Unequal Spacing in Dynamic Panel Data: Identification and Estimation

Yuya Sasaki* Yi Xin

Johns Hopkins University

PRELIMINARY

Comments and suggestions are appreciated.

February 18, 2014

Abstract

This paper provides conditions under which parameters of fixed-effect dynamic models are identified with unequally spaced panel data. Under predeterminedness, weak stationarity, and empirically testable rank conditions, AR(1) parameters are identified if $\tau, \tau+1, \Delta t+\tau, \Delta t+\tau+1 \in \mathcal{T}$ holds for some $\tau \geqslant 0$ and $\Delta t > 0$, where $\mathcal{T}$ is the set of all the time gaps. This result extends to models with multiple covariates, higher-order autoregressions, time-varying trends, and partially linear models. For the NLS Original Cohorts: Older Men, personal interviews took place in 1966, 67, and 69, and the above condition is satisfied with $\mathcal{T} = \{0, 1, 2, 3\}$, i.e., $(\tau, \Delta t) = (0, 2)$. Applying our method to this data set, we obtain estimates of the AR(1) parameter for earning dynamics ranging from .34 to .59.

Keywords: dynamic panel data, unequal spacing

*Yuya Sasaki, Department of Economics, 440 Mergenthaler Hall, 3400 N. Charles St, Baltimore, MD 21218, email: sasaki@jhu.edu. Yi Xin, Department of Economics, 440 Mergenthaler Hall, 3400 N. Charles St, Baltimore, MD 21218, email: yxin4@jhu.edu.

1 Introduction

In economics, numerous empirical questions have been answered through the dynamic panel data model of the form

$$y_{it} = \gamma y_{it-1} + \beta x_{it} + \alpha_i + \varepsilon_{it}, \tag{1.1}$$

where y_{it} is an observed state variable, x_{it} is an observed covariate, α_i is an unobserved individual fixed effect, and ε_{it} is an idiosyncratic error. Among others, method-of-moment approaches (e.g., Anderson and Hsiao, 1981; Arellano and Bond, 1991; Ahn and Schmidt, 1995; Blundell and Bond, 1998; Hahn, 1999) enjoy practical and theoretical advantages to attract a large group of users. These methods exploit the instrumental orthogonality of the first difference $\varepsilon_{it} - \varepsilon_{it-1} = (y_{it} - y_{it-1}) - \gamma(y_{it-1} - y_{it-2}) - \beta(x_{it} - x_{it-1})$ as well as other supplementary moment restrictions. As such, they require observation of y_{it} for at least three consecutive time periods (or alternatively two pairs of two consecutive time periods).

Many panel surveys are conducted with unequal time spacing, and may not provide the required set of time periods. For the NLS Original Cohorts: Older Men, for example, personal interviews were conducted in 1966, 67, 69, 71, 76, 81, and 90.[1] This data set contains neither three consecutive time periods nor two pairs of two consecutive time periods. We thus fail to difference out the fixed effect from equation (1.1), and cannot directly adapt the aforementioned approaches to construct moment restrictions.

Given that the standard method-of-moment approaches are not generally effective once panel data exhibit unequal time spacing, can we develop similarly useful alternative estimation methods? Through this paper, we answer this question by providing conditions under which parameters (γ, β) of the model (1.1) are identified even if panel data are unequally spaced. In addition to the relatively standard assumptions such as predeterminedness, weak stationarity, and empirically testable rank conditions, we require certain patterns of unequal time spacing for the parameters to be identified. It is also shown that many of the unequally spaced panel data sets from the US and the UK satisfy our requirement of spacing patterns.

We are not the first to study unequally spaced panel data. Rosner and Munoz (1988)

[1] They conducted mail or telephone interviews in 1968, 73, 75, 78, 80, and 83, but responses through different media of communication should be carefully distinguished for survey analysis.

use linear interpolation to approximate missing data for dynamic panel models. Jones and Boadi-Boateng (1991) take the parametric maximum likelihood solution for static panel models with serial correlation. Baltagi and Wu (1999) propose a feasible GLS procedure for static panel models with serial correlation. McKenzie (2001) show consistent estimation of dynamic (pseudo) panel models, but this method requires observation of covariates in the missing time periods. Millimet and McDonough (2013) apply a variety of estimation methods for dynamic panel models and report their finite sample performances.

This paper differs from each of these preceding papers in terms of two or more of the following six points. First, most importantly, we prove identification, and specifically propose general spacing patterns as sufficient conditions for identification. Second, we deal with dynamic models which exhibit more complications than static models. Third, parametric distributional assumptions are not imposed. Forth, our approach does not rely on interpolation or imputation. Fifth, our method does not require partial observation in missing time periods. Sixth, our model can allow for arbitrary correlation among the observed state, the unobserved fixed effect, and the observed covariates.

With all these advantages, we admit that our identification result is based on a non-trivial set of assumptions. As mentioned earlier, we assume predeterminedness and weak stationarity. While predeterminedness is often innocuous in applications, the weak stationarity can be restrictive in some applications, particularly with those state variables that grow or accumulate over time. Later in this paper, we provide a remedy to alleviate this stationarity assumption by introducing time-varying means and variances. Our rank condition is empirically testable, and can also be handled by the existing methods of weak-rank-robust inference.

Our key requirement for identification is that $\tau, \tau + 1, \Delta t + \tau, \Delta t + \tau + 1 \in \mathcal{T}$ holds for some $\tau \geqslant 0$ and $\Delta t > 0$, where $\mathcal{T}$ is the set of all the time gaps. None of the preceding papers proposes such general spacing patterns as sufficient (or necessary) condition for identification. This requirement is satisfied with $\mathcal{T} = \{0, 1, 2, 3\}$, i.e., $(\tau, \Delta t) = (0, 2)$, for the NLS Original Cohorts: Older Men, which we picked as an example earlier. This paper contributes to the body of our knowledge and provides a guidance to practitioners by formally ensuring identification of dynamic fixed-effect models under the stylized patterns of unequally spaced panel data.

2 A Basic Model

We first fix index notations for unequally spaced panel data. Let T be the set of all observed time periods. Define the set of survey gaps by $\mathcal{T} = \{|t_1 - t_2| : t_1, t_2 \in T\}$. Also define the set of gap-associated survey years by $T(\tau) = \{t \in T : t+\tau \in T\}$ for each gap $\tau \in \mathcal{T}$, and let $T(\tau) = \phi$ if $\tau \notin \mathcal{T}$. For the NLS Original Cohorts: Older Men, introduced in the previous section, personal interviews were conducted in 1966, 67, 69, 71, 76, 81, and 90. In this case, we have $T = \{66, 67, 69, 71, 76, 81, 90\}$, $\mathcal{T} = \{0, 1, 2, 3, 4, 5, 7, 9, 10, 12, 14, 15, 19, 21, 23, 24\}$, $T(0) = T$, $T(1) = \{66\}$, $T(2) = \{67, 69\}$, $T(3) = \{66\}$, $T(4) = \{67\}$, $T(5) = \{66, 71, 76\}$, and so on.

Let us first consider the following simple first-order autoregressive model for illustration.

$$y_{it} = \gamma y_{it-1} + \beta x_{it} + \alpha_i + \varepsilon_{it} \tag{2.1}$$

where y_{it} is an observed state variable, x_{it} is an observed covariate, α_i is an unobserved individual fixed effect, and ε_{it} is an unobserved idiosyncratic shock for individual i at period t. This baseline model has two parameters, γ and β. The dynamic process (2.1) is equipped with the following set of model assumptions.

Assumption 1 (Predeterminedness). *$E_i[y_{it}\varepsilon_{is}] = 0$ and $E_i[x_{it}\varepsilon_{is}] = 0$ whenever $s > t$.*

Assumption 2 (Weak Stationarity). *For each individual $i = 1, 2, \cdots, N$:*

(i) $\bar{\mu}_i := E_i(x_{it})$ and $\tilde{\mu}_i := E_i(y_{it})$ are t-invariant.

(ii) $\bar{\sigma}_i^2 := Var_i(x_{it})$ and $\tilde{\sigma}_i^2 := Var_i(y_{it})$ are t-invariant.

(iii) $\bar{\psi}_{i\tau} := Cov_i(x_{it}, x_{it+\tau})$ and $\tilde{\psi}_{i\tau} := Cov_i(y_{it}, y_{it+\tau})$ are t-invariant.

(iv) $\Psi_{i\tau} := Cov_i(y_{it}, x_{it+\tau})$ and $\Psi_{i-\tau} := Cov_i(x_{it}, y_{it+\tau})$ are t-invariant.

For a primitive structural model that sufficiently satisfies this weak stationarity assumption, one can consider the linear VAR structure for (y_{it}, x_{it+1}) with sub-unit coefficient restrictions for example, where one of the two equations is (2.1). An obvious disadvantage of Assumption 2 is that it prohibits time-varying means and variances of the state variable, whereas time variations are fairly common for many economic variables, particularly those that grow or accumulate over time. We will relax this restriction by introducing time-varying means and variances later in Section 6.1.

3 Identification under Unequal Spacing

Our identification strategy under unequally spaced dynamic panel data is as follows. Let t_1 and t_2 be two time periods in T such that $t_1 > t_2$. Taking the difference of the dynamic equation (2.1) between these two time periods yields

$$y_{it_1} - y_{it_2} = \gamma(y_{it_1-1} - y_{it_2-1}) + \beta(x_{it_1} - x_{it_2}) + (\varepsilon_{it_1} - \varepsilon_{it_2})$$

Multiplying both sides of this equation by $y_{it_2-1-\tau}$ and $x_{it_2-1-\tau}$, we obtain the following two equations.

$$y_{it_2-1-\tau}(y_{it_1} - y_{it_2}) = \gamma y_{it_2-1-\tau}(y_{it_1-1} - y_{it_2-1}) + \beta y_{it_2-1-\tau}(x_{it_1} - x_{it_2}) + y_{it_2-1-\tau}(\varepsilon_{it_1} - \varepsilon_{it_2})$$

$$x_{it_2-1-\tau}(y_{it_1} - y_{it_2}) = \gamma x_{it_2-1-\tau}(y_{it_1-1} - y_{it_2-1}) + \beta x_{it_2-1-\tau}(x_{it_1} - x_{it_2}) + x_{it_2-1-\tau}(\varepsilon_{it_1} - \varepsilon_{it_2})$$

We take the expectation E_i of the above two equations for each individual i as follows.

$$E_i(y_{it_2-1-\tau}(y_{it_1} - y_{it_2})) = \tag{3.1}$$
$$\gamma E_i(y_{it_2-1-\tau}(y_{it_1-1} - y_{it_2-1})) + \beta E_i(y_{it_2-1-\tau}(x_{it_1} - x_{it_2})) + E_i(y_{it_2-1-\tau}(\varepsilon_{it_1} - \varepsilon_{it_2}))$$

$$E_i(x_{it_2-1-\tau}(y_{it_1} - y_{it_2})) = \tag{3.2}$$
$$\gamma E_i(x_{it_2-1-\tau}(y_{it_1-1} - y_{it_2-1})) + \beta E_i(x_{it_2-1-\tau}(x_{it_1} - x_{it_2})) + E_i(x_{it_2-1-\tau}(\varepsilon_{it_1} - \varepsilon_{it_2}))$$

In order to further simplify (3.1) and (3.2), we now invoke Assumptions 1 and 2. First, the predeterminedness in Assumption 1 implies $E_i(y_{it_2-1-\tau}(\varepsilon_{it_1} - \varepsilon_{it_2})) = 0$ and $E_i(x_{it_2-1-\tau}(\varepsilon_{it_1} - \varepsilon_{it_2})) = 0$ for any $\tau \geq 0$, vanishing the last term on the right-hand side of each of equations (3.1) and (3.2). Second, the weak stationarity in Assumption 2 allows us to define the following t-invariant cross-sectional random variables.

