1 + \sum_{j=1}^{J} \exp\{\widetilde{v}_t(j, \mathbf{x}_t)\}]$, and the inverse of this mapping is $\widetilde{v}_t(a, \mathbf{x}_t) = \ln P_t(a|\mathbf{x}_t) - \ln P_t(0|\mathbf{x}_t)$. Solving these expressions into the Euler equation (21), we obtain the following equation in terms of value differences:

$$
\begin{aligned}
\widetilde{v}_t(a, \mathbf{x}_t) &= [\pi(a, \mathbf{x}_t) - \pi(0, \mathbf{x}_t)] + \beta \, \mathbb{E}_t [\pi(0, a, z_{t+1}) - \pi(0, 0, z_{t+1})] \\
&+ \beta \, \mathbb{E}_t \left[-\ln\left(1 + \sum_{j=1}^{J} \exp\{\widetilde{v}_{t+1}(j, a, z_{t+1})\}\right) + \ln\left(1 + \sum_{j=1}^{J} \exp\{\widetilde{v}_{t+1}(j, 0, z_{t+1})\}\right) \right]
\end{aligned}
\tag{31}
$$

Let $\mathcal{V}$ be the space of the vector of value differences $\widetilde{\mathbf{v}} \equiv \{v(a, \mathbf{x}) : (a, \mathbf{x}) \in (\mathcal{A} - \{0\}) \times \mathcal{X}\}$, such that $\mathcal{V} \subseteq \mathbb{R}^{|\mathcal{X}|J}$. Given that the payoff function $\pi(a, \mathbf{x})$ is bounded on $\mathcal{A} \times \mathcal{X}$, value differences are also bounded and the space $\mathcal{V}$ is a bounded and compact subspace within the Euclidean space $\mathbb{R}^{|\mathcal{X}|J}$. The right-hand-side of equation (31) defines a function $\Gamma_{EE,v}(a, \mathbf{x}_t, \widetilde{\mathbf{v}}_{t+1})$ from the vector of value differences at period $t + 1$, $\widetilde{\mathbf{v}}_{t+1} \in \mathcal{V}$, into the space of values $\mathcal{V}$. Let $\Gamma_{EE,v}(\widetilde{\mathbf{v}})$ be the vector-valued function that consists of the collection of the functions $\Gamma_{EE,v}(a, \mathbf{x}, .)$ for every value of $(a, \mathbf{x})$ in the space $(\mathcal{A} - \{0\}) \times \mathcal{X}$:

$$
\mathbf{\Gamma}_{EE,v}(\widetilde{\mathbf{v}}) \equiv \{\Gamma_{EE,v}(a, \mathbf{x}, \widetilde{\mathbf{v}}) : \text{for any } (a, \mathbf{x}) \in (\mathcal{A} - \{0\}) \times \mathcal{X}\}
\tag{32}
$$

By definition, $\Gamma_{EE,v}(\mathbf{v})$ is a fixed point mapping in the value space $\mathcal{V}$. Using this mapping we can represent the relationship between value differences at periods t and $t + 1$ implied by the Euler Equations as follows:

$$
\widetilde{\mathbf{v}}_t = \mathbf{\Gamma}_{EE,v}(\widetilde{\mathbf{v}}_{t+1})
\tag{33}
$$

In the stationary version of the model, the optimal value differences are time invariant: $\widetilde{\mathbf{v}}_t = \widetilde{\mathbf{v}}_{t+1}$. Therefore, expression (33) describes the vector of optimal value differences as a fixed point of the mapping $\Gamma_{EE,v}$. We denote $\Gamma_{EE,v}$ as the *Euler Equation - value (EE-v) mapping*.

Proposition 5 establishes that the Euler Equation-value mapping is a contraction.

PROPOSITION 5. The Euler Equation - value mapping $\Gamma_{EE,v}$ is a contraction in the complete metric space $(\mathcal{V}, \|.\|_\infty)$, 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 $\mathcal{V}$, we have that $\|\mathbf{\Gamma}_{EE,v}(\widetilde{\mathbf{v}}) - \mathbf{\Gamma}_{EE,v}(\widetilde{\mathbf{w}})\|_\infty \leq \delta \|\widetilde{\mathbf{v}} - \widetilde{\mathbf{w}}\|_\infty.$

Proof: In the Appendix.

A corollary of Proposition 5 is that successive iterations in the EE-v 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: successive approximation to the value function (i.e., iterations in the Bellman equation), policy function (or Newton-Kantorovich) iterations, and hybrids versions of these methods.

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 type of 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 our Euler equation fixed points mappings.

4.1   Empirical Euler-Equation mapping

Given a sample, the researcher can construct an empirical counterpart of the EE mappings defined in Section 3. This empirical EE mapping plays a fundamental role in the estimation of structural parameters and counterfactual experiments that we define below. 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. Let $\mathcal{Z}$ be the set of possible values of $\mathbf{z}$ in the population. Define the empirical set $\mathcal{Z}_N = \{\mathbf{z} \in \mathcal{Z} :$ there is a sample observation (i, t) with $\mathbf{z}_{it} = \mathbf{z}\}$, and the empirical transition probability function $f_{(N)}(\mathbf{z}'|\mathbf{z}_0)$ defined on $\mathcal{Z}_N \times \mathcal{Z}_N$ into $[0, 1]$, such that for any $\mathbf{z}_0 \in \mathcal{Z}_N$, $f_{(N)}(\mathbf{z}'|\mathbf{z}_0) = \sum_{i=1}^{N} 1\{\mathbf{z}_{it+1} = \mathbf{z}'$ and $\mathbf{z}_{it} = \mathbf{z}_0\}/ \sum_{i=1}^{N} 1\{\mathbf{z}_{it} = \mathbf{z}_0\}$. 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 $f_{(N)}(\mathbf{z}'|\mathbf{z}_0)$ converges uniformly to $f_z(\mathbf{z}'|\mathbf{z}_0)$. Let $\mathbb{E}_{\{\mathbf{z}'|\mathbf{z}\}}^{(N)}[.]$ be a sample con-

ditional mean operator from $\mathbb{R}$ into $\mathbb{R}$ such that for any real-valued function $h(\mathbf{z}')$ the operator is defined as:

$$\mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}_0\}}\left[h(\mathbf{z}')\right] \equiv \sum_{\mathbf{z}' \in \mathcal{Z}_N} f_{(N)}(\mathbf{z}'|\mathbf{z}_0)\ h(\mathbf{z}') \tag{34}$$

