Identifying Dynamic Discrete Choice Models off Short Panels*

Peter Arcidiacono   Robert A. Miller

Duke University & NBER   Carnegie Mellon University

October 19, 2015

Abstract

This paper analyzes identifying flow payoffs and counterfactual choice probabilities (CCPs) in single-agent dynamic discrete choice models under three distinct scenarios: stationary; nonstationary where data is sampled every period respondents make decisions; and nonstationary where the time horizon for respondents extends beyond the length of the data (short panels). Our most striking results apply to temporary counterfactual CCPs in nonstationary short panels. They are identified when induced by policy changes affecting payoffs, even though the utility flows are not; counterfactual CCPs induced by innovations to state transitions are generally not identified unless the model has terminal or renewal states.

1 Introduction

Dynamic discrete choice models are increasingly used to explain panel data in labor economics, industrial organization and marketing.[1] It is widely recognized that the interpreting the predictions of policy innovations from structural models critically depend on the assumptions used to identify the model. The central

*Corresponding autthor: Robert Miller, Tepper Business School, Carnegie Mellon University, Pittsburgh, Pa, 15217; phone, 412-268-3701;email, ramiller@cmu.edu

[1] For surveys of this literature see Eckstein and Wolpin (1989), Rust (1994), Pakes (1994), Miller (1997), Aguirregabiria and Mira (2010) and Arcidiacono and Ellickson (2011).

role identification plays in determining the value of estimating structural models has stimulated a small but growing literature on the identification of dynamic discrete choice models in single-agent settings.[2]

Research in this area dates back to Rust (1994), who showed that solutions to stationary infinite horizon dynamic discrete choice models are invariant to a broad class of utility transformations. Magnac and Thesmar (2002) later established that the flow payoffs for a two period model are identified in discrete choice optimization problems when the econometrician knows the joint probability distribution of the choice specific idiosyncratic disturbances and the discount factor, subject to a normalization on the flow payoffs in each of the periods. Norets and Tang (2014) provide conditions for identifying the probability distribution of the choice specific disturbance in stationary binary choice environments in the presence of exclusion restrictions, whereby a set of variables affects the transitions of the states but not the utility flows themselves.[3]

This research has focused on cases where the model is either stationary or where the data covers the full time horizon. Yet many data sets are short panels: they do not cover the full lifetime of the sampled firms, individuals, or products, and the sample respondents are often subjected to aggregate shocks that cannot be averaged out in the cross section. These features pose serious challenges to inference. Conventional wisdom holds that accommodating nonstationarities within dynamic structures complicates inference, explaining why most applied work in this area assumes the data generating process is stationary, or impose other strong restrictions on the aggregate processes. But nonstationarity and aggregate shocks arise naturally: in the human life cycle through aging, business cycles, and the general equilibrium effects of evolving demographics; in industries because of innovation and growth within and external to the market under consideration; and in marketing through the diffusion of new products and more generally over the product life cycle. This paper extends previous work by deriving new conditions on identification

[2] A literature on identification of multi-agent models has also emerged. See, for example, Aguirregabiria and Mira (2015), Aguirregabiria and Suzuki (2014), Bajari, Chernozhukov, Hong, and Nekepelov (2009), and Pesendorfer and Schmidt-Dengler (2008).

[3] Most work in this area, including ours, focuses on the case where all the unobserved variables are independently distributed over time, but Kasahara and Shimotsu (2009) and Hu and Shum (2012) relax this assumption in their analyses of identification. As our results show below, identification in such settings can only be achieved through functional form and exclusion restrictions on observed covariates. Pantano and Xheng (2010) examine identification in the presence of unobserved heterogeneity when subjective expectations data is available.

for dynamic discrete choice models of individual optimization problems, applicable to both stationary and nonstationary settings. In the latter case we distinguish between panels that are short versus panels that are, for want of a better descriptor, long.

Our first set of results builds on Arcidiacono and Miller (2011) which provides a representation of the value function as a mapping of future streams of conditional choice probabilities and flow payoffs associated with any sequence of future choices. Under standard assumptions that the distribution of the choice specific idiosyncratic disturbances and subjective discount factor are known, we show that the flow payoffs are identified up to any normalization on the flow payoff of one of the choices in each state and time period.

A noteworthy feature of our results is establishing links between different observationally equivalent normalizations as well as showing their intertemporal linkages. In long panels the normalization can be made on a flow payoff primitive. However, in short panels, when the time horizon extends beyond the length of the data, this is not possible: terms involving the continuation value must be normalized too, because behavior observed during the sample is not solely attributable to payoffs that occur during the sample but partially reflect decision making and payoffs that occur after the sample ends.[4]

Our second set of results is concerned with the identification of counterfactual conditional choice probabilities. Aguirregabiria (2005, 2010) shows in different settings identification of counterfactual choice probabilities given innovations in utility differences. These findings are echoed in Norets and Tang (2014) in the stationary case of a binary choice model, but they show that for other policy changes normalizing the flow payoff for one of the choices to zero in all states is not innocuous for analyzing counterfactual policy changes. While counterfactual policies that affect the flow payoffs result in the same counterfactual choice probabilities under different normalizations, this is not true when the counterfactual policies affect the state transitions. The reason is that, as we noted above, the state transitions are embedded in the link between one normalization and another. We extend these results with two contributions. First we show that in

[4] The true normalization can be recovered with information from outside sources, or by imposing exclusion restrictions that reduce the size of the parameter space, or functional form restrictions on the flow payoffs. A common exclusion restriction is to include a state variable that affects the state transitions but not the flow payoffs. A common functional form assumption is to restrict flow payoffs to depend on time only through the state variables themselves, thus limiting the channels of nonstationarity to state transitions and the effects of the time horizon.

the short panel case the counterfactual choice probabilities for temporary policy changes are identified if the policy change only affects the flow payoffs. Second, we show that, in general, counterfactual choice probabilities for temporary policy changes affecting the state transitions cannot be identified off short panels, even if all the true normalizations are known for the entire history. Sufficient for identification are that the model has an extended terminal state or renewal property, and the true normalization is known for the terminating or renewal actions, typically a much stronger assumption than finite dependence.[5]

The remainder of the paper proceeds as follows. The next section lays out the decision framework. Section 3 analyzes identification of flow payoffs. Section 4 turns to the identification of conditional choice probabilities under counterfactual regimes. Section 5 concludes.

2 Framework

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

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

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

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

[5] Otherwise, additional exclusion and functional form restrictions are necessary to identify counterfactual conditional choice probabilities when the policy innovation affects the state transitions.

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

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

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

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

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

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:[6]

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

3 Two Theorems on Identification

The objects of identification in the optimization model are the utility flows, the discount factor, the transition matrix of the observed state variables, and the distribution of the unobserved variables,[7] summarized with the notation (u, β, F, G). In this section we build upon Rust (1994), Magnac and Thesmar (2002)

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

[7] Often the distribution of unobserved variables is assumed to be extreme value for tractability. However, Arcidiacono and Miller (2011) showed how generalized extreme value distributions can easily be accommodated within a CCP estimation framework, and recently Chiong, Galichon, and Shum (2013) have proposed simple estimators for a broad range of error distributions.

and Norets and Tang (2014) by considering identification when (β, F, G) are known.[8] In our analysis, let $\mathcal{T}$ denote the last date for which data is available (for a real or synthetic cohort). First we show that u is identified up to a normalization on the flow payoffs for one of the choices in each state when either the environment is stationary or when $\mathcal{T} = T$, that is where the data is long. Then we analyze short panels data sets, meaning $T > \mathcal{T}$.

