

Lee, Nayoung; Moon, Hyungsik Roger; Weidner, Martin

**Working Paper**

## Analysis of interactive fixed effects dynamic linear panel regression with measurement error

cemmap working paper, No. CWP37/11

**Provided in Cooperation with:**

The Institute for Fiscal Studies (IFS), London

*Suggested Citation:* Lee, Nayoung; Moon, Hyungsik Roger; Weidner, Martin (2011) : Analysis of interactive fixed effects dynamic linear panel regression with measurement error, cemmap working paper, No. CWP37/11, Centre for Microdata Methods and Practice (cemmap), London, <https://doi.org/10.1920/wp.cem.2011.3711>

This Version is available at:

<https://hdl.handle.net/10419/64657>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Analysis of interactive fixed effects dynamic linear panel regression with measurement error

---

Nayoung Lee  
Hyungsik Roger Moon  
Martin Weidner

The Institute for Fiscal Studies  
Department of Economics, UCL

**cemmap** working paper CWP37/11

# Analysis of Interactive Fixed Effects Dynamic Linear Panel Regression with Measurement Error\*

Nayoung Lee  
CUHK

Hyungsik Roger Moon  
U of Maryland

Martin Weidner  
UCL and CeMMAP

December 12, 2011

## Abstract

This paper studies a simple dynamic panel linear regression model with interactive fixed effects in which the variable of interest is measured with error. To estimate the dynamic coefficient, we consider the least-squares minimum distance (LS-MD) estimation method.

Keywords: dynamic panel, interactive fixed effects, measurement error, LS-MD estimation.

JEL Classification: C23, C26

## 1 Introduction

This paper studies a *simple* dynamic panel linear regression model with interactive fixed effects in which the variable of interest, say  $Y_{it}^*$ , contains measurement error:

$$Y_{it}^* = \alpha_0 Y_{it-1}^* + \lambda_i^0 f_t^0 + \epsilon_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T, \quad (1)$$

$$Y_{it} = Y_{it}^* + \eta_{it}. \quad (2)$$

Here  $Y_{it}$  is the observed variable and  $\eta_{it}$  represents measurement error. The term  $\lambda_i^0 f_t^0$  describes unobserved interactive fixed effects.<sup>1,2</sup> The goal of the paper is to estimate  $\alpha_0$

---

\*Moon thanks the NSF for financial support.

<sup>1</sup>In this paper, we consider a single factor, that is, the dimensions of  $f_t$  and  $\lambda_i$  equals one. The extension to the multiple factor case is straightforward, but omitted due to the space limitation.

<sup>2</sup>When interpreting  $\lambda_i^0$  as an individual specific fixed effects the term  $f_t^0$  represents the (time varying) linear projection coefficients of  $Y_{it}^*$  on  $\lambda_i^0$  (holding  $Y_{it-1}^*$  constant). This allows the effect of the unobserved

when both the number of individuals  $N$  and the number of time periods  $T$  are large.<sup>3</sup>

The dynamics of the observed variable  $Y_{it}$  can be written as

$$Y_{it} = \alpha_0 Y_{it-1} + \lambda_i^0 f_t^0 + U_{it}, \quad (3)$$

where  $U_{it} = \epsilon_{it} + \eta_{it} - \alpha_0 \eta_{it-1}$ . There are two noticeable features in equations (1) and (3) compared to the widely studied dynamic panel regression model. First, the individual effects take an interactive form instead of the time invariant form. Secondly, the variable of interest  $Y_{it}^*$  is not observed but measured with error. To our knowledge, combining these two features in dynamic linear panel regression has not been studied in the large  $N, T$  panel literature.

We expect two hurdles in estimating  $\alpha_0$ . One is the presence of the interactive fixed effects  $\lambda_i^0 f_t^0$  which might cause a so-called incidental parameter problem in both the cross section and the time dimension. The second one is that the composite error  $U_{it}$  in the observed variable equation (3) is correlated with the lagged dependent variable  $Y_{it-1}$  and we may therefore need to use instrumental variables (IVs).

The main contribution of the paper is to find a valid estimation method that overcomes these two problems. The proposed estimator is a nested two step estimator based on least squares minimization in the first step and minimizing the distance of some of the first step parameter estimates in the second step<sup>4</sup>. Following Moon, Shum and Weidner (2011) (hereafter MSW), we call this method the LS-MD estimation method. This approach was used in estimating endogenous quantile regression models by Chernozhukov and Hansen (2006, 2008) and in estimating the random coefficient logit demand model by MSW.