(i) $Z_{i\tau} := \tilde{\psi}_{i\tau} + \tilde{\mu}_i^2 = E_i(y_{it}y_{it+\tau})$, where $Z_{i\tau}$ is t-invariant.

(ii) $z_{i\tau} := \bar{\psi}_{i\tau} + \bar{\mu}_i^2 = E_i(x_{it}x_{it+\tau})$, where $z_{i\tau}$ is t-invariant.

(iii) $\zeta_{i\tau} := \Psi_{i\tau} + \tilde{\mu}_i\bar{\mu}_i = E_i(y_{it}x_{it+\tau})$, where $\zeta_{i\tau}$ is t-invariant.

(iv) $\zeta_{i-\tau} := \Psi_{i-\tau} + \tilde{\mu}_i\bar{\mu}_i = E_i(x_{it}y_{it+\tau})$, where $\zeta_{i-\tau}$ is t-invariant.

With these properties implied by Assumptions 1 and 2, we can rewrite (3.1) and (3.2) as

$$\begin{aligned} Z_{i\Delta t+\tau+1} - Z_{i\tau+1} &= \gamma(Z_{i\Delta t+\tau} - Z_{i\tau}) + \beta(\zeta_{i\Delta t+\tau+1} - \zeta_{i\tau+1}) \qquad \text{and} \\ \zeta_{i-(\Delta t+\tau+1)} - \zeta_{i-(\tau+1)} &= \gamma(\zeta_{i-(\Delta t+\tau)} - \zeta_{i-\tau}) + \beta(z_{i\Delta t+\tau+1} - z_{i\tau+1}) \end{aligned}$$

respectively, where $\Delta t = t_1 - t_2$ denotes the gap between the two time periods, t_1 and t_2. Taking the cross-sectional means E of each of the above two equations yields

$$Z_{\Delta t+\tau+1} - Z_{\tau+1} = \gamma(Z_{\Delta t+\tau} - Z_{\tau}) + \beta(\zeta_{\Delta t+\tau+1} - \zeta_{\tau+1}) \tag{3.3}$$

$$\zeta_{-(\Delta t+\tau+1)} - \zeta_{-(\tau+1)} = \gamma(\zeta_{-(\Delta t+\tau)} - \zeta_{-\tau}) + \beta(z_{\Delta t+\tau+1} - z_{\tau+1}) \tag{3.4}$$

where $Z_\tau := \mathrm{E}[Z_{i\tau}]$, $z_\tau := \mathrm{E}[z_{i\tau}]$, $\zeta_\tau := \mathrm{E}[\zeta_{i\tau}]$, and $\zeta_{-\tau} := \mathrm{E}[\zeta_{i-\tau}]$ for short-hand notations.

Equations (3.3) and (3.4) involve twelve cross-sectional moments: $Z_{\Delta t+\tau+1}$, $Z_{\tau+1}$, $Z_{\Delta t+\tau}$, Z_τ, $\zeta_{\Delta t+\tau+1}$, $\zeta_{\tau+1}$, $\zeta_{-(\Delta t+\tau+1)}$, $\zeta_{-(\tau+1)}$, $\zeta_{-(\Delta t+\tau)}$, $\zeta_{-\tau}$, $z_{\Delta t+\tau+1}$, and $z_{\tau+1}$. Due to the t-invariance implied by Assumption 2, the first one of these moments, $Z_{\Delta t+\tau+1}$, can be observed as the cross-sectional moment of $y_{it}y_{it+\Delta t+\tau+1}$ for any $t \in T(\Delta t + \tau + 1)$ provided that $T(\Delta t + \tau + 1) \neq \phi$ is true. Likewise, all the cross sectional moments in (3.3) and (3.4) can be observed using unequally spaced panel data if $T(\tau) \neq \phi$, $T(\tau+1) \neq \phi$, $T(\Delta t + \tau) \neq \phi$, and $T(\Delta t + \tau + 1) \neq \phi$ are true.

Once all the cross-sectional moments in (3.3) and (3.4) are observed from unequally spaced panel data, we can solve the system to explicitly identify the structural parameters (γ, β) by

$$\begin{pmatrix} \gamma \\ \beta \end{pmatrix} = \frac{1}{|\Delta|} \begin{pmatrix} (z_{\Delta t+\tau+1} - z_{\tau+1})(Z_{\Delta t+\tau+1} - Z_{\tau+1}) + (\zeta_{\tau+1} - \zeta_{\Delta t+\tau+1})(\zeta_{-(\Delta t+\tau+1)} - \zeta_{-(\tau+1)}) \\ (\zeta_{-\tau} - \zeta_{-(\Delta t+\tau)})(Z_{\Delta t+\tau+1} - Z_{\tau+1}) + (Z_{\Delta t+\tau} - Z_{\tau})(\zeta_{-(\Delta t+\tau+1)} - \zeta_{-(\tau+1)}) \end{pmatrix} \tag{3.5}$$

where $|\Delta| = (Z_{\Delta t+\tau} - Z_\tau)(z_{\Delta t+\tau+1} - z_{\tau+1}) - (\zeta_{-(\Delta t+\tau)} - \zeta_{-\tau})(\zeta_{\Delta t+\tau+1} - \zeta_{\tau+1})$, provided that the following empirically testable rank condition is satisfied.

Assumption 3 (Empirically Testable Rank Condition)**.** $|\Delta| \neq 0$.

This identification result is summarized as a theorem below.

Theorem 1 (Identification)**.** *If Assumptions 1, 2, and 3 are satisfied for (2.1), and unequally spaced panel data have $T(\tau) \neq \phi$, $T(\tau+1) \neq \phi$, $T(\Delta t + \tau) \neq \phi$, and $T(\Delta t + \tau + 1) \neq \phi$, then (γ, β) is identified by the formula (3.5).*

Remark 1. *The condition* $T(\tau) \neq \phi$*,* $T(\tau+1) \neq \phi$*,* $T(\Delta t + \tau) \neq \phi$ *and* $T(\Delta t + \tau + 1) \neq \phi$ *provided in the theorem is equivalent to the alternative writing* $\tau, \tau+1, \Delta t+\tau, \Delta t+\tau+1 \in \mathcal{T}$ *used in the abstract and the introductory section for succinctness.*

Given that $T(0) \neq \phi$ is always true whenever $T \neq \phi$ is true, we can consider two representative cases of unequal spacing structures under which all of these index sets, $T(\tau)$, $T(\tau+1)$, $T(\Delta t + \tau)$, and $T(\Delta t + \tau + 1)$, are nonempty. These two cases are formalized by Definitions 1 and 2 below.

Definition 1 (UK Spacing)**.** *If panel data with unequal spacing has* $T(\tau) \neq \phi$*,* $T(\tau+1) \neq \phi$*, and* $T(\tau+2) \neq \phi$ *for some base gap* τ*, then we call its spacing structure the "UK spacing."*

Example 1 (UK Spacing)**.** *The following is a list of unequally spaced British panel data sets listed in Millimet and McDonough (2013; Table 1).*

1958 National Child Development Study	$T = \{7, 11, 16, 23, 33, 42, 46, 50\}$
1970 British Cohort Study	$T = \{5, 10, 16, 26, 30, 34, 38, 42\}$
Millennium Cohort Study	$T = \{0, 3, 5, 7\}$
National Pupil Database	$T = \{7, 11, 14, 16\}$

Each of these panel data sets has the UK spacing structure. 1958 National Child Development Study has $T(7) \neq \phi$*,* $T(8) \neq \phi$*, and* $T(9) \neq \phi$*. 1970 British Cohort Study has* $T(4) \neq \phi$*,* $T(5) \neq \phi$*, and* $T(6) \neq \phi$*. Millennium Cohort Study has* $T(2) \neq \phi$*,* $T(3) \neq \phi$*, and* $T(4) \neq \phi$*. National Pupil Database has* $T(2) \neq \phi$*,* $T(3) \neq \phi$*, and* $T(4) \neq \phi$*.* □

Definition 2 (US Spacing)**.** *If panel data with unequal spacing has* $T(1) \neq \phi$*,* $T(\Delta t) \neq \phi$*, and* $T(\Delta t + 1) \neq \phi$ *for some gap* $\Delta t \in \mathbb{N}$*, then we call its spacing structure the "US spacing."*

Remark 2. *When* $T = \{1, 2, 4\}$*, the unequal panel data can be characterize as either "UK spacing" or "US spacing". In this case,* $\Delta t = 1$*,* $\tau = 0$*, we only require* $T(1) \neq \phi$ *and* $T(2) \neq \phi$*.*

Example 2 (US Spacing)**.** *The following is a list of unequally spaced American panel data sets listed in Millimet and McDonough (2013; Table 1).*

NLS Original Cohorts: Older Men	$T = \{66, 67, 69, 71, 76, 81, 90\}$
Current Population Survey	$T = \{\cdots, 3, 4, 13, 14, \cdots\}$
Early Childhood Longitudinal Survey-K	$T = \{3, 4, 8, 12, 18\}$
National Longitudinal Survey of Youth 1979	$T = \{\cdots, 93, 94, 96, 98, \cdots\}$
Panel Study of Income Dynamics	$T = \{\cdots, 96, 97, 99, 01, \cdots\}$

Each of these panel data sets has the US spacing structure. NLS Original Cohorts: Older Men has $T(1) \neq \phi$ and $T(2) \neq \phi$. Current Population Survey has $T(1) \neq \phi$, $T(9) \neq \phi$, and $T(10) \neq \phi$. Early Childhood Longitudinal Survey-Kindergarten Cohort[2] *has $T(1) \neq \phi$, $T(4) \neq \phi$, and $T(5) \neq \phi$. National Longitudinal Survey of Youth 1979 has $T(1) \neq \phi$ and $T(2) \neq \phi$. Panel Study of Income Dynamics has $T(1) \neq \phi$ and $T(2) \neq \phi$.* □

These two spacing patterns, namely the UK spacing and the US spacing, satisfy the general condition required in Theorem 1. We state this result as a corollary below.

Corollary 1 (Identification under UK and US Spacing)**.** *If Assumptions 1 and 2 are satisfied for (2.1), then we have the following two results as consequences of Theorem 1.*

(i) If $T(\tau) \neq \phi$, $T(\tau+1) \neq \phi$, and $T(\tau+2) \neq \phi$ (UK Spacing), then (γ, β) is identified by

$$\begin{pmatrix} \gamma \\ \beta \end{pmatrix} = \frac{1}{|\Delta|} \begin{pmatrix} (z_{\tau+2} - z_{\tau+1})(Z_{\tau+2} - Z_{\tau+1}) + (\zeta_{\tau+1} - \zeta_{\tau+2})(\zeta_{-(\tau+2)} - \zeta_{-(\tau+1)}) \\ (\zeta_{-\tau} - \zeta_{-(\tau+1)})(Z_{\tau+2} - Z_{\tau+1}) + (Z_{\tau+1} - Z_{\tau})(\zeta_{-(\tau+2)} - \zeta_{-(\tau+1)}) \end{pmatrix}$$

with $|\Delta| = (Z_{\tau+1} - Z_{\tau})(z_{\tau+2} - z_{\tau+1}) - (\zeta_{-(\tau+1)} - \zeta_{-\tau})(\zeta_{\tau+2} - \zeta_{\tau+1})$ provided $|\Delta| \neq 0$.