The *Empirical EE-value mapping* $\Gamma^{(N)}_{EE,v}(\widetilde{\mathbf{v}})$ is defined as the sample counterpart of the EE-value mapping in equation (31) where we replace the conditional expectation at the population level with its empirical counterpart $\mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}_0\}}$. That is, $\Gamma^{(N)}_{EE,v}(\widetilde{\mathbf{v}}) = \{\Gamma^{(N)}_{EE,v}(a, y, \mathbf{z}; \widetilde{\mathbf{v}}) : (a, y, \mathbf{z}) \in \mathcal{A} \times \mathcal{Y} \times \mathcal{Z}_N\}$ where:

$$\begin{aligned}
\Gamma^{(N)}_{EE,v}(a, y, \mathbf{z}; \widetilde{\mathbf{v}}) &= [\pi(a, y, \mathbf{z}) - \pi(0, y, \mathbf{z})] + \beta\, \mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}_0\}} \left[\pi\left(0, a, \mathbf{z}'\right) - \pi\left(0, 0, \mathbf{z}'\right)\ \right] \\
&+ \beta\, \mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}\}} \left[-\ln\left(1 + \sum_{j=1}^{J} \exp\{\widetilde{v}_{t+1}(j, a, \mathbf{z}')\}\right) + \ln\left(1 + \sum_{j=1}^{J} \exp\{\widetilde{v}_{t+1}(j, 0, \mathbf{z}')\}\right)\right]
\end{aligned} \tag{35}$$

$\Gamma^{(N)}_{EE,v}(\widetilde{\mathbf{v}})$ 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:

$$\widetilde{\mathbf{v}} = \mathbf{\Gamma}^{(N)}_{EE,v}(\widetilde{\mathbf{v}}) \tag{36}$$

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 $\Gamma_{EE,v}$ when the dimension of $\mathcal{Z}$ is large relative to the sample size.

Proposition 6 establishes that the Empirical EE-value mapping is a contraction and it converges uniformly in probability to the true EE-value mapping. We now include explicitly the vector of structural parameters θ as an argument in this mapping.

PROPOSITION 6. The Empirical EE-value mapping $\Gamma^{(N)}_{EE,v}(\widetilde{\mathbf{v}}, \theta)$ is a contraction mapping and it converges uniformly in probability to the population EE-value mapping $\Gamma_{EE,v}(\widetilde{\mathbf{v}}, \theta)$.
Proof: In the Appendix.

We also define the *Empirical EE-probability mapping* $\Gamma^{(N)}_{EE,p}(\mathbf{P})$ as the sample counterpart of the EE-prob mapping in equation (28). That is, $\Gamma^{(N)}_{EE,p}(\mathbf{P}) = \{\Gamma^{(N)}_{EE,p}(a, y, \mathbf{z}; \mathbf{P}) : (a, y, \mathbf{z}) \in \mathcal{A} \times \mathcal{Y} \times \mathcal{Z}_N\}$ where:

$$\Gamma^{(N)}_{EE,p}(a, \mathbf{x}; \mathbf{P}) = \frac{\exp\left\{\pi(a, \mathbf{x}) - \pi(0, \mathbf{x}) + \beta\, \mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}\}} \left[\pi(0, a, \mathbf{z}') - \pi(0, 0, \mathbf{z}') - \ln\left(\frac{P(0|a,\mathbf{z}')}{P(0|0,\mathbf{z}')}\right)\right]\right\}}{\sum_{j=0}^{J} \exp\left\{\pi(j, \mathbf{x}) - \pi(0, \mathbf{x}) + \beta\, \mathbb{E}^{(N)}_{\{\mathbf{z}'|\mathbf{z}\}} \left[\pi(0, j, \mathbf{z}') - \pi(0, 0, \mathbf{z}') - \ln\left(\frac{P(0|j,\mathbf{z}')}{P(0|0,\mathbf{z}')}\right)\right]\right\}} \tag{37}$$

18

Using the same approach as in the proof of Proposition 6, it is straightforward to show that $\Gamma_{EE,p}^{(N)}(\mathbf{P},\theta)$ converges uniformly in probability to the population EE-prob mapping $\Gamma_{EE,p}(\mathbf{P},\theta)$. Since the population EE-prob operator is not necessarily a contraction, this is also the case for its sample counterpart.

4.2 Estimation of structural parameters using Euler equations

Given the empirical EE-prob operator, define the Pseudo Likelihood function:

$$Q_N(\theta,\mathbf{P}) = \sum_{i=1}^{N}\sum_{t=1}^{T} \ln \Gamma_{EE,p}^{(N)}(a_{it},\mathbf{x}_{it};\theta,\mathbf{P}) \tag{38}$$

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 in the nonparametric estimation of the conditional choice probabilities $P_t(a|\mathbf{x}) \equiv \Pr(a_{it} = a \mid \mathbf{x}_{it} = \mathbf{x})$. Let $\widehat{\mathbf{P}}_N \equiv \{\widehat{P}_t(a|\mathbf{x}) : t = 1,2,...,T\}$ be a vector of nonparametric estimates of choice probabilities for any choice alternative a and any value of $\mathbf{x}$ observed in the sample. For instance, $\widehat{P}_t(a|\mathbf{x})$ can be a kernel (Nadaraya-Watson) estimator of the regression of $1\{a_i = a\}$ on $\mathbf{x}_{it}$. Note that 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:

$$\widehat{\theta}_N = \arg\max_{\theta \in \Theta} \; Q_N\left(\theta,\widehat{\mathbf{P}}_N\right) \tag{39}$$

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 recently shown by Ackerberg, Chen, and Hahn (2012) using a computationally simpler parametric-like method as in Newey (1984).[9]

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 other Hotz-Miller type of two-step estimators. 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

---

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

can be many 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 under 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 $\Gamma_{EE,v}(.,\theta^*)$, i.e., $\widetilde{\mathbf{v}}^* = \Gamma_{EE,v}(\widetilde{\mathbf{v}}^*,\theta^*)$.

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-value 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-value mapping,

$$\widetilde{\mathbf{v}}_N^* = \mathbf{\Gamma}_{EE,v}^{(N)}(\widetilde{\mathbf{v}}_N^*,\theta^*) \equiv \left\{ \Gamma_{EE,v}^{(N)}(a,y,\mathbf{z};\widetilde{\mathbf{v}}) : (a,y,\mathbf{z}) \in \mathcal{A} \times \mathcal{Y} \times \mathcal{Z}_N \right\} \tag{40}$$

And the corresponding vector of conditional choice probabilities is $\mathbf{P}_N^* = \Lambda(\widetilde{\mathbf{v}}_N^*)$. This approximate solution has several interesting properties that we describe now.

(a) Consistency. $\widetilde{\mathbf{v}}_N^*$ and $\mathbf{P}_N^*$ are consistent estimators of the true counterfactuals $\widetilde{\mathbf{v}}^*$ and $\mathbf{P}^*$.
PROPOSITION 7. The vector of value differences $\widetilde{\mathbf{v}}_N^*$ that is defined as the fixed point $\widetilde{\mathbf{v}}_N^* = \mathbf{\Gamma}_{EE,v}^{(N)}(\widetilde{\mathbf{v}}_N^*,\theta^*)$ is a root-N consistent and asymptotically normal estimator of $\widetilde{\mathbf{v}}^*$.
Proof: In the Appendix.

