Identifying Dynamic Games with Entry/Exit Decisions*

Fabio A. Miessi Sanches†
University of São Paulo

Daniel Silva Junior‡
London School of Economics

Sorawoot Srisuma§
University of Surrey

February 4, 2014

Abstract

We show commonly assumed restrictions imposed on components of the profit functions, based on past and present entry/exit decisions, allow for a particular arrangement of the discounted payoffs that leads to new identification results subjected to some normalizations. We first show the fixed profit can be identified outside the dynamic programming problem without the knowledge of the discount factor and variable profit. Then we show the variable profit can also be just identified when the discount factor is known. It has also been assumed in the literature that, given an appropriate data set, the variable profit function can also be identified independently outside the dynamic model. In this case we provide a set of sufficient conditions for the identification of the discounting factor when the payoff function is known. Our identification strategy is constructive and can be used to construct corresponding estimators.

JEL Classification Numbers: C14, C25, C61

Keywords: Entry/Exit Games, Identification, Markovian Games.

*We are grateful to Emmanuel Guerre and Martin Pesendorfer for their advice and support. We also thank seminar participants at University of São Paulo, University of Southampton and University of Surrey for comments.

†*E-mail address*: fmiessi@gmail.com

‡*E-mail address*: d.silva-junior@lse.ac.uk

§*E-mail address*: s.srisuma@surrey.ac.uk

1 Introduction

Dynamic structural models provide a useful framework to study counterfactual experiments involving multiple economic agents over time. We consider a class of dynamic discrete games that generalize the single agent Markov decision models surveyed in Rust (1994), as reviewed by Aguirregabiria and Mira (2010) and Bajari, Hong and Nekipelov (2012). In order to perform counterfactual analysis, the primitives of the dynamic models first have to be estimated (and/or partly assumed). The primitives of the games in our framework consist of players' payoff functions, discounting factor and Markov transition law. Much of the literature has focused on the identification and estimation of the payoff functions while taking the other primitives as known or estimated outside the dynamic model.

The prototypical applications for dynamic games involve entry and exit decisions of firms.[1] A common feature of entry/exit games modeled in the literature decomposes the period payoff function into *fixed* and *variable* profits. For the purpose of our identification strategy, in this paper we broadly distinguish two types of profits by whether or not they are explicitly determined based on a firm's entry/exit decision in the previous period. Let $a_{it} = 1$ denote firm i entering the market in period t, and $a_{it} = 0$ denoting a firm exiting the market. In particular, we assume the fixed profit takes one of the two form, either: (i) a fixed entry cost of a new entrant (when $a_{it} = 1, a_{it-1} = 0$); or, (ii) a scrap value when a firm exits the market (when $a_{it} = 0, a_{it-1} = 1$). Our variable profit is defined to capture everything else. We do not consider separately the *fixed operating cost* that is traditionally part of the fixed profit function (incurs when $a_{it} = 1$ independently of a_{it-1}) as it is generally not separately identifiable from our nonparametric variable profit function.[2] Our silent treatment of the fixed operating cost plays a crucial role in interpreting our results, as we shall discuss further below.

The main contribution of our paper is to illustrate that reasonable assumptions, motivated from familiar features of entry and exit games, can lead to new and constructive identification results that can be used for estimation. There are three identification results in this paper. First, we show that the fixed profit can be identified, subject to the normalization of either the fixed cost or scrap value, outside the dynamic programming framework (independent of the value function) without any knowledge of the discounting factor and variable profit function. Second, in addition, if the payoff of an outside option is known, then the variable profit function can also be identified outside the

[1] For examples, see Aguirregabiria and Mira (2007), Beresteanu, Ellickson and Misra (2010), Collard-Wexler (2013), Dunne, Klimek, Roberts and Xu (2013), Fan and Xiao (2012), Gowrisankaran, Lucarelli, Schmidt-Dengler and Town (2010), Lin (2012), Pesendorfer and Schmidt-Dengler (2003), Ryan (2012), Sanches and Silva Jr (2013), Snider (2009) and Suzuki (2013).

[2] An example of a fixed operating cost is land (or equipment) rental that incurs for a firm who is *in* the market even if it is not *actively* producing in a particular period (i.e. $a_{it} = 1$). The operating cost is typically modeled in an empirical application as the intercept term in our variable profit function.

dynamic programming framework when the discounting factor is assumed. Finally, we show the discounting factor can also be identified given the knowledge of other primitives of the model. Our identification result for the discounting factor is not specific to games with entry and exit decisions.

In order to clarify the contribution of our identification results from the onset, we now discuss the implication of our normalization choices. We provide an interpretation based around the fixed operating cost of firms that is contained inside our variable profit function. Firstly, we normalize either the fixed cost or scrap value as they cannot be jointly identified when the fixed operating cost is treated nonparametrically (follows from Proposition 2 of Aguirregabiria and Suzuki (2013)).[3] Next, although the known outside option assumption is a popular identifying restriction employed in the identification literature (e.g. Magnac and Thesmar (2002), Bajari, Chernozhukov, Hong and Nekipelov (2009), Chiong, Galichon and Shum (2013)). However, as alluded in footnote 2, it is not necessarily a trivial assumption and can be interpreted as knowing the operating fixed cost. Lastly, the identification result for the discounting factor is relevant when *our* variable profit function can be identified outside the dynamic model so that the entire payoff function can then be identified independently of the discounting factor. We pursue this result with the motivation from the empirical literature, which often treats the variable profit *of an active firm* to be identifiable given availability of firms' data on prices and quantities, by building a demand system and solving a particular model of competition (also see Berry and Haile (2010,2012)).[4] Therefore we can also interpret the condition that our variable profit function can be identified independently of the discounting factor as a normalization of the fixed operating cost. The appropriateness of the normalizations needed for our results eventually depends on the empirical question of interest, and availability of relevant data. We refer the reader to Aguirregabiria and Suzuki (2013) for further discussions and examples on the interpretation of normalization choices in entry/exit games.

Our identification strategy also offers a new way to estimate games, nonparametrically or otherwise, with attractive features that mimic our identification results. Particularly it is possible to estimate the fixed profit that is invariant to the choice of the discounting factor and functional form of the variable profit function. Our identification results outside the dynamic programming framework implies that estimation can be done without the need to approximate or estimate the value function, which is generally considered the main source of computational cost in estimating dynamic

[3]Collard-Wexler (2013) uses specific information from the concrete industy to assign the value for the scrap cost. Otherwise most empirical applications normalize either the fixed cost or scrap value to zero. Examples that normalize the fixed costs include: Pakes et al. (2007), Ryan (2012), Sweeting (2011) and Igami (2012). The followings normalize scrap values: Aguirregabiria and Mira (2007), Dunne et al. (2013), Snider (2009) and Suzuki (2013).

[4]Empirical examples can be found in Beresteanu et al. (2010), Dunne et al. (2013), Gowrisankaran et al. (2010), Ryan (2012), Snider (2009), Sweeting (2013) and Suzuki (2013).

games.[5] Given appropriate data that can estimate the variable profit outside the dynamic model, we also offer an interesting choice to estimate the discount factor, which is typically assumed to be known in this literature.

Most known identification results for dynamic games, following Magnac and Thesmar (2002), take the discounting factor and Markov transition law (that includes the distribution of the private values) as given and focus on the identification of the entire payoff function (e.g. see Bajari et al. (2009) and Pesendorfer and Schmidt-Dengler (2008)).[6] The only other work we are aware of that uses the specific structure of an entry/exit game to study identification is the recent working paper of Aguirregabiria and Suzuki (2013), whose main concern is on identifying certain counterfactuals for the purpose of policy analysis rather than the model primitives. Our results complement theirs in exploiting particular exclusion restrictions that arise in entry/exit models in order to improve the inference of such games.

Our identification strategy for the payoff function exploits the nonparametric restrictions motivated by idiosyncratic payoffs received by firms who enter/exit a particular market. We show the infinite sum of expected discounted payoffs can then be written additively as sum of the fixed profit from the present period and a conditional expectation of some composite terms, consisting of the (differenced) variable profits from the same period and future discounted expected payoffs. Particularly the composite terms are integrated under the same conditional law under a conditional independence assumption, which describes the conditional choice probabilities that is observed from the data. We then employ a projection method to eliminate the composite term (analogous to partition regression (Frisch and Waugh (1933)), which allows us to identify the normalized fixed profit function independently of the variable profits and discounting factor. The variable profits are generally *over*-identified in our setting. We then use the composite terms, which are also identifiable when fixed profit is known, as additional restrictions that allow us to *just*-identify the variable profit function independently of the value function. We illustrate our identification strategy with the following example.

Example

Consider a simplified entry game based on the empirical application in Pesendorfer and Schmidt-Dengler (2003).[7] Each player i makes a decision, a_{it}, to play 1 (enter the market) or 0 (not enter) at

[5] Well known estimation methodologies in the literature include Aguirregabiria and Mira (2007), Bajari, Benkard and Levin (2007), Pakes, Ostrovsky and Berry (2007), and Pesendorfer and Schmidt-Dengler (2008).

[6] A notable exception is Norets and Tang (2012), who show in a single agent setting that without the distribution of the private values, payoff functions can only be partially identified.