(ii) If $T(1) \neq \phi$, $T(\Delta t) \neq \phi$, $T(\Delta t + 1) \neq \phi$ (US Spacing), then (γ, β) is identified by

$$\begin{pmatrix} \gamma \\ \beta \end{pmatrix} = \frac{1}{|\Delta|} \begin{pmatrix} (z_{\Delta t+1} - z_1)(Z_{\Delta t+1} - Z_1) + (\zeta_1 - \zeta_{\Delta t+1})(\zeta_{-(\Delta t+1)} - \zeta_{-1}) \\ (\zeta_0 - \zeta_{-\Delta t})(Z_{\Delta t+1} - Z_1) + (Z_{\Delta t} - Z_0)(\zeta_{-(\Delta t+1)} - \zeta_{-1}) \end{pmatrix}$$

with $|\Delta| = (Z_{\Delta t} - Z_0)(z_{\Delta t+1} - z_1) - (\zeta_{-\Delta t} - \zeta_0)(\zeta_{\Delta t+1} - \zeta_1)$ provided $|\Delta| \neq 0$.

[2] In this panel data, the first two data points $t = 1$ and 2 are also available, but they correspond to Kindergarten which is structurally different from the grade school periods starting at $t = 3$. We thus exclude the first two periods from the set T of time periods.

4 Estimation under Unequal Spacing

For the dynamic model (2.1) with two parameters (γ, β) with the two constructed cross-sectional moment restrictions (3.3) and (3.4), we achieve the explicit identifying formula (3.5) or its special cases provided in Corollary 1. Therefore, the sample counterparts of the identifying formulas in Corollary 1 suffice for consistent estimation of the parameters, where Z_τ, z_τ, and ζ_τ may be estimated by

$$\hat{Z}_\tau = \frac{1}{N}\sum_{i=1}^{N} \bar{Z}_{i\tau} \qquad \hat{z}_\tau = \frac{1}{N}\sum_{i=1}^{N} \bar{z}_{i\tau} \qquad \hat{\zeta}_\tau = \frac{1}{N}\sum_{i=1}^{N} \bar{\zeta}_{i\tau} \qquad \hat{\zeta}_{-\tau} = \frac{1}{N}\sum_{i=1}^{N} \bar{\zeta}_{i-\tau}$$

for any $\bar{Z}_{i\tau}$, $\bar{z}_{i\tau}$, $\bar{\zeta}_{i\tau}$, and $\bar{\zeta}_{i-\tau}$ such that $Z_\tau = \mathrm{E}[\bar{Z}_{i\tau}]$, $z_\tau = \mathrm{E}[\bar{z}_{i\tau}]$, $\zeta_\tau = \mathrm{E}[\bar{\zeta}_{i\tau}]$ and $\zeta_{-\tau} = \mathrm{E}[\bar{\zeta}_{i-\tau}]$. Due to the t-invariance (Assumption 2), such individual variables, $\bar{Z}_{i\tau}$, $\bar{z}_{i\tau}$, $\bar{\zeta}_{i\tau}$ and $\bar{\zeta}_{i-\tau}$, can be constructed by any linear combination of the form

$$\bar{Z}_{i\tau} = \sum_{t\in T(\tau)} a_t^\tau y_{it} y_{it+\tau} \quad \bar{z}_{i\tau} = \sum_{t\in T(\tau)} b_t^\tau x_{it} x_{it+\tau} \quad \bar{\zeta}_{i\tau} = \sum_{t\in T(\tau)} c_t^\tau y_{it} x_{it+\tau} \quad \bar{\zeta}_{i-\tau} = \sum_{t\in T(\tau)} d_t^\tau x_{it} y_{it+\tau}$$

where $a^\tau = (a_t^\tau)_{t\in T(\tau)}$, $b^\tau = (b_t^\tau)_{t\in T(\tau)}$, $c^\tau = (c_t^\tau)_{t\in T(\tau)}$ and $d^\tau = (d_t^\tau)_{t\in T(\tau)}$ satisfy $\sum_{t\in T(\tau)} a_t^\tau = 1$, $\sum_{t\in T(\tau)} b_t^\tau = 1$, $\sum_{t\in T(\tau)} c_t^\tau = 1$ and $\sum_{t\in T(\tau)} d_t^\tau = 1$. The sample counterpart of the identifying formula (3.5) yields the explicit estimator

$$\begin{pmatrix} \hat{\gamma} \\ \hat{\beta} \end{pmatrix} = \frac{1}{|\hat{\Delta}|} \begin{pmatrix} (\hat{z}_{\Delta t+\tau+1} - \hat{z}_{\tau+1})(\hat{Z}_{\Delta t+\tau+1} - \hat{Z}_{\tau+1}) + (\hat{\zeta}_{\tau+1} - \hat{\zeta}_{\Delta t+\tau+1})(\hat{\zeta}_{-(\Delta t+\tau+1)} - \hat{\zeta}_{-(\tau+1)}) \\ (\hat{\zeta}_{-\tau} - \hat{\zeta}_{-(\Delta t+\tau)})(\hat{Z}_{\Delta t+\tau+1} - \hat{Z}_{\tau+1}) + (\hat{Z}_{\Delta t+\tau} - \hat{Z}_{\tau})(\hat{\zeta}_{-(\Delta t+\tau+1)} - \hat{\zeta}_{-(\tau+1)}) \end{pmatrix}$$

where $|\hat{\Delta}| = (\hat{Z}_{\Delta t+\tau} - \hat{Z}_\tau)(\hat{z}_{\Delta t+\tau+1} - \hat{z}_{\tau+1}) - (\hat{\zeta}_{-(\Delta t+\tau)} - \hat{\zeta}_{-\tau})(\hat{\zeta}_{\Delta t+\tau+1} - \hat{\zeta}_{\tau+1})$.

While this explicit sample-analog estimator may be appealing for ease of implementation, the generalized method of moments (GMM) allows for a more general treatment with potential efficiency gains. The generic GMM restriction is provided by

$$E(\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta}_0)) = \mathbf{0} \tag{4.1}$$

where $\mathbf{g}$ is a vector of functions, $\mathbf{w}_i$ is a vector of observed variables for cross-sectional observation i, and $\boldsymbol{\theta}_\mathbf{0}$ is the true parameter vector. Our restrictions (3.3) and (3.4) under Assumption 2 can be represented by this generic moment restriction (4.1) with the following expressions consisting the rows of $\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta})$.

$$\begin{aligned} g_{1tt't''t'''}(\mathbf{w}_i, \boldsymbol{\theta}) &= [y_{it'''}y_{it'''+\Delta t+\tau+1} - y_{it'}y_{it'+\tau+1}] \\ &\quad -\gamma[y_{it''}y_{it''+\Delta t+\tau} - y_{it}y_{it+\tau}] - \beta[y_{it'''}x_{it'''+\Delta t+\tau+1} - y_{it'}x_{it'+\tau+1}] \quad (4.2) \\ g_{2tt't''t'''}(\mathbf{w}_i, \boldsymbol{\theta}) &= [x_{it'''}y_{it'''+\Delta t+\tau+1} - x_{it'}y_{it'+\tau+1}] \\ &\quad -\gamma[x_{it''}y_{it''+\Delta t+\tau} - x_{it}y_{it+\tau}] - \beta[x_{it'''}x_{it'''+\Delta t+\tau+1} - x_{it'}x_{it'+\tau+1}] \quad (4.3) \end{aligned}$$

for any $(t, t', t'', t''') \in T(\tau) \times T(\tau+1) \times T(\Delta t+\tau) \times T(\Delta t+\tau+1)$, where the parameter vector is $\boldsymbol{\theta} = (\gamma, \beta)$ and the individual data is $\mathbf{w}_i = (x_{it}, y_{it})_{t \in T}$. As such, our moment function $\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta})$ consists of $2 \times |T(\tau)| \times |T(\tau+1)| \times |T(\Delta t+\tau)| \times |T(\Delta t+\tau+1)|$ rows to estimate two parameters (γ, β). This cardinality relation is of course consistent with Theorem 1, which requires $T(\tau) \neq \phi$, $T(\tau+1) \neq \phi$, $T(\Delta t+\tau) \neq \phi$, and $T(\Delta t+\tau+1) \neq \phi$ for identification.

Given the moment restriction (4.1) consisting of expressions of the forms (3.3) and (3.4), the GMM estimator is given by

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} \in \Theta} \left[\frac{1}{N}\sum_{i=1}^{N} \mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta})\right]' \mathbf{W}_N \left[\frac{1}{N}\sum_{i=1}^{N} \mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta})\right]$$

where $\mathbf{W}_N$ is a weighting matrix. We provide low-level sufficient conditions tailored to our model for asymptotic normality of this estimator.

Assumption 4 (Asymptotic Normality)**.** *The following conditions are satisfied.*

(i) Panel data is i.i.d. across i.

(ii) $(\gamma_0, \beta_0) \in int\Theta$, where Θ is compact in $\mathbb{R}^2$.

(iii) $(x_{it}, y_{it})_{t \in T}$ have bounded fourth moments.

(iv) $\mathbf{W}_N \xrightarrow{p} \mathbf{W}$, where $\mathbf{W}$ is a positive definite matrix.

Theorem 2 (Asymptotic Normality)**.** *Suppose that Assumptions 1, 2, and 3 are satisfied for the model (2.1). If Assumption 4 is satisfied, then the GMM estimator based on the moment function $\mathbf{g}$ consisting of (4.2) and (4.3) is asymptotically normal:*

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, (\mathbf{G}'\mathbf{W}\mathbf{G})^{-1}\mathbf{G}'\mathbf{W}\mathbf{S}\mathbf{W}\mathbf{G}(\mathbf{G}'\mathbf{W}\mathbf{G})^{-1})$$

as $N \to \infty$, where $\mathbf{S}$ is finite variance matrix of $\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta}_0)$ and $\mathbf{G}$ is given by

$$\mathbf{G} = E \begin{bmatrix} \vdots & \vdots \\ -[y_{it''} y_{i,t''+\Delta t+\tau} - y_{it} y_{it+\tau}] & -[y_{it'''} x_{it'''+\Delta t+\tau+1} - y_{it'} x_{it'+\tau+1}] \\ \vdots & \vdots \\ -[x_{it''} y_{it''+\Delta t+\tau} - x_{it} y_{it+\tau}] & -[x_{it'''} x_{it'''+\Delta t+\tau+1} - x_{it'} x_{it'+\tau+1}] \\ \vdots & \vdots \end{bmatrix},$$

which is a matrix of dimension $2\,|T(\tau)|\,|T(\tau+1)|\,|T(\Delta t+\tau)|\,|T(\Delta t+\tau+1)|$ by 2.

A proof of this result is provided in Section A.1 in the appendix.

