Nonstationary dynamic models with finite dependence

Peter Arcidiacono
Department of Economics, Duke University and NBER

Robert A. Miller
Tepper School of Business, Carnegie Mellon University

The estimation of nonstationary 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.

Keywords. Dynamic discrete choice, finite dependence, conditional choice probabilities.

JEL classification. C33, C35.

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 computationally cheap way of estimating structural payoff parameters in both single-agent and multiagent settings. These two-step estimators first estimate conditional choice probabilities (CCPs) and then characterize future payoffs as a function of the CCPs 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

---

Peter Arcidiacono: psarcidi@econ.duke.edu
Robert A. Miller: ramiller@cmu.edu

[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).

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

When a finite dependence representation exists, it is possible to relax some of the assumptions about time that are commonly made when estimating 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 terminal or renewal action.[4] 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, Hincapie, and Miller (2018), migration (Bishop (2012), Coate (2016), Ma (forthcoming), Ransom (2018)), participation in the stock market Khorunzhina (2013), agricultural land use Scott (2013), smoking Matsumoto (2014), education Arcidiacono, Aucejo, Maurel, and Ransom (2016), 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.[5]

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

---

[3]See Hotz and Miller (1993), Altug and Miller (1998), Arcidiacono and Miller (2011), Aguirregabiria and Magesan (2013, 2017), and Gayle (2017).

[4]See, for example, Hotz and Miller (1993), Joensen (2009), Scott (2013), Arcidiacono, Bayer, Blevins, and Ellickson (2016), Declerq and Verboven (2018), Mazur (2017), and Beauchamp (2015). The last three exploit one-period finite dependence to estimate dynamic games.

[5]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 (forthcoming). 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.

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, yet the algorithm we propose for dynamic optimization problems 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.

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.

Up until now, research on finite dependence in games has been restricted to models with a terminal action (that ends the process governing the state variables for individual players). Otherwise one-period finite dependence typically 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. In the general case, a bilinear system of equations must be solved, where the number of equations is dictated by the possible states that can be reached a few periods ahead, but in some specializations, including but not limited to terminal and renewal actions, our algorithm reduces to solving a linear system of equations.

The rest of the paper proceeds as follows. Section 2 lays out our framework for analyzing finite dependence in discrete choice dynamic optimization problems and noncooperative equilibrium games. In Section 3, we define finite dependence, and show how

this property can be used in estimation, generalizing existing estimators that exploit finite dependence to order to accommodate the many new applications our algorithm on finite dependence reveals. The fourth section provides a new representation of this property, and uses the representation to demonstrate how to recover finite dependence paths in single agent optimization problems. Section 5 extends the approach to multiagent equilibrium settings. New examples with finite dependence, derived using the algorithm, are provided in Section 6, while Section 7 concludes.

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 $t \in \{1, \dots, T\}$ until $T \leq \infty$, an individual chooses among J mutually exclusive actions. Let d_{jt} equal one if action $j \in \{1, \dots, J\}$ is taken at time t and zero otherwise. The current period payoff for action j at time t depends on the state $x_t \in \mathcal{X}$, a finite set.[6] 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 $\epsilon_t \equiv (\epsilon_{1t}, \dots, \epsilon_{Jt})$ 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 $E[\max\{\epsilon_{1t}, \dots, \epsilon_{Jt}\}] \leq \overline{\epsilon} < \infty$. 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 $d_t \equiv (d_{1t}, \dots, d_{Jt})$ to sequentially maximize the discounted sum of payoffs:

$$E\left\{ \sum_{t=1}^{T} \sum_{j=1}^{J} \beta^{t-1} d_{jt} [u_{jt}(x_t) + \epsilon_{jt}] \right\}, \tag{2.1}$$

where at each period t the expectation is taken over the future values of $x_{t+1}, \dots, x_T$ and $\epsilon_{t+1}, \dots, \epsilon_T$. Expression (2.1) 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 jth 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 integrating over ϵ_t:

$$p_{jt}(x_t) \equiv \int d_{jt}^o(x_t, \epsilon_t) g(\epsilon_t) \, d\epsilon_t. \tag{2.2}$$

---

[6]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.

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

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:

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

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

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

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:[7]

$$v_{jt}(x_t) = u_{jt}(x_t) + \beta \sum_{x_{t+1}=1}^{X} V_{t+1}(x_{t+1}) f_{jt}(x_{t+1}|x_t). \tag{2.3}$$

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 $t \in \{1, \dots, T\}$ and $p \in \Delta^J$, the J dimensional simplex, there exists a real-valued function $\psi_j(p)$ such that

$$\psi_j\big[p_t(x)\big] \equiv V_t(x) - v_{jt}(x). \tag{2.4}$$

To interpret (2.4), 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 precommitting to j versus waiting until ϵ_t is observed and only then making an optimal choice, $V_t(x_t)$, is the constant $\psi_j[p_t(x_t)]$ minus $E[\epsilon_{jt}]$, 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 $j \in \{1, \dots, J\}$ and $\tau \in \{t+1, \dots, T\}$, let any $\omega_\tau(x_\tau, j)$ denote any mapping from the state space $\{1, \dots, X\}$ to R^J satisfying the constraints that*

---

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

$|\omega_{k\tau}(x_\tau, j)| < \infty$ and $\sum_{k=1}^{J} \omega_{k\tau}(x_\tau, j) = 1$. Recursively, define $\kappa_{\tau+1}(x_{\tau+1}|x_t, j)$ as

$$\kappa_{\tau+1}(x_{\tau+1}|x_t, j)$$

$$\equiv \begin{cases} f_{jt}(x_{t+1}|x_t) & \text{for } \tau = t, \\ \sum_{x_\tau=1}^{X} \sum_{k=1}^{J} \omega_{k\tau}(x_\tau, j) f_{k\tau}(x_{\tau+1}|x_\tau) \kappa_\tau(x_\tau|x_t, j) & \text{for } \tau = t+1, \dots, T. \end{cases} \tag{2.5}$$

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

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

$$+ \sum_{x_{\mathcal{T}+1}}^{X} \beta^{\mathcal{T}+1-t} V_{\mathcal{T}+1}(x_{\mathcal{T}+1})\kappa_{\mathcal{T}+1}(x_{\mathcal{T}+1}|x_t, j), \tag{2.6}$$

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

$$v_{jt}(x_t) = u_{jt}(x_t) + \sum_{\tau=t+1}^{\mathcal{T}} \sum_{k=1}^{J} \sum_{x_\tau=1}^{X} \beta^{\tau-t}\big[u_{k\tau}(x_\tau) + \psi_k[p_\tau(x_\tau)]\big]\omega_{k\tau}(x_\tau, j)\kappa_\tau(x_\tau|x_t, j). \tag{2.7}$$

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) proved the theorem when $T = \mathcal{T}$ and $\omega_{k\tau}(x_\tau, j) \geq 0$ for all k and τ. In that case, $\kappa_{\tau+1}(x_{\tau+1}|x_t, j)$ 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.[8]

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 $t \in \{1, \dots, T\}$. The systematic part of payoffs to the nth player not only depends on his own choice in period t, denoted by $d_t^{(n)} \equiv (d_{1t}^{(n)}, \dots, d_{Jt}^{(n)})$, and the state variables x_t, but also the choices of the other players, which we now denote by $d_t^{(\sim n)} \equiv (d_t^{(1)}, \dots, d_t^{(n-1)}, d_t^{(n+1)}, \dots, d_t^{(N)})$. Denote by $U_{jt}^{(n)}(x_t, d_t^{(\sim n)}) + \epsilon_{jt}^{(n)}$ the flow 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 nth player's capital may increase his payoffs and reduce the payoffs to the others. For this reason, the payoff function is superscripted by n.

The players make simultaneous choices in each period. We denote by $P_t(d_t^{(\sim n)}|x_t)$ the joint conditional choice probability that the players aside from n collectively choose

---

[8]The extension to negative weights is also noted in Gayle (2017).

$d_t^{(\sim n)}$ at time t given the state variables x_t. Since $\epsilon_t^{(n)}$ is independently distributed across all the players, $P_t(d_t^{(\sim n)}|x_t)$ has the product representation:

$$P_t\big(d_t^{(\sim n)}|x_t\big) = \prod_{\substack{n'=1 \\ n'\neq n}}^{N} \left(\sum_{j=1}^{J} d_{jt}^{(n')} p_{jt}^{(n')}(x_t) \right). \tag{2.8}$$

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 (2.8). Taking the expectation of $U_{jt}^{(n)}(x_t, d_t^{(\sim n)})$ 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

$$u_{jt}^{(n)}(x_t) = \sum_{d_t^{(\sim n)} \in J^{N-1}} P_t\big(d_t^{(\sim n)}|x_t\big) U_{jt}^{(n)}\big(x_t, d_t^{(\sim n)}\big). \tag{2.9}$$

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.[9] The partition can be expressed as $x_t \equiv (x_t^{(0)}, x_t^{(1)}, \ldots, x_t^{(N)})$, where $x_t^{(0)}$ denotes the states that are exogenously determined by transition probability $f_{0t}(x_{t+1}^{(0)}|x_t^{(0)})$, and $x_t^{(n)} \in \mathcal{X}^{(n)} \equiv \{1, \ldots, X^{(n)}\}$ is the component of the state controlled or influenced by player n. Let $f_{jt}^{(n)}(x_{t+1}^{(n)}|x_t^{(n)})$ 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.[10] 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

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

---

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

[10]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.

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

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

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

$$\omega_\tau^{(n)}(x_\tau, j) \equiv \big(\omega_{1\tau}^{(n)}(x_\tau, j), \dots, \omega_{J\tau}^{(n)}(x_\tau, j)\big)$$

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 $f_t^{(\sim n)}(x_{t+1}^{(\sim n)}|x_t)$, we recursively define $\kappa_{\tau+1}^{(n)}(x_{\tau+1}|x_t, j)$ for the sequence of decision weights $\omega_{k\tau}^{(n)}(x_\tau, j)$ over periods $\tau \in \{t+1, \dots, T\}$ in a similar manner to (2.5) as

$$\kappa_{\tau+1}^{(n)}(x_{\tau+1}|x_t, j)$$
$$\equiv f_{0\tau}\big(x_{\tau+1}^{(0)}|x_\tau^{(0)}\big) \sum_{x_\tau=1}^{X} \sum_{k=1}^{J} f_\tau^{(\sim n)}\big(x_{\tau+1}^{(\sim n)}|x_\tau\big) \omega_{k\tau}^{(n)}(x_\tau, j) f_{k\tau}^{(n)}\big(x_{\tau+1}^{(n)}|x_\tau^{(n)}\big) \kappa_\tau^{(n)}(x_\tau|x_t, j) \quad (2.10)$$

with initializing function:

$$\kappa_{t+1}^{(n)}(x_{t+1}|x_t, j) \equiv f_{jt}^{(n)}\big(x_{t+1}^{(n)}|x_t^{(n)}\big) f_t\big(x_{t+1}^{(\sim n)}|x_t\big) f_{0t}\big(x_{t+1}^{(0)}|x_t^{(0)}\big). \quad (2.11)$$

Letting

$$f_{jt}\big(x_{t+1}|x_t\big) = f_{0t}\big(x_{t+1}^{(0)}|x_t^{(0)}\big) f_t^{(\sim n)}\big(x_{t+1}^{(\sim n)}|x_t\big) f_{jt}^{(n)}\big(x_{t+1}^{(n)}|x_t^{(n)}\big) \quad (2.12)$$

and adding n superscripts to all the other terms in (2.7), it now follows that Theorem 1 applies to this multiagent setting in exactly the same way as in a single agent setting.

3. THE FINITE DEPENDENCE PROPERTY

Theorem 1 shows that the future value term can be expressed relative to any weighted choice sequence as long as the sum of the weights add up to one in each period. Given that many paths can be chosen, it may be possible to line up the distribution of states given two different initial choices at some point in the future, say ρ periods later. If this is the case, then expressing the future value terms relative to these sequences results in the future value terms after ρ periods cancel out once differences in the conditional value function are taken across the two choices. Hence any information that would result in differences between the two choices in the future is already embedded in the conditional choice probabilities. In this section, we formalize the concept of finite dependence. We then show how it can be used in estimation.

3.1 *Defining finite dependence*

Turning first to individual optimization problems, 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

$$\kappa_{t+\rho+1}(x_{t+\rho+1}|x_t, i) = \kappa_{t+\rho+1}(x_{t+\rho+1}|x_t, j) \tag{3.1}$$

for all $x_{t+\rho+1} \in \{1, \ldots, X\}$. That is, the weights associated with each state are the same across the two paths after ρ periods.[11]

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 (3.1) 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 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. 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.

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

$$u_{jt}(x_t) - u_{it}(x_t) = \psi_i[p_t(x_t)] - \psi_j[p_t(x_t)]$$
$$+ \sum_{\tau=t+1}^{t+\rho} \sum_{k=1}^{J} \sum_{x_\tau=1}^{X} \beta^{\tau-t}\{u_{k\tau}(x_\tau) + \psi_k[p_\tau(x_\tau)]\}$$
$$\times [\omega_{k\tau}(x_\tau, i)\kappa_\tau(x_\tau|x_t, i) - \omega_{k\tau}(x_\tau, j)\kappa_\tau(x_\tau|x_t, j)]. \tag{3.2}$$

This equation follows directly from equations (2.4) and (2.7), in Theorem 1.[12]

Extending the definition of finite dependence to dynamic games is straightforward. It is applied to a given equilibrium at date t in state x_t, one with choice i and the other with choice j taken by a given player n. The two sequences of decision weights apply to the future choices of n when the other players follow their equilibrium strategies. Equation (2.12) defines the transition probabilities, while (2.11) and (2.10) determine the transitions of $\kappa_{\tau+1}^{(n)}(x_{\tau+1}|x_t, j)$. Thus in this multiagent setting ρ-period dependence

---

[11]Aguirregabiria and Magesan (2013, 2017) and Gayle (2017) restricted 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.