(b) Low computational cost and no curse of dimensionality. The dimension of the vector $\widetilde{\mathbf{v}}_N^*$ and the mapping $\mathbf{\Gamma}_{EE,v}^{(N)}$ is of the same order as the sample size. In most empirical applications, this dimension is several orders of magnitude smaller than the dimension of the state space and the true $\widetilde{\mathbf{v}}^*$. This substantial reduction in the dimension of the fixed point problem together with the other important computational properties of the EE-value operator (i.e., its contraction property

20

and the no need to compute present values) imply very substantial computational savings. From a practical point of view, the dimension of the operator does not depend on the dimension of the state space $\mathcal{Z}$ but on sample size. Furthermore, our Empirical EE-value 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.

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 four candidate solution algorithms: our EE-value mapping iteration method, the associated EE-policy iteration method, and the standard methods of value function iterations and policy function iterations.[10] 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. These four iterative methods trade these two factors off in different ways. The standard policy iteration mapping is a composite mapping, as the policies are expressed in terms of value functions, which are themselves expressed in terms of the policies. As such, the policy iteration method is very costly per iteration, but the improvement at each iteration is relatively large, so fewer steps are needed. The other three algorithms are not composite mappings, and therefore they are much faster per iteration as there is no intermediate valuation step involved, but they require more steps to convergence. 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: 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 sequential method in Aguirregabiria and Mira

---

[10]For a description of the algorithms of value function iteration and policy function iteration, and their properties, see sections 6.3 and 6.4 in Puterman (1994), and section 5.2 in Rust (1996).

(2002).

Third, given an estimated model and a counterfactual experiment that consists of an increase in the parameter that represents the sunk cost of entry, we present Monte Carlo experiments to evaluate the finite sample properties of four methods to estimate counterfactual choice probabilities. These four methods consist in finding a fixed point of the corresponding empirical operator: EE-value 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.[11] 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 make the normalization $\pi(0, \mathbf{x}_t) = 0$ 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 that varies across firms in the same market; 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).[12]

---

[11] 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).

[12] Let $\{z_j^{(1)}, z_j^{(2)}, ..., z_j^{(K)}\}$ be the support of the state variable z_j, and define the width values $w_j^{(k)} \equiv$

According to the model, the transition of the endogenous state variable induced by the choice probability is the choice probability itself, i.e., $f^P(y_{t+1}|\mathbf{x}_t, P) = P(\mathbf{x}_t) = \Pr(a_t = 1|\mathbf{x}_t)$. The DGP used in our numerical and Monte Carlo experiments is summarized in table 1.

<div align="center">

Table 1
Parameters in the DGP

</div>

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), $\gamma_0^j = 0.0;\quad \gamma_1^j = 0.6;\quad \sigma_e = 1$
Productivity :	ω_t is AR(1), $\gamma_0^\omega = 0.2;\quad \gamma_1^\omega = 0.9;\quad \sigma_e^\omega = 1$
Discount factor	$\beta = 0.95$

This model is a binary choice version of the Multi-armed bandit problem in Example 1. Therefore, taking into account that $\pi(0, \mathbf{x}_t) = 0$, the Euler equation of this model is:

$$\mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}\left[\pi(1, \mathbf{x}_t) - \ln\left(\frac{P(\mathbf{x}_t)}{1 - P(\mathbf{x}_t)}\right) - \beta\,\ln\left(\frac{1 - P(1, \mathbf{z}_{t+1})}{1 - P(0, \mathbf{z}_{t+1})}\right)\right] = 0 \qquad (41)$$

where $\mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}$ denotes the expectation operator over the distribution of $\mathbf{z}_{t+1}$ conditional on $\mathbf{z}_t$. This EE implies the following EE-value mapping: $\widetilde{\mathbf{v}} = \Gamma_{EE,v}(\widetilde{\mathbf{v}})$ with

$$\Gamma_{EE,v}(\mathbf{x}_t, \widetilde{\mathbf{v}}) = \pi(1, \mathbf{x}_t) + \beta\,\mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}\left[\ln(1 + \exp\{\widetilde{v}(1, \mathbf{z}_{t+1})\}) - \ln(1 + \exp\{\widetilde{v}(0, \mathbf{z}_{t+1})\})\right] \quad (42)$$

$\widetilde{v}(y_t, \mathbf{z}_t)$ represents the value difference $v(1, y_t, \mathbf{z}_t) - v(0, y_t, \mathbf{z}_t)$ where we have omitted the action $a_t = 1$ as an argument for notational simplicity given that this is a binary choice model. The EE also implies the following EE-probability mapping: $\mathbf{P} = \Gamma_{EE,p}(\mathbf{P})$ with

$$\Gamma_{EE,p}(\mathbf{x}_t, \mathbf{P}) = \frac{\exp\left\{\pi(1, \mathbf{x}_t) - \beta\,\mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}\left[\ln(1 - P(1, \mathbf{z}_{t+1})) - \ln(1 - P(0, \mathbf{z}_{t+1}))\right]\right\}}{1 + \exp\left\{\pi(1, \mathbf{x}_t) - \beta\,\mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}\left[\ln(1 - P(1, \mathbf{z}_{t+1})) - \ln(1 - P(0, \mathbf{z}_{t+1}))\right]\right\}} \qquad (43)$$

The standard value function mapping is $\mathbf{V} = \Gamma_{VF}(\mathbf{V})$, where

$$\Gamma_{VF}(\mathbf{x}_t, \mathbf{V}) = \ln\left(\exp\left\{\beta\,\mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}[V(0, \mathbf{z}_{t+1})]\right\} + \exp\left\{\pi(1, \mathbf{x}_t) + \beta\,\mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}[V(1, \mathbf{z}_{t+1})]\right\}\right) \quad (44)$$