Example 3 (GMM Estimation under US Spacing)**.** *Suppose that panel data $(y_{it}, x_{it})_{t \in T}$ has the unequal spacing structure given by $T = \{1, 2, 5\}$. In this case, the set of gaps is $\mathcal{T} = \{0, 1, 3, 4\}$. Since $T(1) \neq \phi$, $T(3) \neq \phi$, and $T(4) \neq \phi$, the panel data exhibits the US spacing (Definition 2) with $\tau = 0$ and $\Delta t = 3$. Specifically, $T(0) = T$, $T(1) = \{1\}$, $T(3) = \{2\}$, and $T(4) = \{1\}$. Therefore, the GMM function $\mathbf{g}$ consisting of rows of the forms (4.2) and (4.3) with $t \in T$, $t' = 1$, $t'' = 2$ and $t''' = 1$ is explicitly given by the vector*

$$\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta}) = \begin{pmatrix} [y_{i1}y_{i5} - y_{i1}y_{i2}] - \gamma[y_{i2}y_{i5} - y_{i1}y_{i1}] - \beta[y_{i1}x_{i5} - y_{i1}x_{i2}] \\ [y_{i1}y_{i5} - y_{i1}y_{i2}] - \gamma[y_{i2}y_{i5} - y_{i2}y_{i2}] - \beta[y_{i1}x_{i5} - y_{i1}x_{i2}] \\ [y_{i1}y_{i5} - y_{i1}y_{i2}] - \gamma[y_{i2}y_{i5} - y_{i5}y_{i5}] - \beta[y_{i1}x_{i5} - y_{i1}x_{i2}] \\ [x_{i1}y_{i5} - x_{i1}y_{i2}] - \gamma[x_{i2}y_{i5} - x_{i1}y_{i1}] - \beta[x_{i1}x_{i5} - x_{i1}x_{i2}] \\ [x_{i1}y_{i5} - x_{i1}y_{i2}] - \gamma[x_{i2}y_{i5} - x_{i2}y_{i2}] - \beta[x_{i1}x_{i5} - x_{i1}x_{i2}] \\ [x_{i1}y_{i5} - x_{i1}y_{i2}] - \gamma[x_{i2}y_{i5} - x_{i5}y_{i5}] - \beta[x_{i1}x_{i5} - x_{i1}x_{i2}] \end{pmatrix}$$

of $2\,|T(0)|\,|T(1)|\,|T(3)|\,|T(4)| = 6$ rows. □

5 Finite Sample Performance

5.1 Data Generating Processes

Our Monte Carlo design is based on a linear vector autoregression of (y_{it}, x_{it}). Specifically, the data generating process is given by

$$
\begin{aligned}
y_{it} &= \gamma y_{it-1} + \beta x_{it} + \alpha_i + \varepsilon_{it} \\
x_{it} &= \rho x_{it-1} + \alpha_i + \xi_{it}
\end{aligned}
$$

By construction of the dynamic model, the lagged regressor y_{it-1} is endogenous, i.e., it is dependent on α_i. In addition, the common fixed effect α_i between the two equations makes x_{it} an endogenous covariate too. Throughout our Monte Carlo experiments, we use $\rho = 0.5$. The error terms are generated independently across i and t according to

$$
\alpha_i \sim N(0,1), \qquad \epsilon_{it} \sim N(0,1), \qquad \xi_{it} \sim N(0,1),
$$

After generation of the initial values (y_{it}, x_{it}), we wait for ten time periods to pass so we can ensure that the artificial data 'enter' the stationary distribution required by our baseline identification theorem.[3] For the core parameters (γ, β), we employ the five different DGP scenarios listed in Table 1 (A). In addition, we consider the four patterns of unequal spacing listed in Table 1 (B). We note that each of these spacing patters satisfies our requirement for identification stated in Theorem 1.

5.2 Simulation Results

We run 1,000 Monte Carlo replications with $N = 1,000$ units of cross-sectional observations for each combination of the DGP parameters and spacing patterns listed in Table 1. Results based on the continuously updating GMM estimator (CUE)[4] are summarized in Tables 2, 3, 4, and 5 for the spacing patterns 1, 2, 3, and 4, respectively. In each of these tables, we report the

[3]In our preliminary analysis, we also waited for 100 time periods, but the results are similar. In the interest of time, therefore, we choose to run only ten preliminary iterations to establish the stationarity requirement.

[4]For the spacing pattern 4, the model is just-identified and thus we use the closed-form estimator instead.

standard evaluation criteria, including the mean, the bias, the standard deviation, the mean absolute error, the root mean squared error, the 90% coverage rate, and the 95% coverage rate.

For the US spacing patterns (Tables 2–4), the estimation results under DGPs 1, 2, and 4 tend to behave well overall. The results under DGPs 3 and 5, on the other hand, produce larger biases and variances. The coverage probabilities under DGPs 3 and 5 are also biased relative to the designed probabilities. Likewise, the results under the UK spacing pattern shown in Table 5 do not look better than those under the US spacing patterns. We impute these results to weak satisfaction of the rank condition, i.e., Assumption 3, as opposed to the other assumptions. As mentioned previously, this assumptions is empirically testable. Using the matrix rank test of Kleibergen and Paap (2006), we count the number of times that we reject the null hypothesis of reduced rank against the alternative of full rank. The last column in each of the tables lists the full rank rate (FRR) produced by this number divided by the number of Monte Carlo replications. Observe that the null hypothesis is seldom rejected under the UK spacing.

In light of the possibility that the rank condition may be only weakly satisfied, our natural question is whether there is a way to conduct robust inference. If hypothesis testing is our goal, then we can use some of the existing tools of weak-identification-robust tests. Applying our GMM restrictions to the K test (Kleibergen, 2005; Theorem 1), we produce the 90% and 95% coverage probabilities of (γ, β) robustly against the weak satisfaction of the rank condition. Table 6 shows the Monte Carlo results. Compared to the coverage probabilities displayed in Tables 2–5, the simulated coverage probabilities in Table 6 are far more accurately close to the designed probabilities, even for the UK spacing pattern. These results are based on the joint test for (γ, β), but we remark that similar inference may be conducted for each of the two parameters using Theorem 2 of Kleibergen (2005).

In summary, our Monte Carlo results support our identification and estimation theories. In case results do not appear well, it is due only to a weak satisfaction of the rank condition (Assumption 3), and one can still use robust inference tools to overcome these limitations.

6 Extensions

We focused on the simple baseline model in the previous sections for clarity of exposition to illustrate our identification and estimation approaches. However, this model is too restrictive with several respects to be useful in practice. In the current section, we present how our identification strategy applies to general classes of models by relaxing the previous restrictions.

6.1 Time Varying Means and Variances

One restrictive feature of the baseline result is the requirement of the weak stationarity stated in Assumption 2. This assumption prohibits time-varying means and variances, whereas time variations are fairly common for many economic variables, particularly those that grow or accumulate over time. One way to incorporate time variations in our model is through location-/scale-normalized random variables.

Let y_{it}^* and x_{it}^* denote the true state variable and the true covariate, respectively. For short-hand notations, we define the following cross-sectional moments for each time period t.

$$\mu_t^y = E(y_{it}^*) \qquad \mu_t^y = E(x_{it}^*) \qquad \delta_t^y = Var(y_{it}^*) \qquad \delta_t^x = Var(x_{it}^*). \tag{6.1}$$

We define the location-/scale-normalized versions of the true variables by

$$y_{it} = \frac{y_{it}^* - \mu_t^y}{\sqrt{\delta_t^y}} \qquad x_{it} = \frac{x_{it}^* - \mu_t^x}{\sqrt{\delta_t^x}}, \tag{6.2}$$

and consider the model (2.1) for these normalized random variables equipped with Assumptions 1, 2, and 3. With this modification, the true variables y_{it}^* and x_{it}^* may exhibit time variations in both means and variances. Clearly, the baseline model (2.1) is a special case with $\mu_t^y = \mu_t^y$ and $\mu_t^x = \mu_{t'}^x$ for all $t, t' \in T$, and $\delta_t^y = \delta_t^x = 1$ for all t.

To understand the primitive dynamic process under the current location-/scale-normalization, substitute (6.2) in (2.1) to produce

$$y_{it}^* = \gamma_t y_{it-1}^* + \beta_t x_{it}^* + \alpha_i v_t + w_t + \epsilon_{it}, \tag{6.3}$$

where

$$\gamma_t = \frac{\gamma\sqrt{\delta_t^y}}{\sqrt{\delta_{t-1}^y}}, \quad \beta_t = \frac{\beta\sqrt{\delta_t^y}}{\sqrt{\delta_t^x}}, \quad v_t = \sqrt{\delta_t^y}, \quad w_t = \mu_t^y - \frac{\gamma\mu_{t-1}^y\sqrt{\delta_t^y}}{\sqrt{\delta_{t-1}^y}} - \frac{\beta\mu_t^x\sqrt{\delta_t^y}}{\sqrt{\delta_t^x}}, \tag{6.4}$$

and $\epsilon_{it} = \sqrt{\delta_t^y}\varepsilon_{it}$. Note that this primitive model is a dynamic panel model with interactive fixed effects $\alpha_i v_t$ as well as time-fixed effect w_t, studied in the panel data literature (Holtz-Eakin, Newey, and Rosen, 1988; Ahn, Lee, and Schmidt, 2001; Bai, 2009).

Identification of the primitive structural parameters (γ_t, β_t) follows from identification of (γ, β) and the moments in (6.1). Clearly, the moments in (6.1) are identified for each $t \in T$ by the observed panel data (y_{it}^*, x_{it}^*). Therefore, the normalized random variables (y_{it}, x_{it}) can be constructed for each $t \in T$ by (6.2). Once we construct the normalized random variables (y_{it}, x_{it}) which we assume to follow the baseline process (2.1) with Assumptions 1, 2, and 3, we can apply Theorem 1 to identify (γ, β). Specifically, (γ, β) is identified by

$$\begin{pmatrix}\gamma \\ \beta\end{pmatrix} = \frac{1}{|\Delta|}\begin{pmatrix}(z_{\Delta t+\tau+1} - z_{\tau+1})(Z_{\Delta t+\tau+1} - Z_{\tau+1}) + (\zeta_{\tau+1} - \zeta_{\Delta t+\tau+1})(\zeta_{-(\Delta t+\tau+1)} - \zeta_{-(\tau+1)}) \\ (\zeta_{-\tau} - \zeta_{-(\Delta t+\tau)})(Z_{\Delta t+\tau+1} - Z_{\tau+1}) + (Z_{\Delta t+\tau} - Z_{\tau})(\zeta_{-(\Delta t+\tau+1)} - \zeta_{-(\tau+1)})\end{pmatrix} \tag{6.5}$$

where

$$\begin{aligned}
|\Delta| &= (Z_{\Delta t+\tau} - Z_{\tau})(z_{\Delta t+\tau+1} - z_{\tau+1}) - (\zeta_{-(\Delta t+\tau)} - \zeta_{-\tau})(\zeta_{\Delta t+\tau+1} - \zeta_{\tau+1}) \\
Z_{i\tau} &= \mathrm{E}_i\left[\frac{(y_{it}^* - \mathrm{E}[y_{it}^*])(y_{it+\tau}^* - \mathrm{E}[y_{it+\tau}^*])}{\sqrt{\mathrm{Var}(y_{it}^*)}\sqrt{\mathrm{Var}(y_{it+\tau}^*)}}\right], \qquad Z_\tau = \mathrm{E}(Z_{i\tau}) \\
z_{i\tau} &= \mathrm{E}_i\left[\frac{(x_{it}^* - \mathrm{E}[x_{it}^*])(x_{it+\tau}^* - \mathrm{E}[x_{it+\tau}^*])}{\sqrt{\mathrm{Var}(x_{it}^*)}\sqrt{\mathrm{Var}(x_{it+\tau}^*)}}\right], \qquad z_\tau = \mathrm{E}(z_{i\tau}) \\
\zeta_{i\tau} &= \mathrm{E}_i\left[\frac{(y_{it}^* - \mathrm{E}[y_{it}^*])(x_{it+\tau}^* - \mathrm{E}[x_{it+\tau}^*])}{\sqrt{\mathrm{Var}(y_{it}^*)}\sqrt{\mathrm{Var}(x_{it+\tau}^*)}}\right], \qquad \zeta_\tau = \mathrm{E}(\zeta_{i\tau}) \\
\zeta_{i-\tau} &= \mathrm{E}_i\left[\frac{(x_{it}^* - \mathrm{E}[x_{it}^*])(y_{it+\tau}^* - \mathrm{E}[y_{it+\tau}^*])}{\sqrt{\mathrm{Var}(x_{it}^*)}\sqrt{\mathrm{Var}(y_{it+\tau}^*)}}\right], \qquad \zeta_{-\tau} = \mathrm{E}(\zeta_{i-\tau}).
\end{aligned} \tag{6.6}$$

We summarize this result as a theorem below.

