Identifying Dynamic Discrete Choice Models off Short Panels*

Peter Arcidiacono
Duke University & NBER

Robert A. Miller
Carnegie Mellon University

June 1, 2015

Abstract

This paper analyzes the identification of dynamic discrete choice models in both single agent optimization problems and multiagent games. We provide conditions for identifying flow payoffs in three distinct scenarios: stationary settings; nonstationary settings with a long panel (meaning data is sampled in every period respondents make decisions); and nonstationary settings with a short panel (where the relevant time horizon for respondents extends beyond the length of the data). Turning to the predictions in counterfactual regimes, we establish three further results. Counterfactual choice probabilities for (temporary) policy changes that only affect payoffs are identified (in nonstationary settings with a short panel) from the population generating the sample. In long panels and stationary models, knowing the true utility value of the normalization is required to compute the choice probabilities for a counterfactual policy change induced by an innovation to the state transition probabilities. Even that extra requirement is not sufficient when the panel is short, except in special cases where the model has a specialized form of the finite dependence property. Finally, in nonstationary models with short panels, exclusion restrictions and functional form assumptions can be used to recover counterfactual choice probabilities induced by innovations to state transitions when the flow payoffs are not time dependent.

*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.

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 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 both single-agent and multi-agent settings.

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. Pesendorfer and Schmidt-Dengler (2008) prove that this result does not extend to games because the expected flow payoffs embed state-specific flow payoffs that depend on the actions of the other players, implying that the number of parameters exceeds the number of equations that characterize the empirical content of the game. Absent parametric restrictions, the state-specific utility flows in games are not identified. Still more recently, 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.[2]

This research has focused on cases where the model is either stationary or where the data covers the full

[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).

[2] 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. Pantano and Xheng (2010) examine identification in the presence of unobserved heterogeneity when subjective expectations data is available. In this paper we show that models with unobserved heterogeneity can only be identified by placing structure on the problem through functional form assumptions or exclusion restrictions, which in turn limits the scope for observed heterogeneity to affect behavior.

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 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. These latter processes are not observed by the econometrician. Further, ass has previously been noted by Pesendorfer and Schmidt-Dengler (2008), additional assumptions are required to identify the primitives in games, because individual payoffs depend on the actions of all the players but each player only considers their own choices taking the equilibrium behavior of the others as a given.

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.

Our second set of results is concerned with the identification of counterfactual conditional choice probabilities. Previous work by Norets and Tang (2014) noted that in the stationary case of a binary choice model, normalizing the flow payoff for one of the choices to zero in all states may not be 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 their remarkable result with two contributions. First we show that in 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. Otherwise, additional exclusion and functional form restrictions of the type mentioned above are necessary to identify counterfactual conditional choice probabilities when the policy innovation affects the state transitions.

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 shows how exclusion restrictions and stability of the flow payoffs can be used to identify the true normalization. Section 5 turns to the identification of conditional choice probabilities under counterfactual regimes, and a conclusion briefly wraps up the analysis by highlighting the links between observationally equivalent normalizations and the limitations that implies for identifying from short panels counterfactual CCPs induced by innovations to probability transitions.

2 Framework

This section lays out a general class of dynamic discrete choice models, and then extends it to dynamic games. 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:

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

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:

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

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

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

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

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

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

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:[3]

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

2.1 Extension to dynamic games

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

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

$$P_t\left(d_t^{(\sim n)}|x_t\right) = \prod_{\substack{n'=1\\ n'\neq n}}^{I}\left(\sum_{j=1}^{J}d_{jt}^{(n')}p_{jt}^{(n')}(x_t)\right) \tag{4}$$

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

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 (4). Taking the expectation of $U_{jt}^{(n)}\left(x_t, d_t^{(\sim n)}\right)$ over $d_t^{(\sim n)}$, we define the systematic component of the current utility of player n as a function of the state variables as:

$$u_{jt}^{(n)}\left(x_t\right) = \sum_{d_t^{(\sim n)} \in J^{N-1}} P_t\left(d_t^{(\sim n)}\left|x_t\right.\right) U_{jt}^{(n)}\left(x_t, d_t^{(\sim n)}\right) \tag{5}$$

For future reference we call $u_{jt}^{(n)}\left(x_t\right)$ 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. Denote $F_{jt}\left(x_{t+1}\left|x_t, d_t^{(\sim n)}\right.\right)$ as the probability of x_{t+1} occurring given action j by player n in period t, when its state variables are x_t and the other players choose $d_t^{(\sim n)}$. From the perspective of player n the probability of transitioning from x_t to x_{t+1} given action j is:

$$f_{jt}^{(n)}\left(x_{t+1}\left|x_t\right.\right) = \sum_{d_t^{(\sim n)} \in J^{N-1}} P_t\left(d_t^{(\sim n)}\left|x_t\right.\right) F_{jt}\left(x_{t+1}\left|x_t, d_t^{(\sim n)}\right.\right) \tag{6}$$

Analogous to the reduced form payoffs, this is the (one period) reduced form transition to state x_{t+1} from x_t when player n takes action j in period t. The expressions for the conditional value functions for player n are the same as we described in Subsection 2. Loosely speaking, player n solves a dynamic optimization problem treating the other player's equilibrium actions as nature that help determine both the flow payoffs and the state transitions.

