Practical Methods for Estimation of Dynamic Discrete Choice Models

Peter Arcidiacono

Department of Economics, Duke University, Box 90097, Durham NC 27708; email: psarcidi@econ.duke.edu

Paul B. Ellickson

Simon School of Business Administration, University of Rochester, Rochester, NY 14627; email: paul.ellickson@simon.rochester.edu

CONTENTS

Keywords: dynamic discrete choice, unobserved heterogeneity, conditional choice probabilities

Abstract

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

rich patterns of unobserved state variables.

1 Introduction

Many discrete decisions are made with an eye towards how they will impact 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 analysis of 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. While structural methods are often pitted against their descriptive counterparts, the two can often serve as complements. At their best, structural models can replicate results obtained from randomized experiments or attempts to exploit quasi-randomization *and* tell us how individuals will respond to counterfactual policies. Todd and 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. With the structural model able to 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), showed that, under certain restrictions, estimating these dynamic discrete choice models was both feasible and important for answering key economic questions. Each of these papers 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 identifying the optimal decision rule for each period and explicitly modeling expectations regarding future events. Recent surveys by Aguirregabiria and Mira (2007) and Keane, Todd, and Wolpin (forthcoming) 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 are often substantially easier to program, thereby lowering the barriers to entry that deter many researchers from entering the field. The estimators we discuss will not be as efficient as full solution methods that solve the full dynamic programming program. 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 where the entry costs are relatively low, we hope to broaden the set of individuals who will, when the problem invites it, consider a structural approach. The accessibility of structural models is also important for building credibility among researchers outside the structural audience. Heckman (2010) argues that part of the Angrist and 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. Further, while the models themselves can be quite complicated, their strong link to data makes clear what variation is driving the results.

The methods discussed in this paper build on the seminal work of Hotz and Miller (1993). Their insight was that the data contain a wealth of information regarding individuals expectations about the future. In particular, they showed how the future utility term could 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 non-stationary 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 paper 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 tradeoffs researchers make with regard to the 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 conditional choice probabilities, paying particular attention to what conditions 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 where the future utility term depends only on a few conditional choice probabilities. In section 6, we examine cases in which models estimated via conditional choice probability 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 specified as a linear function of x and a parameter vector θ_1. 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 $\delta(x, \epsilon)$:

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

Since the econometrician only observes x, a distributional assumption is typically made on ϵ, whose 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 $\delta(x, \epsilon) = d$:

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

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 upon observed and unobserved state variables x_t and ϵ_t, which are now subscripted by time.

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

Since 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 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 $f(x_{t+1}|x_t, d_t, \epsilon_t)$. Here, too, the transitions on the state variables are not time-specific, though this will be relaxed later in the paper. In order 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 $t+1$:

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

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 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:

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

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 solve:

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

2.3 Expressing the future utility component

It's rarely practical to work with equation (4) directly. Rather, we use the *value function* to express the future utility term in (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 principal of optimality, the value function can also be defined recursively as follows:

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

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

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

or, following the recursive structure, by:

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

Note that the discounted term on the right hand side of (5) is the expected future utility associated with choice d_t given 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:

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

The conditional value function is the key component to forming the conditional choice probabilities 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 (6):

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

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 (since they are primitives) whereas in the latter the expressions are more complicated (since they are constructed by solving the dynamic programming problem). How much more complicated will depend upon 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 conditional choice probabilities given in (7). The vast majority of the literature focuses on one of two cases:

1. Finite horizon models, in which the future value term is obtained via backwards recursion, and
2. 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 estimating the transition functions governing the observed states, $f(x_{t+1}|x_t, \epsilon_t)$, 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 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, knowing $g(\epsilon_T)$ implies we can (either analytically or numerically) calculate the ex ante value function at T using:

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

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

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

Continuing the backwards 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:

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

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 order to calculate $v_t(x_t, d_t)$, 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 the number of observable states is small, the backwards 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

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

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 where the function was calculated. This is the method used by Keane and Wolpin (1994) and 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 backwards and repeat the steps for $T-1$. Keane, Todd, and Wolpin (forthcoming) provide a review of the tradeoffs 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 next period's state variable by x', the conditional value function becomes:

