Practical Methods for Estimation of Dynamic Discrete Choice Models

Peter Arcidiacono[1] and Paul B. Ellickson[2]

[1]Department of Economics, Duke University, Durham, North Carolina 27708; email: psarcidi@econ.duke.edu

[2]Simon School of Business Administration, University of Rochester, Rochester, New York 14627; email: paul.ellickson@simon.rochester.edu

Annu. Rev. Econ. 2011. 3:363–94

First published online as a Review in Advance on March 29, 2011

The *Annual Review of Economics* is online at economics.annualreviews.org

This article's doi:
10.1146/annurev-economics-111809-125038

JEL codes: C14, C25, C61

1941-1383/11/0904-0363$20.00

Keywords

dynamic structural models, unobserved heterogeneity, conditional choice probabilities

Abstract

Many discrete decisions are made with an eye toward how they will affect future outcomes. Formulating and estimating the underlying models that generate these decisions is difficult. Conditional choice probability (CCP) estimators often provide simpler ways to estimate dynamic discrete choice problems. Recent work shows how to frame dynamic discrete choice problems in a way that is conducive to CCP estimation and demonstrates that CCP estimators can be adapted to handle rich patterns of unobserved state variables.

1. INTRODUCTION

Many discrete decisions are made with an eye toward their impact on future outcomes. Examples can be found in many areas of economics. In labor economics, choices over levels of education are in part driven by how these decisions affect future earnings. In industrial organization, firms pay an entry cost in the hopes of recouping their investment through future profit streams, keeping in mind how their rivals will respond to their actions, both now and in the future. In development economics, the decision to immigrate hinges on a person's beliefs regarding the future gains from doing so.

The effects of these dynamic decisions on particular outcomes can be analyzed using descriptive empirical methods that rely on randomization or quasi-randomization. Here the researcher may be interested in the effect of having a college degree on earnings. In this case, understanding exactly how the decision was made is not relevant except in how it forms the researcher's identification strategy. It is here that randomization, regression discontinuity, and natural experiments provide exogenous sources of variation in the data such that the predicted effect of a college degree on earnings is not being driven by problems of selection.

Structural models offer the opportunity to understand the decisions behind these descriptive results by formally modeling the dynamic discrete choice process. Although structural methods are often pitted against their descriptive counterparts, the two can serve as complements. At their best, structural models can replicate results obtained from randomized experiments or attempts to exploit quasi-randomization and can tell us how individuals will respond to counterfactual policies. Todd & Wolpin (2006) provide an excellent example. They analyze how PROGRESA, a policy intervention by the Mexican government that subsidized educational investments in children, affected human capital decisions using both structural and descriptive techniques. As the structural model can replicate the results from the program evaluation, more confidence can be placed in their counterfactual policy simulations.

The seminal papers of Miller (1984), Pakes (1986), Rust (1987), and Wolpin (1984) show that, under certain restrictions, estimating these dynamic discrete choice models is both feasible and important to answer key economic questions. Each paper exploits Bellman's representation of the dynamic discrete choice problem by breaking the payoff from a particular choice into the component received today and a future utility term that is constructed by assuming that optimal decisions will continue to be made in the future.

Nonetheless, modeling these dynamic decision processes is complicated, requiring calculations of the present discounted value of lifetime utility or profits across all possible choices. Formally modeling this decision process requires that one identify the optimal decision rule for each period and explicitly model expectations regarding future events. Recent surveys by Aguirregabiria & Mira (2010) and Keane et al. (2011) show the complications that arise in the formulation and estimation of dynamic discrete choice problems and provide overviews of the methods that exist to handle them.

In this article we review recent advances that dramatically reduce the computational burden of the structural approach, and that are often substantially easier to program, thereby lowering the barriers to entry that deter many researchers from entering the field. The estimators we discuss are not as efficient as full solution methods that solve the full dynamic programming (DP) problem. At the same time, these techniques open the doors to estimating models that would be computationally intractable or would require much stronger assumptions regarding how individuals form expectations about events far out in the future.

Having techniques that are broadly accessible is important for two reasons. First, it expands the supply of economists who can successfully tackle dynamic discrete choice problems. By highlighting classes of problems in which the entry costs are relatively low, we hope to broaden the set of individuals who will consider a structural approach, when the problem invites it. Second, the accessibility of structural models is important to build credibility among researchers outside the structural audience. Heckman (2010) argues that part of the Angrist & Pischke (2010) criticism of structural methods stems from a lack of transparency in the techniques and, due to the complications involved with estimation, few robustness checks. Working with classes of models that are as complicated as those at the frontier of the literature, but structured so that they are easier to estimate, permits the testing of alternative specifications and eases the burden of replication. Furthermore, although the models themselves can be quite complicated, their strong link to data makes clear what variation is driving the results.

The methods discussed here build on the seminal work of Hotz & Miller (1993). Their insight was that the data contain a wealth of information regarding individuals' expectations about the future. In particular, they show how the future utility term can sometimes be expressed as simple functions of the probabilities that particular choices occur, given the observed state variables. Using the data in this way forms the foundation of conditional choice probability (CCP) estimators.

The continued evolution of CCP techniques has produced several benefits. First, CCP-based estimators are easy to implement, often dramatically simplifying the computer programming required to take a model to data. Second, the computational simplicity of the methods makes it possible to handle both complex problems and rich specifications for unobserved state variables. Third, these methods make problems feasible that would otherwise be out of reach, including the estimation of dynamic games and nonstationary environments in which the full time horizon is not covered in the data and the researcher is unwilling to make assumptions regarding how expectations are formed outside the sample period.

The article proceeds as follows. In Section 2, we use the static discrete choice problem to motivate the structure of dynamic discrete choice. We describe the complications associated with full solution methods and discuss trade-offs researchers make with regard to the

CONTINUOUS TIME MODELS

Although the focus of this paper is on discrete-time treatments of dynamic decision problems, there is also a rich literature that uses continuous-time methods. These models have generally been centered on job search, beginning with the seminal paper of Flinn & Heckman (1982). Work then extended to equilibrium search models, such as Eckstein & Wolpin (1990), with recent contributions accommodating heterogeneity on both the firm and worker side, as well as bargaining (see, e.g., Postel-Vinay & Robin 2002). Continuous-time methods have also been applied to games (see Arcidiacono et al. 2010, Doraszelski & Judd 2010), in which the sequential arrival of moves or outcomes results in far more manageable state transitions than the discrete-time counterpart. This in turn makes it possible to solve for counterfactual equilibria in settings in which the simultaneous structure of discrete time would render the problem intractable.

distribution of the structural errors and how these errors enter the utility function. In Section 3, we show how the future utility component can be represented using CCPs, paying particular attention to the conditions that need to be satisfied for simple representations to result. In Section 4, we turn to estimation, including cases with persistent unobserved state variables. Section 5 extends the analysis to games, focusing in particular on models in which the future utility term depends only on a few CCPs. In Section 6, we examine cases in which models estimated via CCP techniques can be used to conduct counterfactual policy experiments without having to solve the full model. Section 7 concludes.

2. PRELIMINARIES

2.1. The Static Problem

We begin by considering a static discrete choice problem and then show how the problem changes when dynamics are added. Consider an individual who makes a decision d from a finite set of alternatives D. In the standard static discrete choice literature, the utility associated with each alternative is assumed to be the sum of two parts. The first component is a function of state variables, x, that are observed by both the individual and the econometrician. Denote this part of the payoff $u(x,d)$, often specified as a linear function of x and a parameter vector θ. The second component is a choice-specific variable $\epsilon(d)$ that is observed by the individual but not by the econometrician and has support on the real line. Let ϵ denote the vector of all choice-specific unobservables.

The individual then chooses the alternative that yields the highest utility by following the decision rule $\delta(x, \epsilon)$:

$$\delta(x, \epsilon) = \arg\max_{d \in D} \left[u(x, d) + \epsilon(d)\right]. \tag{1}$$

Because the econometrician only observes x, a distributional assumption is typically made on ϵ, whose probability density function (PDF) is then given by $g(\epsilon)$. The probability that the individual chooses d given x, $p(d\,|\,x)$, is found by integrating the decision rule over the regions of ϵ for which $\delta(x, \epsilon) = d$:

$$p(d\,|\,x) = \int I\left[\delta(x, \epsilon) = d\right] g(\epsilon) d\epsilon, \tag{2}$$

where I is the indicator function. Note that adding or multiplying all utilities by a constant will not change the probability that a given alternative is chosen. Thus we require normalizations for level and scale. These normalizations are typically satisfied by setting the observed portion of utility to zero for one of the choices and either normalizing the variance of ϵ to a positive constant or fixing one of the parameters in θ, in which case the variance of ϵ can be estimated.

2.1.1. Payoffs and beliefs in the dynamic problem.

With dynamic discrete choice models, individuals now make decisions in multiple time periods, taking into account how their decisions today impact the value of making subsequent decisions tomorrow.

In each period $t \in \{1, 2, \ldots, T\}$, $T \leq \infty$, the individual again makes a decision d_t from among a finite set of alternatives D_t. The immediate payoff, or flow utility, at period t from choice d_t is the same as the utility from the static problem above. It again depends on observed and unobserved state variables x_t and ϵ_t, which are now subscripted by time.

We again assume that utility is additively separable in the unobservables, with flow payoff $u(x_t, d_t) + \epsilon(d_t)$. Although the assumption of additive separability is made in much of the dynamic discrete choice literature, Keane & Wolpin (1994) and those who have adopted their framework estimate models in which the errors are not additively separable. The additive separability assumption is maintained here as it is crucial to utilize the methods emphasized in this article. The implications of this assumption are discussed throughout the rest of the review. Moreover, the utility function is assumed to be stable over time as there is no t subscript on the function itself.

Because individuals now account for the future impact of their decisions, we need to specify their beliefs over how the state variables transition. We assume that ϵ_t is independent and identically distributed (i.i.d.) over time, again specifying the PDF as $g(\epsilon_t)$. We assume x_t is Markov and denote the PDF of x_{t+1} conditional on x_t, d_t, and ϵ_t as $f(x_{t+1} \mid x_t, d_t, \epsilon_t)$. Here too the transitions on the state variables are not time specific, although this is relaxed below. To make the estimation problem tractable, we follow Rust (1987) in assuming conditional independence. After controlling for both the decision and observed state at t, the unobserved state at t has no effect on the observed state at $t+1$:

$$f(x_{t+1} \mid x_t, d_t, \epsilon_t) = f(x_{t+1} \mid x_t, d_t).$$

This assumption is standard in virtually all dynamic discrete choice papers, subject to the relaxations discussed in Section 4.2.

2.2. The Individual's Problem

We assume that individuals discount the future at a rate $\beta \in \{0, 1\}$, maximizing the present discounted value of their lifetime utilities. They do so by choosing δ^*, a set of decision rules for all possible realizations of the observed and unobserved variables in each time period, whose elements are denoted $\delta_t(x_t, \epsilon_t)$. These optimal decision rules are given by

$$\delta^* = \arg\max_{\delta} \arg E_\delta \left(\sum_{t=1}^{T} \beta^{t-1} \left[u(x_t, d_t) + \epsilon(d_t) \right] \mid x_1, \epsilon_1 \right), \tag{3}$$

where the expectations are taken over the future realizations of x and ϵ induced by δ^*.