Our discussion, and the proofs of the theorems, draw upon the representation of $v_{jt}(x_t)$ given in Theorem 1 of Arcidiacono and Miller (2011). It is based on their Lemma 1, that for every $t \in \{1, \ldots, T\}$ and $p \in \Delta^J$, the J dimensional simplex, there exists a real-valued function $\psi_j(p)$ such that:

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

To interpret (4), note that the value of committing to action j before seeing ϵ_t is $v_{jt}(x_t) + E[\epsilon_{jt}]$. Therefore the expected loss from pre-committing to j versus waiting until ϵ_t is observed and only then making an optimal choice, $V_t(x_t)$, is the constant $E[\epsilon_{jt}]$ plus $\psi_j[p_t(x_t)]$, a composite function that only depends x_t through the conditional choice probabilities. Then, for each choice $j \in \{1, \ldots, J\}$ and $\tau \in \{t, \ldots, T\}$, let any $\omega_\tau(x_\tau, j)$ denote any mapping from the state space $\{1, \ldots, X\}$ to R^J satisfying the constraints that $\sum_{k=1}^J \omega_{k\tau}(x_\tau, j) = 1$ and $|\omega_{k\tau}| < B$ for $B < \infty$. Recursively define $\kappa_\tau(x_{\tau+1}|x_t, j)$ as:

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

Then Arcidiacono and Miller (2011, 2015) show that for $\mathcal{T} < T$:

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

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

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

Representing the conditional value functions in this way facilitates the identification theorems discussed below.

[8] The assumption that (β, F, G) are known is standard. Typically F is identified from the transitions alone by assuming that all the state variables are observed, estimates of β, that calibrate a person's subjective discount factor in a stationary model are obtained from other data, and G is selected largely on the basis of tractability.

3.1 Normalizing utility flows

Rust (1994, Lemma 3.2 on page 3127) showed that the solution to a stationary infinite horizon discrete choice optimization problem is invariant to a broad class of utility transformations. His result can be simply extended to nonstationary optimization problems. Still unanswered is what specifications of the flow payoffs are observationally equivalent. This is more complicated than static discrete choice settings due to adjustments in future flows affecting what flow payoffs are observationally equivalent in earlier periods.

As a first step towards deriving observational equivalence and for future reference, let $\kappa_\tau^*(x_{\tau+1}|x_t, j)$ denote the probability distribution of $x_{\tau+1}$, given a state of x_t and taking action j at t, followed by repeatedly taking the first action from period $t+1$ through to period τ. Formally:

$$\kappa_\tau^*(x_{\tau+1}|x_t, j) \equiv \begin{cases} f_{jt}(x_{t+1}|x_t) & \text{for } \tau = t \\ \sum_{x_\tau=1}^{X} f_{1\tau}(x_{\tau+1}|x_\tau)\kappa_{\tau-1}^*(x_\tau|x_t, j) & \text{for } \tau = t+1, \ldots, T \end{cases} \tag{8}$$

We now show there is an observational equivalent dynamic optimization problem to (u, β, F, G), which we denote by (u^*, β, F, G), where for each (t, x) we arbitrarily select any one choice $l(x,t) \in \{1, \ldots, J\}$ and set the flow utility associated with that choice, $u_{l(x,t),t}^*(x)$, to an arbitrary real value we denote by $c_t(x)$. Similarly in the infinite horizon analogue we select for each x any one choice $l(x) \in \{1, \ldots, J\}$ and set the flow utility associated with that choice, $u_{l(x)}^*(x)$, to an arbitrary real value denoted by $c(x)$.

Theorem 1 *In finite horizon dynamic discrete choice optimization problems let $l(x,t) \in \{1, \ldots, J\}$ and $c_t(x) \in \Re$ respectively denote any arbitrarily defined normalizing action and benchmark flow utility the associated with (x,t), and define for all $j \in \{1, \ldots, J\}$:*

$$u_{jT}^*(x) \equiv u_{jT}(x) - u_{l(x,T),T}(x) + c_T(x) \tag{9}$$

and:

$$u_{jt}^*(x) = u_{jt}(x) + c_t(x) - u_{l(x,t),t}(x) + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \left[u_{1\tau}^*(x_\tau) - u_{1\tau}(x_\tau)\right] \left[\kappa_{\tau-1}^*(x_\tau|x_t, l(x,t)) - \kappa_{\tau-1}^*(x_\tau|x_t, j)\right] \tag{10}$$

When the environment is stationary, define:

$$u_j \equiv \begin{bmatrix} u_j(1) \\ \vdots \\ u_j(X) \end{bmatrix}, \quad u_j^* \equiv \begin{bmatrix} u_j^*(1) \\ \vdots \\ u_j^*(X) \end{bmatrix}, \quad \tilde{u} \equiv \begin{bmatrix} u_{l(1)}(1) \\ \vdots \\ u_{l(X)}(X) \end{bmatrix}, \quad F_j \equiv \begin{bmatrix} f_j(1|1) & \dots & f_j(X|1) \\ \vdots & \ddots & \vdots \\ f_j(1|X) & \dots & f_j(X|X) \end{bmatrix}, \quad c \equiv \begin{bmatrix} c(1) \\ \vdots \\ c(X) \end{bmatrix}$$

Then $[\mathcal{I} - \beta \mathcal{F}_1]$ *is invertible. Define for all* $j \in \{1, \dots, J\}$*:*

$$u_j^* = u_j + c - \tilde{u} + \beta (F_1 - F_j) [\mathcal{I} - \beta F_1]^{-1} (u_1^* - u_1)$$

Then (u^*, β, F, G) *is observationally equivalent to* (u, β, F, G).

A common normalization in empirical work is to set $u_{1t}^*(x) = 0$ for all (t, x) in the finite horizon case and $u_1^*(x) = 0$ for all x in the stationary case. Theorem 2 demonstrates that a normalization like that is necessary to identify the remaining utility parameters. The next section provides conditions under which it is sufficient.

3.2 Sampling from the whole population

Magnac and Thesmar (2002, Theorem 2 and Corollary 3 on pages 807 and 808) established identification of the flow payoff for $T = 2$ finite when $G(\epsilon_t)$ and β are known, $u_1(x)$ is normalized for all x, and the continuation value for one of the actions is also normalized. We extend their results to the case where data on the full time horizon is observed. Norets and Tang (2014) extend their result to stationary environments where there are just two choices. We extend their result to the case where there are any finite number of choices and nonstationary settings.