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

In order to compute the future value term, note that the ex ante value function can be expressed as:

$$\overline{V}(x) = \int \left\{ \max_{d} \left[u(x,d) + \beta \int \overline{V}(x') f(x'|x,d) dx' \right] \right\} g(\epsilon) d\epsilon \qquad (10)$$

For example, by assuming a discrete support for the observed states, we can form an expression (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 values for $\overline{V}(x)$ and substituting these values into the right hand side of (10) to update these guesses. This process is repeated until the difference between the $\overline{V}(x)$'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. With the flow payoff of a particular decision parametrized 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:

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

where, in the stationary infinite-horizon case the t subscripts on $p_t(d_{nt}|x_{nt}, \theta)$ 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:

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

Note that the log likelihood function is the sum of two components: a part associated with the choices and a part 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 $\widehat{\theta}_2$ as given, the data on the choice probabilities can be used to estimate θ_1. While this method is not efficient (since the choice probabilities also contain information on θ_2), it can result in substantial computational savings.

Obtaining $\widehat{\theta}_1$ requires evaluating the future value term of the dynamic discrete choice problem (at each candidate θ_1) either by backwards recursion in the finite horizon case (see 2.5.1) or by solving a fixed point problem (see 2.5.2). Most of the burden of estimating dynamic discrete choice models stems from having to repeatedly solve the dynamic programming problem. The methods introduced later 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 this distribution between distributions that allow for more flexible correlation patterns and distributions where 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 specifying 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, Stern, and Trajtenberg (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 when 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 GEV ERRORS On the other hand, as originally established by McFadden, 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:

$$p_t(d_t|x_t) = \frac{\exp(v_t(x_t,d_t))}{\sum_{d_t'\in D}\exp(v_t(x_t,d_t'))} = \frac{1}{\sum_{d_t'\in D}\exp(v_t(x_t,d_t')-v_t(x_t,d_t))} \tag{11}$$

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

[2] As noted by Rust (1994a), 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 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.

where γ is Euler's constant. This closed form representation of the value function is a huge advantage in estimation since, 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 and Miller (1993). Given 1) structural errors that are additively separable from the flow payoffs, 2) conditional independence of the state transitions, and 3) independence of the structural errors over time, Hotz and Miller proved that *differences* in conditional value functions can always be expressed as functions of the choice probabilities alone. Moreover, the relationship between $p_t(d_t|x_t)$ and $\overline{V}_t(x_t)$ can be quite simple in the GEV case. Arcidiacono and 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.

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

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

The last equality has an intuitive interpretation: the ex ante value of being in state x_t can be expressed as 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 non-negative adjustment term $(-\ln[p(d_t^*|x_t)])$ which adjusts for the fact 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, in order to construct the choice probabilities in (7) and form the likelihood, we need to construct a comparable expression for $v_t(x_t, d_t)$. Using (13) to substitute for the ex-ante value function (at $t+1$) in (8), we can now write

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

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 conditional choice probabilities 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 conditional choice probabilities can be recovered in a first stage. Then, taking these as given, estimation reduces to a static multinomial logit criterion function with a pre-calculated 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 and Miller (1993), Altug and Miller (1997), and Arcidiacono and Miller (2010), we will 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 *which* choice (and hence which conditional value function) to make the future value term relative to. Since discrete choice models are estimated using *differences* in conditional value functions, in many cases a clever choice of which conditional value function to use as this benchmark can allow the difference in the future utility terms across two choices to be characterized by the one-period-ahead conditional choice probabilities alone.

3 CCP Representations

3.1 Models where only one-period-ahead choice probabilities are needed

We begin with examples where 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 and 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 (2004)), and dynamic discrete games with permanent exit (Ericson and 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:

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

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 normalized to zero.

Researchers have used this terminal choice property in a number of empirical applications. Hotz and 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 dropout 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 we will discuss later, dynamic games with permanent exit also share the terminal choice property.