[12]Appealing to (2.4), replace $v_{jt}(x)$ with $V_t(x) - \psi_j[p_t(x)]$ in (2.7) and perform a similar substitution for $v_{it}(x)$. Upon differencing the two equations, the $V_t(x)$ terms drop out.

exists if there is a pair of sequences that give, for all $x_{t+\rho+1} \in \{1, \ldots, X\}$:

$$\kappa_{t+\rho+1}^{(n)}(x_{t+\rho+1}|x_t, i) = \kappa_{t+\rho+1}^{(n)}(x_{t+\rho+1}|x_t, j). \tag{3.3}$$

3.2 *Exploiting finite dependence in estimation*

When finite dependence holds, estimation may be much computationally less demanding. The empirical applications we cited in the introduction illustrate estimators based on finite dependence have appealing computational advantages. Equation (3.2) provides a basis for estimation without resorting to the inverting high dimensional matrices or simulating future paths. In addition, finite dependence has empirical content; it is straightforward to test whether (3.1) is rejected by the data.

To illustrate how to exploit finite dependence in estimation, suppose the data comprise N observations of the state variables and decisions denoted by $\{d_{nt_n}, x_{nt_n}, x_{n,t_n+1}\}_{n=1}^N$ sampled within a time frame of $t \in \{1, \ldots, \mathcal{T}\}$. Say there are M separate instances of finite dependence as defined in (3.1) within that time frame where, for the sake of exposition, each pair of choices includes choice 1. Label the M paths by (j_m, x_m, t_m, ρ_m) for $m \in \{1, \ldots, M\}$.[13] Assume that for each $t \in \{1, \ldots, \mathcal{T}\}$ the probability of the sample selection mechanism drawing $x \in \{1, \ldots, X\}$ is strictly positive.[14]

We then make the standard assumptions in the literature. First, assume the subjective discount factor β, and $g(\epsilon_t)$, the joint probability density function for the unobserved idiosyncratic taste shock ϵ_t, are known.[15] Second, assume $u_{jt}(x)$ can be parameterized by a finite dimensional vector $\theta \equiv (\theta_1, \ldots, \theta_K) \in \Theta$, a closed convex set in $\mathcal{R}^K$, and normalize the first choice to zero, by writing $u_{jt}(x) = \tilde{u}_{jt}(x, \theta)$, where $\tilde{u}_{jt}(x, \theta)$ is a known function with $\tilde{u}_{0t}(x, \theta) = 0$ for all (t, x).[16] Finally, assume that the M instances of finite dependence are sufficient to identify θ.[17]

We propose the following minimum distance CCP estimator for θ, new to the literature:

1. For all $t \in \{1, \ldots, \mathcal{T}\}$ and $x \in \{1, \ldots, X\}$, define the cell estimators of $p_{jt}(x)$ as

$$\widehat{p}_{jt}(x) \equiv \frac{\sum_{n=1}^N 1\{d_{nt(n)j} = 1\}1\{t_n = t\}1\{x_{nt(n)} = x\}}{\sum_{n=1}^N 1\{t_n = t\}1\{x_{nt(n)} = x\}}$$

---

[13]For example, in models with a renewal action or a terminal choice, every other choice at every state exhibits one- period dependence so in these cases $H = X(J - 1)\mathcal{T}$.

[14]This assumption is made for expositional simplicity: the state space could be redefined to be time specific, including only those states that are reached with strictly positive probability in each period $t \in \{1, \ldots, \mathcal{T}\}$.

[15]Both assumptions can be relaxed without losing identification depending on how restrictive are the assumptions on the functional form of $u_{jt}(x)$.

[16]Note that $u_{jt}(x)$ can be represented as a $(J - 1)X\mathcal{T}$ dimensional vector, so this parameterization amounts to imposing at most $(J - 1)X\mathcal{T} - K$ restrictions on that vector. For more details on identifying such models in nonstationary settings, see Arcidiacono and Miller (forthcoming).

[17]A necessary condition for identification is then $M > K$.

and estimate the $XJ\mathcal{T}$ CCP vector $p \equiv (p_{11}(1), \ldots, p_{J\mathcal{T}}(X))'$ with $\widehat{p}$ formed from $\widehat{p}_{jt}(x)$. Also, if the state transitions are unknown, estimate $f_{jt}(x)$ with $\widehat{f}_{jt}(x)$ in this first stage, for example, with a cell estimator (similar to the CCP estimator).

2. Let $\quad y(p, f) \equiv (y_1(p, f), \ldots, y_H(p, f))'\quad$ and $\quad Z(p, f, \theta) \equiv (Z_1(p, f, \theta), \ldots,$ $Z_M(p, f, \theta))'$ where

$$
y_m(p, f) \equiv \psi_1\big[p_{t(m)}(x_m)\big] - \psi_{j(m)}\big[p_{t(m)}(x_m)\big]
$$
$$
+ \sum_{\tau=t_m+1}^{t_m+\rho_m} \sum_{k=1}^{J} \sum_{x_\tau=1}^{X} \beta^{\tau-t(m)} \psi_k\big[p_\tau(x_\tau)\big]
$$
$$
\times \big[\omega_{k\tau}(x_\tau, 1)\kappa_\tau(x_\tau|x_m, 1) - \omega_{k\tau}(x_\tau, j_m)\kappa_\tau(x_\tau|x_m, j_m)\big],
$$
$$
Z_m(p, f, \theta) \equiv \widetilde{u}_{j(m), t(m)}(x_m, \theta)
$$
$$
- \sum_{\tau=t_m+1}^{t_m+\rho_m} \sum_{k=1}^{J} \sum_{x_\tau=1}^{X} \beta_{k\tau}^{\tau-t(m)} \widetilde{u}_{k\tau}(x_\tau, \theta)
$$
$$
\times \big[\omega_{k\tau}(x_\tau, 1)\kappa_\tau(x_\tau|x_m, 1) - \omega_{k\tau}(x_\tau, j_h)\kappa_\tau(x_\tau|x_m, j_m)\big].
$$

3. Let W denote an M dimensional positive definite matrix and choose θ to minimize:

$$
\big[y(\widehat{p}, \widehat{f}) - Z(\widehat{p}, \widehat{f}, \theta)\big]' W\big[y(\widehat{p}) - Z(\widehat{p}, \widehat{f}, \theta)\big]. \tag{3.4}
$$

It is straightforward to show that $\widehat{\theta}$, the solution to (3.4), is $\sqrt{N}$ consistent and asymptotically normal, because (3.1) implies $y_t(x, p, f) = Z_t(x, p, f, \theta)$ at the true parameter values.[18] For example, denoting a consistent estimate of the inverse of the asymptotic covariance matrix of $(\widehat{p}', \widehat{f}')'$ by $\widehat{W}$, and setting $W = \widehat{W}$, the asymptotic covariance matrix of $\widehat{\theta}$ is $[Z(\widehat{p}, \widehat{f}, \theta)/\partial\theta' \widehat{W} Z(\widehat{p}, \widehat{f}, \theta)/\partial\theta]^{-1}$. When W is diagonal matrix, (3.4) reduces to nonlinear least squares in this case. A second specialization is to assume $\widetilde{u}_{jt}(x, \theta)$ is linear in θ. In this case, the solution to (3.4) has a closed form and when $W = \widehat{W}$:

$$
\widehat{\theta} = \big\{\big[\partial Z(\widehat{p}, \widehat{f}, \theta)/\partial\theta\big]' \widehat{W}\big[\partial Z(\widehat{p}, \widehat{f}, \theta)/\partial\theta\big]\big\}^{-1}\big[\partial Z(\widehat{p}, \widehat{f}, \theta)/\partial\theta\big]' \widehat{W} y(\widehat{p}, \widehat{f}).
$$

Finally, the estimator carries over to the games case with minimal notational changes.

4. Finite dependence in individual optimization problems

We now turn to determining when finite dependence holds. 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

---

[18]Note that a full solution approach, based on solving the underlying dynamic programming problem for each value of $\theta \in \Theta$ does not exist when $\mathcal{T} < T$ unless the econometrician makes strong assumptions about the functional form utility takes in all periods $\tau \in \{\mathcal{T}+1, \ldots, T\}$ beyond the end of the data.

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 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 $j \in \{1, \ldots, J\}$ with a terminal or renewal choice leads to same value of state variables after two periods, because for all x_{t+2}:

$$\sum_{x_{t+1}=1}^{X} f_{1,t+1}(x_{t+2}|x_{t+1})f_{jt}(x_{t+1}|x_t) = \sum_{x_{t+1}=1}^{X} f_{1,t+1}(x_{t+2}|x_{t+1})f_{1t}(x_{t+1}|x_t). \tag{4.1}$$

Therefore, equation (3.1) is satisfied at $t + 2$ for all $j \in \{1, \ldots, J\}$ and $x \in \mathcal{X}$ by setting weights $\omega_{k,t+1}(x_{t+1}, j) = 1$ if $k = 1$ and zero otherwise.

4.1 *One-period dependence in optimization problems with two choices*

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 (2.5) implies that one-period dependence holds in this specialization at x_t if and only if there exists a weighting rule such that $\kappa_{t+2}(x'|x_t, 1) = \kappa_{t+2}(x'|x_t, 2)$ 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

$$\sum_{x_{t+1}=1}^{X} \left\{ \left[f_{2,t+1}(x'|x_{t+1}) - f_{1,t+1}(x'|x_{t+1}) \right] \right.$$

$$\times \left[\omega_{t+1}(x_{t+1}, 2)f_{2t}(x_{t+1}|x_t) - \omega_{t+1}(x_{t+1}, 1)f_{1t}(x_{t+1}|x_t) \right] \right\}$$

$$= \sum_{x_{t+1}=1}^{X} f_{1,t+1}(x'|x_{t+1})\left[f_{1t}(x_{t+1}|x_t) - f_{2t}(x_{t+1}|x_t) \right] \tag{4.2}$$

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 $\kappa_{t+2}(x'|x_t, j)$ 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. Therefore, 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 (4.2) 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 $\mathcal{A}_{1,t+1} \cup \mathcal{A}_{2,t+1}$ 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 $x \in \mathcal{A}_{1,t+1} \cup \mathcal{A}_{2,t+1}$ 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 (4.2) reduces to a linear system of $A_{t+2} - 1$ equations with $A_{1,t+1} + A_{2,t+1}$ unknowns.[19]

Denote by $\mathsf{K}_{j,t+1}(\mathcal{A}_{j,t+1})$ 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 $t + 1$.[20] A typical element of $\mathsf{F}_{k,t+1}(\mathcal{A}_{j,t+1})$ is $f_{k,t+1}(x'|x)$ 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 (4.2), 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,

$$\mathcal{A}_{1,t+1} = \mathcal{A}_{2,t+1} = \mathcal{A}_{t+2} = \mathcal{X}, \qquad \Omega_{t+1}(\mathcal{A}_{j,t+1}, j) = \Omega_{t+1}(\mathcal{X}, j),$$

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

---

[19] 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

[20] 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.

so we can write

$$\Omega_{t+1}(\mathcal{X}, j) \circ \mathsf{K}_{j,t+1}(\mathcal{X}) = \left[\omega_{t+1}(1, j) f_{jt}(1|x_t) \quad \cdots \quad \omega_{t+1}(X, j) f_{jt}(X|x_t) \right]',$$

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

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

Stacking the equations in (4.2) for all $x' \in \{1, \ldots, X-1\}$, the left-hand side of the stack is a linear combination of four expressions, each taking the form:

$$\begin{bmatrix} \sum_{x_{t+1}=1}^{X} f_{k,t+1}(1|x_{t+1}) \omega_{t+1}(x_{t+1}, j) f_{jt}(x_{t+1}|x_t) \\ \vdots \\ \sum_{x_{t+1}=1}^{X} f_{k,t+1}(X-1|x_{t+1}) \omega_{t+1}(x_{t+1}, j) f_{jt}(x_{t+1}|x_t) \end{bmatrix}$$
$$= \left[\mathsf{F}_{k,t+1}(\mathcal{X}) \right]' \left[\Omega_{t+1}(\mathcal{X}, j) \circ \mathsf{K}_{j,t+1}(\mathcal{X}) \right]. \tag{4.3}$$

Note that when $k = 2$, equation (4.3) 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 $\tilde{x} \notin \mathcal{A}_{j,t+1}$, that is, when $f_{jt}(\tilde{x}|x_t) = 0$, we remove the element $\omega_{t+1}(\tilde{x}, j) f_{jt}(\tilde{x}|x_t)$ from $\Omega_{t+1}(\mathcal{X}, j) \circ \mathsf{K}_{jt}(\mathcal{X})$ and the $\tilde{x}$th row in $\mathsf{F}_{k,t+1}(\mathcal{X})$. This reduces the dimension of $\Omega_{t+1}(\mathcal{X}, j) \circ \mathsf{K}_{j,t+1}(\mathcal{X})$ to $\mathcal{A}_{j,t+1}$ and the dimension of $\mathsf{F}_{k,t+1}(\mathcal{X})$ from $X \times (X-1)$ to $\mathcal{A}_{j,t+1} \times (X-1)$. Similarly, if $\hat{x} \notin \mathcal{A}_{t+2}$, in words if $\hat{x}$ is unattainable given either initial choice regardless of the weighting rules at $t + 1$, then we remove the $\hat{x}$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 $\mathcal{A}_{j,t+1} \times (\mathcal{A}_{t+2} - 1)$ matrix.