[7] Our example is a slightly more general version of the Monte Carlo design in Pesendorfer and Schmidt-Dengler (2008, Section 7).

time t. The player's per period profit is determined by

$$u_i\left(a_{it}, a_{-it}, x_t, \varepsilon_{it}\right) = \pi_i\left(a_{it}, a_{-it}, x_t\right) - \varepsilon_{it} \cdot \mathbf{1}\left[a_{it} = 1\right],$$

where π_i is the systematic profit function, $x_t = \left(a_{1t-1}, a_{2t-1}\right)$ represents the actions of both players in the previous period and ε_{it} denotes the privately observed idiosyncratic profit shock that is independently drawn (across players, states and time) when she enters the market. The profit function can be decomposed further into variable profit, π_i^v, fixed cost, F_i, and scrap value, W_i:

$$\pi_i\left(a_{it}, a_{-it}, x_t\right) = \pi_i^v\left(a_{it}, a_{-it}\right) - F_i \cdot \mathbf{1}\left[a_{it} = 1, a_{it-1} = 0\right] + W_i \cdot \mathbf{1}\left[a_{it} = 0, a_{it-1} = 1\right].$$

Note that the variable profit, fixed cost and scrap value do not depend directly on past actions. In equilibrium each player chooses the optimal action, $\alpha_i\left(x_t, \varepsilon_{it}\right) \in \{0, 1\}$, that satisfies

$$\alpha_i\left(x_t, \varepsilon_{it}\right) \in \arg\max_{a_i \in \{0,1\}} \left\{ \begin{array}{c} E_{\sigma_i}\left[u_i\left(a_{it+\tau}, a_{-it+\tau}, x_{t+\tau}, \varepsilon_{it+\tau}\right) | x_t, a_{it} = a_i\right] \\ +\beta E_{\sigma_i}\left[m_{i,\sigma_i}\left(x_{t+1}\right) | x_t, a_{it} = a_i\right] \end{array} \right\},$$

where β denotes the discount factor, the conditional expectations are defined with respect to the equilibrium distribution of a_{-it} conditioning on x_t, which we denote by σ_i, and $m_{i,\sigma_i}\left(x_t\right)$ is the equilibrium expected discounted payoff for a given initial value of x_t, i.e.

$$\begin{aligned} m_{i,\sigma_i}\left(x_t\right) &= \sum_{\tau=t}^{\infty} \beta^{\tau-t} E_{\sigma_i}\left[u_i\left(a_{it+\tau}, a_{-it+\tau}, x_{t+\tau}, \varepsilon_{it+\tau}\right) | x_t\right] \\ &= E_{\sigma_i}\left[u_i\left(a_{it}, a_{-it}, x_t, \varepsilon_{it}\right) | x_t\right] + \beta E_{\sigma_i}\left[m_{i,\sigma_i}\left(x_{t+1}\right) | x_t\right]. \end{aligned} \tag{1}$$

The framework here implicitly assumes certain independence condition as well as a stationary solution concept of an equilibrium that we shall describe explicitly for the general game below. Henceforth we omit the dependence of σ_i for notational simplicity.

The equilibrium strategy is characterized by $\alpha_i\left(x_t, \varepsilon_{it}\right) = \mathbf{1}\left[\varepsilon_{it} \leq \Delta v_i\left(x_t\right)\right]$, where

$$\begin{aligned} \Delta v_i\left(x_t\right) &= E\left[\pi_i\left(a_{it}, a_{-it}, x_t\right) | x_t, a_{it} = 1\right] + \beta E\left[m_i\left(x_{t+1}\right) | x_t, a_{it} = 1\right] \\ &\quad - \left(E\left[\pi_i\left(a_{it}, a_{-it}, x_t\right) | x_t, a_{it} = 0\right] + \beta E\left[m_i\left(x_{t+1}\right) | x_t, a_{it} = 0\right]\right) \\ &= E\left[\pi_i^v\left(1, a_{-it}\right) - \pi_i^v\left(0, a_{-it}\right) | x_t\right] + \beta E\left[m_i\left(1, a_{-it}\right) - m_i\left(0, a_{-it}\right) | x_t\right] \\ &\quad - F_i \cdot \mathbf{1}\left[a_{it-1} = 0\right] - W_i \cdot \mathbf{1}\left[a_{it-1} = 1\right]. \end{aligned}$$

The key feature in the last equality is that the expectations above are taken under the same conditional law, of a_{-it} given x_t. Let $\Delta\pi_i^v\left(a_{-it}\right)$ denote $\pi_i^v\left(1, a_{-it}\right) - \pi_i^v\left(0, a_{-it}\right)$, and similarly, $\Delta m_i\left(a_{-it}\right)$ denote $m_i\left(1, a_{-it}\right) - m_i\left(0, a_{-it}\right)$. Then it follows for any $x_t = \left(a_i, a_{-i}\right)$ that

$$\begin{aligned} \Delta v_i\left(a_i, a_{-i}\right) &= P_{-i}\left(0 | a_i, a_{-i}\right)\left(\Delta\pi_i^v(0) + \beta\Delta m_i(0)\right) \\ &\quad + P_{-i}\left(1 | a_i, a_{-i}\right)\left(\Delta\pi_i^v(1) + \beta\Delta m_i(1)\right) \\ &\quad - F_i \cdot \mathbf{1}\left[a_i = 0\right] - W_i \cdot \mathbf{1}\left[a_i = 1\right], \end{aligned}$$

where we use $P_i(a_i|x)$ to denote $\Pr[\alpha_i(x_t, \varepsilon_{it}) = a_i | x_t = x]$. Then we can summarize all combinations of $\Delta v_i(a_i, a_{-i})$ using by a matrix equation,

$$\Delta \mathbf{v}_i = \mathbf{Z}_i \boldsymbol{\lambda}_i + \mathbf{d}_f F_i + \mathbf{d}_w W_i, \text{ representing} \tag{2}$$

$$\begin{bmatrix} \Delta v_i(0,0) \\ \Delta v_i(0,1) \\ \Delta v_i(1,0) \\ \Delta v_i(1,1) \end{bmatrix} = \begin{bmatrix} P_{-i}(0|0,0) & P_{-i}(1|0,0) \\ P_{-i}(0|0,1) & P_{-i}(1|0,1) \\ P_{-i}(0|1,0) & P_{-i}(1|1,0) \\ P_{-i}(0|1,1) & P_{-i}(1|1,1) \end{bmatrix} \begin{bmatrix} \lambda_{1i} \\ \lambda_{2i} \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix} F_i + \begin{bmatrix} 0 \\ 0 \\ -1 \\ -1 \end{bmatrix} W_i,$$

where, in particular, we have

$$\boldsymbol{\lambda}_i = \begin{bmatrix} \lambda_{1i} \\ \lambda_{2i} \end{bmatrix} = \begin{bmatrix} \Delta \pi_i^v(0) + \beta \Delta m_i(0) \\ \Delta \pi_i^v(1) + \beta \Delta m_i(1) \end{bmatrix}. \tag{3}$$

Let $\mathbf{M}_{\mathbf{Z}_i}$ be the orthogonal projection matrix whose null space is the column space of $\mathbf{Z}_i$, which we denote by $\mathcal{CS}(\mathbf{Z}_i)$. If $\mathbf{d}_f \notin \mathcal{CS}(\mathbf{Z}_i)$, then it follows that

$$F_i = \left(\mathbf{d}_f^\top \mathbf{M}_{\mathbf{Z}_i} \mathbf{d}_f\right)^{-1} \mathbf{d}_f^\top \mathbf{M}_{\mathbf{Z}_i} (\Delta \mathbf{v}_i - \mathbf{d}_w W_i), \tag{4}$$

and similarly, if $\mathbf{d}_w \notin \mathcal{CS}(\mathbf{Z}_i)$ then

$$W_i = \left(\mathbf{d}_w^\top \mathbf{M}_{\mathbf{Z}_i} \mathbf{d}_w\right)^{-1} \mathbf{d}_w^\top \mathbf{M}_{\mathbf{Z}_i} (\Delta \mathbf{v}_i - \mathbf{d}_f F_i). \tag{5}$$

Note that we do not need $\mathbf{Z}_i$ to have full column rank. Simple sufficient conditions for the expressions above for F_i and W_i to hold can be given in terms of variation in the entry decisions. For instance, constrained by the law of total probability, it is easy to verify that necessary conditions for $\mathbf{d}_f \in \mathcal{CS}(\mathbf{Z}_i)$ include $P_{-i}(1|0,0) = P_{-i}(1|0,1)$ and $P_{-i}(1|1,0) = P_{-i}(1|1,1)$, thus any violation will identify F_i; analogous conditions hold for W_i. Note that we require the distribution of ε_{it} in order for $\Delta \mathbf{v}_i$ to be recovered from the choice probabilities (Hotz and Miller's (1993) inversion). Therefore F_i (W_i) can be written explicitly in terms of choice probabilities when W_i (F_i) is normalized, independently of β and π_i^v.

Next, suppose that $\mathbf{Z}_i$ has full column rank, then

$$\boldsymbol{\lambda}_i = \left(\mathbf{Z}_i^\top \mathbf{Z}_i\right)^{-1} \mathbf{Z}_i^\top (\Delta \mathbf{v}_i - \mathbf{d}_f F_i - \mathbf{d}_w W_i). \tag{6}$$

The condition for full column rank of $\mathbf{Z}_i$ can again be interpreted in terms of variation in entry probabilities. This time the necessary and sufficient condition is that for $P_{-i}(1|a_i, a_{-i})$ to differ for some pair of (a_i, a_{-i}). We consider recovering π_i^v from λ_i (see (3)) when we, in addition, assume the outside option to be known. Particularly we impose $\pi_i^v(0, a_{-i}) = 0$ for all a_{-i}. With an abuse of notation we simply use $\pi_i^v(a_{-i})$ to denote they payoff of player i when she enters the market, i.e. $\pi_i^v(a_i, a_{-i}) = \pi_i^v(a_{-i}) \cdot \mathbf{1}[a_i = 1]$. Similarly, we let $m_i(a_{-i})$ denote $m_i(1, a_{-i})$, so that the relation in (3) can be re-written as

$$\begin{bmatrix} m_i(0,0) \\ m_i(0,1) \end{bmatrix} = \begin{bmatrix} m_i(0) + (\pi_i^v(0) - \lambda_{1i})/\beta \\ m_i(1) + (\pi_i^v(1) - \lambda_{2i})/\beta \end{bmatrix},$$

and $m_i(a_i, a_{-i}) = m_i(a_{-i}) + \frac{\pi_i^v(a_{-i}) - \lambda_{1i}\cdot\mathbf{1}[a_{-i}=0] - \lambda_{2i}\cdot\mathbf{1}[a_{-i}=1]}{\beta} \cdot \mathbf{1}[a_i = 0]$. We utilize the recursive form of the ex-ante value function, m_i (see (1)),

$$\begin{aligned} m_i(x_t) &= \zeta_i(x_t) + E[\pi_i^v(a_{it}, a_{-it})|x_t] + \beta E[m_i(x_{t+1})|x_t], \text{ where} \\ \zeta_i(x_t) &= E[-F_i \cdot \mathbf{1}[a_{it} = 1, a_{it-1} = 0] + W_i \cdot \mathbf{1}[a_{it} = 0, a_{it-1} = 1] - \varepsilon_{it} \cdot \mathbf{1}[a_{it} = 1]|x_t]. \end{aligned}$$