Another class of models that only require one-period-ahead conditional choice probabilities are settings in which there exists a choice that makes the choice in the previous period irrelevant. An example of this 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 and 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

$$f(x_{t+1}|x_t, d_t^*)f(x_{t+2}|x_{t+1}, R) = f(x_{t+1}|x_t, d_t')f(x_{t+2}|x_{t+1}, R) \tag{15}$$

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 replacing plus the ex ante value function at $t+2$:

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

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] Since 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 which can either be estimated in a prior stage (e.g. $p_t(d_t|x_t)$ and $f(x_{t+1}|x_t, d_t)$) and 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 only destroys 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 is not affected by the

[3]Note that in Rust's (1987) empirical specification the argument is even simpler since normalizing the baseline utility of replacing the bus engine to zero for all mileage states (which is required since 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 $\{x_{t+1}, x'_{t+1}\}$.

replacement choice itself. If the price of bus engines fall and these lower prices persist, this will affect the probability of replacing. At the same time, there is still no effect of past replacement choices on the future value of a new engine conditional on knowing we are in a low price period.

3.2 Models where multiple-period-ahead choice probabilities are needed

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 and Miller (1997) 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 non-work 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 conditional choice probabilities are needed in estimation. Note that 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 and Miller's (1997) framework and the renewal examples there was an action which, when taken at $t+1$ or taken from $t+1$ to $t+\rho$, would undo the dependence of the state on the choice at t. Arcidiacono and 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 conditional choice probabilities during the time of the sequence are needed in estimation.

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 and Miller (2010), we can rewrite the expression for $v_t(x_t, d_t)$ given in (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 (8) with the one-period-ahead expression in (13). Next, substitute in for $v_{t+1}(x_{t+1}, d_{t+1}^*)$ using the one-period-ahead expression for (8). Continual repeating for ρ periods, each time expressing the value function relative to the next choice in the sequence yields:

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

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

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

Note that the last line of (18) disappears due to finite dependence. Namely, because the two choice sequences led to the same states in expectation, the last line of (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 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:

$$p_t(d^*_t|x_t) = \frac{\exp(v_t(x_t, d^*_t))}{\sum_{d_t} \exp(v_t(x_t, d_t))} = \frac{\exp(v_t(x_t, d^*_t) - v_t(x_t, A))}{\sum_{d_t} \exp(v_t(x_t, d_t) - v_t(x_t, A))}$$

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 (20), the last term in (18) cancels out. Hence, finite dependence must hold for each possible choice d_t when compared to the anchored choice A. While 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 make clear how finite dependence applies to some more complex problems tackled in the literature, we consider a simplified version of Keane and Wolpin's (1997) model of career decisions. Keane and Wolpin estimate a dynamic model of human capital accumulation in which individuals choose each period whether

[4] Note that in situations where ρ and/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.

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 where, 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 later, 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. 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 $t + 1$ will be given by:

$$x_{t+1} = \begin{bmatrix} h_{Bt} + d_{Bt} & d_{Bt} & h_{Wt} + d_{Wt} & d_{Wt} \end{bmatrix}$$

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

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

$$x_{t+3} = \left[\begin{array}{cccc} h_{Bt}+1 & 0 & h_{Wt} & 0 \end{array} \right]$$

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 that begins with staying home. Namely, the sequences $(\{B, H, B\}, \{H, B, B\})$ would both lead to the same state at $t+3$, though it would be a different state than the one that resulted from the pair of sequences employed before. Similarly, when comparing white collar to staying home, the following pairs of choices all 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 hinged on there being a choice path initiated by the anchor choice and a different choice path initiated by another choice that each led 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 mattered in the this example because the occupation in the previous period was a state variable, entering as a cost of switching occupations. However, the two sequences of choices could be lined up so that, after the choice sequences terminate, the last occupation was the same, implying the effect of the timing of the choice could be undone. Had the cost of switching occupations depended on the full history of previous choices, 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 CCP's simple. If we make the assumption 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 conditional choice probabilities. In this case, we would choose one of the pairs of sequences that resulted in finite dependence between blue collar and staying at home and choose another pair of sequences that resulted in finite dependence between white collar and staying at home. We would then use these to form the differenced conditional valuation functions between staying at home relative to blue collar and white collar respectively.

