Nonstationary Dynamic Models with Finite Dependence*

Peter Arcidiacono — Duke University & NBER
Robert A. Miller — Carnegie Mellon University

June 13, 2016

Abstract

The estimation of non-stationary dynamic discrete choice models typically requires making assumptions far beyond the length of the data. We extend the class of dynamic discrete choice models that require only a few-period-ahead conditional choice probabilities, and develop algorithms to calculate the finite dependence paths. We do this both in single agent and games settings, resulting in expressions for the value functions that allow for much weaker assumptions regarding the time horizon and the transitions of the state variables beyond the sample period.

1 Introduction

Estimation of dynamic discrete choice models is complicated by the calculation of expected future payoffs. These complications are particularly pronounced in games where the equilibrium actions and future states of the other players must be margined out to derive a player's best response. Originating with Hotz and Miller (1993), two-step methods provide a way of cheaply estimating structural payoff parameters in both single-agent and multi-agent settings. These two-step estimators first estimate

*We thank the referees, Victor Aguirregabiria, Shakeeb Khan, Jean-Marc Robin, and seminar participants at Duke, Sciences Po, Toulouse, and Toronto for helpful comments. We acknowledge support from National Science Foundation Grant Awards SES0721059 and SES0721098.

conditional choice probabilities (CCP's) and then characterize future payoffs as a function of the CCP's when estimating the structural payoff parameters.[1]

CCP estimators fall into two classes: those that exploit finite dependence, and those that do not.[2] The former entails expressing the future value term or its difference across two alternatives as a function of just a few-period ahead conditional choice probabilities and flow payoffs.[3] Intuitively, ρ period finite dependence holds when there exist two sequences of choices that lead off from different initial choices but generate the same distribution of state variables $\rho + 1$ periods later.[4]

Employing a finite dependence representation makes it possible to relax some of the assumptions about time that are commonly made when employing dynamic discrete choice models. Nonstationary infinite horizon models can be estimated when finite dependence holds. In finite horizon models, assumptions about the length of the time horizon and the evolution of the state variables beyond the sample period, can be relaxed. For example, a dynamic model of schooling requires making assumptions regarding the age of retirement, and also the functional form of utilities of older workers, although the data available to researchers might only track individuals into their twenties or thirties. Furthermore, estimation is fast because conditional choice probabilities need only be computed for a few periods ahead of the current choices.

Many papers have used the finite dependence property in estimation, often employing either a

[1]See Arcidiacono and Ellickson (2011) for a review.

[2]CCP estimators that do not rely on finite dependence include those of Hotz, Miller, Sanders, and Smith (1994), Aguirregabiria and Mira (2002, 2007), Bajari, Benkard, and Levin (2007), and Pesendorfer and Schmidt-Dengler (2008).

[3]See Hotz and Miller (1993), Altug and Miller (1998), Arcidiacono and Miller (2011), Aguirregabiria and Magesan (2013), and Gayle (2013).

[4]The sequences of choices need not be optimal and may involve mixing across choices within a period.

terminal or renewal action.[5] More general forms of finite dependence, whether a feature of the data or imposed by the authors, have been applied in models of fertility and female labor supply (Altug and Miller 1998, Gayle and Golan 2012, Gayle and Miller 2016), migration (Bishop, 2012, Coate 2013, Ma 2013, Ransom 2014), participation in the stock market (Khorunzhina 2013), agricultural land use (Scott, 2013), smoking (Matsumoto 2014), education (Arcidiacono, Aucejo, Maurel, and Ransom 2014), occupational choice (James 2014), and housing choices (Khorunzhina and Miller 2016). These papers demonstrate the advantage of exploiting finite dependence in estimation: it is not necessary to solve the value function within a nested fixed point algorithm, nor invert matrices the size of the state space.[6]

The current method for determining whether finite dependence holds or not is to guess and verify. The main contribution of this paper is to provide a systematic way of determining whether finite dependence holds when there are a (large but) finite number of states. To accomplish this, we slightly generalize the definition of finite dependence given in Arcidiacono and Miller (2011). Key to the generalization is recognizing that the ex-ante value function can be expressed as a weighted average of the conditional value functions of all the alternatives plus a function of the conditional choice probabilities, where all the weights sum to one but some may be negative or greater than one. As one of our examples shows, this slight generalization enlarges the class of models that can be cheaply estimated by exploiting this more inclusive definition of the finite dependence property.

Determining whether finite dependence holds for a pair of initial choices is a nonlinear problem,

[5]See, for example Hotz and Miller (1993), Joensen (2009), Scott (2013), Arcidiacono, Bayer, Blevins, and Ellickson (forthcoming), Declerq and Verboven (2014), Mazur (2014), and Beauchamp (2015). The last three exploit one period finite dependence to estimate dynamic games.

[6]The finite dependence property has also been directly imposed on the decision making process in models to economize on the state space. See for example Bishop (2012) and Ma (2013). Assuming players do not use all the information at their disposal reduces the state space players use to solve their optimization problems. This approach provides a parsimonious way of modeling bounded rationality when the state space is high dimensional.

yet the algorithm we propose only has a finite number of steps. We partition candidate paths for demonstrating finite dependence in say ρ periods; paths that reach the same set of states reached with a nonzero weight are collected together. Partitioning by whether a weight is zero or not, rather than the value of the weight, reduces an uncountable infinity of paths to a finite set. Each element in the partition maps into a linear system of equations, and we check the rank of the system, also a finite number of operations. The size of the linear system is based on the number of states attainable in $\rho - 1$ periods from the initial state, not the total number of states in the model. The algorithm proceeds iteratively, by checking the determinants of selected elements in the partition. If one (or more) of the elements has a nonzero determinant, then the pair of choices exhibits ρ period finite dependence; otherwise it does not. Once finite dependence is established, another linear operation (on a finite number of equations) yields a set of weights that can be used in any CCP estimator that exploits finite dependence.

Many estimators exploiting finite dependence have an intrinsically linear structure. From the standpoint of computational efficiency and accuracy, the advantages of linear over nonlinear solution methods are well known. For example the Monte Carlo applications given in our previous work (Arcidiacono and Miller, 2011) compare CCP estimators exploiting the finite dependence and linearity with nonlinear Maximum Likelihood estimators. We find the CCP estimators are much cheaper to compute and are almost as precise as MLE even in low dimensional problems, where nonlinear methods are least likely to be computationally burdensome.

In game settings, finite dependence is applicable to each player individually. Here finite dependence relates to transition matrices for the state variables when a designated player places arbitrary weight on each of her possible future decisions (so long as the weights sum to one within a period) and the other players follow their equilibrium strategies. Consequently, finite dependence in games cannot be ascertained from the transition primitives alone (as in the individual optimization case).

Indeed, whether or not finite dependence holds might also hinge on which equilibrium is played, not a paradoxical result, because different equilibria for the same game sometimes reveal different information about the primitives, so naturally require different estimation approaches.

Research on finite dependence in games has been restricted, up until now, to cases where there is a terminal or renewal action (that ends or restarts the process governing the state variables for individual players). Absent these two cases, one-period finite dependence fails to hold, because the equilibrium actions of the other players depend on what the designated agent has already done. Hence the distribution of the state variables, which the other players partly determine, depends on the actions of the designated player two periods earlier. These stochastic connections, a vital feature of many strategic interactions, has limited empirical research in estimating games with nonstationarities. We develop an algorithm to solve for finite dependence in a broader class of games than those characterized by terminal and renewal actions. As in the single agent case, the algorithm entails solving a linear system of equations where the number of equations is dictated by the possible states that can be reached a few periods ahead.

The rest of the paper proceeds as follows. Section 2 lays out our framework for analyzing finite dependence in discrete choice dynamic optimization and games. In Section 3 we define finite dependence, provide a new representation of this property, and use the representation to demonstrate how to recover finite dependence paths in both single agent and multi-agent settings. New examples of finite dependence, derived using the algorithm, are provided in Section 4, while Section 5 concludes with some remarks on outstanding questions that future research might address.

2 Framework

This section first lays out a general class of dynamic discrete choice models. Drawing upon our previous work (Arcidiacono and Miller, 2011), we extend our representation of the conditional value functions which plays an overarching role in our analysis, and then modify our framework to accommodate games with private information.

2.1 Dynamic optimization discrete choice

In each period $t \in \{1, \ldots, T\}$ until $T \leq \infty$, an individual chooses among J mutually exclusive actions. Let d_{jt} equal one if action $j \in \{1, \ldots, J\}$ is taken at time t and zero otherwise. The current period payoff for action j at time t depends on the state $x_t \in \{1, \ldots, X\}$.[7] If action j is taken at time t, the probability of x_{t+1} occurring in period $t+1$ is denoted by $f_{jt}(x_{t+1}|x_t)$.

The individual's current period payoff from choosing j at time t is also affected by a choice-specific shock, ϵ_{jt}, which is revealed to the individual at the beginning of the period t. We assume the vector $\epsilon_t \equiv (\epsilon_{1t}, \ldots, \epsilon_{Jt})$ has continuous support, is drawn from a probability distribution that is independently and identically distributed over time with density function $g(\epsilon_t)$, and satisfies $E[\max\{\epsilon_{1t}, \ldots, \epsilon_{Jt}\}] \leq M < \infty$. The individual's current period payoff for action j at time t is modeled as $u_{jt}(x_t) + \epsilon_{jt}$.

The individual takes into account both the current period payoff as well as how his decision today will affect the future. Denoting the discount factor by $\beta \in (0, 1)$, the individual chooses the vector $d_t \equiv (d_{1t}, \ldots, d_{Jt})$ to sequentially maximize the discounted sum of payoffs:

$$E\left\{\sum_{t=1}^{T}\sum_{j=1}^{J}\beta^{t-1}d_{jt}\left[u_{jt}(x_t) + \epsilon_{jt}\right]\right\} \tag{1}$$

where at each period t the expectation is taken over the future values of $x_{t+1}, \ldots, x_T$ and $\epsilon_{t+1}, \ldots, \epsilon_T$. Expression (1) is maximized by a Markov decision rule which gives the optimal action conditional on t, x_t, and ϵ_t. We denote the optimal decision rule at t as $d_t^o(x_t, \epsilon_t)$, with j^{th} element $d_{jt}^o(x_t, \epsilon_t)$. The probability of choosing j at time t conditional on x_t, $p_{jt}(x_t)$, is found by taking $d_{jt}^o(x_t, \epsilon_t)$ and

[7] Our analysis is based on the assumption that x_t belongs to a finite set, an assumption that is often made in this literature. See Aguirregabiria and Mira (2002) for example. However it is worth mentioning that finite dependence can be applied without making that assumption. See Altug and Miller (1998) for example.

integrating over ϵ_t:

$$p_{jt}(x_t) \equiv \int d^o_{jt}(x_t, \epsilon_t)\, g(\epsilon_t)\, d\epsilon_t \tag{2}$$

We then define $p_t(x_t) \equiv (p_{1t}(x_t), \ldots, p_{Jt}(x_t))$ as the vector of conditional choice probabilities.

Denote $V_t(x_t)$, the ex-ante value function in period t, as the discounted sum of expected future payoffs just before ϵ_t is revealed and conditional on behaving according to the optimal decision rule:

$$V_t(x_t) \equiv E\left\{\sum_{\tau=t}^{T}\sum_{j=1}^{J} \beta^{\tau-t} d^o_{j\tau}(x_\tau, \epsilon_\tau)\left(u_{j\tau}(x_\tau) + \epsilon_{j\tau}\right)\right\}$$

Given state variables x_t and choice j in period t, the expected value function in period $t+1$, discounted one period into the future, is $\beta \sum_{x_{t+1}=1}^{X} V_{t+1}(x_{t+1}) f_{jt}(x_{t+1}|x_t)$. Under standard conditions, Bellman's principle applies and $V_t(x_t)$ can be recursively expressed as:

$$V_t(x_t) = \sum_{j=1}^{J} \int d^o_{jt}(x_t, \epsilon_t)\left[u_{jt}(x_t) + \epsilon_{jt} + \beta \sum_{x_{t+1}=1}^{X} V_{t+1}(x_{t+1}) f_{jt}(x_{t+1}|x_t)\right] g(\epsilon_t)\, d\epsilon_t$$

We then define the choice-specific conditional value function, $v_{jt}(x_t)$, as the flow payoff of action j without ϵ_{jt} plus the expected future utility conditional on following the optimal decision rule from period $t+1$ on:[8]

$$v_{jt}(x_t) = u_{jt}(x_t) + \beta \sum_{x_{t+1}=1}^{X} V_{t+1}(x_{t+1}) f_{jt}(x_{t+1}|x_t) \tag{3}$$

Our analysis is based on a representation of $v_{jt}(x_t)$ that slightly generalizes Theorem 1 of Arcidiacono and Miller (2011). Both results are based on their Lemma 1, that for every $t \in \{1, \ldots, T\}$ and $p \in \Delta^J$, the J dimensional simplex, there exists a real-valued function $\psi_j(p)$ such that:

$$\psi_j[p_t(x)] \equiv V_t(x) - v_{jt}(x) \tag{4}$$

To interpret (4), note that the value of committing to action j at period t before seeing ϵ_t and behaving optimally thereafter is $v_{jt}(x_t) + E[\epsilon_{jt}]$. Therefore the expected loss from pre-committing to j versus waiting until ϵ_t is observed and only then making an optimal choice, $V_t(x_t)$, is the constant

[8] For ease of exposition we refer to $v_{jt}(x_t)$ as the conditional value function in the remainder of the paper.

$\psi_j[p_t(x_t)]$ minus $E\left[\epsilon_{jt}\right]$, a composite function that only depends on x_t through the conditional choice probabilities. This result leads to the following theorem, proved using an induction.