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

And the standard policy function iteration mapping is $\mathbf{P} = \Gamma_{PF}(\mathbf{P})$, where

$$\Gamma_{PF}(\mathbf{x}_t, \mathbf{P}) = \frac{\exp\left\{\pi\left(1, \mathbf{x}_t\right) + \beta\ \mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}\left[W(1, \mathbf{z}_{t+1}; \mathbf{P})\right]\right\}}{\exp\left\{\beta\ \mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}\left[W(0, \mathbf{z}_{t+1}; \mathbf{P})\right]\right\} + \exp\left\{\pi\left(1, \mathbf{x}_t\right) + \beta\ \mathbb{E}_{\{\mathbf{z}_{t+1}|\mathbf{z}_t\}}\left[W(1, \mathbf{z}_{t+1}; \mathbf{P})\right]\right\}} \tag{45}$$

and $W(\mathbf{x}_t; \mathbf{P})$ is the valuation operator of policy $\mathbf{P}$. The vector of valuations $\mathbf{W}(\mathbf{P}) \equiv \{\mathbf{W}(\mathbf{x}; \mathbf{P}) : \mathbf{x} \in \mathcal{X}\}$ is defined as the solution to the linear system of equations:

$$\left[\mathbf{I} - \beta\ \mathbf{F}^P(\mathbf{P})\right]\ \mathbf{W}(\mathbf{P}) = \Pi^P(\mathbf{P}) \tag{46}$$

where $\mathbf{I}$ is the identity matrix; $\mathbf{F}^P(\mathbf{P})$ is the transition probability matrix for $\mathbf{x}_t$ induced by the vector of choice probabilities $\mathbf{P}$; and $\Pi^P(\mathbf{P})$ is the vector of expected payoffs associated to $\mathbf{P}$ such that an element of this vector is $(1 - P(\mathbf{x}_t))\left[-\ln(1 - P(\mathbf{x}_t))\right] + P(\mathbf{x}_t)\left[\pi\left(1, \mathbf{x}_t\right) - \ln P(\mathbf{x}_t)\right]$. To solve for this system of linear equation we use the QR algorithm based on a QR decomposition of matrix $\mathbf{I} - \beta\ \mathbf{F}^P(\mathbf{P})$.

5.2   Comparing solution methods

We compare the computing times of four methods for the exact solution of the model: (a) successive iterations in the value function (Bellman equation); (b) policy function iterations; (c) iterations in the EE-value operator; and (d) iterations in the EE-probability operator. Methods (a) and (b) are the most standard algorithms for the solution of DP problems. Methods (c) and (d) 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. Table 2 presents the time per iteration, number of iterations, and total computation time as a function of the state space dimensionality. 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 table 2 are the average over these initial values.

For every dimension of the state space, iterating in the EE-value mapping is always the most efficient algorithm, followed closely by iterating in the EE-probability mapping. The computational gains relative to the standard methods are very substantial, as shown in the last column of the table that reports the ratio between the computing times of the policy iteration and EE-value methods. These gains increase with the dimension of the state space.

Table 2
Comparison of Standard and EE Solution Methods

# states	Number Iterations				Seconds Per Iteration				Total Time (seconds)				Time Ratios	
$\|\mathcal{X}\|$	EE-v	EE-p	PF	VF	EE-v	EE-p	PF	VF	EE-v	EE-p	PF	VF	$\frac{PF}{EE\text{-}v}$	$\frac{VF}{EE\text{-}v}$
64	13	12.8	4.9	310.2	$6*10^{-4}$	$5*10^{-4}$	0.011	$2*10^{-4}$	0.008	0.006	0.053	0.069	7.0	8.6
486	13	12.2	5	276.5	0.005	0.007	0.604	0.003	0.060	0.081	3.018	0.800	50.3	13.3
2048	13	12	5	280.8	0.031	0.051	16.28	0.032	0.406	0.609	81.44	8.852	200.6	21.8
6250	13	12	5	288.8	0.270	0.477	150.0	0.318	3.515	5.723	750.0	91.79	213.4	26.1
15552	13	12	5	291.3	1.347	2.730	916.3	1.660	17.50	32.72	4581	483.66	261.7	27.6
$2*10^5$	13	13	5	298	21.8	22.0	-	22.1	284	290	$6.5*10^5$	$1.5*10^4$	2302	54.4

Note: EE-v = EE-value mapping iterations; EE-p = EE-probability mapping iterations; PF = Policy iterations;
VF = Value function (Bellman equation) iterations.

The comparison of the EE-value and the policy iteration methods illustrates that the computational gains come from the fact that a single iteration in the EE-value operator is much cheaper than one iteration in the policy iteration method. In fact, for larger state spaces, the total time required to solve the model using the EE-value operator is smaller than the time required for one iteration of the standard policy operator. The computational saving, per iteration, of EE-value relative to policy iteration comes from avoiding the computation of infinite-period forward present values. This computational saving, in CPU time, increases in a convex way with the dimension of the state space, and it becomes very substantial for large state spaces. For a state space with $|\mathcal{X}| = 15,552$ points, the CPU seconds for one EE-value iteration is 1.35 seconds, while one policy iteration requires 916.36 seconds. Note that this dimension of the state space is still quite small relative to the dimensions that we find in actual applications.

The EE-value and EE-prob methods share a well-known property of the standard policy iteration method (see Rust, 1996): the number of iterations to convergence is very stable with respect to the dimension of the state space. In this numerical example, the number of iterations of the PF, EE-value, and EE-prob methods remain roughly constant at 5, 13, and 12 iterations, respectively. This property has two important implications. First, for a large enough state space, the EE methods are computationally more efficient than the policy iteration method: their respective number of

iterations stay constant but the cost per iteration in the policy iteration method increases faster with $|\mathcal{X}|$ than the cost per iteration of the EE methods. A second implication has to do with the comparison between EE and value function (or Bellman equation) iteration methods. The cost per iteration of the EE methods is very similar to value function iterations because they do not compute present values but only one-period forward expectations. However, a key difference between EE and value function iterations is in the behavior of the number of iterations to convergence when the state space increases. For the method of value function iterations, it is well known that the number of iterations monotonically increases with the dimension of the state space. Given that the strength of the contraction stays fixed, for large state space value function iteration is substantially more expensive than the EE solution methods.

The number of iterations that a method 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. By definition, a fixed-point operator $\Gamma(.)$ is a contraction if there is some constant $c < 1$ such that, for any two points $\mathbf{V}$ and $\mathbf{W}$ in the domain of the operator, we have that $\|\Gamma(\mathbf{V}) - \Gamma(\mathbf{W})\|/\|\mathbf{V} - \mathbf{W}\| \leq c$ for some distance function $\|.\|$. The Lipschitz constant of mapping Γ is defined as the smallest constant c that satisfies this condition for any pair $\mathbf{V}$ and $\mathbf{W}$ in the domain of Γ. For instance, for the value function mapping Γ_{VF} and using the sup-norm (or uniform norm), the Lipschitz constant is defined as:

$$L(\Gamma_{VF}) \equiv \sup_{\mathbf{V},\mathbf{W}\in\mathbb{R}^{|\mathcal{X}|}} \left[\frac{\sup_{\mathbf{x}\in\mathcal{X}} |\Gamma_{VF}(\mathbf{x},\mathbf{V}) - \Gamma_{VF}(\mathbf{x},\mathbf{W})|}{\sup_{\mathbf{x}\in\mathcal{X}} |V(\mathbf{x}) - W(\mathbf{x})|} \right] \tag{47}$$