Substituting these transformations into (4.2), we now express the system of $\mathcal{A}_{t+2} - 1$ equations with $\mathcal{A}_{1,t+1} + \mathcal{A}_{2,t+1}$ unknowns in matrix form. Define the $\mathcal{A}_{t+2} - 1$ dimensional vector $\mathcal{K}_{t+1}$, and the $(\mathcal{A}_{t+2} - 1) \times (\mathcal{A}_{1,t+1} + \mathcal{A}_{2,t+1})$ matrix H_{t+1}, respectively as

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

$$\mathsf{H}_{t+1} \equiv \begin{bmatrix} \mathsf{F}_{2,t+1}(\mathcal{A}_{2,t+1}) - \mathsf{F}_{1,t+1}(\mathcal{A}_{2,t+1}) \\ \mathsf{F}_{1,t+1}(\mathcal{A}_{1,t+1}) - \mathsf{F}_{2,t+1}(\mathcal{A}_{1,t+1}) \end{bmatrix}.$$

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

$$\mathcal{K}_{t+1} = \mathsf{H}_{t+1} \begin{bmatrix} \Omega_{t+1}(\mathcal{A}_{2,t+1}, 2) \circ \mathsf{K}_{2,t+1}(\mathcal{A}_{2,t+1}) \\ \Omega_{t+1}(\mathcal{A}_{1,t+1}, 1) \circ \mathsf{K}_{1,t+1}(\mathcal{A}_{1,t+1}) \end{bmatrix} \equiv \mathsf{H}_{t+1} \mathsf{D}_{t+1}. \tag{4.4}$$

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. A solution to (4.4) for D_{t+1} exists if and only if the rank of H_{t+1} equals the rank of the augmented matrix $\mathsf{H}_{t+1}^* \equiv [\mathcal{K}_{t+1} \vdots \mathsf{H}_{t+1}]$ 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 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 (4.4) by the matching K vectors, yielding

$$\begin{bmatrix} \Omega_{t+1}(\mathcal{A}_{2,t+1}, 2) \\ \Omega_{t+1}(\mathcal{A}_{1,t+1}, 1) \end{bmatrix} = \mathsf{H}_{t+1}^{-1} \mathcal{K}_{t+1} \circ / \begin{bmatrix} \mathsf{K}_{2,t+1}(\mathcal{A}_{2,t+1}) \\ \mathsf{K}_{1,t+1}(\mathcal{A}_{1,t+1}) \end{bmatrix}, \tag{4.5}$$

where $\circ/$ 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 (4.4) 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 (4.5) is solved for the weights characterizing finite dependence.[21]

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 $\overline{\mathsf{H}}_{t+1}$. 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 $\overline{\mathsf{D}}_{t+1}$. We now eliminate $A_{t+2} - R_{t+1} - 1$ rows of $\overline{\mathsf{H}}_{t+1}$ to form an R_{t+1} dimensional square matrix with rank R_{t+1} denoted by $\overline{\mathsf{H}}_{t+1}$. Strictly for notational purposes, so without loss of generality, we reorder the equations defining (4.4) so that the linearly independent equations are the bottom ones. This allows us to partition $\overline{\mathsf{H}}_{t+1}' \equiv [\overline{\mathsf{H}}_{t+1}' \vdots \underline{\mathsf{H}}_{t+1}']$ and $\mathcal{K}_{t+1}' \equiv [\overline{\mathcal{K}}_{t+1}' \vdots \underline{\mathcal{K}}_{t+1}']$, where $\underline{\mathsf{H}}_{t+1}$ is $(A_{t+2} - 1 - R_{t+1}) \times R_{t+1}$, while $\underline{\mathcal{K}}_{t+1}'$ is $(A_{t+2} - 1 - R_{t+1}) \times 1$ and $\overline{\mathcal{K}}_{t+1}$ is $R_{t+1} \times 1$. Inverting $\overline{\mathsf{H}}_{t+1}$, we obtain $\overline{\mathsf{D}}_{t+1} = \overline{\mathsf{H}}_{t+1}^{-1} \overline{\mathcal{K}}_{t+1}$. Thus a solution to (4.4) attains in this knife edged case if and only if $\overline{\mathsf{D}}_{t+1}$ solves $A_{t+2} - R_{t+1} - 1$ additional equations $\underline{\mathcal{K}}_{t+1} = \underline{\mathsf{H}}_{t+1} \overline{\mathsf{H}}_{t+1}^{-1} \overline{\mathcal{K}}_{t+1}$.

---

[21]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 $x = X$.[22] 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:

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

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

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

Substituting these expressions into (4.5), 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/nonterminal action to achieve one-period dependence:

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

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.

4.2 *Solving nonlinear systems to attain ρ-period dependence*

Analyzing the existence of finite dependence for $\rho > 1$ introduces nonlinearity into the system. For convenience, we relabel the two initial choices i and j in equation (3.1) as 1 and 2, and the initial state as x_t. 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.[23] 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.[24] Define $\mathsf{K}_{j\tau}(\mathcal{A}_{j\tau})$ 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. 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

---

[22]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 $f_{2,t+1}(X|X) = 1$.

[23]For example, suppose $\mathcal{X} \equiv \{1, 2, 3\}$ and $x_t = 3$. Also assume $f_{1,t+1}(1|3) = 3/4$, and $f_{1,t+1}(2|3) = 1/4$. Then the the first two states are attainable in $t+1$ from taking the first choice but the third is not.

[24]In our simple example, $A_{1,t+1} = 2$ and $\mathcal{A}_{1,t+1} = \{1, 2\}$.

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, with $\widetilde{\mathsf{F}}_{k\tau}(\mathcal{A}_{j\tau})$ containing all the columns of the transition matrix. The matrix comprises elements $f_{k\tau}(x'|x)$ for each $x \in \mathcal{A}_{j\tau}$ and $x' \in \mathcal{A}_{\tau+1}$.

The $A_{\tau+1}$ system of equations exhibits ρ-period dependence, that is $\kappa_{\tau+1}(x_{\tau+1}|x_t, 1) = \kappa_{\tau+1}(x_{\tau+1}|x_t, 2)$ with $\tau = t + \rho$, if and only if there exist vectors $\Omega_{k\tau}(\mathcal{A}_{j\tau}, 1)$ and $\Omega_{k\tau}(\mathcal{A}_{j\tau}, k)$ for each $k \in \{2, \dots, J\}$ solving:

$$\mathcal{K}_{\tau+1} \equiv \mathsf{F}_{1\tau}(\mathcal{A}_{1\tau})\mathsf{K}_{1\tau}(\mathcal{A}_{1\tau}) - \mathsf{F}_{1\tau}(\mathcal{A}_{2\tau})\mathsf{K}_{2\tau}(\mathcal{A}_{2\tau}) = \mathsf{H}_{\tau}\mathsf{D}_{\tau}, \tag{4.6}$$

where 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}, are respectively defined by[25]

$$\mathsf{H}_{\tau} \equiv \begin{bmatrix} \mathsf{F}_{2\tau}(\mathcal{A}_{2\tau}) - \mathsf{F}_{1\tau}(\mathcal{A}_{2\tau}) \\ \vdots \\ \mathsf{F}_{J\tau}(\mathcal{A}_{2\tau}) - \mathsf{F}_{1\tau}(\mathcal{A}_{2\tau}) \\ \mathsf{F}_{1\tau}(\mathcal{A}_{1\tau}) - \mathsf{F}_{2\tau}(\mathcal{A}_{1\tau}) \\ \vdots \\ \mathsf{F}_{1\tau}(\mathcal{A}_{1\tau}) - \mathsf{F}_{J\tau}(\mathcal{A}_{1\tau}) \end{bmatrix}', \qquad \mathsf{D}_{\tau} \equiv \begin{bmatrix} \Omega_{2\tau}(\mathcal{A}_{2\tau}, 2) \circ \mathsf{K}_{2\tau}(\mathcal{A}_{2\tau}) \\ \vdots \\ \Omega_{J\tau}(\mathcal{A}_{2\tau}, 2) \circ \mathsf{K}_{2\tau}(\mathcal{A}_{2\tau}) \\ \Omega_{2\tau}(\mathcal{A}_{1\tau}, 1) \circ \mathsf{K}_{1\tau}(\mathcal{A}_{1\tau}) \\ \vdots \\ \Omega_{J\tau}(\mathcal{A}_{1\tau}, 1) \circ \mathsf{K}_{1\tau}(\mathcal{A}_{1\tau}) \end{bmatrix}. \tag{4.7}$$

Appealing to Hadley (1961, pp. 168–169) yields necessary and sufficient conditions for the existence of a solution to this linear system, which we state as a theorem.

THEOREM 2. *Define the $(A_{\tau+1} - 1) \times \{(J - 1)[A_{1\tau} + A_{2\tau}] + 1\}$ matrix $\mathsf{H}_{\tau}^* \equiv [\mathsf{H}_{\tau} \vdots \mathcal{K}_{t+1}]$, 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}^*.*

Theorem 2 shows that establishing one-period dependence when there are more than two choices is a straightforward extension of the case in which $J = 2$. However, nonlinearity in the weights enter (4.6) when $\rho > 1$ because $\mathsf{K}_{j\tau}(\mathcal{A}_{j\tau})$ depends on $\Omega_{ks}(\mathcal{A}_{2s}, j)$, the weight on action $k \in \{2, \dots, J\}$ for every period $s < \tau$ given initial choice $j \in \{1, 2\}$. Denote $\widetilde{\mathsf{K}}_{j\tau}(\mathcal{A}_{\tau})$ as the A_{τ} vector containing the weights for transitioning to each of the A_{τ} states, that is, the attainable states from *either* path-given the initial choice of j. $\mathsf{K}_{j\tau}(\mathcal{A}_{j\tau})$ is then the nonzero entries of $\widetilde{\mathsf{K}}_{j\tau}(\mathcal{A}_{\tau})$. The following recursive structure is then evident:

$$\widetilde{\mathsf{K}}_{j\tau}(\mathcal{A}_{\tau}) = \begin{bmatrix} \widetilde{\mathsf{F}}_{2,\tau-1}(\mathcal{A}_{j,\tau-1}) \\ \vdots \\ \widetilde{\mathsf{F}}_{J,\tau-1}(\mathcal{A}_{j,\tau-1}) \end{bmatrix}' \begin{bmatrix} \Omega_{2,\tau-1}(\mathcal{A}_{j,\tau-1}, j) \circ \mathsf{K}_{j,\tau-1}(\mathcal{A}_{j,\tau-1}) \\ \vdots \\ \Omega_{J,\tau-1}(\mathcal{A}_{j,\tau-1}, j) \circ \mathsf{K}_{j,\tau-1}(\mathcal{A}_{j,\tau-1}) \end{bmatrix}. \tag{4.8}$$

---

[25]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.

Taking the nonzero elements out of $\widetilde{K}_{j\tau}(\mathcal{A}_\tau)$ to form $K_{j\tau}(\mathcal{A}_{j\tau})$ and substituting in for $K_{j\tau}(\mathcal{A}_{j\tau})$ using (4.8) in (4.7) and (4.6) demonstrates that cross products of elements in $\Omega_{2,\tau-1}(\mathcal{A}_{j,\tau-1}, j)$ and $\Omega_{2\tau}(\mathcal{A}_{2\tau}, 2)$ enter (4.8). Formally, the system is bilinear, not linear.

To see that the system is bilinear, suppose $J = 2$ and write $\omega_\tau(x_\tau, j) \equiv \omega_{2\tau}(x_\tau, j)$: expanding (4.6) term by term proves that two-period dependence exists for some given x_t if and only if:

$$\sum_{x_{t+2}=1}^{X} \sum_{x_{t+1}=1}^{X} f_{1,t+2}(x_{t+3}|x_{t+2}) f_{1,t+1}(x_{t+2}|x_{t+1}) \big[f_{1t}(x_{t+1}|x_t) - f_{2t}(x_{t+1}|x_t) \big]$$

$$= \sum_{x_{t+2}=1}^{X} \sum_{x_{t+1}=1}^{X} \big[f_{2,t+2}(x_{t+3}|x_{t+2}) - f_{1,t+2}(x_{t+3}|x_{t+2}) \big]$$

$$\times \big[f_{2,t+1}(x_{t+2}|x_{t+1}) - f_{1,t+1}(x_{t+2}|x_{t+1}) \big]$$

$$\times \big[\omega_{t+2}(x_{t+2}, 2) \omega_{t+1}(x_{t+1}, 2) f_{2t}(x_{t+1}|x_t)$$

$$- \omega_{t+2}(x_{t+2}, 1) \omega_{t+1}(x_{t+1}, 1) f_{1t}(x_{t+1}|x_t) \big]$$

$$+ \sum_{x_{t+2}=1}^{X} \sum_{x_{t+1}=1}^{X} \big[f_{2,t+2}(x_{t+3}|x_{t+2}) - f_{1,t+2}(x_{t+3}|x_{t+2}) \big]$$

$$\times f_{1,t+1}(x_{t+2}|x_{t+1}) \omega_{t+2}(x_{t+1}, 1) f_{1t}(x_{t+1}|x_t)$$

$$\times \big[\omega_{t+2}(x_{t+1}, 2) f_{2t}(x_{t+1}|x_t) - \omega_{t+2}(x_{t+1}, 1) f_{1t}(x_{t+1}|x_t) \big]$$

$$+ \sum_{x_{t+2}=1}^{X} \sum_{x_{t+1}=1}^{X} f_{1,t+2}(x_{t+3}|x_{t+2}) \big[f_{2,t+1}(x_{t+2}|x_{t+1}) - f_{1,t+1}(x_{t+2}|x_{t+1}) \big]$$

$$\times \big[\omega_{t+1}(x_{t+1}, 2) f_{2t}(x_{t+1}|x_t) - \omega_{t+1}(x_{t+1}, 1) f_{1t}(x_{t+1}|x_t) \big] \tag{4.9}$$

for all $x_{t+3} \in \mathcal{X}$. Since products of weights appear in (4.9), bilinear solution techniques are required to solve this problem. More generally, cross products to the power of ρ enter into the equation system defining ρ-period dependence.

