Journal of Econometrics 188 (2015) 236–249

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Heterogeneity and selection in dynamic panel data

Yuya Sasaki *

Johns Hopkins University, MD, United States

ARTICLE INFO

Article history:
Received 26 March 2013
Received in revised form
12 March 2015
Accepted 6 May 2015
Available online 21 May 2015

JEL classification:
C23

Keywords:
Dynamic panel data
Heterogeneity
Selection
Initial conditions problem

ABSTRACT

The data generating process (DGP) for generic dynamic panel data consists of a law of state dynamics g, a selection or attrition rule h, and an initial condition F. I study nonparametric identifiability of this complete DGP (g, h, F) using short unbalanced panel data, allowing for nonseparability between observed states and unobserved heterogeneity in each of g, h and F. For $T \geqslant 3$, the DGP is identified by using a proxy variable. For $T \geqslant 6$, the three additional periods construct a proxy, and thus the DGP is identified without an auxiliary variable.

1. Introduction

Dynamic econometric models describe how one's current experiences are causally related to her past experiences. The commonly observed persistence in economic states and economic choices are imputed to two distinct factors. First, past experiences reinforce the proneness to similar experiences or the propensity to make related choices — state dependence. Second, innate characteristics determine perpetual proneness to certain experiences — heterogeneity. Whether we can distinguish between these two causal paths has been discussed in a broad literature (e.g. Feller, 1943), particularly in econometrics (e.g., Heckman, 1981a,b, 1991).[1] In a closely related matter, the problem of handling the conditional distribution of unobserved heterogeneity on initial states (Heckman, 1981b), known as the initial conditions problem, is explored from various angles in the dynamic econometrics literature.

Bates and Neyman (1952) show that state dependence and heterogeneity can be distinguished by observing multiple periods for individuals, i.e., panel data. Dynamic panel data thus have proven to be useful in economic analyses since Balestra and Nerlove (1966). Reduced-form dynamic panel models usually contain additive fixed effects, and therefore rule out potential interactions between state dependence and heterogeneity. However, such interactions are ubiquitous in economics. For example, preference parameters seldom appear additively separably from state variables. To accommodate generic structures, econometric models should therefore factor in observed states and unobserved heterogeneity nonseparably.

Allowing for the nonseparability is still insufficient to completely describe the common data generating process (DGP) in practice. Empirical panel data are almost always unbalanced, which may occur as a result of economic decisions based the states and heterogeneity (e.g. Roy, 1951). When the unbalancedness is so caused, the selection process is a non-negligible component of the DGP in structural frameworks. Econometricians therefore account for these endogenous dynamic selection behaviors, as well as to control for the correlated unobserved heterogeneity, in order to identify the true causal effects of interest.

Reflecting the nonseparable state dependence and heterogeneity in both the dynamic process and the selection process, a complete DGP for generic dynamic panel data is defined by a triple (g, h, F) which consists of a law of state dynamics g, a selection rule h, and an initial condition F. An instance of such a DGP is $(g, h, F_{Y_1 U})$, where

$$\begin{cases} Y_t = g(Y_{t-1}, U, \mathcal{E}_t) \quad t = 2, \dots, T & \text{(Law of State Dynamics)} \\ D_t = h(Y_t, U, V_t) \quad t = 1, \dots, T-1 & \text{(Selection Rule)} \\ F_{Y_1 U} & \text{(Initial Condition).} \end{cases}$$

The first equation defines the law of dynamics for an observed state variable Y_t, such as income, as a first-order process with nonseparable unobserved heterogeneity U. The second equation

* Correspondence to: Department of Economics, 440 Mergenthaler Hall, 3400 N. Charles St, Baltimore, MD 21218, United States.
E-mail address: sasaki@jhu.edu.

[1] See also Chen et al. (1999) for semiparametric identification and estimation.

http://dx.doi.org/10.1016/j.jeconom.2015.05.002
0304-4076/

models a binary choice of selection D_t, such as labor force participation, as a Markov decision process with the state and heterogeneity. The initial condition F_{Y_1U} models the dependence of the initial state Y_1 on unobserved heterogeneity U, and features the aforementioned initial conditions problem (Heckman, 1981b).[2] This initial condition is a necessary ingredient of the DGP, because the DGP would not generate dynamic panel data without it. The period-specific shocks $(\mathcal{E}_t, V_t)$ are assumed to be exogenous, while the fixed effect U is not.

A leading example of selection in panel data is attrition, where individuals drop out of the panel upon $D_t = 0$. In this case, (Y_2, D_2) is observed if $D_1 = 1$; (Y_3, D_3) is observed if $D_1 = D_2 = 1$; and so on. Heckman and Navarro (2007) study this class of dynamic selection; also see Abbring and Heckman (2007) and Abbring (2010) for the background.

The panel data literature often focuses on inference of g under various structural and statistical restrictions on g, h, and F. In order to forecast potential effects of a policy by counterfactual analyses, however, it is essential to identify the complete policy-invariant DGP (g, h, F_{Y_1U}) involving not only g, but also h and F_{Y_1U}.[3] To this end, we first show that this complete DGP (g, h, F_{Y_1U}) is nonparametrically identified (up to normalization of error distributions as in Matzkin (2003, 2007)) using $T \geqslant 3$ periods of unbalanced panel data and a proxy variable. We then show that a proxy can be constructed from three additional periods of the state variable Y_t, and hence the DGP is identified using $T \geqslant 6$ periods of unbalanced panel data without any auxiliary variable.