Keane and Wolpin (1997) assume the structural errors are normal and, in the case of the work choices, these errors are not additively separable. There are good reasons for making 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 (forthcoming) has both a Type I extreme value preference error as well as an error associated with the wage. Hence, the framework described here would still apply conditional on knowing the CCP's associated with both the observed states and the wage errors. As we discuss in section 4.2, some progress has been made on obtaining conditional choice probabilities in the presence of unobserved states and at that point we will return to this issue.

3.4 Future events

When finite dependence applies, it substantially weakens the assumptions needed 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 conditional choice probabilities.

Further, the transitions on the observed state variables could actually be time-dependent: rather than $f(x_{t+1}|x_t, d_t)$, we could have $f_t(x_{t+1}|x_t, d_t)$, making no assumptions regarding how the state variables transitioned 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 conditional choice probabilities. For example, in the occupational choice case, three-period-ahead conditional choice probabilities and transition functions were needed. Hence, we would only use the last three periods of data in forming conditional choice probabilities 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 and 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 non-stationary time-dependent processes for the transitions of the state variables. Knowing how these processes evolve beyond the sample period is unnecessary to estimate the model since 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 making 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 conditional choice probabilities. The other option is to have expectations regarding next period's states based on this period's state. At this point, there will be an issue of coverage of the CCP's 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 CCP frameworks

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

To show the first, 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 and Miller (1993), this expected structural error can be expressed as a function of the CCP's. 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:

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

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

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

Now, we express each of the components of (23) in vector or matrix form:

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

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

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

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

[6] While finite horizon problems without finite dependence can still be estimated using CCP techniques, the advantages of using CCP's are not as large as the full backwards recursion problem will still need to be solved.

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

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

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

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

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

Another alternative is to use the approach proposed by Hotz, Miller, Sanders and Smith (1997), which involves using 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 conditional choice probabilities, 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 provides an approximation to the expected future utility for each initial choice. The paths are drawn far enough out into the future such that the discounting renders future terms past this point irrelevant. Bajari, Benkard and Levin (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]

[7]Note that using CCP methods does not require assuming a stationary environment, but

4 CCP Estimation

We first consider the case where, apart from the iid ϵ's, there are no additional unobserved state variables. We then describe how to incorporate additional, non-iid unobservables into CCP estimation using the algorithms developed by Arcidiacono and 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 conditional choice probabilities 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 and Miller 1993, Aguirregabiria and Mira 2002 and 2007).

4.1 Estimation without unobserved states

Without unobserved heterogeneity, estimation occurs in two stages: 1) recovering the CCP's and transition functions for the state variables, and 2) forming the value function using the CCP's and estimating the structural parameters.

4.1.1 Stage 1: obtaining conditional choice probabilities and transition functions

Given unlimited data, both the CCP's and transition functions could be estimated non-parametrically using simple bin estimators. For example, the estimated probability of choice d_t given state x_t, $\widehat{p}_t(d|x_t)$ could be

will, in general, require the estimation of first stage CCPs that are fully flexible functions of time. This can obviously be quite data-intensive. However, since they do not require solving for a fixed point, CCP methods do raise the possibility of handling non-stationary, infinite horizon problems, a point which we will return to in section 5.2.

found using:

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

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 (27) infeasible. In this case, some smoothing will need to occur. Hence, non-parametric kernels, basis functions, or flexible logits could be employed. In each of these cases, collecting more data would allow for more flexible interactions of the state variables, with the promise of more interactions ensuring consistency.

A known parametric form is typically assumed for the state transition functions. Therefore, in the event that data is 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 non-parametric estimates with the CCPs implied by the structural model. We will return to this point in section 4.1.3.

