Journal of Econometrics 188 (2015) 219–235

Contents lists available at ScienceDirect

Journal of Econometrics

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

Estimation of heterogeneous autoregressive parameters with short panel data

Sophocles Mavroeidis[a], Yuya Sasaki[b,*], Ivo Welch[c]

[a] Department of Economics, University of Oxford, United Kingdom
[b] Department of Economics, Johns Hopkins University, United States
[c] Anderson Graduate School of Management, UCLA, United States

ARTICLE INFO

Article history:
Received 12 February 2013
Received in revised form
12 December 2014
Accepted 5 May 2015
Available online 19 May 2015

Keywords:
Panel data
Employment dynamics
Heterogeneous autoregressive parameters
Initial conditions
Maximum likelihood

ABSTRACT

This paper presents a maximum likelihood approach to estimation of cross sectional distributions of heterogeneous autoregressive (AR) parameters with short panel data. We construct a panel likelihood by integrating the unknown cross sectional density of heterogeneous AR parameters with respect to a known time-series data generating kernel. The solution to this extremal criterion recovers the unknown density of heterogeneous AR parameters. Applying our method to a model of employment dynamics with the firm-level data of Arellano and Bond (1991), we find that adjustment rates of employment are significantly heterogeneous across firms.

1. Introduction

Dynamic panel models are useful for understanding dynamic causality under correlated unobserved heterogeneity. Empirical studies often presume the constant-coefficient model

$$Y_{i,t} = \alpha_i + \beta Y_{i,t-1} + \varepsilon_{i,t} \quad \text{where}$$
$$\varepsilon_{i,t} \overset{i.i.d.}{\sim} N(0, \sigma^2), \quad t = 2, \ldots, T \tag{1.1}$$

with additive fixed effects α_i. Various assumptions and methods are proposed in an extensive body of the literature for estimation of this class of models.

In an extended model, all the AR parameters (α, β, σ) are potentially heterogeneous across individuals or firms i. A nondegenerate distribution of β_i reflects heterogeneous adjustment rates or heterogeneous growth rates. With this cross sectional variation in the AR parameters, we consider the following random-parameter extension to (1.1):

$$Y_{i,t} = \alpha_i + \beta_i Y_{i,t-1} + \varepsilon_{i,t} \quad \text{where} \quad \varepsilon_{i,t} \overset{i.i.d.}{\sim} N(0, \sigma_i^2) \tag{1.2}$$

for $i = 1, \ldots, N$ and $t = 2, \ldots, T$. This paper proposes an estimation approach for the cross-sectional distributions of the heterogeneous AR parameters $(\alpha_i, \beta_i, \sigma_i)$ using short panel data.

To this goal, we suggest the following perspective of dynamic panel data. A realization $(\alpha_i, \beta_i, \sigma_i, Y_{i,1}) \overset{i.i.d.}{\sim} F_{\alpha\beta\sigma Y_1}$ of the cross sectional draws of the AR parameters and the initial state generates an individual's time series $\{Y_{i,1}, \ldots, Y_{i,T}\}$ through the dynamic model (1.2). Therefore, N cross sectional realizations of $(\alpha_i, \beta_i, \sigma_i, Y_{i,1}) \overset{i.i.d.}{\sim} F_{\alpha\beta\sigma Y_1}$ produce dynamic panel data $\{Y_{i,1}, \ldots, Y_{i,T}\}_{i=1}^N$ through the dynamic model (1.2). We can thus consider the cross-sectional distribution $F_{\alpha\beta\sigma Y_1}$ as the primitive of the data generating process. Since the marginal distribution of $Y_{i,1}$ is directly identifiable and the primitive can be decomposed as $f_{\alpha\beta\sigma Y_1} = f_{\alpha\beta\sigma|Y_1} \cdot f_{Y_1}$, identification of $f_{\alpha\beta\sigma|Y_1}$ suffices for identification of the primitive density $f_{\alpha\beta\sigma Y_1}$. Moreover, the following examples show why it is meaningful to estimate this conditional density $f_{\alpha\beta\sigma|Y_1}$.