## 2 LS-MD Estimation

The properties of the quasi-maximum likelihood estimator (QMLE), which minimizes the sum of squared residuals, for large  $N, T$  linear panel regressions with interactive fixed effects were discussed in Bai (2009), and Moon and Weidner (2010). However,

---

individual characteristic  $\lambda_i^0$  on  $Y_{it}^*$  to be time varying. Alternatively, one can interpret  $f_t^0$  as a common time specific shock (a common factor) and  $\lambda_i^0$  then describes reaction to the common shock (a factor loading).

<sup>3</sup>We consider large  $N, T$  approximations to characterize the bias due to the incidental parameter problems, see e.g. Bai (2009) and Hahn and Kuersteiner (2004).

<sup>4</sup>An alternative approach would be to use the the common correlated effect methods suggested by Harding and Lamarche (2011). Both approaches have their own merits and weaknesses. Comparing these different methods is not the interest of this paper.

this estimation method cannot be used to estimate model (3) since the regressor  $Y_{it-1}$  is endogenous w.r.t. the error  $U_{it}$  through the lagged measurement error  $\eta_{it-1}$ . In this case, we may use instrumental variables. Since  $U_{it}$  has an  $MA(1)$  type serial dependence structure, we have  $E(U_{it}Y_{it-1-s}) = 0$  for all  $s \geq 1$ . This suggests to choose  $Z_{it} = (Z_{1,it}, \dots, Z_{L,it})' = (Y_{it-2}, \dots, Y_{it-1-L})'$  for the IVs of the endogenous regressor  $Y_{it-1}$ ,  $t = 1, \dots, T$ . The question, then, is how to use the instrumental variables  $Z_{it}$  to estimate  $\alpha_0$  in the presences of interactive fixed effects  $\lambda_i^0 f_t^0$  when both  $N$  and  $T$  are large.

The estimation method we consider in this paper is a two-step least-squares minimum distance (LS-MD) estimation. This was recently proposed by MSW for estimating the BLP demand model. A similar multi-step estimation idea was also used in Chernozhukov and Hansen (2006, 2008) in estimating endogenous quantile regressions with IVs.

The LS-MD estimation consists of the following two steps: Step 1: For given  $\alpha$ , we solve the least squares problem augmented by the instrumental variables  $Z_{it}$ , that is, we run the OLS regression of  $Y_{it} - \alpha Y_{it-1}$  on  $Z_{it}$  with interactive fixed effects  $\lambda_i f_t$  and solve

$$\left( \hat{\gamma}(\alpha), \hat{\lambda}(\alpha), \hat{f}(\alpha) \right) = \arg \min_{(\gamma, \lambda, f)} \sum_{i=1}^N \sum_{t=1}^T (Y_{it} - \alpha Y_{it-1} - \gamma' Z_{it} - \lambda_i f_t)^2,$$

where  $\gamma = (\gamma_1, \dots, \gamma_L)'$ ,  $\lambda = (\lambda_1, \dots, \lambda_N)'$  and  $f = (f_1, \dots, f_T)'$ . Step 2: For some positive definite weight matrix  $W_{NT}^\gamma$ , we estimate  $\alpha$  by minimizing the length of  $\hat{\gamma}(\alpha)$  as

$$\hat{\alpha} = \arg \min_{\alpha} \hat{\gamma}(\alpha)' W_{NT}^\gamma \hat{\gamma}(\alpha).$$

The idea of the LS-MD method is that since  $Z_{it}$  is excluded in the regression equation (3) the coefficient of  $Z_{it}$  should be zero when  $\alpha = \alpha_0$ . When there is no interactive fixed effect one can show that the LS-MD estimator is equivalent to the conventional 2SLS estimator for an appropriate weight matrix  $W_{NT}^\gamma$ .