Two critical differences distinguish noncooperative discrete choice dynamic games from their single agents counterparts, and both are relevant for studying identification. Whereas $u_{jt}(x_t)$ denotes the primitive utility flow term in single agent optimization problems, $u_{jt}^{(n)}(x_t)$ defined by (5) is a reduced form parameter that depends on the actions of the other players. In dynamic games, the flow payoff $u_{jt}^{(n)}(x_t)$ is not a primitive but an expected utility found by integrating $U_{jt}^{(n)}\left(x_t, d_t^{(\sim n)}\right)$ over the joint probability distribution $P_t\left(d_t^{(\sim n)}\left|x_t\right.\right)$ induced by the current actions of the other players simultaneously making their equilibrium choices that are partly determined by their private information. Similarly, $f_{jt}\left(x_{t+1}\left|x_t\right.\right)$ is the primitive defining the state transition probabilities in single agent optimization problems, but $f_{jt}^{(n)}\left(x_{t+1}\left|x_t\right.\right)$, defined by (6), is a reduced form parameter that depends on the conditional choice probabilities of the other players, $P_t\left(d_t^{(\sim n)}\left|x_t\right.\right)$, as well as the primitive $F_{jt}\left(x_{t+1}\left|x_t, d_t^{(\sim n)}\right.\right)$.

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,[4] summarized with the notation (u, β, F, G). In this section we build upon Rust (1994), Magnac and Thesmar (2002) and Norets and Tang (2014) in single agent settings and Pesendorfer Schmidt-Dengler (2008) in games settings, by considering identification when (β, F, G) are known.[5] 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:

$$\psi_j[p_t(x)] \equiv V_t(x) - v_{jt}(x) \tag{7}$$

To interpret (7), 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:

Theorem 1 ***(Arcidiacono and Miller 2011, page 1835)*** *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*

[4] 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.

[5] 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.

constraints that $\sum_{k=1}^{J} \omega_{k\tau}(x_\tau, j) = 1$.[6] *Recursively define* $\kappa_\tau(x_{\tau+1}|x_t, j)$ *as:*

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

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

$$\begin{aligned} v_{jt}(x_t) &= u_{jt}(x_t) + \sum_{\tau=t+1}^{\mathcal{T}} \sum_{k=1}^{J} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \left[u_{k\tau}(x_\tau) + \psi_k[p_\tau(x_\tau)]\right] \omega_{k\tau}(x_\tau, j)\kappa_{\tau-1}(x_\tau|x_t, j) \\ &\quad + \sum_{x_{\mathcal{T}+1}}^{X} \beta^{\mathcal{T}+1-t} V_{\mathcal{T}+1}(x_{\mathcal{T}+1})\kappa_{\mathcal{T}}(x_{\mathcal{T}+1}|x_t, j) \end{aligned} \tag{9}$$

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

$$v_{jt}(x_t) = u_{jt}(x_t) + \sum_{\tau=t+1}^{\mathcal{T}} \sum_{k=1}^{J} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \left[u_{k\tau}(x_\tau) + \psi_k[p_\tau(x_\tau)]\right] \omega_{k\tau}(x_\tau, j)\kappa_{\tau-1}(x_\tau|x_t, j) \tag{10}$$

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 and dynamic games. 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:

$$\kappa^*_\tau(x_{\tau+1}|x_t, j) \equiv \begin{cases} f_{jt}(x_{t+1}|x_t) & \text{for } \tau = t \\ \sum_{x_\tau=1}^{X} f_{1\tau}(x_{\tau+1}|x_\tau)\kappa^*_{\tau-1}(x_\tau|x_t, j) & \text{for } \tau = t+1, \dots, T \end{cases} \tag{11}$$

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 $l(x, t) \in \{1, \dots, J\}$ and set the flow utility associated with that choice, $u^*_{l(x,t),t}(x)$, to an arbitrary real value we denote by $c_t(x)$.

[6] Arcidiacono and Miller (2011) state the theorem for $T = \mathcal{T}$ and $\omega_{k\tau}(x_\tau, j) \geq 0$. However the induction proof also covers the case in which $T > \mathcal{T}$, and as Arcidiacono and Miller (2014) show, the constraints on $\omega_{k\tau}(x_\tau, j)$ (which do not play any role in this analysis) are redundant, .

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 $c(x)$.[7]

Theorem 2 *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 $j \in \{1, \ldots, J\}$:*

$$u^*_{jT}(x) \equiv u_{jT}(x) - u_{l(x,T),T}(x) + c_T(x) \tag{12}$$

and:

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

When the environment is stationary, define:

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

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

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

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

[7] We defer until Section 6 discussion about recovering the primitive utility flows $U^{(n)}_{jt}\left(x_t, d^{(\sim n)}_t\right)$ in games from the reduced form flows $u^{(n)}_{jt}(x_t)$. It is however worth noting here that there are $(N-1)(J-1)$ values of $U^{(n)}_{jt}\left(x_t, d^{(\sim n)}_t\right)$ for each value of (n, j, t, x_t) uniquely determining $u^{(n)}_{jt}(x_t)$, so many more normalizations are possible.

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. Proving these generalizations is essentially an exercise in applying Theorem 1.

Theorem 2 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 (10) 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:

$$v_{jt}(x_t) - v_{1t}(x_t) = u_{jt}(x_t) + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \psi_1[p_\tau(x_\tau)] \left[\kappa^*_{\tau-1}(x_\tau | x_t, j) - \kappa^*_{\tau-1}(x_\tau | x_t, 1)\right] \tag{14}$$

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 (7):

$$v_{jt}(x_t) - v_{1t}(x_t) = \psi_1[p_t(x_t)] - \psi_j[p_t(x_t)] \tag{15}$$

Theorem 3 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 3 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 3 *For all j, t, and x_t, the flow payoff $u_{jt}(x_t)$ can be expressed as:*

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

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:

$$u_j = \Psi_j - \Psi_1 + \beta (F_1 - F_j) [\mathcal{I} - \beta F_1]^{-1} \; \Psi_1 \tag{17}$$