Theorem 1 *For each choice $j \in \{1, \ldots, J\}$ and $\tau \in \{t+1, \ldots, T\}$, let any $\omega_\tau(x_\tau, j)$ denote any mapping from the state space $\{1, \ldots, X\}$ to R^J satisfying the constraints that $|\omega_{k\tau}(x_\tau, j)| < \infty$ and $\sum_{k=1}^{J} \omega_{k\tau}(x_\tau, j) = 1$. Recursively define $\kappa_\tau(x_{\tau+1}|x_t, j)$ as:*

$$\kappa_\tau(x_{\tau+1}|x_t, j) \equiv \begin{cases} f_{jt}(x_{t+1}|x_t) & \textit{for } \tau = t \\ \sum_{x_\tau=1}^{X} \sum_{k=1}^{J} \omega_{k\tau}\left(x_\tau, j\right) f_{k\tau}(x_{\tau+1}|x_\tau) \kappa_{\tau-1}(x_\tau|x_t, j) & \textit{for } \tau = t+1, \ldots, T \end{cases} \tag{5}$$

Then for $\mathcal{T} < T$:

$$\begin{aligned} v_{jt}(x_t) &= u_{jt}(x_t) + \sum_{\tau=t+1}^{\mathcal{T}} \sum_{k=1}^{J} \sum_{x_\tau=1}^{X} \beta^{\tau-t}\left[u_{k\tau}(x_\tau) + \psi_k[p_\tau(x_\tau)]\right] \omega_{k\tau}(x_\tau, j) \kappa_{\tau-1}(x_\tau|x_t, j) \\ &\quad + \sum_{x_{\mathcal{T}+1}}^{X} \beta^{\mathcal{T}+1-t} V_{\mathcal{T}+1}(x_{\mathcal{T}+1}) \kappa_{\mathcal{T}}(x_{\mathcal{T}+1}|x_t, j) \end{aligned} \tag{6}$$

and for $\mathcal{T} = T$:

$$v_{jt}(x_t) = u_{jt}(x_t) + \sum_{\tau=t+1}^{\mathcal{T}} \sum_{k=1}^{J} \sum_{x_\tau=1}^{X} \beta^{\tau-t}\left[u_{k\tau}(x_\tau) + \psi_k[p_\tau(x_\tau)]\right] \omega_{k\tau}(x_\tau, j) \kappa_{\tau-1}(x_\tau|x_t, j) \tag{7}$$

For the purposes of this work it is convenient to interpret $\mathcal{T}$ as the final period in the sample; typically $\mathcal{T} < T$. Arcidiacono and Miller (2011) prove the theorem when $T = \mathcal{T}$ and $\omega_{k\tau}(x_\tau, j) \geq 0$ for all k and τ. In that case, $\kappa_\tau(x_{\tau+1}|x_t, j)$ is the probability of reaching $x_{\tau+1}$ by following the sequence defined by $\omega_\tau(x_\tau, j)$ and the value function representation extending over the whole decision-making horizon.[9]

2.2 Extension to dynamic games

This framework extends naturally to dynamic games. In the games setting, we assume that there are N players making choices in periods $t \in \{1, \ldots, T\}$. The systematic part of payoffs to the

[9]The extension to negative weights is also noted in Gayle (2013).

n^{th} player not only depends on his own choice in period t, denoted by $d_t^{(n)} \equiv \left(d_{1t}^{(n)}, \ldots, d_{Jt}^{(n)}\right)$, and the state variables x_t, but also the choices of the other players, which we now denote by $d_t^{(\sim n)} \equiv \left(d_t^{(1)}, \ldots, d_t^{(n-1)}, d_t^{(n+1)}, \ldots, d_t^{(N)}\right)$. Denote by $U_{jt}^{(n)}\left(x_t, d_t^{(\sim n)}\right) + \epsilon_{jt}^{(n)}$ the current utility of player n in period t, where $\epsilon_{jt}^{(n)}$ is an identically and independently distributed random variable that is private information to player n. Although the players all face the same observed state variables, these state variables typically affect players in different ways. For example, adding to the n^{th} player's capital may increase his payoffs and reduce the payoffs to the others. For this reason the payoff function is superscripted by n.

Each period the players make simultaneous choices. We denote by $P_t\left(d_t^{(\sim n)} | x_t\right)$ the joint conditional choice probability that the players aside from n collectively choose $d_t^{(\sim n)}$ at time t given the state variables x_t. Since $\epsilon_t^{(n)}$ is independently distributed across all the players, $P_t\left(d_t^{(\sim n)} | x_t\right)$ has the product representation:

$$P_t\left(d_t^{(\sim n)} | x_t\right) = \prod_{\substack{n'=1 \\ n' \neq n}}^{N} \left(\sum_{j=1}^{J} d_{jt}^{(n')} p_{jt}^{(n')}(x_t)\right) \tag{8}$$

We assume each player acts like a Bayesian when forming his beliefs about the choices of the other players and that a Markov-perfect equilibrium is played. Hence, the beliefs of the players match the probabilities given in equation (8). Taking the expectation of $U_{jt}^{(n)}\left(x_t, d_t^{(\sim n)}\right)$ over $d_t^{(\sim n)}$, we define the systematic component of the current utility of player n as a function of the state variables as:

$$u_{jt}^{(n)}(x_t) = \sum_{d_t^{(\sim n)} \in J^{N-1}} P_t\left(d_t^{(\sim n)} | x_t\right) U_{jt}^{(n)}\left(x_t, d_t^{(\sim n)}\right) \tag{9}$$

For future reference we call $u_{jt}^{(n)}(x_t)$ the reduced form payoff to player n from taking action j in period t when the state is x_t.

The values of the state variables at period $t+1$ are determined by the period t choices by all the players as well as the values of the period t state variables. We consider a model in which the state variables can be partitioned into those that are affected by only one of the players, and those that are exogenous. For example, to explain the number and size of firms in an industry, the state variables for the model might be indicators of whether each potential firm is active or not, and a scalar to measure firm capital or capacity; each firm controls their own state variables, through their entry and exit choices, as well as their investment decisions.[10] The partition can be expressed as $x_t \equiv \left(x_t^{(0)}, x_t^{(1)}, \ldots, x_t^{(N)}\right)$, where $x_t^{(0)}$ denotes the states that are exogenously determined by transition probability $f_{0t}\left(x_{t+1}^{(0)} \middle| x_t^{(0)}\right)$, and $x_t^{(n)} \in \mathcal{X}^{(n)} \equiv \left\{1, \ldots, X^{(n)}\right\}$ is the component of the state controlled or influenced by player n. Let $f_{jt}^{(n)}\left(x_{t+1}^{(n)} \middle| x_t^{(n)}\right)$ denote the probability that $x_{t+1}^{(n)}$ occurs at time $t+1$ when player n chooses j at time t given $x_t^{(n)}$. Many models in industrial organization exploit this specialized structure because it provides a flexible way for players to interact while keeping the model simple enough to be empirically tractable.[11] Since the transitions of the exogenous variables do not substantively effect our analysis, we ignore them for the rest of the paper to conserve on notation.

Denote the state variables associated with all the players aside from n as:

$$x_t^{(\sim n)} \equiv \left(x_t^{(1)}, \ldots, x_t^{(n-1)}, x_t^{(n+1)} \ldots, x_t^{(N)}\right) \in \mathcal{X}^{(\sim n)} \equiv \mathcal{X}^{(1)} \times \ldots \times \mathcal{X}^{(n-1)} \times \mathcal{X}^{(n+1)} \times \ldots \times \mathcal{X}^{(N)}$$

Under this specification the reduced form transition generated by their equilibrium choice probabilities is defined as:

$$f_t^{(\sim n)}\left(x_{t+1}^{(\sim n)} \left| x_t\right.\right) \equiv \prod_{\substack{n'=1 \\ n' \neq n}}^{N} \left[\sum_{k=1}^{J} p_{kt}^{(n')}(x_t) f_{kt}^{(n')}\left(x_{t+1}^{(n')} \middle| x_t^{(n')}\right)\right]$$

As in Subsection 2.1, consider for all $\tau \in \{t, \ldots, T\}$ any sequence of decision weights:

$$\omega_\tau^{(n)}(x_\tau, j) \equiv \left(\omega_{1\tau}^{(n)}(x_\tau, j), \ldots, \omega_{J\tau}^{(n)}(x_\tau, j)\right)$$

[10] The second example in Arcidiacono and Miller (2011) also belongs to this class of models.

[11] All the empirical applications of structural modeling of which we are aware have this property, including those based on Ericson and Pakes (1995). For example, firms affect their own product quality through their own investment decisions, but do not directly affect the product quality of other players. Thus each firm's decisions affect the product quality of other players only through the effect on the decisions of the other players.

subject to the constraints $\sum_{k=1}^{J} \omega_{k\tau}^{(n)}(x_\tau, j) = 1$ and starting value $\omega_{jt}^{(n)}(x_t, j) = 1$. Given the equilibrium actions of the other players impounded in $f_t^{(\sim n)}\left(x_{t+1}^{(\sim n)}|x_t\right)$, we recursively define $\kappa_\tau^{(n)}(x_{\tau+1}|x_t, j)$ for the sequence of decision weights $\omega_{k\tau}^{(n)}(x_\tau, j)$ over periods $\tau \in \{t+1, \ldots, T\}$ in a similar manner to Equation (5) as:

$$\kappa_\tau^{(n)}(x_{\tau+1}|x_t, j) \equiv f_{0\tau}\left(x_{\tau+1}^{(0)}\left|x_\tau^{(0)}\right.\right) \sum_{x_\tau=1}^{X} \sum_{k=1}^{J} f_\tau^{(\sim n)}\left(x_{\tau+1}^{(\sim n)}|x_\tau\right) \omega_{k\tau}^{(n)}(x_\tau, j)\, f_{k\tau}^{(n)}\left(x_{\tau+1}^{(n)}\left|x_\tau^{(n)}\right.\right) \kappa_{\tau-1}^{(n)}(x_\tau|x_t, j)$$

with initializing function:

$$\kappa_t^{(n)}(x_{t+1}|x_t, j) \equiv f_{jt}^{(n)}\left(x_{t+1}^{(n)}\left|x_t^{(n)}\right.\right) f_t\left(x_{t+1}^{(\sim n)}|x_t\right) f_{0t}\left(x_{t+1}^{(0)}\left|x_t^{(0)}\right.\right)$$

Letting:

$$f_{jt}(x_{t+1}|x_t) = f_{0t}\left(x_{t+1}^{(0)}\left|x_t^{(0)}\right.\right) f_t^{(\sim n)}\left(x_{t+1}^{(\sim n)}|x_t\right) f_{jt}^{(n)}\left(x_{t+1}^{(n)}\left|x_t^{(n)}\right.\right) \tag{10}$$

and adding n superscripts to all the other terms in (7), it now follows that Theorem 1 applies to this multi-agent setting in exactly the same way as in a single agent setting.

3 Finite dependence

If there were transition matrices satisfying the equality $\kappa_{\mathcal{T}}^*(x_{\mathcal{T}+1}|x_t, 1) = \kappa_{\mathcal{T}}^*(x_{\mathcal{T}+1}|x_t, j)$, then (6) implies differences in the conditional value functions $v_{jt}(x_t) - v_{1t}(x_t)$ could be expressed as a weighted sum of flow payoffs and $\psi_k(\cdot)$ terms that occur between t and $\mathcal{T}$. Finite dependence is the natural generalization of an equality like $\kappa_{\mathcal{T}}^*(x_{\mathcal{T}+1}|x_t, 1) = \kappa_{\mathcal{T}}^*(x_{\mathcal{T}+1}|x_t, j)$. It captures the notion that the differential effects on the state variable from taking two distinct actions in period t might be obliterated, say ρ periods later, if certain corrective paths are followed that are specific to the initial action.

3.1 Defining finite dependence

Consider two sequences of decision weights that begin at date t in state x_t, one with choice i and the other with choice j. We say that the pair of choices $\{i, j\}$ exhibits ρ-period dependence if there exist sequences of decision weights from i and j for x_t such that :

$$\kappa_{t+\rho}(x_{t+\rho+1}|x_t, i) = \kappa_{t+\rho}(x_{t+\rho+1}|x_t, j) \tag{11}$$

for all $x_{t+\rho+1}$. That is, the weights associated with each state are the same across the two paths after ρ periods.[12]

Several comments on this definition are in order. First, finite dependence trivially holds in all finite horizon problems. However the property of ρ-period dependence only merits attention when $\rho < T - t$. To avoid repeatedly referencing the trivial case of $\rho = T - t$, we will henceforth write finite dependence holds only when (11) applies for $\rho < T - t$. Second, finite dependence is defined with respect to a pair of choices conditional on the value of the state variable, not the whole model. The main reason for this narrow definition is that finite dependence might hold for some choice pairs but not others, and for some certain states but not others. Even in this case, we can reduce the computational burden of estimating the model by exploiting finite dependence on the pairs of choices where it holds. Third, as explained in Arcidiacono and Miller (2016), finite dependence between just two choices in a single agent setting where there are J choices each period, helps in identifying counterfactual regimes generated by temporary changes in the transition matrix. Finally a more general definition of finite dependence would encompass mixed choices to start the sequence, not just pure strategies; our analysis easily extends to the more general case.

[12] Aguirregabiria and Magesan (2013, 2016) and Gayle (2013) restrict their analyses to cases where there is one period finite dependence, thus ruling out labor supply applications such as Altug and Miller (1998), as well as games that do not have a terminal choice.

Under finite dependence, differences in current utility $u_{jt}(x_t) - u_{it}(x_t)$ can be expressed as:

$$\begin{aligned} u_{jt}(x_t) - u_{it}(x_t) &= \psi_i[p_t(x_t)] - \psi_j[p_t(x_t)] \\ &+ \sum_{\tau=t+1}^{t+\rho} \sum_{k=1}^{J} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \left\{ u_{k\tau}(x_\tau) + \psi_k[p_\tau(x_\tau)] \right\} \begin{bmatrix} \omega_{k\tau}(x_\tau, i)\kappa_{\tau-1}(x_\tau|x_t, i) \\ -\omega_{k\tau}(x_\tau, j)\kappa_{\tau-1}(x_\tau|x_t, j) \end{bmatrix} \end{aligned} \tag{12}$$

This equation follows directly from Equations (4) and (7), in Theorem 1.[13] As the empirical applications of finite dependence illustrate, equations like (12) provide the basis for estimation without resorting to the inversion of high dimension matrices or long simulations. Aside from its computational benefits, finite dependence has a second attractive feature–empirical content–because it is straightforward to test whether (11) is rejected by the data.

3.2 One-period dependence in optimization problems with two choices

As foreshadowed in the Introduction, the algorithm for determining ρ-period dependence for $\rho > 1$ iterates between two procedures: checking the rank of a matrix, and listing the elements of the matrix. The procedure is simpler to establish one-period dependence as there are no intermediate decisions between the initial choice and the choice of weights that generate finite dependence. Hence, checking the rank of a particular matrix is sufficient for determining one-period dependence.

There is a second reason for investigating one-period dependence before analyzing the more general case. Because the guess and verify method is essentially the only method researchers have to determine finite dependence, almost all empirical applications of finite dependence have exploited two special cases of one-period dependence, models with two choices where one of them is either a terminal or a renewal choice. Terminal choices end the optimization problem or game by preventing any future decisions; irreversible sterilization against future fertility (Hotz and Miller, 1993), and firm exit from an industry (Aguirregabiria and Mira, 2007; Pakes, Ostrovsky, and Berry, 2007) are

[13] Appealing to (4), replace $v_{jt}(x)$ with $V_t(x) - \psi_j[p_t(x)]$ in (7) and perform a similar substitution for $v_{it}(x)$. Upon differencing the two equations, the $V_t(x)$ terms drop out.

examples. The defining feature of a renewal choice is that it resets the states that were influenced by past actions. Turnover and job matching (Miller, 1984), or replacing a bus engine (Rust, 1987), are illustrative of renewal actions. In such models, following any choice with a terminal or renewal choice yields the same value of the state variable after two periods. Therefore the key difference between terminal and renewal actions is that the former end the dynamic sequence, turning the optimization problem into a stopping problem. Designate the first choice as the terminal or renewal choice. Following any choice $j \in \{1, \dots, J\}$ with a terminal or renewal choice leads to same value of state variables after two periods, because for all x_{t+2}:

$$\sum_{x_{t+1}=1}^{X} f_{1,t+1}(x_{t+2}|x_{t+1}) f_{jt}(x_{t+1}|x_t) = \sum_{x_{t+1}=1}^{X} f_{1,t+1}(x_{t+2}|x_{t+1}) f_{1t}(x_{t+1}|x_t) \tag{13}$$

Therefore Equation (11) is satisfied at $t+2$ for all $j \in \{1, \dots, J\}$ and $x \in \mathcal{X}$ by setting weights $\omega_{k,t+1}(x_{t+1}, j) = 1$ if $k = 1$ and zero otherwise.