For one example, note that one can identify the mean of the heterogeneous adjustment rates by

$$\mathrm{E}[\beta_i] = \int\int\int\int b \cdot f_{\alpha\beta\sigma|Y_1}(a, b, s \mid y_1)\, dadbds\, dF_{Y_1}(y_1)$$

once $f_{\alpha\beta\sigma|Y_1}$ is identified. Hence, one can consistently estimate it by the sample counterpart

$$\widehat{\mathrm{E}}[\beta_i] = \frac{1}{N}\sum_{i=1}^{N}\int\int\int b \cdot \hat{f}_{\alpha\beta\sigma|Y_1}(a, b, s \mid Y_{i,1})\, dadbds$$

* Corresponding author.
E-mail addresses: sophocles.mavroeidis@economics.ox.ac.uk (S. Mavroeidis), sasaki@jhu.edu (Y. Sasaki), ivo.welch@anderson.ucla.edu (I. Welch).

http://dx.doi.org/10.1016/j.jeconom.2015.05.001
0304-4076/

once an estimate $\hat{f}_{\alpha\beta\sigma|Y_1}$ is obtained. Similarly, the variance of the heterogeneous adjustment rates may be consistently estimated by

$$\widehat{\mathrm{Var}}(\beta_i) = \frac{1}{N}\sum_{i=1}^{N}\int\int\int \left(b - \widehat{\mathrm{E}}[\beta_i]\right)^2 \cdot \hat{f}_{\alpha\beta\sigma|Y_1}(a, b, s \mid Y_{i,1})\, dadbds. \tag{1.3}$$

For another example, suppose that β_i has subunit support and hence $(\alpha_i, \beta_i, \sigma_i, Y_{i,t})$ is jointly stationary. In this case, $(\alpha_i, \beta_i, \sigma_i \mid Y_{i,t}) \overset{\text{Law}}{=} (\alpha_i, \beta_i, \sigma_i \mid Y_{i,1})$ holds for all t, and thus $f_{\alpha\beta\sigma|Y_1}$ can be used for the purpose of forecasting as follows:

$$\mathrm{E}[Y_{i,t+1} \mid Y_{i,t}] = \int\int\int (a + bY_{i,t}) \cdot f_{\alpha\beta\sigma|Y_1}(a, b, s \mid Y_{i,t})\, dadbds. \tag{1.4}$$

The sample-counterpart conditional mean $\widehat{\mathrm{E}}[Y_{i,t+1}|Y_{i,t}]$ is obtained by simply replacing $f_{\alpha\beta\sigma|Y_1}$ by the estimate $\hat{f}_{\alpha\beta\sigma|Y_1}$ in this formula (1.4). These examples show that, even if it is not solving the incidental parameters problem *per se*,[1] our consistent estimation of the cross-sectional conditional distribution $F_{\alpha\beta\sigma|Y_1}$ can be useful for a variety of research objectives.

This paper is not the first to consider heterogeneous dynamic panel models. Many empirical papers consider heterogeneity along specific dimensions of observed attributes, such as OECD versus the rest in growth economics, and large versus small firms in production analysis. A more general approach is to consider heterogeneous models across unknown groups (e.g., Lin and Ng (2012) study this type of model in static setups). Random coefficient models handle finer resolutions of unknown groups, namely up to individual levels of heterogeneity. With large N and large T, Pesaran and Smith (1995) discuss consistent estimation of long-run effect, Pesaran et al. (1999) propose a long-run restriction for consistent estimation of random coefficients, and Pesaran and Yamagata (2008) propose modified Swamy tests of coefficient homogeneity. Hsiao et al. (1999) propose a Bayesian estimator[2] which is asymptotically equivalent to the mean group estimator studied by Pesaran and Smith (1995). Hsiao and Pesaran (2008) provide a comprehensive survey of both static and dynamic contexts of random-coefficient panel data models.

Our contribution to this literature is at least five-fold. First, our method is based on large N and small T, whereas the aforementioned frequentist studies on random-coefficient dynamic panel models consider large-T asymptotics for consistency. Our small T setup is useful for a wide array of micro-econometric applications as short periods (e.g., $T \approx 5$) are quite common in longitudinal surveys of households and firms.[3] Second, compared to the Bayesian estimator of Hsiao et al. (1999) that imposes a fixed initial value y_1, our approach of estimating the local distribution $f_{\alpha\beta\sigma|Y_1}$ allows for a full flexibility in the marginal distribution of the initial values.[4] Third, unlike the previous studies which are focused on inference of the means of short-run and long-run effects, our method is capable of making an inference for other distributional features such as the variance (1.3) of the heterogeneous adjustment rates. Fourth, our estimand $f_{\alpha\beta\sigma|Y_1}$ can be used for forecasting, e.g., (1.4), under the additional assumption of stationarity as previously mentioned. The last but not least contribution is our empirical illustration that evidences non-trivial heterogeneity. Applying our method to the firm-level data of Arellano and Bond (1991), we find that $\mathrm{Var}(\beta_i)$, as a measure of heterogeneity in adjustment rates, is significantly different from zero for employment dynamics.