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 (16) and (17) 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 (16) , 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 4 *If $u_{1t}(x_t) = 0$ for all t and x_t, then:*

$$
\begin{aligned}
u_{jt}(x_t) \quad = \quad & \psi_1[p_t(x_t)] - \psi_j[p_t(x_t)] + \sum_{\tau=t+1}^{\mathcal{T}} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \psi_1[p_\tau(x_\tau)] \left[\kappa^*_{\tau-1}(x_\tau|x_t,1) - \kappa^*_{\tau-1}(x_\tau|x_t,j) \right] \\
& + \sum_{x_{\mathcal{T}+1}=1}^{X-1} \beta^{\mathcal{T}-t} V_{\mathcal{T}+1}(x_{\mathcal{T}+1})] \left[\kappa^*_{\mathcal{T}}(x_{\mathcal{T}+1}|x_t,1) - \kappa^*_{\mathcal{T}}(x_{\mathcal{T}+1}|x_t,j) \right]
\end{aligned} \tag{18}
$$

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 (18) 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:

$$V_{\mathcal{T}+1}(x_{\mathcal{T}+1}) = u_{1\mathcal{T}+1}(x_{\mathcal{T}+1})+\psi_1\left[p_{\mathcal{T}+1}(x_{\mathcal{T}+1})\right]+\sum_{\tau=\mathcal{T}+1}^{T}\sum_{x_{\tau=1}}^{X}\beta^{\tau-\mathcal{T}}\psi_1\left[p_\tau(x_\tau)\right]\left[\kappa^*_{\tau-1}(x_\tau|x_t,1)-\kappa^*_{\tau-1}(x_\tau|x_t,j)\right] \tag{19}$$

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:

$$u_{1\mathcal{T}+1}(x_{\mathcal{T}+1}) = -\psi_1\left[p_{\mathcal{T}+1}(x_{\mathcal{T}+1})\right]-\sum_{\tau=\mathcal{T}+2}^{T}\sum_{x_{\tau=1}}^{X}\beta^{\tau-\mathcal{T}-1}\psi_1\left[p_\tau(x_\tau)\right]\left[\kappa^*_{\tau-1}(x_\tau|x_t,1)-\kappa^*_{\tau-1}(x_\tau|x_t,j)\right] \tag{20}$$

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 Recovering the Remaining Primitives

The results in the previous section suggest that identification of the normalized payoffs can be achieved in the short panel case only by making a normalization that relies on data beyond the sample period. When the researcher instead wants to work with a known normalization, either because the normalization is in fact known or is more palatable based on economic reasons, identification of the normalized flow payoffs can be achieved through the use of exclusion restrictions on the arguments in the utility function, or by functional form. We show how to establish identification when there are exclusions restrictions (variables that affect the state transitions but do not enter the flow payoffs), and when the payoff flow functions are stable over time (flow payoffs depend on the state and choice but not directly on time itself). Finally, we show that stability in the flow payoffs can be used to recover normalized payoffs in games.

4.1 Exclusion restrictions

One means of restoring identification of the normalized flow payoffs is to reduce the parameter space by imposing exclusion restrictions. For example, assuming some variables affect the state transitions, but not the flow payoffs. The main result of this subsection is to show that an exclusion restriction on a variable that only takes two values may suffice to restore the identification of the flow payoffs, along with the ex-ante value function at $\mathcal{T}$, even when there is only one period of data on choices followed by one set of transitions.

For example, consider a non-stationary setting with three choices. It is convenient to work with the matrix analog of (18), denoting u_{jt}, Ψ_{jt}, and V_t as the matrix analogs of $u_{jt}(x)$, $\psi_j[p_t(x_t)]$, and $V_t(x_t)$. Denote F_{jt} as the $X \times X$ transition matrix given choice j at time t. Finally, normalize $u_{1t}(x_t) = 0$ for all t and x_t. We then have the following system of equations with only period of data:

$$\begin{bmatrix} u_{2t} \\ u_{3t} \end{bmatrix} = \begin{bmatrix} \Psi_{1t} - \Psi_{2t} \\ \Psi_{1t} - \Psi_{3t} \end{bmatrix} + \beta \begin{bmatrix} F_{1t} - F_{2t} \\ F_{1t} - F_{3t} \end{bmatrix} \begin{bmatrix} V_{t+1} \\ V_{t+1} \end{bmatrix} \tag{21}$$

Absent exclusion restrictions, we have a system of $2X$ equations with $3X - 1$ unknowns, $2X$ unknowns corresponding to the flow payoffs for each choice and $X - 1$ unknowns corresponding to the elements of V_{t+1} as one of the elements of V_{t+1} can be normalized.

Now suppose X can be partitioned into X_1 and X_2 where X_2 contains two values and where x_2 affects the state transitions but not the flow payoffs:

$$u_{jt}(x_1, x_2) = u_{jt}(x_1)$$

Letting superscripts indicate the value of x_2 at time t, we have the following system of equations:

$$\begin{bmatrix} u_{2t}^{(1)} \\ u_{2t}^{(2)} \\ u_{3t}^{(1)} \\ u_{3t}^{(2)} \end{bmatrix} = \begin{bmatrix} \Psi_{1t}^{(1)} - \Psi_{2t}^{(1)} \\ \Psi_{1t}^{(2)} - \Psi_{2t}^{(2)} \\ \Psi_{1t}^{(1)} - \Psi_{3t}^{(1)} \\ \Psi_{1t}^{(2)} - \Psi_{3t}^{(2)} \end{bmatrix} + \beta \begin{bmatrix} F_{1t}^{(1)} - F_{2t}^{(1)} \\ F_{1t}^{(2)} - F_{2t}^{(2)} \\ F_{1t}^{(1)} - F_{3t}^{(1)} \\ F_{1t}^{(2)} - F_{3t}^{(2)} \end{bmatrix} \begin{bmatrix} V_{t+1} \\ V_{t+1} \end{bmatrix} \tag{22}$$

which now has $4X_1$ equations and $4X_1 - 1$ unknowns.