We begin a systematic search for finite dependence by analyzing the special case of one-period dependence where there are two choices. Formally, the definition of $\kappa_{t+1}(x'|x_t, j)$ given by Equation (5) implies that one-period dependence holds in this specialization at x_t if and only if there exists a weighting rule such that $\kappa_{t+1}(x'|x_t, 1) = \kappa_{t+1}(x'|x_t, 2)$ for all $x' \in \mathcal{X}$. Since $J = 2$ and the weights sum to one, we can economize on subscripts by setting $\omega_{t+1}(x_{t+1}, j) \equiv \omega_{2,t+1}(x_{t+1}, j)$, the weight on the second action. Thus $\omega_{t+1}(x_{t+1}, j)$ must solve:

$$\begin{aligned} &\sum_{x_{t+1}=1}^{X} \left\{ \begin{array}{l} [f_{2,t+1}(x'|x_{t+1}) - f_{1,t+1}(x'|x_{t+1})] \\ \times [\omega_{t+1}(x_{t+1}, 2) f_{2t}(x_{t+1}|x_t) - \omega_{t+1}(x_{t+1}, 1) f_{1t}(x_{t+1}|x_t)] \end{array} \right\} \\ = &\sum_{x_{t+1}=1}^{X} f_{1,t+1}(x'|x_{t+1}) \left[f_{1t}(x_{t+1}|x_t) - f_{2t}(x_{t+1}|x_t) \right] \end{aligned} \tag{14}$$

for all $x' \in \mathcal{X}$. Nominally this is a linear system of $X-1$ equations in $\omega_{t+1}(x_{t+1}, 1)$ and $\omega_{t+1}(x_{t+1}, 2)$; if the $X-1$ equations are satisfied for all but one of the state variables, the equation associated with the remaining state will automatically be satisfied since summing $\kappa_{t+1}(x'|x_t, j)$ over x' equals one.

The dimension of $\omega_{t+1}(x_{t+1}, j)$ is X for each $j \in \{1, 2\}$. Therefore there are fewer equations than unknowns. However, if a state is not reached at $t + 1$, then changing the weight placed on an action at that state cannot help in obtaining finite dependence. Threrefore we need only consider states at $t + 1$ that can be reached with positive probability from at least one of the initial choices.

The fact that some of the states may not be reached at $t + 1$ regardless of the initial choice effectively reduces the number of relevant unknowns in the system. Another feature of the system reduces the relevant number of equations. The equations associated with states at $t + 2$ that cannot be reached given either initial choice are automatically satisfied: given either initial choice, the weight on these states at $t + 2$ is zero.

We can incorporate these two features into the system of equations given by (14) as follows. Suppose $A_{j,t+1}$ states can be reached with positive probability in period $t + 1$ from state x_t with choice j at time t, and denote their set by $\mathcal{A}_{j,t+1} \subseteq \mathcal{X}$. Thus $x \in \mathcal{A}_{j,t+1}$ if and only if $f_{jt}(x|x_t) > 0$. Let $\mathcal{A}_{t+2} \subseteq \mathcal{X}$ denote the states that can be reached with positive probability in period $t + 2$ from any element in the union $\mathcal{A}_{1,t+1} \bigcup \mathcal{A}_{2,t+1}$ with either action at $t + 1$. Thus $x' \in \mathcal{A}_{t+2}$ if and only if $f_{k,t+1}(x'|x) > 0$ for some $x \in \mathcal{A}_{1,t+1} \bigcup \mathcal{A}_{2,t+1}$ and $k \in \{1, 2\}$. Finally, denote by A_{t+2} the number of states in $\mathcal{A}_{t+2}(x_t)$. It now follows that the matrix-equivalent of Equation (14) reduces to a linear system of $A_{t+2} - 1$ equations with $A_{1,t+1} + A_{2,t+1}$ unknowns.[14]

Denote by $\mathsf{K}_{jt}(\mathcal{A}_{j,t+1})$ the $A_{j,t+1}$ dimensional vector of nonzero probabilities in the string: $f_{jt}(1|x_t), \ldots, f_{jt}(X|x_t)$. It gives the one period transition probabilities to $\mathcal{A}_{j,t+1}$ from x_t when choice j is made. Let $\mathsf{F}_{k,t+1}(\mathcal{A}_{j,t+1})$ denote the first $A_{t+2} - 1$ columns of the $A_{j,t+1} \times A_{t+2}$ transition matrix from $\mathcal{A}_{j,t+1}$ to $\mathcal{A}_{t+2}$ when choice k is made in period $t + 1$.[15] A typical element of $\mathsf{F}_{k,t+1}(\mathcal{A}_{j,t+1})$ is

[14]We can remove one equation from the A_{t+2} system because if the weights associated with each state match for $A_{t+2} - 1$ states, they must also match for the remaining state

[15]We focus on the first $A_{t+2} - 1$ columns because the last column must be given by one minus the sum of the previous columns.

$f_{k,t+1}(x'|x)$ where $x \in \mathcal{A}_{j,t+1}$ and $x' \in \mathcal{A}_{t+2}$. Note that some elements of $\mathsf{F}_{k,t+1}(\mathcal{A}_{j,t+1})$ may be zero. Finally, let $\Omega_{t+1}(\mathcal{A}_{j,t+1}, j)$ denote an $A_{j,t+1}$ dimensional vector of weights on each of the attainable states at $t+1$ for taking the second choice at that time given initial choice j, comprising elements $\omega_{t+1}(x, j)$ for each $x \in \mathcal{A}_{j,t+1}$.

To see how these matrices relate to (14), momentarily consider what would happen if all the states were attainable at both $t+2$ and $t+1$ given an initial state x_t and initial choice j. In this case:

$$\mathcal{A}_{1,t+1} = \mathcal{A}_{2,t+1} = \mathcal{A}_{t+2} = \mathcal{X}, \quad \Omega_{t+1}(\mathcal{A}_{j,t+1}, j) = \Omega_{t+1}(\mathcal{X}, j), \quad \mathsf{K}_{jt}(\mathcal{A}_{j,t+1}) = \mathsf{K}_{jt}(\mathcal{X})$$

so we can write:

$$\Omega_{t+1}(\mathcal{X}, j) \circ \mathsf{K}_{jt}(\mathcal{X}) = \begin{bmatrix} \omega_{t+1}(1, j) f_{jt}(1|x_t) & \dots & \omega_{t+1}(X, j) f_{jt}(X|x_t) \end{bmatrix}'$$

where $\circ$ refers to element-by-element multiplication. Also $\mathsf{F}_{k,t+1}(\mathcal{A}_{j,t+1})$ becomes the $t+1$ transition matrix given choice k, less one column, say:

$$\mathsf{F}_{k,t+1}(\mathcal{A}_{j,t+1}) = \mathsf{F}_{k,t+1}(\mathcal{X}) = \begin{bmatrix} f_{k,t+1}(1|1) & \dots & f_{k,t+1}(X-1|1) \\ \vdots & \ddots & \dots \\ f_{k,t+1}(1|X) & \dots & f_{k,t+1}(X-1|X) \end{bmatrix}$$

Stacking the equations in (14) for all $x' \in \{1, \dots, X-1\}$, the left hand side of the stack is a linear combination of four expressions, each taking the form:

$$\begin{bmatrix} \sum_{x_{t+1}=1}^{X} f_{k,t+1}(1|x_{t+1}) \omega_{t+1}(x_{t+1}, j) f_{jt}(x_{t+1}|x_t) \\ \vdots \\ \sum_{x_{t+1}=1}^{X} f_{k,t+1}(X-1|x_{t+1}) \omega_{t+1}(x_{t+1}, j) f_{jt}(x_{t+1}|x_t) \end{bmatrix} = [\mathsf{F}_{k,t+1}(\mathcal{X})]' [\Omega_{t+1}(\mathcal{X}, j) \circ \mathsf{K}_{jt}(\mathcal{X})] \quad (15)$$

Note that when $k = 2$, Equation (15) is the weight for each element of $\mathcal{X}$ when the initial choice j is followed by the second choice.

Typically not all states in $\mathcal{X}$ are attainable at period $t+1$ given initial choice j. For all $\widetilde{x} \notin \mathcal{A}_{j,t+1}$, that is when $f_{jt}(\widetilde{x}|x_t) = 0$, we remove the element $\omega_{t+1}(\widetilde{x}, j) f_{jt}(\widetilde{x}|x_t)$ from $\Omega_{t+1}(\mathcal{X}, j) \circ \mathsf{K}_{jt}(\mathcal{X})$ and the $\widehat{x}^{th}$ row in $\mathsf{F}_{k,t+1}(\mathcal{X})$. This reduces the dimension of $\Omega_{t+1}(\mathcal{X},j)\circ\mathsf{K}_{jt}(\mathcal{X})$ to $A_{j,t+1}$ and the dimension of $\mathsf{F}_{k,t+1}(\mathcal{X})$ from $X\times(X-1)$ to $A_{j,t+1}\times(X-1)$. Similarly, if $\widehat{x}\notin\mathcal{A}_{t+2}$, in words if $\widehat{x}$ is unattainable given either initial choice regardless of the weighting rules at $t+1$, then we remove the $\widehat{x}^{th}$ column of $\mathsf{F}_{k,t+1}(\mathcal{X})$, which is a vector of zeros. The transition matrix $\mathsf{F}_{k,t+1}(\mathcal{A}_{j,t+1})$ is then a $A_{j,t+1}\times(A_{t+2}-1)$ matrix.

Substituting these transformations into (14) we now express the system of $A_{t+2}-1$ equations with $A_{1,t+1}+A_{2,t+1}$ unknowns in matrix form. Define the $A_{t+2}-1$ dimensional vector $\mathcal{K}_{t+1}$, and the $(A_{t+2}-1)\times(A_{1,t+1}+A_{2,t+1})$ matrix H_{t+1}, respectively as:

$$\mathcal{K}_{t+1}\equiv\begin{bmatrix}\mathsf{F}_{1,t+1}(\mathcal{A}_{1,t+1})\\ \\ -\mathsf{F}_{1,t+1}(\mathcal{A}_{2,t+1})\end{bmatrix}'\begin{bmatrix}\mathsf{K}_{1t}(\mathcal{A}_{1,t+1})\\ \\ \mathsf{K}_{2t}(\mathcal{A}_{2,t+1})\end{bmatrix},\ \mathsf{H}_{t+1}\equiv\begin{bmatrix}\mathsf{F}_{2,t+1}(\mathcal{A}_{2,t+1})-\mathsf{F}_{1,t+1}(\mathcal{A}_{2,t+1})\\ \\ \mathsf{F}_{1,t+1}(\mathcal{A}_{1,t+1})-\mathsf{F}_{2,t+1}(\mathcal{A}_{1,t+1})\end{bmatrix}$$

Then one period dependence holds if and only if there exists an $(A_{1,t+1}+A_{2,t+1})$ vector of unknowns denoted by D_{t+1} solving:

$$\mathcal{K}_{t+1}=\mathsf{H}_{t+1}\begin{bmatrix}\Omega_{t+1}(\mathcal{A}_{2,t+1},2)\circ\mathsf{K}_{2t}(\mathcal{A}_{2,t+1})\\ \\ \Omega_{t+1}(\mathcal{A}_{1,t+1},1)\circ\mathsf{K}_{1t}(\mathcal{A}_{1,t+1})\end{bmatrix}\equiv\mathsf{H}_{t+1}\mathsf{D}_{t+1}\tag{16}$$

Note that if the weights placed on all the states in $\mathcal{A}_{j,t+1}$ but one are the same across the two paths then the weights placed on the remaining state must be the same as well. Appealing to Hadley (1961, pages 108-109), a solution to (16) for D_{t+1} exists if and only if the rank of H_{t+1} equals the rank of the augmented matrix $\mathsf{H}^*_{t+1}\equiv\begin{bmatrix}\mathcal{K}_{t+1}\vdots\mathsf{H}_{t+1}\end{bmatrix}$ formed by augmenting H_{t+1} with the extra column $\mathcal{K}_{t+1}$.