### 3 Asymptotic Results

**Assumption 3.1** (i) The unobserved error terms  $\{\epsilon_{it}\} \sim iid(0, \sigma_\epsilon^2)$  and  $\{\eta_{it}\} \sim iid(0, \sigma_\eta^2)$  across  $i$  and over  $t$  and  $E|\epsilon_{it}|^\kappa, E|\eta_{it}|^\kappa < \infty$  for some  $\kappa > 8$ . Also,  $\{\epsilon_{it}\}$  and  $\{\eta_{it}\}$  are independent. (ii) Assume that  $f_t^0$  are strictly stationary and ergodic with  $\sup_t |f_t| < \infty$  and  $\frac{1}{T} \sum_{t=1}^T (f_t^0)^2 \rightarrow_p \Sigma_f > 0$ , and  $\lambda_i$  are iid with  $\sup_i |\lambda_i| < \infty$  and  $\frac{1}{N} \sum_{i=1}^N (\lambda_i^0)^2 \rightarrow_p \Sigma_\lambda > 0$ .

Also assume that  $\{f_t^0\}, \{\lambda_i^0\}, \{\epsilon_{it}\}, \{\eta_{it}\}$  are independent. (iii)  $W_{NT}^\gamma \rightarrow_p W^\gamma > 0$ . (vi)  $|\alpha_0| < 1$  and  $\alpha_0 \neq 0$ .

The iid assumptions of  $\epsilon_{it}$  and  $\eta_{it}$  are made for simplicity of the analysis. Later, an extension to a non-iid case will be discussed. Assumption 3.1(i) also assumes that the measurement error  $\eta_{it}$  is classical in the sense that  $\eta_{it}$  has zero mean and is uncorrelated with  $Y_{it}^*$ . Later we discuss how to extend our method to some special cases of non-classical measurement error. Assumption 3.1(ii) assumes that the factors are strong, which is standard in the factor analysis literature. Assumption 3.1(vi) assumes that  $\alpha_0 \neq 0$ , otherwise the IVs become irrelevant.

Before we present the next assumption, we introduce some further notation. We use  $[a_{it}]$  to denote an  $N \times T$  matrix with elements  $a_{it}$ . For a full column rank matrix  $A$ , let  $\mathbb{P}_A = A(A'A)^{-1}A'$  and  $\mathbb{M}_A = I - \mathbb{P}_A$ . We use notation  $Y = [Y_{it}]$ ,  $Y_{-k} = [Y_{it-k}]$ ,  $Z = [Z_{it}]$ ,  $U = [U_{it}]$ ,  $\epsilon = [\epsilon_{it}]$ ,  $\eta = [\eta_{it}]$ , and  $\eta_{-1} = [\eta_{it-1}]$ . Define  $\lambda^0 = (\lambda_1^0, \dots, \lambda_N^0)'$  and  $f^0 = (f_1^0, \dots, f_T^0)'$ . We also define the  $NT$ -vectors  $y_{-1} = \text{vec}(Y_{-1})$  and  $z = \text{vec}(Z)$ .

**Assumption 3.2** Assume that there exists a positive constant  $c > 0$  such that  $\frac{1}{NT}y'_{-1}\mathbb{P}_z y_{-1} - \max_\lambda \frac{1}{NT}y'_{-1}\mathbb{P}_{I_T \otimes \tilde{\lambda}} y_{-1} > c$  with probability approaching one as  $N, T \rightarrow \infty$ , where  $\tilde{\lambda} = (\lambda^0, \lambda)$ .

Assumption 3.2 is a relevance condition on the instruments. It demands that the explanatory power of the instruments  $Z_{it}$  for the endogenous regressor  $Y_{it-1}$ , given by  $\frac{1}{NT}y'_{-1}\mathbb{P}_z y_{-1}$ , is larger than the joint explanatory power for  $Y_{it-1}$  of the true factor loading  $\lambda^0$  together with any other factor loading  $\lambda$ , given by  $\frac{1}{NT}y'_{-1}\mathbb{P}_{I_T \otimes \tilde{\lambda}} y_{-1}$ . If there are no interactive fixed effects included in the model, then the assumption simplifies to the standard relevance condition  $\frac{1}{NT}y'_{-1}\mathbb{P}_z y_{-1} > 0$ , which is satisfied for  $\alpha \neq 0$ .