Since $u_{jt}^{(1)} = u_{jt}^{(2)}$, taking differences and rearranging terms yields:

$$\begin{bmatrix} \Psi_{2t}^{(1)} - \Psi_{1t}^{(1)} + \Psi_{1t}^{(2)} - \Psi_{2t}^{(2)} \\ \Psi_{3t}^{(1)} - \Psi_{1t}^{(1)} + \Psi_{1t}^{(2)} - \Psi_{3t}^{(2)} \end{bmatrix} = \beta \begin{bmatrix} F_{1t}^{(1)} - F_{2t}^{(1)} - F_{1t}^{(2)} + F_{2t}^{(2)} \\ F_{1t}^{(1)} - F_{3t}^{(1)} - F_{1t}^{(2)} + F_{3t}^{(2)} \end{bmatrix} V_{t+1} \tag{23}$$

Hence if the rank of the $2X_1 \times 2X_1$ matrix:

$$\begin{bmatrix} F_{1t}^{(1)} - F_{2t}^{(1)} - F_{1t}^{(2)} + F_{2t}^{(2)} \\ F_{1t}^{(1)} - F_{3t}^{(1)} - F_{1t}^{(2)} + F_{3t}^{(2)} \end{bmatrix}$$

is at least $2X_1 - 1$, we can solve for the value functions. Given the value functions, we obtain the flow payoffs using (1) and (2).

4.2 Stable utility functions

Stability of the flow payoffs can also restore identification of the normalized flow payoffs. We define stability of the flow payoffs as $u_{jt}(x) = u_{jt'}(x)$ for all $\{t, t'\}$ and for all $j \in [1, \ldots, J]$. In this case the non-stationarity comes from either the state transitions or the time horizon. Identification is achieved by solving for both the flow payoffs and the value functions in the last period, similar to Section 4.1.

To illustrate how stability of the flow payoffs may restore identification in the short panel case, suppose there are only two choices each period, and the data covers two periods, t and $t+1$. We now assume $u_{2t}(x) = u_{2t+1}(x) = u_2(x)$ for all $x \in \{1, \ldots, X\}$ and adopt the normalization $u_{1t}(x) = u_{1t+1}(x) = 0$.

We again adopt the matrix representation in Section 4.1. We can then express u_2, the vector of flow payoffs for action 2 in every state, relative to choosing action 1 in the next period. Similarly, we can also express u_2 relative to choosing action 1 in the next period and in the period after that. Hence, given conditional choice probabilities in two periods, t and $t+1$, we can express u_2 as:

$$u_2 = \Psi_{1t+1} - \Psi_{2t+1} + \beta(F_{1t+1} - F_{2t+1})V_{t+2} \tag{24}$$

$$u_2 = \Psi_{1t} - \Psi_{2t} + \beta(F_{1t} - F_{2t})\Psi_{1t+1} + \beta^2(F_{1t} - F_{2t})F_{1t+1}V_{t+2} \tag{25}$$

Taking differences and rearranging terms yields:

$$\Psi_{1t} - \Psi_{2t} - \Psi_{1t+1} + \Psi_{2t+1} + \beta(F_{1t} - F_{2t})\Psi_{1t+1} = \beta(F_{1t+1} - F_{2t+1} - \beta F_{1t}F_{1t+1} + \beta F_{2t}F_{1t+1})V_{t+2} \tag{26}$$

Since adding a constant to the future value terms does not affect choice probabilities, we need the rank of

$$(F_{1t+1} - F_{2t+1} - \beta F_{1t}F_{1t+1} + \beta F_{2t}F_{1t+1})$$

to be $X-1$ to identify the $X-1$ differenced value functions. Note that the rank condition is on differences in state transition matrices–it does not depend on the conditional choice probabilities or the flow payoffs–implying that the rank condition is straightforward to check given reasonably-sized problems.

4.3 Recovering primitives in dynamic games

Additional structure is also needed to identify the primitives–as opposed to the reduced form payoffs–of dynamic games. To illustrate how nonstationarity aids in the recovery of the primitive flow payoffs in games settings (up to a normalization), we first consider an entry/exit game. An incumbent (new) firm, labelled $n \in \{1, 2\}$, remains in (enters) the market, by setting $d_{2t}^{(n)} = 1$, or exits (stays out forever), by setting $d_{1t}^{(n)} = 1$. Opportunities for new firms only arise when there are less than 2 firms in the market. The market continues at least until after the sample period, $T > \mathcal{T}$.

The transitory payoff shock to player n from making choice j is denoted by $\epsilon_{jt}^{(n)}$. Current period payoffs for entering or remaining in the market depend on whether there is a rival in the market ($d_{2t}^{(\sim n)}(x_t) = 1$) or not ($d_{2t}^{(\sim n)}(x_t) = 0$), whether the firm is an incumbent ($d_{2t-1}^{(n)}(x_t) = 1$) or not ($d_{2t-1}^{(n)}(x_t) = 0$), and the value of a discrete market state variable denoted by $x_{1t} \in \{1, \ldots X_1\}$. Let $x_t \equiv \left(d_{t-1}^{(n)}, d_{t-1}^{(\sim n)}, x_{1t}\right)$. We assume the flow utility from entry (remaining in the industry) does not depend on time. Formally $U_{2t}^{(n)}(x_t) = U_2^{(n)}(x_t)$ for $t < T$. We normalize the payoff from exiting or staying out to zero, so $U_{1t}^{(n)}(x_t) = u_{1t}^{(n)}(x_t) = 0$. The model is nonstationary because the transition of the market state variable is time dependent, with transition probability given by $F_t(x_{1t+1}|x_{1t}) > 0$ for all $x_{1t+1} \in X_1$.