Denote the rank of H_{t+1} by R_{t+1} and the rank of of $\mathsf{H}^*_{t+1}$by R^*_{t+1}. Clearly $R_{t+1}\leq R^*_{t+1}\leq R_{t+1}+1$ and $R_{t+1}\leq\min\{A_{t+2}-1,A_{1,t+1}+A_{2,t+1}\}$. There are two cases to consider:

1. Suppose $R_{t+1}=A_{1,t+1}+A_{2,t+1}$. If in addition $R_{t+1}=A_{t+2}-1$, implying H_{t+1} is square, we solve for the weights by inverting H_{t+1} and then element-by-element dividing both sides of

(16) by the matching K vectors, yielding:

$$\begin{bmatrix} \Omega_{t+1}(\mathcal{A}_{2,t+1},2) \\ \\ \Omega_{t+1}(\mathcal{A}_{1,t+1},1) \end{bmatrix} = \mathsf{H}_{t+1}^{-1}\mathcal{K}_{t+1}. \Big/ \begin{bmatrix} \mathsf{K}_{2t}(\mathcal{A}_{2,t+1}) \\ \\ \mathsf{K}_{1t}(\mathcal{A}_{1,t+1}) \end{bmatrix} \tag{17}$$

where ./ refers to element-by-element division. If $R_{t+1} > A_{t+2} - 1$, we successively eliminate $A_{1,t+1} + A_{2,t+1} - A_{t+2} + 1$ linearly dependent columns of H_{t+1} to form a square matrix of rank $A_{t+2} - 1$. We now remove the corresponding elements in D_{t+1} in (16) so that the reduced $A_{t+2} - 1$ dimensional vector conforms with the square matrix, by deleting the elements that would have been multiplied by the columns removed from H_{t+1}, effectively giving zero weight to the second action for the removed elements. Finally an analogous equation to (17) is solved for the weights characterizing finite dependence.[16]

2. Alternatively $R_{t+1} < A_{1,t+1} + A_{2,t+1}$. First we successively eliminate $A_{1,t+1} + A_{2,t+1} - R_{t+1}$ linearly dependent columns of H_{t+1} to form an $(A_{t+2} - 1) \times R_{t+1}$ matrix denoted by $\underline{\overline{\mathsf{H}}}_{t+1}$. This operation corresponds to reducing the vector length of D_{t+1} from $A_{1,t+1} + A_{2,t+1}$ to R_{t+1} by effectively setting $A_{1,t+1} + A_{2,t+1} - R_{t+1}$ weights to zero. Denote the $R_{t+1} \times 1$ vector of weights not eliminated by $\underline{\mathsf{D}}_{t+1}$. We now eliminate $A_{t+2} - R_{t+1} - 1$ rows of $\underline{\overline{\mathsf{H}}}_{t+1}$ to form an R_{t+1} dimensional square matrix with rank R_{t+1} denoted by $\underline{\mathsf{H}}_{t+1}$. Strictly for notational purposes, so without loss of generality, we reorder the equations defining (16) so that the linearly independent equations are the bottom ones. This allows us to partition $\underline{\overline{\mathsf{H}}}'_{t+1} \equiv \left[\overline{\mathsf{H}}'_{t+1} \vdots \underline{\mathsf{H}}'_{t+1}\right]$ and $\mathcal{K}'_{t+1} \equiv \left[\overline{\mathcal{K}}'_{t+1} \vdots \underline{\mathcal{K}}'_{t+1}\right]$, where $\overline{\mathsf{H}}_{t+1}$ is $(A_{t+2} - 1 - R_{t+1}) \times R_{t+1}$, while $\overline{\mathcal{K}}'_{t+1}$ is $(A_{t+2} - 1 - R_{t+1}) \times 1$ and $\underline{\mathcal{K}}_{t+1}$ is $R_{t+1} \times 1$. Inverting $\underline{\mathsf{H}}_{t+1}$ we obtain $\underline{\mathsf{D}}_{t+1} = \underline{\mathsf{H}}_{t+1}^{-1}\underline{\mathcal{K}}_{t+1}$. Thus a solution to (16) attains in this knife edged case if and only if $\underline{\mathsf{D}}_{t+1}$ solves $A_{t+2} - R_{t+1} - 1$ additional equations $\overline{\mathcal{K}}_{t+1} = \overline{\mathsf{H}}_{t+1}\underline{\mathsf{H}}_{t+1}^{-1}\underline{\mathcal{K}}_{t+1}$.

[16] The set of weights generated by this procedure depends on which linearly dependent columns are removed. Therefore the weight vectors satisfying finite dependence are not unique.

To illustrate the algorithm in the renewal and terminal state models mentioned above, let $\mathcal{X} \equiv \{1, 2, \ldots, X\}$, and suppose the first choice denotes the terminal or renewal choice which returns the state variable x to the value one, while the second increases x by one unit for all $x < X$ and returns X when $x = X$.[17] Because the transitions are deterministic $A_{1,t+1} = A_{2,t+1} = 1$, with $\mathcal{A}_{1,t+1} = \{1\}$ and $\mathcal{A}_{2,t+1} = \{x_t + 1\}$. Also $A_{t+2} = 3$, with $\mathcal{A}_{t+2} = \{1, 2, x_t + 2\}$. It now follows that in this example:

$$\mathsf{F}_{1,t+1}(\mathcal{A}_{1,t+1}) = \mathsf{F}_{1,t+1}(\mathcal{A}_{2,t+1}) = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \mathsf{F}_{2,t+1}(\mathcal{A}_{1,t+1}) = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad \mathsf{F}_{2,t+1}(\mathcal{A}_{2,t+1}) = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

$$\mathsf{H}_{t+1} = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} \quad \text{or} \quad \mathsf{H}_{t+1}^{-1} = \begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix}$$

Substituting these expressions into (17), and noting that $\Omega_{t+1}(\mathcal{A}_{j,t+1}, j) = \omega_{t+1}(x, j)$ because $\mathsf{K}_{1t}(\mathcal{A}_{1,t+1}) = \mathsf{K}_{2t}(\mathcal{A}_{1,t+1}) = 1$, demonstrates that zero weight is placed on the non-renewal/non-terminal action to achieve one-period dependence:

$$\begin{bmatrix} \omega_{t+1}(x,2) \\ \omega_{t+1}(x,1) \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} . \Big/ \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The limitations of the guess and verify approach become evident when such a widely used class of models in empirical analysis is revealed to have such a simple structure. The class of models exhibiting even one-period finite dependence is much larger than terminal and renewal models, and the method developed here provides a systematic way of discovering them.

3.3 Extension to ρ-period dependence with J choices

We now extend our framework to analyzing the existence of finite dependence for $\rho > 1$ and $J \geq 2$. Given specified decision weights between $t + 1$ and $t + \rho - 1$, two initial choices i and j in Equation (11) relabeled as 1 and 2 for convenience, and an initial state x_t, we provide a new set of necessary and sufficient conditions for whether $\kappa_{t+\rho}(x_{t+\rho+1}|x_t, 1) = \kappa_{t+\rho}(x_{t+\rho+1}|x_t, 2)$.

[17] More formally, $f_{1,t+1}(1|x_t) = 1$, for all t and x_t, while for all t, $f_{2,t+1}(x_t+1|x_t) = 1$ if $x_t < X$ and $f_{2,t+1}(X|X) = 1$.

As we have shown, checking for one-period dependence reduces to solving a linear system of equations. However the equations for determining finite dependence when $\rho > 1$ are highly nonlinear. For example suppose $J = 2$, and we ask whether $\rho = 2$ for some given x_t. Writing $\omega_\tau(x_\tau, j) \equiv \omega_{2\tau}(x_\tau, j)$ and telescoping (5) two periods forward proves that $\kappa_{t+2}(x_{t+3}|x_t, 2) = \kappa_{t+2}(x_{t+3}|x_t, 1)$ if and only if:

$$\begin{aligned}
&\sum_{x_{t+2}=1}^{X}\sum_{x_{t+1}=1}^{X} f_{1,t+2}(x_{t+3}|x_{t+2}) f_{1,t+1}(x_{t+2}|x_{t+1}) \left[f_{1t}(x_{t+1}|x_t) - f_{2t}(x_{t+1}|x_t)\right] \qquad (18)\\
= &\sum_{x_{t+2}=1}^{X}\sum_{x_{t+1}=1}^{X} \left[f_{2,t+2}(x_{t+3}|x_{t+2}) - f_{1,t+2}(x_{t+3}|x_{t+2})\right]\left[f_{2,t+1}(x_{t+2}|x_{t+1}) - f_{1,t+1}(x_{t+2}|x_{t+1})\right]\\
&\times\left[\omega_{t+2}(x_{t+2}, 2)\,\omega_{t+1}(x_{t+1}, 2)\, f_{2t}(x_{t+1}|x_t) - \omega_{t+2}(x_{t+2}, 1)\,\omega_{t+1}(x_{t+1}, 1)\, f_{1t}(x_{t+1}|x_t)\right]\\
&+\sum_{x_{t+2}=1}^{X}\sum_{x_{t+1}=1}^{X} \left[f_{2,t+2}(x_{t+3}|x_{t+2}) - f_{1,t+2}(x_{t+3}|x_{t+2})\right] f_{1,t+1}(x_{t+2}|x_{t+1})\omega_{t+2}(x_{t+1}, 1)\, f_{1t}(x_{t+1}|x_t)\\
&\times\left[\omega_{t+2}(x_{t+1}, 2)\, f_{2t}(x_{t+1}|x_t) - \omega_{t+2}(x_{t+1}, 1)\, f_{1t}(x_{t+1}|x_t)\right]\\
&+\sum_{x_{t+2}=1}^{X}\sum_{x_{t+1}=1}^{X} f_{1,t+2}(x_{t+3}|x_{t+2})\left[f_{2,t+1}(x_{t+2}|x_{t+1}) - f_{1,t+1}(x_{t+2}|x_{t+1})\right]\\
&\times\left[\omega_{t+1}(x_{t+1}, 2)\, f_{2t}(x_{t+1}|x_t) - \omega_{t+1}(x_{t+1}, 1)\, f_{1t}(x_{t+1}|x_t)\right]
\end{aligned}$$

Therefore two-period dependence holds if and only if there exist weights solving (18) for all $x_{t+3} \in \mathcal{X}$. Since products of weights appear in (18), nonlinear solution techniques are required to solve this problem. More generally the equations for ρ-period dependence involve ρ-tuple products of weights.

We exploit the special structure of this nonlinear problem by dividing it into two parts, each having a finite number of operations. The second part is a linear inversion problem that applies to the period $t + \rho$, essentially the same as the case described above when $\rho = 1$ and $J = 2$. The first part delineates the subsets of nodes in $\mathcal{X}$ that can be reached by period $t+\rho$ with nonzero weight by a path from each of the two initial choices being considered. This part also involves a finite number of steps. Having established existence, we can obtain weights satisfying (11) as a by-product.

Analogous to the one-period finite dependence case, for any $\tau \in \{t + 1, \ldots, t + \rho - 1\}$ we say $x_\tau \in \{1, \ldots, X\}$ is attainable by a sequence of decision weights from initial choice $j \in \{1, 2\}$ if the weight on x_τ is nonzero. Let $A_{j\tau} \in \{1, \ldots, X\}$ denote the number of attainable states, and $\mathcal{A}_{j\tau} \subseteq \mathcal{X}$ the set of attainable states for the sequence beginning with choice j. Define $\mathsf{K}_{\tau-1}(\mathcal{A}_{j\tau})$ as an $A_{j\tau}$ vector containing the weights for transitioning to each of the $A_{j\tau}$ attainable states given the choice sequence beginning with j and state x_t. Denote $\Omega_{k\tau}(\mathcal{A}_{j\tau}, j)$ as a vector giving the weight placed on choice $k \in \{2, \ldots, J\}$ for each of the $A_{j\tau}$ possible states at τ. Similarly let $A_{\tau+1} \in \{1, \ldots, X\}$ denote the number of states that are attainable by at least one of the sequences beginning either with choice 1 or 2, and denote by $\mathcal{A}_{\tau+1} \subseteq \mathcal{X}$ the corresponding set. Given an initial state and choice, we denote by $\mathsf{F}_{k\tau}(\mathcal{A}_{j\tau})$ the first $A_{\tau+1} - 1$ columns of the $A_{j\tau} \times A_{\tau+1}$ transition matrix from $\mathcal{A}_{j\tau}$ to $\mathcal{A}_{\tau+1}$ when k is chosen at period τ. The matrix comprises elements $f_{k\tau}(x'|x)$ for each $x \in \mathcal{A}_{j\tau}$ and $x' \in \mathcal{A}_{\tau+1}$. Finally define the $(A_{\tau+1} - 1) \times (J-1)[A_{1\tau} + A_{2\tau}]$ matrix H_τ, and the $(J-1)[A_{1\tau} + A_{2\tau}]$ vector D_τ, respectively by:

$$\mathsf{H}_\tau \equiv \begin{bmatrix} \mathsf{F}_{2\tau}(\mathcal{A}_{2\tau}) - \mathsf{F}_{1\tau}(\mathcal{A}_{2\tau}) \\ \vdots \\ \mathsf{F}_{J\tau}(\mathcal{A}_{2\tau}) - \mathsf{F}_{1\tau}(\mathcal{A}_{2\tau}) \\ \mathsf{F}_{1\tau}(\mathcal{A}_{1\tau}) - \mathsf{F}_{2\tau}(\mathcal{A}_{1\tau}) \\ \vdots \\ \mathsf{F}_{1\tau}(\mathcal{A}_{1\tau}) - \mathsf{F}_{J\tau}(\mathcal{A}_{1\tau}) \end{bmatrix}', \ \mathsf{D}_\tau \equiv \begin{bmatrix} \Omega_{2\tau}(\mathcal{A}_{2\tau}, 2) \circ \mathsf{K}_{\tau-1}(\mathcal{A}_{2\tau}) \\ \vdots \\ \Omega_{J\tau}(\mathcal{A}_{2\tau}, 2) \circ \mathsf{K}_{\tau-1}(\mathcal{A}_{2\tau}) \\ \Omega_{2\tau}(\mathcal{A}_{1\tau}, 1) \circ \mathsf{K}_{\tau-1}(\mathcal{A}_{1\tau}) \\ \vdots \\ \Omega_{J\tau}(\mathcal{A}_{1\tau}, 1) \circ \mathsf{K}_{\tau-1}(\mathcal{A}_{1\tau}) \end{bmatrix}. \tag{19}$$

The $A_{\tau+1}$ system of equations to be solved can now be expressed as:

$$\mathsf{H}_\tau \mathsf{D}_\tau = \mathsf{F}_{1\tau}(\mathcal{A}_{1\tau})\mathsf{K}_{\tau-1}(\mathcal{A}_{1\tau}) - \mathsf{F}_{1\tau}(\mathcal{A}_{2\tau})\mathsf{K}_{\tau-1}(\mathcal{A}_{2\tau}) \equiv \mathcal{K}_{t+1} \tag{20}$$

Note that one of the equations is redundant because if all other states have the same weight assigned to them across the two paths then the last one must be lined up as well, implying that if the rank of H_τ is $A_{\tau+1} - 1$ then finite dependence holds in ρ periods. More generally, again appealing to Hadley (1961, pages 108-109), we obtain the following necessary and sufficient conditions for the existence of a solution to this linear system.

Theorem 2 *Define the* $(A_{\tau+1}-1)\times\{(J-1)[A_{1\tau}+A_{2\tau}]+1\}$ *matrix* $\mathsf{H}_{\tau}^{*}\equiv\left[\mathsf{H}_{\tau}\vdots\mathcal{K}_{t+1}\right]$*, obtained by adding an extra column* $\mathcal{K}_{t+1}$ *to* H_{τ}*. Finite dependence from* x_t *with respect to choices* i *and* j *is achieved in* $\rho=\tau-t$ *periods if and only if there exist weights from* $t+1$ *to* $\tau-1$ *such that the rank of* H_{τ} *equals the rank of* H_{τ}^{*}*.*

There are an infinite number of weighting schemes, each of which might conceivably establish finite dependence. This fact explains why researchers have opted for guess and verify methods when designing models that exhibit this computationally convenient property. Our next theorem, however, proved by construction in the Appendix, shows that an exhaustive search for a set of weights that establish finite dependence can be achieved in a finite number of steps. The key to the proof is that although the definition of H_{τ} does indeed depend on the weights, many sets of weights produce the same $\mathcal{A}_{1\tau}$ and $\mathcal{A}_{2\tau}$ (and hence the same $\mathcal{A}_{\tau+1}$). Since the inversion of H_{τ} hinges on the attainable states, and the sets of all possible attainable states is finite, a finite number of operations is needed to establish whether a finite dependence path exists.