In any period t, the individual's maximization problem can be decomposed into two parts: the utility received at t plus the discounted future utility from behaving optimally in the future. Hence the subset of decision rules in δ^* that cover periods t onward also solves

$$\max_{\delta} \left[u(x_t, d_t) + \epsilon(d_t) + E_\delta \left(\sum_{t'=t+1}^{T} \beta^{t'-t} \left[u(x_{t'}, d_{t'}) + \epsilon(d_{t'}) \right] \right) \right]. \tag{4}$$

2.3. Expressing the Future Utility Component

It is rarely practical to work with Equation 4 directly. Rather, we use the value function to express the future utility term in Equation 4 in a simpler fashion. The value function at time t, which represents the expected present discounted value of lifetime utility from following δ^*, given x_t and ϵ_t, can be written as

$$V_t(x_t, \epsilon_t) \equiv \max_{\delta} E_\delta \left(\sum_{t'=t}^{T} \beta^{t'-t} \left[u(x_{t'}, d_{t'}) + \epsilon(d_{t'}) \,|\, x_t, \epsilon_t \right] \right).$$

By Bellman's principle of optimality, the value function can also be defined recursively as follows:

$$\begin{aligned} V_t(x_t, \epsilon_t) &= \max_{d_t} \left\{ u(x_t, d_t) + \epsilon_t + \beta E\left[V_{t+1}(x_{t+1}, \epsilon_{t+1} \,|\, x_t, d_t) \right] \right\} \\ &= \sum_{d_t} I\left[\delta_t(x_t, \epsilon_t) = d_t \right] \left\{ u(x_t, d_t) + \epsilon(d_t) + \right. \\ &\qquad \left. \beta \iint \left[V_{t+1}(x_{t+1}, \epsilon_{t+1}) g(\epsilon_{t+1}) d\epsilon_{t+1} \right] f(x_{t+1} \,|\, x_t, d_t) dx_{t+1} \right\}. \end{aligned}$$

Because ϵ_t is unobserved, we further define the ex ante value function (or integrated value function), $\bar{V}_t(x_t)$, as the continuation value of being in state x_t just before ϵ_t is revealed. $\bar{V}_t(x_t)$ is then given by integrating $V_t(x_t, \epsilon_t)$ over ϵ_t,

$$\bar{V}_t(x_t) \equiv \int V_t(x_t, \epsilon_t) g(\epsilon_t) d\epsilon_t,$$

or, following the recursive structure, by

$$\begin{aligned} \bar{V}_t(x_t) = \sum_{d_t} \int I\left[\delta_t(x_t, \epsilon_t) = d_t \right] \left[u(x_t, d_t) + \epsilon(d_t) + \right. \\ \left. \beta \int \bar{V}_{t+1}(x_{t+1}) f(x_{t+1} \,|\, x_t, d_t) dx_{t+1} \right] g(\epsilon_t) d\epsilon_t. \end{aligned} \tag{5}$$

Note that the discounted term on the right-hand side of Equation 5 is the expected future utility associated with choice d_t given that the current state is x_t.

2.4. Dynamic Decisions and Probabilities

With the future value term in hand, we now define the conditional value function $v_t(x_t, d_t)$ as the present discounted value (net of ϵ_t) of choosing d_t and behaving optimally from period $t+1$ on

$$v_t(x_t, d_t) \equiv u(x_t, d_t) + \beta \int \bar{V}_{t+1}(x_{t+1}) f(x_{t+1} \,|\, x_t, d_t) dx_{t+1}.$$

The conditional value function is the key component to forming the CCPs, which are needed to form the likelihood of seeing the data. The individual's optimal decision rule at t solves

$$\delta_t(x_t, \epsilon_t) = \arg\max_{d_t} \left[v_t(x_t, d_t) + \epsilon_t) \right]. \tag{6}$$

Following the logic of the static discrete choice problem, the probability of observing d_t conditional on x_t, $p_t(d_t \,|\, x_t)$, is found by integrating out ϵ_t from the decision rule in Equation 6:

$$p_t(d_t \,|\, x_t) = \int I\{\delta_t(x_t, \epsilon_t) = d_t\} g(\epsilon_t) d\epsilon_t = \int I \left\{ \arg\max_{d_t \in D_t} \left[v_t(x_t, d_t) + \epsilon_t(d_t) \right] = d_t \right\} g(\epsilon_t) d\epsilon_t. \tag{7}$$

Therefore, conditional on being able to form $v_t(x_t)$, estimation is no different than the approach taken for static discrete choice models. The main difference between static and dynamic discrete choice is that, in the former, the payoffs will generally be expressed as linear functions of the state variables (because they are primitives), whereas in the latter, the expressions are more complicated (because they are constructed by solving the DP problem). How much more complicated will depend on the number of choices, the number of possible observed states, and the distribution of the structural errors.

2.5. Estimation Using Full Solution Methods

Standard methods for estimating dynamic discrete choice models involve forming likelihood functions derived from the CCPs given in Equation 7. The vast majority of the literature focuses on one of two cases: (*a*) finite-horizon models, in which the future value term is obtained via backward recursion, and (*b*) stationary infinite-horizon models, in which the value functions are computed using contraction mappings.

We now describe how to calculate the future value terms in both cases. For both the finite- and infinite-horizon cases, $u(x_t, d_t)$ is assumed to be a known function of x_t, d_t, and a parameter vector θ_1. As in static choice models, $u(x_t, d_t)$ must be normalized to zero for one of the options. Estimation of the structural model also requires the estimation of the transition functions governing the observed states, $f(x_{t+1} \mid x_t, \epsilon_t)$, typically parameterized by a vector θ_2. For the purposes of calculating the future value term, the parameters governing these transitions are then treated as known, as is the distribution of the unobservables, $g(\epsilon)$. With the future value term in place, we can then show how to form the log likelihood function and explain the computational burden associated with the full solution estimation.

2.5.1. Finite-horizon value functions. In finite-horizon problems, the decision in the last period, T, is static, so the conditional value function at T is just the flow payoff function:[1]

$$v_T(x_T, d_T) = u(x_T, d_T).$$

With $u(x_T, d_T)$ parameterized by θ_1, knowledge of $g(\epsilon_T)$ implies that we can (either analytically or numerically) calculate the ex ante value function at T using

$$\bar{V}_T(x_T) = \int \sum_{d_T} I[\delta_T(x_T, \epsilon_T) = d_T][v_T(x_T, d_T) + \epsilon_T(d_T)]g(\epsilon_T)d\epsilon_T.$$

The conditional value function at $T-1$ is then given by

$$v_{T-1}(x_{T-1}, d_{T-1}) = u(x_{T-1}, d_{T-1}) + \beta \int \bar{V}_T(x_T) f(x_T \mid x_{T-1}, d_{T-1}) dx_T.$$

Continuing the backward recursion, once we have $v_{T-1}(x_{T-1}, d_{T-1})$, we can form the future value term for $T-2$, and so on. At time t, the conditional value function is then given by

$$v_t(x_t, d_t) = u(x_t, d_t) + \beta \int \bar{V}_{t+1}(x_{t+1}) f(x_{t+1} \mid x_t, d_t) dx_{t+1}. \tag{8}$$

[1]Depending on the time horizon, the terminal payoff function may be specified differently here than at other points in time to approximate the value of future decisions.

Moving back each period requires integrating out over both the observed and unobserved state variables, which can cause computational time to increase markedly as the size of the state space grows large or the time horizon becomes long. For example, in the calculation of $v_t(x_t, d_t)$, the future value component requires integrating out over the value of $\{x_{t+1}, \ldots, x_T\}$ and $\{\epsilon_{t+1}, \ldots, \epsilon_T\}$.

In some cases, such as when the error distribution leads to a closed-form solution for the ex ante value function and when the number of observable states is small, the backward recursion is simply a sum that can be evaluated directly. However, when the ex ante value function does not have a closed form or the number of states is large, we could instead evaluate the future utility terms at a subset of the possible observed and unobserved states. At states for which the value function is not calculated directly, the value function can be interpolated using the values at the points at which the function was calculated. This is the method used by Keane & Wolpin (1994), and it does not require the flow utility errors to be additively separable. For example, at $T-1$, choose a finite subset of x_T and ϵ_T at which to evaluate $V_T(x_T, \epsilon_T)$. Then fit a relationship between the calculated values of $V_T(x_T, \epsilon_T)$ and a flexible function of x_T and ϵ_T and use these values to interpolate $V_T(x_T, \epsilon_T)$ over the rest of the state space. Taking $V_T(x_T, \epsilon_T)$ as given, move backward and repeat the steps for $T-1$. Keane et al. (2011) provide a review of the trade-offs of the different evaluation methods when the state space is large.

2.5.2. Infinite-horizon value functions.

When the horizon is infinite but the environment is stationary, we can remove the t subscripts from both the conditional and ex ante value functions. Denoting the current value of the observed state variable by x and the next period's state variable by x', the conditional value function becomes

$$\begin{aligned} v(x,d) &= u(x,d) + \beta \int \bar{V}(x') f(x' \mid x, d) dx' \\ &= u(x,d) + \beta \int \left[\int \left\{ \max_{d' \in D} \left[v(x', d') + \epsilon(d') \right] \right\} f(x' \mid x, d) dx' \right] g(\epsilon) d\epsilon. \end{aligned} \tag{9}$$

To compute the future value term, note that the ex ante value function can be expressed as

$$\bar{V}(x) = \int \left\{ \max_{d} \left[u(x,d) + \beta \int \bar{V}(x') f(x' \mid x, d) dx' \right] \right\} g(\epsilon) d\epsilon. \tag{10}$$

For example, by assuming a discrete support for the observed states, we can form an expression (Equation 10) for each value of x and simply stack the equations. Rust (1987) establishes that this set of equations has a unique fixed point, which can be found by guessing the values for $\bar{V}(x)$ and substituting these values into the right-hand side of Equation 10 to update these guesses. This process is repeated until the difference between the $\bar{V}(x)$'s across iterations is sufficiently small.

2.5.3. Forming the likelihood.

For both finite- and infinite-horizon problems, the log likelihood function is formed by calculating the probabilities of the decisions observed in the data. Let d_{nt}, x_{nt}, and ϵ_{nt} indicate the choice, observed state, and unobserved state at time t for individual n, respectively. With the flow payoff of a particular decision parameterized by the vector θ_1, $u(x_{nt}, d_t, \theta_1)$, and the transitions on the observed states parameterized by a vector θ_2, the conditional value functions, decision rules, and choice probabilities will also depend on $\theta \equiv \{\theta_1, \theta_2\}$. The likelihood contribution of the choice for individual n at time t is then given by

$$\begin{aligned} p_t(d_{nt} \mid x_{nt}, \theta) &= \int I(\delta(x_{nt}, \epsilon_{nt}, \theta) = d_{nt}) g(\epsilon_{nt}) d\epsilon_{nt} \\ &= \int I\left\{ \arg\max_{d_t} [v_t(x_{nt}, d_t, \theta) + \epsilon_{nt}(d_t)] = d_{nt} \right\} g(\epsilon_{nt}) d\epsilon_{nt}, \end{aligned}$$

where, in the stationary infinite-horizon case, the t subscripts on $p_t(d_{nt} \mid x_{nt}, \theta)$ and $v_t(x_{nt}, d_t, \theta)$ are removed.

Forming these probabilities for each individual and each time period yields the components necessary for maximum likelihood estimation. With $\mathcal{N}$ individuals for $\mathcal{T}$ periods, estimates of θ_1 and θ_2 are obtained via

$$\hat{\theta} = \arg\max_{\theta} \sum_{n=1}^{\mathcal{N}} \sum_{t=1}^{\mathcal{T}} \left(\ln [p_t(d_{nt} \mid x_{nt}, \theta)] + \ln [f(x_{nt+1} \mid x_{nt}, d_{nt}, \theta_2)] \right).$$

Note that the log likelihood function is the sum of two components: one associated with the choices and one associated with the state transitions. Because the log likelihood function is additively separable, a consistent estimate of θ_2 can be obtained using information on the state transitions alone. Then, taking $\hat{\theta}_2$ as given, the data on the choice probabilities can be used to estimate θ_1. Although this method is not efficient (as the choice probabilities also contain information on θ_2), it can result in substantial computational savings.

Obtaining $\hat{\theta}_1$ requires that one evaluate the future value term of the dynamic discrete choice problem (at each candidate θ_1) either by backward recursion in the finite-horizon case (see Section 2.5.1) or by solving a fixed-point problem (see Section 2.5.2). Most of the burden of estimating dynamic discrete choice models stems from having to repeatedly solve the DP problem. The methods introduced below avoid its solution entirely.

2.6. Choices Over the Error Distribution

The choice of the distribution of the structural errors, $g(\epsilon)$, will also affect how costly it is to evaluate the future value terms. Researchers face trade-offs when choosing between distributions that allow for more flexible correlation patterns and distributions in which both the ex ante value function and the probabilities of making particular decisions have a closed form. In practice, the relevant choice is typically between the multivariate normal and the generalized extreme value (GEV) distributions.