Theorem 1 implies that without loss of generality we can normalize $u_{1t}(x) = 0$ for all (t, x) in the finite horizon case, or $u_1(x) = 0$ for all x in the stationary case and set $d_{1\tau}^*(x_\tau) = 1$ for all τ. Equation (7) then simplifies to:

$$v_{jt}(x_t) = u_{jt}(x_t) + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \psi_1[p_\tau(x_\tau)] \kappa_{\tau-1}^*(x_\tau | x_t, j)$$

Subtracting $v_{1t}(x_t)$ from $v_{jt}(x_t)$ yields:

$$v_{jt}(x_t) - v_{1t}(x_t) = u_{jt}(x_t) + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \psi_1[p_\tau(x_\tau)] \left[\kappa_{\tau-1}^*(x_\tau | x_t, j) - \kappa_{\tau-1}^*(x_\tau | x_t, 1) \right] \tag{11}$$

An alternative expression for this difference can be obtained by differencing the expressions for $\psi_1(x_t)$ and $\psi_t(x_t)$ given in Equation (4):

$$v_{jt}(x_t) - v_{1t}(x_t) = \psi_1[p_t(x_t)] - \psi_j[p_t(x_t)] \tag{12}$$

Theorem 2 below uses the two expressions for $v_{jt}(x_t) - v_{1t}(x_t)$ to form expressions for $u_{jt}(x_t)$ as a function of the transition probabilities, the conditional choice probabilities, and the discount factor. Further, Theorem 2 shows how the problem simplifies in the stationary case where the time subscripts are dropped from the flow payoffs and the state transition probabilities.

Theorem 2 *For all j, t, and x_t, the flow payoff $u_{jt}(x_t)$ can be expressed as:*

$$u_{jt}(x_t) = \psi_1[p_t(x_t)] - \psi_j[p_t(x_t)] + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \psi_1[p_\tau(x_\tau)] \left[\kappa^*_{\tau-1}(x_\tau|x_t, 1) - \kappa^*_{\tau-1}(x_\tau|x_t, j)\right] \tag{13}$$

When the environment is stationary, let $\mathcal{I}$ denote the X dimensional identity matrix and define:

$$u_j \equiv \begin{bmatrix} u_j(1) \\ \vdots \\ u_j(X) \end{bmatrix}, \quad F_j \equiv \begin{bmatrix} f_j(1|1) & \dots & f_j(X|1) \\ \vdots & \ddots & \vdots \\ f_j(1|X) & \dots & f_j(X|X) \end{bmatrix}, \quad \Psi_j \equiv \begin{bmatrix} \psi_j[p(1)] \\ \vdots \\ \psi_j\,[p(X)] \end{bmatrix}$$

Then $[\mathcal{I} - \beta F_1]$ is invertible and for all j:

$$u_j = \Psi_j - \Psi_1 + \beta\left(F_1 - F_j\right)\left[\mathcal{I} - \beta F_1\right]^{-1} \ \Psi_1 \tag{14}$$

Given the assumptions made at the beginning of this section regarding the state transitions, conditional choice probabilities, the discount factor, and the distribution of the structural errors, everything on the right hand side of both (13) and (14) is known implying both systems are exactly identified. These equations therefore yield asymptotically efficient estimators of the unrestricted utility flows. They are defined by substituting sample analogues for the conditional choice probabilities into the mappings that represent the utility flows; they are efficient because the mapping of the conditional choice probabilities on to the current utility flows is the one to one, and the relative frequencies observed in the data are the maximum estimates of the conditional choice probabilities.

3.3 Short panels

We next consider cases where the sampling period, $\mathcal{T}$, falls short of the time horizon T. Since choices and state transitions are not observed after period $\mathcal{T}$, the corresponding conditional choice probabilities and state transition matrices are not identified beyond that period either. Rather than express $u_{jt}(x_t)$ relative to the (normalized) first choice for the full horizon as in (13), we express u_{jt} relative to the normalized choice until period $\mathcal{T}$ and then use the value function at $\mathcal{T}+1$. This yields an expression for $u_{jt}(x_t)$ that provides the basis for the following corollary giving the degree of underidentification.

Corollary 3 *If $u_{1t}(x_t) = 0$ for all t and x_t, then:*

$$\begin{aligned} u_{jt}(x_t) &= \psi_1[p_t(x_t)] - \psi_j[p_t(x_t)] + \sum_{\tau=t+1}^{\mathcal{T}} \sum_{x_\tau=1}^{X} \beta^{\tau-t}\psi_1[p_\tau(x_\tau)] \left[\kappa^*_{\tau-1}(x_\tau|x_t,1) - \kappa^*_{\tau-1}(x_\tau|x_t,j)\right] \\ &\quad + \sum_{x_{\mathcal{T}+1}=1}^{X-1} \beta^{\mathcal{T}-t} V_{\mathcal{T}+1}(x_{\mathcal{T}+1})] \left[\kappa^*_{\mathcal{T}}(x_{\mathcal{T}+1}|x_t,1) - \kappa^*_{\mathcal{T}}(x_{\mathcal{T}+1}|x_t,j)\right] \end{aligned} \tag{15}$$

Given β and $G(\epsilon)$, the degree of underidentification for the first $\mathcal{T}$ flow payoffs is at most $X-1$.

The last term in Equation (15) gives the underidentification result. Since the choice probabilities and state transition matrices are identified from the data up to $\mathcal{T}$, and $u_{jt}(x_t)$ is a linear mapping of $V_{\mathcal{T}+1}(x_{\mathcal{T}+1})$, the utility flows would be exactly identified if $V_{\mathcal{T}+1}(x_{\mathcal{T}+1})$ was known. The corollary shows that the degree of under-identification is less than or equal to the number of states in the state space.

Following Magnac and Thesmar (2002), we could normalize the value functions in the last period to zero. At that point we would treat the sample as if the time horizon was $\mathcal{T}$ rather than T. The difficulty with such a normalization is that the primitives which justify it are unknown. For example, we could set the payoffs to action 1 in periods $\mathcal{T}+2$ to T to zero. In this case $V_{\mathcal{T}+1}$ can be expressed as:

$$V_{\mathcal{T}+1}(x_{\mathcal{T}+1}) = u_{1\mathcal{T}+1}(x_{\mathcal{T}+1}) + \psi_1\left[p_{\mathcal{T}+1}(x_{\mathcal{T}+1})\right] + \sum_{\tau=\mathcal{T}+1}^{T} \sum_{x_{\tau=1}}^{X} \beta^{\tau-\mathcal{T}}\psi_1\left[p_\tau(x_\tau)\right] \left[\kappa^*_{\tau-1}(x_\tau|x_t,1) - \kappa^*_{\tau-1}(x_\tau|x_t,j)\right] \tag{16}$$

The normalization of $V_{\mathcal{T}+1}(x_{\mathcal{T}+1}) = 0$ is then achieved by normalizing $u_{1\mathcal{T}+1}(x_{\mathcal{T}+1})$ such that:

$$u_{1\mathcal{T}+1}(x_{\mathcal{T}+1}) = -\psi_1\left[p_{\mathcal{T}+1}(x_{\mathcal{T}+1})\right] - \sum_{\tau=\mathcal{T}+2}^{T} \sum_{x_{\tau=1}}^{X} \beta^{\tau-\mathcal{T}-1}\psi_1\left[p_\tau(x_\tau)\right] \left[\kappa^*_{\tau-1}(x_\tau|x_t,1) - \kappa^*_{\tau-1}(x_\tau|x_t,j)\right] \tag{17}$$

But this normalization on the primitives depends on state transitions and conditional choice probabilities that lie beyond the sample period, implying linking this normalization to alternative normalizations (such as $u_{1t}(x_t) = 0$ for all periods and for all x_t) is not possible.