Theorem 3 *For each* $\tau\in\{t+1,\ldots,\rho\}$ *the rank of* H_{τ} *and* H_{τ}^{*} *can be determined in a finite number of operations.*

Theorem 3 applies to any dynamic discrete choice problem described in Section 2. However the number of calculations required to determine ρ-period dependence is specific to the number of choices, J, in periods between $t+1$ and $t+\rho$, the number of states in each of those periods, and the transition matrices. As ρ increases, so too will the sets of possible attainable states, increasing computational complexity in finding the finite dependence path. Increasing the number of choices, J, also will increase the sets of possible attainable states. At the same time, increasing J gives more control to line up the states. When examining finite dependence for a pair of initial choices, the minimum ρ must be weakly decreasing as more choices are available as one could always set the weight on these additional choices to zero. Finally, the complexity of the state space does not necessarily require more calculations to determine finite dependence for two reasons. First, it is only the states that can be reached in ρ periods from the current state that are relevant for determining finite dependence. Second, as the sets of attainable states increase, the researcher also has more options for finding paths that exhibit finite dependence.

3.4 Finite dependence in games

The methods developed above are directly applicable to dynamic games off short panels, that is, after modifying the notation with the (n) superscripts as appropriate. Nevertheless, establishing finite dependence in games is more onerous. Finite dependence in a game is player specific; in principle finite dependence might hold for some players but not for others. Furthermore, the transitions of the state variables depend on the decisions of all the players, not just player n. Thus, finite dependence in games is ultimately a property that derives not just from the game primitives, but is defined with respect to an equilibrium. For this reason, games of incomplete information generally do not exhibit one period finite dependence. If two alternative choices of n at time t affect the equilibrium choices the other players make in the next period at $t+1$ (or later), it is generally not feasible to line up the states across both paths emanating from the respective choices by the beginning of period $t+2$.

The existence of finite dependence in games for a given player n can be established if two conditions are met by the model and the equilibrium played out in the data. First, by taking a sequence of weighted actions, player n can induce, say after ρ periods, the other players to match up the distributions of $x_{t+\rho+1}^{(\sim n)}$, conditional on $x_t^{(\sim n)}$ by following their equilibrium strategies, meaing the distribution of $x_{t+\rho+1}^{(\sim n)}$ does not depend on whether the sequence started with the choice j or k. Whether this condition is satisfied or not depends on the reduced form transitions $f_t^{(\sim n)}\left(x_{t+1}^{(\sim n)}|x_t\right)$. Second, given the distribution of states for player n at $t+\rho$ from the two sequences, one period finite dependence applies to $x_{t+\rho+1}^{(n)}$, meaning that the player is able to line up his own state after executing the weighted sequences across the two paths to line up the states of the other players. This condition is determined by primitives alone, namely matrices formed from the $f_{jt}^{(n)}\left(x_{t+1}^{(n)}\middle|x_t^{(n)}\right)$ transitions.

3.5 Establishing finite dependence in games

From (11), finite dependence at τ for this class of games requires:

$$\begin{aligned}&\sum_{x_\tau=1}^{X}\sum_{k=1}^{J} f_\tau^{(\sim n)}\left(x_{\tau+1}^{(\sim n)}|x_\tau\right) f_{k\tau}^{(n)}\left(x_{\tau+1}^{(n)}\middle|x_\tau^{(n)}\right)\omega_{k\tau}^{(n)}(x_\tau,j)\,\kappa_{\tau-1}^{(n)}(x_\tau|x_t,j)\\ =&\sum_{x_\tau=1}^{X}\sum_{k=1}^{J} f_\tau^{(\sim n)}\left(x_{\tau+1}^{(\sim n)}|x_\tau\right) f_{k\tau}^{(n)}\left(x_{\tau+1}^{(n)}\middle|x_\tau^{(n)}\right)\omega_{k\tau}^{(n)}(x_\tau,i)\,\kappa_{\tau-1}^{(n)}(x_\tau|x_t,i)\end{aligned} \tag{21}$$

We provide a set of sufficient conditions for (21) to hold that are relatively straightforward to check. They are based on the intuition that from periods $t+1$ through $\tau-1$ player n takes actions that indirectly induce the other other players to align $x_{\tau+1}^{(\sim n)}$ through their equilibrium choices, and that at date τ player n takes an action that aligns $x_{\tau+1}^{(n)}$.

One necessary condition for τ dependence can be derived by summing (21) over the $x_{\tau+1}^{(n)}$ outcomes. Noting that:

$$\begin{aligned}&\sum_{x_{\tau+1}^{(n)}=1}^{X^{(n)}}\sum_{x_\tau=1}^{X} f_\tau^{(\sim n)}\left(x_{\tau+1}^{(\sim n)}|x_\tau\right)\left[\sum_{k=1}^{J}\omega_{k\tau}^{(n)}(x_\tau,j)\, f_{k\tau}^{(n)}\left(x_{\tau+1}^{(n)}\middle|x_\tau^{(n)}\right)\right]\kappa_{\tau-1}^{(n)}(x_\tau|x_t,j)\\ =&\sum_{x_\tau=1}^{X} f_\tau^{(\sim n)}\left(x_{\tau+1}^{(\sim n)}|x_\tau\right)\left[\sum_{k=1}^{J}\omega_{k\tau}^{(n)}(x_\tau,j)\right]\left[\sum_{x^{(n)}=1}^{X^{(n)}} f_{k\tau}^{(n)}\left(x_{\tau+1}^{(n)}\middle|x_\tau^{(n)}\right)\right]\kappa_{\tau-1}^{(n)}(x_\tau|x_t,j)\\ =&\sum_{x_\tau=1}^{X} f_\tau^{(\sim n)}\left(x_{\tau+1}^{(\sim n)}|x_\tau\right)\kappa_{\tau-1}^{(n)}(x_\tau|x_t,j)\end{aligned} \tag{22}$$

we simplify the sum (21) over $x_{\tau+1}^{(n)}$ using (22) to obtain:

$$\sum_{x_\tau=1}^{X} f_\tau^{(\sim n)}\left(x_{\tau+1}^{(\sim n)}|x_\tau\right)\left[\kappa_{\tau-1}^{(n)}(x_\tau|x_t,j)-\kappa_{\tau-1}^{(n)}(x_\tau|x_t,i)\right]=0 \tag{23}$$

This proves that in our framework whether (23) holds or not depends on the weights assigned to n in periods $t+1$ though $\tau-1$, but not on the weights chosen in period τ. Furthermore, since the state transitions $f_{k\tau}^{(n)}\left(x_{\tau+1}^{(n)}\left|x_{\tau}^{(n)}\right.\right)$ for $x_{\tau+1}^{(n)}$ do not depend on $x_{\tau}^{(\sim n)}$, whether the weights at τ attain finite dependence next period is not affected by how they vary with $x_{\tau}^{(\sim n)}$. Accordingly, we now express $\omega_{k\tau}^{(n)}(x_\tau, j)$ as a function of $x_{\tau}^{(n)}$ only, writing $\omega_{k\tau}^{(n)}\left(x_{\tau}^{(n)}, j\right)$.

A second necessary condition for τ period dependence is that, in conjunction with the weights in periods preceding τ, the weights ascribed for period τ must line up the unconditional weight distribution of $x_{\tau+1}^{(n)}$ for the two initial choices. Noting that:

$$
\begin{aligned}
&\sum_{x_{\tau+1}^{(\sim n)}}^{X^{(\sim n)}}\sum_{x_\tau=1}^{X}\sum_{k=1}^{J} f_{k\tau}^{(n)}\left(x_{\tau+1}^{(n)}\left|x_{\tau}^{(n)}\right.\right)\left[f_{\tau}^{(\sim n)}\left(x_{\tau+1}^{(\sim n)}\left|x_\tau\right.\right)\omega_{k\tau}^{(n)}\left(x_{\tau}^{(n)}, j\right)\kappa_{\tau-1}^{(n)}(x_\tau|x_t, j)\right] \\
= \;& \sum_{x_\tau=1}^{X}\sum_{k=1}^{J} f_{k\tau}^{(n)}\left(x_{\tau+1}^{(n)}\left|x_{\tau}^{(n)}\right.\right)\omega_{k\tau}^{(n)}\left(x_{\tau}^{(n)}, j\right)\kappa_{\tau-1}^{(n)}(x_\tau|x_t, j)\sum_{x_{\tau+1}^{(\sim n)}}^{X^{(\sim n)}} f_{\tau}^{(\sim n)}\left(x_{\tau+1}^{(\sim n)}\left|x_\tau\right.\right) \\
= \;& \sum_{x_\tau=1}^{X}\sum_{k=1}^{J} f_{k\tau}^{(n)}\left(x_{\tau+1}^{(n)}\left|x_{\tau}^{(n)}\right.\right)\omega_{k\tau}^{(n)}\left(x_{\tau}^{(n)}, j\right)\kappa_{\tau-1}^{(n)}(x_\tau|x_t, j) \qquad (24)
\end{aligned}
$$

we simplify the sum (21) over $x_{\tau+1}^{(\sim n)}$ using (24) to obtain:

$$
0 = \sum_{x_\tau=1}^{X}\sum_{k=1}^{J} f_{k\tau}^{(n)}\left(x_{\tau+1}^{(n)}\left|x_{\tau}^{(n)}\right.\right)\left[\omega_{k\tau}^{(n)}\left(x_{\tau}^{(n)}, j\right)\kappa_{\tau-1}^{(n)}(x_\tau|x_t, j) - \omega_{k\tau}^{(n)}\left(x_{\tau}^{(n)}, i\right)\kappa_{\tau-1}^{(n)}(x_\tau|x_t, i)\right] \qquad (25)
$$

for all $x_{\tau+1}^{(n)} \in \mathcal{X}^{(n)}$. Equation (25) is remarkably similar to an expression obtained for single agent optimization problems found by substituting (5) into (11). Immediately before finite dependence is obtained, n must align the weights on his own state variables for next period, so that the two unconditional weight distributions of $x_{\tau+1}^{(n)}$ for initial choices i and j match.

To derive a rank condition under which (23) holds, it is notationally convenient to focus on the first two choices as before. Suppose (23) holds at $\tau+1$. Then there must be decision weights at $\tau-1$ with the following property: the states that result in τ lead the other players to make (equilibrium) decisions at τ so that each of their own states have the same weight across the two paths at $\tau+1$. Formally, let $\mathcal{A}_{j\tau-1}^{(n)} \subseteq \mathcal{X}$ denote the set of attainable states at $\tau-1$ for the weight sequence beginning with n choosing $j \in \{1, 2\}$. Let $\mathcal{A}_{\tau}^{(n)} \subseteq \mathcal{X}$ denote the set of attainable states at $\tau-1$ for the weight sequence beginning with n either choosing 1 or 2. Let $\mathcal{A}_{\tau+1}^{(\sim n)} \subseteq \mathcal{X}^{(\sim n)}$ denote the states of the other players at $\tau+1$ given the sequence, when n begins by choosing either 1 or 2. Let $A_{\tau+1}^{(\sim n)}$ denote the number of elements in $\mathcal{A}_{\tau+1}^{(\sim n)}$. Let $\mathsf{F}_{k\tau-1}^{(n)}\left(\mathcal{A}_{j\tau-1}^{(n)}\right)$ denote the transition matrix from $\mathcal{A}_{j\tau-1}^{(n)}$ to $\mathcal{A}_{\tau}^{(n)}$ given choice k at time $\tau-1$ when competitors play their equilibrium strategies. Let $\mathsf{P}_{\tau}^{(\sim n)}\left(\mathcal{A}_{\tau}^{(n)}\right)$ denote the transpose of the first $A_{\tau+1}^{(\sim n)}-1$ columns of the transition matrix from $\mathcal{A}_{\tau}^{(n)}$ to the set of competitor states $\mathcal{A}_{\tau+1}^{(\sim n)}$. Finally define $\mathsf{H}_{\tau}^{(\sim n)}$ as:

$$\mathsf{H}_{\tau}^{(\sim n)} \equiv \mathsf{P}_{\tau}^{(\sim n)}\left(\mathcal{A}_{\tau}^{(n)}\right)\begin{bmatrix} \mathsf{F}_{2,\tau-1}^{(n)}\left(\mathcal{A}_{2,\tau-1}^{(n)}\right)-\mathsf{F}_{1\tau-1}^{(n)}\left(\mathcal{A}_{2,\tau-1}^{(n)}\right) \\ \vdots \\ \mathsf{F}_{J,\tau-1}^{(n)}\left(\mathcal{A}_{2,\tau-1}^{(n)}\right)-\mathsf{F}_{1\tau-1}^{(n)}\left(\mathcal{A}_{2,\tau-1}^{(n)}\right) \\ \mathsf{F}_{1,\tau-1}^{(n)}\left(\mathcal{A}_{1,\tau-1}^{(n)}\right)-\mathsf{F}_{2\tau-1}^{(n)}\left(\mathcal{A}_{1,\tau-1}^{(n)}\right) \\ \vdots \\ \mathsf{F}_{1,\tau-1}^{(n)}\left(\mathcal{A}_{1,\tau-1}^{(n)}\right)-\mathsf{F}_{J\tau-1}^{(n)}\left(\mathcal{A}_{1,\tau-1}^{(n)}\right) \end{bmatrix}' \tag{26}$$

Finite dependence requires weighting rules at $\tau-1$ so that when the other players take equilibrium actions at τ on the two paths the states of the other players are lined up at $\tau+1$. The effects of these equilibrium actions on the state operate through $\mathsf{P}_{\tau}^{(\sim n)}\left(\mathcal{A}_{\tau}^{(n)}\right)$ in (26). Equation (26) parallels the expression for H_{τ} given in (19) for the games setting where the states to be matched at $\tau+1$ are the states of the other players rather then the state of the decision-maker. Following the same logic as Theorem 2 yields sufficient conditions for finite dependence in games.[18] The following theorem combines the two conditions–one lining up the states of the other players and the other lining up the player's own states.

Theorem 4 *If the rank of* $\mathsf{H}_{\tau}^{(\sim n)}$ *is* $A_{\tau+1}^{(\sim n)}-1$, *and there exists weights at* τ *such that* (25) *holds for all* $x_{\tau+1}^{(n)}$, *then* $\rho=\tau-t$ *period dependence is attained for initial choices 1 and 2.*

[18]The knife edged case omitted here can also be derived in a similar manner to obtain necessary and sufficient conditions similar to those given in Theorem 2.

4 Applications

This section provides two illustrations, new to the literature, that apply our finite dependence representation. The first is a job search model. Establishing finite dependence in a search model would seem difficult given that there is no guarantee one will receive another job offer in the future if an offer is turned down today and hence lining up, for example, future experience levels would seem difficult. We show that our representation applies directly to this case. The second is a coordination game where we apply the results of Theorem 4 to show that we can achieve two-period finite dependence in a strategic setting.