Incorporating the relationship involving λ_i into the equation above for m_i, yields on the LHS and then on the RHS respectively:

$$\begin{aligned} &m_i(a_{-it}) + \frac{\pi_i^v(a_{-it}) - \lambda_{1i}\cdot\mathbf{1}[a_{-it} = 0] - \lambda_{2i}\cdot\mathbf{1}[a_{-it} = 1]}{\beta} \cdot \mathbf{1}[a_{it} = 0], \text{ and} \\ &\zeta_i(x_t) + E[\pi_i^v(a_{-it})|x_t] + \beta E[m_i(a_{-it})|x_t] - \lambda_{1i} P_i(0|x_t) P_{-i}(0|x_t) - \lambda_{2i} P_i(0|x_t) P_{-i}(1|x_t), \end{aligned}$$

which can be vectorized across combinations of $x_t = (a_i, a_{-i})$ as follows,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} m_i(0) \\ m_i(1) \end{bmatrix} + \frac{1}{\beta} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \pi_i^v(0) \\ \pi_i^v(1) \end{bmatrix} - \frac{1}{\beta} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_{1i} \\ \lambda_{2i} \end{bmatrix}, \text{ and}$$

$$\begin{aligned} &\begin{bmatrix} \zeta_i(0,0) \\ \zeta_i(0,1) \\ \zeta_i(1,0) \\ \zeta_i(1,1) \end{bmatrix} + \begin{bmatrix} P_{-i}(0|0,0) & P_{-i}(1|0,0) \\ P_{-i}(0|0,1) & P_{-i}(1|0,1) \\ P_{-i}(0|1,0) & P_{-i}(1|1,0) \\ P_{-i}(0|1,1) & P_{-i}(1|1,1) \end{bmatrix} \begin{bmatrix} \pi_i^v(0) + \beta m_i(0) \\ \pi_i^v(1) + \beta m_i(1) \end{bmatrix} \\ &- \begin{bmatrix} P_i(0|0,0) P_{-i}(0|0,0) & P_i(0|0,0) P_{-i}(1|0,0) \\ P_i(0|0,1) P_{-i}(0|0,1) & P_i(0|0,1) P_{-i}(1|0,1) \\ P_i(0|1,0) P_{-i}(0|1,0) & P_i(0|1,0) P_{-i}(1|1,0) \\ P_i(0|1,1) P_{-i}(0|1,1) & P_i(0|1,1) P_{-i}(1|1,1) \end{bmatrix} \begin{bmatrix} \lambda_{1i} \\ \lambda_{2i} \end{bmatrix}. \end{aligned}$$

Combining them gives:

$$\begin{aligned}
\mathbf{e}_i &= \mathbf{B}_i\boldsymbol{\pi}_i^v + \mathbf{C}_i\mathbf{m}_i, \text{ where } \boldsymbol{\pi}_i^v = \left[\pi_i^v(0), \pi_i^v(1)\right]^\top, \mathbf{m}_i = \left[m_i(0), m_i(1)\right]^\top, \qquad (7)\\
\mathbf{B}_i &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} - \beta \begin{bmatrix} P_{-i}(0|0,0) & P_{-i}(1|0,0) \\ P_{-i}(0|0,1) & P_{-i}(1|0,1) \\ P_{-i}(0|1,0) & P_{-i}(1|1,0) \\ P_{-i}(0|1,1) & P_{-i}(1|1,1) \end{bmatrix},\\
\mathbf{C}_i &= \beta\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} - \beta \begin{bmatrix} P_{-i}(0|0,0) & P_{-i}(1|0,0) \\ P_{-i}(0|0,1) & P_{-i}(1|0,1) \\ P_{-i}(0|1,0) & P_{-i}(1|1,0) \\ P_{-i}(0|1,1) & P_{-i}(1|1,1) \end{bmatrix}\right),\\
\mathbf{e}_i &= \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} - \beta \begin{bmatrix} P_i(0|0,0)P_{-i}(0|0,0) & P_i(0|0,0)P_{-i}(1|0,0) \\ P_i(0|0,1)P_{-i}(0|0,1) & P_i(0|0,1)P_{-i}(1|0,1) \\ P_i(0|1,0)P_{-i}(0|1,0) & P_i(0|1,0)P_{-i}(1|1,0) \\ P_i(0|1,1)P_{-i}(0|1,1) & P_i(0|1,1)P_{-i}(1|1,1) \end{bmatrix}\right)\begin{bmatrix} \lambda_{1i} \\ \lambda_{2i} \end{bmatrix} + \beta \begin{bmatrix} \zeta_i(0,0) \\ \zeta_i(0,1) \\ \zeta_i(1,0) \\ \zeta_i(1,1) \end{bmatrix}.
\end{aligned}$$

Following the argument above, and assuming the knowledge of β, $\mathbf{B}_i, \mathbf{C}_i, \mathbf{e}_i$ are functions of the choice probabilities. By inspection $\mathbf{B}_i$ always have full column rank by the dominant diagonal theorem (applied to the first two rows of $\mathbf{B}_i$). Furthermore, by comparing the last two rows, the null space of $\mathbf{B}_i$ and $\mathbf{C}_i$ can only intersect at $\mathbf{0}$. Let $\mathbf{M}_{\mathbf{C}_i}$ be a projection matrix whose null space is $\mathcal{CS}(\mathbf{C}_i)$, then we have:

$$\boldsymbol{\pi}_i^v = \left(\mathbf{B}_i^\top\mathbf{M}_{\mathbf{C}_i}\mathbf{B}_i\right)^{-1}\mathbf{B}_i^\top\mathbf{M}_{\mathbf{C}_i}\mathbf{e}_i.$$

Therefore π_i^v is *just*-identified.

The identifying strategy employed in the Example can be generalized substantially. Our identification results are semiparametric in the sense that we do not impose any parametric structure on the profit function, however we require the distribution of the state variables in order to apply Hotz and Miller's inversion result.

The discounting factor is a primitive of the model that is generally assumed to be known for the purpose of identification in dynamic games. Perhaps relatedly, most estimation methodologies and empirical applications in the literature also treat β as a calibrating parameter and do not estimate them along with other primitives in the model. Motivated by the prospect of identifying the fixed profit independently of β, as well as the variable profit function (given an appropriate data set and other identifying assumption), we assume the entire payoff function to be known and show the discounting factor can also be identified given the knowledge of other primitives of the model. Our result shares similarities with Proposition 4 in Magnac and Thesmar (2002), who give conditions for a positive identification result of the discounting factor in a two-period model with a single decision maker when the period payoff function satisfies a particular exclusion restriction.[8] Although we do not explore a weaker condition for the identification of the discounting factor (by imposing additional structures on the payoff functions), however, our framework is arguably more complicated as it is set in the context of a dynamic game with an infinite time horizon.

The remainder of the paper is organized as follows. We define the theoretical model and modeling assumptions in Section 2. We give our identification results in Section 3. Section 4 concludes.

2 Model and Assumptions

We consider a game with I players, indexed by $i \in \mathcal{I} = \{1, \ldots, I\}$, over an infinite time horizon. The variables of the game in each period are action and state variables. The action set of each player is $A = \{0, 1\}$. Let $\mathbf{a}_t = (a_{1t}, \ldots, a_{It}) \in \mathbf{A} = \times_{i=1}^{I} A$. We will also occasionally abuse the notation and write $\mathbf{a}_t = (a_{it}, \mathbf{a}_{-it})$ where $\mathbf{a}_{-it} = (a_{1t}, \ldots, a_{i-1t}, a_{i+1t} \ldots, a_{It}) \in \mathbf{A} \backslash A$. Player i's information set is represented by the state variables $s_{it} \in S$, where $s_{it} = (x_t, \varepsilon_{it})$ such that $x_t \in X$ is common knowledge to all players and $\varepsilon_{it} \in \mathbb{R}$ denotes private information only observed by player i. We define $s_t = (x_t, \boldsymbol{\varepsilon}_t)$ and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{It})$. Future states are uncertain. Players' actions and states today affect future states. The evolution of the states is summarize by a Markov transition law $P(s_{t+1}|s_t, \mathbf{a}_t)$. Each player has a payoff function, $u_i : \mathbf{A} \times S \to \mathbb{R}$, which is time separable. Future period's payoffs are discounted at the rate $\beta \in (0, 1)$.

The setup described above and the following assumptions, which we shall assume throughout the paper, are standard in the modeling of dynamic discrete games.

Assumption M1 (Additive Separability): *For all i, $u_i(a_{it}, \mathbf{a}_{-it}, x_t, \varepsilon_{it}) = \pi_i(a_{it}, \mathbf{a}_{-it}, x_t) - \varepsilon_{it} \cdot \mathbf{1}[a_{it} = 1]$.*

Assumption M2 (Conditional Independence I): *The transition distribution of the states has the following factorization: $P(x_{t+1}, \boldsymbol{\varepsilon}_{t+1}|x_t, \boldsymbol{\varepsilon}_t, \mathbf{a}_t) = Q(\boldsymbol{\varepsilon}_{t+1}) G(x_{t+1}|x_t, \mathbf{a}_t)$, where Q is the cumulative distribution function of $\boldsymbol{\varepsilon}_t$ and G denotes the transition law of x_{t+1} conditioning on $\mathbf{a}_t$ and x_t.*

Assumption M3 (Independent Private Values): *The private information is independently distributed across players, and each is absolutely continuous with respect to the Lebesgue measure whose density is bounded on $\mathbb{R}^{K+1}$. So that $Q(\boldsymbol{\varepsilon}_t) = \prod_{i=1}^{I} Q_i(\varepsilon_{it})$, where Q_i denotes the cumulative distribution function of ε_{it}.*

[8] See Restriction R in Magnac and Thesmar (2002, page 809).

Assumption M4 (Discrete Public Values): *The support of* x_t *is finite so that* $X = \{x^1, \ldots, x^J\}$ *for some* $J < \infty$.