Theorem 3 (Identification). *If Assumptions 1, 2, and 3 are satisfied for (2.1) with the location-/scale-normalized variables (y_{it}, x_{it}) defined in (6.2), and unequally spaced panel data have $T(\tau) \neq \phi$, $T(\tau+1) \neq \phi$, $T(\Delta t + \tau) \neq \phi$, and $T(\Delta t + \tau + 1) \neq \phi$, then (γ, β) is identified by the formula (6.5) with (6.6).*

Since the cross-sectional moments in (6.1) are identified for each $t \in T$, the time-varying coefficients $(\gamma_t, \beta_t, v_t, w_t)$ in the primitive model (6.3) are identified by (6.4) for each $t \in T$. This result is stated as a corollary below.

Corollary 2. *If Assumptions 1, 2, and 3 are satisfied for (2.1) with the location-/scale-normalized variables* (y_{it}, x_{it}) *defined in (6.2), and unequally spaced panel data have* $T(\tau) \neq \phi$, $T(\tau+1) \neq \phi$, $T(\Delta t + \tau) \neq \phi$, *and* $T(\Delta t + \tau + 1) \neq \phi$, *then* $(\gamma_t, \beta_t, v_t, w_t)$ *is identified by the formula (6.4) for each* $t \in T$.

To estimate the parameters, we stack the moment restrictions (6.1) to the GMM criterion. In other words, with the parameter vector $\boldsymbol{\theta} = (\gamma, \beta, (\mu_t^y)_{t\in T}, (\mu_t^x)_{t\in T}, (\delta_t^y)_{t\in T}, (\delta_t^x)_{t\in T})$, the vector $\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta})$ of moment functions consists of (4.2) and (4.3) concatenated with the additional $4 \times |T|$ rows of the form

$$\begin{aligned} g_{3t} &= y_{it}^* - \mu_t^y & \qquad g_{5t} &= (y_{it}^* - \mu_t^y)^2 - \delta_t^y \\ g_{4t} &= x_{it}^* - \mu_t^x & \qquad g_{6t} &= (x_{it}^* - \mu_t^x)^2 - \delta_t^x \end{aligned}$$

for all $t \in T$. We thus obtain $2 \times |T(\tau)| \times |T(\tau+1)| \times |T(\Delta t + \tau)| \times |T(\Delta t + \tau + 1)| + 4\,|T|$ restrictions to estimate $2 + 4\,|T|$ parameters.

Example 3′ (GMM Estimation under US Spacing with Time-Varying Means and Variances)**.** We continue with Example 3, but now relax the stationarity assumption by allowing time-varying means and variances. In this case, the vector $\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta})$ of moment functions can be constructed by

$$\begin{pmatrix} [y_{i1}y_{i5} - y_{i1}y_{i2}] - \gamma[y_{i2}y_{i5} - y_{i1}y_{i1}] - \beta[y_{i1}x_{i5} - y_{i1}x_{i2}] \\ [y_{i1}y_{i5} - y_{i1}y_{i2}] - \gamma[y_{i2}y_{i5} - y_{i2}y_{i2}] - \beta[y_{i1}x_{i5} - y_{i1}x_{i2}] \\ [y_{i1}y_{i5} - y_{i1}y_{i2}] - \gamma[y_{i2}y_{i5} - y_{i5}y_{i5}] - \beta[y_{i1}x_{i5} - y_{i1}x_{i2}] \\ [x_{i1}y_{i5} - x_{i1}y_{i2}] - \gamma[x_{i2}y_{i5} - x_{i1}y_{i1}] - \beta[x_{i1}x_{i5} - x_{i1}x_{i2}] \\ [x_{i1}y_{i5} - x_{i1}y_{i2}] - \gamma[x_{i2}y_{i5} - x_{i2}y_{i2}] - \beta[x_{i1}x_{i5} - x_{i1}x_{i2}] \\ [x_{i1}y_{i5} - x_{i1}y_{i2}] - \gamma[x_{i2}y_{i5} - x_{i5}y_{i5}] - \beta[x_{i1}x_{i5} - x_{i1}x_{i2}] \end{pmatrix}$$

of $2\,|T(0)|\,|T(1)|\,|T(3)|\,|T(4)| = 6$ rows obtained in Example 3, concatenated with

$$\begin{aligned} \big(&y_{i1}^* - \mu_1^y, \quad y_{i2}^* - \mu_2^y, \quad y_{i5}^* - \mu_5^y, \quad (y_{i1}^* - \mu_1^y)^2 - \delta_1^y, \quad (y_{i2}^* - \mu_2^y)^2 - \delta_2^y, \quad (y_{i5}^* - \mu_5^y)^2 - \delta_5^y, \\ &x_{i1}^* - \mu_1^x, \quad x_{i2}^* - \mu_2^x, \quad x_{i5}^* - \mu_5^x, \quad (x_{i1}^* - \mu_1^x)^2 - \delta_1^x, \quad (x_{i2}^* - \mu_2^x)^2 - \delta_2^x, \quad (x_{i5}^* - \mu_5^x)^2 - \delta_5^x\big)' \end{aligned}$$

of additional $4\,|T| = 12$ rows. Thus, moment function $\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta})$ has the total of 18 rows to estimate the total of 14 parameters, $\boldsymbol{\theta} = (\gamma, \beta, \mu_1^y, \mu_2^y, \mu_5^y, \mu_1^x, \mu_2^x, \mu_5^x, \delta_1^y, \delta_2^y, \delta_5^y, \delta_1^x, \delta_2^x, \delta_5^x)$. □

We remark that the asymptotic distributions of the time-varying parameters $(\gamma_t, \beta_t, v_t, w_t)$ of the primitive process (6.3) can be obtained by applying the Delta method to the transformation formulas provided in (6.4).

6.2 Multiple Covariates and Higher Order Process

The identification and estimation theories developed for the baseline model (2.1) can be extended to models with multivariate covariates and higher order autoregressive process. This extension does not induce any theoretical complication relative to the baseline model. However, higher-order processes require somewhat richer data structure while still allowing for unequal spacing. Consider the model with K-variate covariates and p-th order autoregressive process:

$$y_{it} = \sum_{j=1}^{p} \gamma_j y_{it-j} + X_{it}\boldsymbol{\beta} + \alpha_i + \varepsilon_{it} \tag{6.7}$$

where $X_{it} = (x_{1it}, x_{2it}, \cdots, x_{Kit})$ denotes the covariate vector for individual i at time t. This vector is accompanied by the parameter vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots, \beta_K)'$.

We can derive identification in the same way as the one we presented in Section 3 for the baseline model (2.1). To avoid cumbersome subscript notations, we proceed with the case of $\Delta t = 1$ and $\tau = 0$ in this section. We remark that similar arguments will also follow with other choices of Δt and τ. Thus, take the first difference, $\Delta t = 1$, of (6.7) to get

$$y_{it} - y_{it-1} = \sum_{j=1}^{p} \gamma_j (y_{it-j} - y_{it-1-j}) + (X_{it} - X_{it-1})\boldsymbol{\beta} + \varepsilon_{it}$$

Multiply this equation separately by the vectors $Y' = (y_{it-2}, y_{it-3}, \cdots, y_{it-p}, y_{it-p-1})'$ and $X'_{it-2} = (x_{1it-2}, x_{2it-2}, \cdots, x_{Kit-2})'$, and taking the first moment for each individual i yield

$$\begin{aligned} E_i[Y'(y_{it} - y_{it-1})] &= \sum_{j=1}^{p} \gamma_j E_i[Y'(y_{it-j} - y_{it-1-j})] + E_i[Y'(X_{it} - X_{it-1})]\boldsymbol{\beta} \\ E_i[X'_{it-2}(y_{it} - y_{it-1})] &= \sum_{j=1}^{p} \gamma_j E_i[X'_{it-2}(y_{it-j} - y_{it-1-j})] + E_i[X'_{it-2}(X_{it} - X_{it-1})]\boldsymbol{\beta} \end{aligned}$$

under Assumption 1. Taking the cross-sectional expectations on these equations under Assumption 2, we obtain the system of $p+K$ equations:

$$
\begin{aligned}
Z_2 - Z_1 &= \sum_{j=1}^{p} \gamma_j (Z_{|2-j|} - Z_{|1-j|}) + \sum_{k=1}^{K} (\zeta_{k,2} - \zeta_{k,1}) \beta_k \\
&\vdots \\
Z_{p+1} - Z_p &= \sum_{j=1}^{p} \gamma_j (Z_{|p+1-j|} - Z_{|p-j|}) + \sum_{k=1}^{K} (\zeta_{k,p+1} - \zeta_{k,p}) \beta_k \\
\zeta_{1,-2} - \zeta_{1,-1} &= \sum_{j=1}^{p} \gamma_j (\zeta_{1,j-2} - \zeta_{1,j-1}) + \sum_{k=1}^{K} (z_{1k,2} - z_{1k,1}) \beta_k \\
&\vdots \\
\zeta_{K,-2} - \zeta_{K,-1} &= \sum_{j=1}^{p} \gamma_j (\zeta_{K,j-2} - \zeta_{K,j-1}) + \sum_{k=1}^{K} (z_{Kk,2} - z_{Kk,1}) \beta_k
\end{aligned}
\tag{6.8}
$$

consisting of the t-invariant cross-sectional moments $Z_1, \cdots, Z_{p+1}$ defined in Section 3, as well as the newly defined t-invariant cross-sectional moments:

$$
z_{\kappa k,\tau} = E(x_{\kappa it} x_{kit+\tau}) \qquad \zeta_{k,\tau} = E(y_{it} x_{kit+\tau}) \qquad \zeta_{k,-\tau} = E(x_{kit} y_{it+\tau})
$$

The cross-sectional moments, $Z_1, \cdots, Z_{p+1}$, $\zeta_{k,-2}, \zeta_{k,-1}, \cdots, \zeta_{k,p-1}$, $z_{\kappa k,1}$, and $z_{\kappa k,2}$, included in the system 6.8 are fully observed if $T(1) \neq \phi$, $\cdots$, $T(p+1) \neq \phi$ are true. Given that all the relevant cross-sectional moments are observed, the $(p+K)$-dimensional parameter vector can be identified through the system (6.8) of $p+K$ equations, provided that the following empirically testable rank condition is satisfied.

Assumption 3′ (Empirically Testable Rank Condition)**.** (6.8) admits a unique solution.

Theorem 4 (Identification)**.** *If Assumptions 1, 2, and 3′ are satisfied for (6.7), and unequally spaced panel data have* $T(1) \neq \phi$, $\cdots$, *and* $T(p+1) \neq \phi$, *then the parameter vector* $(\gamma_1, \cdots, \gamma_p, \boldsymbol{\beta}')'$ *is identified by the solution to the system (6.8).*

6.3 Partially Linear Semiparametric Models

This paper has focused on parametric autoregressive equations so far. We now demonstrate how our approach under unequally spaced panels can similarly handle partially linear semiparametric models. Consider the first-order autoregressive model of the form

$$y_{it} = \gamma y_{it-1} + m(x_{it}) + \alpha_i + \varepsilon_{it} \tag{6.9}$$

where m is an unknown function that takes a scalar covariate x_{it}. We allow the covariate x_{it} to be correlated with the fixed effect α_i.

This class of models has been studied in the literature under the setting of regularly spaced panel data. With exogenous covariates, Baltagi and Li (2002a) propose a kernel-based instrumental variable estimator, while Baltagi and Li (2002b) propose a series estimator. Lee (2013) considers a model without the strict exogeneity. Our major contribution to the existing literature is the capability of identifying this class of models even under unequal spacing in panel data. To this end, we strengthen Assumption 2, and in addition assume that the nonparametric function m is square integrable as follows.