4.1.2 Stage 2: estimating the structural parameters Since the previously estimated CCP 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 a 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:

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

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 and Mira (2002) use (26) to construct a CCP representation,

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

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:

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

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 CCP's impacts the results. Aguirregabiria and Mira (2002) show that the model can be used to update the CCP's in stationary infinite horizon settings, mitigating the small sample bias. To understand how their approach works, recall that the nested fixed point algorithm solved for the value function within the maximization routine. Here the steps are effectively flipped: given the CCP's, estimate the parameters, then update the CCP's 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 $\widehat{\theta}$ is:

$$p(d|x,\widehat{\theta}) = \frac{\exp(v(x,d))}{\sum_{d'}\exp(v(x,d'))} \tag{28}$$

Hence, given estimates of the structural parameters, (28) can be used to update the CCP's. Then, the value function can be updated directly using (26). By calculating the matrix inversions outside of the likelihood maximization, significant computational gains can be obtained. Aguirregabiria and 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. Backwards 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 conditional choice probabilities to forward simulate the value function, as developed by Hotz, Miller, Sanders, and Smith (1994) and later extended and applied to games by Bajari, Benkard, and Levin (2007), calculates 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, when 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 approximates the value function, though 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 which 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 and Miller (1997) in the context of female labor supply. Bajari, Benkard, and Levin (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 until discounting makes the increment to the value function small, we instead use the conditional choice probabilities 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 and Walker (forthcoming). The choice set is over fifty locations where 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. Since her model (and the Kennan and 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

4.2.1 Mixture distributions and the EM algorithm Following Heckman and Singer (1984), the standard approach to accounting 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 where the unobserved state does not vary over time, although Arcidiacono and 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:

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

Since s_n is unobserved, we integrate it out of the likelihood. Denote $\pi(s|x_{n1})$ 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:

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

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

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

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

The first order conditions to this problem are:

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

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 (31) can then be rewritten as:

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

Since $q\left(s|d_n, x_n; \widehat{\theta}, \widehat{\pi}\right)$ 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 $\widehat{\pi}(s|x_1)$, the estimated population probability of being in state s given first-period data x_1

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

Dempster, Laird, and Rubin (1977) note that the same first order condition given in (31) would hold if $q(s|d_n, x_n; \widehat{\theta}, \widehat{\pi})$ was taken as given and the maximization problem was:

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

Since $\left(\widehat{\theta}, \widehat{\pi}\right)$ are not known pre-estimation, Dempster, Laird, and Rubin developed the Expectation-Maximization (EM) algorithm which yields a solution to the first order conditions in (31) upon convergence. Namely, given initial values $\theta^{(1)}$ and $\pi^{(1)}$, the $(m+1)$th iteration is given by:

1. **Expectation Step** Update the conditional probabilities of being in each unobserved state according to:

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

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

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

2. **Maximization Step** Taking $q^{(m+1)}(s|d_n, x_n)$ as given, obtain $\theta^{(m+1)}$ from:

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

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 $q^{(m+1)}(s|d_n, x_n)$ as population weights. With s treated as observed, additive separability at the maximization step is reintroduced. Arcidiacono and Jones (2003) show that the maximization step can once again be carried out in stages. For example, we can express (38) as:

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

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

Note that the EM algorithm requires repeating the maximization step multiple times. For many dynamic discrete choice problems, this renders the EM algorithm an unattractive solution. There are two cases, however, when the EM algorithm is particularly helpful. First, if estimating in stages substantially speeds up estimation, then the methods of Arcidiacono and Jones (2003) can result in large computational savings.[8] Second, as we will see in the next section, it provides a natural way of integrating unobserved heterogeneity into CCP estimation.

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

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

Denoting d_n and x_n as the set of decisions and observed states for n and applying the law of iterated expectations, (40) can be expressed as:

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

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

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

[8]See, for example, Arcidiacono (2004), Arcidiacono (2005), Arcidiacono, Sieg, and Sloan (2007), Beffy, Fougere, and Maurel (forthcoming), and Fiorini (forthcoming).