4.1 A search model

The following simple search model shows why negative weights are useful in establishing finite dependence, and uses the algorithm to exhibit an even less intuitive path to achieve finite dependence. Each period $t \in \{1, \ldots, T\}$ an individual may stay home by setting $d_{1t} = 1$, or apply for temporary employment setting $d_{2t} = 1$. Job applicants are successful with probability λ_t, and the value of the position depends on the experience of the individual denoted by $x \in \{1, \ldots, X\}$. If the individual works his experience increases by one unit, and remains at the current level otherwise. The preference primitives are given by the current utility from staying home, denoted by $u_1(x_t)$, and the utility from working, $u_2(x_t)$. Thus the dynamics of the model arise only from accumulating job experience, while nonstationarities arise from time subscripted offer arrival weights.

Constructing a finite dependence path The guess and verify approach is useful for verifying this model satisfies one-period finite dependence: we simply construct two paths that generate the same probability distribution of x_{t+2} conditional on x_t. Denote $\omega_\tau(x_t, j)$ as the weight placed on action 2 at time τ given initial choice j. Then set:

$$\omega_{t+1}(x_t, 2) = \omega_{t+1}(x_t + 1, 2) = 0,\ \omega_{t+1}(x_t, 1) = \lambda_t/\lambda_{t+1}$$

The distribution of x_{t+2} from following either path is the same: $x_{t+2} = x_t$ with probability $f_{2t}(x_t|x_t) = 1 - \lambda_t$, and $x_{t+2} = x_t + 1$ with probability $f_{2t}(x_t + 1|x_t) = \lambda_t$.

Applying the finite dependence path, the difference in conditional value functions can then be expressed as:

$$\begin{aligned} v_{2t}(x_t) - v_{1t}(x_t) &= \lambda_t \left[u_2(x_t) - u_1(x_t) + \beta u_{1t}(x_t + 1) - \beta u_{2t}(x_t)\right] \\ &\quad + \beta\left[\lambda_t \psi_1\left[p_{t+1}(x_t + 1)\right] + \lambda_t\left(\frac{1}{\lambda_{t+1}} - 1\right)\psi_1\left[p_{t+1}(x_t)\right] - \frac{\lambda_{t+1}}{\lambda_t}\psi_2\left[p_{t+1}(x_t)\right]\right] \end{aligned} \tag{27}$$

Note that if $\lambda_t > \lambda_{t+1}$ then $\omega_{t+1}(x_t, 1) > 1$, demonstrating that negative weights and weights exceeding one can be used to establish finite dependence.

Applying Theorem 2 While Section 4.1 provides a constructive example of forming a finite dependence path, it is also useful to show how the results from Section 3.2 apply. We now use the results from Section 3.2 to derive another finite dependence path.

To do so, we first define relevant terms in Equation (16). $\mathcal{A}_{1,t+1}$ and $\mathcal{A}_{2,t+1}$ are given by $\{x_t\}$ and $\{x_t, x_t + 1\}$. If the individual stays home the state remains unchanged, and if the individual applies for temporary employment he may be employed, or not. Thus $\mathsf{K}_{1t}(\mathcal{A}_{1,t+1})$ and $\mathsf{K}_{2t}(\mathcal{A}_{2,t+1})$ are $[1]$ and $\begin{bmatrix} 1-\lambda & \lambda \end{bmatrix}'$. The relevant transition matrices are given by:

$$\mathsf{F}_{1,t+1}(\mathcal{A}_{1,t+1}) = \begin{bmatrix} 1 & 0 \end{bmatrix},\ \mathsf{F}_{1,t+1}(\mathcal{A}_{2,t+1}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathsf{F}_{2,t+1}(\mathcal{A}_{1,t+1}) = \begin{bmatrix} 1-\lambda_{t+1} & \lambda_{t+1} \end{bmatrix},\ \mathsf{F}_{2,t+1}(\mathcal{A}_{2,t+1}) = \begin{bmatrix} 1-\lambda_{t+1} & \lambda_{t+1} \\ 0 & 1-\lambda_{t+1} \end{bmatrix}$$

The last column, giving the transitions to state $x_t + 2$, is omitted because if the probabilities are aligned in all but one attainable state, then the remaining probability must match up as well.

The system of equations in (16) has two equations (one for the probability of state x_t; another for the probability of state x_{t+1}), plus three choice variables. The three choice variables are the weights on the probability of choosing work conditional on either (i) work in the first period but no job ($x_{t+1} = x_t$), (ii) work in the first period and obtaining a job ($x_{t+1} = x_t + 1$), and (iii) not working in the first period ($x_{t+1} = x_t$). We then have the following expression for the first term on the left-hand-side of (16):

$$\begin{bmatrix} \mathsf{F}_{2t+1}(\mathcal{A}_{2,t+1}) - \mathsf{F}_{1t+1}(\mathcal{A}_{2,t+1}) \\ \mathsf{F}_{1t+1}(\mathcal{A}_{1,t+1}) - \mathsf{F}_{2t+1}(\mathcal{A}_{1,t+1}) \end{bmatrix}' = \begin{bmatrix} -\lambda_{t+1} & 0 & \lambda_{t+1} \\ \lambda_{t+1} & -\lambda_{t+1} & -\lambda_{t+1} \end{bmatrix} \tag{28}$$

To reduce the system to two equations and two unknowns, we set the weight on looking for a job to zero conditional on being in state x_t at $t + 1$ and having chosen not to look for work at t. The last column of (28) can then be eliminated. Noting that:

$$\begin{bmatrix} -\lambda_{t+1} & 0 \\ \lambda_{t+1} & -\lambda_{t+1} \end{bmatrix}^{-1} = \begin{bmatrix} -1/\lambda_{t+1} & 0 \\ -1/\lambda_{t+1} & -1/\lambda_{t+1} \end{bmatrix}$$

the solution to the system, given $\omega_{t+1}(x_t, 1) = 0$, is then:

$$\begin{bmatrix} \omega_{t+1}(x_t, 2) \\ \omega_{t+1}(x_t + 1, 2) \end{bmatrix} = \begin{bmatrix} -1/\lambda_{t+1} & 0 \\ -1/\lambda_{t+1} & -1/\lambda_{t+1} \end{bmatrix} \begin{bmatrix} \lambda_t \\ -\lambda_t \end{bmatrix} . \Big/ \begin{bmatrix} 1 - \lambda_t \\ \lambda_t \end{bmatrix} = \begin{bmatrix} \frac{-\lambda_t}{(1-\lambda_t)\lambda_{t+1}} \\ 0 \end{bmatrix}$$

Finite dependence can then be achieved by setting:

$$\omega_{t+1}(x_t, 1) = \omega_{t+1}(x_t + 1, 2) = 0,\ \omega_{t+1}(x_t, 2) = -\lambda_t \left[(1 - \lambda_t)\lambda_{t+1}\right]^{-1}.$$

Here the path that begins with not looking for work involves not looking for work in period 2 either. By placing negative weight on looking for work conditional on (i) looking for work in period t and (ii) not finding work at period t, we can cancel out the gains from successful search in period t. Hence we arrive at the state x_t along both choice paths.

4.2 A coordination game

Applications of finite dependence in the empirical literature on games are scarce. One exception are models with exit decisions, which have the terminal state property. Although finite dependence is usually not exploited in these models (but see Beauchamp, 2015 and Mazur, 2014), Collard-Wexler (2013), Dunne et al. (2013), and Ryan (2012) all exhibit the finite dependence property that could be used to simplify estimation.

Finite dependence holds for a much broader class of games than those with terminal choices. To illustrate this point, we analyze a simple example of a two player coordination game. Each player $n \in \{1, 2\}$ chooses whether or not to compete in a market at time t, competing by setting $d_{2t}^{(n)} = 1$, not competing by setting $d_{1t}^{(n)} = 1$. In this model $x_t \equiv \left(x_t^{(n)}, x_t^{(\sim n)}\right)$ and $x_t^{(n)} = d_{2,t-1}^{(n)}$, so the state variable transition matrix is deterministic and time invariant. Conditional on the lagged participation of the other player, we assume an individual's choices depend on his own lagged partcipation, implying $p_{2,t+1}^{(n)}(1, d_{2t}^{(\sim n)}) \neq p_{2,t+1}^{(n)}(2, d_{2t}^{(\sim n)})$.[19] This assumption can be tested with data generated from an equilibrium for the game. Summarizing, the dynamics of the game arise purely from the effect of decisions made by both players in the previous period on current payoffs. Nonstationarity arises from the flow payoffs that may depend on time and hence the corresponding choice probabilities.

This model exhibits two period dependence. Let $\omega_{t+2}^{(n)}(x_{t+2}^{(n)}, j)$ denote the weight for action 2 given $x_{t+2}^{(n)} \in \{1, 2\}$ and action $j \in \{1, 2\}$ taken at time t.[20] To satisfy (25) we set the $t+2$ choice weight to be the same across both paths. For example let $\omega_{t+2}^{(n)}(x_{t+2}^{(n)}, j) = 1$ implying $x_{t+3}^{(n)} = 1$ for both paths. All that remains is to find two weighting sequences for n, one for each initial choice at t, such that when the other player makes his equilibrium choice at $t+2$, the distribution of $d_{t+2}^{(\sim n)}$, and

[19] Restating this assumption as it applies to the other player: $p_{2,t+1}^{(\sim n)}(d_{2t}^{(n)}, 1) \neq p_{2,t+1}^{(\sim n)}(d_{2t}^{(n)}, 2)$.

[20] Recall from our general discussion of finite dependence in games that the choice of n at $t+2$ has no effect on the other player's choice at that time because it is not one of his state variables at $t+2$.

hence the distribution of $x_{t+3}^{(\sim n)}$, is the same for both sequences. In this model the rank condition for $\mathsf{H}_{\tau}^{(\sim n)}$ is easy to check because $x_{t+3}^{(\sim n)} \equiv d_{2,t+2}^{(n)}$ only takes two values. Another parsimonious feature of this model is that the initial choice of a player at t is by definition identical to his own contribution to the state variable in $t+1$; in symbols $\omega_{t+1}^{(n)}(x_{t+1}, j) \equiv \omega_{t+1}^{(n)}\left(\left(d_{2t}^{(n)}, d_{2t}^{(\sim n)}\right), d_{2t}^{(n)}\right)$. To eliminate this notational redundancy we now define $\underline{\omega}_{t+1}^{(n)}(x_{t+1}) \equiv \omega_{t+1}^{(n)}(x_{t+1}, j)$. Theorem 5 establishes two period dependence by specifying a $\underline{\omega}_{t+1}^{(n)}(x_{t+1})$, that in conjunction with setting $\omega_{t+2}^{(n)}(x_{t+2}^{(n)}, j) = 1$, achieves finite dependence.

Theorem 5 *The coordination game exhibits two period dependence for all x_t. For $i \in \{1, 2\}$ define:*

$$C_i \equiv p_{2,t+2}^{(\sim n)}(2,1) - p_{2,t+2}^{(\sim n)}(1,1) + p_{2,t+1}^{(\sim n)}(2,i)\left[p_{2,t+2}^{(\sim n)}(2,2) + p_{2,t+2}^{(\sim n)}(1,1) - p_{2,t+2}^{(\sim n)}(2,1) - p_{2,t+2}^{(\sim n)}(1,2)\right]$$

If $C_1 = 0$ then $C_2 \neq 0$. For $C_i \neq 0$ set $\underline{\omega}_{t+1}^{(n)}(x_{t+1}) = 0$ for all $x_{t+1} \neq (2, i)$ and:

$$\underline{\omega}_{t+1}^{(n)}(2,i) = \mathsf{P}_{t+2}^{(\sim n)}\left(\mathcal{A}_{t+2}^{(n)}\right)\begin{bmatrix} \mathsf{F}_{1,t+1}^{(n)}(\mathcal{A}_{1,t+1}^{(n)}) \\ -\mathsf{F}_{1,t+1}^{(n)}(\mathcal{A}_{2,t+1}^{(n)}) \end{bmatrix}' . \Big/ \left[p_{2t}^{(\sim n)}(x_t) C_i\right].$$

5 Conclusion

CCP methods provide a cheap way of estimating dynamic discrete choice models in both single-agent and multi-agent settings. This paper precisely delineates and expands the class of models that exhibit the finite dependence property used in CCP estimators, whereby only a-few-period-ahead conditional choice probabilities are used in estimation. Our approach applies a wide class of problems lacking stationarity, and is free of assumptions about the structure of the model and the beliefs of players regarding events that occur after the (short) panel has ended. For example these methods enable estimation of nonstationary infinite horizon games even when there are no terminal or renewal actions. Finally, our analysis leaves several questions unanswered: What are the computational benefits and costs associated with implementing this algorithm to determine finite dependence in situations where "guess and verify" is infeasible? Since there is no presumption that a unique set of weights exists when finite dependence hold, a point illustrated in the search example, which set of weights should be used in estimation?[21] We defer these topics to future research.

6 Appendix: Proofs

Proof of Theorem 1. With (bounded) negative weights the finite horizon results of Theorem 1 of Arcidiacono and Miller (2011) is easily adapted, since the proof of whether the positivity or negativity of the weights is not used in that proof. ■

Proof of Theorem 3. Denote by $\mathcal{A} \equiv \left\{x_{\mathcal{A}}^{(1)}, \ldots, x_{\mathcal{A}}^{(A)}\right\}$ where $x_{\mathcal{A}}^{(a)} \in \mathcal{X}$ for all $a \in \{1, \ldots, A\}$. Thus $\mathcal{A} \in \mathcal{S}$, the set containing 2^X elements of all subsets of $\mathcal{X}$. Also define the set $\mathcal{A}$ attains at τ by:

$$\mathcal{B} \equiv \left\{x_{\mathcal{B}}^{(b)} \in \mathcal{X} \text{ such that } f_{j\tau}(x_{\mathcal{B}}^{(b)} \,|x) \neq 0 \text{ for some } x \in \mathcal{A} \text{ and some } j = 1, \ldots, J\right\}$$

Thus $\mathcal{B} = \left\{x_{\mathcal{B}}^{(1)}, \ldots, x_{\mathcal{B}}^{(B)}\right\}$ for some $B \leq X$. For each $a \in \{1, \ldots, A\}$ define the $(J-1) \times 1$ weight vector:

$$\omega_\tau\left(x_{\mathcal{A}}^{(a)}\right) = \left(\omega_{1\tau}\left(x_{\mathcal{A}}^{(a)}\right), \ldots, \omega_{J-1,\tau}\left(x_{\mathcal{A}}^{(a)}\right)\right)'$$

where $\left|\omega_{j\tau}\left(x_{\mathcal{A}}^{(a)}\right)\right| < \infty$ and $\omega_{J\tau}\left(x_{\mathcal{A}}^{(a)}\right) \equiv 1 - \sum_{j=1}^{J-1} \omega_{j\tau}\left(x_{\mathcal{A}}^{(a)}\right)$. Let $\mathcal{K}_{\mathcal{A}} \equiv \left(\mathcal{K}_{\mathcal{A}}^{(1)}, \ldots, \mathcal{K}_{\mathcal{A}}^{(A)}\right)'$ denote an $A \times 1$ weight vector over the states in $\mathcal{A}$, that is satisfying $\sum_{x=1}^{A} \mathcal{K}_{\mathcal{A}}^{(a)} = 1$ with $\left|\mathcal{K}_{\mathcal{A}}^{(a)}\right| < \infty$ and $\mathcal{K}_{\mathcal{A}}^{(x)} \neq 0$. We also define:

$$\mathcal{K}_{\mathcal{B}}^{(b)} \equiv \sum_{a=1}^{A}\sum_{j=1}^{J} f_{j\tau}(x_{\mathcal{B}}^{(b)} \left| x_{\mathcal{A}}^{(a)}\right.)\omega_{j\tau}\left(x_{\mathcal{A}}^{(a)}\right)\mathcal{K}_{\mathcal{A}}^{(a)}$$

and note that:

$$\sum_{b=1}^{B}\mathcal{K}_{\mathcal{B}}^{(b)} = \sum_{b=1}^{B}\sum_{a=1}^{A}\sum_{j=1}^{J} f_{j\tau}(x_{\mathcal{B}}^{(b)} \left| x_{\mathcal{A}}^{(a)}\right.)\omega_{j\tau}\left(x_{\mathcal{A}}^{(a)}\right)\mathcal{K}_{\mathcal{A}}^{(a)} = \sum_{a=1}^{A}\sum_{j=1}^{J}\omega_{j\tau}\left(x_{\mathcal{A}}^{(a)}\right)\mathcal{K}_{\mathcal{A}}^{(a)} = \sum_{a=1}^{A}\mathcal{K}_{\mathcal{A}}^{(a)} = 1 \quad (29)$$

[21] Weighting future utility terms differently affects the asymptotic covariance matrix of the estimator, as well as its finite sample properties. Consequently choosing amongst alternative weighting schemes that attain finite dependence is application specific.