2. The main idea

Dynamic panel data restricted to a single individual i is a time series $\{Y_{i,1}, \ldots, Y_{i,T}\}$. The standard time series methods would provide a consistent estimate of $(\alpha_i, \beta_i, \sigma_i)$ for each individual i with large T. For example, the maximum likelihood estimator

$$(\hat{\alpha}_i, \hat{\beta}_i, \hat{\sigma}_i) = \arg\max_{(a,b,s)} \log \underbrace{\left[s^{1-T} \prod_{t=2}^{T} \phi\left(\frac{Y_{i,t} - a - bY_{i,t-1}}{s}\right)\right]}_{\text{Individual's time series likelihood}}, \tag{2.1}$$

where ϕ denotes the density function of the standard normal distribution, is consistent in the limit $T \to \infty$, provided appropriate restrictions on the time-series dependence. Under our short panel setting, however, we instead exploit the asymptotics in the limit as $N \to \infty$, while we continue to use the individual's time series likelihood function of (2.1) as a kernel.

The cross sectional distribution $F_{Y_T\cdots Y_2|Y_1}$ of time series is observable in panel data of length $T < \infty$. Its density $f_{Y_T\cdots Y_2|Y_1}$ can be decomposed into the cross sectional density $f_{\alpha\beta\sigma|Y_1}$ of heterogeneous AR parameters and the time series likelihood of (2.1) in the following manner:

$$\underbrace{f_{Y_T\cdots Y_2|Y_1}(y_t, \ldots, y_2 \mid y_1)}_{\substack{\text{Cross sectional distribution of time series} \\ \text{Observed}}}$$
$$= \int\int\int f_{\alpha\beta\sigma|Y_1}(a, b, s \mid y_1) f_{Y_T\cdots Y_2|Y_1\alpha\beta\sigma} \times (y_T, \ldots, y_2 \mid y_1, a, b, s)\, dadbds$$
$$= \int\int\int \underbrace{f_{\alpha\beta\sigma|Y_1}(a, b, s \mid y_1)}_{\substack{\text{Primitive of the model} \\ \text{Unknown}}} \times \underbrace{\left[s^{1-T} \prod_{t=2}^{T} \phi\left(\frac{y_t - a - by_{t-1}}{s}\right)\right]}_{\substack{\text{Individual's time series likelihood} \\ \text{Known}}} dadbds. \tag{2.2}$$

The integrand in the last line of (2.2) consists of two factors. The first factor, the unknown cross sectional density $f_{\alpha\beta\sigma|Y_1}$ of heterogeneous AR parameters, is the primitive of the model that we seek to estimate. The second factor, the individual's time series likelihood function similar to the one in (2.1), is known up to the unknown AR parameters (a, b, s).

We focus on a given point y_1 of the initial state Y_{i1} for the moment, and define the linear integral operator $L : \mathcal{L}^2(F_{\alpha\beta\sigma|Y_1=y_1}) \to \mathcal{L}^2(F_{Y_T\cdots Y_2|Y_1=y_1})$ by

$$(L\xi)(y_T, \ldots, y_2) := \int\int\int \xi(a, b, s) \times \left[s^{1-T} \prod_{t=2}^{T} \phi\left(\frac{y_t - a - by_{t-1}}{s}\right)\right] dadbds.$$

[1] For incidental parameters, information does not accumulate as sample size increases (Neyman and Scott, 1948). Examples of such parameters are fixed effects in the panel data literature with small T (Lancaster, 2000). For the model (1.2), both α_i and β_i are incidental parameters since no information accumulates if T is fixed.

[2] Also related to this Bayesian estimator is the MCMC literature that proposes hierarchical distributional specifications for random-coefficient static panel models, e.g., Chib and Carlin (1999) and Chib (2008).

[3] It is well known that time series estimates suffer from finite-sample biases of order T^{-1} (Kiviet and Phillips, 1993). This rate of bias is problematic for longitudinal survey data with $T \approx 5$.

[4] Handling the initial states is a key issue in panel data analysis (Heckman, 1981). Hsiao (1986) lists a number of alternative assumptions about the initial value. Blundell and Bond (1998) and Hahn (1999) propose to impose restrictions on the initial condition for identifying restrictions and efficiency gains. Wooldridge (2005) and Honoré and Tamer (2006) advance this initial conditions problems in the contexts of nonlinear binary response models.