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 the linear inversion problem to which Theorem 2 applies. 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. Having established existence, we can obtain weights satisfying (3.1) as a by-product.

There are an infinite number of weighting schemes, each of which might conceivably establish finite dependence, a fact that might explain why researchers have opted for guess and verify methods when designing models exhibiting 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 $\tau \in \{t + 1, \dots, t + \rho\}$, 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.

5. FINITE DEPENDENCE IN GAMES

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 (2017), Collard-Wexler (2013), Dunne, Klimek, Roberts, and Xu (2013), and Ryan (2012)) all exhibit the finite dependence property that could be used to simplify estimation.

In principle, the methods developed above are directly applicable to dynamic games off short panels, that is, after defining $f_{jt}(x_{t+1}|x_t)$ with (2.12). Let $\mathsf{F}_{k\tau}^{(n)}(\mathcal{A}_{j\tau})$ denote the first $A_\tau - 1$ columns of the transition matrix from $\mathcal{A}_{j\tau}$ to $\mathcal{A}_{\tau+1}$ given choice k by player n at time τ when everyone else plays their equilibrium strategy and let $\widetilde{\mathsf{F}}_{k\tau}^{(n)}(\mathcal{A}_{j\tau})$ denote the transition matrix containing all the columns. These are defined analogously to $\mathsf{F}_{k\tau}(\mathcal{A}_{j\tau})$ and $\widetilde{\mathsf{F}}_{k\tau}(\mathcal{A}_{j\tau})$ in the individual optimization case. Also let $\Omega_{k,\tau}^{(n)}(\mathcal{A}_{2\tau}, j)$ denote a vector of weights on choice k for each of the $A_{j\tau}$ states in $\mathcal{A}_{j\tau}$. Finally, let $\mathsf{K}_{j\tau}^{(n)}(\mathcal{A}_{j\tau})$ denote the τ-period transition probabilities to $\mathcal{A}_{j\tau}$ when n initially chooses j and follows the weights when everybody else plays their equilibrium strategies. Analogous to (4.6),

ρ period dependence holds for the first two actions if

$$
\begin{bmatrix}
\mathsf{F}_{2\tau}^{(n)}(\mathcal{A}_{2\tau}) - \mathsf{F}_{1\tau}^{(n)}(\mathcal{A}_{2\tau}) \\
\vdots \\
\mathsf{F}_{J\tau}^{(n)}(\mathcal{A}_{2\tau}) - \mathsf{F}_{1\tau}^{(n)}(\mathcal{A}_{2\tau}) \\
\mathsf{F}_{1\tau}^{(n)}(\mathcal{A}_{1\tau}) - \mathsf{F}_{2\tau}^{(n)}(\mathcal{A}_{1\tau}) \\
\vdots \\
\mathsf{F}_{1\tau}^{(n)}(\mathcal{A}_{1\tau}) - \mathsf{F}_{J\tau}^{(n)}(\mathcal{A}_{1\tau})
\end{bmatrix}'
\begin{bmatrix}
\Omega_{2,\tau}^{(n)}(\mathcal{A}_{2\tau}, 2) \circ \mathsf{K}_{2\tau}^{(n)}(\mathcal{A}_{2\tau}) \\
\vdots \\
\Omega_{J\tau}^{(n)}(\mathcal{A}_{2\tau}, 2) \circ \mathsf{K}_{2\tau}^{(n)}(\mathcal{A}_{2\tau}) \\
\Omega_{2,\tau}^{(n)}(\mathcal{A}_{1\tau}, 1) \circ \mathsf{K}_{1\tau}^{(n)}(\mathcal{A}_{1\tau}) \\
\vdots \\
\Omega_{J\tau}^{(n)}(\mathcal{A}_{1\tau}, 1) \circ \mathsf{K}_{1\tau}^{(n)}(\mathcal{A}_{1\tau})
\end{bmatrix}
$$

$$
= \begin{bmatrix}
\mathsf{F}_{1\tau}^{(n)}(\mathcal{A}_{1\tau}) \\
-\mathsf{F}_{1\tau}^{(n)}(\mathcal{A}_{2\tau})
\end{bmatrix}'
\begin{bmatrix}
\mathsf{K}_{1\tau}^{(n)}(\mathcal{A}_{1\tau}) \\
\mathsf{K}_{2\tau}^{(n)}(\mathcal{A}_{2\tau})
\end{bmatrix}. \tag{5.1}
$$

In practice, establishing finite dependence is generally more onerous in games than in individual optimization problems. Finite dependence in a game is player specific; in principle finite dependence might hold for some players but not for others. Furthermore, the transition of the state variables for any one player taking a particular action depends on the equilibrium decisions of all the other players. Thus, finite dependence in games is ultimately a property that derives not just from the game primitives, but also equilibrium play. Consequently, games do not typically exhibit one-period finite dependence: if two different choices of n at time t affect the other players' equilibrium choices in $t + 1$ (or later), it is generally not feasible to line up all the states $x_{t+2} \equiv (x_{t+2}^{(0)}, x_{t+2}^{(1)}, \ldots, x_{t+2}^{(N)})$ across both paths emanating from the respective initial choices of n within two periods.

A key feature of the incomplete information games settings we consider is that at t, when the players other than n collectively choose $d_t^{(\sim n)}$, they condition on the lagged choice of n (i.e., how $d_{t-1}^{(n)}$ affects $x_t^{(n)}$), but not on $d_t^{(n)}$, the current choice of n. Our approach to determining finite dependence in games exploits this feature in the following way. First, we obtain necessary and sufficient conditions for player n to take a sequence of weighted actions inducing, say after $\rho - 1$ periods, the other players to take actions at $t + \rho$ that match up the weight distributions of $x_{t+\rho+1}^{(\sim n)}$, conditional on $x_t^{(\sim n)}$, meaning:[26]

$$
\kappa_{t+\rho+1}\big(x_{t+\rho+1}^{(\sim n)} | x_t, i\big) = \kappa_{t+\rho+1}\big(x_{t+\rho+1}^{(\sim n)} | x_t, j\big). \tag{5.2}
$$

Clearly, (5.2) is a necessary condition for (3.3) to hold. Second, with one last choice of weight pairs at $t + \rho$, player n lines up the joint distribution of the states of all the players, setting $\omega_{k,t+\rho}^{(n)}(x_{t+\rho}, i)$ and $\omega_{k,t+\rho}^{(n)}(x_{t+\rho}, j)$, and incorporating the restrictions that give (5.2), so that (3.3) simultaneously holds.

---

[26]This inducement is based on the other players following their equilibrium strategies.

5.1 *Finite dependence for state components controlled by other players*

From (3.1), finite dependence at τ requires

$$\sum_{x_\tau=1}^{X}\sum_{k=1}^{J} f_\tau^{(\sim n)}\big(x_{\tau+1}^{(\sim n)}|x_\tau\big) f_{k\tau}^{(n)}\big(x_{\tau+1}^{(n)}|x_\tau^{(n)}\big)\omega_{k\tau}^{(n)}(x_\tau, j)\kappa_\tau^{(n)}(x_\tau|x_t, j)$$

$$= \sum_{x_\tau=1}^{X}\sum_{k=1}^{J} f_\tau^{(\sim n)}\big(x_{\tau+1}^{(\sim n)}|x_\tau\big) f_{k\tau}^{(n)}\big(x_{\tau+1}^{(n)}|x_\tau^{(n)}\big)\omega_{k\tau}^{(n)}(x_\tau, i)\kappa_\tau^{(n)}(x_\tau|x_t, i). \tag{5.3}$$

Necessary and sufficient conditions for (5.3) to hold are found in the same way as finite dependence is determined for individual optimization problems. They are based on the intuition that from periods t through $\tau - 1$ player n takes pairs of weighted actions starting with i and j that induce the other other players to align the probability distributions for $x_{\tau+1}^{(\sim n)}$ through their equilibrium choices.

A necessary condition for τ dependence comes from summing (5.3) over the $x_{\tau+1}^{(n)}$ outcomes. Noting that

$$\sum_{x_{\tau+1}^{(n)}=1}^{X^{(n)}}\sum_{x_\tau=1}^{X} f_\tau^{(\sim n)}\big(x_{\tau+1}^{(\sim n)}|x_\tau\big)\left[\sum_{k=1}^{J}\omega_{k\tau}^{(n)}(x_\tau, j) f_{k\tau}^{(n)}\big(x_{\tau+1}^{(n)}|x_\tau^{(n)}\big)\right]\kappa_\tau^{(n)}(x_\tau|x_t, j)$$

$$= \sum_{x_\tau=1}^{X} f_\tau^{(\sim n)}\big(x_{\tau+1}^{(\sim n)}|x_\tau\big)\left[\sum_{k=1}^{J}\omega_{k\tau}^{(n)}(x_\tau, j)\right]\left[\sum_{x^{(n)}=1}^{X^{(n)}} f_{k\tau}^{(n)}\big(x_{\tau+1}^{(n)}|x_\tau^{(n)}\big)\right]\kappa_\tau^{(n)}(x_\tau|x_t, j)$$

$$= \sum_{x_\tau=1}^{X} f_\tau^{(\sim n)}\big(x_{\tau+1}^{(\sim n)}|x_\tau\big)\kappa_\tau^{(n)}(x_\tau|x_t, j) \tag{5.4}$$

we simplify the sum (5.3) over $x_{\tau+1}^{(n)}$ using (5.4) to obtain

$$\sum_{x_\tau=1}^{X} f_\tau^{(\sim n)}\big(x_{\tau+1}^{(\sim n)}|x_\tau\big)\big[\kappa_\tau^{(n)}(x_\tau|x_t, j) - \kappa_\tau^{(n)}(x_\tau|x_t, i)\big] = 0. \tag{5.5}$$

This proves that whether (5.5) holds or not depends on the weights assigned to n in periods $t + 1$ though $\tau - 1$, but not on the period τ weights.

To derive a rank condition under which (5.5) holds, it is notationally convenient to focus on the first two choices as before. Suppose (5.5) 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 $\mathcal{A}_{j,\tau-1} \subseteq \mathcal{X}$ denote the set of attainable states at $\tau - 1$ for the weight sequence beginning with n choosing $j \in \{1, 2\}$. Let $\mathcal{A}_\tau \subseteq \mathcal{X}$ denote the set of attainable states at τ 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 attainable states of the other players at $\tau + 1$ given the two weight sequences. Let $A_{\tau+1}^{(\sim n)}$ denote the number of elements in

$\mathcal{A}_{\tau+1}^{(\sim n)}$. Let $\mathsf{P}_\tau^{(\sim n)}(\mathcal{A}_\tau)$ denote the transpose of the first $A_{\tau+1}^{(\sim n)} - 1$ columns of the transition matrix from $\mathcal{A}_\tau$ to the set of competitor states $\mathcal{A}_{\tau+1}^{(\sim n)}$. Finally, define $\mathsf{H}_\tau^{(\sim n)}$ and $\mathcal{K}_{\tau+1}^{(\sim n)}$ as

$$
\mathsf{H}_\tau^{(\sim n)} \equiv \mathsf{P}_\tau^{(\sim n)}(\mathcal{A}_\tau)
\begin{bmatrix}
\widetilde{\mathsf{F}}_{2,\tau-1}^{(n)}(\mathcal{A}_{2,\tau-1}) - \widetilde{\mathsf{F}}_{1,\tau-1}^{(n)}(\mathcal{A}_{2,\tau-1}) \\
\vdots \\
\widetilde{\mathsf{F}}_{J,\tau-1}^{(n)}(\mathcal{A}_{2,\tau-1}) - \widetilde{\mathsf{F}}_{1\tau-1}^{(n)}(\mathcal{A}_{2,\tau-1}) \\
\widetilde{\mathsf{F}}_{1,\tau-1}^{(n)}(\mathcal{A}_{1,\tau-1}) - \widetilde{\mathsf{F}}_{2,\tau-1}^{(n)}(\mathcal{A}_{1,\tau-1}) \\
\vdots \\
\widetilde{\mathsf{F}}_{1,\tau-1}^{(n)}(\mathcal{A}_{1,\tau-1}) - \widetilde{\mathsf{F}}_{J,\tau-1}^{(n)}(\mathcal{A}_{1,\tau-1})
\end{bmatrix}',
\tag{5.6}
$$

$$
\mathcal{K}_{\tau+1}^{(\sim n)} \equiv \mathsf{P}_\tau^{(\sim n)}(\mathcal{A}_\tau)\left[\widetilde{\mathsf{F}}_{1,\tau-1}^{(\sim n)}(\mathcal{A}_{1,\tau-1})\mathsf{K}_{1,\tau-1}^{(n)}(\mathcal{A}_{1,\tau-1}) - \widetilde{\mathsf{F}}_{1,\tau-1}^{(\sim n)}(\mathcal{A}_{2,\tau-1})\mathsf{K}_{2,\tau-1}^{(n)}(\mathcal{A}_{2,\tau-1})\right]. \tag{5.7}
$$

Finite dependence requires weighting rules from $t + 1$ to $\tau - 1$ 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 $\mathsf{P}_\tau^{(\sim n)}(\mathcal{A}_\tau)$ in (5.6). Thus the similarity of $\mathsf{H}_\tau^{(\sim n)}$ and H_τ is evident from comparing (5.6) with (4.7); likewise the similarities between $\mathcal{K}_{\tau+1}^{(\sim n)}$ and $\mathcal{K}_{\tau+1}$ are obvious from (4.6) and (5.7). Following the same logic as Theorem 2, we obtain the following result.