Depending on $\mathcal{K}_{\mathcal{A}}$, and also the choice of $\omega_{\tau\mathcal{A}} \equiv \left(\omega_\tau\left(x_{\mathcal{A}}^{(1)}\right), \ldots, \omega_\tau\left(x_{\mathcal{A}}^{(A)}\right)\right)'$, some elements of $\mathcal{K}_{\mathcal{B}} \equiv \left(\mathcal{K}_{\mathcal{B}}^{(1)}, \ldots, \mathcal{K}_{\mathcal{B}}^{(B)}\right)'$ may be zero. We say that $\mathcal{A}$ reaches $\mathcal{A}^* \subseteq \mathcal{A}'$ at τ for the vector weighting $\mathcal{K}_{\mathcal{A}}$ if, for some choice of $\omega_{\tau\mathcal{A}}$, every element in $\mathcal{A}^*$ is attained (has nonzero weight), and every element in the complement of $\mathcal{A}^*$ is not attained (has zero weight).

Theorem 2, and its proof in the text, shows that only a finite number of operations are required to determine whether or not finite dependence dependence can be achieved in one period from two given sets $\mathcal{A}_{1,t+\rho}$ and $\mathcal{A}_{2,t+\rho}$. In particular, it is evident from the construction of H_τ, that the operations do not depend on the $\omega_{\tau,\mathcal{A}_{1,t+\rho}}$ and $\omega_{\tau,\mathcal{A}_{2,t+\rho}}$, the respective weights on elements in $\mathcal{A}_{1,t+\rho}$ and $\mathcal{A}_{2,t+\rho}$. Given $j \in \{1,2\}$, and a sequence of weights defined from $t+1$ to $t+\rho$, a unique sequence of sets is determined: say $\{\mathcal{A}_{j\tau}\}_{\tau=t+2}^{\rho}$. Although there are an uncountable number of paths, since $\mathcal{A}_{j\tau} \in \mathcal{S}$ and $\mathcal{S}$ contains (only) 2^X elements, there are at most $2^{(\rho-1)X}$ sets that any weight sequence can successively reach, from $\mathcal{A}_{j,t+1} \equiv \{x \in X : f_{jt}(x|x_t) > 0\}$ up to and including $\mathcal{A}_{j,t+\rho}$. Therefore the proof is completed by showing that a finite number of operations suffice to determine whether or not a given $\mathcal{A} \subseteq \mathcal{A}'_{j,\tau+1}$can be reached from any $\mathcal{A}_{j\tau} \in \mathcal{S}$, for all possible (nonzero) weights $\mathcal{K}_{\mathcal{A}}$.

To determine whether $\mathcal{A}$ reaches $\mathcal{A}^*$ at τ we extend similar arguments given in the text for checking whether $\rho = 2$ in the special case where $J = 2$. Without loss of generality we focus on the case where $\mathcal{A}^*$ is might be reached because the first A^* elements of $\mathcal{K}_{\mathcal{A}^*}$ are nonzero and the remaining $B^* - A^*$ are zero. (The other cases are covered by a reordering of the states.) Thus $\mathcal{K}_{\mathcal{B}} \equiv \left(\mathcal{K}_{\mathcal{B}}^{(1)}, \ldots, \mathcal{K}_{\mathcal{B}}^{(B)}\right)'$ is a weighting for $\mathcal{A}^*$ if and only if:

$$\mathcal{K}_{\mathcal{B}}^{(b)} = \begin{cases} 1 - \sum_{b=2}^{A^*} \mathcal{K}_{\mathcal{B}}^{(b)} \text{ for } b = 1 \\ \text{any nonzero value for } b \in \{2, \ldots, A^*\} \text{ subject to the constraint } \sum_{b=2}^{A^*} \mathcal{K}_{\mathcal{B}}^{(b)} \neq 1 \\ 0 \text{ for } b \in \{A^*+1, \ldots, B\} \end{cases} \tag{30}$$

The existence of a solution to an unconstrained linear system, comprising $B - 1$ equations in $(J-1)\,A$ unknowns, determines whether $\mathcal{A}$ reaches $\mathcal{A}^*$ at τ or not. The unknown variables in the linear system are the A choice weight vectors $\omega_\tau\left(x_{\mathcal{A}}^{(a)}\right)$, each of dimension $J-1$. The $B-1$ equations correspond to the nonzero weights placed on the states $\left\{x_{\mathcal{B}}^{(2)},\ldots,x_{\mathcal{B}}^{(A^*)}\right\}$ and the zero weighting placed on the last $B-A^*$ states, which belong to $\mathcal{B}$ but not $\mathcal{A}^*$. All choice weights satisfying the equations corresponding to $\left\{x_{\mathcal{B}}^{(2)},\ldots,x_{\mathcal{B}}^{(A^*)}\right\}$ also satisfy the first state in $\mathcal{B}$ by (29)and (30).

Given $\mathcal{K}_{\mathcal{B}}^{(b)}$ satisfying (30) , a solution to this linear system exists if there exists A choice weight vectors $\omega_\tau\left(x_{\mathcal{A}}^{(a)}\right)$ for each $b\in\{2,\ldots,B\}$ solving:

$$\mathcal{K}_{\mathcal{B}}^{(b)}=\sum_{a=1}^{A}f_{J\tau}(x_{\mathcal{B}}^{(b)}\left|x_{\mathcal{A}}^{(a)}\right.)\mathcal{K}_{\mathcal{A}}^{(a)}+\sum_{a=1}^{A}\sum_{j=1}^{J-1}\left[f_{j\tau}(x_{\mathcal{B}}^{(b)}\left|x_{\mathcal{A}}^{(a)}\right.)-f_{J\tau}(x_{\mathcal{B}}^{(b)}\left|x_{\mathcal{A}}^{(a)}\right.)\right]\omega_{j\tau}\left(x_{\mathcal{A}}^{(a)}\right)\mathcal{K}_{\mathcal{A}}^{(a)} \tag{31}$$

Let $F_{j\tau}(\mathcal{A})$ denote the $A\times(B-1)$ transition matrix for $\mathcal{A}$ into all but the first states in $\mathcal{B}$ for choice $j\in\{1,2,\ldots,J-1\}$. Define $[\mathcal{K}_{\mathcal{A}}\circ\omega_\tau(\mathcal{A})]$ as the $A(J-1)\times1$ vector formed from the element-by-element product $\mathcal{K}_{\mathcal{A}}^{(a)}\omega_{j\tau}\left(x_{\mathcal{A}}^{(a)}\right)$. Denote the $(B-1)\times A(J-1)$ concatenated matrix of transitions by:

$$\begin{aligned}F_\tau(\mathcal{A})' &\equiv \left[\begin{array}{c}F_{1\tau}(\mathcal{A})'\cdots F_{J-1,\tau}(\mathcal{A})'\end{array}\right]\\ &= \left[\begin{array}{ccccccc} f_{1\tau}(x_{\mathcal{B}}^{(2)}|x_{\mathcal{A}}^{(1)}) & \cdots & f_{1\tau}(x_{\mathcal{B}}^{(2)}|x_{\mathcal{A}}^{(A)}) & \cdots & f_{J-1,\tau}(x_{\mathcal{B}}^{(2)}|x_{\mathcal{A}}^{(1)}) & \cdots & f_{J-1,\tau}(x_{\mathcal{B}}^{(2)}|x_{\mathcal{A}}^{(A)})\\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots\\ f_{1\tau}(x_{\mathcal{B}}^{(B)}|x_{\mathcal{A}}^{(1)}) & \cdots & f_{1\tau}(x_{\mathcal{B}}^{(B)}|x_{\mathcal{A}}^{(A)}) & \cdots & f_{J-1,\tau}(x_{\mathcal{B}}^{(B)}|x_{\mathcal{A}}^{(1)}) & \cdots & f_{J-1,\tau}(x_{\mathcal{B}}^{(B)}|x_{\mathcal{A}}^{(A)})\end{array}\right]\end{aligned}$$

Defining $\mathcal{K}_{\mathcal{B}}^*$ as a $(B-1)\times1$ vector formed from all but the first element of $\mathcal{K}_{\mathcal{B}}$ satisfying (30) then (31) may be expressed in matrix notation as:

$$\mathcal{K}_{\mathcal{B}}^*=F_{J\tau}(\mathcal{A})'\mathcal{K}_{\mathcal{A}}+\left[F_\tau(\mathcal{A})'-F_{J\tau}(\mathcal{A})'\right][\mathcal{K}_{\mathcal{A}}\circ\omega_\tau(\mathcal{A})] \tag{32}$$

Appealing to Hadley (1961, pages 108-109), for a given $\mathcal{K}_{\mathcal{B}}^*$, a solution to (32) in $[\mathcal{K}_{\mathcal{A}}\circ\omega_\tau^*(\mathcal{A})]$ exists if and only if the rank of $\left[F_\tau(\mathcal{A})'-F_{J\tau}(\mathcal{A})'\right]$ equals the rank of the augmented matrix formed by adding the column $\left[\mathcal{K}_{\mathcal{B}}^*-F_{J\tau}(\mathcal{A})'\mathcal{K}_{\mathcal{A}}\right]$ to $\left[F_\tau(\mathcal{A})'-F_{J\tau}(\mathcal{A})'\right]$. By construction the augmented matrix either has the same rank as, or one plus the rank of $\left[F_\tau(\mathcal{A})'-F_{J\tau}(\mathcal{A})'\right]$. Determining the rank of a finite dimensional matrix requires only a finite number of operations. Since each of the finite number of steps described above involves a finite number of operations, the theorem is proved.

■

Proof of Theorem 4. The proof follows steps similar to that of Theorem 2. Define $\mathsf{K}_{\tau-2}^{(n)}\left(\mathcal{A}_{j\tau-1}^{(n)}\right)$ as an $A_{j,\tau-1}^{(n)}$ vector containing the probabilities of transitioning to each of the $A_{j\tau-1}^{(n)}$ attainable states given the choice sequence beginning with j by player n and state x_t. Denote $\Omega_{k,\tau-1}^{(n)}\left(\mathcal{A}_{j,\tau-1}^{(n)}, j\right)$ as a vector giving the weight placed on choice $k \in \{2, \ldots, J\}$ by player n for each of the $A_{j,\tau-1}^{(n)}$ possible states at $\tau - 1$. Let $\mathsf{D}_{j,\tau-1}^{(n)}\left(\mathcal{A}_{j,\tau-1}^{(n)}\right)$ be a $(J-1)A_{j,\tau-1}^{(n)}$ vector defined by:

$$\mathsf{D}_{j,\tau-1}^{(n)}\left(\mathcal{A}_{j,\tau-1}^{(n)}\right) = \begin{bmatrix} \Omega_{2,\tau-1}^{(n)}\left(\mathcal{A}_{j,\tau-1}^{(n)}, j\right) \circ \mathsf{K}_{\tau-2}^{(n)}\left(\mathcal{A}_{j,\tau-1}^{(n)}\right) \\ \vdots \\ \Omega_{J,\tau-1}^{(n)}\left(\mathcal{A}_{j,\tau-1}^{(n)}, j\right) \circ \mathsf{K}_{\tau-2}^{(n)}\left(\mathcal{A}_{j,\tau-1}^{(n)}\right) \end{bmatrix}$$

where $\circ$ refers to element-by-element multiplication. The matrix representation of the finite dependence condition given in (23) for state $x_{\tau+1}^{(\sim n)}$ is then given by the $A_{\tau+1}^{(\sim n)}$ system of equations:

$$\mathsf{H}_{\tau}^{(\sim n)} \begin{bmatrix} \mathsf{D}_{2,\tau-1}^{(n)}\left(\mathcal{A}_{2,\tau-1}^{(n)}\right) \\ \mathsf{D}_{1,\tau-1}^{(n)}\left(\mathcal{A}_{1,\tau-1}^{(n)}\right) \end{bmatrix} = \mathsf{P}_{\tau}^{(\sim n)}\left(\mathcal{A}_{\tau}^{(n}\right) \begin{bmatrix} \mathsf{F}_{1,\tau-1}^{(n)}\left(\mathcal{A}_{1,\tau-1}^{(n)}\right) \\ -\mathsf{F}_{1,\tau-1}^{(n)}\left(\mathcal{A}_{2,\tau-1}^{(n)}\right) \end{bmatrix}' \begin{bmatrix} \mathsf{K}_{\tau-2}^{(n)}\left(\mathcal{A}_{1,\tau-1}^{(n)}\right) \\ \mathsf{K}_{\tau-2}^{(n)}\left(\mathcal{A}_{2,\tau-1}^{(n)}\right) \end{bmatrix} \tag{33}$$

Note that one of the equations is redundant because if all other competitor states have the same weight assigned to them across the two paths then the last one must be lined up as well. Hence if the rank of $\mathsf{H}_{\tau}^{(\sim n)}$ is $A_{\tau+1}^{(\sim n)} - 1$ there exists decisions weights such that the distribution of $x_{t+\rho}^{(\sim n)}$ is the same for both initial choices. If in addition player n can select weights at $\tau + \rho$ such that the distribution of $x_{t+\rho}^{(n)}$ is the same for both initial choices then finite dependence is attained. ■