2.6.1. Normal errors. There are two advantages to choosing a normal distribution. First, it has a more flexible correlation structure and is therefore able to capture richer patterns of substitution across choices. The GEV class requires that one specify the pattern (but not the magnitude) of the correlations a priori, and it may not always be clear how to form a GEV distribution that captures the flexibility that researchers would like. However, Bresnahan et al. (1997) show how to accommodate errors that are correlated across multiple nests, so flexible correlation patterns are certainly possible. Nonetheless, most dynamic discrete choice models that employ GEV errors choose the type I extreme value distribution.[2] Second, it is easy to draw from a normal distribution. This is only true in the GEV case in which the errors follow a type I extreme value distribution, implying multinomial logit expressions for the choice probabilities. Indeed, a whole article in *Econometric Theory* (Cardell 1997) is dedicated to showing how to draw errors that generate nested logit probabilities.

2.6.2. Generalized extreme value errors. Alternatively, as originally established by McFadden (1978), the GEV distribution delivers several advantages when working with discrete choice problems. First, there are closed-form expressions for the choice probabilities, $p_t(x_t)$, easing the computational burden of any estimation approach that employs this distribution. Second, as noted by Rust (1987), the expectations of future utility conditional on the states also admit a closed-form solution. Consider the case in which the errors follow a type I extreme value distribution, yielding the popular dynamic logit model. The probability of an arbitrary choice d_t and the ex ante value function are then given by

$$p_t(d_t \mid x_t) = \frac{\exp[v_t(x_t, d_t)]}{\sum_{d'_t \in D} \exp[v_t(x_t, d'_t)]} = \frac{1}{\sum_{d'_t D} \exp[v_t(x_t, d'_t) - v_t(x_t, d_t)]}, \tag{11}$$

$$\bar{V}_t(x_t) = \ln\left\{ \sum_{d'_t \in D} \exp[v_t(x_t, d'_t)] \right\} + \gamma, \tag{12}$$

where γ is Euler's constant. This closed-form representation of the value function is a huge advantage in estimation as, without it, numerical integration over the structural errors would need to take place for every future value term, substantially slowing computation. Clearly, this advantage will be shared by any estimation method that exploits GEV distributional assumptions.

The third advantage of working with the GEV distribution concerns the mapping from choice probabilities back to ex ante value functions, first explored by Hotz & Miller (1993). Given (*a*) structural errors that are additively separable from the flow payoffs, (*b*) conditional independence of the state transitions, and (*c*) independence of the structural errors over time, Hotz & Miller proved that differences in conditional value functions can always be expressed as functions of the choice probabilities alone. Moreover, the relationship between $p_t(d_t \mid x_t)$ and $\bar{V}_t(x_t)$ can be quite simple in the GEV case. Arcidiacono & Miller (2010) derive the representations for the GEV case, which admit analytic solutions when the errors lead to nested logit probabilities and can be solved numerically otherwise.

[2]As noted by Rust (1994b), the dynamic logit model inherits the main computational benefit of the extreme value specification (i.e., closed-form solutions) without suffering from its main drawback in the static choice setting (the independence from irrelevant alternatives). This is because the choice probabilities in the dynamic logit model depend on differences in choice-specific value functions, rather than static utilities, and these value functions will generally include the characteristics of all the alternatives.

In the dynamic logit example, the ex ante value function given in Equation 12 can be rewritten with respect to the conditional value function associated with an arbitrarily selected choice, say d_t^*:

$$\begin{aligned}\bar{V}_t(x_t) &= \ln\left(\exp[v_t(x_t, d_t^*)]\left\{\frac{\sum_{d_t'\in D}\exp[v_t(x_t, d_t')]}{\exp[v_t(x_t, d_t^*)]}\right\}\right) + \gamma \\ &= \ln\left\{\sum_{d_t'\in D}\exp\left[v_t(x_t, d_t') - v_t(x_t, d_t^*)\right]\right\} + v_t(x_t, d_t^*) + \gamma \\ &= -\ln[p(d_t^* \mid x_t)] + v_t(x_t, d_t^*) + \gamma.\end{aligned} \tag{13}$$

The last equality has an intuitive interpretation: The ex ante value of being in state x_t can be expressed as the sum of the conditional value from making an arbitrary choice d_t^* $[v_t(x_t, d_t^*)]$, the mean of the type I extreme value distribution (γ), and a nonnegative adjustment term ($-\ln[p(d_t^* \mid x_t)]$), which adjusts for the possibility that d_t^* may not be the optimal choice. Notice that as the probability of selecting d_t^* goes to one, the adjustment term goes to zero.

2.7. Conditional Choice Probabilities in the Dynamic Problem

Of course, to construct the choice probabilities in Equation 7 and form the likelihood, we need to construct a comparable expression for $v_t(x_t, d_t)$. Using Equation 13 to substitute for the ex ante value function (at $t+1$) in Equation 8, we can now write

$$v_t(x_t, d_t) = u(x_t, d_t) + \beta\int(v_{t+1}(x_{t+1}, d_{t+1}^*) - \ln[p_{t+1}(d_{t+1}^* \mid x_{t+1})])f(x_{t+1} \mid x_t, d_t)dx_{t+1} + \beta\gamma, \tag{14}$$

where d_{t+1}^* is an arbitrary choice in period $t+1$. The future value term now has three components: the function characterizing the transitions of the state variables, the CCPs for the arbitrary choice d_{t+1}^*, and the conditional value function associated with d_{t+1}^*. The first two can often be estimated separately in a first stage. It still remains to be shown how to deal with the remaining conditional value function.

However, absent this last issue, estimation is now quite simple. As discussed in more detail in Section 4.1, the transitions on the state variables and consistent estimates of the CCPs can be recovered in a first stage. Then, taking these as given, estimation reduces to a static multinomial logit criterion function with a precalculated offset term. At this point, the dynamic discrete choice problem can be estimated using standard statistical software (e.g., Stata).

Using the arguments of Hotz & Miller (1993), Altug & Miller (1998), and Arcidiacono & Miller (2010), we now show that, for a large class of problems, dealing with the remaining conditional value function is surprisingly straightforward. The key to the argument is that the researcher can choose to which choice (and hence which conditional value function) to make the future value term relative. Because discrete choice models are estimated using differences in conditional value functions, in many cases a clever choice of conditional value function for this benchmark can allow the difference in the future utility terms across two choices to be characterized by the one-period-ahead CCPs alone.

3. CONDITIONAL CHOICE PROBABILITY REPRESENTATIONS

3.1. Models Requiring Only One-Period-Ahead Choice Probabilities

We begin with examples in which only one-period-ahead choice probabilities are needed for estimation. To illustrate, suppose the choice set D_t includes an action that, when taken, implies that no further decisions are made. Hotz & Miller (1993) use the example of sterilization in a dynamic model of fertility choices. This structure is also shared by many optimal stopping problems, including classic models of search (McCall 1970), modern treatments of durable demand without replacement (Melnikov 2001, Nair 2007), and dynamic discrete games with permanent exit (Ericson & Pakes 1995). The central feature is that once the terminal action is chosen, the agent's decision problem is no longer dynamic, allowing the future value term to be replaced with a known parametric form (or normalized to zero). To see how this works, let $d = R$ indicate the terminal choice. Expressing the future value term relative to choice R, Equation 14 becomes

$$v_t(x_t, d_t) = u(x_t, d_t) + \beta \int (v_{t+1}(x_{t+1}, R) - \ln [p_{t+1}(R \mid x_{t+1})]) f(x_{t+1} \mid x_t, d_t) dx_{t+1} + \beta\gamma.$$

The key is to realize that the conditional value function for choice R does not have a future value component because choosing R terminates the dynamic decision process. The expression for $v_{t+1}(x_{t+1}, R)$ is then just a component of static utility, which is typically assumed to follow a known parametric form or normalize to zero.

Researchers have used this terminal choice property in a number of empirical applications. Hotz & Miller (1993) examine fertility choices, assuming that once the individual chooses sterilization, no further fertility decisions are made. Joensen (2009) examines the choice of whether to continue one's education and, if so, how much to work. In her case the decision to drop out of school terminates the dynamic decision problem. Murphy (2010) examines the choice to develop a parcel of land, treating development as a terminal choice. Finally, as discussed below, dynamic games with permanent exit also share the terminal choice property.

Another class of models that only require one-period-ahead CCPs are settings in which there exists a choice that makes the choice in the previous period irrelevant. An example is the capital replacement problem in Rust (1987), which is formulated as a regenerative optimal stopping problem. In Rust's specification, the value of a new engine does not depend on the mileage that accumulated on the old engine. Engine replacement effectively resets the clock. Again, this renewal structure applies in other contexts as well, including models of durable demand with replacement (Gowrisankaran & Rysman 2009) and certain inventory problems.

To formalize this argument, label the renewal choice as R. By taking the renewal action at $t + 1$, the effect of the choice at t on the state at $t + 2$ is removed so that

$$\int f(x_{t+1} \mid x_t, d_t^*) f(x_{t+2} \mid x_{t+1}, R) dx_{t+1} = \int f(x_{t+1} \mid x_t, d_t') f(x_{t+2} \mid x_{t+1}, R) dx_{t+1} \qquad (15)$$

holds for all $\{d_t^*, d_t'\}$ and x_{t+2}. To see how the renewal property can be exploited in estimation, recall Equation 14. Substitute in for $v_{t+1}(x_{t+1}, R)$ with the flow payoff of replacement plus the ex ante value function at $t + 2$:

$$
\begin{aligned}
v_t(x_t, d_t) &= u(x_t, d_t) + \beta \int (v_{t+1}(x_{t+1}, R) - \ln [p_{t+1}(R \,|\, x_{t+1})]) f(x_{t+1} \,|\, x_t, d_t) dx_{t+1} + \beta\gamma \\
&= u(x_t, d_t) + \beta \int (u(x_{t+1}, R) - \ln [p_{t+1}(R \,|\, x_{t+1})]) f(x_{t+1} \,|\, x_t, d_t) dx_{t+1} + \beta\gamma \\
&\quad + \beta^2 \iint V_{t+2}(x_{t+2}) f(x_{t+2} \,|\, x_{t+1}, R) f(x_{t+1} \,|\, x_t, d_t) dx_{t+2} dx_{t+1}.
\end{aligned} \tag{16}
$$

The last term, which represents the continuation value at time $t+2$, conditional on choosing the renewal action at time $t+1$, is constant across all choices made at time t.[3] Because discrete choice estimation works with differences in expected payoffs (here, conditional value functions), this last term will drop out of the likelihood, again leaving only expressions that can either be estimated in a prior stage [e.g., $p_t(d_t \,|\, x_t)$ and $f(x_{t+1} \,|\, x_t, d_t)$] or include at most one set of flow payoff parameters. Note that the renewal action does not destroy the dependence of the state on previous states, but destroys only the dependence of the state on the previous choice. For example, in the bus engine problem, there may be a state that affects whether the engine will be replaced but that is not affected by the replacement choice itself. If the price of bus engines falls and these lower prices persist, this will affect the probability of replacement. At the same time, there is still no effect of past replacement choices on the future value of a new engine conditional on the knowledge that we are in a low-price period.

3.2. Models Requiring Multiple-Period-Ahead Choice Probabilities

In certain choice settings, it might take more than just a single action at time t to reset the system. In some cases, the same renewal action might need to be repeated for a fixed number of periods. In other settings, it might require a particular sequence of actions.

Altug & Miller (1998) consider the first case, focusing on the example of female labor supply with human capital accumulation and depreciation. In their model, a woman who stays out of the workforce for ρ consecutive periods effectively resets her human capital (i.e., there is full depreciation if she stays out of the labor force for ρ periods). Hence expressing the future value term for every conditional value function relative to choosing the nonwork option for ρ consecutive periods results in the same level of human capital once the decision sequences are complete, regardless of the initial choice. Consequently, only ρ-period-ahead CCPs are needed in estimation. A similar structure arises in many marketing models of state-dependent demand. In these models, consumers face a switching cost when choosing a different product than the one they consumed last period, but this cost is reset to zero whenever the consumer chooses the outside good.

In both Altug & Miller's (1998) framework and the renewal examples, there is an action that, when taken at $t+1$ or taken from $t+1$ to $t+\rho$, can undo the dependence of the state on the choice at t. Arcidiacono & Miller (2010) generalize this concept, noting that future utility terms can be expressed relative to the value of any sequence of choices. When sequences of choices exist such that, given different initial choices, the same state results in expectation, only CCPs during the time of the sequence are needed in estimation.