At time t every player observes s_{it}, each then chooses a_{it} simultaneously. We consider a Markovian framework where players' behaviors are stationary across time and players are assumed to play pure strategies. More specifically, for some $\alpha_i : S \to A$, $a_{it} = \alpha_i(s_{it})$ for all i, t, so that whenever $s_{it} = s_{i\tau}$ then $\alpha_i(s_{it}) = \alpha_i(s_{i\tau})$ for any τ. The beliefs are also time invariant. Player i's beliefs, σ_i, is a distribution of $\mathbf{a}_t = (\alpha_1(s_{1t}), \ldots, \alpha_I(s_{It}))$ conditional on x_t for some pure Markov strategy profile $(\alpha_1, \ldots, \alpha_I)$. The decision problem for each player is to solve, for any s_i,

$$\begin{aligned}\max_{a_i \in \{0,1\}} \{E[u_i(a_{it}, \mathbf{a}_{-it}, s_i) \,|s_{it} &= s_i, a_{it} = a_i] + \beta E[V_i(s_{it+1}) \,|s_{it} = s_i, a_{it} = a_i]\}, \qquad (8)\\ \text{where } V_i(s_i) &= \sum_{\tau=t}^{\infty} \beta^{\tau-t} E[u_i(\mathbf{a}_\tau, s_{i\tau}) \,|s_{it} = s_i].\end{aligned}$$

The expectation operators in the display above integrate out variables with respect to the probability distribution induced by the equilibrium beliefs and Markov transition law. V_i denotes the value function. Note that the transition law for future states is completely determined by the primitives and the beliefs. Any strategy profile that solves the decision problems for all i and is consistent with the beliefs satisfies is an equilibrium strategy. It is well-known that players' best responses are pure strategies almost surely and Markov perfect equilibria for games under M1 - M4 (e.g. see Aguirregabiria and Mira (2007) and Pesendorfer and Schmidt-Dengler (2008)). However, there may be multiple equilibria. In this paper we do not address issues associated with multiple equilibria and consider identification based on the joint distribution of the observables, namely (a_{it}, x_t, x_{t+1}), consistent with a single equilibrium play. The primitives of the game under this setting are $(\{\pi_i\}_{i=1}^I, \beta, Q, G)$. Throughout the paper we shall assume G and Q to be known.

We now formally introduce the specific structures for games with entry and exit decisions as assumptions. We assume N1 - N3 hold for the remainder of this section.

Assumption N1 (Decomposition of Observed States): $x_t = (x_t^v, \mathbf{a}_{t-1}) \in X = X^v \times \mathbf{A}$, *where the support of* x_t^v *has* $J^v < \infty$ *elements.*

Assumption N2 (Decomposition of Profits): $\pi_i(a_{it}, \mathbf{a}_{-it}, x_t^v, \mathbf{a}_{t-1}) = \pi_i^v(a_{it}, \mathbf{a}_{-it}, x_t^v) - F_i(x_t^v) \cdot \mathbf{1}[a_{it} = 1, a_{it-1} = 0] + W_i(x_t^v) \cdot \mathbf{1}[a_{it} = 0, a_{it-1} = 1]$.

Assumption N3 (Conditional Independence II): *The distribution of* x_{t+1}^v *conditional on* $\mathbf{a}_t$ *and* x_t^v *is independent of* $\mathbf{a}_{t-1}$.

Assumption N1 requires that entry decisions from the previous period are in the set of observed state variables. Assumption N2 assumes the period profit function can be decomposed into variable profit, fixed cost and scrap value. As explained in the introduction, we do not include the fixed operating cost in the fixed profit function, which is implicitly contained within π_i^v.[9] N2 also imposes an exclusion structure so that actions from the previous period cannot directly affect the period profit function beyond the indicator whether the firm is losing or receiving profit by entering or exiting the market. However, past actions can affect profit through the evolution of state variables and decisions on current period's action. Otherwise N2 is quite general as empirical works in the literature often do not even allow fixed cost and scrap value to depend on any state variables. Assumption N3 imposes that present period actions and (non-action) observable states are sufficient statistics for x_{t+1}^v, in particular past action values do not contain any additional information. Note that the restriction imposed in N3 is not implied by M2. We expect N1 - N3 to be reasonable in many empirical applications.

The following proposition says that some normalization has to be made in order to identify the fixed profit.

PROPOSITION 1: *For all i, x^v, $F_i(x^v)$ and $W_i(x^v)$ cannot be jointly identified.*

Since we do not specify π_i^v, our variable profit function is treated as nonparametric and includes the specification considered in Aguirregabiria and Suzuki (2013) as a special case. Our Proposition 1 then follows from the non-identification result in their paper (Proposition 2). However, if we assume the knowledge of either the fixed cost or scrap value then the other function can be nonparametrically identified outside the dynamic programming model if there is sufficient variation in the behavior of firms.

3 Main Results

We present our identification results first for the payoffs and then for the discounting factor.

3.1 Payoff Function

We first introduce some additional notations and representation lemmas.

[9] For example, Aguirregabiria and Suzuki (2013) specify:

$$\pi_i^v(a_{it}, \mathbf{a}_{-it}, x_t^v) = (\widetilde{\pi}_i^v(\mathbf{a}_{-it}, x_t^v) - f_i(x_t^v)) \cdot \mathbf{1}[a_{it} = 1],$$

so that $\widetilde{\pi}_i^v$ denotes the variable profit function net off the fixed operating cost, denoted by f_i. It is clear that $\widetilde{\pi}_i^v$ and f_i cannot be jointly identified nonparametrically without further restrictions.

We denote the ex-ante expected payoffs by $m_i\left(x_t^v,\mathbf{a}_{t-1}\right)=E\left[V_i\left(s_{it}\right)|x_t^v,\mathbf{a}_{t-1}\right]$, which is defined recursively as a solution to

$$m_i\left(x_t^v,\mathbf{a}_{t-1}\right)=E\left[\pi_i\left(\mathbf{a}_t,x_t^v,\mathbf{a}_{t-1}\right)|x_t^v,\mathbf{a}_{t-1}\right]-E\left[\varepsilon_{it}\cdot\mathbf{1}\left[a_{it}=1\right]|x_t^v,\mathbf{a}_{t-1}\right]+E\left[m_i\left(x_{t+1}^v,\mathbf{a}_t\right)|x_t^v,\mathbf{a}_{t-1}\right], \tag{9}$$

and the choice specific expected payoffs for choosing action a_i prior to adding the period unobserved state variable is

$$\begin{aligned} v_i\left(a_i,x_t^v,\mathbf{a}_{t-1}\right) &= E\left[\pi_i\left(a_{it},\mathbf{a}_{-it},x_t^v,\mathbf{a}_{t-1}\right)|a_{it}=a_i,x_t^v,\mathbf{a}_{t-1}\right] \\ &\quad +\beta E\left[m_i\left(x_{t+1}^v,\mathbf{a}_t\right)|a_{it}=a_i,x_t^v,\mathbf{a}_{t-1}\right], \end{aligned} \tag{10}$$

for any $a_i\in A$. Also define $\widetilde{m}_i\left(\mathbf{a}_t,x_t^v\right)=E\left[m_i\left(x_{t+1}^v,\mathbf{a}_t\right)|\mathbf{a}_t,x_t^v\right]$, and $\Delta v_i\left(x_t^v,\mathbf{a}_{t-1}\right)=v_i\left(1,x_t^v,\mathbf{a}_{t-1}\right)-v_i\left(0,x_t^v,\mathbf{a}_{t-1}\right)$. And, $\Delta\pi_i^v\left(\mathbf{a}_{-it},x_t^v\right)=\pi_i^v\left(1,\mathbf{a}_{-it},x_t^v\right)-\pi_i^v\left(0,\mathbf{a}_{-it},x_t^v\right)$, $\Delta\widetilde{m}_i\left(\mathbf{a}_{-it},x_t^v\right)=\widetilde{m}_i\left(1,\mathbf{a}_{-it},x_t^v\right)-\widetilde{m}_i\left(0,\mathbf{a}_{-it},x_t^v\right)$. For any matrix Z, we let $\mathcal{CS}\left(Z\right)$ denote the column space of Z, and let $\mathbf{M}_Z$ denote the orthogonal projection matrix, whose null space is the column space of Z, so that the range of $I-\mathbf{M}_Z$ is $\mathcal{CS}\left(Z\right)$.

Lemma 1: *Under M1 - M4 and N1 - N3, we have for all i:*

$$\Delta v_i\left(x_t^v,\mathbf{a}_{t-1}\right)=E\left[\lambda_i\left(\mathbf{a}_{-it},x_t^v\right)|\,x_t^v,\mathbf{a}_{t-1}\right]-F_i\left(x_t^v\right)\cdot\mathbf{1}\left[a_{it-1}=0\right]-W_i\left(x_t^v\right)\cdot\mathbf{1}\left[a_{it-1}=1\right], \tag{11}$$

where

$$\lambda_i\left(\mathbf{a}_{-it},x_t^v\right)=\Delta\pi_i^v\left(\mathbf{a}_{-it},x_t^v\right)+\beta\Delta\widetilde{m}_i\left(\mathbf{a}_{-it},x_t^v\right). \tag{12}$$

Lemma 1 says that the normalized choice specific expected payoffs can be decomposed into a sum of the fixed profits at time t and a conditional expectation of other composite terms, λ_i. In particular the conditional law for the expectation in (11), which is that of $\mathbf{a}_{-it}$ given x_t^v, is identifiable from the data. Since a conditional expectation operator is a linear operator, and $\mathbf{A}$ is a finite set with 2^I elements, we can then represent (11) by a matrix equation indexed by x_t^v. We state this as a lemma. We let $\Delta\mathbf{v}_i\left(x_t^v\right)$ denote a 2^I-dimensional vector of normalized expected discounted payoffs from entering, i.e. $\left\{\Delta v_i\left(x_t^v,\mathbf{a}\right)\right\}_{\mathbf{a}\in\mathbf{A}}$.