THEOREM 4. *Given an initial period and state* (t, x_t), *and initial choices* 1 *and* 2, *(5.5) holds for all* $x^{(\sim n)} \in X^{(\sim n)}$ *if and only if there exists a pair of weight sequences defining* $\mathsf{H}_\tau^{(\sim n)}$ *and* $\mathcal{K}_{\tau+1}^{(\sim n)}$ *such that* $\mathsf{H}_\tau^{(\sim n)}$ *and* $[\mathsf{H}_\tau^{(\sim n)} \vdots \mathcal{K}_{\tau+1}^{(\sim n)}]$ *have the same rank.*

5.2 *Aligning the joint distributions*

For one specialization, checking the conditions of Theorem 4 suffices to determine whether (3.3) holds or not. Suppose that for each $x \in \mathcal{A}_\tau$, there is an action $d^{(n)}(x)$ yielding some fixed $\overline{x}^{(n)} \in X^{(n)}$ for sure.[27] Then satisfying the conditions of Theorem 4 imply the conditions of Theorem 2 are met too. In this specialization, the joint distribution across the two paths is aligned in $\tau + 1$ because (i) $\kappa_\tau^{(\sim n)}(x_{\tau+1}^{(\sim n)}|x_t, j)$, the marginal weight distribution of the other players' states is aligned, and (ii) the state of player n does not vary across the states of the other players. Thus verifying finite dependence reduces to finding conditions that satisfy (5.2) in this case. Renewal and terminal actions provide examples because the renewal or terminal state, $\overline{x}^{(n)} \in \mathcal{X}^{(n)}$, can be reached from any $x^{(n)} \in \mathcal{X}^{(n)}$ in one period with certainty. Section 6.2 illustrates our step-by-step procedure for establishing finite dependence in a coordination game.

A second special case occurs when the rank of $\mathsf{H}_\tau^{(\sim n)}$ is $A_{\tau+1}^{(\sim n)} - 1$ and the set of weights is unique. We first derive the unique set of weights $\omega_{k\tau}^{(n)}(x_\tau, 1)$ and $\omega_{k\tau}^{(n)}(x_\tau', 2)$ for $\tau \in$ in this linear subproblem; then following the approach in the preceding subsection, we show below, as a special case of a more general result, that whether a set of weights

---

[27]More generally, there exists one action $d^{(n)}(x)$, or some weighted mixture of actions, that when applied to either sequence, yields the same weight distribution over $x^{(n)} \in X^{(n)}$ for all $x \in \mathcal{A}_\tau^{(n)}$.

exists establishing τ- period finite dependence or not, reduces to solving a second linear problem in $\omega_{k\tau}^{(n)}(x_\tau, 1)$ and $\omega_{k\tau}^{(n)}(x'_\tau, 2)$ for $k \in \{2, \ldots, J\}$ and $x_\tau \in \mathcal{A}_{1,\tau+1}$ and $x'_\tau \in \mathcal{A}_{2,\tau+1}$, similar to those analyzed in the single agent problems.

If the set of weights equalizing the marginal distributions for $x_{t+\rho+1}^{(\sim n)} \in X^{(\sim n)}$ is not unique, then an uncountable number do, since any convex combination of say two sets of weights also equalize the marginal distributions. Since the weights determining the solution to the states of the other players also help determine the conditional distribution for $x_{t+\rho+1}^{(n)}$, the selection of a solution for the other players may impact whether the conditional distributions for $x_{t+\rho+1}^{(n)}$ can be aligned or not.

To treat both cases formally, let $\Omega_{k,\tau-1}^{(n)}(\mathcal{A}_{2,\tau-1}, j)$ denote an $A_{j,\tau-1}$ dimensional row vector of unknown weights assigning a real number to choice $k \in \{2, \ldots, J\}$ for each state in $\mathcal{A}_{j,\tau-1}$ at $\tau - 1$, given strictly positive weights $\mathsf{K}_{2,\tau-1}^{(n)}(\mathcal{A}_{2,\tau-1})$. Denote $\mathbb{1}_{J-1}$ as a $(J-1)$ column vector of ones. Define $\Omega_{\tau-1}^{(n)}$ and $\mathsf{K}_{\tau-1}^{(n)}$ as

$$\Omega_{\tau-1}^{(n)} \equiv \begin{bmatrix} \Omega_{2,\tau-1}^{(n)}(\mathcal{A}_{2,\tau-1}, j) \\ \vdots \\ \Omega_{J,\tau-1}^{(n)}(\mathcal{A}_{2,\tau-1}, j) \\ \Omega_{2,\tau-1}^{(n)}(\mathcal{A}_{1,\tau-1}, j) \\ \vdots \\ \Omega_{J,\tau-1}^{(n)}(\mathcal{A}_{1,\tau-1}, j) \end{bmatrix}', \qquad \mathsf{K}_{\tau-1}^{(n)} \equiv \begin{bmatrix} \mathbb{1}_{J-1} \otimes \mathsf{K}_{2,\tau-1}^{(n)}(\mathcal{A}_{2,\tau-1}) \\ \mathbb{1}_{J-1} \otimes \mathsf{K}_{1,\tau-1}^{(n)}(\mathcal{A}_{1,\tau-1}) \end{bmatrix}, \qquad (5.8)$$

where $\otimes$ is the Kronecker product.

If there exists weights $\Omega_{\tau-1}^{(n)}$ solving

$$\mathcal{K}_{\tau+1}^{(\sim n)} = \mathsf{H}_\tau^{(\sim n)}\big(\Omega_{\tau-1}^{(n)} \circ \mathsf{K}_{\tau-1}^{(n)}\big) \qquad (5.9)$$

then a necessary condition for finite dependence embodied in (5.5), relating the weights of all the players aside from n, is satisfied. In the special case where $(A_{1,\tau-1} + A_{2,\tau-1})(J - 1) = A_{\tau+1}^{(\sim n)} - 1$ and $\mathsf{H}_\tau^{(\sim n)} = \overline{\mathsf{H}}_\tau^{(\sim n)}$ inverts, from (5.9),

$$\Omega_{\tau-1}^{(n)} = \big[\mathsf{H}_\tau^{(\sim n)}\big]^{-1} \mathcal{K}_{\tau+1}^{(\sim n)} \circ / \mathsf{K}_{\tau-1}^{(n)}. \qquad (5.10)$$

More generally, let $\overline{\mathsf{D}}_{\tau-1}^{(n)}$ denote an $A_{\tau+1}^{(\sim n)} - 1$ dimensional vector given by $\overline{\Omega}_{\tau-1}^{(n)} \circ \overline{\mathsf{K}}_{\tau-1}^{(n)}$ and $\underline{\mathsf{D}}_{\tau-1}^{(n)}$ a vector of dimension $(A_{1,\tau-1} + A_{2,\tau-1})(J - 1) - (A_{\tau+1}^{(\sim n)} - 1)$ given by $\underline{\Omega}_{\tau-1}^{(n)} \circ \underline{\mathsf{K}}_{\tau-1}^{(n)}$. Also partition $\mathsf{H}_\tau^{(\sim n)}$ at the $(A_{\tau+1}^{(\sim n)} - 1)$th column, writing $\mathsf{H}_\tau^{(\sim n)} = [\overline{\mathsf{H}}_\tau^{(\sim n)} \ \underline{\mathsf{H}}_\tau^{(\sim n)}]$, where $\overline{\mathsf{H}}_\tau^{(\sim n)}$ conforms to $\overline{\mathsf{D}}_{\tau-1}^{(n)}$, and $\underline{\mathsf{H}}_\tau^{(\sim n)}$ to $\underline{\mathsf{D}}_{\tau-1}^{(n)}$. Then (5.9) can be expressed as

$$\mathcal{K}_{\tau+1}^{(\sim n)} = \overline{\mathsf{H}}_\tau^{(\sim n)} \overline{\mathsf{D}}_{\tau-1}^{(n)} + \underline{\mathsf{H}}_\tau^{(\sim n)} \underline{\mathsf{D}}_{\tau-1}^{(n)}. \qquad (5.11)$$

From (5.11), it is evident that whether a solution to $\overline{\mathsf{D}}_{\tau-1}^{(n)}$ exists hinges on $\overline{\mathsf{H}}_\tau^{(\sim n)}$, but not on the values of $\underline{\mathsf{D}}_{\tau-1}^{(n)}$. For example, a sufficient condition for a solution to the first step

is that the rank of $\overline{\mathsf{H}}_\tau^{(\sim n)}$ equals $\mathcal{A}_\tau^{(\sim n)}$.[28] Moreover, from (5.11) when the rank condition for $\overline{\mathsf{H}}_\tau^{(\sim n)}$ is satisfied, $\overline{\mathsf{D}}_{\tau-1}^{(n)}$ varies with $\underline{\mathsf{D}}_{\tau-1}^{(n)}$; specifically

$$\overline{\mathsf{D}}_{\tau-1}^{(n)} = \left[\overline{\mathsf{H}}_\tau^{(\sim n)}\right]^{-1}\left(\mathcal{K}_{\tau+1}^{(\sim n)} - \underline{\mathsf{H}}_\tau^{(\sim n)}\underline{\mathsf{D}}_{\tau-1}^{(n)}\right). \tag{5.12}$$

To establish finite dependence at period $\tau + 1$ for the joint system, it suffices to show that there exists some $\underline{\mathsf{D}}_{\tau-1}^{(n)}$ and a set of weights on the choices made by n in period τ solving the joint system when we incorporate the effects of $\underline{\mathsf{D}}_{\tau-1}^{(n)}$ operating through $\mathsf{K}_{j,\tau-1}^{(n)}$. Modify (4.8) by superscripting with n the $\widetilde{\mathsf{F}}_{j,\tau-1}^{(n)}(\mathcal{A}_{j,\tau-1})$ transitions, as well as the weight vectors $\widetilde{\mathsf{K}}_{j\tau}(\mathcal{A}_\tau)$ and $\mathsf{K}_{j,\tau-1}(\mathcal{A}_{j,\tau-1})$, to indicate the player for whom finite dependence is being checked. Also replace the vector formed from the elements $\Omega_{k,\tau-1}(\mathcal{A}_{j,\tau-1}, j) \circ \mathsf{K}_{j,\tau-1}(\mathcal{A}_{j,\tau-1})$ with $\mathsf{D}_{\tau-1}^{(n)}$. Then (4.8) becomes

$$\widetilde{\mathsf{K}}_{j\tau}^{(n)}(\mathcal{A}_{j\tau}) = \begin{bmatrix} \widetilde{\mathsf{F}}_{2,\tau-1}^{(n)}(\mathcal{A}_{j,\tau-1}) \\ \vdots \\ \widetilde{\mathsf{F}}_{J,\tau-1}^{(n)}(\mathcal{A}_{j,\tau-1}) \end{bmatrix}' \mathsf{D}_{\tau-1}^{(n)}. \tag{5.13}$$

Form the vector $\widehat{\mathsf{D}}_{\tau-1}^{(n)}$ by replacing the elements in $\overline{\mathsf{D}}_{\tau-1}^{(n)}$ with the linear mappings defined in (5.12), and substitute for $\mathsf{K}_{j\tau}^{(n)}(\mathcal{A}_{j\tau})$ using the nonzero elements of (5.13) into (5.1). These operations yield a bilinear system of equations to be solved in $\Omega_\tau^{(n)}$ and $\underline{\mathsf{D}}_{\tau-1}^{(n)}$. We then check for a solution by minimizing a quadratic norm of the equation system. Finite dependence is achieved when the quadratic norm attains a value of zero for some $\Omega_\tau^{(n)}$ and $\underline{\mathsf{D}}_{\tau-1}^{(n)}$. The production quality game considered in Section 6.3 illustrates this stepwise procedure.

6. Applications

This section provides three 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, and in the process highlight the practical importance of using negative weights. 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 when the conditions for the specialization discussed in the previous section hold. Third, we analyze a product quality game that does not satisfy the conditions for the specialization and defies a guess and verify approach.

6.1 *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

---

[28]Necessary and sufficient conditions are found in the same way as the single agent optimization case.

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.

6.1.1 *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

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

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

$$
\begin{aligned}
v_{2t}(x_t) &- v_{1t}(x_t) \\
&= \lambda_t \big[u_2(x_t) - u_1(x_t) + \beta u_1(x_t + 1) - \beta u_2(x_t) \big] \\
&+ \beta \bigg[\lambda_t \psi_1 \big[p_{t+1}(x_t + 1) \big] + \lambda_t \bigg(\frac{1}{\lambda_{t+1}} - 1 \bigg) \psi_1 \big[p_{t+1}(x_t) \big] - \frac{\lambda_{t+1}}{\lambda_t} \psi_2 \big[p_{t+1}(x_t) \big] \bigg]. \quad (6.1)
\end{aligned}
$$

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.

6.1.2 *Applying Theorem 2*　While Section 6.1.1 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 (4.4). $\mathcal{A}_{1,t+1}$ 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 $\mathsf{K}_{1,t+1}(\mathcal{A}_{1,t+1})$ and $\mathsf{K}_{2,t+1}(\mathcal{A}_{2,t+1})$ are $[1]$ and $[1 - \lambda \quad \lambda]'$. The relevant transition matrices are given by

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

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

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 (4.4) 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 (4.4):

$$\begin{bmatrix} F_{2t+1}(\mathcal{A}_{2,t+1}) - F_{1t+1}(\mathcal{A}_{2,t+1}) \\ F_{1t+1}(\mathcal{A}_{1,t+1}) - F_{2t+1}(\mathcal{A}_{1,t+1}) \end{bmatrix}' = \begin{bmatrix} -\lambda_{t+1} & 0 & \lambda_{t+1} \\ \lambda_{t+1} & -\lambda_{t+1} & -\lambda_{t+1} \end{bmatrix}. \tag{6.2}$$

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 (6.2) 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

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

Finite dependence can then be achieved by setting:

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

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.