[3] In Rust's (1987) empirical specification, the argument is even simpler as normalizing the baseline utility of replacing the bus engine to zero for all mileage states (which is required as the only covariate in the utility/cost function is mileage) implies that $v_{t+1}(x_{t+1}, R) = v_{t+1}(x'_{t+1}, R)$ for all $\{x_{t+1}, x'_{t+1}\}$.

To see this, it is first convenient to give an expression for the cumulative probability of being in a particular state given a particular decision sequence and initial state. Consider an individual in state x_t and a candidate sequence of decisions from t to $t+\rho$ periods: $\{d_t^*, d_{t+1}^*, \ldots, d_{t+\rho}^*\}$. For $\tau \in \{t, \ldots, t+\rho\}$, denote $\kappa_\tau^*(x_{\tau+1} | x_t)$ as the cumulative probability of being in state $x_{\tau+1}$ given the decision sequence and initial state, defined recursively as

$$\kappa_\tau^*(x_{\tau+1} \,|\, x_t) = \begin{cases} f(x_{\tau+1} \,|\, x_\tau, d_\tau^*) & \text{if } \tau = t \\ \int f(x_{\tau+1} \,|\, x_\tau, d_\tau^*)\kappa_{\tau-1}^*(x_\tau \,|\, x_t)dx_\tau & \text{otherwise.} \end{cases} \tag{17}$$

Applying the arguments of Arcidiacono & Miller (2010), we can rewrite the expression for $v_t(x_t, d_t)$ given in Equation 8 such that the future utility term is expressed relative to the choices in the sequence $\{d_t^*, d_{t+1}^*, \ldots, d_{t+\rho}^*\}$. Namely, we can substitute for the value function in Equation 8 with the one-period-ahead expression in Equation 13. Next, substitute in for $v_{t+1}(x_{t+1}, d_{t+1}^*)$ using the one-period-ahead expression for Equation 8. Continually repeating for ρ periods, each time expressing the value function relative to the next choice in the sequence yields

$$\begin{aligned} v_t(x_t, d_t^*) = u(x_t, d_t^*) &+ \sum_{\tau=t+1}^{t+\rho} \int \beta^{\tau-t}\big[u(x_\tau, d_\tau^*) - \ln[p_\tau(x_\tau, d_\tau^*)] + \gamma\big]\kappa_{\tau-1}^*(x_\tau \,|\, x_t)dx_\tau \\ &+ \int \beta^{\rho+1} V_{t+\rho+1}(x_{t+\rho+1})\kappa_{t+\rho}^*(x_{t+\rho+1} \,|\, x_t)dx_{t+\rho+1}. \end{aligned} \tag{18}$$

Once again, the representation has an intuitive interpretation. The first and third terms, which are standard, represent the flow utility associated with choice d_t^* and the continuation value that will be obtained ρ periods in the future. The second term collects the flow utilities accrued over the ρ-period sequence along with a term that compensates for the possibility that the imposed sequence may not be optimal given the draws of the ϵ's.

Now consider an alternative sequence of decisions $\{d_t', d_{t+1}', \ldots, d_{t+\rho}'\}$. Define $\kappa_\tau'(x_{t+1} | x_t)$ as the cumulative probability of x_{t+1} given this alternative sequence, defined recursively by replacing d_τ^* with d_τ' in the right-hand side of Equation 17. Suppose these sequences of decisions lead the individual to the same state in expectation, in which case

$$\kappa_{t+\rho}^*(x_{t+\rho+1} \,|\, x_t) = \kappa_{t+\rho}'(x_{t+\rho+1} \,|\, x_t) \tag{19}$$

holds for all $x_{t+\rho+1}$.

Arcidiacono & Miller (2010) state that two choices exhibit ρ-period dependence if sequences exist following each of these choices such that Equation 19 holds. Forming the expression for $v_t(x_t, d_t')$ in which the future value term is expressed relative to the choices $\{d_{t+1}', \ldots, d_\rho'\}$ and subtracting this expression from Equation 18 yields

$$\begin{aligned} v_t(x_t, d_t^*) - v_t(x_t, d_t') = u(x_t, d_t^*) &- u(x_t, d_t') \\ &+ \sum_{\tau=t+1}^{t+\rho} \int \beta^{\tau-t}(u(x_\tau, d_\tau^*) - \ln[p_\tau(x_\tau, d_\tau^*)])\kappa_{\tau-1}^*(x_\tau \,|\, x_t)dx_\tau \\ &- \sum_{\tau=t+1}^{t+\rho} \int \beta^{\tau-t}(u(x_\tau, d_\tau') - \ln[p_\tau(x_\tau, d_\tau')])\kappa_{\tau-1}'(x_\tau \,|\, x_t)dx_\tau. \end{aligned} \tag{20}$$

Note that the last line of Equation 18 disappears because of finite dependence. Namely, because the two choice sequences lead to the same states in expectation, the last line of Equation 18 will be the same in the expression for $v_t(x_t, d'_t)$.[4]

In estimation, we always work with differences in conditional value functions. In particular, we can express the probability of making any choice as a function of the conditional value functions differenced with respect to the conditional value function associated with a baseline, or anchor, choice. Denote this anchor choice as A, which is equivalent to d'_t in the discussion above. When the structural errors are distributed type 1 extreme value, for example, the probability of an arbitrary choice d^*_t can be written as

$$p_t(d^*_t \mid x_t) = \frac{\exp\left[v_t(x_t, d^*_t)\right]}{\sum_{d_t} \exp\left[v_t(x_t, d_t)\right]} = \frac{\exp\left[v_t(x_t, d^*_t) - v_t(x_t, A)\right]}{\sum_{d_t} \exp\left[v_t(x_t, d_t) - v_t(x_t, A)\right]}.$$

When calculating $v_t(x_t, d_t) - v_t(x_t, A)$, for each d_t we want to express $v_t(x_t, A)$ such that, when differences are taken as in Equation 20, the last term in Equation 18 cancels out. Hence finite dependence must hold for each possible choice d_t when compared with the anchor choice A. Although using finite dependence in estimation may seem restrictive, the structure economists place on dynamic discrete choice models often leads to these requirements holding.

3.3. Example: Occupational Choice

To demonstrate how finite dependence applies to some more complex problems in the literature, we consider a simplified version of Keane & Wolpin's (1997) model of career decisions. Keane & Wolpin estimate a dynamic model of human capital accumulation in which individuals choose each period whether to obtain more education, stay at home, or work in one of three occupations (blue collar, white collar, or military service). A key feature of their model is that individuals invest more in human capital at earlier ages, either by accruing more experience in one of the occupations or by obtaining more education, with an eye toward increasing their future wages.

We focus here on a simpler setting in which, at time t, an individual decides whether to stay home, $d_t = H$, or work in either the blue- or white-collar occupation, $d_t = B$ and $d_t = W$, respectively.[5] The per-period utility function for working in either occupation depends on the occupation's wage as well as whether the individual worked in that occupation in the previous period. This latter variable can be interpreted as an occupation-switching cost. We discuss the form of the payoff function below, focusing first on how certain choice sequences result in finite dependence.

The effect on the future of the choice today occurs partly through human capital accumulation, which in turn affects future wages. The choice to work in the blue- (white-) collar sector increases the individual's blue- (white-) collar experience by one unit and also turns on an indicator for whether the most recent decision was to work in that sector.

[4]In situations in which ρ or the state space is sufficiently large (e.g., games), these integrals may be difficult (or impossible) to evaluate analytically. In such cases, forward simulation is an obvious alternative. This is the approach taken by Bishop (2008) and is discussed in more detail in Section 4.1.4.

[5]Finite dependence will still hold if the education and military options are included.

Denote h_{Bt} and h_{Wt} as the number of periods of experience in the blue- and white-collar sectors, respectively. To capture the occupation switching cost, denote $d_{Bt} = 1$ if $d_t = B$ and $d_{Bt} = 0$ otherwise (and likewise for W). Given a choice at t, the observed state variables at $t+1$ are given by

$$x_{t+1} = [\, h_{Bt} + d_{Bt} \quad d_{Bt} \quad h_{Wt} + d_{Wt} \quad d_{Wt} \,].$$

As we are going to be working with differences in conditional value functions, it is useful to first set the anchor choice. The anchor choice fixes the conditional value function from which we will be differencing. Here we set the anchor choice to staying home, H, although the particular choice in this setup does not matter. We then compare the conditional value functions for blue collar and white collar with the conditional value function for staying home.

Consider the comparison between staying home and working in the blue-collar sector. Are there choice sequences from these initial choices that lead the state variables to be the same a few periods ahead? The answer is yes: Decision sequences $\{B, H, H\}$ and $\{H, B, H\}$ result in the individual having one additional unit of blue-collar experience, with the most recent decision staying home. Under both sequences, x_{t+3} is given by

$$x_{t+3} = [\, h_{Bt} + 1 \quad 0 \quad h_{Wt} \quad 0 \,],$$

so the future value terms will indeed be the same across the two initial choices.

Note that any choice at $t+2$ would have worked provided that it was the same across the sequence beginning with the blue-collar choice and the sequence beginning with staying home. Namely, the sequences ($\{B, H, B\}$, $\{H, B, B\}$) would both lead to the same state at $t+3$, although it would be a different state than the one that resulted from the pair of sequences employed above. Similarly, when comparing white collar to staying home, the following pairs of choices lead to the same state at $t+3$: ($\{W, H, H\}, \{H, W, H\}$), ($\{W, H, B\}, \{H, W, B\}$), and ($\{W, H, W\}, \{H, W, W\}$).

This example hinges on there being a choice path initiated by the anchor choice and a different choice path initiated by another choice that each lead to the same state. As shown in Section 3.2, these paths do not need to deterministically lead to the same state, but once the sequences are complete, each state must have the same probability of occurring under both sequences. Where finite dependence breaks down is when both the timing of the decision matters and the effect of the timing cannot be undone. Timing matters in this example because the occupation in the previous period is a state variable, entering as a cost of switching occupations. However, the two sequences of choices can be lined up so that, after the choice sequences terminate, the last occupation is the same, implying that the effect of the timing of the choice can be undone. If the cost of switching occupations depends on the full history of previous choices, then finite dependence would no longer hold.

Regardless of the specification of the utility function, the problem exhibits finite dependence. However, specifying the utility function in a particular way makes the mapping from the differenced future utility term to the CCPs simple. If we assume that the utility function is additively separable in the structural errors and the structural errors follow the type I extreme value distribution, the differenced conditional value functions can be expressed as known functions of the two-period-ahead CCPs. In this case, we would choose one of the pairs of sequences that results in finite dependence between blue collar and staying at home and choose another pair of sequences that results in finite dependence between white collar and staying at home. We would then use these to form the differenced conditional value functions between staying at home relative to blue collar and white collar, respectively.

Keane & Wolpin (1997) assume that the structural errors are normal and that, in the case of the work choices, these errors are not additively separable. There are good reasons for these assumptions. Namely, suppose the structural error associated with the utility of a particular occupation operates through wages. Standard practice is to estimate wages in logs, implying that the additive error in the log wage equation will enter multiplicatively in wages themselves. One could, however, have two sources of structural errors. For example, Sullivan (2011) has both a type I extreme value preference error and an error associated with the wage. Hence the framework described here would still apply conditional on knowledge of the CCPs associated with both the observed states and the wage errors. As we discuss in Section 4.2, some progress has been made in obtaining CCPs in the presence of unobserved states.

3.4. Future Events

When finite dependence applies, it substantially weakens the assumptions regarding how individuals form expectations far into the future in order to estimate the model. Consider again the occupational choice model discussed in Section 3.3. As long as the time horizon extends beyond $t + 2$, the expressions would be exactly the same regardless of the length of the time horizon. All the information embedded in the individual's time horizon is captured in the CCPs.

Furthermore, the transitions on the observed state variables could actually be time dependent: Rather than $f(x_{t+1} \mid x_t, d_t)$, we could have $f_t(x_{t+1} \mid x_t, d_t)$, making no assumptions regarding how the state variables transition beyond the sample period. We can still recover the parameters governing the utility function; we would just only use data for which we had the relevant CCPs. For example, in the occupational choice case, three-period-ahead CCPs and transition functions are needed. Hence we would only use the last three periods of data to form CCPs and transition functions.