Heterogeneity and selection are the two major sources of bias in panel data analysis (Hsiao, 2003, Ch. 1), and are often separately treated in the panel data literature. However, they are simultaneously relevant to a wide array of applications because (nonseparable) heterogeneity shows up in most economic models and dynamic selection is supposed in most empirical panel data. In this paper, we handle both of these jointly important issues in a unified framework.

Nonseparable panel models can be interpreted as nonparametric mixture models. The recent dynamic panel data literature demonstrates identification of heterogeneous dynamic processes as nonparametric mixture components (e.g. Kasahara and Shimotsu, 2009; Hu and Shum, 2010; Shiu and Hu, 2011).[4] This paper complements these references by showing nonparametric identification of the mixture components for the complete DGP involving the dynamic selection rule h as well as the nonparametric dynamic law g of the usual interest, using endogenously unbalanced panel data. To this end, a new method of handling missing data is proposed.

Selection has been studied in the panel data literature at least since Hausman and Wise (1979).[5] Heckman and Navarro (2007) nonparametrically identify nonseparable panel models with selection, using exogenous variables which exhibit sufficient variations. Davezies and D'Haultfoeuille (2011) study a related problem using instrumental variables. We propose an alternative approach which relies only on an endogenously evolving state variable Y_t. Our method requires no auxiliary variable when unbalanced panel data contain $T \geqslant 6$ periods.

While many applications focus on the dynamic law g as the object of primary interest, the selection rule h also helps to explain important causal effects in a variety of economic problems. The selection rule h can be interpreted as a reduced-form optimal stopping policy for such decisions as school dropout, retirement, exit and capital replacement after which econometricians do not observe states Y_t or selection D_t in panel data. The following example illustrates this connection between the structural optimal stopping models and the model studied in this paper, where the fixed effect U is common between the g and h functions by construction.[6]

Example 1 (*A Structural Optimal Stopping Model*)**.** Suppose that an economic agent knows her current utility or profit as a function π of state y_t and heterogeneity u. Let v_t^d denote a selection-specific private shock for each choice $d \in \{0, 1\}$, which is known to the agent. She also knows her exit value as a function $\overline{\nu}$ of state y_t and heterogeneity u. Using the dynamic function g, define the value function ν as the fixed point of the Bellman equation

$$\nu(y_t, u) = \mathrm{E}[\max\{\pi(y_t, u) + V_t^1 + \beta \mathrm{E}[\nu(g(y_t, u, \mathcal{E}_{t+1}), u)], \\ \pi(y_t, u) + V_t^0 + \beta\overline{\nu}(y_t, u)\}],$$

where β denotes the rate of time preference. It follows by construction that the reduced-form self-selection function h is given by

$$h(y_t, u, v_t) := \mathbb{1}\{\underbrace{\beta \mathrm{E}[\nu(g(y_t, u, \mathcal{E}_{t+1}), u)]}_{\text{Continuation value}} - \underbrace{\beta\overline{\nu}(y_t, u)}_{\text{Exit Value}} \\ \geqslant \underbrace{v_t^0 - v_t^1}_{\overset{\|}{v_t}}\}.$$

If $h(Y_t, U, V_t) = 0$, then the agent exits ($D_t = 0$) and subsequent states Y_t are unobserved. □

When attrition $D_t = 0$ is associated with hazards or ends of some duration, identification of the selection rule h entails identification of the mixed conditional hazard model. In this sense, our objective is also related to the literature on duration analysis (e.g. Lancaster, 1979; Elbers and Ridder, 1982; Heckman and Singer, 1984; Honoré, 1990; Ridder, 1990; Horowitz, 1999; Ridder and Woutersen, 2003; Abbring, 2012).

2. An overview

In this section, we present an informal sketch of the identification strategy focusing on a simple case in order to intuitively illustrate the logic of the main identification result of this paper. Section 3 follows it up with formal results for more general cases. For a quick view of the main results with no intuitive discussion, readers can skip Sections 2.2 and 2.3 to directly go to Section 3 after Section 2.1 without loss of logic flow.

2 Wooldridge (2005) and Honoré and Tamer (2006) advance this problem in the contexts of discrete outcome models. Blundell and Bond (1998) and Hahn (1999) use semiparametric distributions to obtain identifying restrictions and efficiency gain. This paper proposes restrictions for its nonparametric point identification.

3 Exogenous policy variables may enter g, h, and/or F_{Y_1U}, but we keep them implicit for notational simplicity. Identification of the policy-invariant DGP (g, h, F_{Y_1U}) in the reduced form is necessary and sufficient — see Marschak's (1953) maxim and Hurwicz (1962), discussed in Heckman and Vytlacil (2007).

4 Interpretation of nonseparable panel models as nonparametric mixtures was also pointed out by Evdokimov (2009) in the context of static models. Nonseparable static panel models are also studied by Altonji and Matzkin (2005), Chernozhukov et al. (2010), and Hoderlein and White (2009).

5 Existing solution methods include, but are not limited to, use of additional data such as refreshment samples (Ridder, 1992; Hirano et al., 2001; Bhattacharya, 2008), matching (Kyriazidou, 1997), and weighting (Hellerstein and Imbens, 1999; Moffitt et al., 1999; Wooldridge, 2002b). Das (2004) studies selection for nonparametric additively separable panel models.

6 Identification under additive parametric models is feasible even if U is not common between the outcome and selection equations — see Kyriazidou (1997) for example.