Lemma 2: *Under M1 - M4 and N1 - N3, we have for all i:*

$$\Delta\mathbf{v}_i\left(x_t^v\right)=\mathbf{Z}_i\left(x_t^v\right)\boldsymbol{\lambda}_i\left(x_t^v\right)+\mathbf{d}_fF_i\left(x_t^v\right)+\mathbf{d}_wW_i\left(x_t^v\right), \tag{13}$$

where $\mathbf{Z}_i\left(x_t^v\right)$ *is a* 2^I *by* 2^{I-1} *matrix of conditional probabilities* $\left\{\Pr\left[\mathbf{a}_{-it}=\mathbf{a}_{-i}|x_t^v,\mathbf{a}_{t-1}\right]\right\}_{\mathbf{a}_{-i}\in\mathbf{A}_{-i}}$, $\boldsymbol{\lambda}_i\left(x_t^v\right)=\left\{\lambda_i\left(\mathbf{a}_{-i},x_t^v\right)\right\}_{\mathbf{a}_{-i}\in\mathbf{A}_{-i}}$ *denotes a* 2^{I-1} *by* 1 *vector of* $\left\{\Delta\pi_i^v\left(\mathbf{a}_{-i},x_t^v\right)+\beta\Delta\widetilde{m}_i\left(\mathbf{a}_{-i},x_t^v\right)\right\}_{\mathbf{a}_{-i}\in\mathbf{A}_{-i}}$, $-\mathbf{d}_f$ *and* $-\mathbf{d}_w$ *are* 2^I *by* 1 *vectors of dummy variables such that* $\boldsymbol{\iota}+\mathbf{d}_f+\mathbf{d}_w=\mathbf{0}$.

Lemma 2 generalizes the representation for the flow profits in the Example (see equation (2)). The important structures of the terms on the RHS of (13) that are of practical interest are: (i) $\mathbf{Z}_i\left(x_t^v\right)$ is a matrix choice probabilities, (ii) $\boldsymbol{\lambda}_i\left(x_t^v\right)$ contains all the terms that are defined using the discounting factor and the entire profit function, and (iii) $-\mathbf{d}_f$ and $-\mathbf{d}_w$ are known vectors of dummy variables that sums to a vector of ones. We can then use the projection technique to isolate the fixed profits.

Theorem 1 (Identification of Fixed Profits): *Under M1 - M4 and N1 - N3, and for all i, x^v: (i) if $\mathbf{d}_f \notin \mathcal{CS}\left(\mathbf{Z}_i\left(x^v\right)\right)$ then,*

$$F_i\left(x^v\right)=\left(\mathbf{d}_f^{\top} \mathbf{M}_{\mathbf{Z}_i\left(x^v\right)} \mathbf{d}_f\right)^{-1} \mathbf{d}_f^{\top} \mathbf{M}_{\mathbf{Z}_i\left(x^v\right)}\left(\Delta \mathbf{v}_i\left(x^v\right)-\mathbf{d}_w W_i\left(x^v\right)\right); \quad (14)$$

and (ii) if $\mathbf{d}_w \notin \mathcal{CS}\left(\mathbf{Z}_i\left(x^v\right)\right)$ then ,

$$W_i\left(x^v\right)=\left(\mathbf{d}_w^{\top} \mathbf{M}_{\mathbf{Z}_i\left(x^v\right)} \mathbf{d}_w\right)^{-1} \mathbf{d}_w^{\top} \mathbf{M}_{\mathbf{Z}_i\left(x^v\right)}\left(\Delta \mathbf{v}_i\left(x^v\right)-\mathbf{d}_f F_i\left(x^v\right)\right). \quad (15)$$

Since $\Delta \mathbf{v}_i\left(x^v\right)$ can be identified from the data, our sufficient conditions for the nonparametric identification of either $F_i\left(x^v\right)$ or $W_i\left(x^v\right)$ require the knowledge of the other term, and also enough variation in enter/exit probabilities over past actions formalized by the requirement that $\mathcal{CS}\left(\mathbf{Z}_i\left(x^v\right)\right)$ does not contain either $\mathbf{d}_f$ or $\mathbf{d}_w$. As shown in the Example, sufficient condition that ensures $\mathbf{d}_f$ does not lie in $\mathcal{CS}\left(\mathbf{Z}_i\left(x^v\right)\right)$ can be interpreted as there being sufficient variation in the entry/exit probabilities past decisions. Note that we do not even require $\mathbf{Z}_i\left(x^v\right)$ to have full column rank in order to apply Theorem 1. In any case $\mathbf{Z}_i\left(x^v\right)$ is identified by the data, independently of any model primitives.

From Lemma 2, we can also identify $\boldsymbol{\lambda}_i$ if there is sufficient variation in the enter/exit decisions of firms. The precise condition for required can be given in terms of a testable rank condition of a matrix of choice probabilities.

Assumption R (Variation in Entry/Exit Decisions): *For all i, x^v, $\mathbf{Z}_i\left(x^v\right)$ has full column rank.*

Lemma 3: *Under M1 - M4, N1 - N3 and R, for all i, x^v, if $\mathbf{Z}_i\left(x^v\right)$ has full column rank then we have*

$$\boldsymbol{\lambda}_i\left(x^v\right)=\left(\mathbf{Z}_i\left(x^v\right)^{\top} \mathbf{Z}_i\left(x^v\right)\right)^{-1} \mathbf{Z}_i\left(x^v\right)^{\top}\left(\Delta \mathbf{v}_i\left(x^v\right)-\mathbf{d}_f F_i\left(x^v\right)-\mathbf{d}_w W_i\left(x^v\right)\right).$$

For the identification of the variable profit function, we introduce another normalization assumption.

Assumption N4 (Known Outside Option): *The variable profit for player i is normalized to zero if $a_{it} = 0$, i.e. $\pi_i^v(0, \mathbf{a}_{-it}, x_t^v) = 0$ for all i.*

For the remainder of this subsection we take the fixed profit function to be known and also assume N4 and R. We also define (with an abuse of notation): $\pi_i^v(1, \mathbf{a}_{-i}, x^v) = \pi_i^v(\mathbf{a}_{-i}, x^v)$ and $\widetilde{m}_i(1, \mathbf{a}_{-i}, x^v) = \widetilde{m}_i(\mathbf{a}_{-i}, x^v)$.

We are interested in identifying the primitives $\{\pi_i^v(\mathbf{a}_{-i}, x^v)\}_{(\mathbf{a}_{-i}, x^v) \in \mathbf{A}_{-i} \times X^v}$, consisting of $J^v 2^{I-1}$ elements. Existing identification results in the literature use the choice specific value function that relates the payoff functions with other observed or known objects through a system of linear equations to study identification (e.g. Pesendorfer and Schmidt-Dengler (2008) and Bajari et al. (2009)). Note that the number of equations is determined by the observable states, which equals $J^v 2^I$ in our case (from the support of $(\mathbf{a}_{t-1}, x_t^v)$). Therefore we generally expect π_i^v to be *over*-identified using existing methods. Instead, we continue to treat $\{\widetilde{m}_i(\mathbf{a}, x^v)\}_{(\mathbf{a}, x^v) \in \mathbf{A} \times X^v}$ as nuisance parameters, and recover π_i^v from λ_i , and the ex-ante value function (see (9)). In particular, although the ex-ante value function provides $J^v 2^I$ equations, there are $J^v 2^I$ additional unknowns coming from $\widetilde{m}_i$. However, we have $J^v 2^{I-1}$ linear restrictions imposed by the relationships between $\pi_i^v, \widetilde{m}_i$ and λ_i that allows us to *just* identify π_i^v. Lemma 4 states the aforementioned relationship (cf. (12)).

Lemma 4: *Under M1 - M4, N1 - N3 and R, we have for all i:*

$$\widetilde{m}_i(a_{it}, \mathbf{a}_{-it}, x_t^v) = \widetilde{m}_i(\mathbf{a}_{-it}, x_t^v) + \frac{(\pi_i^v(\mathbf{a}_{-it}, x_t^v) - \lambda_i(\mathbf{a}_{-it}, x_t^v))}{\beta} \cdot \mathbf{1}[a_{it} = 0]. \tag{16}$$

Lemma 5: *Under M1 - M4, N1 - N3 and R, we have for all i:*

$$\begin{aligned} \widetilde{m}_i(\mathbf{a}_t, x_t^v) \quad = \quad & E[E\left[\pi_i^v\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right) + \beta \widetilde{m}_i\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right) | x_{t+1}^v, \mathbf{a}_t\right] | x_t^v, \mathbf{a}_t] \\ & -E[E\left[\lambda_i\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right) \cdot \mathbf{1}[a_{it+1} = 0] | x_{t+1}^v, \mathbf{a}_t\right] | x_t^v, \mathbf{a}_t] \\ & +E\left[\zeta_i\left(x_{t+1}^v, \mathbf{a}_t\right) | x_t^v, \mathbf{a}_t\right], \text{ where} \end{aligned} \tag{17}$$

$\zeta_i\left(x_{t+1}^v, \mathbf{a}_t\right) = -E\left[F_i\left(x_{t+1}^v\right) \cdot \mathbf{1}[a_{it+1} = 1, a_{it} = 0] | x_{t+1}^v, \mathbf{a}_t\right] + E\left[W_i\left(x_{t+1}^v\right) \cdot \mathbf{1}[a_{it+1} = 0, a_{it} = 1] | x_{t+1}^v, \mathbf{a}_t\right]$ $-E\left[\varepsilon_{it+1} \cdot \mathbf{1}[a_{it+1} = 1] | x_{t+1}^v, \mathbf{a}_t\right]$.

We see that λ_i and ζ_i, as well as the conditional laws that define the conditional expectations in equations (16) and (17) are identified from the choice transition probabilities. Furthermore, as previously, we can represent the conditional expectations as matrices since the support of $(\mathbf{a}_t, x_t^v)$ is finite (cf. Lemma 2). The next theorem then follows from equating the two equations in Lemmas 4 and 5. Let $\boldsymbol{\pi}_i^v$ denote $\{\boldsymbol{\pi}_i^v(\mathbf{a}_{-i}, x^v)\}_{(\mathbf{a}_{-i}, x^v) \in \mathbf{A}_{-i} \times X^v}$, a $J 2^{I-1}$ by 1 vector of primitives of interest, and similarly $\widetilde{\mathbf{m}}_i = \{\widetilde{\mathbf{m}}_i(\mathbf{a}_{-i}, x^v)\}_{(\mathbf{a}_{-i}, x^v) \in \mathbf{A}_{-i} \times X^v}$.