6.2 *A coordination game*

To illustrate why finite dependence holds for a much broader class of games than those with terminal choices, we first consider the following simple 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 $d_t^{(n)} = 2$, not competing by setting $d_t^{(n)} = 1$. Let the superscript $\sim n$ refer to the rival player of n; we define the state space $x_t \equiv (x_t^{(n)}, x_t^{(\sim n)})$ from the nth player's perspective, and assume $x_t^{(n)} = d_{t-1}^{(n)}$. Therefore, the state variable transition matrix is deterministic and time invariant. Let $p_{2t}^{(n)}(x_t)$ denote the equilibrium probability of n competing at date t when the state variable is x_t, and analogously denote the probability of noncompeting by $p_{1t}^{(n)}(x_t) = 1 - p_{2t}^{(n)}(x_t)$. To prevent this game from degenerating to a single agent optimization problem, we assume $p_{2,t+1}^{(n)}(2, 1) \neq p_{2,t+1}^{(n)}(2, 2)$; in

equilibrium, the rival's actions affect the player's choice through the state variables. Conditional on the lagged participation of the other player, we also assume an individual's choices depend on his own lagged participation, implying $p_{2,t+1}^{(n)}(1,2) \neq p_{2,t+1}^{(n)}(2,2)$. Both assumptions 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.

We prove two-period dependence by construction. 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 initial action $j \in \{1, 2\}$ taken at time t, and set $\omega_{t+2}^{(n)}(x_{t+2}^{(n)}, j) = 1$, implying $x_{t+3}^{(n)} = 1$ for both initial actions $j \in \{1, 2\}$. This ensures $x_{t+3}^{(n)}$ is the same for both paths by setting the $t + 2$ choice weight to be the same across both paths.[29] All that remains is to find two weighting sequences for n, one for each initial choice $j \in \{1, 2\}$ at t, such that when the other player makes his equilibrium choice at $t + 2$, the distribution of $d_{t+2}^{(\sim n)}$, and hence the distribution of $x_{t+3}^{(\sim n)}$, is the same for both sequences. In this model, the rank condition for $H_\tau^{(\sim n)}$ is easy to check because $x_{t+3}^{(\sim n)} \equiv d_{t+2}^{(\sim n)}$ only takes two values. Theorem 5 establishes two period dependence by specifying a $\omega_{t+1}^{(n)}(x_{t+1}^{(n)}, j)$, 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.*

6.3 *A product quality game*

We now consider a game where the solution cannot readily be solved by hand. In the process, we outline an algorithm that, while not covering all cases, makes it easier to find finite dependence paths in games settings.

Setup The game we consider has two players $n \in \{1, 2\}$. In each period $t \in \{1, \ldots, T\}$, the players simultaneously decide whether to increase their product quality from $x_t^{(n)}$ to $x_t^{(n)} + 1$ subject to a maximal product quality of $\overline{x}$. With some probability π nature reduces the product quality of both players to a minimum denoted by $\underline{x}$. The expected profit for player n for maintaining product quality (choosing $j = 1$), and increasing product quality (choosing $j = 2$), net of an independent shock, are given by

$$u_1^{(n)}(x_t) = \ln[x_t^{(n)}]\{\alpha_0 + \alpha_1 E_t[\gamma^{(\sim n)}(x_t)]\},$$
$$u_2^{(n)}(x_t) = \ln[\min\{x_t^{(n)} + 1, \overline{x}\}]\{\alpha_0 + \alpha_1 E_t[\gamma^{(\sim n)}(x_t)]\} + \alpha_2$$

respectively, where

$$E_t[\gamma^{(\sim n)}(x_t)] = p_{1t}^{(\sim n)}(x_t) \ln[x_t^{(\sim n)}] + p_{2t}^{(\sim n)}(x_t) \ln[\min\{x_t^{(\sim n)} + 1, \overline{x}\}]$$

---

[29]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 $t + 2$.

is the expected logged product quality of the rival, α_0 gives the baseline returns to product quality, α_1 measures how profit is diminished by rivalry, and α_2 is the cost of increasing product quality. We assume that the payoff shock associated with each action is distributed Type 1 extreme value, and induce nonstationarity into the game by imposing a finite horizon.

In the numerical specification we analyze, the time horizon is set to $T = 20$, the maximal product quality to $\overline{x} = 25$, the minimal product quality to $\underline{x} = 2$, and the probability that both products become worthless to $\pi = 0.05$. Regarding preferences, we set the discount factor to $\beta = 0.9$, the baseline flow return from product quality to $\alpha_0 = 0.35$, the coefficient of the rival's product quality to $\alpha_1 = -0.15$, and the cost of increasing product quality to $\alpha_2 = -3$.

To solve for a symmetric pure strategy Markov perfect equilibrium, we first calculate the probabilities of taking each action in the period T states by solving a fixed-point problem in probability space. The period T solution (the equilibrium for the static model) gives us the expected future utility at period $T - 1$ for each of the possible choices. We then solve a fixed-point problem to obtain the choice probabilities in period $T - 1$, continuing this procedure until the first period.

Algorithm   We now show how the finite dependence properties of this game can be investigated using the techniques developed in this paper. Specifically, we check whether the game satisfies two-period dependence at $(t, x_t^{(n)}, x_t^{(\sim n)}) = (1, 4, 5)$, that is investigating finite dependence in the first period when the product quality of player n is 4 and the product quality of her rival is 5. Following the decomposition argument in Section 5.1, we first obtain the weighted choices of n in period 2 that induce finite dependence for $x_3^{(\sim n)}$, aligning the two marginal distributions of the rival's states in period 3. We then derive the period 3 weights that line up the two joint distributions for x_4, the states for both players in period 4.

Should nature destroy the product quality of both firms, the state is automatically reset independently of past actions. Hence to determine whether the game exhibits finite dependence, we only need to consider paths on which nature has no debilitating consequences. The description of the algorithm as applied to this example accordingly ignores this aspect of nature. A program solving this example is provided in the Online Supplementary Material (Arcidiacono and Miller (2019)).

1. Form the vector $\mathsf{K}_{j,t+1}^{(n)}(\mathcal{A}_{j,t+1})$ with dimension $\mathcal{A}_{j,t+1}$ and elements given by the probabilities associated with each states in $\mathcal{A}_{j,t+1}$. In our example, when player n chooses action 2 at time t, the state transitions from $(4, 5)$ to either $(5, 5)$ or $(5, 6)$, depending whether her rival takes action 1 or 2, which implies

$$\mathsf{K}_{2,t+1}^{(n)}(\mathcal{A}_{2,t+1}) = \begin{bmatrix} p_{1t}^{(\sim n)}(4, 5) & p_{2t}^{(\sim n)}(4, 5) \end{bmatrix}'.$$

2. Recall $\mathcal{A}_{t+2}$ are the attainable states at $t + 2$ for the two initial choices at t and any decision at $t + 1$. Form $\mathsf{F}_{k,t+1}^{(n)}(\mathcal{A}_{j,t+1})$, the transition matrix from $\mathcal{A}_{j,t+1}$ to $\mathcal{A}_{t+2}$, for all

(j, k), given choice k by n at $t + 1$ and equilibrium choices by the rival. Its columns, give the probabilities of transitioning from each of the states associated with $\mathsf{K}^{(n)}_{j,t+1}(\mathcal{A}_{j,t+1})$ to one of the possible states in $\mathcal{A}_{t+2}$; its rows refer to the states in $\mathcal{A}_{t+2}$. Our example features nine attainable states at $t + 2$ because each player can make the second choice zero times, once, or twice, and thus

$$\mathcal{A}_{t+2} = \big\{(4, 5), (5, 5), (6, 5), (4, 6), (5, 6), (6, 6), (4, 7), (5, 7), (6, 7)\big\}. \tag{6.3}$$

The rows of $\mathsf{F}^{(n)}_{k,t+1}(\mathcal{A}_{2,t+1})$ correspond to the possible states at $t + 1$ given action 2 was taken at t. Hence the dimension of $\mathsf{F}^{(n)}_{k,t+1}(\mathcal{A}_{2,t+1})$ is 2×9 and, ordering the columns following the $\mathcal{A}_{t+2}$ list of elements above:

$$\mathsf{F}^{(n)}_{2,t+1}\big(\mathcal{A}^{(n)}_{2,t+1}\big) = \begin{bmatrix} 0 & 0 & p^{(\sim n)}_{1,t+1}(5, 5) & 0 & 0 & p^{(\sim n)}_{2,t+1}(5, 5) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p^{(\sim n)}_{1,t+1}(5, 6) & 0 & 0 & p^{(\sim n)}_{2,t+1}(5, 6) \end{bmatrix}.$$

The other transition matrices $\mathsf{F}^{(n)}_{1,t+1}(\mathcal{A}_{1,t+1})$, $\mathsf{F}^{(n)}_{1,t+1}(\mathcal{A}_{2,t+1})$, and $\mathsf{F}^{(n)}_{2,t+1}(\mathcal{A}_{1,t+1})$, are formed in a similar way.

3. Form $\mathsf{P}^{(\sim n)}_{t+2}(\mathcal{A}_{t+2})$, the transpose of the transition matrix from $\mathcal{A}_{t+2}$ to $\mathcal{A}^{(\sim n)}_{t+3}$, truncated by a row (reflecting the linear dependence from summing the probabilities). A row of this matrix has the probabilities of transitioning to a particular other player state at $t + 3$ from each of the attainable states at $t + 2$. In our example, $\mathcal{A}^{(\sim n)}_{t+3} = \{5, 6, 7, 8\}$, and (without loss of generality) we drop the row associated with $x^{(\sim n)}_{t+3} = 8$. With reference to (6.3), the 3×9 dimensional matrix $\mathsf{P}^{(\sim n)}_{t+2}(\mathcal{A}_{t+2})$ takes the form:

$$\begin{bmatrix} p^{(\sim n)}_{1t+1}(4, 5) & p^{(\sim n)}_{1t+1}(5, 5) & p^{(\sim n)}_{1t+1}(6, 5) & 0 & 0 & 0 & 0 & 0 & 0 \\ p^{(\sim n)}_{2t+1}(4, 5) & p^{(\sim n)}_{2t+1}(5, 5) & p^{(\sim n)}_{2t+1}(6, 5) & p^{(\sim n)}_{1t+1}(4, 6) & p^{(\sim n)}_{1t+1}(5, 6) & p^{(\sim n)}_{1t+1}(6, 6) & 0 & 0 & 0 \\ 0 & 0 & 0 & p^{(\sim n)}_{2t+1}(4, 6) & p^{(\sim n)}_{2t+1}(5, 6) & p^{(\sim n)}_{2t+1}(6, 6) & p^{(\sim n)}_{1t+1}(4, 7) & p^{(\sim n)}_{1t+1}(5, 7) & p^{(\sim n)}_{1t+1}(6, 7) \end{bmatrix}.$$

4. Form $\mathcal{K}^{(\sim n)}_{t+3}$ defined by (5.7). In our example, $\mathcal{K}^{(\sim n)}_{t+3}$ is the 3×1 vector taking numerical value:

$$\mathcal{K}^{(\sim n)}_{t+3} = \mathsf{P}^{(\sim n)}_{t+2}(\mathcal{A}_\tau)\big[\mathsf{F}^{(n)}_{1,t+1}(\mathcal{A}_{1,t+1})\mathsf{K}^{(n)}_{1,t+1}(\mathcal{A}_{1,t+1}) - \mathsf{F}^{(n)}_{1,t+1}(\mathcal{A}_{2,t+1})\mathsf{K}^{(n)}_{2,t+1}(\mathcal{A}_{2,t+1})\big]$$

$$= \begin{bmatrix} -0.7081 & -0.7592 & 0.0899 \end{bmatrix}'. \tag{6.4}$$

5. Form $\mathsf{H}^{(\sim n)}_{t+2}$, the $(\mathcal{A}^{(\sim n)}_{t+3} - 1) \times \mathcal{A}_{t+2}$ dimensional matrix defined by (5.6). In our example, $\mathsf{P}^{(\sim n)}_{t+2}(\mathcal{A}_{t+2})$ is a 3×9 matrix whereas $\mathsf{F}^{(n)}_{2,t+1}(\mathcal{A}_{2,t+1}) - \mathsf{F}^{(n)}_{1,t+1}(\mathcal{A}_{2,t+1})$ and $\mathsf{F}^{(n)}_{1,t+1}(\mathcal{A}_{1,t+1}) - \mathsf{F}^{(n)}_{2,t+1}(\mathcal{A}_{1,t+1})$ are both 2×9 matrices, so $\mathsf{H}^{(\sim n)}_{t+2}$ is 3×4. Substituting the values of the equilibrium CCPs computed as the solution to the model into (6.5)

yields

$$
\begin{aligned}
\mathsf{H}_{t+2}^{(\sim n)} &= \mathsf{P}_{t+2}^{(\sim n)}\big(\mathcal{A}_{t+2}^{(n)}\big)
\begin{bmatrix}
\mathsf{F}_{2,t+1}^{(n)}\big(\mathcal{A}_{2,t+1}^{(n)}\big) - \mathsf{F}_{1,t+1}^{(n)}\big(\mathcal{A}_{2,t+1}^{(n)}\big) \\
\mathsf{F}_{1,t+1}^{(n)}\big(\mathcal{A}_{1,t+1}^{(n)}\big) - \mathsf{F}_{2,t+1}^{(n)}\big(\mathcal{A}_{1,t+1}^{(n)}\big)
\end{bmatrix}' \\
&=
\begin{bmatrix}
0.0892 & 0 & -0.0611 & 0 \\
-0.0139 & 0.0890 & -0.0802 & -0.0776 \\
-0.0753 & -0.0441 & 0.1413 & -0.0256
\end{bmatrix}.
\end{aligned}
\tag{6.5}
$$

If the rank of $\mathsf{H}_{t+2}^{(\sim n)}$ is less than A_{t+3}, then two-period dependence does not hold. Using (6.5) it is straightforward to verify that $\mathsf{H}_{t+2}^{(\sim n)}$ is rank three in our example.