None of primitive flow payoffs depend on t. However, the reduced form payoffs from entry are defined by:

$$u_{2t}^{(n)}(x_t) = \sum_j p_{jt}^{(\sim n)}(x_t) U_2^{(n)}(j, d_{2t-1}^{(n)}, x_{1t}) \tag{27}$$

and do depend on time through the conditional choice probabilities. From Equation (6), the joint probability of a rival choosing action j in period t, and x_{1t+1} occurring next period, is $f_{2t}^{(n)}(j, x_{1t+1}|x_{1t}) \equiv p_{jt}^{(\sim n)}(x_t) F_t(x_{1t+1}|x_{1t})$. Adapting Equation (36) we obtain:

$$u_{2t}^{(n)}(x_t) = \psi_1^{(n)}[p_t^{(n)}(x_t)] - \psi_2^{(n)}[p_t^{(n)}(x_t)] - \beta \sum_{x_{1t+1}=1}^{X_1} \sum_{j=1}^{2} \psi_1^{(n)}[p_{t+1}^{(n)}(1, j, x_{1t+1})] f_{2t}^{(n)}(j, x_{1t+1}|x_t) \tag{28}$$

Equation (28) identifies $u_{2t}^{(n)}(x_t)$ from the conditional choice and transition probabilities, by exploiting the terminal state property coupled to the normalization of the exit utility.

To recover the primitive utility flow $U_2^{(n)}(j, d_{2t-1}^{(n)}, x_{1t})$ we substitute the expression for $u_{2t}^{(n)}(x_t)$ from (27) into (28) to obtain:

$$\sum_j p_{jt}^{(\sim n)}(x_t) U_2^{(n)}(j, d_{2t-1}^{(n)}, x_{1t}) = \psi_1^{(n)}[p_t^{(n)}(x_t)] - \psi_2^{(n)}[p_t^{(n)}(x_t)] - \beta \sum_{x_{1t+1}=1}^{X_1} \sum_{j=1}^{2} \psi_1^{(n)}[p_{t+1}^{(n)}(1, j, x_{1t+1})] f_{2t}^{(n)}(j, x_{1t+1}|x_t) \tag{29}$$

For any given β and x_t, Equation (29) is linear in $U_2^{(n)}(1, d_{2t-1}^{(n)}, x_{1t})$ and $U_2^{(n)}(2, d_{2t-1}^{(n)}, x_{1t})$. The terms on the right hand side of Equation (29) as well as $p_{1t}^{(\sim n)}(x_t)$ and $p_{2t}^{(\sim n)}(x_t)$ on the left hand side are identified from the conditional choice probabilities and state transitions, and vary over time. If $x_t = x_{t'} = x$ for some $t \neq t'$, then $p_{jt}^{(n)}(x) \neq p_{jt'}^{(n)}(x)$ and $f_{2t}^{(n)}(j, x_{t+1}|x) \neq f_{2t'}^{(n)}(j, x_{t'+1}|x)$, so the resulting two equations evaluated at time periods t and t' can be solved to recover $U_2^{(n)}(1, d_{2t-1}^{(n)}, x_{1t})$ and $U_2^{(n)}(2, d_{2t-1}^{(n)}, x_{1t})$. If three or more periods have common state variables, then over-identifying restrictions emerge that could be used to recover, for example, β. To summarize:

Theorem 5 *In the entry/exit model* $U_2^{(n)}(j, k, x_1)$ *are identified for all firms* $i \in \{1, \ldots, I\}$ *, rival actions* $j \in \{1, 2\}$ *, incumbency status* $k \in \{0, 1\}$ *and market states* $x_1 \in \{1, \ldots X_1\}$ *if* $\mathcal{T} \geq 2$.

5 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) then 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:

$$\tilde{p}_{jt}(x) = \int \prod_{k=1}^{J} 1 \left\{ \begin{array}{l} \epsilon_{jt} - \epsilon_{kt} + \Delta_{jt}(x) - \Delta_{kt}(x) + u_{jt}(x) - u_{kt}(x) \\ + \sum_{x'=1}^{X} \beta \widetilde{V}_{t+1}(x') \left[f_{jt}(x'|x) - f_{kt}(x'|x) + \Lambda_{jt}(x'|x) - \Lambda_{kt}(x'|x) \right] \end{array} \right\} g(\epsilon_t)\, d\epsilon_t \quad (30)$$

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

$$\widetilde{V}_{t+1}(x') = \psi_1[\widetilde{p}_{t+1}(x')] + u_{1,t+1}(x') + \Delta_{1,t+1}(x') + \sum_{x''=1}^{X} \beta \widetilde{V}_{t+2}(x'') \left[f_{1,t+1}(x''|x') + \Lambda_{1,t+1}(x''|x') \right] \quad (31)$$

which exploits equation (7) 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.

5.1 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 (13) in Theorem 2:

$$\widetilde{u}^*_{jt}(x) = \widetilde{u}_{jt}(x) + c_t(x) - \widetilde{u}_{l(x,t),t}(x) + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \left[\widetilde{u}^*_{1\tau}(x_\tau) - \widetilde{u}_{1\tau}(x_\tau) \right] \left[\kappa^*_{\tau-1}(x_\tau|x_t, l(x,t)) - \kappa^*_{\tau-1}(x_\tau|x_t, j) \right] \quad (32)$$

Noting that $\widetilde{u}_{jt}(x) \equiv u_{jt}(x) + \Delta_{jt}(x)$ we difference (32) and (13), in other words the difference of (13) before and after the payoff innovation, to obtain:

$$\begin{aligned} \widetilde{u}^*_{jt}(x) - u^*_{jt}(x) &= \Delta_{jt}(x) - \Delta_{l(x,t),t}(x) \\ &\quad + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \left[\widetilde{u}^*_{1\tau}(x_\tau) - u^*_{1\tau}(x_\tau) - \Delta_{1\tau}(x_\tau) \right] \left[\kappa^*_{\tau-1}(x_\tau|x_t, l(x,t)) - \kappa^*_{\tau-1}(x_\tau|x_t, j) \right] \\ &= \Delta_{jt}(x) - \Delta_{l(x,t),t}(x) \\ &\quad + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \Delta_{l(x,\tau),\tau}(x) \left[\kappa^*_{\tau-1}(x_\tau|x_t, l(x,t)) - \kappa^*_{\tau-1}(x_\tau|x_t, j) \right] \end{aligned} \quad (33)$$

where the second line in the equation follows from setting $j = 1$ and applying a backwards induction argument. Equation (33) 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 (18) 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 6 *Given any temporary payoff innovation in which $\Delta_{jt}(x) = 0$ for all $t > \mathcal{T}$ then:*