Theorem 2 (Identification of Variable Profits): *Under M1 - M4, N1 - N3 and R, there exist matrices* $\mathbf{B}_i$ *and* $\mathbf{C}_i$*,and a vector* $\mathbf{e}_i$ *whose elements are explicit functions of* β*, choice and transition probabilities such that for all* i*:*

$$\mathbf{B}_i \boldsymbol{\pi}_i^v + \mathbf{C}_i \widetilde{\mathbf{m}}_i = \mathbf{e}_i. \tag{18}$$

Furthermore, $\mathbf{B}_i^\top \mathbf{M}_{\mathbf{C}_i} \mathbf{B}_i$ *is non-singular so that*

$$\boldsymbol{\pi}_i^v = \left(\mathbf{B}_i^\top \mathbf{M}_{\mathbf{C}_i} \mathbf{B}_i\right)^{-1} \mathbf{B}_i^\top \mathbf{M}_{\mathbf{C}_i} \mathbf{e}_i.$$

The exact forms for $\mathbf{B}_i, \mathbf{C}_i$ *and* $\mathbf{e}_i$ *are given in the Appendix.*

Equation (18) of Theorem 2 generalizes the contents in display (7) of the Example. The interesting feature for this result is that we can *just*-identified π_i^v independently of the future expected payoffs ($\widetilde{m}_i$). This particular approach may be useful in practice if we parameterize π_i^v, and the state space is large. In particular, we can still use (18) to form estimating equations based on the $J^v 2^{I-1}$ restrictions, which is half the size of the original state space. In contrast, the standard methodologies for games that we cite in the introduction will typically require an approximation or estimation of the value function with $J^v 2^I$ states.

3.2 Discounting Factor

When the variable profit function can be identified outside the dynamic model, our Theorem 1 implies it is possible to identify the entire payoff function without the knowledge of β. We now consider the inference on β and take all other primitives of the model as known (i.e., keep $(\{\pi_i\}_{i=1}^I, Q, G)$ fixed). The result in this section is not specific to games with entry/exit decisions, hence we do not need to distinguish past entry decisions in the state variable, and each player can have more than two action choice. Therefore in this section we do not impose Assumptions N1 - N3.

The parameter space for the model is now $\mathfrak{B} \subseteq (0,1)$ and we are interested in the discounting factor that is consistent with the data generating process, which we denote by β_0. We begin with an updated expression for the choice specific expected payoffs for choosing action a_i prior to adding the period unobserved state variable, where we now explicitly denote the dependence on the parameter β, so that for any a_i (cf. (10)):

$$v_i(a_i, x_t; \beta) = E\left[\pi_i(a_{it}, \mathbf{a}_{-it}, x_t) \,|a_{it} = a_i, x_t\right] + \beta g_i(a_i, x_t; \beta), \tag{19}$$

where $g_i(a_i, x_t; \beta) = E\left[V_i(s_{it+1}; \beta) \,|a_{it} = a_i, x_t\right]$, and $V_i(s_{it}; \beta) = \sum_{\tau=t}^{\infty} \beta^{\tau-t} E\left[u_i(\mathbf{a}_\tau, s_{i\tau}) \,|s_{it} = s_i\right]$. Note that the expectations are taken with respect to the observed choice and transition probabilities that are consistent with β_0. We can normalize the terms in (19) with action 0, so that:

$$\Delta v_i\left(x_t;\beta\right)=E\left[\Delta\pi_i\left(\mathbf{a}_{-it},x_t\right)|x_t\right]+\beta\Delta g_i\left(x_t;\beta\right), \tag{20}$$

where $\Delta v_i\left(x_t;\beta\right) = v_i\left(1,x_t;\beta\right)-v_i\left(0,x_t;\beta\right), \Delta\pi_i\left(\mathbf{a}_{-it},x_t\right) = \pi_i\left(1,\mathbf{a}_{-it},x_t\right)-\pi_i\left(0,\mathbf{a}_{-it},x_t\right)$ and $\Delta g_i\left(x_t;\beta\right) = g_i\left(1,x_t;\beta\right)-g_i\left(0,x_t;\beta\right)$. Since $\Delta v_i\left(x_t;\beta_0\right)$ is identified from the data, we take any given β to be a *structure* of the (pseudo-)model, and an implied expected payoffs, denoted by $\mathcal{V}_\beta = \left\{\Delta v_i\left(x_t;\beta\right)\right\}_{i\in\mathcal{I}}$, to be a *reduced form.*[10,11] We can then define identification using the notion of observational equivalence in terms of the expected payoffs (cf. Magnac and Thesmar (2002)).

Definition I1 (Observational Equivalence): Any distinct β and β' in $\mathfrak{B}$ are *observationally equivalent* if and only if $\mathcal{V}_\beta=\mathcal{V}_{\beta'}$.

Definition I2 (Identification): An element in $\mathfrak{B}$, say β, is *identified* if and only if β' and β are not observationally equivalent for all $\beta'\neq\beta$ in $\mathfrak{B}$.

Since $E\left[\pi_i\left(a_{it},\mathbf{a}_{-it},x_t\right)|a_{it}=a_i,x_t\right]$ does not depend on β, identification only depends on $\beta\Delta g_i\left(x_t;\beta\right)$. The following lemma expresses $\Delta g_i\left(x_t;\beta\right)$ in terms of β and other components that can be identified from the choice and transition probabilities. In what follows, we let: $\Delta H_i(x_t)$ denote a J by 1 vector of $\left\{\Pr\left[x_{t+1}=x|x_t,a_{it}=1\right]-\Pr\left[x_{t+1}=x|x_t,a_{it}=0\right]\right\}_{x\in X}$, $\mathbf{L}$ be a J by J stochastic matrix of transition probabilities of x_{t+1} conditioning on x_t, $\mathbf{R}$ is a J by J matrix of conditional choice probabilities such that $\mathbf{R}\boldsymbol{\pi}_i$ represents a J by 1 vector of $\left\{E\left[\pi_i\left(\mathbf{a}_t,x_t\right)|x_t=x\right]\right\}_{x\in X}$, and $\Delta\underline{v}_i\left(x_t;\beta_0\right)=\Delta H_i(x_t)\times\left(I-\beta\mathbf{L}\right)^{-1}\underline{\mathbf{r}}_i$ where $\underline{\mathbf{r}}_i$ represents a J by 1 vector of $\left\{E\left[\sum_{a'\in A}\varepsilon_{it}\left(a'\right)\mathbf{1}\left[a_{it}=a'\right]|x_t=x\right]\right\}_{x\in X}$.

Lemma 6: *Under M1 - M4, we have for all i:*

$$\Delta g_i\left(x_t;\beta\right)=\Delta H_i(x_t)\times\left(I-\beta\mathbf{L}\right)^{-1}\mathbf{R}\boldsymbol{\pi}_i+\Delta\underline{v}_i\left(x_t;\beta_0\right). \tag{21}$$

Note that hence $\Delta v_i\left(x_t;\beta_0\right)$, and $\Delta\underline{v}_i\left(x_t;\beta_0\right)$, are also identifiable from the observed data. Therefore β_0 is identifiable if for any $\beta\neq\beta'$, there exists some i such that $\beta\Delta g_i\left(x_t;\beta\right)\neq\beta'\Delta g_i\left(x_t;\beta'\right)$ with a positive probability. The relation in (21) can be written in a matrix for across possible values of x_t. Let $\Delta\mathbf{H}_i$ denote $\left[\Delta H_i(x^1)^\top,\ldots,\Delta H_i(x^J)^\top\right]^\top$, a J by J matrix, and $\Delta\mathbf{g}_i\left(\beta\right)$ denote $\left\{\Delta g_i\left(x;\beta\right)\right\}_{x\in X}$, a J by 1 vector, and similarly $\Delta\underline{\mathbf{v}}_i\left(\beta\right)$ denotes $\left\{\Delta\underline{v}_i\left(x;\beta\right)\right\}_{x\in X}$.

Lemma 7: *Under M1 - M4, we have for all i:*

$$\Delta\mathbf{g}_i\left(\beta\right)=\Delta\mathbf{H}_i\times\left(I-\beta\mathbf{L}\right)^{-1}\mathbf{R}\boldsymbol{\pi}_i+\Delta\underline{\mathbf{v}}_i\left(\beta_0\right). \tag{22}$$

[10] This is a pseudo-model in the sense that we only work with the equilibrium beliefs that generate the data.

[11] It is equivalent to define the reduced forms in terms of expected payoffs is equivalent to defining them in terms of conditional choice probabilities; following Hotz and Miller's (1993) inversion theorem.

Therefore β_0 is identified if and only if there is no other β in $\mathfrak{B}$ such that $\beta \Delta \mathbf{H}_i \times (I - \beta \mathbf{L})^{-1} \mathbf{R} \boldsymbol{\pi}_i$ $= \beta' \Delta \mathbf{H}_i \times (I - \beta' \mathbf{L})^{-1} \mathbf{R} \boldsymbol{\pi}_i$. The conditions in our next theorem is sufficient for β_0 to be identified.

Theorem 3 (Identification of Discounting Factor): *Under M1 - M4, if* $\mathbf{R}\boldsymbol{\pi}_i \neq \mathbf{0}$ *and* $\Delta \mathbf{H}_i$ *is invertible for some* i*, then* β_0 *is identified.*

4 Concluding Remarks

It is well known that (point) identification of the payoff functions in dynamic games generally relies on strong assumptions of other primitives of the model. We show the decomposition of the payoff functions based on past actions, as often seen in empirical work, provides natural restrictions that allow us to identify either the fixed cost or scrap value independently of the discounting factor and other profits under weak conditions. Particularly, our identification strategy is based on a re-arrangement from the usual representation of the choice specific value function that allows us isolate the fixed profit from the variable profit and future expected payoffs. We then exploit the same equation to provide a condition to *just*-identify the variable profit function that can be useful in practice as it relies on solving linear equations with half the size of the state space. Otherwise, we emphasize that existing identification results can also be used for the identification of the variable profit, as we rely on the same knowledge of other primitives of the game. We also derive an identification result for the discounting factor, which is quite rare in the literature, behind the motivation that it is possible to identify the entire profit function outside the dynamic model.