Assumption 2′ (Stationarity)**.** For each individual $i = 1, 2, \cdots, N$: the distribution of (y_{it}, x_{it}) is jointly stationary.

Assumption 5 (Square Integrability)**.** $\int m(x)^2 dx$ *exists.*

Note that the Lebesgue $L^2(\mathbb{R})$ space is a separable Hilbert space, and therefore Assumption 5 guarantees that m can be represented by an orthogonal basis. For instance, we consider the sequence $h = (h_1(x), h_2(x), h_3(x), \cdots)$ of the Hermite polynomials.[5] It then follows that there exists a unique sequence $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \cdots)'$ such that

$$m(x) \ = \ \sum_{q=1}^{\infty} \beta_q h_q(x)$$

holds for all $x \in \mathbb{R}$. Substitute this series representation in (6.9) to write the model as

$$y_{it} = \gamma y_{it-1} + \sum_{q=1}^{\infty} \beta_q h_q(x_{it}) + \alpha_i + \varepsilon_{it}. \tag{6.10}$$

Let t_1 and t_2 be two time periods in T such that $t_1 > t_2$. Taking the difference of the dynamic equation (6.10) between these two time periods yields

$$y_{it_1} - y_{it_2} = \gamma(y_{it_1-1} - y_{it_2-1}) + \sum_{q=1}^{\infty} \beta_q (h_q(x_{it_1}) - h_q(x_{it_2})) + (\varepsilon_{it_1} - \varepsilon_{it_2})$$

[5]It consists of $h_1(x) = x$, $h_2(x) = x^2 - 1$, $h_3(x) = x^3 - 3x$, and so on. The general Hermite polynomials has the leading element $h_0(x) = 1$, but we omit the intercept term because the additive fixed effect captures it.

Multiplying both sides of this equation by $y_{it_2-1-\tau}$ and $h_r(x_{it_2-1-\tau})$ for $r = 1, 2, 3, \cdots$, we obtain the following equations.

$$\begin{aligned}
y_{it_2-1-\tau}(y_{it_1} - y_{it_2}) &= \gamma y_{it_2-1-\tau}(y_{it_1-1} - y_{it_2-1}) \\
&\quad + \sum_{q=1}^{\infty} \beta_q y_{it_2-1-\tau}(h_q(x_{it_1}) - h_q(x_{it_2})) + y_{it_2-1-\tau}(\varepsilon_{it_1} - \varepsilon_{it_2}) \\
h_r(x_{it_2-1-\tau})(y_{it_1} - y_{it_2}) &= \gamma h_r(x_{it_2-1-\tau})(y_{it_1-1} - y_{it_2-1}) \\
&\quad + \sum_{q=1}^{\infty} \beta_q h_r(x_{it_2-1-\tau})(h_q(x_{it_1}) - h_q(x_{it_2})) + h_r(x_{it_2-1-\tau})(\varepsilon_{it_1} - \varepsilon_{it_2})
\end{aligned}$$

where r runs across $1, 2, 3, \cdots$. We take the expectation E_i of the above two equations for each individual i as follows.

$$E_i(y_{it_2-1-\tau}(y_{it_1} - y_{it_2})) = \tag{6.11}$$
$$\gamma E_i(y_{it_2-1-\tau}(y_{it_1-1} - y_{it_2-1})) + \sum_{q=1}^{\infty} \beta_q E_i(y_{it_2-1-\tau}(h_q(x_{it_1}) - h_q(x_{it_2})))$$
$$E_i(h_r(x_{it_2-1-\tau})(y_{it_1} - y_{it_2})) = \tag{6.12}$$
$$\gamma E_i(h_r(x_{it_2-1-\tau})(y_{it_1-1} - y_{it_2-1})) + \sum_{q=1}^{\infty} \beta_q E_i(h_r(x_{it_2-1-\tau})(h_q(x_{it_1}) - h_q(x_{it_2})))$$

where r runs across $1, 2, 3, \cdots$, and $E_i(y_{it_2-1-\tau}(\varepsilon_{it_1} - \varepsilon_{it_2})) = 0$ and $E_i(h_r(x_{it_2-1-\tau})(\varepsilon_{it_1} - \varepsilon_{it_2})) = 0$ follow from Assumption 1.

The strong stationarity in Assumption $2'$ allows us to define the following t-invariant cross-sectional random variables for $r = 1, 2, 3, \cdots$ and $q = 1, 2, 3, \cdots$.

(i) $Z_{i\tau} := E_i(y_{it} y_{it+\tau})$, where $Z_{i\tau}$ is t-invariant.

(ii) $z_{i\tau}^{rq} := E_i(h_r(x_{it}) h_q(x_{it+\tau}))$, where $z_{i\tau}^{rq}$ is t-invariant.

(iii) $\zeta_{i\tau}^{q} := E_i(y_{it} h_q(x_{it+\tau}))$, where $\zeta_{i\tau}^{q}$ is t-invariant.

(iv) $\zeta_{i-\tau}^{r} := E_i(h_r(x_{it}) y_{it+\tau})$, where $\zeta_{i-\tau}^{r}$ is t-invariant.

With these properties implied by Assumptions 1 and $2'$, we can rewrite (6.11) and (6.12) as

$$\begin{aligned}
Z_{i\Delta t+\tau+1} - Z_{i\tau+1} &= \gamma(Z_{i\Delta t+\tau} - Z_{i\tau}) + \sum_{q=1}^{\infty} \beta_q (\zeta_{i\Delta t+\tau+1}^{q} - \zeta_{i\tau+1}^{q}) \qquad \text{and} \\
\zeta_{i-(\Delta t+\tau+1)}^{r} - \zeta_{i-(\tau+1)}^{r} &= \gamma(\zeta_{i-(\Delta t+\tau)}^{r} - \zeta_{i-\tau}^{r}) + \sum_{q=1}^{\infty} \beta_q (z_{i\Delta t+\tau+1}^{rq} - z_{i\tau+1}^{rq})
\end{aligned}$$

respectively, where $\Delta t = t_1 - t_2$ denotes the gap between the two time periods, t_1 and t_2. Taking the cross-sectional means E of each of the above two equations yields

$$
\begin{aligned}
Z_{\Delta t+\tau+1} - Z_{\tau+1} &= \gamma(Z_{\Delta t+\tau} - Z_{\tau}) + \sum_{q=1}^{\infty} \beta_q (\zeta^q_{\Delta t+\tau+1} - \zeta^q_{\tau+1}) && (6.13)\\
\zeta^r_{-(\Delta t+\tau+1)} - \zeta^r_{-(\tau+1)} &= \gamma(\zeta^r_{-(\Delta t+\tau)} - \zeta^r_{-\tau}) + \sum_{q=1}^{\infty} \beta_q (z^{rq}_{\Delta t+\tau+1} - z^{rq}_{\tau+1}) && (6.14)
\end{aligned}
$$

for $r = 1, 2, 3, \cdots$, where $Z_\tau := \mathrm{E}[Z_{i\tau}]$, $z^{rq}_\tau := \mathrm{E}[z^{rq}_{i\tau}]$, $\zeta^q_\tau := \mathrm{E}[\zeta^q_{i\tau}]$, and $\zeta^r_{-\tau} := \mathrm{E}[\zeta^r_{i-\tau}]$ for short-hand notations.

Equations (6.13) and (6.14) involve cross-sectional moments of the form $Z_{\Delta t+\tau+1}$, $Z_{\tau+1}$, $Z_{\Delta t+\tau}$, Z_τ, $\zeta^q_{\Delta t+\tau+1}$, $\zeta^q_{\tau+1}$, $\zeta^r_{-(\Delta t+\tau+1)}$, $\zeta^r_{-(\tau+1)}$, $\zeta^r_{-(\Delta t+\tau)}$, $\zeta^r_{-\tau}$, $z^{rq}_{\Delta t+\tau+1}$, and $z^{rq}_{\tau+1}$. Due to the t-invariance implied by Assumption 2′, the first one of these moments, $Z_{\Delta t+\tau+1}$, can be observed as the cross-sectional moment of $y_{it} y_{it+\Delta t+\tau+1}$ for any $t \in T(\Delta t + \tau + 1)$ provided that $T(\Delta t + \tau + 1) \neq \phi$ is true. Likewise, all the cross sectional moments in (6.13) and (6.14) can be observed using unequally spaced panel data if $T(\tau) \neq \phi$, $T(\tau + 1) \neq \phi$, $T(\Delta t + \tau) \neq \phi$, and $T(\Delta t + \tau + 1) \neq \phi$ are true.

Define the operator $K_{\Delta t,\tau} : l^2(\mathbb{Z}_+) \to l^2(\mathbb{Z}_+)$ for each $\boldsymbol{\phi} \in l^2(\mathbb{Z}_+)$ by

$$
K_{\Delta t,\tau}(\boldsymbol{\phi}) = \begin{pmatrix}
\phi_0(Z_{\Delta t+\tau} - Z_\tau) + \sum_{q=1}^{\infty} \phi_q(\zeta^q_{\Delta t+\tau+1} - \zeta^q_{\tau+1}) \\
\phi_0(\zeta^1_{-(\Delta t+\tau)} - \zeta^1_{-\tau}) + \sum_{q=1}^{\infty} \phi_q(z^{1q}_{\Delta t+\tau+1} - z^{1q}_{\tau+1}) \\
\phi_0(\zeta^2_{-(\Delta t+\tau)} - \zeta^2_{-\tau}) + \sum_{q=1}^{\infty} \phi_q(z^{2q}_{\Delta t+\tau+1} - z^{2q}_{\tau+1}) \\
\vdots
\end{pmatrix}.
$$

Once all the cross-sectional moments in (6.13) and (6.14) are observed for all $r = 1, 2, 3, \cdots$ and $q = 1, 2, 3, \cdots$ from unequally spaced panel data, we can solve the system to explicitly identify the structural parameters $(\gamma, \boldsymbol{\beta}')$ by

$$
\begin{pmatrix} \gamma \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \gamma \\ \beta_1 \\ \beta_2 \\ \vdots \end{pmatrix} = K^{-1}_{\Delta t,\tau} \begin{pmatrix}
Z_{\Delta t+\tau+1} - Z_{\tau+1} \\
\zeta^1_{-(\Delta t+\tau+1)} - \zeta^1_{-(\tau+1)} \\
\zeta^2_{-(\Delta t+\tau+1)} - \zeta^2_{-(\tau+1)} \\
\vdots
\end{pmatrix}, \qquad (6.15)
$$

provided that the following rank condition is satisfied.

Assumption 3″ (Rank Condition)**.** The operator $K_{\Delta t,\tau} : l^2(\mathbb{Z}_+) \to l^2(\mathbb{Z}_+)$ is invertible.

We remark that unlike the empirically testable rank condition in Assumption 3 used for the parametric autoregressive model (2.1), the current rank condition in Assumption 3″ is not testable. This identification result is summarized as a theorem below.

Theorem 5 (Identification)**.** *If Assumptions 1, 2′, 3″, and 5 are satisfied for (6.9), and unequally spaced panel data have* $T(\tau) \neq \phi$, $T(\tau+1) \neq \phi$, $T(\Delta t+\tau) \neq \phi$, *and* $T(\Delta t+\tau+1) \neq \phi$, *then* $(\gamma, \boldsymbol{\beta}')$ *is identified by the formula (6.15).*

This identification strategy features the Fredholm equation of the first kind, which suffers from the problem called the ill-posedness when data of finite sample is taken into the identifying equation. Carrasco, Florens, and Renault (2007; Ch. 3, 4) propose regularized solutions to the problem. Furthermore, the sieve semiparametric approach (Chen, 2007) is useful to obtain asymptotic normality of γ with the infinite-dimensional nuisance parameter $\boldsymbol{\beta}$ representing m.