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

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

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

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:

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

It is important to note that the results on policy innovations that affect flow payoffs only apply to single agent settings. In games settings, any payoff innovation for one player also affects the state transitions through the other players' conditional choice probabilities. Hence the negative results in the next section apply to the games setting regardless of whether the policy innovation affects flow payoffs or the structural state transitions.

5.2 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 (30) and (31):

$$\tilde{p}_{jt}(x) = \int \prod_{k=1}^{J} 1 \left\{ \begin{array}{l} \epsilon_{jt} - \epsilon_{kt} + u_{jt}(x) - u_{kt}(x) \\ + \sum_{\tau=t+1}^{T} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \left[\psi_1[\widetilde{p}_\tau(x_\tau)] + u_{1,\tau}(x_\tau)\right] \left[\widetilde{\kappa}^*_{\tau-1}(x_\tau|x,k) - \widetilde{\kappa}^*_{\tau-1}(x_\tau|x_t,j)\right] \end{array} \right\} g(\epsilon_t)\, d\epsilon_t \tag{34}$$

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 (11) 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.

5.2.1 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.

5.3 Restoring Identification

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 we 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. Then we shows how exclusion and functional form restrictions can be brought to bear on the problem ot achieve identification.

5.3.1 Renewal and terminal choices

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 $j \in \{1, \ldots, J\}$ 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 $t+1$:

$$\sum_{x_{t+1}=1}^{X} f_{1,t+1}(x_{t+2}|x_{t+1}) f_{jt}(x_{t+1}|x_t) = \sum_{x_{t+1}=1}^{X} f_{1,t+1}(x_{t+2}|x_{t+1}) f_{1t}(x_{t+1}|x_t) \tag{35}$$

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 (18). 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 (18) simplifies to:

$$u_{jt}(x_t) = \psi_1[p_t(x_t)] - \psi_j[p_t(x_t)] + \sum_{\tau=t+1}^{t+\rho} \sum_{x_\tau=1}^{X} \beta^{\tau-t} \psi_1[p_\tau(x_\tau)] \left[\kappa^*_{\tau-1}(x_\tau|x_t, 1) - \kappa^*_{\tau-1}(x_\tau|x_t, j)\right] \tag{36}$$

as after ρ periods the probabilities of being in each of the states are the same across the two paths.[8] In 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.

5.3.2 Recovering the true normalization through further restrictions

When either (i) there is no terminal or renewal or action or (ii) the flow payoff of the terminal or renewal action is not known, then restrictions can be placed on the utility flows to identify the counterfactual CCPs under regimes where the state transitions are changed. Similar to section 4.2, working with stable flow payoff functions can sometimes be used to recover all flow payoffs subject to normalizing the payoff of one choice in one state to zero, which is innocuous. We do, however, need three periods of data, rather than two, to recover the normalization.

[8] 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.

Consider again the two-choice case in section 4.2. Given three periods of data, we can express the flow payoffs for choice 2 as:

$$u_2 = u_1 + \Psi_{1t+2} - \Psi_{2t+2} + \beta(F_{1t+2} - F_{2t+2})V_{t+3} \tag{37}$$

$$u_2 = u_1 + \Psi_{1t+1} - \Psi_{2t+1} + \beta(F_{1t+1} - F_{2t+1})(u_1 + \Psi_{1t+2}) + \beta^2(F_{1t+1} - F_{2t+1})F_{1t+2}V_{t+3} \tag{38}$$

$$\begin{aligned} u_2 = {} & u_1 + \Psi_{1t} - \Psi_{2t} + \beta(F_{1t} - F_{2t})(u_1 + \Psi_{1t+1}) + \beta^2(F_{1t} - F_{2t})F_{1t+1}(u_1 + \Psi_{1t+2}) \\ & + \beta^3(F_{1t} - F_{2t})F_{1t+1}F_{1t+2}V_{t+3} \end{aligned} \tag{39}$$

Denote F^*_{jt} as F_{jt} with the last column removed. Similarly, denote u^*_1 and V^*_{t+3} as u_1 and V_{t+3} with the last element of each removed, effectively setting these last two elements to zero. Subtracting (37) from (38) and (39) and rearranging terms yields the following matrix form:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} u^*_1 \\ V^*_{t+3} \end{bmatrix} = \begin{bmatrix} E \\ F \end{bmatrix} \tag{40}$$

where:

$$\begin{aligned}
A &= \beta(F^*_{1t+1} - F^*_{2t+1}) \\
B &= \beta(F^*_{2t+2} - F^*_{1t+2}) + \beta^2(F_{1t+1} - F_{2t+1})F^*_{1t+2} \\
C &= \beta(F^*_{1t} - F^*_{2t}) + \beta^2(F_{1t} - F_{2t})F^*_{1t+1} \\
D &= \beta(F^*_{2t+2} - F^*_{1t+2}) + \beta^3(F_{1t} - F_{2t})F_{1t+1}F^*_{1t+2} \\
E &= \psi_{2t+1} - \psi_{1t+1} + \psi_{1t+2} - \psi_{2t+2} + \beta(F_{2t+1} - F_{1t+1})\psi_{1t+2} \\
F &= \psi_{2t} - \psi_{1t} + \psi_{1t+2} - \psi_{2t+2} + \beta(F_{2t} - F_{1t})\psi_{1t+1} + \beta^2(F_{2t} - F_{1t})F_{1t+1}\psi_{1t+2}
\end{aligned}$$

a $2 \times X$ equations with $2 \times X - 2$ unknowns. Hence if the rank of $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ equals $2 \times X - 2$, then we can recover u^*_1 and V^*_{t+3} which can then be substituted into (37) to recover u_2.