The idea behind Theorem 1 is quite general. It can be extended to games with multiple action choice and sequential game with an entry decision in the first step (e.g. dynamic oligopoly models) with more notations. Our Theorems 2 and 3 are more specific to entry games, particularly they make use of an additional assumption that we interpret as normalizing the fixed operating cost. Therefore the relevance of Theorems 2 and 3 depends on specific empirical question and availability of data.

Our conditions for identification can be used to form estimating equations. More specifically, equations (13), (18) and (20) can be used to nonparametrically estimate either the fixed cost or scrap value, variable profit and discounting factor respectively. The same equations can also be used for parametric estimation, for instance, by parameterizing the primitive functions and form the criterion functions for asymptotic least squares estimators for dynamic games in the same spirit as Pesendorfer and Schmidt-Dengler (2008) and Sanches, Silva and Srisuma (2013).

Our results also immediately accommodate more general models with unobserved heterogeneity as long as the choice and transition probabilities can be nonparametrically identified (Kasahara and

Shimotsu (2009)). And we expect the idea behind our identification results to be valid more generally when the observed state variables contains continuously distributed variables, by replacing various matrices with linear operators. However, sufficient conditions for identification become harder to check. Furthermore, the estimation problem also becomes more complicated as it involves estimating infinite dimensional parameters (e.g. see Bajari et al. (2009), and Srisuma and Linton (2012)). We leave the formal extensions in these directions for future research.

Appendix

Proofs of Lemmas

Proof of Lemma 1: Using the law of iterated expectation, under M3 $E\left[V_i\left(s_{it+1}\right)|a_{it}=a_i,x_t^v,\mathbf{a}_{t-1}\right]=E\left[m_i\left(x_{t+1}^v,\mathbf{a}_t\right)|a_{it}=a_i,x_t^v,\mathbf{a}_{t-1}\right]$, which simplifies further, after another application of the law of iterated expectation and N3, to $E\left[\widetilde{m}_i\left(a_i,\mathbf{a}_{-it},x_t^v\right)|x_t^v,\mathbf{a}_{t-1}\right]$. The remainder of the proof of Lemma 1 then follows from the definitions of the terms defined above.■

Proof of Lemma 2: Immediate.■

Proof of Lemma 3: Immediate.■

Proof of Lemma 4: Immediately from re-arranging equation (12).■

Proof of Lemma 5: By definition of the ex-ante value, equation (9), expressed in its recursive form, we have at time $t+1$

$$\begin{aligned}
m_i\left(x_{t+1}^v,\mathbf{a}_t\right) &= E\left[\pi_i^v\left(\mathbf{a}_{t+1},x_{t+1}^v\right)|x_{t+1}^v,\mathbf{a}_t\right]+\beta E\left[m_i\left(x_{t+2}^v,\mathbf{a}_{t+1}\right)|x_{t+1}^v,\mathbf{a}_t\right]+\zeta_i\left(x_{t+1}^v,\mathbf{a}_t\right), \text{ where} \\
\zeta_i\left(x_{t+1}^v,\mathbf{a}_t\right) &= -E\left[F_i\left(x_{t+1}^v\right)\cdot\mathbf{1}\left[a_{it+1}=1,a_{it}=0\right]+W_i\left(x_{t+1}^v\right)\cdot\mathbf{1}\left[a_{it+1}=0,a_{it}=1\right]|x_{t+1}^v,\mathbf{a}_t\right] \\
&\quad -E\left[\varepsilon_{it+1}\cdot\mathbf{1}\left[a_{it+1}=1\right]|x_{t+1}^v,\mathbf{a}_t\right].
\end{aligned} \tag{23}$$

Then taking condition expectation of $m_i\left(x_{t+1}^v,\mathbf{a}_t\right)$, conditioning on $\mathbf{a}_t,x_t^v$. An application of the law of iterated expectation, under N3, gives $E\left[m_i\left(x_{t+2}^v,\mathbf{a}_{t+1}\right)|x_{t+1}^v,\mathbf{a}_t\right]=E\left[\widetilde{m}_i\left(\mathbf{a}_{t+1},x_{t+1}^v\right)|x_{t+1}^v,\mathbf{a}_t\right]$. Then using (16), and substituting these two terms into gives the result.■

Proof of Lemma 6: Immediate.■

Proof of Lemma 7: Immediate.■

Proofs of Theorems

Proof of Theorem 1: Follows from Lemma 2 using straightforward algebra.■

Proof of Theorem 2: First we equate (16) and (17), which yields

$$\begin{aligned}
&\beta\widetilde{m}_i\left(\mathbf{a}_{-it},x_t^v\right)+\left(\pi_i^v\left(\mathbf{a}_{-it},x_t^v\right)-\lambda_i\left(\mathbf{a}_{-it},x_t^v\right)\right)\cdot\mathbf{1}\left[a_{it}=0\right] \\
=\ &\beta E[E\left[\pi_i^v\left(\mathbf{a}_{-it+1},x_{t+1}^v\right)+\beta\widetilde{m}_i\left(\mathbf{a}_{-it+1},x_{t+1}^v\right)|x_{t+1}^v,\mathbf{a}_t\right]|x_t^v,\mathbf{a}_t] \\
&-\beta E[E\left[\lambda_i\left(\mathbf{a}_{-it+1},x_{t+1}^v\right)\cdot\mathbf{1}\left[a_{it+1}=0\right]|x_{t+1}^v,\mathbf{a}_t\right]|x_t^v,\mathbf{a}_t]+E\left[\zeta_i\left(x_{t+1}^v,\mathbf{a}_t\right)|x_t^v,\mathbf{a}_t\right].
\end{aligned}$$

Collecting terms and re-arranging leads to

$$\begin{aligned}
&\left[\begin{array}{c}\left\{\pi_i^v\left(\mathbf{a}_{-it}, x_t^v\right)\cdot\mathbf{1}\left[a_{it}=0\right]-\beta E[E\left[\pi_i^v\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right)|x_{t+1}^v, \mathbf{a}_t\right]|x_t^v, \mathbf{a}_t]\right\} \\ +\left\{\beta\widetilde{m}_i\left(\mathbf{a}_{-it}, x_t^v\right)-\beta^2 E[E\left[\widetilde{m}_i\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right)|x_{t+1}^v, \mathbf{a}_t\right]|x_t^v, \mathbf{a}_t]\right\}\end{array}\right] \\
=&\left[\begin{array}{c}\left\{\lambda_i\left(\mathbf{a}_{-it}, x_t^v\right)\cdot\mathbf{1}\left[a_{it}=0\right]-\beta E[E\left[\lambda_i\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right)\cdot\mathbf{1}\left[a_{it+1}=0\right]|x_{t+1}^v, \mathbf{a}_t\right]|x_t^v, \mathbf{a}_t]\right\} \\ +E\left[\zeta_i\left(x_{t+1}^v, \mathbf{a}_t\right)|x_t^v, \mathbf{a}_t\right]\end{array}\right],
\end{aligned}$$

which can be represented in a linear functional notation (cf. (7)):

$$\begin{aligned}
\mathcal{B}_i\pi_i^v+\mathcal{C}_i\widetilde{m}_i &= \eta_i, \text{ where for any } a_i, \mathbf{a}_{-i}, x^v \\
\mathcal{B}_i\pi_i^v\left(a_i, \mathbf{a}_{-i}, x^v\right) &= \pi_i^v\left(\mathbf{a}_{-i}, x^v\right)\cdot\mathbf{1}\left[a_i=0\right]-\beta E[E\left[\pi_i^v\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right)|x_{t+1}^v, \mathbf{a}_t\right]|x_t^v=x^v, \mathbf{a}_t=\left(a_i, \mathbf{a}_{-i}\right)] \\
\mathcal{C}_i\widetilde{m}_i\left(a_i, \mathbf{a}_{-i}, x^v\right) &= \beta\widetilde{m}_i\left(\mathbf{a}_{-i}, x^v\right)-\beta^2 E[E\left[\widetilde{m}_i\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right)|x_{t+1}^v, \mathbf{a}_t\right]|x_t^v=x^v, \mathbf{a}_t=\left(a_i, \mathbf{a}_{-i}\right)] \\
\eta_i\left(a_i, \mathbf{a}_{-i}, x^v\right) &= \lambda_i\left(x^v;\mathbf{a}_{-i}\right)\cdot\mathbf{1}\left[a_i=0\right] \\
&\quad -\beta E[E\left[\lambda_i\left(x_{t+1}^v;\mathbf{a}_{-it+1}\right)\cdot\mathbf{1}\left[a_{it+1}=0\right]|x_{t+1}^v, \mathbf{a}_t\right]|x_t^v=x^v, \mathbf{a}_t=\left(a_i, \mathbf{a}_{-i}\right)] \\
&\quad +E\left[\zeta_i\left(x_{t+1}^v, \mathbf{a}_t\right)|x_t^v=x^v, \mathbf{a}_t=\left(a_i, \mathbf{a}_{-i}\right)\right].
\end{aligned}$$