4 Identifying the Effects of Policy Innovations

One of the main rationales for estimating structural models is their policy invariance; they yield robust predictions about the effects of changes in the primitives on equilibrium in different regimes. Aguirregaberia (2005) proved that in stationary infinite horizon models the CCPs for a counterfactual policy regimes involving only payoff innovations are identified. However Norets and Tang (2014) showed that the normalizations on the flow payoffs in a stationary two-choice setting are innocuous for some policy changes but not others. In their model, the CCPs for counterfactual policies that adjust the relative payoff a particular action-state combination are identified without knowing the underlying normalization, but that counterfactual predictions for policies changing the state variable transitions are not. This section investigates necessary and sufficient conditions for generalizing their results to choice sets larger than two, nonstationary settings, short panels, or temporary versus permanent policy innovations. A critical distinction we make throughout is whether the true normalization is known or not. We first consider payoff innovations, and then transition innovations. A fourth subsection illustrates how restrictions on payoffs can restore identification of CCPs for counterfactual innovations to state transitions.

To conduct the analysis we denote the true payoffs in the sampled regime by $u_{jt}(x)$, the true payoffs in the counterfactual regime by $\widetilde{u}_{jt}(x)$, and a payoff innovation by $\Delta_{jt}(x) \equiv \widetilde{u}_{jt}(x) - u_{jt}(x)$. We let $u^*_{jt}(x)$ denote any normalization that is observationally equivalent to $u_{jt}(x)$. Similarly, transition innovations are denoted by $\Lambda_{jt}(x'|x) \equiv \widetilde{f}_{jt}(x'|x) - f_{jt}(x'|x)$, where $f_{jt}(x'|x)$ is the observed transition for the sampled regime. Since $f_{jt}(x'|x)$ and $\widetilde{f}_{jt}(x'|x)$ are both probability transitions, it immediately follows that $-f_{jt}(x'|x) \leq \Lambda_{jt}(x'|x) \leq 1$ for all (j,t,x) and $\sum_{x'} \Lambda_{jt}(x'|x) = 0$ for all (t,x). Finally we let $p_t(x)$ denote the CCPs for the sampled regime and $\tilde{p}_{jt}(x)$ denote the CCPs for the counterfactual regime. Results for identifying $\tilde{p}_{jt}(x)$ are based on:

$$\tilde{p}_{jt}(x) = \int \prod_{k=1}^{J} 1 \left\{ \begin{array}{l} \epsilon_{jt} - \epsilon_{kt} + \Delta_{jt}(x) - \Delta_{kt}(x) + u_{jt}(x) - u_{kt}(x) \\ + \sum_{x'=1}^{X} \beta \widetilde{V}_{t+1}(x') \left[f_{jt}(x'|x) - f_{kt}(x'|x) + \Lambda_{jt}(x'|x) - \Lambda_{kt}(x'|x) \right] \end{array} \right\} g(\epsilon_t)\, d\epsilon_t \quad (18)$$

which is a direct implication of Equations (2) and (3), and the recursion:

$$\widetilde{V}_{t+1}(x') = \psi_1[\widetilde{p}_{t+1}(x')] + u_{1,t+1}(x') + \Delta_{1,t+1}(x') + \sum_{x''=1}^{X} \beta \widetilde{V}_{t+2}(x'') \left[f_{1,t+1}(x''|x') + \Lambda_{1,t+1}(x''|x') \right] \quad (19)$$

which exploits equation (4) for the first choice. From these two equations it is obvious that the counterfactual conditional choice probabilities, $\tilde{p}_{jt}(x)$, are not identified off short panels if the innovations go beyond $\mathcal{T}$ without imposing further restrictions on the payoffs and state transitions that occur after the sample ends. We therefore limit our analysis to temporary policy innovations that expire at $\mathcal{T}$ or earlier.

4.1 Counterfactual policies that affect payoffs

A starting point for investigating payoff innovations, which take the form $\Delta_{jt}(x) \equiv \widetilde{u}_{jt}(x) - u_{jt}(x)$, is to consider the link between the true payoffs and any normalization in the counterfactual regime, that is $\widetilde{u}_{jt}(x)$ and $\widetilde{u}^*_{jt}(x)$. From Equation (10) in Theorem 1:

$$\widetilde{u}^*_{jt}(x) = \widetilde{u}_{jt}(x) + c_t(x) - \widetilde{u}_{l(x,t),t}(x) + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \left[\widetilde{u}^*_{1\tau}(x_\tau) - \widetilde{u}_{1\tau}(x_\tau)\right] \left[\kappa^*_{\tau-1}(x_\tau|x_t, l(x,t)) - \kappa^*_{\tau-1}(x_\tau|x_t, j)\right] \tag{20}$$

Noting that $\widetilde{u}_{jt}(x) \equiv u_{jt}(x) + \Delta_{jt}(x)$ we difference (20) and (10), in other words the difference of (10) before and after the payoff innovation, to obtain:

$$\begin{aligned}
\widetilde{u}^*_{jt}(x) - u^*_{jt}(x) &= \Delta_{jt}(x) - \Delta_{l(x,t),t}(x) \\
&\quad + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \left[\widetilde{u}^*_{1\tau}(x_\tau) - u^*_{1\tau}(x_\tau) - \Delta_{1\tau}(x_\tau)\right] \left[\kappa^*_{\tau-1}(x_\tau|x_t, l(x,t)) - \kappa^*_{\tau-1}(x_\tau|x_t, j)\right] \\
&= \Delta_{jt}(x) - \Delta_{l(x,t),t}(x) \\
&\quad + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \Delta_{l(x,\tau),\tau}(x) \left[\kappa^*_{\tau-1}(x_\tau|x_t, l(x,t)) - \kappa^*_{\tau-1}(x_\tau|x_t, j)\right]
\end{aligned} \tag{21}$$

where the second line in the equation follows from setting $j = 1$ and applying a backwards induction argument. Equation (21) shows that for every observationally equivalent utility values that applies to the current regime, there is a corresponding set of utility values that are observationally equivalent to the true values in the counterfactual regime simply recovered by an adjustment involving the payoff innovations only. For example normalizing payoffs by the first choice in both regimes by setting $l(x,t) = 1$ and $c_t(x) = \Delta_{1t}(x) = 0$ for all (t,x), the equation reduces to the familiar form $\widetilde{u}^*_{jt}(x) = u^*_{jt}(x) + \Delta_{jt}(x)$ for all $j \neq 1$.

These remarks prompt the key result of this subsection. The CCP's for a counterfactual regime defined by temporary payoff innovations (that expire before the sample period ends) occurring within the sample frame can be computed from the CCPs for the current regime, that is without making any normalization and without directly estimating the utility parameters. Intuitively, a normalization of the utilities, $u^*_{jt}(x)$, can be computed as a function of the CCPs in the sample periods, using (15) and setting $V_{\mathcal{T}+1}(x_{\mathcal{T}+1}) = 0$ for example. Consequently the arguments of the previous paragraph imply that $\widetilde{u}^*_{jt}(x)$ is a mapping of the CCP's and the elements of the payoff innovation $\Delta_{jt}(x)$. Solving the backwards recursion optimization problem we thus obtain the CCPs for the counterfactual regime.