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 $\|\Gamma(\mathbf{V}) - \Gamma(\mathbf{W})\|/\|\mathbf{V} - \mathbf{W}\|$ only 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:

$$\widetilde{L}_{\mathbf{V}_0}(\Gamma) \equiv \max_{k\in\{0,1,...,\mathcal{I}_{\mathbf{V}_0}\}} \frac{\|\Gamma(\mathbf{V}_{k+1}) - \Gamma(\mathbf{V}_k)\|}{\|\mathbf{V}_{k+1} - \mathbf{V}_k\|} \tag{48}$$

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 3 reports the Lipschitz constants of the four mappings for the different dimensions of the state space. There are four results that we want to underline. First, the Lipschitz constants are very stable across the different dimensions of the state space. Second, there is a very substantial difference in the degree of contraction of the VF operator and the other three mappings. As shown in table 2, the weaker contraction property of the VF mapping is the reason why this method requires much larger number of iterations than the other methods. Third, the PF mapping has the smallest Lipschitz constant (i.e., stronger contraction). Again, this property is also reflected in table 2 in the small number of iterations that this method needs for convergence. Finally, it is worth to note that the Lipschitz constant for the EE-prob mapping is smaller than one, and in fact quite small. This result contrasts with our example in section 3.2.1 where we show that this mapping is not necessarily a contraction. Note there may be models and parameterizations where the EE-prob is a contraction mapping, and this seems the case of our entry-exit model. But note also that the EE-value operator is always a stronger contraction than the EE-prob.

Table 3
Lipschitz Constants of the mappings

Dimension $|\mathcal{X}|$	EE-value	EE-prob	Policy iter	Value iter
64	0.198	0.418	0.137	0.950
486	0.182	0.307	0.128	0.950
2032	0.181	0.240	0.107	0.950
6250	0.180	0.236	0.103	0.950
15552	0.180	0.248	0.094	0.950

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 table 2 (e.g., 15552 points), one single policy iteration takes almost one hundred times longer than the whole time to convergence of the EE-value method. In other words, even for this moderate dimension, the

optimal 'hybrid' algorithm is the pure value function iteration method.[13]

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 $|\mathcal{X}| = 6250$. 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.

We compute four estimators: (a) two-step PML-EE, as defined in equations (38) and (39); (b) sequential or K-step PML-EE, that consists in applying recursively the two-step estimator using updated choice probabilities, $\widehat{\mathbf{P}}_N^{(K)} = \Gamma_{EE,p}^{(N)}(\widehat{\theta}^{(K)}, \widehat{\mathbf{P}}_N^{(K-1)})$; (c) two-step PML-PF; and (d) MLE. Following Aguirregabiria and Mira (2002), we compute the MLE using the sequential (K-step) PML-PF algorithm. As shown in that paper, upon convergence the K-step estimator provides the Maximum Likelihood estimator. For the computation of the MLE, this algorithm is several orders of magnitude more efficient that the nested fixed point algorithm.[14]

First-step choice probabilities were estimated using a simple raw frequency estimator. For the number of steps of the recursive estimators, we set $K = 15$, but in almost all the $1,000$ Monte Carlo simulations we achieved convergence of the estimators with less that 15 iterations. 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 4 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

---

[13]For instance, note that when the state space has 15552 points, the VF method takes 483 seconds to converge while one single iteration of PF takes 916 seconds.

[14]The estimation results for the PML-EE estimators are based on the Euler equation for $y_{t+2} = 0$. We have also computed PML-EE estimators the Euler equation for $y_{t+2} = 1$. Although these two Euler equations are equivalent at the population level, in finite samples they generate different estimates. The reason is that first-step choice probabilities and structural parameters appear differently in these equations. Our Monte Carlo experiments show that the biased and mean square error of the two estimators are very similar. Here we report results only for the estimators that uses the Euler equation for $y_{t+2} = 0$.

Mira (2002), this estimator is asymptotically efficient (i.e., asymptotically equivalent to the MLE), but there is a well-known potential for finite sample bias with this class of two-step estimators due to the fact that it relies on potentially imprecise first step estimates of the conditional choice probabilities. However, for this Monte Carlo experiment, there are only marginal 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 ML 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 6378 CPU seconds. Therefore, in this experiment and within the class of PF pseudo maximum likelihood estimators, the two-step method seems the best estimator as we value both statistical and computational efficiency.

Second, compared to the (asymptotically efficient) two-step PML-PF estimator, the two-step Euler Equations estimator has more MAB (i.e., 26% greater) and RMSE (i.e., 25% greater). One could potentially ask, 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 bothering 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 Euler Equations estimator with either the two-step PML-PF or the ML estimators. As we see in the table, in terms of MAB and RMSE, the K-step EE estimator is superior to the two-step PML-PF estimator and it is essentially indistinguishable from the MLE. However the K-step EE estimator is thousands of times faster than either the two-step PML-PF or ML estimators, and there is no real trade-off to be made. The K-step EE estimator is statistically as good as the PF estimators, but more than 2000 times faster.

Table 4

Monte Carlo Experiments. Estimation of Structural Parameters

Sample size: $N = 1,000$ & $T = 2$; Monte Carlo samples $= 1,000$

Parameter (True value)	Mean Absolute Bias				Root Mean Squared Error			
	2-step PF	2-step EE	MLE	K-step EE	2-step PF	2-step EE	MLE	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 %)
$\theta_2^{VP}(-1)$	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 Ratio $\frac{MAB \text{ PF}}{MAB \text{ EE}}$	0.557 74.1 %	0.751	0.535 99.1 %	0.540				
Total RMSE Ratio $\frac{RMSE \text{ PF}}{RMSE \text{ EE}}$					0.703 75.1 %	0.935	0.680 99.4 %	0.684
Time (in secs) Ratio $\frac{Time \text{ PF}}{Time \text{ EE}}$	1067.80 4898 %	0.218	7345.79 2252 %	3.261	1067.80 4898 %	0.218	7345.79 2252 %	3.261

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

Overall, the experiment illustrates that the limit/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 the K-step EE estimator and the MLE estimator, 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.

5.4 Estimation of counterfactuals

We now study how standard methods and Euler Equation-based methods perform in answering an economically relevant counterfactual policy question. 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 would require from a few weeks (for value function iterations) to more than one year (for policy function iterations) of computing time. 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-value mapping.

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.