6. Partition $\mathsf{H}_{t+2}^{(\sim n)}$ into $[\overline{\mathsf{H}}_{t+2}^{(\sim n)} \vdots \underline{\mathsf{H}}_{t+2}^{(\sim n)}]$ where $\overline{\mathsf{H}}_{t+2}^{(\sim n)}$ denotes a square matrix with dimension and rank A_{t+2} and $\underline{\mathsf{H}}_{t+2}^{(\sim n)}$ denotes a matrix comprising the remaining columns. In our example, any one of the four columns could be removed to yield a matrix of rank 3. Accordingly, we omit the first column of $\mathsf{H}_{t+2}^{(\sim n)}$ corresponding to state $(5,5)$ to obtain the 3×3 matrix $\overline{\mathsf{H}}_{t+2}^{(\sim n)}$, and the 3×1 vector $\underline{\mathsf{H}}_{t+2}^{(\sim n)} = [0.0892 \ -0.0139 \ -0.0753]'$.

7. Let $\overline{\mathsf{D}}_{t+1}^{(n)}$ denote an $A_{t+1} - 1$ dimensional real vector with generic component $\overline{\mathsf{D}}_{kj,t+1}^{(n)}$ for $k = 2$ and $j \in \{1,2\}$. Similarly, let $\underline{\mathsf{D}}_{t+1}^{(n)}$ denote an $A_{t+1} - (A_{t+3}^{(\sim n)} - 1)$ dimensional weight vector with generic elements $\underline{\mathsf{D}}_{kj,t+1}^{(n)}$, and solve for $\overline{\mathsf{D}}_{t+1}^{(n)}$ as a linear mapping in $\underline{\mathsf{D}}_{t+1}^{(n)}$ using (5.12). In our example, $\overline{\mathsf{D}}_{t+1}^{(n)}$ is 3×1 and $\underline{\mathsf{D}}_{t+1}^{(n)}$ is a real number. Substituting the numerical values for $\underline{\mathsf{H}}_\tau^{(\sim n)}$, $\overline{\mathsf{H}}_{t+2}^{(\sim n)}$, and $\mathcal{K}_{t+3}^{(\sim n)}$, (5.12) simplifies to

$$
\overline{\mathsf{D}}_{t+1}^{(n)} =
\begin{bmatrix}
-0.2643 \\
2.7522 \\
1.4614
\end{bmatrix}
+
\begin{bmatrix}
0.2560 \\
0.4073 \\
0.1114
\end{bmatrix}
\underline{\mathsf{D}}_{t+1}^{(n)}.
\tag{6.6}
$$

8. Substitute (5.12) and $\underline{\mathsf{D}}_{t+1}^{(n)}$ into (5.13) to obtain an expression for $\mathsf{K}_{j,t+2}^{(n)}(\mathcal{A}_{j,t+2})$ in terms of $\underline{\mathsf{D}}_{t+1}^{(n)}$ and $\mathcal{K}_{t+3}^{(\sim n)}$. In our example, this yields

$$
\begin{aligned}
\widetilde{\mathsf{K}}_{j,t+2}^{(n)}(\mathcal{A}_{j,t+2}) &= \mathsf{F}_{2,t+1}^{(n)}(\underline{\mathcal{A}}_{j,t+1})'\underline{\mathsf{D}}_{t+1}^{(n)} \\
&\quad + \mathsf{F}_{2,t+1}^{(n)}(\overline{\mathcal{A}}_{j,t+1})'\big[\overline{\mathsf{H}}_{t+2}^{(\sim n)}\big]^{-1}\big(\mathcal{K}_{t+3}^{(\sim n)} - \underline{\mathsf{H}}_{t+2}^{(\sim n)}\underline{\mathsf{D}}_{t+1}^{(n)}\big).
\end{aligned}
\tag{6.7}
$$

9. Substitute the linear expressions for $\mathsf{K}_{j,t+1}^{(n)}(\mathcal{A}_{j,t+1})$ into (5.1) and check for a solution to the resulting bilinear equation system. In our example, (5.1) specializes to the case where $J = 2$. There is two-period dependence if and only if the criterion function (6.8) defined below attains a minimal value of zero for some $\underline{\mathsf{D}}_{t+1}^{(\sim n)}$ and

TABLE 1. Weights that generate finite dependence.

| Time | State | Probability of $d_{2\tau}^{(n)} = 1$ | |
		$d_{1t}^{(n)} = 1$	$d_{2t}^{(n)} = 1$
$t+1$	$(4, 5)$	4.7859	
	$(4, 6)$	2.0178	
	$(5, 5)$		1.3928
	$(5, 6)$		1.0139
$t+2$	$(4, 5)$	1	
	$(4, 6)$	1	
	$(4, 7)$	1	
	$(5, 5)$	0	4.7475
	$(5, 6)$	0	11.0759
	$(5, 7)$	0	45.7515
	$(6, 5)$		0
	$(6, 6)$		0
	$(6, 7)$		0

$\Omega_{t+2}^{(n)}$:

$$
\begin{aligned}
&\Big\{\big[\mathsf{F}_{2,t+2}^{(n)}(\mathcal{A}_{2,t+2}) - \mathsf{F}_{1,t+2}^{(n)}(\mathcal{A}_{2,t+2})\big]'\big[\Omega_{2,t+2}^{(n)}(\mathcal{A}_{2,t+2}, 2) \circ \mathsf{K}_{2,t+2}^{(n)}(\mathcal{A}_{2,t+2})\big] \\
&\quad + \big[\mathsf{F}_{1,t+2}^{(n)}(\mathcal{A}_{1,t+2}) - \mathsf{F}_{2,t+2}^{(n)}(\mathcal{A}_{1,t+2})\big]'\big[\Omega_{2,t+2}^{(n)}(\mathcal{A}_{1,t+2}, 1) \circ \mathsf{K}_{1,t+2}^{(n)}(\mathcal{A}_{1,t+2})\big] \\
&\quad + \big[\mathsf{F}_{1,t+2}^{(n)}(\mathcal{A}_{2,t+2})\big]'\big[\mathsf{K}_{2,t+2}^{(n)}(\mathcal{A}_{2,t+2})\big] - \big[\mathsf{F}_{1,t+2}^{(n)}(\mathcal{A}_{1,t+2})\big]'\big[\mathsf{K}_{1,t+2}^{(n)}(\mathcal{A}_{1,t+2})\big]\Big\}^2. \qquad (6.8)
\end{aligned}
$$

In this example, we exploit the bilinear property of (6.8) by solving for $\Omega_{t+2}^{(n)}$ as a linear system in the scalar $\underline{\mathsf{D}}_{t+1}^{(\sim n)}$, then substituting the solution for $\Omega_{t+2}^{(n)}$ back into (6.8), and finally resolving the resulting system in the scalar $\underline{\mathsf{D}}_{t+1}^{(\sim n)}$. Weights giving a zero value to (6.8) are displayed in Table 1. Thus two-period dependence is established by construction.

7. Conclusion

CCP methods provide a computationally cheap way of estimating dynamic discrete choice models in both single-agent and multiagent 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 to 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. Thus our methods provide an approach to estimating nonstationary infinite horizon games even when there are no terminal actions.

<center>APPENDIX: PROOFS</center>

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 positivity or neg-
ativity of the weights is not used in that proof.   □

PROOF OF THEOREM 3. Denote by $\mathcal{A} \equiv \{x_{\mathcal{A}}^{(1)}, \ldots, x_{\mathcal{A}}^{(A)}\}$ 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

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

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

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

where $|\omega_{j\tau}(x_{\mathcal{A}}^{(a)})| < \infty$ and $\omega_{J\tau}(x_{\mathcal{A}}^{(a)}) \equiv 1 - \sum_{j=1}^{J-1} \omega_{j\tau}(x_{\mathcal{A}}^{(a)})$. Let $\mathcal{K}_{\mathcal{A}} \equiv (\mathcal{K}_{\mathcal{A}}^{(1)}, \ldots, \mathcal{K}_{\mathcal{A}}^{(A)})'$ de-
note 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
$|\mathcal{K}_{\mathcal{A}}^{(a)}| < \infty$ and $\mathcal{K}_{\mathcal{A}}^{(x)} \neq 0$. We also define

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

and note that

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

$$= \sum_{a=1}^A \sum_{j=1}^J \omega_{j\tau}(x_{\mathcal{A}}^{(a)}) \mathcal{K}_{\mathcal{A}}^{(a)} = \sum_{a=1}^A \mathcal{K}_{\mathcal{A}}^{(a)} = 1. \tag{A.1}$$

Depending on $\mathcal{K}_{\mathcal{A}}$, and also the choice of $\omega_{\tau\mathcal{A}} \equiv (\omega_\tau(x_{\mathcal{A}}^{(1)}), \ldots, \omega_\tau(x_{\mathcal{A}}^{(A)}))'$, some ele-
ments of $\mathcal{K}_{\mathcal{B}} \equiv (\mathcal{K}_{\mathcal{B}}^{(1)}, \ldots, \mathcal{K}_{\mathcal{B}}^{(B)})'$ 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 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 $\{\mathcal{A}_{j\tau}\}_{\tau=t+2}^\rho$. Al-
though 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 $\mathcal{K}_{\mathcal{B}} \equiv (\mathcal{K}_{\mathcal{B}}^{(1)}, \ldots, \mathcal{K}_{\mathcal{B}}^{(B)})'$ is a weighting for $\mathcal{A}^*$ if and only if

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

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 $\omega_\tau(x_{\mathcal{A}}^{(a)})$, each of dimension $J - 1$. The $B - 1$ equations correspond to the nonzero weights placed on the states $\{x_{\mathcal{B}}^{(2)}, \ldots, x_{\mathcal{B}}^{(A^*)}\}$ 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 $\{x_{\mathcal{B}}^{(2)}, \ldots, x_{\mathcal{B}}^{(A^*)}\}$ also satisfy the first state in $\mathcal{B}$ by (A.1) and (A.2).

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

$$
\mathcal{K}_{\mathcal{B}}^{(b)} = \sum_{a=1}^{A} f_{J\tau}(x_{\mathcal{B}}^{(b)} | x_{\mathcal{A}}^{(a)}) \mathcal{K}_{\mathcal{A}}^{(a)} + \sum_{a=1}^{A} \sum_{j=1}^{J-1} [f_{j\tau}(x_{\mathcal{B}}^{(b)} | x_{\mathcal{A}}^{(a)}) - f_{J\tau}(x_{\mathcal{B}}^{(b)} | x_{\mathcal{A}}^{(a)})] \omega_{j\tau}(x_{\mathcal{A}}^{(a)}) \mathcal{K}_{\mathcal{A}}^{(a)}. \tag{A.3}
$$

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 $A(J - 1) \times 1$ vector formed from the element-by-element product $\mathcal{K}_{\mathcal{A}}^{(a)} \omega_{j\tau}(x_{\mathcal{A}}^{(a)})$. Denote the $(B - 1) \times A(J - 1)$ concatenated matrix of transitions by

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

Defining $\mathcal{K}_{\mathcal{B}}^*$ as a $(B - 1) \times 1$ vector formed from all but the first element of $\mathcal{K}_{\mathcal{B}}$ satisfying (A.2) then (A.3) may be expressed in matrix notation as

$$
\mathcal{K}_{\mathcal{B}}^* = F_{J\tau}(\mathcal{A})' \mathcal{K}_{\mathcal{A}} + [F_\tau(\mathcal{A})' - F_{J\tau}(\mathcal{A})'][\mathcal{K}_{\mathcal{A}} \circ \omega_\tau(\mathcal{A})]. \tag{A.4}
$$

Appealing to Hadley (1961, pp. 168–169), for a given $\mathcal{K}_\mathcal{B}^*$, a solution to (A.4) in $[\mathcal{K}_\mathcal{A} \circ \omega_\tau^*(\mathcal{A})]$ exists if and only if the rank of $[F_\tau(\mathcal{A})' - F_{J\tau}(\mathcal{A})']$ equals the rank of the augmented matrix formed by adding the column $[\mathcal{K}_\mathcal{B}^* - F_{J\tau}(\mathcal{A})'\mathcal{K}_\mathcal{A}]$ to $[F_\tau(\mathcal{A})' - F_{J\tau}(\mathcal{A})']$. By construction, the augmented matrix either has the same rank as, or one plus the rank of $[F_\tau(\mathcal{A})' - F_{J\tau}(\mathcal{A})']$. Since determining the rank of a finite dimensional matrix requires only a finite number of operations, and there are only a finite number of steps, the theorem is proved. □

PROOF OF THEOREM 5. The proof is by construction. In this game, each player $n \in \{1, 2\}$ controls two states, namely the choices of the previous period "in" or "out," so from (5.9) a sufficient condition for two-period dependence is the existence of a solution to

$$
\mathsf{H}_{t+2}^{(\sim n)} \begin{bmatrix} \Omega_{2,t+1}^{(n)}(\mathcal{A}_{2,t+1}, 2) \circ \mathsf{K}_{2,t+1}(\mathcal{A}_{2,t+1}, 2) \\ \Omega_{2,t+1}^{(n)}(\mathcal{A}_{1,t+1}, 1) \circ \mathsf{K}_{1,t+1}(\mathcal{A}_{1,t+1}, 1) \end{bmatrix}
$$
$$
= \mathsf{P}_{t+2}^{(\sim n)}(\mathcal{A}_{t+2}^{(n)}) \begin{bmatrix} \mathsf{F}_{1,t+1}^{(n)}(\mathcal{A}_{1,t+1}) \\ -\mathsf{F}_{1,t+1}^{(n)}(\mathcal{A}_{2,t+1}) \end{bmatrix}' \begin{bmatrix} \mathsf{K}_{1,t+1}(\mathcal{A}_{1,t+1}) \\ \mathsf{K}_{2,t+1}(\mathcal{A}_{2,t+1}) \end{bmatrix}, \tag{A.5}
$$

where the definitions of $\mathsf{H}_\tau^{(\sim 2)}$, given in (5.6), $\mathsf{K}_{j,t+1}(\mathcal{A}_{2,t+1})$ and $\Omega_{2,t+1}^{(n)}(\mathcal{A}_{2,t+1}, j)$, given above (5.1) and $\mathsf{P}_{t+2}^{(\sim n)}(\mathcal{A}_{t+2})$, given above (5.6) specialize to[30]