Theorem 4 *Given any temporary payoff innovation in which $\Delta_{jt}(x) = 0$ for all $t > \mathcal{T}$ then:*

$$\tilde{p}_{j\mathcal{T}}(x) = \int \prod_{k=1}^{J} 1\left\{\epsilon_{j\mathcal{T}} - \epsilon_{k\mathcal{T}} + \Delta_{j\mathcal{T}}(x) - \Delta_{k\mathcal{T}}(x) + \psi_k[p_{\mathcal{T}}(x)] - \psi_j[p_{\mathcal{T}}(x)]\right\} g(\epsilon_{\mathcal{T}})\, d\epsilon_{\mathcal{T}}$$

For all $t < \tau$ the CCPs for the counterfactual regime can be expressed as:

$$\tilde{p}_{jt}(x) = \int \prod_{k=1}^{J} 1\left\{\epsilon_{jt} - \epsilon_{kt} + \widetilde{v}_{jt}(x) - \widetilde{v}_{kt}(x)\right\} g(\epsilon_t)\, d\epsilon_t$$

where $\widetilde{v}_{k\mathcal{T}}(x') = v_{k\mathcal{T}}(x')$ and for $t < \mathcal{T}$ the difference $\widetilde{v}_{jt}(x) - \widetilde{v}_{kt}(x)$ is recursively defined by:

$$\begin{aligned}\widetilde{v}_{jt}(x) - \widetilde{v}_{kt}(x) \quad = \quad & \Delta_{jt}(x) - \Delta_{kt}(x) + \psi_k[p_t(x)] - \psi_j[p_t(x)] \\ & + \sum_{x'=1}^{X-1} \beta\left[\widetilde{v}_{1,t+1}(x') - v_{1,t+1}(x') + \psi_1[\widetilde{p}_{t+1}(x')] - \psi_1[p_{t+1}(x')]\right]\left[f_{jt}(x'|x) - f_{kt}(x'|x)\right]\end{aligned}$$

4.2 Counterfactual policies that affect state transitions

Recovering counterfactual conditional choice probabilities that result from changes in the state transitions is much more problematic. Setting $\Delta_{jt}(x) = 0$ for all (j, t, x) to focus on policy innovations arising from changes in transitions, and recursively substituting for $\widetilde{V}_{t+1}(x')$ we obtain from Equations (18) and (19):

$$\tilde{p}_{jt}(x) = \int \prod_{k=1}^{J} 1 \left\{ \begin{array}{l} \epsilon_{jt} - \epsilon_{kt} + u_{jt}(x) - u_{kt}(x) \\ + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t}\left[\psi_1[\widetilde{p}_\tau(x_\tau)] + u_{1,\tau}(x_\tau)\right]\left[\widetilde{\kappa}^*_{\tau-1}(x_\tau|x,k) - \widetilde{\kappa}^*_{\tau-1}(x_\tau|x_t,j)\right] \end{array} \right\} g(\epsilon_t)\, d\epsilon_t \tag{22}$$

where $\widetilde{\kappa}^*_{\tau-1}(x_\tau|x,k)$ and $\widetilde{\kappa}^*_{\tau-1}(x_\tau|x_t,j)$ are defined similarly to $\kappa^*_{\tau-1}(x_\tau|x,k)$ and $\kappa^*_{\tau-1}(x_\tau|x_t,j)$ from Equation (8) by replacing $f_{j,t+1}(x'|x)$ with $\widetilde{f}_{j,t+1}(x'|x)$ as appropriate. The presence of the $u_{1,\tau}(x_\tau)$ terms show that they cannot be derived without knowing the normalization, regardless of the sample length.

However the case of short samples is more dire, because knowing the normalization is generally not sufficient to identify the effects of a temporary innovation. That is, even if the policy expires within the sample period so $\psi_1[\widetilde{p}_\tau(x_\tau)] = \psi_1[p_\tau(x_\tau)]$ for $\tau > \mathcal{T}$, the weights placed on the different states in the last period will have changed, which implies that the weights on the $\psi_1[\widetilde{p}_\tau(x_\tau)]$ will have changed as well and these terms are not known outside the sample period.

4.3 A two-period, two-choice example

A simple example illustrates the importance of the normalization for counterfactual policies that affect the state transitions and how even knowing the normalization does not help when the panel is short. Consider a two period model, $T = 2$, of the decision to smoke, $d_{2t} = 1$, or not, $d_{1t} = 1$, where the relevant state variable is whether the individual is in good health, $x_t = 1$, or bad health, $x_t = 2$. Suppose all individuals begin in good health and remain in good health at $t = 2$ is they do not smoke at $t = 1$. Suppose further that the probability of transitioning from good health to bad at $t + 1$ is given by π should the individual smoke at t.

Suppose that the true normalization is that the flow payoff for not smoking is 0 when in good health and c in bad health regardless of the period. Suppose, however, that the econometrician adopts the observationally equivalent normalization that the flow payoff in all periods is 0 for not smoking regardless of the state of the individual's health. Under the two normalizations and given data on both time periods, $u_{21}(1)$ and $u^*_{21}(1)$ are given by:

$$\begin{aligned} u_{21}(1) &= \psi_1[p_1(1)] - \psi_2[p_1(1)] + \beta\pi\left(\psi_1[p_2(1)] - \psi_1[p_2(2)]\right) - \beta\pi c \\ u^*_{21}(1) &= \psi_1[p_1(1)] - \psi_2[p_1(1)] + \beta\pi\left(\psi_1[p_2(1)] - \psi_1[p_2(2)]\right) \end{aligned}$$

implying $u^*_{21}(1) = u_{21}(1) + \beta\pi c$.

Now consider a new regime where the probability of transitioning to bad health conditional on smoking is given by π'. Note that this does not change the probability of smoking in the last period conditional on the state. Denote the counterfactual probability of smoking in the first period under the correct normalization as $p'_{21}(1)$ and the corresponding probability under the alternative normalization as $p^*_{21}(1)$. We now show that these two probabilities are different when $\pi \neq \pi'$. Note that the counterfactual probabilities under each normalization solve:

$$\begin{aligned} \psi_1\left[p'_1(1)\right] - \psi_2\left[p'_1(1)\right] &= u_{21}(1) + \beta\pi'\left(\psi_1[p_2(2)] - \psi_1[p_2(1)]\right) + \beta\pi' c \\ \psi_1\left[p^*_1(1)\right] - \psi_2\left[p^*_1(1)\right] &= u^*_{21}(1) + \beta\pi'\left(\psi_1[p_2(2)] - \psi_1[p_2(1)]\right) \end{aligned}$$

Substituting in for the flow payoffs and rearranging terms yields:

$$\begin{aligned}\psi_1\left[p_1'(1)\right]-\psi_2\left[p_1'(1)\right] &= \psi_1\left[p_1(1)\right]-\psi_2\left[p_1(1)\right]+\beta(\pi'-\pi)\left(\psi_1[p_2(2)]-\psi_1[p_2(1)]\right)+\beta(\pi'-\pi)c \\ \psi_1\left[p_1^*(1)\right]-\psi_2\left[p_1^*(1)\right] &= \psi_1\left[p_1(1)\right]-\psi_2\left[p_1(1)\right]+\beta(\pi'-\pi)\left(\psi_1[p_2(2)]-\psi_1[p_2(1)]\right)\end{aligned}$$

The expressions on the right hand side are identical in the two equations except the last term in the first equation is missing from the second as the incorrect normalization has embedded in it state transitions; state transitions that have now changed under the counterfactual policy. Hence, the counterfactual choice probabilities must differ across the normalizations as well.