To formalize the argument, 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:

$$Q^* = \mathbb{E}\left(\sum_{i=1}^{N} a_{it} \, \exp\left(\omega_{it}\right)\right) = N \sum_{\mathbf{z},\omega} q(\mathbf{z},\omega) \, \exp\left(\omega\right) \, f^*(\mathbf{z},\omega) \tag{49}$$

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

$$q(\mathbf{z},\omega) = (1 - q(\mathbf{z},\omega)) \, P(0,\mathbf{z},\omega) + q(\mathbf{z},\omega) \, P(1,\mathbf{z},\omega) \tag{50}$$

Rearranging we get, $q(\mathbf{z},\omega) = P(0,\mathbf{z},\omega)/[1 - P(1,\mathbf{z},\omega) + P(0,\mathbf{z},\omega)]$. The effect of an increase in the entry cost on $q(\mathbf{z},\omega)$ is:

$$\frac{\partial q(\mathbf{z},\omega)}{\partial EC} = \frac{\frac{\partial P(0,\mathbf{z},\omega)}{\partial EC}(1 - P(1,\mathbf{z},\omega)) + \frac{\partial P(1,\mathbf{z},\omega)}{\partial EC}P(0,\mathbf{z},\omega)}{[1 - P(1,\mathbf{z},\omega) + P(0,\mathbf{z},\omega)]^2} \tag{51}$$

The derivative $\partial P(0,\mathbf{z},\omega)/\partial EC$ represents the effect of entry cost on the probability of entry for a potential entrant. This effect is negative. The derivative $\partial P(1,\mathbf{z},\omega)/\partial EC$ represents the effect of entry cost on the probability of staying in the market for an incumbent firm. This effect is positive. 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, $q(\mathbf{z},\omega)$. If increasing the entry cost causes disproportionately more productive incumbents to stay in the market, higher entry cost may actually increase output.[15]

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

---

[15]Specifically, the sign of $\frac{\partial^2 q(\mathbf{z},\omega)}{\partial EC \partial \omega}$ is also ambiguous. That is, it is not clear if an increase in the entry cost effects the steady state probability of observing a low productivity firms more than the steady state probability of observing a high productivity firm. In principle these effects could have opposite sign. This makes it unclear how *average* output depends on the entry cost. Although the number of firms increases, the new firms may be disproportionately less productive.

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., $200,000$ 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 $|\mathcal{Z}_1| = |\mathcal{Z}_2| = |\mathcal{Z}_3| = |\mathcal{Z}_4| = |\Omega| = 10$, such that the number of points in the complete state space is $|\mathcal{X}| = 2 * 10^5$. 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-value mapping. Using standard computing equipment, the total computing time was: 284 seconds (less than five minutes) for the EE-v method; 4.3 hours for the VF method; and 7.5 days for the PF method. Based on the ratios of computing times that we report in table 2, we have also extrapolated the computing times of these methods for a state space with $|\mathcal{X}| = 2 * 15^5$ points (approximately 1.5 million points): 1.4 hours for the EE-v method; 5.2 days for the VF method; and 549 days for the PF method.[16]

Table 5 presents predictions from the exact solution of the model under the factual and counterfactual scenarios. We calculate: (1) the probability of being active: $q^* = \sum_{\mathbf{z},\omega} q(\mathbf{z},\omega) \, f^*(\mathbf{z},\omega)$; (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: $\Pr(a_t = y_t) = \sum_{\mathbf{z},\omega}[q(z,\omega) \, P(1,\mathbf{z},\omega) + (1 - q(z,\omega)) \, (1 - P(0,\mathbf{z},\omega))] \, f^*(\mathbf{z},\omega)$; and (5) output per firm: $q^* = \sum_{\mathbf{z},\omega} q(\mathbf{z},\omega) \, \exp(\omega) \, f^*(\mathbf{z},\omega)$. Column (1) shows that, in this experiment, the effect of an increase in θ_0^{EC} from 1 to 2.5 is a net decrease in the probability of being active by 6.5 percentage points. The increase in entry cost causes a decrease in entry, partially offset by a decrease in exit which we see in columns (2) and (3). 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 output, from 0.529 to 0.423 units.

---

[16]We estimate this computing times by running a regression between the logarithm of the ratio of CPU time and the logarithm of the size of the state space using the numbers reported in Table 2. Then, we use the estimated regressions to predict CPU times for a state space with $2 * 15^5$ points. The estimated regression of log(CPU time PF/ CPU EE-v) on log(size) has intercept -0.4792 (s.e. $= 0.9067$) slope 0.6735 (s.e. $= 0.1240$), and R-square 0.912. The estimated regression of log(CPU time VF/ CPU EE-v) on log(size) has intercept 1.2343 (s.e. $= 0.1742$) slope 0.2263 (s.e. $= 0.0231$), and R-square 0.970.

Table 5
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%)

5.4.3 Estimation of counterfactuals

The dimension of the state space in this experiment, with $200,000$ 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 the same model and state space as in previous section 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.

To study the statistical properties of the different mappings, we simulate many sets of data, and compare the average (over the samples) answer to the counterfactual question delivered by each method. To make a fair comparison, we keep the time required to solve the fixed points roughly equal. 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 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-value 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 $N_{PF} < N_{EE}$ and $N_{VF} < N_{EE}$. 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-value 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 6 presents the number of points in the spaces $\mathcal{Z}_{N_{EE}}$ and $\mathcal{Z}_{N_{PF}}$ and the average (across the Monte Carlo simulations) computing times of the different methods.

Table 6
Monte Carlo Results: Average Computation Times

	VF	PF	EE-prob	EE-value
# 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

5.4.4 Monte Carlo results

We estimate the policy effects (1) to (5) that we describe in Table 5. We implement four different iterative methods: EE-prob, EE-value, Value function iterations, and Policy function iterations. We use 500 Monte Carlo simulations from the DGP. Table 7 presents the Root Mean Squared Error (RMSE) and the Mean Absolute Bias (MAB) based on the 500 Monte Carlo simulations. The statistical properties of the approximation methods based on the EE-prob and EE-value 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.

Table 7
Monte Carlo Results: Counterfactual Estimates
Effect of Entry Cost on Number of firms, Entry, Exit, 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[1]					
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-value 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[1]					
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-value 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.

6   Conclusion

In the context of dynamic discrete choice models, this paper derives optimality conditions that involve payoffs at only two consecutive periods. We describe these conditions as Euler equations. We show that these Euler Equations (EE) imply a fixed point mapping in the space of conditional choice value functions, that we denote as the EE-value operator. We show that this operator is a contraction. Based on these results, we propose new methods for solving dynamic discrete choice models and for estimating structural parameters and the effects of counterfactual policies in these models. The EE-value mapping only involves one period forward expectations and has strong contraction properties. These two properties imply very substantial computational savings relative to the standard methods for solving these models, i.e., value function iterations and policy function iterations.