Proof of Theorem 5. In this game each player $n \in \{1, 2\}$ controls two states, namely the choices of the previous period "in" or "out", so from (33) two period dependence requires a solution to:

$$\mathsf{H}_{t+2}^{(\sim n)} \begin{bmatrix} \Omega_{2,t+1}^{(n)}(\mathcal{A}_{2,t+1}^{(n)}, 2) \circ \mathsf{K}_{2t}(\mathcal{A}_{2,t+1}^{(n)}, 2) \\ \Omega_{2,t+1}^{(n)}(\mathcal{A}_{1,t+1}^{(n)}, 1) \circ \mathsf{K}_{1t}(\mathcal{A}_{1,t+1}^{(n)}, 1) \end{bmatrix} = \mathsf{P}_{t+2}^{(\sim n)}\left(\mathcal{A}_{t+2}^{(n)}\right) \begin{bmatrix} \mathsf{F}_{1,t+1}^{(n)}(\mathcal{A}_{1,t+1}^{(n)}) \\ -\mathsf{F}_{1,t+1}^{(n)}(\mathcal{A}_{2,t+1}^{(n)}) \end{bmatrix}' \begin{bmatrix} \mathsf{K}_{1t}(\mathcal{A}_{1,t+1}^{(n)}) \\ \mathsf{K}_{2t}(\mathcal{A}_{2,t+1}^{(n)}) \end{bmatrix} \tag{34}$$

where the definition of $\mathsf{H}_\tau^{(\sim 2)}$ given in (26) specializes to:[22]

$$\mathsf{H}_{t+2}^{(\sim n)} \equiv \mathsf{P}_{t+2}^{(\sim n)}\left(\mathcal{A}_{t+2}^{(n)}\right)\begin{bmatrix} \mathsf{F}_{2,t+1}^{(n)}(\mathcal{A}_{2,t+1}^{(n)}) - \mathsf{F}_{1t+1}^{(n)}(\mathcal{A}_{2,t+1}^{(n)}) \\ \mathsf{F}_{1,t+1}^{(n)}(\mathcal{A}_{1,t+1}^{(n)}) - \mathsf{F}_{2t+1}^{(n)}(\mathcal{A}_{1,t+1}^{(n)}) \end{bmatrix}'$$

$$\Omega_{2,t+1}^{(n)}(\mathcal{A}_{2,t+1}^{(n)}, j) = \left[\underline{\omega}_{t+1}^{(n)}(j,2), \underline{\omega}_{t+1}^{(n)}(j,1)\right]$$

$$\mathsf{K}_{2t}(\mathcal{A}_{2,t+1}^{(n)}) = \mathsf{K}_{1t}(\mathcal{A}_{2,t+1}^{(n)}) = \begin{bmatrix} p_{2t}^{(\sim n)}(x_t) & p_{1t}^{(\sim n)}(x_t) \end{bmatrix}'$$

$$\mathsf{P}_{t+2}^{(\sim n)}\left(\mathcal{A}_{t+2}^{(n)}\right) = \begin{bmatrix} p_{2,t+2}^{(\sim n)}(2,2) & p_{2,t+2}^{(\sim n)}(2,1) & p_{2,t+2}^{(\sim n)}(1,2) & p_{2,t+2}^{(\sim n)}(1,1) \end{bmatrix}$$

$$\begin{bmatrix} \mathsf{F}_{1,t+1}^{(n)}(\mathcal{A}_{1,t+1}^{(n)}) \\ -\mathsf{F}_{1,t+1}^{(n)}(\mathcal{A}_{2,t+1}^{(n)}) \end{bmatrix}' = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ p_{2,t+1}^{(\sim n)}(1,2) & p_{2,t+1}^{(\sim n)}(1,1) & -p_{2,t+1}^{(\sim n)}(2,2) & -p_{2,t+1}^{(\sim n)}(1,2) \\ p_{1,t+1}^{(\sim n)}(1,2) & p_{1,t+1}^{(\sim n)}(1,1) & -p_{1,t+1}^{(\sim n)}(2,2) & -p_{1,t+1}^{(\sim n)}(2,1) \end{bmatrix}$$

$$\begin{bmatrix} \mathsf{F}_{2,t+1}^{(n)}\left(\mathcal{A}_{2,t+1}\right) - \mathsf{F}_{1,t+1}^{(n)}\left(\mathcal{A}_{2,t+1}\right) \\ \mathsf{F}_{1,t+1}^{(n)}\left(\mathcal{A}_{1,t+1}\right) - \mathsf{F}_{2,t+1}^{(n)}\left(\mathcal{A}_{1,t+1}\right) \end{bmatrix}' = \begin{bmatrix} p_{2,t+1}^{(\sim n)}(2,2) & p_{2,t+1}^{(\sim n)}(2,1) & -p_{2,t+1}^{(\sim n)}(1,2) & -p_{2,t+1}^{(\sim n)}(1,1) \\ p_{1,t+1}^{(\sim n)}(2,2) & p_{1,t+1}^{(\sim n)}(2,1) & -p_{1,t+1}^{(\sim n)}(1,2) & -p_{1,t+1}^{(\sim n)}(1,1) \\ -p_{2,t+1}^{(\sim n)}(2,2) & -p_{2,t+1}^{(\sim n)}(2,1) & p_{2,t+1}^{(\sim n)}(1,2) & p_{2,t+1}^{(\sim n)}(1,1) \\ -p_{1,t+1}^{(\sim n)}(2,2) & -p_{1,t+1}^{(\sim n)}(2,1) & p_{1,t+1}^{(\sim n)}(1,2) & p_{1,t+1}^{(\sim n)}(1,1) \end{bmatrix} \tag{35}$$

Substituting the expressions above into the left hand side of (34) yields:

$$\begin{bmatrix} p_{2,t+2}^{(\sim n)}(2,2) \\ p_{2,t+2}^{(\sim n)}(2,1) \\ p_{2,t+2}^{(\sim n)}(1,2) \\ p_{2,t+2}^{(\sim n)}(1,1) \end{bmatrix}' \begin{bmatrix} p_{2,t+1}^{(\sim n)}(2,2) & p_{2,t+1}^{(\sim n)}(2,1) & -p_{2,t+1}^{(\sim n)}(1,2) & -p_{2,t+1}^{(\sim n)}(1,1) \\ p_{1,t+1}^{(\sim n)}(2,2) & p_{1,t+1}^{(\sim n)}(2,1) & -p_{1,t+1}^{(\sim n)}(1,2) & -p_{1,t+1}^{(\sim n)}(1,1) \\ -p_{2,t+1}^{(\sim n)}(2,2) & -p_{2,t+1}^{(\sim n)}(2,1) & p_{2,t+1}^{(\sim n)}(1,2) & p_{2,t+1}^{(\sim n)}(1,1) \\ -p_{1,t+1}^{(\sim n)}(2,2) & -p_{1,t+1}^{(\sim n)}(2,1) & p_{1,t+1}^{(\sim n)}(1,2) & p_{1,t+1}^{(\sim n)}(1,1) \end{bmatrix} \begin{bmatrix} \underline{\omega}_{t+1}^{(n)}(2,2)p_{2t}^{(\sim n)}(x_t) \\ \underline{\omega}_{t+1}^{(n)}(2,1)p_{2t}^{(\sim n)}(x_t) \\ \underline{\omega}_{t+1}^{(n)}(1,2)p_{2t}^{(\sim n)}(x_t) \\ \underline{\omega}_{t+1}^{(n)}(1,1)p_{2t}^{(\sim n)}(x_t) \end{bmatrix} \tag{36}$$

[22]Since matching the weight on one state automatically matches the weight on the other, we can eliminate the last row of $\mathsf{P}_{t+2}^{(\sim n)}\left(\mathcal{A}_{t+2}^{(n)}\right)$.

Since $p_{2t}^{(\sim n)}(x_t) > 0$ we can establish two period dependence by equating (36) with the right hand side of (34) and solving for the unknowns. By inspection (36) is 1×1, and (34) reduces to a single equation, with four unknowns that conform to the 1×4 row vector $\mathsf{H}_{t+2}^{(\sim n)}$. We consider two possibilities, labeled $i \in \{1, 2\}$. Both set three of the four unkowns to zero, either the first element in in vector on the far right of (36) or the second. Setting $\underline{\omega}_{t+1}^{(n)}(x_{t+1}) = 0$ for all $x_{t+1} \neq (2, i)$, making use of the fact that $p_{1,t+1}^{(\sim n)}(2, 2) = 1 - p_{2,t+1}^{(\sim n)}(2, 2)$, and appealing to the definition of C_i given in the theorem, simplifies (36) to $C_i p_{2t}^{(\sim n)}(x_t)\omega_{t+1}^{(n)}(2, i)$.

To show that $C_2 \neq 0$ if $C_1 = 0$ and complete the proof of the theorem, note:

$$C_2 - C_1 = \left[p_{2,t+1}^{(\sim n)}(2,2) - p_{2,t+1}^{(\sim n)}(2,1)\right]\left[p_{2,t+2}^{(\sim n)}(2,2) - p_{2,t+2}^{(\sim n)}(2,1) + p_{2,t+2}^{(\sim n)}(1,1) - p_{2,t+2}^{(\sim n)}(1,2)\right] \quad (37)$$

If the second bracketed term is zero, then:

$$\begin{aligned} & p_{2,t+2}^{(\sim n)}(2,1) - p_{2,t+2}^{(\sim n)}(2,2) = p_{2,t+2}^{(\sim n)}(1,1) - p_{2,t+2}^{(\sim n)}(1,2) \\ \Rightarrow \quad & C_1 = C_2 = p_{2,t+2}^{(\sim n)}(1,1) - p_{2,t+2}^{(\sim n)}(1,2) \neq 0 \end{aligned}$$

because by assumption $p_{2,t+2}^{(\sim n)}(1,1) \neq p_{2,t+2}^{(\sim n)}(1,2)$. If the second bracketed term is nonzero, then $C_2 \neq C_1$ by (37) because $p_{2,t+1}^{(\sim n)}(2,1) \neq p_{2,t+1}^{(\sim n)}(2,2)$. Therefore $C_2 \neq 0$ if $C_1 = 0$. Given $C_i \neq 0$, we set $\underline{\omega}_{t+1}^{(n)}(x_{t+1}) = 0$ for all $x_{t+1} \neq (2, i)$ and solve for $\underline{\omega}_{t+1}^{(n)}(2, i)$ by substituting (37) into (34) as indicated in the statement of the theorem. ■

References

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

[2] Aguirregabiria, V. and A. Magesan (2016): "Solution and Estimation of Dynamic Discrete Choice Structural Models Using Euler Equations", working paper.

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

[4] Aguirregabiria, V., and P. Mira (2007): “ Sequential Estimation of Dynamic Discrete Games,” *Econometrica,* 75, 1-54.

[5] Altug, S., and R. Miller (1998): “The Effect of Work Experience on Female Wages and Labour Supply,” *Review of Economic Studies*, 62, 45-85.

[6] Arcidiacono, P., E. Aucejo, A. Maurel, and T. Ransom (2014): “College Attrition and the Dynamics of Information Revelation,” working paper.

[7] Arcidiacono, P., P. Bayer, J. Blevins, and P. Ellickson (forthcoming): “Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition.” *Review of Economic Studies.*

[8] Arcidiacono, P. and P. Ellickson (2011): “Practical Methods for Estimation of Dynamic Discrete Choice Models”, *Annual Review of Economics,* 3, 363-394.

[9] Arcidiacono, P., and R. Miller (2011): “Conditional Choice Probability Estimation of Dynamic Discrete Choice Model with Unobserved Heterogeneity”, *Econometrica*, 79, 1823-1867.

[10] Bajari, P., L. Benkard and J. Levin (2007): “Estimating Dynamic Models of Imperfect Competition”, *Econometrica,* 75, 1331-1371.

[11] Beauchamp, A. (2015): “Regulation, Imperfect Competition, and the U.S. Abortion Market,” *International Economic Review*, 56, 963-996.

[12] Bishop, K. (2012): “A Dynamic Model of Location Choice and Hedonic Valuation”. working paper.

[13] Coate, P. (2013): “Parental Influence on Labor Market Outcomes and Location Decisions of Young Workers,” working paper.

[14] Collard-Wexler, A. (2013): “Demand Fluctuations in the Ready-Mix Concrete Industry”, *Econometrica*, 81, 1003–1037

[15] Declerq K. and F. Verboven (2014): “Enrollment and degree completion in higher education without ex ante admission standards,” working paper.

[16] Dunne, T., S. Klimek, M. Roberts, and D. Xu (2013): “Entry, Exit, and The Determinants of Market Structure”, *Rand Journal of Economics*, 44, 462-487.

[17] Ericson, R., and A. Pakes (1995): “ Markov-Perfect Industry Dynamics: A Framework for Empirical Work,” *Review of Economic Studies*, 62, 53-82.

[18] Gayle, G. and L. Golan (2012): “Estimating a Dynamic Adverse Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968-97,” *Review of Economic Studies*, 79, 227-267.

[19] Gayle G. and R. Miller (2014): ”Life-Cycle Fertility and Human Capital Accumulation”, working paper.

[20] Gayle, W. (2013): “CCP Estimation of Dynamic Discrete/Continuous Choice Models with Generalized Finite Dependence and Correlated Unobserved Heterogeneity”, working paper.

[21] Hadley, G. (1961): *Linear Algebra*, Addison-Wesley.

[22] Hotz, V. J., and R. Miller (1993): “ Conditional Choice Probabilities and Estimation of Dynamic Models,” *Review of Economic Studies*, 60, 497-529.

[23] Hotz, V. J., R. Miller, S. Sanders, and J. Smith (1994): “ A Simulation Estimator for Dynamic Models of Discrete Choice,” *Review of Economic Studies*, 61, 265-289.

[24] James, J. (2014): "Learning and Occupational Sorting", working paper.

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

[26] Khorunzhina, N. (2013): "Structural Estimation of Stock Market Participation Costs," *Journal of Economic Dynamics & Control,* 37, 2928-2942.

[27] Khorunzhina, N. and R. Miller (2016): "American Dream Delayed: Shifting Determinants of Homeownership in the U.S.," working paper.

[28] Ma, L. (2013): "Learning in a Hedonic Framework: Valuing Brownfield Remediation," working paper.

[29] Matsumoto, B. (2014): "Lighting the Fires: Explaining Youth Smoking Initiation and Experimentation in the Context of a Rational Addiction Model with Learning," working paper.

[30] Mazur, J. (2014): "Can Stricter Bankruptcy Laws Discipline Capital Investment?" working paper.

[31] Miller, R. (1984): "Job Matching and Occupational Choice," *Journal of Policitical Economy*, 92, 1086 -1120.

[32] 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, 373-399.

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

[34] Ransom, T. (2014): "The Effect of Business Cycle Fluctuations on Migration Decisions," working paper.

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

[36] Ryan, S. (2012): "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, 80, 1019–1062.

[37] Scott, P. (2013): "Dynamic Discrete Choice Estimation of Agricultural Land Use," working paper.