But now suppose the true normalization is known in both periods but where data is only available on the first period smoking decisions. It is not possible to recover the counterfactual choice probabilities in the new regime even when the new regime only changes the first period transitions on the state variables. Namely, the weights placed on the different states in the second period have changed but the conditional choice probabilities in the second period are unavailable implying we do not have the correct adjustment terms. This stands in contrast to the case where the (temporary) policy affects the flow payoff of smoking as in the case the weights on the second periods conditional on smoking or not remain unchanged in the new regime, effectively allowing the future value terms to difference out across the regimes which can then be used to recover the counterfactual choice probabilities.

4.4 Renewal and terminal choices

There are essentially two directions to pursue in seeking to restore the identification of counterfactual policy changes that involve innovations to the transition functions. First exclusion and functional form restrictions can be brought to bear on the problem to achieve identification. Second we now show that when there is a terminal or renewal choice, and the value of utility from taking the terminal or renewal choice is known, then the counterfactual probabilities are identified. 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.

Let the first choice denote the terminal or renewal choice. In such models, following any choice $j \in \{1, \dots, J\}$ with a terminal or renewal choice leads to same value of state variables after two periods. Thus for all $t < T$ and x_t the probability distribution of x_{t+2} conditional on x_t does not depend on the choice made in period t if the terminal or renewal choice is taken in period $t+1$:

$$\sum_{x_{t+1}=1}^{X} f_{1,t+1}(x_{t+2}|x_{t+1}) f_{jt}(x_{t+1}|x_t) = \sum_{x_{t+1}=1}^{X} f_{1,t+1}(x_{t+2}|x_{t+1}) f_{1t}(x_{t+1}|x_t) \tag{23}$$

Since $\kappa^*_{\mathcal{T}}(x_{\mathcal{T}+1}|x_t, 1) = \kappa^*_{\mathcal{T}}(x_{\mathcal{T}+1}|x_t, j)$, the $V_{\mathcal{T}+1}(x_{\mathcal{T}+1})$ term drops out of Equation (15). Since we are considering the case where the flow payoff of the terminal or renewal choice is known in all states identification of $u_{jt}(x_t)$ is restored. Note that the one-period-ahead choice probabilities are needed, implying the flow payoffs for regimes involving changes in the state transitions can only be recovered until $\mathcal{T} - 1$.

But note that, since the normalization was known, these are the correct flow payoffs. Note further that, for policies expiring at or before $\mathcal{T} - 1$, we have the conditional choice probabilities for the new regime at $\mathcal{T}$. Hence for policies that affect the state transitions at or before $\mathcal{T} - 1$ the combination of a terminal or renewal action coupled with knowing the flow payoff associated with the terminal or renewal action permits the identification of the conditional choice probabilities in a counterfactual regime where the state transitions differ until $\mathcal{T}$.

The stopping and renewal problems discussed above can be simply extended by assuming there is some action, say the first, that if repeatedly taken after period $t+1$ for ρ periods, removes the dependence of the state on the action at time t implying $\kappa_{t+\rho}(x_{t+\rho+1}|x_t, j) = \kappa_{t+\rho}(x_{t+\rho+1}|x_t, 1)$ and the flow payoffs of that action are known. For simplicity suppose the known flow payoff of the action was such that $u_{1t}(x_t) = 0$, Equation (15) simplifies to:

$$u_{jt}(x_t) = \psi_1[p_t(x_t)] - \psi_j[p_t(x_t)] + \sum_{\tau=t+1}^{t+\rho} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \psi_1[p_\tau(x_\tau)] \left[\kappa^*_{\tau-1}(x_\tau|x_t, 1) - \kappa^*_{\tau-1}(x_\tau|x_t, j)\right] \tag{24}$$

as after ρ periods the probabilities of being in each of the states are the same across the two paths.[9] In

[9] This form of finite dependence has been applied in several empirical studies. The dynamic labor supply models of Altug and Miller (1998) and Gayle and Golan (2011) embodies it in their assumption that hours worked more than a finite number of years ago does not add to human capital; in the model of fertility of Gayle and Miller (2006) the relevant assumption is that the age of older offspring is immaterial to their parents after reaching some finite threshold.

this case, conditional choice probabilities are needed ρ periods ahead, implying flow payoffs can only be recovered until $\mathcal{T} - \rho$. Similarly, counterfactual choice probabilities for regimes where the state transitions differ from the observed regime until period $\mathcal{T}$ can also be recovered.

5 Conclusion

This paper establishes conditions for identifying dynamic discrete choice models, both for long panels where the sample period covers the full time horizon or the model is stationary, and for short panels where the sample period is shorter than the time horizon of the individual sample respondents. For a known disturbance structure and discount factor, dynamic discrete choice models of individual optimization are identified from long panels up to any normalization on one choice-specific flow payoff for each period in each state when the model is non-stationary.

Whether flow payoffs are identified relative to a normalization is often not relevant for policy innovations. In the short panel case, temporary policy innovations that affect flow payoffs can be identified even when the flow payoffs are not. When policy innovations affect state transitions, the true normalization (up to a constant) must be recovered in order to obtain the counterfactual choice probabilities. Even this is not enough in the short panel case unless there is an extended terminal or renewal action and the normalized utilities of this action is known. Otherwise restrictions on the flow payoffs are necessary to achieve identification, as is the case when the disturbance structure is unknown.

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

6 Appendix

Proof of Theorem 1.

It is convenient to prove the finite horizon and stationary cases separately, the nonstationary case first. Let $l(x,t) \in \{1,\ldots,J\}$ and $c_t(x)$ respectively denote the normalizing action and benchmark flow utility the associated with (t,x). We set $u^*_{l(x,t),t}(x) = c_t(x)$ and for all $j \neq l(x,t)$ define:

$$u^*_{jT}(x) \equiv u_{jT}(x) - u_{l(x,T),T}(x) + c_T(x) \tag{25}$$

and:

$$u^*_{jt}(x) = u_{jt}(x) + c_t(x) - u_{l(x,t)t}(x) + \sum_{\tau=t+1}^{T}\sum_{x_\tau=1}^{X}\beta^{\tau-t}\left[u^*_{1\tau}(x_\tau) - u_{1\tau}(x_\tau)\right]\left[\kappa^*_{\tau-1}(x_\tau|x_t, l(x,t)) - \kappa^*_{\tau-1}(x_\tau|x_t, j)\right] \tag{26}$$

In the final period T, supposing $x_T = x$ the agent optimally sets $d_{jT} = 1$ if:

$$u_{jT}(x) + \epsilon_{jT} \geq \max_{k\in\{1,\ldots,J\}}\{u_{kT}(x) + \epsilon_{kT}\}$$

inequalities that are satisfied if and only if:

$$u^*_{jT}(x) + \epsilon_{jT} \geq \max_{k\in\{1,\ldots,J\}}\{u^*_{kT}(x) + \epsilon_{kT}\}$$

as required by the theorem, and establishing the result for $T = 1$.