Given that the EM algorithm provides estimates of the conditional probabilities of being in each unobserved state, we can use the sample analog of (42) evaluated at the current parameter estimates and then update the conditional choice probabilities using the EM algorithm.[9] Given initial values $\theta^{(1)}$, $\pi^{(1)}$, and the vector of conditional choice probabilities $p^{(1)}$, the $(m+1)$th iteration is given by:

1. **Expectation Step**

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

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

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

2. **Maximization Step**

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

4.3 Unobserved heterogeneity in a first stage

Updating the conditional choice probabilities in this way suggests the two-stage approach described in Arcidiacono and Miller (2010). Namely, in the first stage

[9] Arcidiacono and Miller (2010) also suggest an alternative way of updating the conditional choice probabilities in stationary environments. Namely, the CCP's can be updated from the model using:

$$p^{(m+1)}(j|x, s) = p(j|x, s, p^{(m)}; \theta_1^{(m)})$$

This is similar in spirit to the approach of Aguirregabiria and Mira (2007). This method has the advantage of not having to smooth the conditional choice probabilities when there are few individuals in a particular state.

allow the conditional choice probabilities to be completely flexible. Instead of using:

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

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

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

The first stage maximization step is then only over θ_2:

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

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.

4.4 Choices of the error distributions with unobserved states

While we have, for ease of exposition, explained the Arcidiacono and Miller approach to unobserved heterogeneity in cases where the unobserved states are permanent, their method applies to the cases where the unobserved state transitions over time. Further, due to the use of the EM algorithm, allowing the unobserved states to transition over time has little effect on computational times. While it is certainly possible to program full solution methods where the unobserved states transition over time (see Keane and 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, in Keane and Wolpin (1997), the errors associated with wages are assumed to be known to the individual, but enter multiplicatively. However, they are also assumed to be 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 tradeoff between knowing the full wage error, but assuming serial independence, versus 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. While 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 and 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 that last few years, has quickly become one of the fastest growing areas of empirical industrial organization. CCP estimation of discrete games was first proposed in Rust (1994b). Building on the logic of Hotz and Miller (1993), Rust suggested an estimation strategy that involved substituting non-parametric 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". In a series of contemporaneous papers (Aguirregabiria and Mira (2007), Bajari et al. (2007), Pakes et al. (2007), and Pesendorfer and Schmidt-Dengler (2008)), several authors have recently generalized and extended this approach, focusing on infinite horizon games with stationary Markov Perfect Equilibrium (MPE).

The use of CCP's 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 dynamic programming 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 GMM 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 non-parametric estimation of the CCP's effectively completes the model by conditioning on the equilibrium that was actually played.

5.1 Dynamic discrete games

While CCP methods can be applied to models for which the controls are continuous and/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 and Mira (2007). Consider a discrete game played by I players in each of many markets. Note that, 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 $d_t^{(i)} \in D$, as well as 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 identically and independently distributed 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 fact 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 $P\left(d_t^{(-i)}|x_t\right)$ denote the probability that firm i's competitors choose $d_t^{(-i)}$ conditional on x_t. Since $\epsilon_t^{(i)}$ is independently distributed across all the firms, $P\left(d_t^{(-i)}|x_t\right)$ can be written as the following product:

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

where $p^{(j)}\left(d_t^{(j)}|x_t\right)$ is player j's conditional choice probability. These CCP's represent "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). While existence of equilibrium follows directly from Brouwer's theorem (see Aguirregabiria and Mira (2007)), uniqueness is unlikely to hold given the inherent non-linearity of the underlying reaction functions. However, given a stationary rational expectations MPE, 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

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

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

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 $F\left(x_{t+1} \middle| x_t, d_t^{(i)}, d_t^{(-i)}\right)$ 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:

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

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:

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

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 and 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 and Schmidt-Dengler (2008)). Aguirregabiria and 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. (1997), 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 in either Aguirregabiria and Mira (2007) or Arcidiacono and Miller (2010).

5.2 Finite dependence in games

Notably, games represent a setting where 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 either be 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 CCP's.

Beresteanu, Ellickson, and Misra (2009) employ a version of this representation in their analysis of dynamic competition between retail chains. In their setting, two types of firms (supermarkets and Wal-Mart 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 *non-stationary* games or models of social interactions. Clearly, non-stationarity coupled with an infinite horizon raises issues concerning existence of equilibria. Nonetheless, the data does 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 non-stationary 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 CCP's 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 dynamic programing (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 re-solving the DP. While the structural model may only need to be solved once in order 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 re-solving 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 conditional choice probabilities observed in the data, they will be the same at $t+1$ regardless of whether the policy was in effect at t or not. 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 CCP's. The original CCP's will only be invalid during the period in which the policy is in place.