Note that $\mathcal{B}_i$ and $\mathcal{C}_i$ are linear operators. Since the support of $\mathbf{a}_t$ and x_t^v are finite, $\mathcal{B}_i\pi_i^v+\mathcal{C}_i\widetilde{m}_i=\eta_i$ can be equivalently represented in a matrix form. For instance,

$$E[E\left[\pi_i^v\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right)|x_{t+1}^v, \mathbf{a}_t\right]|x_t^v=x^v, \mathbf{a}_t=\left(a_i, \mathbf{a}_{-i}\right)]=\left[\begin{array}{c}\mathbf{H}_i^1\mathbf{P}_i^{\#\mathbf{a}}\boldsymbol{\pi}_i^v \\ \vdots \\ \mathbf{H}_i^{\#\mathbf{a}}\mathbf{P}_i^{\#\mathbf{a}}\boldsymbol{\pi}_i^v \\ \vdots \\ \mathbf{H}_i^{2^I}\mathbf{P}_i^{2^I}\boldsymbol{\pi}_i^v\end{array}\right]=\mathbf{H}_i\mathbf{P}_i\boldsymbol{\pi}_i^v.$$

where $\mathbf{P}_i^{\#\mathbf{a}}$ is a J by $J2^{I-1}$ matrix of conditional probabilities of $\mathbf{a}_{-it+1}$ given $\left(x_t^v, \mathbf{a}_t\right)$ so that $\mathbf{P}_i^{\#\mathbf{a}}\boldsymbol{\pi}_i^v$ gives $\left\{E\left[\boldsymbol{\pi}_i^v\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right)|x_{t+1}=x^v, \mathbf{a}_t=\#\mathbf{a}\right]\right\}_{x^v\in X^v}$ for some action profile that we enumerate as $\#\mathbf{a}$, and $\mathbf{H}_i^{\#\mathbf{a}}$ is a J by J matrix of conditional probabilities of x_{t+1}^v given $\left(x_t^v, \mathbf{a}_t\right)$ so that $\mathbf{H}_i^{\#\mathbf{a}}\mathbf{P}_i^{\#\mathbf{a}}\boldsymbol{\pi}_i^v$ gives $\left\{E\left[E\left[\boldsymbol{\pi}_i^v\left(\mathbf{a}_{-it+1}, x_{t+1}^v\right)|x_{t+1}, \mathbf{a}_t\right]|x_t=x^v, \mathbf{a}_t=(\#)\,\mathbf{a}\right]\right\}_{x^v\in X^v}$. Therefore, by inspection, we obtain equation (18), namely

$$\mathbf{B}_i\boldsymbol{\pi}_i^v+\mathbf{C}_i\widetilde{\mathbf{m}}_i=\mathbf{e}_i,$$

where $\mathbf{B}_i=\mathbf{D}_i-\beta\mathbf{H}_i\mathbf{P}_i, \mathbf{C}_i=\beta\left(\mathbf{I}-\beta\mathbf{H}_i\mathbf{P}_i\right)$, $\mathbf{I}$ is an identity matrix of order $J2^I$, $\mathbf{D}_i$ is a matrix of 0s and 1s, and $\mathbf{e}_i$ is a vector of terms identifiable directly from the data. Note that $\mathbf{H}_i\mathbf{P}_i$ is a (right stochastic) matrix of conditional probabilities of size $J2^I$ by $J2^{I-1}$. Then it follows that $\mathbf{B}_i$ and $\mathbf{C}_i$ both have full column rank (each with rank $J2^{I-1}$) by dominant diagonal theorem (also applies for $\mathbf{B}_i$, particularly to the sub-block of $\left(\mathbf{D}_i-\beta\mathbf{H}_i\mathbf{P}_i\right)$, of size $J2^{I-1}$, that corresponds to $\left(a_{it}, \mathbf{a}_{-it}\right)=\left(0, \mathbf{a}_{-i}\right)$).

Finally the null space of $\mathbf{B}_i$ and $\mathbf{C}_i$ only intersect at $\mathbf{0}$ due to the presence of $\beta \in (0,1)$, therefore $\mathbf{B}_i^\top \mathbf{M}_{\mathbf{C}_i} \mathbf{B}_i$ is invertible, and projecting out $\widetilde{\mathbf{m}}_i$ gives $\boldsymbol{\pi}_i^v = \left(\mathbf{B}_i^\top \mathbf{M}_{\mathbf{C}_i} \mathbf{B}_i\right)^{-1} \mathbf{B}_i^\top \mathbf{M}_{\mathbf{C}_i} \mathbf{e}_i$.■

PROOF OF THEOREM 3: Take any $\beta, \beta' \in (0,1)$ such that $\beta \neq \beta'$. We have from Lemma 7:

$$\beta \Delta \mathbf{g}_i(\beta) - \beta' \Delta \mathbf{g}_i(\beta') = \left(\beta \Delta \mathbf{H}_i \times (I - \beta \mathbf{L})^{-1} - \beta' \Delta \mathbf{H}_i \times (I - \beta' \mathbf{L})^{-1}\right) \mathbf{R} \boldsymbol{\pi}_i.$$

We consider the terms in the parenthesis on the RHS of the equation above,

$$\begin{aligned}
& \beta \Delta \mathbf{H}_i \times (I - \beta \mathbf{L})^{-1} - \beta' \Delta \mathbf{H}_i \times (I - \beta' \mathbf{L})^{-1} \\
= \; & (\beta - \beta') \Delta \mathbf{H}_i \times (I - \beta \mathbf{L})^{-1} + \beta' \Delta \mathbf{H}_i \left((I - \beta \mathbf{L})^{-1} - (I - \beta' \mathbf{L})^{-1}\right) \\
= \; & (\beta - \beta') \Delta \mathbf{H}_i \times (I - \beta \mathbf{L})^{-1} + \beta' (\beta - \beta') \Delta \mathbf{H}_i \times (I - \beta' \mathbf{L})^{-1} \mathbf{L} (I - \beta \mathbf{L})^{-1} \\
= \; & (\beta - \beta') \Delta \mathbf{H}_i \times \left(I + \beta' (I - \beta' \mathbf{L})^{-1} \mathbf{L}\right) (I - \beta \mathbf{L})^{-1} \\
= \; & (\beta - \beta') \Delta \mathbf{H}_i \times (I - \beta' \mathbf{L})^{-1} (I - \beta \mathbf{L})^{-1},
\end{aligned}$$

so that

$$\beta \Delta \mathbf{g}_i(\beta) - \beta' \Delta \mathbf{g}_i(\beta') = (\beta - \beta') \Delta \mathbf{H}_i \times (I - \beta' \mathbf{L})^{-1} (I - \beta \mathbf{L})^{-1} \mathbf{R} \boldsymbol{\pi}_i.$$

If $\mathbf{R} \boldsymbol{\pi}_i \neq \mathbf{0}$, then $(I - \beta' \mathbf{L})^{-1} (I - \beta \mathbf{L})^{-1} \mathbf{R} \boldsymbol{\pi}_i \neq \mathbf{0}$ since both $(I - \beta' \mathbf{L})^{-1}$ and $(I - \beta \mathbf{L})^{-1}$ are non-singular by dominant diagonal theorem. Therefore $\Delta \mathbf{H}_i (I - \beta' \mathbf{L})^{-1} (I - \beta \mathbf{L})^{-1} \mathbf{R} \boldsymbol{\pi}_i$ cannot be a zero vector if $\Delta \mathbf{H}_i$ has full column rank, hence $\beta \Delta g_i(x; \beta)$ must differ from $\beta' \Delta g_i(x; \beta')$ for some x in X.■

References

[1] Aguirregabiria, V. and P. Mira (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, **75**, 1-53.

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

[3] Aguirregabiria, V. and J. Suzuki (2013): "Identification and Counterfactuals in Dynamic Models of Market Entry and Exit," *Working Paper, University of Toronto.*

[4] Bajari, P., C.L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, **75**, 1331-1370.

[5] Bajari, P., V. Chernozhukov, H. Hong and D. Nekipelov (2009): "Identification and Efficient Estimation of a Dynamic Discrete Game," *Working Paper, University of Minnesota.*

[6] Beresteanu, A., P. Ellickson and S. Misra (2010): "The Dynamics of Retail Oligopolies," *Working Paper, University of Rochester, Simon School of Business.*

[7] Berry, S. T. and P. A. Haile (2010): "Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers," *Cowles Foundation Discussion Paper No. 1718.*

[8] Berry, S. T. and P. A. Haile (2012): "Identification in Differentiated Products Markets Using Market Level Data," *Cowles Foundation Discussion Paper No. 1744R.*

[9] Chiong, K., A. Galichon, and M. Shum (2013): "Duality in Dynamic Discrete Choice Models," *working paper, California Institute of Technology.*

[10] Collard-Wexler, A. (2013): "Demand Fluctuations and Plant Turnover in the Ready-Mix Concrete Industry," *forthcoming in Econometrica.*

[11] Dunne, T., S. Klimek, M. J. Roberts and Y. D. Xu (2013): "Entry, Exit, and the Determinants of Market Structure," *forthcoming in RAND Journal of Economics..*

[12] Fan, Y. and M. Xiao (2012): "Competition and Subsidies in the Deregulated U.S. Local Telephone Industry," *Working Paper, University of Michigan, Department of Economics.*

[13] Frisch, R. and F.V. Waugh (1933): "Partial Time Regressions as Compared with Individual Trends," *Econometrica*, **1**, 387-401.

[14] Gowrisankaran, G., C. Lucarelli, P. Schmidt-Dengler and R. Town (2010): "Government Policy and the Dynamics of Market Structure: Evidence from Critical Access Hospitals," *Eleventh CEPR Conference on Applied Industrial Organization, Toulouse School of Economics.*

[15] Hotz, V., and R.A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, **60**, 497-531.

[16] Igami, M. (2012): "Estimating the Innovator's Dilemma: Structural Analysis of Creative Destruction in the Hard Disk Drive Industry," *Working Paper, Yale University.*

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

[18] Lin, H. (2012): "Quality Choice and Market Structure: A Dynamic Analysis of Nursing Home Oligopolies," *Working Paper, Indiana University, Kelley School of Business.*

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

[20] Norets, A. and X. Tang (2013): "Semiparametric Inference in Dynamic Binary Choice Models," *forthcoming in Review of Economic Studies.*

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

[22] Pesendorfer, M. and P. Schmidt-Dengler (2003): "Identification and Estimation of Dynamic Games," *NBER Working Paper.*

[23] Pesendorfer, M., and P. Schmidt-Dengler (2008): "Asymptotic Least Squares Estimator for Dynamic Games," *Review of Economics Studies*, **75**, 901-928.

[24] Ryan, S. P. (2012): "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, **80**, 1019-1061.

[25] Sanches, F.M. and D. Silva Jr (2013): "Bank Privatization and the Supply of Banking Services: Evidence from a Dynamic Structural Model," *Working Paper, LSE.*

[26] Sanches, F.M., D. Silva Jr and S. Srisuma (2013): "An Alternative Asymptotic Least Squares Estimator for Dynamic Games," *mimeo.*

[27] Snider, C. (2009): "Predatory Incentives and Predation Policy: The American Airlines Case," *Working Paper, UCLA, Department of Economics.*

[28] Srisuma, S. and O.B. Linton (2012): "Semiparametric Estimation of Markov Decision Processes with Continuous State Space," *Journal of Econometrics*, **166**, 320-341.

[29] Suzuki, J. (2013): "Land Use Regulation as a Barrier to Entry: Evidence from the Texas Lodging Industry," *forthcoming in International Economic Review.*

[30] Sweeting, A. (2013): "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry," *forthcoming in Econometrica.*