To give an example of how recovering the normalization works in practice, we now extend the two-period smoking example. Now let individuals make decisions until T but data is only observed at t and $t + 1 < T$. Again there are two states: good health, $x_t = 1$, and bad health, $x_t = 2$. We now work with a more general health transition matrix such that $\pi_{jt}(x_t)$ gives the probability of being in bad health at period $t + 1$ given choice j and state x_t at time t.

There are three unknown flow payoffs: the flow payoff of smoking in either health state and the flow payoff for not smoking in the bad health state. Recall that normalizing the flow payoff for one choice in one state is innocuous, in this case the normalization is that the flow payoff for not smoking in the good health state is zero. We also need to recover the ex-ante value function at $t+2$ for one of the states where we choose to recover it for the bad health state. Again, setting the ex-ante value function at $t+2$ to zero for the good health state is innocuous.

Expressing the conditional value functions out to $t+2$ and differencing yields four equations and four unknowns:

$$
\begin{aligned}
\psi_{1t+1}(1)-\psi_{2t+1}(1) &= u_2(1)+\beta\left(\pi_{2t+1}(1)-\pi_{1t+1}(1)\right)V_{t+2}(2)\\
\psi_{1t+1}(2)-\psi_{2t+1}(2) &= u_2(2)-u_1(2)+\beta\left(\pi_{2t+1}(2)-\pi_{1t+1}(2)\right)V_{t+2}(2)\\
\psi_{1t}(1)-\psi_{2t}(1) &= u_2(1)+\beta(\pi_{2t}(1)-\pi_{1t}(1))(u_1(2)+\psi_{1t+1}(2))\\
&\quad+\beta^2\left(\pi_{2t}(1)-\pi_{1t}(1)\right)\left(\pi_{1t+1}(2)-\pi_{1t+1}(1)\right)V_{t+2}(2)\\
\psi_{1t}(2)-\psi_{2t}(2) &= u_2(2)-u_1(2)+\beta(\pi_{2t}(2)-\pi_{1t}(2))(u_1(2)+\psi_{1t+1}(2))\\
&\quad+\beta^2\left(\pi_{2t}(2)-\pi_{1t}(2)\right)\left(\pi_{1t+1}(2)-\pi_{1t+1}(1)\right)V_{t+2}(2)
\end{aligned}
$$

In matrix notation we have:

$$
\begin{bmatrix} 1 & 0 & 0 & a \\ 0 & 1 & -1 & b \\ 1 & 0 & c & d \\ 0 & 1 & e & f \end{bmatrix}
\begin{bmatrix} u_2(1) \\ u_2(2) \\ u_1(2) \\ V_{t+2}(2) \end{bmatrix}
=
\begin{bmatrix} \psi_{1t+1}(1)-\psi_{2t+1}(1) \\ \psi_{1t+1}(2)-\psi_{2t+1}(2) \\ \psi_{1t}(1)-\psi_{2t}(1)+\beta(\pi_{2t}(1)-\pi_{1t}(1))(\psi_{1t+1}(1)-\psi_{1t+1}(2)) \\ \psi_{1t}(2)-\psi_{2t}(2)+\beta(\pi_{2t}(2)-\pi_{1t}(2))(\psi_{1t+1}(1)-\psi_{1t+1}(2)) \end{bmatrix}
\tag{41}
$$

where:

$$
\begin{aligned}
a &= \beta\left(\pi_{2t+1}(1)-\pi_{1t+1}(1)\right)\\
b &= \beta\left(\pi_{2t+1}(2)-\pi_{1t+1}(2)\right)\\
c &= \beta\left(\pi_{2t}(1)-\pi_{1t}(1)\right)\\
d &= \beta^2\left(\pi_{2t}(1)-\pi_{1t}(1)\right)\left(\pi_{1t+1}(2)-\pi_{1t+1}(1)\right)\\
e &= \beta\left(\pi_{2t}(2)-\pi_{1t}(2)\right)-1\\
f &= \beta^2\left(\pi_{2t}(2)-\pi_{1t}(2)\right)\left(\pi_{1t+1}(2)-\pi_{1t+1}(1)\right)
\end{aligned}
$$

If the determinant of the left-hand-side matrix is non-zero, then we are exactly identified. The determinant of this matrix is:

$$\beta^2 \left[(\pi_{2t+1}(1) - \pi_{1t+1}(1))(\pi_{2t}(2) - \pi_{1t}(2)) - (\pi_{2t+1}(2) - \pi_{1t+1}(2))(\pi_{2t}(1) - \pi_{1t}(1)) \right]$$

It is then clear what role non-stationarity plays in identifying the normalization. Namely, if the state transitions are the same at t and $t+1$ then the expression is zero.

6 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. Our results show that in nonstationary settings, relatively few restrictions on the role of time dependence suffice to identify the utility flows in dynamic models of discrete choice optimization and dynamic games.

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. Alternatively, stable utility functions or exclusion restrictions can be used to recover both the underlying payoffs as well as counterfactual choice probabilities.

7 Appendix

Proof of Theorem 2.