However, during this period, updated policy functions (i.e. valid CCP's) can be constructed using the new per period payoff functions by solving a backwards recursion problem that terminates at t. The future value terms for periods after t (which are state-specific) can be constructed from the original CCP's, 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 conditional choice probabilities in hand, we can forward simulate the dynamics of home ownership after the policy expires, exploiting its impact on the resulting distribution of states. While there are certainly examples of policy interventions that would permanently alter the relevant CCP's (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 we noted in the introduction, there is clearly no shortage of empirical applications in which dynamics play a central role. Moreover, high-quality panel data has 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. While 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.

Literature Cited

1. Altug, S., and R. A. Miller (1998): "The Effect of Work Experience on Female Wages and Labour Supply," *Review of Economic Studies*, 62, 45-85.

2. Aguirregabiria, V., and P. Mira (2002):"Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models, *Econometrica*, 70, 1519-1543.

3. Aguirregabiria, V., and P. Mira (2007): " Sequential Estimation of Dynamic Discrete Games," *Econometrica,* 75, 1-54.

4. Aguirregabiria, V., and P. Mira (2010): "Dynamic Discrete Choice Structural Models: A Survey," *Journal of Econometrics*, 156(1), 38-67.

5. Angrist, J. and J.S. Pischke (2010):"The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics," *The Journal of Economic Perspectives*, 24(1), 3-30.

6. Arcidiacono, P. (2004):"Ability Sorting and the Returns to College Major," *Journal of Econometrics*, 121(1-2), 343-375.

7. Arcidiacono, P. (2005): "Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica*, 73, 1477-1524.

8. Arcidiacono, P., and J. B. Jones (2003): "Finite Mixture Distributions, Sequential Likelihood, and the EM Algorithm," *Econometrica*, 71, 933-946.

9. Arcidiacono, P. and Miller (2010):"CCP Estimation of Dynamic Discrete Choice Model with Unobserved Heterogeneity," working paper.

10. Arcidiacono, P., H. Seig, and F. Sloan (2007):"Living Rationally Under the Volcano? An Empirical Analysis of Heavy Drinking and Smoking." *International Economic Review*, 48(1).

11. Bajari, P., L. Benkard, and J. Levin (2007): " Estimating Dynamic Models of Imperfect Competition," *Econometrica,* 75, 1331-1371.

12. Beauchamp, A. (2010): "Abortion Supplier Dynamics", working paper.

13. Beffy, M., D. Fougere, and A. Maurel (forthcoming):"Choosing the Field of Study in Post-Secondary Education: Do Expected Earnings Matter?" *Review of Economics and Statistics*

14. Benkard, C.L., A. Bodoh-Creed, and J. Lazarev (2010):"Simulating the Dynamic Effects of Horizontal Mergers: U.S. Airlines" working paper.

15. Beresteanu A., P.B. Ellickson, and S. Misra (2010): "The Dynamics of Retail Oligopoly, working paper.

16. Bishop, K.C. (2008):"A Dynamic Model of Location Choice and Hedonic Valuation," working paper.

17. Bresnahan, T. F., S. Stern, and M. Trajtenberg (1997): "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *RAND Journal of Economics*, 28, 17-44.

18. Cardell, N.S. (1997):"Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, 13(2), 185-213.

19. Collard-Wexler, A. (2006): "Demand Fluctuations and Plant Turnover in the Ready-Mix Concrete Industry," working paper.

20. Dempster, A. P., N. M. Laird, and D. B. Rubin (1977): "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of Royal Statistical Society Series B*, 39,1-38.

21. Dunne, T., Klimek S., Roberts, M. and Y. Xu (2006): "Entry and Exit in Geographic Markets," working paper.

22. Eckstein, Z. and K.I. Wolpin (1990):"Estimating a Market Equilibrium Search Modle from Panel Data on Individuals," *Econometrica*, 58(4), 783-808.