For the representation of $v_{jt}(x_t)$ provided by (7), set $d^*_{1\tau}(x_\tau, k) = 1$ for all $\tau = \{t+1,\ldots,T\}$ $k \in \{1,\ldots,J\}$ and $x_\tau \in \{1,\ldots,X\}$. Supposing $x_t = x$ in period t, the decision maker optimally sets $d_{jt} = 1$ if:

$$j = \arg\max_{k\in\{1,\ldots,J\}}\left\{u_{kt}(x) + \epsilon_{kt} + \sum_{\tau=t+1}^{T}\sum_{x_\tau=1}^{X}\beta^{\tau-t}\left[u_{1\tau}(x_\tau) + \psi_1[p_\tau(x_\tau)]\right]\kappa^*_{\tau-1}(x_\tau|x_t,k)\right\}$$

Subtracting the constant:

$$u_{l(x,t)t}(x) + \sum_{\tau=t+1}^{T}\sum_{x_\tau=1}^{X}\beta^{\tau-t}u_{1\tau}(x_\tau)\kappa^*_{\tau-1}(x_\tau|x_t, l(x,t))$$

does not change the optimal choice, so $d_{jt} = 1$ is optimal if:

$$j = \arg\max_{k\in\{1,\ldots,J\}}\left\{u_{kt}(x) - u_{l(x,t)t}(x) + \epsilon_{kt} + \sum_{\tau=t+1}^{T}\sum_{x_\tau=1}^{X}\beta^{\tau-t}\left\{\begin{array}{c} u_{1\tau}(x_\tau)\left[\kappa^*_{\tau-1}(x_\tau|x_t,k) - \kappa^*_{\tau-1}(x_\tau|x_t,l(x,t))\right] \\ +\psi_k[p_\tau(x_\tau)]\kappa^*_{\tau-1}(x_\tau|x_t,k) \end{array}\right\}\right\} \tag{27}$$

From (26):

$$
\begin{aligned}
& u_{jt}(x)-u_{l(x,t)t}(x)-\sum_{\tau=t+1}^{T}\sum_{x_\tau=1}^{X}\beta^{\tau-t}u_{1\tau}(x_\tau)\left[\kappa^*_{\tau-1}(x_\tau|x_t,l(x,t))-\kappa^*_{\tau-1}(x_\tau|x_t,j)\right] \\
= \;& u^*_{jt}(x)-c_t(x)-\sum_{\tau=t+1}^{T}\sum_{x_\tau=1}^{X}\beta^{\tau-t}u^*_{1\tau}(x_\tau)\left[\kappa^*_{\tau-1}(x_\tau|x_t,l(x,t))-\kappa^*_{\tau-1}(x_\tau|x_t,j)\right]
\end{aligned}
$$

Substitute the second line into the maximand of (27). Then $d_{jt}=1$ is optimal if:

$$
\begin{aligned}
j \;=\; & \underset{k\in\{1,\dots,J\}}{\arg\max}\left\{u^*_{kt}(x)-c_t(x)+\epsilon_{kt}+\sum_{\tau=t+1}^{T}\sum_{x_\tau=1}^{X}\beta^{\tau-t}\left\{\begin{array}{c} u^*_{1\tau}(x_\tau)\left[\kappa^*_{\tau-1}(x_\tau|x_t,k)-\kappa^*_{\tau-1}(x_\tau|x_t,l(x,t))\right] \\ +\psi_k[p_\tau(x_\tau)]\kappa^*_{\tau-1}(x_\tau|x_t,k)\end{array}\right\}\right\} \\
= \;& \underset{k\in\{1,\dots,J\}}{\arg\max}\left\{u^*_{kt}(x)+\epsilon_{kt}+\sum_{\tau=t+1}^{T}\sum_{x_\tau=1}^{X}\beta^{\tau-t}\left[u^*_{1\tau}(x_\tau)+\psi_k[p_\tau(x_\tau)]\right]\kappa^*_{\tau-1}(x_\tau|x_t,k)\right\}
\end{aligned}
$$

as required, where the last line follows because the dropped terms do not depend on the choice. This proves the result for all finite T.

We now turn to infinite horizon stationary models. We start by defining a c_x an i_x and a u_{i_x} for each x analogously to the finite horizon case and set:

$$
u^*_j(x)=u_j(x)+c_x-u_{i_x}(x)+\sum_{\tau=1}^{\infty}\sum_{x_\tau=1}^{X}\beta^{\tau}\left[u^*_1(x_\tau)-u_1(x_\tau)\right]\left[\kappa^*_{\tau-1}(x_\tau|x_t,l(x,t))-\kappa^*_{\tau-1}(x_\tau|x_t,j)\right]
$$

or in matrix notation:

$$
u^*_j=u_j+c-\tilde{u}+\beta\left(F_1-F_j\right)\left[\mathcal{I}-\beta F_1\right]^{-1}\left(u^*_1-u_1\right)
$$

which is the result in the text. ■

Proof of Theorem 2. Substituting in for $v_{jt}(z_t)-v_{1t}(z_t)$ in (11) with the corresponding expression in (12) implies:

$$
\psi_1[p_t(z_t)]-\psi_j[p_t(z_t)]=u_{jt}(z_t)+\sum_{\tau=t+1}^{T}\sum_{z_\tau=1}^{Z}\beta^{\tau-t}\psi_1[p_\tau(z_\tau)]\left[\kappa^*_{\tau-1}(z_\tau|z_t,j)-\kappa^*_{\tau-1}(z_\tau|z_t,1)\right]
$$

Solving for $u_{jt}(z_t)$ completes the first part of the theorem:

$$
u_{jt}(z_t)=\psi_1[p_t(z_t)]-\psi_j[p_t(z_t)]+\sum_{\tau=t+1}^{T}\sum_{z_\tau=1}^{Z}\beta^{\tau-t}\psi_1[p_\tau(z_\tau)]\left[\kappa^*_{\tau-1}(z_\tau|z_t,1)-\kappa^*_{\tau-1}(z_\tau|z_t,j)\right] \tag{28}
$$

To prove the second part, note that the two decision sequences set the initial choices such that $d_{jt}=1$ or $d_{1t}=1$ and then both decision sequences set $d_{1t'}=1$ for all $t'>t$. From the definition of F_1, the columns of F_1^{τ} gives the probabilities of being in each state after τ periods conditional choosing alternative 1 in each of those periods. The rows indicate how these probabilities differ given the initial state. Hence, for $\tau \geq 1$, the (z, z') element of F_1^{τ} is $\kappa^*_{t+\tau-1}(z'|z, 1)$. Similarly, the (z, z') element of $F_j F^{\tau}$ is $\kappa^*_{t+\tau-1}(z'|z, j)$.

Using the matrix notation defined in the theorem, we can express u_j as:

$$u_j = \Psi_j - \Psi_1 + \sum_{\tau=1}^{\infty} \beta^{\tau} (F_1 - F_j) F_1^{\tau-1} \Psi_1 = \Psi_j - \Psi_1 + \beta (F_1 - F_j) \left(\sum_{\tau=0}^{\infty} \beta^{\tau} F_1^{\tau} \right) \Psi_1 \tag{29}$$

Noting that $\beta f_j(z'|z) > 0$ for all (j, z, z') and $\beta \sum_{z'=1}^{Z} f_j(z'|z) = \beta < 1$ for all (j, z), the existence of $[\mathcal{I} - \beta F_1]^{-1}$ follows from Hadley (page 118, 1961) with:

$$Q \equiv \sum_{\tau=0}^{\infty} \beta^{\tau} F_1^{\tau} = \mathcal{I} + \beta Q F_1 = [\mathcal{I} - \beta F_1]^{-1}$$