We use Monte Carlo experiments to illustrate the relative computational gains associated with our approach. 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-value 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-value 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 $e(a, \mathbf{P}) \equiv e^{(v)}(a, \Lambda^{-1}(\mathbf{P})) = \mathbb{E}[\widetilde{\varepsilon}_a \mid \Lambda^{-1}(a, \mathbf{P}) + \widetilde{\varepsilon}_a \geq \Lambda^{-1}(j, \mathbf{P}) + \widetilde{\varepsilon}_j$ 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 $\widetilde{\mathbf{v}} = (\widetilde{v}_1, \widetilde{v}_2, ..., \widetilde{v}_J)'$,*

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

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

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

38

Therefore,

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

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 $\sum_{j=1}^{J} P_j \left[\widetilde{v}_j + e(j, \mathbf{P})\right]$ and taking into account that $e(j, \mathbf{P}) = e^{(v)}(j, \Lambda^{-1}(\mathbf{P}))$, we have that:

$$\frac{\partial W(\mathbf{P})}{\partial P_a} = \widetilde{v}_a + e(a, \mathbf{P}) + \sum_{j=1}^{J} P_j \frac{\partial e(j, \mathbf{P})}{\partial P_a}$$

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

$$= \widetilde{v}_a + e(a, \mathbf{P}) + \mathbf{P}' \left[\frac{\partial \mathbf{e}^{(v)}(\widetilde{\mathbf{v}})}{\partial \widetilde{\mathbf{v}}'}\bigg|_{\widetilde{\mathbf{v}}=\mathbf{\Lambda}^{-1}(\mathbf{P})}\right] \frac{\partial \Lambda^{-1}(\mathbf{P})}{\partial P_a}$$

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,

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

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

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

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

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

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:

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

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 $\partial W_t(\mathbf{P}_t^*, \mathbf{x}_t)/\partial\mathbf{P}_t =$

$\widetilde{\mathbf{v}}_t(\mathbf{x}_t) - \boldsymbol{\Lambda}^{-1}(\mathbf{P}_t^*) = 0$. By definition of the inverse mapping $\boldsymbol{\Lambda}^{-1}(\mathbf{P}_t)$, we have that $\mathbf{P}_t^*(\mathbf{x}_t) = \boldsymbol{\Lambda}(\widetilde{\mathbf{v}}_t(\mathbf{x}_t))$.

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:

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

The Jacobian matrix $\partial \boldsymbol{\Lambda}(\widetilde{\mathbf{v}})/\widetilde{\mathbf{v}}'$ 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:

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

Proof of Proposition 5. Using the definition of the Social Surplus function, we can represent the Euler Equation-Value mapping using the following expression:

$$\Gamma_{EE,v}(a, \mathbf{x}, \widetilde{\mathbf{v}}) = c(a, \mathbf{x}) + \beta \left[-\sum_{\mathbf{x}'} S\left(\widetilde{\mathbf{v}}(\mathbf{x}')\right) \left[f(\mathbf{x}'|a, \mathbf{x}) - f(\mathbf{x}'|0, \mathbf{x})\right]\right] \tag{A.23}$$

Using this expression and the weak contraction property of the Social Surplus function in Lemma 3, we can show the contraction property of the *EE-V mapping* $\mathbf{\Gamma}_{EE,v}$. For any $\widetilde{\mathbf{v}}$ and $\widetilde{\mathbf{w}}$ in $\mathcal{V}$,

$$
\begin{aligned}
\|\mathbf{\Gamma}_{EE,v}(\widetilde{\mathbf{v}}) - \mathbf{\Gamma}_{EE,v}(\widetilde{\mathbf{w}})\|_{\infty} &= \max_{a,\mathbf{x}} |\Gamma_{EE,v}(a,\mathbf{x},\widetilde{\mathbf{v}}) - \Gamma_{EE,v}(a,\mathbf{x},\widetilde{\mathbf{w}})| \\
&= \beta \max_{a,\mathbf{x}} \left| \sum_{\mathbf{x}'} [S(\widetilde{\mathbf{w}}(\mathbf{x}')) - S(\widetilde{\mathbf{v}}(\mathbf{x}'))] \; [f(\mathbf{x}'|a,\mathbf{x}) - f(\mathbf{x}'|0,\mathbf{x})] \right| \\
&\leq \beta \max_{\mathbf{x}'} |S(\widetilde{\mathbf{v}}(\mathbf{x}')) - S(\widetilde{\mathbf{w}}(\mathbf{x}'))| \\
&\leq \beta \max_{\mathbf{x}'} \|\widetilde{\mathbf{v}}(\mathbf{x}') - \widetilde{\mathbf{w}}(\mathbf{x}')\|_{\infty} \\
&= \beta \; \|\widetilde{\mathbf{v}} - \widetilde{\mathbf{w}}\|_{\infty} \qquad\qquad \blacksquare
\end{aligned}
\tag{A.24}
$$

Proof of Proposition 6.

(A) $\Gamma_{EE,v}^{(N)}(\widetilde{\mathbf{v}},\theta)$ **is a contraction.** This proof is a straightforward extension of the proof of proposition 5. 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 $\widetilde{v}$ and $\widetilde{w}$ in $\mathcal{V}_N$, using the same inequalities as in (A.24) but replacing $\mathbf{\Gamma}_{EE,v}$ with $\mathbf{\Gamma}_{EE,v}^{(N)}$, and f with f_N, we have that:

$$
\begin{aligned}
\left\| \mathbf{\Gamma}_{EE,v}^{(N)}(\widetilde{\mathbf{v}},\theta) - \mathbf{\Gamma}_{EE,v}^{(N)}(\widetilde{\mathbf{w}},\theta) \right\|_{\infty} &= \max_{a,\mathbf{x}} \left| \Gamma_{EE,v}^{(N)}(a,\mathbf{x},\widetilde{\mathbf{v}}) - \Gamma_{EE,v}^{(N)}(a,\mathbf{x},\widetilde{\mathbf{w}}) \right| \\
&= \beta \max_{a,\mathbf{x}} \left| \sum_{\mathbf{x}'} [S(\widetilde{\mathbf{w}}(\mathbf{x}')) - S(\widetilde{\mathbf{v}}(\mathbf{x}'))] \; [f_N(\mathbf{x}'|a,\mathbf{x}) - f_N(\mathbf{x}'|0,\mathbf{x})] \right| \\
&\leq \beta \; \|\widetilde{\mathbf{v}} - \widetilde{\mathbf{w}}\|_{\infty} \qquad\qquad \blacksquare
\end{aligned}
\tag{A.25}
$$