23. Erdem, T., and M.P. Keane (1996): "Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15(1), 1-20.

24. Fiorini, M. (forthcoming):"Fostering Educational Enrollment Through Subsidies: the Issue of Timing," *Journal of Applied Econometrics*.

25. Flinn, C. and J. Heckman (1982):"New methods for analyzing structural models of labor force dynamics," *Journal of Econometrics*, 18(1), 115-168.

26. Gowrisankaran G., and M. Rysman (2009):"Dynamics of Consumer Demand for New Durable Goods," working paper.

27. Heckman, J.J. (2010):"Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy," *Journal of Economic Literature*, 48(2), 356-398.

28. Heckman, J.J. and B. Singer (1984):"A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52(2), 271-320.

29. Hotz, V. J., and R. A. Miller (1993): " Conditional Choice Probabilities and Estimation of Dynamic Models," *Review of Economic Studies*, 61, 265-289.

30. Hotz, V. J., R. A. Miller, S. Sanders and J. Smith (1994): "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies,* 60, 265-289.

31. Joensen, J.S. (2009): "Academic and Labor Market Success: The Impact of Student Employment, Abilities, and Preferences," working paper.

32. Keane, M.P., P.E. Todd, and K.I. Wolpin (forthcoming): "The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications" *Handbook of Labor Economics.*

33. Keane, M.P., and K.I. Wolpin (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence" *The Review of Economics and Statistics*, 76, 648-672.

34. Keane, M.P., and K.I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy,* 105, 473-522.

35. Keane, M.P., and K.I. Wolpin (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," *International Economic Review,* 51(3), 851-892.

36. Kennan, J. and J.R. Walker (forthcoming): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica.*

37. McCall, J.J. (1970): "Economics of Information and Job Search," *Quarterly Journal of Economics*, 84(1), 113-126.

38. McFadden, D. (1978): "Modelling the Choice of Residential Location," *Spatial Interaction and Planning Models,* editors F. Snickars and J. Weibull, North Holland Amsterdam, 75-96.

39. Melnikov, O. (2001): "Demand for Differentiated Durable Products: The Case of the US Computer Printer Market," working paper.

40. Miller, R. A. (1984): "Job Matching and Occupational Choice," *Journal of Political Economy,* 92, 1086-1020.

41. Murphy, A. (2010): "A Dynamic Model of Housing Supply," working paper.

42. Nair, H. (2007). "Intertemporal Price Discrimination with Forward-looking

Consumers: Application to the US Market for Console Video-Games." *Quantitative Marketing and Economics*, 5(3), 239-292.

43. Pakes, A. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica,* 54, 755-784.

44. Pakes, A., M. Ostrovsky, and S. Berry (2007): " Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry/Exit Examples)," *RAND Journal of Economics*, 38(2), 373-399.

45. Pesendorfer, M., and P. Schmidt-Dengler (2008): " Asymptotic Least Square Estimators for Dynamic Games," *Review of Economic Studies*, 75, 901-908.

46. Postel-Vinay, F. and J. Robin (2002):"Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica*, 70(6), 2295-2350.

47. Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999-1033.

48. Rust, J. (1994a):"Structural Estimation of Markov Decision Processes" Chapter 51 in *Handbook of Econometrics,* 4, editors R. Engle and D. McFadden, Elsevier Science, Amsterdam, 3081-3143.

49. Rust, J. (1994b):"Estimation of dynamic structural models, problems and prospects: Discrete decision processes." In Sims, C. (Ed.) *Advances in Econometrics. Sixth World Congress.* Cambridge University Press.

50. Ryan, S. (2006): "The Costs of Environmental Regulation in a Concentrated Industry," working paper.

51. Sullivan, P. (forthcoming)."A dynamic analysis of educational attainment, occupational choices and job search," *International Economic Review.*

52. Todd, P.E. and K.I. Wolpin (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," *American Economic Review*, 96(50), 1384-1417.

53. Wolpin, K. (1984): "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," *Journal of Political Economy,* 92, 852-874.