Substituting the expression for Q into (29) we obtain:

$$u_j = \Psi_j - \Psi_1 + \beta (F_1 - F_j) [\mathcal{I} - \beta F_1]^{-1} \ \Psi_1$$

which proves the theorem. ■

Proof of Theorem 4. In the counterfactual regime, dynamic optimization requires the agent to choose the action that maximizes $\epsilon_{jt} + \widetilde{v}_{jt}(x)$ over $j \in \{1, \ldots, J\}$ which implies:

$$\tilde{p}_{jt}(x) = \int \prod_{k=1}^{J} 1 \{\epsilon_{jt} - \epsilon_{kt} + \widetilde{v}_{jt}(x) - \widetilde{v}_{kt}(x)\} g(\epsilon_t) d\epsilon_t \tag{30}$$

But:

$$\begin{aligned}
\widetilde{v}_{j\mathcal{T}}(x) - \widetilde{v}_{k\mathcal{T}}(x) &= u_{j\mathcal{T}}(x) - u_{k\mathcal{T}}(x) + \Delta_{j\mathcal{T}}(x) - \Delta_{k\mathcal{T}}(x) + \sum_{x'=1}^{X-1} \beta V_{\mathcal{T}+1}(x')] [f_{j\mathcal{T}}(x'|x) - f_{k\mathcal{T}}(x'|x)] \\
&= \Delta_{j\mathcal{T}}(x) - \Delta_{k\mathcal{T}}(x) + v_{j\mathcal{T}}(x) - v_{k\mathcal{T}}(x) \\
&= \Delta_{j\mathcal{T}}(x) - \Delta_{k\mathcal{T}}(x) + \psi_k[p_{\mathcal{T}}(x)] - \psi_j[p_{\mathcal{T}}(x)] \qquad (31)
\end{aligned}$$

Substituting (31) into (30) yields:

$$\tilde{p}_{j\mathcal{T}}(x) = \int \prod_{k=1}^{J} 1 \left\{\epsilon_{j\mathcal{T}} - \epsilon_{k\mathcal{T}} + \Delta_{j\mathcal{T}}(x) - \Delta_{k\mathcal{T}}(x) + \psi_k[p_{\mathcal{T}}(x)] - \psi_j[p_{\mathcal{T}}(x)]\right\} g(\epsilon_{\mathcal{T}}) d\epsilon_{\mathcal{T}}$$

Now we exploit the fact that for all t:

$$\widetilde{V}_t(x) = V_t(x) + \widetilde{v}_{kt}(x) - v_{kt}(x) + \psi_k[\widetilde{p}_t(x)] - \psi_k[p_t(x)]$$

which implies:

$$\begin{aligned}
\widetilde{v}_{jt}(x) - \widetilde{v}_{kt}(x) &= u_{jt}(x) - u_{kt}(x) + \Delta_{jt}(x) - \Delta_{kt}(x) + \sum_{x'=1}^{X-1} \beta \widetilde{V}_{t+1}(x')] \left[f_{jt}(x'|x) - f_{kt}(x'|x) \right] \\
&= u_{jt}(x) - u_{kt}(x) + \Delta_{jt}(x) - \Delta_{kt}(x) \\
&\quad + \sum_{x'=1}^{X-1} \beta V_{t+1}(x') \left[f_{jt}(x'|x) - f_{kt}(x'|x) \right] + \sum_{x'=1}^{X-1} \beta \left[\widetilde{V}_{t+1}(x') - V_{t+1}(x') \right] \left[f_{jt}(x'|x) - f_{kt}(x'|x) \right] \\
&= \Delta_{jt}(x) - \Delta_{kt}(x) + \psi_k[p_t(x)] - \psi_j[p_t(x)] + \sum_{x'=1}^{X-1} \beta \left[\widetilde{V}_{t+1}(x') - V_{t+1}(x') \right] \left[f_{jt}(x'|x) - f_{kt}(x'|x) \right]
\end{aligned}$$

Substituting $\widetilde{v}_{1,t+1}(x) + \psi_1[\widetilde{p}_{t+1}(x)] - \psi_1[p_{t+1}(x)]$ for $\widetilde{V}_{t+1}(x)$ and $v_{k,t+1}(x) + \psi_k[p_{t+1}(x)]$ for $V_{t+1}(x)$ in the equation above, and then appealing to an induction argument, completes the proof. ■

References

[1] Aguirregabiria, V. (2005): "Nonparametric identification of Behavioral Responses to Counterfactual Policy Interventions in Dynamic Discrete Choice Processes," *Economic Letters,* 87, 393-398.

[2] Aguirregabiria, V. (2010): "Another Look at the Identification of Dynamic Discrete Decision Processes, with an Application to Retirement Behavior", *Journal of Business and Economic Statistics*, Vol. 28(3), 201-218.

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

[4] Aguirregabiria, V., and P. Mira (2015): "Identification of Games of Incomplete Information with Multiple Equilibria and Unobserved Heterogeneity", working paper.

[5] Aguirregabiria, V. and J. Suzuki (2014): "Identification and Counterfactuals in Dynamic Models of Market Entry and Exit", *Quantitative Marketing and Economics*, Vol. 12(3), 267-304.

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

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

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

[9] Arcidiacono, P. and R.A. Miller (2015): “Nonstationary Dynamic Models with Finite Dependence,” working paper.

[10] Bajari, P., V. Chernozhukov, H. Hong, and D. Nekipelov (2009): “Nonparametric and Semiparametric Analysis of a Dynamic Discrete Game”, working paper.

[11] Chiong, K.X., A. Galichon, and M. Shum (2013): “Estimating Dynamic Discrete Choice Models via Convex Analysis,” working paper.

[12] Eckstein, Z. and K.I. Wolpin (1989):“The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey,” *The Journal of Human Resources*, 24, 562-598.

[13] 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.

[14] Gayle G. and R. A. Miller (2006): "Life-Cycle Fertility and Human Capital Accumulation", working paper.

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

[16] Hu, Y. and M. Shum (2012): “Nonparametric identification of dynamic models with unobserved state variables,” *Journal of Econometrics*, 171, 32-44.

[17] Kasahara, H., and K. Shimotsu (2009): “Nonparametric Identification and Estimation of Finite Mixture Models of Dynamic Discrete Choices," *Econometrica*, 77, 135-175.

[18] Magnac T. and D. Thesmar (2002):“Identifying Dynamic Discrete Decision Processes,” *Econometrica*, 70, 801-816.

[19] Miller, R.A. (1984): “Job Matching and Occupational Choice,” *Journal of Political Economy*, 92(6), pp. 1086 -1120.

[20] Miller, R.A. (1997): "Estimating models of dynamic optimization with microeconomic data," in *Handbook of Applied Econometrics*, M. Pesaran and P. Schmidt, editors, Basil Blackwell, Vol. 2, pp. 246-299.

[21] Norets, A. and X. Tang (2014): "Semiparametric Inference in Dynamic Binary Choice Models", *Review of Economic Studies*, Vol. 81(3): 1229-1262.

[22] Pakes, A. (1994): "The Estimation of Dynamic Structural Models: Problems and Prospects, Part II. Mixed Continuous-Discrete Control Models and Market Interactions," Chapter 5 of *Advances in Econometrics: Proceedings of the 6th World Congress of the Econometric Society*, edited by J.J. Laffont and C. Sims., pp. 171-259.

[23] Pantano, J. and Y. Zheng (2010): "Using Subjective Expectations Data to Allow for Unobserved Heterogeneity in Hotz-Miller Estimation Strategies", working paper.

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

[25] Rust, J. (1994):"Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics, Volume 4*, ed. by R.E. Engle and D. McFadden. Amsterdam: Elsevier-North Holland, Chapter 51, 3081-3143.