7 Empirical Application

The panel autoregression with covariates of the form (2.1) or its extensions presented in Section 6 is one of popular models of earning dynamics, and has been used at least since Ashenfelter (1978).[6] The method-of-moment approaches are particularly useful for parameter estimation and testing, but they are not generally effective for unequally spaced panel data as argued in the introductory section. As a consequence, researchers may well tend to use only 'regular' panel data to study earning dynamics. One the other hand, we may conceivably find new empirical evidences by using unequally spaced panel data that have not been used by other researchers due to technical limitations.

In this section, we apply our methods to the NLS Original Cohorts: Older Men. Personal interviews were conducted in 1966, 67, 69, 71, 76, 81, and 90, and thus the set of gap years is given by $\mathcal{T} = \{0, 1, 2, 3, 4, 5, 7, 9, 10, 12, 14, 15, 19, 21, 23, 24\}$. This set satisfies the condition of identification for first-order autoregressive model with $(\tau, \Delta t) = (0, 2)$. In fact, the condition is still satisfied with $\mathcal{T} = \{0, 1, 2, 3\}$ and $(\tau, \Delta t) = (0, 2)$ even if we drop the years 1971, 76, 81, and 90. In this light, it was decided that we only use the survey responses in 1966, 67, and 69

[6] He uses this type of model for the objective of evaluating the effects of active labor market programs.

for two purposes. First, focusing on a shorter time span mitigates the effect of macroeconomic structural changes that may arise in the long run. Second, since a non-trivial subsample hits the retirement age after 1970, focusing on earlier time periods allows us to suffer less from the endogenous self-selection problems associated with post-retirement labor-leisure choice.

The earnings reported in years 1966, 67, and 69 includes wages, salary, commissions, or tips from all jobs in years 1965, 66, and 68, respectively. The observed individual characteristics include age, race, and education. The age is imputed from the reported birth year. The education measures the highest grade completed at the moment of the survey in 1966. After cleaning the data, we obtain the effective sample size of $N = 2,998$. Table 7 shows summary statistics.

While our models can accommodate additive covariates, most of the covariates being time-invariant will be differenced out like the additive fixed effects, and the associated parameters are unidentified. Therefore, instead of using them as additive covariates, we use the observed individual characteristics to form subpopulations, and estimate the model parameters for each of the constructed subpopulations. Using the method presented in Section 4, we first estimate the AR(1) coefficient for the baseline model (2.1) assuming the weak stationarity. Results are summarized in panel (A) of Table 8. The point estimates for γ range narrowly from .25 to .29. The rank conditions are likely to be satisfied according to the p-values for the reduced-rank test. Despite these robust results, an obvious disadvantage of the current analysis is the assumption of stationarity, as earnings are hardly considered to follow stationary processes.

Following the methodology suggested in Section 6.1, we next estimate the AR(1) parameters for the extended model (6.3) which is free from the previous stationarity assumption. Results are summarized in panel (B) of Table 8. The estimates for γ are the estimates of the AR(1) coefficient for the location-/scale-normalized log earnings, i.e., y_{it} defined in (6.2) as opposed to y_{it}^*. The estimates for μ_t^y and δ_t^y are the estimates of the t-varying means and variances of the actual log earnings defined in (6.1). Using the relations in (6.4) and these point estimates, we recover estimates for the AR(1) coefficients, γ_{66} and γ_{68}, for the actual log earnings y_{it}^*. The standard errors are computed using the Delta method. The rate γ_{66} for 1966 ranges from .35 to .53, and the rate γ_{68} for 1968 ranges from .34 to .59. These imputed AR(1) coefficients tend to be larger for whites, those with twelve or more years of education, and/or younger individuals. Our point estimates do not differ from the typical estimates obtained in earlier studies, and thus confirm the existing conclusions.[7] The attempt that we made at *actually* checking this result using the previously unused data set makes a modest but important contribution to solidifying the empirical knowledge of researchers.

8 Conclusions

Despite the abundance of theoretical and methodological solutions available in the 'regular' dynamic panel data literature, formal identification of fixed-effect dynamic models has not been addressed under unequally spaced panel data. In this paper, we propose that certain spacing patterns guarantee identification. Specifically, if $\mathcal{T} = \{|t_1 - t_2| : t_1, t_2 \in T\}$ denotes the set of survey gaps, then we obtain identification of the AR(1) parameters provided that $\tau, \tau+1, \Delta t+\tau, \Delta t+\tau+1 \in \mathcal{T}$ holds for some $\tau \geqslant 0$ and $\Delta t > 0$. This result extends to models with multiple covariates, higher-order dynamic processes, time-varying trends, and partially linear semiparametric models. The extension to models with time-varying trends is useful to alleviate the restrictive stationarity assumption required for the baseline model. We propose a GMM estimator and derive its asymptotic distribution following the standard argument. This paper contributes to the body of our knowledge and provides a guidance to practitioners by formally ensuring identification of dynamic fixed-effect models under the stylized patterns of unequally spaced panel data. In addition, our proposed estimators are readily available for empirical practitioners.[8]

References

Ahn, S.C., Y.H. Lee, and P. Schmidt (2001) "GMM Estimation of Linear Panel Data Models with Time-Varying Individual Effects," *Journal of Econometrics,* 101, 219–255.

[7] For instance, Hirano (2002) gives the range from .32 to .74, and Hu (2002) gives the range from .50 to .61.

[8] We are happy to share our matlab code upon request.

Ahn, S.C. and P. Schmidt (1995) “Efficient Estimation of Models for Dynamic Panel Data,” *Journal of Econometrics,* 68, 5–27.

Anderson, T.W. and C. Hsiao (1981) “Estimation of Dynamic Models with Error Components,” *Journal of the American Statistical Association,* 76, 598–606.

Arellano, M. and S. Bond (1991) “Some Tests of Speci.cation for Panel Data: Monte Carlo Evidence and an Application to Employment Equations,” *Review of Economic Studies,* 58, 277–297.

Ashenfelter, O. (1978) “Estimating the Effect of Training Programs on Earnings,” *Review of Economics and Statistics,* 60, 47–57.

Bai, J. (2009) “Panel Data Models with Interactive Fixed Effects,” *Econometrica,* 77, 1229–1279.

Baltagi, B.H. and Q. Li (2002a) “On Instrumental Variable Estimation of Semiparametric Dynamic Panel Data Models,” *Economics Letters,* 76, 1–9.

Baltagi, B.H. and D. Li (2002b) “Series Estimation of Partially Linear Panel Data Models with Fixed Effects,” *Annals of Economics and Finance,* 3, 103–116.

Baltagi, B.H. and P.X. Wu (1999) “Unequally Spaced Panel Data Regressions with AR(1) Disturbances,” *Econometric Theory,* 15, 814–823.

Blundell, R. and S. Bond (1998) “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models,” *Journal of Econometrics,* 87, 115–143.

Carrasco, M., J.-P. Florens, and E. Renault (2007) “Linear Inverse Problems in Structural Econometrics Estimation Based on Spectral Decomposition and Regularization,” in J.J. Heckman and E.E. Leamer (ed.) *Handbook of Econometrics*, 6, 5633–5751.

Chen, X. (2007) “Large Sample Sieve Estimation of Semi-Nonparametric Models,” in J.J. Heckman and E.E. Leamer (ed.) *Handbook of Econometrics*, 6, 5549–5632.

Hahn, J. (1999) “How Informative Is the Initial Condition in the Dynamic Panel Model with Fixed Effects?” *Journal of Econometrics,* 93, 309–326.

Hirano, K. (2002) “Semiparametric Bayesian Inference in Autoregressive Panel Data Models,” *Econometrica,* Vol. 70, pp. 781-799.

Holtz-Eakin, D., W. Newey and H.S. Rosen (1988) “Estimating Vector Autoregressions with Panel Data,” *Econometrica,* 56, 1371–1395.

Hu, L. (2002) “Estimation of a Censored Dynamic Panel Data Model,” *Econometrica,* Vol. 70, pp. 2499-2517.

Jones, R.H. and F. Boadi-Boateng (1991) “Unequally Spaced Longitudinal Data with AR(1) Serial Correlation,” *Biometrics,* 47, 161–175.

Kleibergen, F. (2005) “Testing Parameters in GMM Without Assuming that They Are Identified,” *Econometrica,* 73, pp. 1103–1123.

Kleibergen, F. and R. Paap (2006) “Generalized Reduced Rank Tests Using the Singular Value Decomposition,” *Journal of Econometrics,* 133, pp. 97–126.

Lee, Y. (2013) “Nonparametric Estimation of Dynamic Panel Models with Fixed Effects,” Working Paper. University of Michigan.

McKenzie, D.J. (2001) “Estimation of AR(1) Models with Unequally-Spaced Pseudo-Panels,” *Econometrics Journal,* 4, 89–108.

Millimet, D.L. and I.K. McDonough (2013) “Dynamic Panel Data Models with Irregular Spacing: With Applications to Early Childhood Development” IZA Discussion Paper No. 7359.

Newey, W.K. and D. McFadden (1994) “Large Sample Estimation and Hypothesis Testing,” in R.F. Engle and D.L. McFadden (ed.) *Handbook of Econometrics*, 4, 2111–2245.

Rosner, B. and A. Munoz (1988) “Autoregressive Modelling for the Analysis of Longitudinal Data with Unequally Spaced Examinations,” *Statistics in Medicine,* 7, 59–71.

A Appendix

A.1 Proof of Theorem 2

Proof. We show that our assumptions imply the high-level assumptions for the asymptotic normality of GMM estimators (Newey and McFadden, 1994; Theorem 3.4). Identification of $(\gamma_0, \beta_0) \in \Theta$ and $\mathrm{E}[\mathbf{g}(\mathbf{w}, \boldsymbol{\theta}_0)] = 0$ hold under Assumptions 1, 2, and 3 by Theorem 1. The cross-sectional i.i.d. is directly stated in Assumption 4 (i). Similarly, $(\gamma_0, \beta_0) \in \mathrm{int}\Theta$ and the compactness of Θ are directly stated in Assumption 4 (ii). The vector $\mathbf{g}$ of moment functions is continuously differentiable with its definitions given by (4.2) and (4.3). $\mathrm{E}[\|\mathbf{g}(\mathbf{w}, \boldsymbol{\theta}_0)\|^2] < \infty$ and $E[sup_{\boldsymbol{\theta}\in\Theta}\|\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta})\|] < \infty$ follow from the concrete expressions of $\mathbf{g}$ given by (4.2) and (4.3) under the compact parameter space of Assumption 4 (ii) and the bounded fourth moment of Assumption 4 (iii). Similarly, $E[sup_{\boldsymbol{\theta}\in\Theta}\|D_{\boldsymbol{\theta}}\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta})\|] < \infty$ follows from the concrete expressions of $\mathbf{g}$ given by (4.2) and (4.3) under Assumption 4 (iii). The random vector $\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta}_0)$ has finite variance matrix $\mathbf{S}$ under Assumption 4 (iii). The assumption that $\mathbf{W}_N \xrightarrow{p} \mathbf{W}$ is directly stated in Assumption 4 (iv). Lastly, the non-singularity of $\mathbf{G}'\mathbf{W}\mathbf{G}$ is implied by the positive-definiteness of $\mathbf{W}$ in Assumption 4 (iv) and by the fact that Assumption 3 implies that the $(2\,|T(\tau)|\,|T(\tau+1)|\,|T(\Delta t+\tau)|\,|T(\Delta t+\tau+1)|) \times 2$ matrix $\mathbf{G} = \mathrm{E}[D_{\boldsymbol{\theta}}\mathbf{g}(\mathbf{w}_i, \boldsymbol{\theta})]|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$ given with rows of the form

$$\begin{matrix} -[y_{it''}y_{i,t''+\Delta t+\tau} - y_{it}y_{it+\tau}] & \qquad -[y_{it'''}x_{it'''+\Delta t+\tau+1} - y_{it'}x_{it'+\tau+1}] \\ -[x_{it''}y_{it''+\Delta t+\tau} - x_{it}y_{it+\tau}] & \qquad -[x_{it'''}x_{it'''+\Delta t+\tau+1} - x_{it'}x_{it'+\tau+1}] \end{matrix}$$

has rank $\dim(\boldsymbol{\theta}) = 2$. Therefore, by Newey and McFadden (1994; Theorem 3.4), the stated conclusion holds. $\square$

A.2 Tables

(A) Data Generating Processes

	γ	β
DGP 1	0.75	0.25
DGP 2	0.50	0.50
DGP 3	0.25	0.75
DGP 4	0.75	0.75
DGP 5	0.25	0.25

(B) Patterns of Unequal Spacing

	Available Periods
Pattern 1: US Spacing	1, 2, 6
Pattern 2: US Spacing	1, 2, 5
Pattern 3: US Spacing	1, 2, 4
Pattern 4: UK Spacing	1, 3, 6, 10

Table 1: (A) Five pairs of data generating parameters and (B) four patterns of unequal spacing.