Murphy (2010) provides an example of this in the context of new housing supply. Once a parcel of land is developed, no further decisions are made, so the terminal state property of Hotz & Miller (1993) applies, and the dynamics are fully captured by the one-period-ahead choice probabilities of developing the parcel. His model is infinite horizon and contains many nonstationary time-dependent processes for the transitions of the state variables. Knowing how these processes evolve beyond the sample period is not necessary to estimate the model because the expectations about how these processes affect the developer's decision in the future are fully captured by the one-period-ahead probability of developing.

Note that this still requires assumptions regarding what the individual knows about the one-period-ahead state variables. One option is to have the individuals know exactly how the state variables will evolve. In this case, we have the correct CCPs. The other option is to have expectations regarding the next period's states based on this period's state. At this point, there will be an issue of coverage of the CCPs as some states that the individual will be forming expectations over will not be seen in the data. We return to this issue in Section 4.1.1.

3.5. Alternative Conditional Choice Probability Frameworks

When finite dependence does not hold, there are certain cases—namely stationary, infinite-horizon settings—in which CCP estimation may still prove particularly advantageous.[6] There are two approaches one can take here.

To illustrate the first approach, consider a stationary infinite-horizon setting in which the observed states have finite support on the integers $\{1,\ldots,X\}$. Let $\epsilon^*(d\,|\,x)$ represent the expected structural error conditional on choice d being optimal in state x. As shown in Hotz & Miller (1993), this expected structural error can be expressed as a function of the CCPs. In the type 1 extreme value case, it is given by $\gamma - \ln[p(d\,|\,x)]$, where γ is Euler's constant. The ex ante value function is then

$$\bar{V}(x) = E\left\{\max_d \left[v(x,d) + \epsilon(d)\right]\right\} \tag{21}$$

$$= \sum_d p(d\,|\,x)\left[v(x,d) + \epsilon^*(d\,|\,x)\right] \tag{22}$$

$$= \sum_d p(d\,|\,x)\left[u(x,d) + \beta\sum_{x'}\bar{V}(x')f(x'\,|\,x,d) + \epsilon^*(d\,|\,x)\right]. \tag{23}$$

Now we express each of the components of Equation 23 in vector or matrix form:

$$\bar{V} = \begin{bmatrix}\bar{V}(1)\\ \vdots \\ \bar{V}(X)\end{bmatrix}, \qquad U(d) = \begin{bmatrix}u(d\,|\,1)\\ \vdots \\ u(d\,|\,X)\end{bmatrix}, \qquad \epsilon^*(d) = \begin{bmatrix}\epsilon^*(d\,|\,1)\\ \vdots \\ \epsilon^*(d\,|\,X)\end{bmatrix},$$

$$P(d) = \begin{bmatrix}p(d\,|\,1)\\ \vdots \\ p(d\,|\,X)\end{bmatrix}, \qquad F(d) = \begin{bmatrix}f(1\,|\,1,d) & \ldots & f(X\,|\,1,d)\\ \vdots & \ddots & \vdots \\ f(1\,|\,X,d) & \ldots & f(X\,|\,X,d)\end{bmatrix},$$

implying that the vector of ex ante value functions can be expressed as

$$\bar{V} = \sum_d P(d)*[U(d) + \beta F(d)\bar{V} + \epsilon^*(d)], \tag{24}$$

where $*$ refers to element-by-element multiplication. Rearranging the terms yields

$$\bar{V} - \beta\sum_d P(d)*[F(d)\bar{V}] = \sum_j P(d)*[U(d) + \epsilon^*(d)]. \tag{25}$$

Denoting I as an $X\times X$ identity matrix and λ as a $1\times X$ vector of ones, and solving for $\bar{V}$ yields

$$\bar{V} = \left(I - \beta\sum_d [P(d)\lambda]*F(d)\right)^{-1}\left(\sum_d P(d)*[U(d) + \epsilon^*(d)]\right). \tag{26}$$

[6]Although finite-horizon problems without finite dependence can still be estimated using CCP techniques, the advantages of using CCPs are not as large because the full backward recursion problem will still need to be solved.

Expressing the value function in this way serves as the basis for Aguirregabiria & Mira's (2002) pseudo-likelihood estimator as well as the games approaches of Aguirregabiria & Mira (2007), Pakes et al. (2007), and Pesendorfer & Schmidt-Dengler (2008).

Another alternative is to use the approach proposed by Hotz et al. (1994), which involves forward simulation to construct future values by directly summing up the relevant future utility contributions. This approach can be particularly advantageous when the state space is large, as it essentially uses Monte Carlo simulation to avoid enumerating the full set of future outcomes. The idea here is that, given the CCPs, we can draw future paths of both the choices and the states, collecting both flow utility parameters and the structural errors, which have been selected as part of the choice problem. Drawing many paths and averaging provide an approximation to the expected future utility for each initial choice. The paths are drawn far enough into the future so that the discounting renders future terms past this point irrelevant. Bajari et al. (2007) extend this approach to include games with continuous controls and auxiliary payoff variables (e.g., prices and quantities). Because forward simulation can also be useful in situations with finite dependence, we refrain from discussing this further until Section 4.1.4.[7]

4. CONDITIONAL CHOICE PROBABILITY ESTIMATION

We first consider the case in which, apart from the i.i.d. ϵ's, there are no additional unobserved state variables. We then describe how to incorporate additional, non-i.i.d. unobservables into CCP estimation using the algorithms developed by Arcidiacono & Miller (2010). Accounting for unobserved state variables can control for dynamic selection, allowing the individual's choices and the transitions of the state variables to be correlated with each other and across time. Note that, despite the CCPs only being approximated, these CCP estimators of the structural parameters are $\sqrt{N}$ consistent and asymptotically normal under standard regularity conditions (see, for example, Hotz & Miller 1993; Aguirregabiria & Mira 2002, 2007).

4.1. Estimation Without Unobserved States

Without unobserved heterogeneity, estimation occurs in two stages: (*a*) the recovery of the CCPs and transition functions for the state variables and (*b*) the formation of the value function using the CCPs and the estimation of the structural parameters.

4.1.1. Stage 1: obtaining conditional choice probabilities and transition functions. Given unlimited data, both the CCPs and transition functions could be estimated nonparametrically using simple bin estimators. For example, the estimated probability of choice d_t given state x_t, $\hat{p}_t(d\,|\,x_t)$, could be found using

$$p_t(d_t\,|\,x_t) = \frac{\sum_{n=1}^{\mathcal{N}} I(d_{nt} = j, x_{nt} = x)}{\sum_{n=1}^{\mathcal{N}} I(x_{nt} = x)}. \tag{27}$$

[7]The use of CCP methods does not require assuming a stationary environment but in general does require the estimation of first-stage CCPs that are fully flexible functions of time. This can obviously be quite data intensive. However, as they do not require solving for a fixed point, CCP methods do raise the possibility of handling nonstationary, infinite-horizon problems, a point we return to in Section 5.2.

Similar expressions could be formed for the transition probabilities over the observed state variables.

Data limitations, particularly when the state space is large (or continuous), will often make Equation 27 infeasible. In this case, some smoothing will need to occur. Hence nonparametric kernels, basis functions, or flexible logits could be employed. In each case, collecting more data would allow for more flexible interactions of the state variables, with the promise of more interactions to ensure consistency.

A known parametric form is typically assumed for the state transition functions. Therefore, in the event that data are sparse, the structural parameters that index these functions can be informed by the structure of the model. This is not the case for the CCPs, which should not be treated as structural objects in the first stage. This can introduce small sample bias into the second-stage structural parameter estimates. This bias can be mitigated by updating the initial nonparametric estimates with the CCPs implied by the structural model. We return to this point in Section 4.1.3.

4.1.2. Stage 2: estimating the structural parameters. As the previously estimated CCPs and state transition functions are now taken as given, this stage can often be quite simple. For example, when finite dependence holds, the only components of the future utility term that are not estimated in the first stage are the flow payoff terms associated with the finite-dependence sequences. In the case of renewal or terminating actions, the payoff for these actions may be normalized to zero, so estimation is as simple as a multinomial or binary logit with an offset term. For example, from Equation 16, the offset term for $v_t(x_t, d_t)$ would be

$$-\beta \int \ln\left[p_{t+1}(R \mid x_{t+1})\right] f(x_{t+1} \mid x_t, d_t),$$

which, for a fixed β, can be calculated outside the model. In other cases, the flow payoff terms that are accumulated over the relevant sequences must be multiplied by the relevant transitions of the state variables and discounted. Under finite dependence, the number of flow utility terms that must be collected here is only as large as the sequence itself, which can be quite small.

When finite dependence does not hold but the setting is infinite horizon and stationary, Aguirregabiria & Mira (2002) use Equation 26 to construct a CCP representation,

$$\bar{V} = \left(I - \beta \sum_d [P(d)\lambda] {*} F(d)\right)^{-1} \left(\sum_d P(d) {*} \left[U(d) + \epsilon^*(d)\right]\right),$$

solving the matrix inversion and $\epsilon^*(d)$ portions in a first stage. This $(X \times 1)$ vector of ex ante value functions is then used to form the conditional value functions associated with each choice. Let $v(d)$ represent the $(X \times 1)$ vector of conditional value functions for each observed state given choice d, which can be written

$$v(d) = U(d) + \beta F(d) \left(I - \beta \sum_d [P(d)\lambda] {*} F(d)\right)^{-1} \left(\sum_d P(d) {*} \left[U(d) + \epsilon^*(d)\right]\right),$$

where the only unknown parameters are contained in $U(d)$. Given the conditional value functions, we can then form the likelihood for the choices in the data.

4.1.3. Improving the precision of the conditional choice probabilities. With limited data, concerns can arise as to how the accuracy of the CCPs impacts the results. Aguirregabiria & Mira (2002) show that the model can be used to update the CCPs in stationary infinite-horizon settings, mitigating the small sample bias. To understand how their approach works, recall that the nested fixed-point algorithm solves for the value function within the maximization routine. Here the steps are effectively flipped: Given the CCPs, estimate the parameters and then update the CCPs using the new parameter estimates. For example, suppose type 1 extreme value errors are used. The probability of choice d given observed state x and structural parameters $\hat{\theta}$ is

$$p(d \mid x, \hat{\theta}) = \frac{\exp[v(x, d)]}{\sum_{d'} \exp[v(x, d')]}. \tag{28}$$

Hence, given estimates of the structural parameters, Equation 28 can be used to update the CCPs. Then the value function can be updated directly using Equation 26. By calculating the matrix inversions outside the likelihood maximization, significant computational gains can be obtained. Aguirregabiria & Mira (2002) provide Monte Carlo results illustrating the reductions in small sample bias, along with the improvements in computational speed.

4.1.4. Dealing with large state spaces. One of the biggest benefits of CCP estimation comes into play when the state space is very large. Backward recursion requires either evaluating or interpolating across all values of the state space. This is because, in the process of evaluating the value function, it is unclear what states will be reached. Using the CCPs to forward simulate the value function, as developed by Hotz et al. (1994) and later extended and applied to games by Bajari et al. (2007), one can calculate the value function at states that are more likely to occur.

Forward simulation works in the following way. Given the individual's current state, use the estimated conditional probability functions to draw a choice and then, conditional on the choice, draw a realized state using the estimated transition function. This process is continued until the time horizon is reached or, in the case of infinite-horizon models, the increment to the value function is sufficiently small due to discounting. Note that the process requires drawing or, alternatively, having analytic expressions for the expectations of the structural errors that are consistent with the choice path. Taking many paths and averaging approximate the value function, although Hotz et al. (1994) show that, because they are able to make the expressions linear, only one path is necessary for consistency. Of course, more paths will make the estimates more precise.

Another advantage of the Hotz et al. forward simulation approach is that it can easily be adapted to handle continuous state variables (such as wages)—it is essentially using Monte Carlo simulation to approximate continuation values at states that are not observed in the data. The extension of CCP methods to continuous controls is more complex, as it involves, for example, augmenting the participation equation with Euler equations characterizing hours worked. This is the approach taken by Altug & Miller (1998) in the context of female labor supply. Bajari et al. (2007) explore an alternative approach that avoids Euler equations entirely. Both approaches place rather strong restrictions on the way in which unobserved state variables impact the continuous decisions, making this a fruitful area for future research.