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

$$
\Omega_{2,t+1}^{(n)}(\mathcal{A}_{2,t+1}, j) = \left[\underline{\omega}_{t+1}^{(n)}(j, 2), \underline{\omega}_{t+1}^{(n)}(j, 1) \right]',
$$

$$
\mathsf{K}_{2,t+1}(\mathcal{A}_{2,t+1}) = \mathsf{K}_{1,t+1}(\mathcal{A}_{2,t+1}) = \left[p_{2t}^{(\sim n)}(x_t) \quad p_{1t}^{(\sim n)}(x_t) \right]',
$$

$$
\mathsf{P}_{t+2}^{(\sim n)}(\mathcal{A}_{t+2}) = \left[p_{2,t+2}^{(\sim n)}(2, 2) \quad p_{2,t+2}^{(\sim n)}(2, 1) \quad p_{2,t+2}^{(\sim n)}(1, 2) \quad p_{2,t+2}^{(\sim n)}(1, 1) \right]
$$

and in this example:

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

---

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

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

$$
= \begin{bmatrix} p_{2,t+1}^{(\sim n)}(2,2) & p_{2,t+1}^{(\sim n)}(2,1) & -p_{2,t+1}^{(\sim n)}(1,2) & -p_{2,t+1}^{(\sim n)}(1,1) \\ p_{1,t+1}^{(\sim n)}(2,2) & p_{1,t+1}^{(\sim n)}(2,1) & -p_{1,t+1}^{(\sim n)}(1,2) & -p_{1,t+1}^{(\sim n)}(1,1) \\ -p_{2,t+1}^{(\sim n)}(2,2) & -p_{2,t+1}^{(\sim n)}(2,1) & p_{2,t+1}^{(\sim n)}(1,2) & p_{2,t+1}^{(\sim n)}(1,1) \\ -p_{1,t+1}^{(\sim n)}(2,2) & -p_{1,t+1}^{(\sim n)}(2,1) & p_{1,t+1}^{(\sim n)}(1,2) & p_{1,t+1}^{(\sim n)}(1,1) \end{bmatrix}. \tag{A.6}
$$

Noting $\omega_{t+1}^{(n)}(x_{t+1}, j) \equiv \omega_{t+1}^{(n)}((j, d_{2t}^{(\sim n)}), j)$ we now define $\underline{\omega}_{t+1}^{(n)}(x_{t+1}) \equiv \omega_{t+1}^{(n)}(x_{t+1}, j)$ to eliminate the notational redundancy, and substitute the expressions above into the left-hand side of (A.5) to obtain

$$
\begin{bmatrix} p_{2,t+2}^{(\sim n)}(2,2) \\ p_{2,t+2}^{(\sim n)}(2,1) \\ p_{2,t+2}^{(\sim n)}(1,2) \\ p_{2,t+2}^{(\sim n)}(1,1) \end{bmatrix}' \begin{bmatrix} p_{2,t+1}^{(\sim n)}(2,2) & p_{2,t+1}^{(\sim n)}(2,1) & -p_{2,t+1}^{(\sim n)}(1,2) & -p_{2,t+1}^{(\sim n)}(1,1) \\ p_{1,t+1}^{(\sim n)}(2,2) & p_{1,t+1}^{(\sim n)}(2,1) & -p_{1,t+1}^{(\sim n)}(1,2) & -p_{1,t+1}^{(\sim n)}(1,1) \\ -p_{2,t+1}^{(\sim n)}(2,2) & -p_{2,t+1}^{(\sim n)}(2,1) & p_{2,t+1}^{(\sim n)}(1,2) & p_{2,t+1}^{(\sim n)}(1,1) \\ -p_{1,t+1}^{(\sim n)}(2,2) & -p_{1,t+1}^{(\sim n)}(2,1) & p_{1,t+1}^{(\sim n)}(1,2) & p_{1,t+1}^{(\sim n)}(1,1) \end{bmatrix}
$$

$$
\times \begin{bmatrix} \underline{\omega}_{t+1}^{(n)}(2,2) p_{2t}^{(\sim n)}(x_t) \\ \underline{\omega}_{t+1}^{(n)}(2,1) p_{2t}^{(\sim n)}(x_t) \\ \underline{\omega}_{t+1}^{(n)}(1,2) p_{2t}^{(\sim n)}(x_t) \\ \underline{\omega}_{t+1}^{(n)}(1,1) p_{2t}^{(\sim n)}(x_t) \end{bmatrix}. \tag{A.7}
$$

Since $p_{2t}^{(\sim n)}(x_t) > 0$, we can establish two-period dependence by equating (A.7) with the right-hand side of (A.5) and solving for the unknowns. By inspection, (A.7) is 1×1, and (A.5) reduces to a single equation, with four unknowns that conform to the 1×4 row vector $\mathsf{H}_{t+2}^{(\sim n)}$.

To complete the proof, it is useful to define for $i \in \{1, 2\}$ the expression:

$$
C_i \equiv p_{2,t+2}^{(\sim n)}(2,1) - p_{2,t+2}^{(\sim n)}(1,1)
$$
$$
+ p_{2,t+1}^{(\sim n)}(2,i) \big[p_{2,t+2}^{(\sim n)}(2,2) + p_{2,t+2}^{(\sim n)}(1,1) - p_{2,t+2}^{(\sim n)}(2,1) - p_{2,t+2}^{(\sim n)}(1,2) \big]. \tag{A.8}
$$

We now prove $C_2 \neq 0$ if $C_1 = 0$. Note that

$$
C_2 - C_1 = \big[p_{2,t+1}^{(\sim n)}(2,2) - p_{2,t+1}^{(\sim n)}(2,1) \big]
$$
$$
\times \big[p_{2,t+2}^{(\sim n)}(2,2) + p_{2,t+2}^{(\sim n)}(1,1) - p_{2,t+2}^{(\sim n)}(2,1) - p_{2,t+2}^{(\sim n)}(1,2) \big]. \tag{A.9}
$$

If the second bracketed term is zero, then $C_1 = C_2$ from (A.9), and hence from (A.8) $C_1 \neq 0$ because by assumption $p_{2,t+2}^{(\sim n)}(2,1) \neq p_{2,t+2}^{(\sim n)}(1,1)$. Therefore, if $C_1 = 0$ the bracketed term is nonzero. In that case, $C_2 \neq C_1$ by (A.9) because $p_{2,t+1}^{(\sim n)}(2,1) \neq p_{2,t+1}^{(\sim n)}(2,2)$ by assumption.

We consider two possibilities, in which $\omega_{t+2}^{(n)}(x_{t+2}^{(n)}, j) = 1$ for $j \in \{1, 2\}$ and $\underline{\omega}_{t+1}^{(n)}(1, i) = 0$ for $i \in \{1, 2\}$ for both possibilities. Also set $\omega_{t+1}^{(n)}(2, 2) = 0$ if $C_1 = 0$, and set $\omega_{t+1}^{(n)}(2, 1) =$

0 if $C_1 \neq 0$. Using (A.8) and noting $p_{1,t+1}^{(\sim n)}(2,2) = 1 - p_{2,t+1}^{(\sim n)}(2,2)$, simplify (A.7) to $C_i p_{2t}^{(\sim n)}(x_t) \omega_{t+1}^{(n)}(2,i)$. Solving for the only nonzero weight, take the quotient of the scalar (A.7) and $C_i p_{2t}^{(\sim n)}(x_t)$ to obtain

$$\underline{\omega}_{t+1}^{(n)}(2,i) = P_{t+2}^{(\sim n)}\big(\mathcal{A}_{t+2}^{(n)}\big) \begin{bmatrix} F_{1,t+1}^{(n)}(\mathcal{A}_{1,t+1}) \\ -F_{1,t+1}^{(n)}(\mathcal{A}_{2,t+1}) \end{bmatrix}' \begin{bmatrix} K_{1,t+1}^{(n)}(\mathcal{A}_{1,t+1}) \\ K_{2,t+1}^{(n)}(\mathcal{A}_{2,t+1}) \end{bmatrix} \circ \Big/ \big[p_{2t}^{(\sim n)}(x_t) C_i \big], \quad \text{(A.10)}$$

where the matrices in (A.10) are given above. Thus $\underline{\omega}_{t+1}^{(n)}(2,1)$ is determined by setting $i=1$ in (A.10) when $C_1 \neq 0$ and $\underline{\omega}_{t+1}^{(n)}(2,2)$ is determined by setting $i=2$ in (A.10) when $C_1 = 0$. Two-period dependence can now be established by direct verification. □

References

Aguirregabiria, V. and A. Magesan (2013), "Euler equations for the estimation of dynamic discrete choice structural models." *Advances in Econometrics*, 31, 3–44. [854, 861]

Aguirregabiria, V. and A. Magesan (2017), "Solution and estimation of dynamic discrete choice structural models using Euler equations." Working paper. [854, 861]

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. [853, 856]

Aguirregabiria, V. and P. Mira (2007), "Sequential estimation of dynamic discrete games." *Econometrica*, 75, 1–54. [853, 864]

Altug, S. and R. Miller (1998), "The effect of work experience on female wages and labour supply." *Review of Economic Studies*, 62, 45–85. [854, 856, 861]

Arcidiacono, P., E. Aucejo, A. Maurel, and T. Ransom (2016), "College attrition and the dynamics of information revelation." Working paper. [854]

Arcidiacono, P., P. Bayer, J. Blevins, and P. Ellickson (2016), "Estimation of dynamic discrete choice models in continuous time with an application to retail competition." *Review of Economic Studies*, 83, 889–931. [854]

Arcidiacono, P. and P. Ellickson (2011), "Practical methods for estimation of dynamic discrete choice models." *Annual Review of Economics*, 3, 363–394. [853]

Arcidiacono, P. and R. Miller (2011), "Conditional choice probability estimation of dynamic discrete choice model with unobserved heterogeneity." *Econometrica*, 79, 1823–1867. [854, 856, 857, 858, 859, 884]

Arcidiacono, P. and R. Miller (2019), "Supplement to 'Nonstationary dynamic models with finite dependence'." *Quantitative Economics Supplemental Material*, 10, https://doi.org/10.3982/QE626. [880]

Arcidiacono, P. and R. Miller (forthcoming), "Identifying dynamic discrete choice models off short panels." *Journal of Econometrics*. [862]

Bajari, P., L. Benkard, and J. Levin (2007), "Estimating dynamic models of imperfect competition." *Econometrica*, 75, 1331–1371. [853]

Beauchamp, A. (2015), "Regulation, imperfect competition, and the U.S. abortion market." *International Economic Review*, 56, 963–996. [854, 871]

Bishop, K. (2012), "A dynamic model of location choice and hedonic valuation." Working paper. [854]

Coate, P. (2016), "Parental influence on labor market outcomes and location decisions of young workers." Working paper. [854]

Collard-Wexler, A. (2013), "Demand fluctuations in the ready-mix concrete industry." *Econometrica*, 81, 1003–1037. [871]

Declerq, K. and F. Verboven (2018), "Enrollment and degree completion in higher education without ex ante admission standards." *Economics of Education Review*, 66, 223–244. [854]

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. [871]

Ericson, R. and A. Pakes (1995), "Markov-perfect industry dynamics: A framework for empirical work." *Review of Economic Studies*, 62, 53–82. [859]

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. [854]

Gayle, G., A. Hincapie, and R. Miller (2018), "Life-cycle fertility and human capital accumulation." Working paper. [854]

Gayle, W. (2017), "CCP estimation of dynamic discrete/continuous choice models with generalized finite dependence and correlated unobserved heterogeneity." Working paper. [854, 858, 861]

Hadley, G. (1961), *Linear Algebra*. Addison-Wesley. [869, 886]

Hotz, V. J. and R. Miller (1993), "Conditional choice probabilities and estimation of dynamic models." *Review of Economic Studies*, 60, 497–529. [853, 854, 864]

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. [853]

James, J. (2014), "Learning and occupational sorting." Working paper. [854]

Joensen, J. (2009), "Academic and labor market success: The impact of student employment, abilities, and preferences." Working paper. [854]

Khorunzhina, N. (2013), "Structural estimation of stock market participation costs." *Journal of Economic Dynamics & Control*, 37, 2928–2942. [854]

Khorunzhina, N. and R. Miller (2016), "American dream delayed: Shifting determinants of homeownership in the U.S." Working paper. [854]

Ma, L. (forthcoming), "Learning in a hedonic framework: Valuing Brownfield remediation." *International Economic Review*. [854]

Matsumoto, B. (2014), "Lighting the fires: Explaining youth smoking initiation and experimentation in the context of a rational addiction model with learning." Working paper. [854]

Mazur, J. (2017), "Can stricter bankruptcy laws discipline capital investment?" Working paper. [854, 871]

Miller, R. (1984), "Job matching and occupational choice." *Journal of Policitical Economy*, 92, 1086–1120. [864]

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. [864]

Pesendorfer, M. and P. Schmidt-Dengler (2008), "Asymptotic least square estimators for dynamic games." *Review of Economic Studies*, 75, 901–908. [853]

Ransom, T. (2018), "Labor market frictions and moving costs of the employed and unemployed." Working paper. [854]

Rust, J. (1987), "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher." *Econometrica*, 55, 999–1033. [864]

Ryan, S. (2012), "The costs of environmental regulation in a concentrated industry." *Econometrica*, 80, 1019–1062. [871]

Scott, P. (2013), "Dynamic discrete choice estimation of agricultural land use." Working paper. [854]