(B) $\Gamma_{EE,v}^{(N)}(\widetilde{\mathbf{v}},\theta)$ **converges uniformly to** $\mathbf{\Gamma}_{EE,v}(\widetilde{\mathbf{v}},\theta)$. We need to show that, for any value of $(a,\mathbf{x})$, $\sup_{\widetilde{\mathbf{v}},\theta} \left| \mathbf{\Gamma}_{EE,v}^{(N)}(a,\mathbf{x},\widetilde{\mathbf{v}},\theta) - \mathbf{\Gamma}_{EE,v}(a,\mathbf{x},\widetilde{\mathbf{v}},\theta) \right| \longrightarrow_p 0$. 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,

$$
\Pr\left(\left| \mathbf{\Gamma}_{EE,v}^{(N)}(a,\mathbf{x},\widetilde{\mathbf{v}},\theta) - \mathbf{\Gamma}_{EE,v}(a,\mathbf{x},\widetilde{\mathbf{v}},\theta) \right| > \varepsilon \right) < \delta
\tag{A.26}
$$

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 $\mathbf{\Gamma}_{EE,v}^{(N)}(\gamma) = \sum_{\mathbf{z}' \in \mathcal{Z}_N} h(\mathbf{z}',\gamma) \; f_{(N)}(\mathbf{z}')$, 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 $\mathbf{\Gamma}_{EE,v}^{(N)}(\gamma) = \sum_{\mathbf{z}' \in \mathcal{Z}} h(\mathbf{z}', \gamma) f_{(N)}(\mathbf{z}')$; 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,

$$\Pr \left(\sup_{\mathbf{z}' \in \mathcal{Z}} |f_{(N)}(\mathbf{z}') - f(\mathbf{z}')| > \varepsilon_f \right) < \delta_f \tag{A.27}$$

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

$$\begin{aligned}
\left| \mathbf{\Gamma}_{EE,v}^{(N)}(\gamma) - \mathbf{\Gamma}_{EE,v}(\gamma) \right| &= \left| \sum_{\mathbf{z}' \in \mathcal{Z}} \left[f_{(N)}(\mathbf{z}') - f(\mathbf{z}') \right] h(\mathbf{z}', \gamma) \right| \\
&\leq h^* \sum_{\mathbf{z}' \in \mathcal{Z}} \left| f_{(N)}(\mathbf{z}') - f(\mathbf{z}') \right| \\
&\leq h^* \sup_{\mathbf{z}' \in \mathcal{Z}} \left| f_{(N)}(\mathbf{z}') - f(\mathbf{z}') \right|
\end{aligned} \tag{A.28}$$

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

$$\begin{aligned}
\Pr \left(\left| \mathbf{\Gamma}_{EE,v}^{(N)}(\gamma) - \mathbf{\Gamma}_{EE,v}(\gamma) \right| > \varepsilon \right) &\leq \Pr \left(h^* \sup_{\mathbf{z}' \in \mathcal{Z}} \left| f_{(N)}(\mathbf{z}') - f(\mathbf{z}') \right| > \varepsilon \right) \\
&= \Pr \left(\sup_{\mathbf{z}' \in \mathcal{Z}} \left| f_{(N)}(\mathbf{z}') - f(\mathbf{z}') \right| > \frac{\varepsilon}{h^*} \right)
\end{aligned} \tag{A.29}$$

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 $\Pr \left(\left| \mathbf{\Gamma}_{EE,v}^{(N)}(\gamma) - \mathbf{\Gamma}_{EE,v}(\gamma) \right| > \varepsilon \right) < \delta$. ∎

Proof of Proposition 7

(A) Consistency. For notational simplicity, we omit now θ^* as an argument in functions $\mathbf{\Gamma}_{EE,v}(\widetilde{v}, \theta^*)$ and $\mathbf{\Gamma}_{EE,v}^{(N)}(\widetilde{v}, \theta^*)$. 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 $\widetilde{v}_N^* = \mathbf{\Gamma}_{EE,v}^{(N)}(\widetilde{v}_N^*)$. Given that, (a) $\mathbf{\Gamma}_{EE,v}(.)$ is a continuous function, and (b) $\mathbf{\Gamma}_{EE,v}^{(N)}$ 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:

$$\begin{aligned}
\widetilde{v}_N^* - \widetilde{v}^* &= \mathbf{\Gamma}_{EE,v}^{(N)}(\widetilde{v}_N^*) - \mathbf{\Gamma}_{EE,v}(\widetilde{v}^*) \\
&= \sum_{\mathbf{z} \in \mathcal{Z}} h(\mathbf{z}, \widetilde{v}_N^*) f_{(N)}(\mathbf{z}) - \sum_{\mathbf{z} \in \mathcal{Z}} h(\mathbf{z}, \widetilde{v}^*) f(\mathbf{z})
\end{aligned} \tag{A.30}$$

43

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:

$$h(\mathbf{z}, \widetilde{v}_N^*) = h(\mathbf{z}, \widetilde{v}^*) + \frac{\partial h(\mathbf{z}, \widetilde{v}^*)}{\partial \widetilde{v}'} (\widetilde{v}_N^* - \widetilde{v}^*) + o_p(1) \tag{A.31}$$

Solving this expression into (A.30) and using the consistency of $f_{(N)}$ (i.e., $f_{(N)}(\mathbf{z}) = f_{(N)}(\mathbf{z}) + o_p(1)$), we obtain:

$$\widetilde{v}_N^* - \widetilde{v}^* = \sum_{\mathbf{z} \in \mathcal{Z}} h(\mathbf{z}, \widetilde{v}^*) \left(f_{(N)}(\mathbf{z}) - f(\mathbf{z}) \right) + \left[\sum_{\mathbf{z} \in \mathcal{Z}} f(\mathbf{z}) \frac{\partial h(\mathbf{z}, \widetilde{v}^*)}{\partial \widetilde{v}'} \right] (\widetilde{v}_N^* - \widetilde{v}^*) + o_p(1) \tag{A.32}$$

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

$$\widetilde{v}_N^* - \widetilde{v}^* = [\mathbf{I} - \mathbf{D}]^{-1} \mathbf{H} \left(\mathbf{f}_{(N)} - \mathbf{f} \right) + o_p(1) \tag{A.33}$$

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:

$$\sqrt{N} \left(\mathbf{f}_{(N)} - \mathbf{f} \right) \to_d N \left(0, \ [\mathbf{I} - \mathbf{D}]^{-1} \mathbf{H} \, \mathbf{V}_f \, \mathbf{H}' \, [\mathbf{I} - \mathbf{D}']^{-1} \right) \qquad \blacksquare \tag{A.34}$$

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

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

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

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

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

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

[9] Cooper, R., J. Haltiwanger, and J. Willis (2010): "Euler Equation Estimation for Discrete Choice Models: A Capital Accumulation Application," NBER Working Paper 15675.

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

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

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

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

[14] 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.

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

[16] 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.

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

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

[19] 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.

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

[21] Pakes, A. (1994): "Dynamic structural models, problems and prospects," in C. Sims (ed.) Advances in Econometrics. Sixth World Congress, Cambridge University Press.

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

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

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

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

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

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

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

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

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

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