Forward simulation is particularly powerful when coupled with finite dependence. Specifically, rather than taking draws out for the full time horizon or waiting until discounting makes the increment to the value function small, we instead use the CCPs associated with paths that lead to finite dependence. The transitions on the state variables are then drawn from these paths.

Bishop (2008) considers a dynamic model of migration similar to Kennan & Walker (2011). The choice set is over 50 locations in which each location's state variables evolve according to their own processes. Even having one binary state variable that evolves over time for each location leads to $2^{50} > 1.12E+15$ possibilities for the one-period-ahead states. The actual size of her state space is $1.12E+184$. Because her model (and the Kennan & Walker model) exhibits finite dependence, she is able to form the future value terms by forward simulating the transitions of the state variables given the choice sequences that lead to finite dependence. With finite dependence being achieved after three periods, she is able to simulate many paths of the state variables given the three-period choice sequences.

4.2. Estimation with Unobserved States

We now turn to the case in which some of the states are unobserved and persist over time. In this way, dynamic discrete choice models can be adapted to handle dynamic selection. We show the standard ways of handling unobserved heterogeneity as well as how CCP methods can be adapted to these cases.

4.2.1. Mixture distributions and the Expectation-Maximization algorithm. Following Heckman & Singer (1984), the standard approach to account for unobserved heterogeneity in dynamic discrete models is to employ finite mixture distributions. Now, in addition to x_{nt}, there is an unobserved state variable s_{nt}, which takes on one of S values, $s_{nt} \in \{1, \ldots, S\}$. To keep the exposition simple, we focus on the case in which the unobserved state does not vary over time, although Arcidiacono & Miller (2010) allow for time variation in s_{nt}. The joint likelihood of d_{nt} and x_{nt+1}, conditional on x_{nt} and unobserved state s, is

$$\mathcal{L}_t(d_{nt}, x_{nt+1} \mid x_{nt}, s; \theta) = p_t(d_{nt} \mid x_{nt}, s; \theta_1) f_t(x_{nt+1} \mid d_{nt}, x_{nt}, s; \theta_2). \tag{29}$$

Because s_n is unobserved, we integrate it out of the likelihood. Denote $\pi(s \mid x_{n1})$ as the probability of being in unobserved state s given the data at the first observed time period, x_{n1}. The likelihood of the observed data for n is then given by

$$\sum_{s=1}^{S} \pi(s \mid x_{n1}) \mathcal{L}_t(d_{nt}, x_{nt+1} \mid x_{nt}, s, \theta).$$

Maximizing the log likelihood of the observed data now requires that one solve for both θ and π, where π refers to all possible values of $\pi(s \mid x_{n1})$:

$$(\hat{\theta}, \hat{\pi}) = \arg\max_{\theta, \pi} \sum_{n=1}^{\mathcal{N}} \ln \left[\sum_{s=1}^{S} \pi(s \mid x_{n1}) \prod_{t=1}^{\mathcal{T}} \mathcal{L}_t(d_{nt}, x_{nt+1} \mid x_{nt}, s; \theta) \right]. \tag{30}$$

Note that the log likelihood function is no longer additively separable, implying that direct maximization of Equation 30 can no longer be done in stages.

The first-order conditions to this problem are

$$\frac{\partial L}{\partial \theta}=0=\sum_{n=1}^{\mathcal{N}}\frac{\sum_{s}\sum_{t'}\pi(s\,|\,x_{n1})\prod_{t\neq t'}\mathcal{L}_t(d_{nt},x_{nt+1}\,|\,x_{nt},s;\theta)\dfrac{\partial \mathcal{L}_{t'}(d_{nt'},x_{nt'+1}\,|\,x_{nt'},s;\theta)}{\partial\theta}}{\sum_{s=1}^{S}\pi(s\,|\,x_{n1})\prod_{t=1}^{\mathcal{T}}\mathcal{L}_t(d_{nt},x_{nt+1}\,|\,x_{nt},s;\theta)}$$
$$=\sum_{n=1}^{\mathcal{N}}\frac{\sum_{s=1}^{S}\sum_{t'}\pi(s\,|\,x_{n1})\prod_{t=1}^{\mathcal{T}}\mathcal{L}_t(d_{nt},x_{nt+1}\,|\,x_{nt},s;\theta)\dfrac{\partial \ln \mathcal{L}_{t'}(d_{nt'},x_{nt'+1}\,|\,x_{nt'},s;\theta)}{\partial\theta}}{\sum_{s=1}^{S}\pi(s\,|\,x_{n1})\prod_{t=1}^{\mathcal{T}}\mathcal{L}_t(d_{nt},x_{nt+1}\,|\,x_{nt},s;\theta)}. \tag{31}$$

Bayes' rule implies that the conditional probability of n being in unobserved state s given the data for n and the parameters $\{\theta,\pi\}$, $q(s\,|\,d_n,x_n;\theta,\pi)$ is

$$q(s\,|\,d_n,x_n;\theta,\pi)=\frac{\pi(s\,|\,x_{n1})\prod_{t=1}^{\mathcal{T}}\mathcal{L}_t(d_{nt},x_{nt+1}\,|\,x_{nt},s;\theta)}{\sum_{s'}\pi(s'\,|\,x_{n1})\prod_{t=1}^{\mathcal{T}}\mathcal{L}_t(d_{nt},x_{nt+1}\,|\,x_{nt},s';\theta)}. \tag{32}$$

The first-order condition given in Equation 31 can then be rewritten as

$$0=\sum_{n=1}^{\mathcal{N}}\sum_{s=1}^{S}\sum_{t=1}^{\mathcal{T}}q(s\,|\,d_n,x_n;\theta,\pi)\frac{\partial \ln \mathcal{L}_t(d_{nt},x_{nt+1}\,|\,x_{nt},s;\theta)}{\partial\theta}. \tag{33}$$

Because $q(s\,|\,d_n,x_n;\hat{\theta},\hat{\pi})$ is the probability of n being in unobserved state s conditional on the data for n, averaging across all individuals with $x_{1n}=x_1$ must correspond to $\hat{\pi}(s\,|\,x_1)$, the estimated population probability of being in state s given first-period data x_1:

$$\hat{\pi}(s\,|\,x_1)=\frac{\sum_n q\left(s\,|\,d_n,x_n;\hat{\theta},\hat{\pi}\right)I(x_{1n}=x_1)}{\sum_n I(x_{1n}=x_1)}. \tag{34}$$

Dempster et al. (1977) note that the same first-order condition given in Equation 31 would hold if $q(s\,|\,d_n,x_n;\hat{\theta},\hat{\pi})$ were taken as given and the maximization problem were

$$\hat{\theta}=\arg\max_{\theta}\sum_{n=1}^{\mathcal{N}}\sum_{t=1}^{\mathcal{T}}q(s\,|\,d_n,x_n;\hat{\theta},\hat{\pi})\ln\left[\mathcal{L}_t(d_{nt},x_{nt},s,\theta)\right]. \tag{35}$$

Because $(\hat{\theta},\hat{\pi})$ are not known pre-estimation, Dempster et al. developed the expectation-maximization (EM) algorithm, which yields a solution to the first-order conditions in Equation 31 upon convergence. Namely, given initial values $\theta^{(1)}$ and $\pi^{(1)}$, the $(m+1)$-th iteration is given by the following two-step process. In the expectation step, update the conditional probabilities of being in each unobserved state according to

$$q^{(m+1)}(s\,|\,d_n,x_n)=\frac{\pi^{(m)}(s\,|\,x_{n1})\prod_{t=1}^{\mathcal{T}}\mathcal{L}_t\left[d_{nt},x_{nt+1}\,|\,x_{nt},s;\theta^{(m)}\right]}{\sum_{s'}\pi^{(m)}(s'\,|\,x_{n1})\prod_{t=1}^{\mathcal{T}}\mathcal{L}_t\left[d_{nt},x_{nt+1}\,|\,x_{nt},s';\theta^{(m)}\right]} \tag{36}$$

and update the population probability of being in each unobserved state, given values for the first-period state variables, using

$$\pi^{(m+1)}(s\,|\,x_1)=\frac{\sum_n q^{(m+1)}(s\,|\,d_n,x_n)I(x_{1n}=x_1)}{\sum_n I(x_{1n}=x_1)}. \tag{37}$$

In the maximization step, taking $q^{(m+1)}(s \mid d_n, x_n)$ as given, obtain $\theta^{(m+1)}$ from

$$\theta^{(m+1)} = \arg\max_{\theta} \sum_{n=1}^{\mathcal{N}} \sum_{t=1}^{\mathcal{T}} \sum_{s=1}^{S} q^{(m+1)}(s \mid d_n, x_n) \ln\left[\mathcal{L}_t(d_{nt}, x_{nt+1} \mid x_{nt}, s; \theta)\right]. \tag{38}$$

These steps are repeated until convergence, with each step increasing the log likelihood of the original problem.

The maximization step of the EM algorithm has a nice interpretation. Namely, it operates as though s were observed, using $q^{(m+1)}(s \mid d_n, x_n)$ as population weights. With s treated as observed, additive separability at the maximization step is reintroduced. Arcidiacono & Jones (2003) show that the maximization step can once again be carried out in stages. For example, we can express Equation 38 as

$$\begin{aligned}\theta^{(m+1)} = \arg\max\nolimits_{\theta} \sum\nolimits_{n=1}^{\mathcal{N}} \sum\nolimits_{t=1}^{\mathcal{T}} \sum\nolimits_{s=1}^{S} q^{(m+1)}(s \mid d_n, x_n)\big(&\ln\left[p_t(d_{nt} \mid x_{nt}, s; \theta)\right] \\ &+ \ln\left[f_t(x_{nt+1} \mid d_{nt}, x_{nt}, s; \theta_2)\right]\big).\end{aligned} \tag{39}$$

At the maximization step, we first obtain $\theta_2^{(m+1)}$ from the second half of Equation 39 and then estimate θ_1 from the first half of Equation 39, taking $\theta_2^{(m+1)}$ as given.

Note that the EM algorithm requires that one repeat the maximization step multiple times. For many dynamic discrete choice problems, this renders the EM algorithm an unattractive solution. There are two cases, however, in which the EM algorithm is particularly helpful. First, if estimating in stages substantially speeds up estimation, then the methods of Arcidiacono & Jones (2003) can result in large computational savings (see, e.g., Arcidiacono 2004, 2005; Arcidiacono et al. 2007; Beffy et al. 2011; Fiorini 2011). Second, as shown the next section, it provides a natural way of integrating unobserved heterogeneity into CCP estimation.