CUE

DGP		True	MC Mean	MC Bias	MC SD	MAE	RMSE	90% CR	95% CR	FRR
1	γ	0.75	0.7497	-0.0003	0.0384	0.0304	0.0384	0.9010	0.9490	1.0000
	β	0.25	0.2463	-0.0037	0.1027	0.0817	0.1027	0.9110	0.9590	
2	γ	0.50	0.5007	0.0007	0.0447	0.0362	0.0447	0.8880	0.9420	1.0000
	β	0.50	0.4994	-0.0006	0.1055	0.0835	0.1054	0.9050	0.9500	
3	γ	0.25	0.2521	0.0021	0.0511	0.0413	0.0512	0.8520	0.9110	1.0000
	β	0.75	0.7465	-0.0035	0.1163	0.0921	0.1163	0.8800	0.9380	
4	γ	0.75	0.7461	-0.0039	0.0431	0.0339	0.0433	0.8930	0.9350	1.0000
	β	0.75	0.7561	0.0061	0.1530	0.1201	0.1530	0.9000	0.9500	
5	γ	0.25	0.2497	-0.0003	0.0491	0.0395	0.0490	0.8600	0.9140	1.0000
	β	0.25	0.2537	0.0037	0.0885	0.0685	0.0885	0.9030	0.9560	

Table 2: Monte Carlo simulation results for the US spacing pattern 1 ($t = 1, 2, 6$).

CUE

DGP		True	MC Mean	MC Bias	MC SD	MAE	RMSE	90% CR	95% CR	FRR
1	γ	0.75	0.7495	-0.0005	0.0422	0.0340	0.0422	0.9030	0.9430	1.0000
	β	0.25	0.2483	-0.0017	0.1018	0.0804	0.1018	0.8870	0.9460	
2	γ	0.50	0.5007	0.0007	0.0471	0.0379	0.0471	0.8820	0.9310	1.0000
	β	0.50	0.5008	0.0008	0.1052	0.0829	0.1051	0.9030	0.9440	
3	γ	0.25	0.2529	0.0029	0.0519	0.0417	0.0519	0.8450	0.9000	1.0000
	β	0.75	0.7473	-0.0027	0.1164	0.0929	0.1164	0.8920	0.9430	
4	γ	0.75	0.7480	-0.0020	0.0456	0.0361	0.0456	0.9050	0.9350	1.0000
	β	0.75	0.7514	0.0014	0.1357	0.1077	0.1357	0.9010	0.9510	
5	γ	0.25	0.2524	0.0024	0.0494	0.0400	0.0495	0.8580	0.9220	1.0000
	β	0.25	0.2521	0.0021	0.0955	0.0748	0.0955	0.9070	0.9510	

Table 3: Monte Carlo simulation results for the US spacing pattern 2 ($t = 1, 2, 5$).

CUE

DGP		True	MC Mean	MC Bias	MC SD	MAE	RMSE	90% CR	95% CR	FRR
1	γ	0.75	0.7501	0.0001	0.0511	0.0406	0.0511	0.8920	0.9210	1.0000
	β	0.25	0.2492	-0.0008	0.1013	0.0801	0.1013	0.8930	0.9440	
2	γ	0.50	0.5005	0.0005	0.0507	0.0409	0.0507	0.8840	0.9420	1.0000
	β	0.50	0.5034	0.0034	0.1078	0.0850	0.1078	0.9040	0.9410	
3	γ	0.25	0.2513	0.0013	0.0533	0.0414	0.0533	0.8530	0.9010	1.0000
	β	0.75	0.7504	0.0004	0.1236	0.0989	0.1235	0.8810	0.9410	
4	γ	0.75	0.7482	-0.0018	0.0534	0.0423	0.0534	0.8860	0.9340	1.0000
	β	0.75	0.7506	0.0006	0.1225	0.0967	0.1224	0.9080	0.9490	
5	γ	0.25	0.2508	0.0008	0.0516	0.0408	0.0516	0.8520	0.9080	1.0000
	β	0.25	0.2533	0.0033	0.1095	0.0887	0.1095	0.9020	0.9480	

Table 4: Monte Carlo simulation results for the US spacing pattern 3 ($t = 1, 2, 4$).

GMM (Just Identified)

DGP		True	MC Mean	MC Bias	MC SD	MAE	RMSE	90% CR	95% CR	FRR
1	γ	0.75	-0.9987	-1.7487	61.9761	7.2989	61.9698	0.9280	0.9580	0.0150
	β	0.25	3.0046	2.7546	145.5485	12.9666	145.5018	0.9860	0.9960	
2	γ	0.50	-1.5413	-2.0413	65.1228	5.0320	65.1222	0.9190	0.9460	0.0300
	β	0.50	0.3391	-0.1609	41.6676	5.5829	41.6471	0.9610	0.9770	
3	γ	0.25	-0.7701	-1.0201	26.5555	2.4351	26.5618	0.9510	0.9800	0.0320
	β	0.75	0.2446	-0.5054	33.8003	3.3229	33.7871	0.9660	0.9860	
4	γ	0.75	1.2339	0.4839	63.0417	7.1836	63.0120	0.9100	0.9420	0.0210
	β	0.75	0.0020	-0.7480	135.2561	14.4443	135.1906	0.9780	0.9920	
5	γ	0.25	0.8350	0.5850	28.7486	3.8446	28.7402	0.9660	0.9800	0.0110
	β	0.25	0.9519	0.7019	24.4595	3.7175	24.4573	0.9750	0.9850	

Table 5: Monte Carlo simulation for the UK spacing pattern 4 ($t = 1, 3, 6, 10$).

DGP	US Spacing						UK Spacing	
	$t = 1, 2, 6$		$t = 1, 2, 5$		$t = 1, 2, 4$		$t = 1, 3, 6, 10$	
	90% KS	95% KS	90% KS	95% KS	90% KS	95% KS	90% KS	95% KS
1	0.9150	0.9550	0.9060	0.9500	0.8770	0.9360	0.8940	0.9390
2	0.8980	0.9560	0.8910	0.9530	0.8940	0.9380	0.8990	0.9500
3	0.9030	0.9470	0.8870	0.9450	0.8950	0.9470	0.8990	0.9500
4	0.8970	0.9480	0.9120	0.9580	0.8870	0.9440	0.9000	0.9520
5	0.8870	0.9390	0.8900	0.9530	0.8860	0.9440	0.8990	0.9460

Table 6: Monte Carlo simulation results for the coverage probabilities of the K statistic.

(A) Summary Statistics

Year	Log Earnings	Age	White	Education	Sample Size
1965	8.5849	50.9316	0.6828	9.4043	2998
	(0.7514)	(4.3073)	(0.4655)	(3.8957)	
1966	8.6452				
	(0.7239)				
1968	8.7508				
	(0.7418)				

(B) Composition

	Age at 1965 $\leqslant 50$		Age at 1965 > 50	
Education $\geqslant 12$	White	517	White	386
	Non-White	101	Non-White	72
Education < 12	White	504	White	640
	Non-White	344	Non-White	434

Table 7: Summary statistics of the NLS Original Cohorts: Older Men. (A) The displayed numbers are the sample means. The numbers in parentheses are the standard deviations. (B) The size of the eight subsamples categorized by the observed characteristics.

(A) Basic Model with Stationarity

	Full Sample	White	Non-White	Educ⩾12	Educ<12	Age⩽50	Age>50
γ	0.2693	0.2719	0.2652	0.2908	0.2533	0.2666	0.2719
	(0.0162)	(0.0199)	(0.0274)	(0.0212)	(0.0219)	(0.0213)	(0.0244)
Rank p-Val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

(B)Extended Model with Time-Varying Means and Variances

	Full Sample	White	Non-White	Educ⩾12	Educ<12	Age⩽50	Age>50
γ	0.4364	0.4741	0.3622	0.5276	0.4030	0.5296	0.3502
	(0.0693)	(0.0971)	(0.0910)	(0.1279)	(0.0754)	(0.0908)	(0.0951)
μ^y_{65}	8.5849	8.7776	8.1700	8.9589	8.3755	8.6556	8.5172
	(0.0137)	(0.0142)	(0.0259)	(0.0172)	(0.0174)	(0.0185)	(0.0201)
μ^y_{66}	8.6452	8.8304	8.2466	8.9928	8.4507	8.7220	8.5718
	(0.0132)	(0.0135)	(0.0254)	(0.0174)	(0.0166)	(0.0171)	(0.0198)
μ^y_{68}	8.7508	8.9227	8.3808	9.0985	8.5562	8.8284	8.6766
	(0.0135)	(0.0144)	(0.0255)	(0.0195)	(0.0165)	(0.0188)	(0.0192)
δ^y_{65}	0.5644	0.4124	0.6397	0.3195	0.5794	0.5008	0.6160
	(0.0284)	(0.0274)	(0.0554)	(0.0269)	(0.0370)	(0.0394)	(0.0406)
δ^y_{66}	0.5239	0.3734	0.6152	0.3256	0.5294	0.4298	0.6029
	(0.0276)	(0.0263)	(0.0541)	(0.0286)	(0.0361)	(0.0308)	(0.0442)
δ^y_{68}	0.5500	0.4256	0.6174	0.4091	0.5234	0.5200	0.5675
	(0.0315)	(0.0347)	(0.0563)	(0.0566)	(0.0353)	(0.0478)	(0.0412)
γ_{66}	0.4204	0.4512	0.3551	0.5326	0.3852	0.4906	0.3464
	(0.0654)	(0.0900)	(0.0882)	(0.1264)	(0.0706)	(0.0819)	(0.0927)
γ_{68}	0.4472	0.5062	0.3628	0.5914	0.4007	0.5825	0.3397
	(0.0780)	(0.1172)	(0.0975)	(0.1741)	(0.0805)	(0.1173)	(0.0966)
Rank p-Val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 8: Estimation results. The displayed numbers indicate the point estimates, and the numbers in parentheses indicate standard errors. Panel (A) shows estimates for the baseline model that assumes stationarity. Panel (B) shows estimates for the extended model with time-varying means and variances. The bottom row of each panel shows p-values for rank tests.