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:

$$u^*_{jT}(x) \equiv u_{jT}(x) - u_{l(x,T),T}(x) + c_T(x) \tag{42}$$

and:

$$u^*_{jt}(x) = u_{jt}(x) + c_t(x) - u_{l(x,t)t}(x) + \sum_{\tau=t+1}^{T}\sum_{x_\tau=1}^{X}\beta^{\tau-t}\left[u^*_{1\tau}(x_\tau) - u_{1\tau}(x_\tau)\right]\left[\kappa^*_{\tau-1}(x_\tau|x_t,l(x,t)) - \kappa^*_{\tau-1}(x_\tau|x_t,j)\right] \tag{43}$$

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 (10), 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:

$$j = \arg\max_{k\in\{1,\ldots,J\}}\left\{u_{kt}(x) - u_{l(x,t)t}(x) + \epsilon_{kt} + \sum_{\tau=t+1}^{T}\sum_{x_\tau=1}^{X}\beta^{\tau-t}\left\{\begin{array}{c} u_{1\tau}(x_\tau)\left[\kappa^*_{\tau-1}(x_\tau|x_t,k) - \kappa^*_{\tau-1}(x_\tau|x_t,l(x,t))\right] \\ +\psi_k[p_\tau(x_\tau)]\kappa^*_{\tau-1}(x_\tau|x_t,k)\end{array}\right\}\right\} \tag{44}$$

From (43):

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

Substitute the second line into the maximand of (44) . Then $d_{jt} = 1$ is optimal if:

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

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 3. Substituting in for $v_{jt}(z_t) - v_{1t}(z_t)$ in (14) with the corresponding expression in (15) 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:

$$
u_{jt}(z_t) = \psi_1[p_t(z_t)] - \psi_j[p_t(z_t)] + \sum_{\tau=t+1}^{T}\sum_{z_\tau=1}^{Z}\beta^{\tau-t}\psi_1[p_\tau(z_\tau)]\left[\kappa^*_{\tau-1}(z_\tau|z_t,1) - \kappa^*_{\tau-1}(z_\tau|z_t,j)\right] \qquad (45)
$$

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^τ_1 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:

$$u_j = \Psi_j - \Psi_1 + \sum_{\tau=1}^{\infty} \beta^\tau \left(F_1 - F_j\right) F_1^{\tau-1} \Psi_1 = \Psi_j - \Psi_1 + \beta \left(F_1 - F_j\right) \left(\sum_{\tau=0}^{\infty} \beta^\tau F_1^\tau \right) \Psi_1 \tag{46}$$

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 (46) we obtain:

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

which proves the theorem. ■

Proof of Theorem 5. Denote $P^{(\sim n)}$ as a 2×2 matrix given by:

$$P^{(\sim n)} = \begin{bmatrix} p_{1t}^{(\sim n)}(x) & p_{2t}^{(\sim n)}(x) \\ p_{1t+1}^{(\sim n)}(x) & p_{2t+1}^{(\sim n)}(x) \end{bmatrix} \tag{47}$$

Noting that x_t provides all the relevant state variables expect for the choice of the competitors, define $U_2^{(i)}$ as:

$$U_2^{(n)} = \begin{bmatrix} U_2^{(n)}(1, x) \\ U_2^{(n)}(2, x) \end{bmatrix} \tag{48}$$

Finally, define A as:

$$A = \begin{bmatrix} \psi_{1t}\left[p_t^{(n)}(x)\right] - \psi_{2t}\left[p_t^{(n)}(x)\right] - \beta \sum_j \sum_{x_{1t+1}} p_{jt}^{(\sim n)} \psi_2 \left[p_{t+1}^{(n)}(j, 1, x_{1t+1})\right] f_t(x_{1t+1}|x) \\ \psi_1\left[p_{t+1}^{(n)}(x)\right] - \psi_2\left[p_{t+1}^{(n)}(x)\right] - \beta \sum_j \sum_{x_{1t+2}} p_{jt+1}^{(\sim n)} \psi_2 \left[p_{t+2}^{(n)}(j, 1, x_{1t+2})\right] f_{t+1}(x_{1t+2}|x) \end{bmatrix} \tag{49}$$

The system of equation is then:

$$P^{(\sim n)} U_2^{(n)} = A \tag{50}$$

Since by assumption the choice probabilities vary between t and $t+1$, the rank of $P^{(\sim n)}$ is two, implying we can invert $P^{(\sim n)}$ and solve for $U_2^{(n)}$. ■

Proof of Theorem 6. 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:

$$\tilde{p}_{jt}(x) = \int \prod_{k=1}^{J} 1 \left\{ \epsilon_{jt} - \epsilon_{kt} + \widetilde{v}_{jt}(x) - \widetilde{v}_{kt}(x) \right\} g\left(\epsilon_t\right) d\epsilon_t \tag{51}$$

But:

$$
\begin{aligned}
\widetilde{v}_{j\mathcal{T}}(x) - \widetilde{v}_{k\mathcal{T}}(x) &= u_{j\mathcal{T}}(x) - u_{k\mathcal{T}}(x) + \Delta_{j\mathcal{T}}(x) - \Delta_{k\mathcal{T}}(x) + \sum_{x'=1}^{X-1} \beta V_{\mathcal{T}+1}(x')\left[f_{j\mathcal{T}}(x'|x) - f_{k\mathcal{T}}(x'|x)\right] \\
&= \Delta_{j\mathcal{T}}(x) - \Delta_{k\mathcal{T}}(x) + v_{j\mathcal{T}}(x) - v_{k\mathcal{T}}(x) \\
&= \Delta_{j\mathcal{T}}(x) - \Delta_{k\mathcal{T}}(x) + \psi_k[p_{\mathcal{T}}(x)] - \psi_j[p_{\mathcal{T}}(x)] \qquad (52)
\end{aligned}
$$

Substituting (52) into (51) yields:

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

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

[3] 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.

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

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

[6] Arcidiacono, P. and R.A. Miller (2014): "Nonstationary Dynamic Models with Finite Dependence," working paper.

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

[8] 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.

[9] 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.

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

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

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

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

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

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

[16] 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.

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

[18] 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.

[19] 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.

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

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

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

[23] 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.