4.2.2. Linking the expectation-maximization algorithm to conditional choice probability estimation. Because the EM algorithm treats the unobserved state as known in maximization, Arcidiacono & Miller (2010) show that it is easily adapted to CCP estimation. For ease of notation, denote $d_{njt} = 1$ if individual n chooses j at time t. We can then express the CCP $p_t(j \mid x_t, s)$ as

$$p_t(j \mid x_t, s) = \frac{Pr(j, s \mid x_t)}{Pr(s \mid x_t)} = \frac{E\left[d_{njt} I(s_n = s) \mid x_{nt} = x\right]}{E\left[I(s_n = s) \mid x_{nt} = x_t\right]}. \tag{40}$$

Denoting d_n and x_n as the set of decisions and observed states for n and applying the law of iterated expectations, we can express Equation 40 as

$$p_t(j \mid x_t, s) = \frac{E\left[d_{njt} E\{I(s_n = s) \mid d_n, x_n\} \mid x_{nt} = x\right]}{E\left[E\{I(s_n = s) \mid d_n, x_n\} \mid x_{nt} = x_t\right]}, \tag{41}$$

but note that the inner expectations of both the numerator and the denominator are the conditional probabilities of being in each unobserved state, $q(s \mid d_n, x_n; \theta, \pi)$, implying

$$p_t(j \mid x_t, s) = \frac{E\left[d_{njt} q(s \mid d_n, x_n) \mid x_{nt} = x\right]}{E\left[q(s \mid d_n, x_n) \mid x_{nt} = x_t\right]}. \tag{42}$$

Given that the EM algorithm provides estimates of the conditional probabilities of being in each unobserved state, we can use the sample analog of Equation 42 evaluated at the current parameter estimates and then update the CCPs using the EM algorithm.[8] Given initial values $\theta^{(1)}$, $\pi^{(1)}$, and the vector of CCPs $p^{(1)}$, the $(m+1)$-th iteration is given by the following two-step process. In the expectation step,

$$q^{(m+1)}(s \mid d_n, x_n) = \frac{\pi^{(m)}(s \mid x_{n1})\prod_{t=1}^{\mathcal{T}}\mathcal{L}_t\left[d_{nt}, x_{nt+1} \mid x_{nt}, s, p^{(m)}; \theta^{(m)}\right]}{\sum_{s'}\pi^{(m)}(s' \mid x_{n1})\prod_{t=1}^{\mathcal{T}}\mathcal{L}_t\left[d_{nt}, x_{nt+1} \mid x_{nt}, s', p^{(m)}; \theta^{(m)}\right]}, \tag{43}$$

$$\pi^{(m+1)}(s \mid x_1) = \frac{\sum_n q^{(m+1)}(s \mid d_n, x_n) I(x_{1n} = x_1)}{\sum_n I(x_{1n} = x_1)}, \tag{44}$$

$$p_t^{(m+1)}(j \mid x_t, s) = \frac{\sum_{n=1}^{\mathcal{N}} q^{(m+1)}(s \mid d_n, x_n) d_{njt} I(x_{nt} = x_t)}{\sum_{n=1}^{\mathcal{N}} q^{(m+1)}(s \mid d_n, x_n) I(x_{nt} = x_t)}. \tag{45}$$

In the maximization step,

$$\theta^{(m+1)} = \arg\max_{\theta} \sum_{n=1}^{\mathcal{N}}\sum_{t=1}^{\mathcal{T}}\sum_{s=1}^{S} q^{(m+1)}(s \mid d_n, x_n) \ln\left\{\mathcal{L}_t[d_{nt}, x_{nt+1} \mid x_{nt}, s, p^{(m)}; \theta]\right\}. \tag{46}$$

4.3. Unobserved Heterogeneity in a First Stage

Updating the CCPs in this way suggests the two-stage approach described in Arcidiacono & Miller (2010). Namely, in the first stage, allow the CCPs to be completely flexible. Instead of using

$$\mathcal{L}_t[d_{nt}, x_{nt+1} \mid x_{nt}, s, p^{(m)}; \theta^{(m)}] = p_t[d_{nt} \mid x_{nt}, s; \theta_1^{(m)}]\, f_t\, [x_{nt+1} \mid d_{nt}, x_{nt}, s; \theta_2^{(m)}]$$

in Equation 43, replace $p_t[d_{nt} \mid x_{nt}, s, p^{(m)}; \theta_1^{(m)}]$ with $p_t^{(m)}(d_{nt} \mid x_{nt}, s)$ using Equation 45. The update for the conditional probability of being in an unobserved state is then given by

$$q^{(m+1)}(s \mid d_n, x_n) = \frac{\pi^{(m)}(s \mid x_{n1})\prod_{t=1}^{\mathcal{T}} f_t[x_{nt+1} \mid d_{nt}, x_{nt}, s; \theta_2^{(m)}] p_t^{(m)}(d_{nt} \mid x_{nt}, s)}{\sum_{s'}\pi^{(m)}(s' \mid x_{n1})\prod_{t=1}^{\mathcal{T}} f_t[x_{nt+1} \mid d_{nt}, x_{nt}, s'; \theta_2^{(m)}] p_t^{(m)}(d_{nt} \mid x_{nt}, s')}. \tag{47}$$

The first-stage maximization step is then only over θ_2:

$$\theta_2^{(m+1)} = \arg\max_{\theta} \sum_{n=1}^{\mathcal{N}}\sum_{t=1}^{\mathcal{T}}\sum_{s=1}^{S} q^{(m+1)}(s \mid d_n, x_n) \ln\,[f_t(x_{nt+1} \mid d_{nt}, x_{nt}, s; \theta_2)]. \tag{48}$$

At convergence, we will have recovered estimates of the conditional probabilities of being in each unobserved state, the choice probabilities conditional on both observed and unobserved states, and estimates of θ_2. In the second stage, any CCP method could be used to recover the parameters governing the utility payoffs.

[8]Arcidiacono & Miller (2010) also suggest an alternative way to update the CCPs in stationary environments. Namely, the CCPs can be updated from the model using $p^{(m+1)}(j \mid x, s) = p\left[j \mid x, s, p^{(m)}; \theta_1^{(m)}\right]$. This is similar in spirit to the approach of Aguirregabiria & Mira (2007). This method has the advantage of not having to smooth the CCPs when there are few individuals in a particular state.

4.4. Choices of the Error Distributions with Unobserved States

Although above we explain the Arcidiacono and Miller approach to unobserved heterogeneity in cases in which the unobserved states are permanent for ease of exposition, their method applies to cases in which the unobserved state transitions over time. Furthermore, due to the use of the EM algorithm, allowing the unobserved states to transition over time has little effect on computational times. Although it is certainly possible to program full solution methods in which the unobserved states transition over time (see Keane & Wolpin 2010 for the one paper we are aware that does so), the computational requirements will increase dramatically. This is because we now have to integrate out over the possible unobserved state transitions, and this will need to be done within the likelihood maximization routine.

Recall that one of the criticisms of CCP methods is the requirement of the structural errors entering additively in the utility function. For example, Keane & Wolpin (1997) assume that the errors associated with wages are known to the individual, but enter multiplicatively. However, they also assume that the errors are independent over time. An alternative is to assume that the portion of the wage that is known to the individual at the time of the decision operates through the observed variables and through discrete, unobserved states that transition over time. Hence there may be a computational trade-off between knowing the full wage error, but assuming serial independence, and allowing for persistence of the error (letting the unobserved state transition over time), but not knowing the full wage error at the time of the decision. Although in principle full solution methods could be used to incorporate serially correlated continuous variables, this is virtually never done, with the one exception being Erdem & Keane (1996).

5. GAMES

The simplifications that arise from both CCP estimation and the representations emphasized here extend beyond single-agent problems. Arguably, the biggest impact of CCP estimation has been on the estimation of dynamic discrete games, which, over the past few years, has quickly become one of the fastest growing areas of empirical industrial organization. CCP estimation of discrete games was first proposed by Rust (1994a). Building on the logic of Hotz & Miller (1993), Rust suggested an estimation strategy that involved substituting nonparametric estimates of rival firms' reaction functions into each firm's own optimal response function, effectively turning a complex equilibrium problem into a collection of simpler games-against-nature one. In a series of contemporaneous papers (Aguirregabiria & Mira 2007, Bajari et al. 2007, Pakes et al. 2007, Pesendorfer & Schmidt-Dengler 2008), several authors have recently generalized and extended this approach, focusing on infinite-horizon games with stationary Markov perfect equilibrium.

The use of CCPs in the games context yields two important benefits. First, as in the single-agent case, the repeated solution of a fixed-point problem is avoided. This is especially useful for games because the fixed-point problem is doubly nested, involving both the solution of each agent's DP problem and the corresponding equilibrium conditions, a computational problem that increases exponentially both in the number of states and heterogeneous players. Second, unlike the single-agent problem, the game may (and likely will) admit more than one fixed point, and this multiplicity problem can complicate the correct specification of the likelihood (or generalized method of moments criterion function) because the model is then incomplete. However, if the researcher is willing to assume that only one equilibrium is played in the data, first-stage nonparametric estimation of the CCPs effectively completes the model by conditioning on the equilibrium that was actually played.

5.1. Dynamic Discrete Games

Although CCP methods can be applied to models for which the controls are continuous or additional payoff variables are observed, for ease of illustration we consider a pure discrete control example with latent payoffs, closest in structure to the model proposed by Rust and later extended by Aguirregabiria & Mira (2007). Consider a discrete game played by I players in each of many markets. In addition to the common state variables x_t, the systematic component of the i-th firm's payoff now depends on both its own choice in period t, now denoted by $d_t^{(i)} \in D$, and the actions of its rivals, denoted by $d_t^{(-i)} \equiv \left[d_t^{(1)}, \ldots, d_t^{(i-1)}, d_t^{(i+1)}, \ldots, d_t^{(I)}\right]$. We continue to assume additive separability and conditional independence. The current (flow) payoff of firm i in period t is then given by $U^{(i)}\left[x_t, d_t^{(i)}, d_t^{(-i)}\right] + \epsilon_t\left[d_t^{(i)}\right]$, where $\epsilon_t\left[d_t^{(i)}\right]$ is an i.i.d. random variable that is privately observed by firm i, making this a game of incomplete information. Also, the payoff function is now superscripted by i to account for the possibility that the state variables might impact different firms in different ways (e.g., own versus other characteristics).

We assume that moves (i.e., choices) are taken simultaneously in each period, and let $P\left[d_t^{(-i)} \mid x_t\right]$ denote the probability that firm i's competitors choose $d_t^{(-i)}$ conditional on x_t. Because $\epsilon_t^{(i)}$ is independently distributed across all the firms, $P\left[d_t^{(-i)} \mid x_t\right]$ can be written as the following product:

$$P\left[d_t^{(-i)} \mid x_t\right] = \prod_{j \neq i}^{I} p^{(j)}\left[d_t^{(j)} \mid x_t\right], \tag{49}$$

where $p^{(j)}\left[d_t^{(j)} \mid x_t\right]$ is player j's CCP. These CCPs represent the best-response probability functions, constructed by integrating firm j's decision rule (i.e., strategy) over its private information draw, and characterize the firm's equilibrium behavior from the point of view of each of its rivals (as well as the econometrician). Although the existence of equilibrium follows directly from Brouwer's theorem (see Aguirregabiria & Mira 2007), uniqueness is unlikely to hold given the inherent nonlinearity of the underlying reaction functions. However, given a stationary rational expectations Markov perfect equilibrium, the beliefs of firm i will match the probabilities given in Equation 49. Taking the expectation of $U^{(i)}\left[x_t, d_t^{(i)}, d_t^{(-i)}\right]$ over $d_t^{(-i)}$, the systematic component firm i's current payoff is then given by

$$u^{(i)}\left[x_t, d_t^{(i)}\right] = \sum_{d_t^{(-i)} \in D^{I-1}} P\left[d_t^{(-i)} \mid x_t\right] U^{(i)}\left[x_t, d_t^{(i)}, d_t^{(-i)}\right], \tag{50}$$

which is straightforward to construct, up to the parameterized utility function, using first-stage estimates of the relevant CCPs.

Having dealt with the flow payoffs, we now must construct the continuation values. Note that, from the perspective of firm i, the period $t+1$ state variables are updated from the prior period's values by their own actions (which they know) as well as the actions of their rivals (which they have beliefs over). Let $F\left[x_{t+1} \mid x_t, d_t^{(i)}, d_t^{(-i)}\right]$ represent the probability of x_{t+1} occurring given own action $d_t^{(i)}$, current state x_t, and rival actions $d_t^{(-i)}$. The probability of transitioning from x_t to x_{t+1} given $d_t^{(i)}$ is then given by

$$f^{(i)}\left[x_{t+1} \mid x_t, d_t^{(i)}\right] = \sum_{d_t^{(-i)} \in D^{I-1}} P\left[d_t^{(-i)} \mid x_t\right] F\left[x_{t+1} \mid x_t, d_t^{(i)}, d_t^{(-i)}\right]. \tag{51}$$

Notice that the expression for the conditional value function for firm i matches that of Equation 14 subject to the condition that we are now in a stationary environment. In particular, Equation 14 is now simply

$$\begin{aligned} v^{(i)}\left[x_t, d_t^{(i)}\right] &= u^{(i)}\left[x_t, d_t^{(i)}\right] \\ &+ \beta \int \left(v^{(i)}\left[x_{t+1}, d_{t+1}^{(i)}\right] - \ln\left\{p^{(i)}\left[d_{t+1}^{(i)} \mid x_{t+1}\right]\right\}\right) f^{(i)}\left[x_{t+1} \mid x_t, d_t^{(i)}\right] dx_{t+1} + \beta\gamma, \end{aligned} \tag{52}$$

where γ is Euler's constant.

At this point, any of the estimation methods that apply to the stationary infinite-horizon setting can be used. In particular, Aguirregabiria & Mira (2007) use a matrix inversion as in Equation 26 to directly solve for the continuation value [this method is also employed by Pakes et al. (2007) and Pesendorfer & Schmidt-Dengler (2008)]. Aguirregabiria & Mira (2007) also show how to improve small sample performance by iterating on the fixed-point mapping (in probability space). Alternatively, Bajari et al. (2007) use the forward simulation technique from Hotz et al. (1994), which they extend to accommodate continuous controls and additional information on per-period payoffs. Note that unobserved heterogeneity can also be included using the methods developed by either Aguirregabiria & Mira (2007) or Arcidiacono & Miller (2010).

5.2. Finite Dependence in Games

Notably, games represent a setting in which finite dependence is unlikely to hold, due to the complicating presence of strategic interactions. Even if a given firm's actions are locally reversible (e.g., plants can be both built and scrapped), strategic reactions by the firm's rivals will likely make it difficult to ensure that the effect of each choice can be undone relative to the anchor choice. However, if one of the choices is to exit the market (with no possibility of re-entry), then the terminal state property holds, allowing the continuation value from exiting to be either normalized to zero or parameterized as a component of the utility function (capturing a scrap value, for example). In this case, the game can once again be estimated using only one-period-ahead CCPs.

Beresteanu et al. (2010) employ a version of this representation in their analysis of dynamic competition between retail chains. In their setting, two types of firms (supermarkets and Walmart-style supercenters) open and close stores, engaging in per-period price competition that depends on the number of stores per capita that each firm operates. Per-period profits are modeled using a logit demand system, there is free entry and exit, and scrap values are assumed to depend on the current size of the chain. The size of their state space renders alternative approaches intractable.

The simplifying structure implied by the exit option raises the possibility of estimating nonstationary games or models of social interactions. Clearly, nonstationarity coupled with an infinite horizon raises issues concerning the existence of equilibria. Nonetheless, the data do contain information on the probabilities with which certain choices are made given observed states. If we assume that players know these probabilities, estimation can proceed just as before. Beauchamp (2010) estimates a nonstationary entry/exit game of abortion providers, allowing the demand for abortion to depend explicitly on time. He recovers entry and exit probabilities (given the observed states) from the data and expresses the firm's future value term relative to the conditional value of exiting. In this case, the only future CCPs that are needed are the firm's own exit probabilities conditional on the possible one-period-ahead states.

6. POLICY EVALUATION

The computational advantages of the CCP approach stem from avoiding the solution of the full DP problem when estimating the structural parameters of the underlying model. This is sometimes perceived as a weakness when it comes to conducting counterfactual policy simulations, which typically involves fully resolving the DP. Although the structural model may only need to be solved once to conduct policy simulations (as opposed to the multiple times required to estimate the model using a full solution approach), this contradicts the spirit of this article, which is focused on keeping both programming and computation simple.

Short-run interventions, however, are well suited to the CCP framework, providing another opportunity to avoid resolving the full DP. Consider, for example, the recent subsidy for first-time homebuyers. Suppose that this policy is known to be in place only for t periods and came as a complete surprise. Absent general equilibrium effects, the policy should only impact decisions that take place after period t through its lasting effect on the state variables. That is, conditional on the state that obtains at $t+1$, whether we calculate the value functions by solving the full DP or by using the CCPs observed in the data, they will be the same at $t+1$ regardless of whether the policy was in effect at t. So long as the effect of the expired policy operates exclusively through its impact on the resulting states (and provided the relevant values of the state variables are spanned by the conditions that existed in the market prior to the policy intervention), once the policy expires, individuals will revert to the behavior observed under the prior structure, thereby restoring the validity of the original CCPs. The original CCPs will only be invalid during the period in which the policy is in place.

However, during this period, updated policy functions (i.e., valid CCPs) can be constructed using the new per-period payoff functions by solving a backward recursion problem that terminates at t. The future value terms for periods after t (which are state specific) can be constructed from the original CCPs, eliminating the need to consider the agent's full time horizon or solve a contraction mapping. In the case of the homebuyer subsidy, ownership patterns could still continue to evolve differently over time even after the policy expires, but this would operate only through the way in which the policy impacted the individual's state variables, which were in turn affected by the choices made when the policy was in place. With the CCPs in hand, we can forward simulate the dynamics of home ownership after the policy expires, exploiting its impact on the resulting distribution of states. Although there are certainly examples of policy interventions that would permanently alter the relevant CCPs (e.g., a shift to a new equilibrium), in many cases of interest, the impact on the optimal policy functions is likely to be short lived.

7. CONCLUSION

As noted in Section 1, there is clearly no shortage of empirical applications in which dynamics play a central role. Moreover, high-quality panel data have never been more abundant, due to the recent revolutions in information technology and data-storage costs. Nonetheless, the computational burden and complexity associated with structural estimation of dynamic discrete choice models remain a formidable barrier to entry. We hope that by highlighting methods for which this burden is low, but the scope of application high, we have reduced these costs. Although a curse of complexity will continue to exist, as talented researchers continue to push the boundaries of what can be estimated, we believe that the core methods should be accessible to all researchers.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Aguirregabiria V, Mira P. 2002. Swapping the nested fixed point algorithm: a class of estimators for discrete Markov decision models. *Econometrica* 70:1519–43

Aguirregabiria V, Mira P. 2007. Sequential estimation of dynamic discrete games. *Econometrica* 75:1–54

Aguirregabiria V, Mira P. 2010. Dynamic discrete choice structural models: a survey. *J. Econom.* 156(1):38–67

Altug S, Miller RA. 1998. The effect of work experience on female wages and labour supply. *Rev. Econ. Stud.* 62:45–85

Angrist J, Pischke JS. 2010. The credibility revolution in empirical economics: how better research design is taking the con out of econometrics. *J. Econ. Perspect.* 24(1):3–30

Arcidiacono P. 2004. Ability sorting and the returns to college major. *J. Econom.* 121(1–2):343–75

Arcidiacono P. 2005. Affirmative action in higher education: How do admission and financial aid rules affect future earnings? *Econometrica* 73:1477–524

Arcidiacono P, Bayer P, Blevins J, Ellickson P. 2010. *Estimation of dynamic discrete choice models in continuous time*. Work. Pap., Duke Univ.

Arcidiacono P, Jones JB. 2003. Finite mixture distributions, sequential likelihood, and the EM algorithm. *Econometrica* 71:933–46

Arcidiacono P, Miller RA. 2010. *CCP estimation of dynamic discrete choice model with unobserved heterogeneity.* Work. Pap., Duke Univ.

Arcidiacono P, Sieg H, Sloan F. 2007. Living rationally under the volcano? An empirical analysis of heavy drinking and smoking. *Int. Econ. Rev.* 48(1):37–65

Bajari P, Benkard L, Levin J. 2007. Estimating dynamic models of imperfect competition. *Econometrica* 75:1331–71

Beauchamp A. 2010. *Abortion supplier dynamics*. Work. Pap., Boston Coll.

Beffy M, Fougere D, Maurel A. 2011. Choosing the field of study in post-secondary education: Do expected earnings matter? *Rev. Econ. Stat.* In press

Beresteanu A, Ellickson PB, Misra S. 2010. *The dynamics of retail oligopoly.* Work. Pap., Univ. Rochester

Bishop KC. 2008. *A dynamic model of location choice and hedonic valuation.* Work. Pap., Washington Univ.

Bresnahan TF, Stern S, Trajtenberg M. 1997. Market segmentation and the sources of rents from innovation: personal computers in the late 1980s. *Rand J. Econ.* 28:17–44

Cardell NS. 1997. Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econom. Theory* 13(2):185–213

Dempster AP, Laird NM, Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B* 39:1–38

Doraszelski U, Judd K. 2010. *Avoiding the curse of dimensionality in dynamic stochastic games*. Work. Pap., Harvard Univ.

Eckstein Z, Wolpin KI. 1990. Estimating a market equilibrium search model from panel data on individuals. *Econometrica* 58(4):783–808

Erdem T, Keane MP. 1996. Decision-making under uncertainty: capturing dynamic brand choice processes in turbulent consumer goods markets. *Mark. Sci.* 15(1):1–20

Ericson R, Pakes A. 1995. Markov-perfect industry dynamics: a framework for empirical work. *Rev. Econ. Stud.* 62:53–82

Fiorini M. 2011. Fostering educational enrollment through subsidies: the issue of timing. *J. Appl. Econ.* In press

Flinn C, Heckman J. 1982. New methods for analyzing structural models of labor force dynamics. *J. Econom.* 18(1):115–68

Gowrisankaran G, Rysman M. 2009. *Dynamics of consumer demand for new durable goods*. Work. Pap., Boston Univ.

Heckman JJ. 2010. Building bridges between structural and program evaluation approaches to evaluating policy. *J. Econ. Lit.* 48(2):356–98

Heckman JJ, Singer B. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52(2):271–320

Hotz VJ, Miller RA. 1993. Conditional choice probabilities and estimation of dynamic models. *Rev. Econ. Stud.* 61:265–89

Hotz VJ, Miller RA, Sanders S, Smith J. 1994. A simulation estimator for dynamic models of discrete choice. *Rev. Econ. Stud.* 60:265–89

Joensen JS. 2009. *Academic and labor market success: the impact of student employment, abilities, and preferences*. Work. Pap., Univ. Stockholm

Keane MP, Todd PE, Wolpin KI. 2011. The structural estimation of behavioral models: discrete choice dynamic programming methods and applications. In *Handbook of Labor Economics*, ed. O Ashenfelter, D Card. Amsterdam: North-Holland. In press

Keane MP, Wolpin KI. 1994. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. *Rev. Econ. Stat.* 76:648–72

Keane MP, Wolpin KI. 1997. The career decisions of young men. *J. Polit. Econ.* 105:473–522

Keane MP, Wolpin KI. 2010. The role of labor and marriage markets, preference heterogeneity, and the welfare system in the life cycle decisions of black, Hispanic, and white women. *Int. Econ. Rev.* 51(3):851–92

Kennan J, Walker JR. 2011. The effect of expected income on individual migration decisions. *Econometrica*. In press

McCall JJ. 1970. Economics of information and job search. *Q. J. Econ.* 84(1):113–26

McFadden D. 1978. Modelling the choice of residential location. In *Spatial Interaction and Planning Models*, ed. F Snickars, J Weibull, pp. 75–96. Amsterdam: North-Holland

Melnikov O. 2001. *Demand for differentiated durable products: the case of the US computer printer market*. Work. Pap., Cornell Univ.

Miller RA. 1984. Job matching and occupational choice. *J. Polit. Econ.* 92:1086–120

Murphy A. 2010. *A dynamic model of housing supply.* Work. Pap., Washington Univ.

Nair H. 2007. Intertemporal price discrimination with forward-looking consumers: application to the US market for console video-games. *Quant. Mark. Econ.* 5(3):239–92

Pakes A. 1986. Patents as options: some estimates of the value of holding European patent stocks. *Econometrica* 54:755–84

Pakes A, Ostrovsky M, Berry S. 2007. Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *Rand J. Econ.* 38(2):373–99

Pesendorfer M, Schmidt-Dengler P. 2008. Asymptotic least square estimators for dynamic games. *Rev. Econ. Stud.* 75:901–8

Postel-Vinay F, Robin J. 2002. Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica* 70(6):2295–350

Rust J. 1987. Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher. *Econometrica* 55:999–1033

Rust J. 1994a. Estimation of dynamic structural models, problems and prospects: discrete decision processes. In *Advances in Econometrics: Sixth World Congress*, ed. C Sims, pp. 119–70. Cambridge, UK: Cambridge Univ. Press

Rust J. 1994b. Structural estimation of Markov decision processes. In *Handbook of Econometrics*, Vol. 4, ed. R Engle, D McFadden, pp. 3081–143. Amsterdam: North-Holland

Sullivan P. 2011. A dynamic analysis of educational attainment, occupational choices and job search. *Int. Econ. Rev.* In press

Todd PE, Wolpin KI. 2006. Assessing the impact of a school subsidy program in Mexico: using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *Am. Econ. Rev.* 96(50):1384–417

Wolpin K. 1984. An estimable dynamic stochastic model of fertility and child mortality. *J. Polit. Econ.* 92:852–74

Annual Review of Economics

Volume 3, 2011

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Economics* articles may be found at http://econ.annualreviews.org