Suppose that Assumption 3.1 holds, and consider the special case where  $f_t^0$  has mean zero and is distributed independently over  $t$ . Then, Assumption 3.2 is equivalent to<sup>5</sup>

$$\alpha_0^2 > \frac{1 + \frac{\sigma_\epsilon^2}{\Sigma_\lambda \Sigma_f} + \frac{\sigma_\eta^2}{\Sigma_\lambda \Sigma_f}}{\left(1 + \frac{\sigma_\epsilon^2}{\Sigma_\lambda \Sigma_f}\right)^2 + \frac{\sigma_\eta^2}{\Sigma_\lambda \Sigma_f}}. \quad (4)$$

Thus, by imposing an appropriate lower bound on  $|\alpha_0|$  one can guarantee that the lagged values of  $Y_{it}$  are sufficiently relevant instruments. The conclusion that an appropriate lower

<sup>5</sup>For the proof of this, we refer to the supplementary appendix which is available at <http://ihome.cuhk.edu.hk/~b121322/Research.html>.

bound on  $|\alpha_0|$  is sufficient for the relevance assumption Assumption 3.2 can be extended to cases where  $f_t^0$  is correlated across  $t$ , but in general it is not possible to give such a convenient analytic expression as in (4) for the lower bound.<sup>6</sup> Note that the lower bound in (4) goes to zero when  $\Sigma_\lambda \Sigma_f$  becomes small relative to  $\sigma_\varepsilon^2$ , i.e. the bound is not restrictive when the relative influence of the factors on  $Y_{it}$  is small.

**Theorem 3.1** *Under Assumption 3.1 and 3.2 we have  $\hat{\alpha} \rightarrow_p \alpha_0$  as  $N, T \rightarrow \infty$ .*<sup>78</sup>

To present the limiting distribution of  $\hat{\alpha}$ , we need to introduce some further notation. Define the  $NT$ -vectors  $y_{-1}^{\lambda f}$  and  $z_l^{\lambda f}$  by  $y_{-1}^{\lambda f} = \text{vec}(\mathbb{M}_{\lambda_0} Y_{-1} \mathbb{M}_{f_0})$ , and  $z_l^{\lambda f} = \text{vec}(\mathbb{M}_{\lambda_0} Z_l \mathbb{M}_{f_0})$ , where  $l = 1, \dots, L$ . Let  $u = \text{vec}(U)$  and  $z^{\lambda f} = (z_1^{\lambda f}, \dots, z_L^{\lambda f})'$ .

Define  $G = \text{plim}_{N, T \rightarrow \infty} \frac{1}{NT} y_{-1}^{\lambda f'} z^{\lambda f} = \frac{\sigma_\varepsilon^2}{1 - \alpha_0^2} (\alpha_0, \alpha_0^2, \dots, \alpha_0^L)'$ , and

$$\begin{aligned} W &= \text{plim}_{N, T \rightarrow \infty} \left( \frac{1}{NT} z^{\lambda f'} z^{\lambda f} \right)^{-1} W_{NT}^\gamma \left( \frac{1}{NT} z^{\lambda f'} z^{\lambda f} \right)^{-1} \\ &= \left\{ \frac{\sigma_\varepsilon^2}{1 - \alpha_0^2} \begin{bmatrix} 1 & \dots & \alpha_0^{L-1} \\ \vdots & \ddots & \vdots \\ \alpha_0^{L-1} & \dots & 1 \end{bmatrix} + \sigma_\eta^2 I_L \right\} W^\gamma \left\{ \frac{\sigma_\varepsilon^2}{1 - \alpha_0^2} \begin{bmatrix} 1 & \dots & \alpha_0^{L-1} \\ \vdots & \ddots & \vdots \\ \alpha_0^{L-1} & \dots & 1 \end{bmatrix} + \sigma_\eta^2 I_L \right\}. \end{aligned}$$

Notice that under Assumption 3.1, the limits  $G$  and  $W$  are well defined. Also, notice that under Assumption 3.1, we have  $GWG' > 0$ .

Define

$$c^{(2)} = \left[ C^{(2)}(Z_l, u) \right]_{l=1, \dots, L},$$

---

<sup>6</sup>A non-zero mean of  $f_t^0$  can result in situations where Assumption 3.2 is not satisfied for any value of  $\alpha_0$ . The assumption that  $f_t^0$  is mean zero would not be restrictive if we would include a conventional individual specific fixed effect in the model, in addition to the interactive fixed effect — or equivalently (from an asymptotic perspective), one can demean  $Y_{it}$  separately for each  $i$  before estimating the model with only interactive effects.

<sup>7</sup>The proof is omitted due to space limitation. It is a special case of MSW where their  $\delta(\alpha) = Y - \alpha Y_{-1}$  and the conditions in Assumptions 3.1 and 3.2 are sufficient for the consistency conditions in MSW (see the supplementary appendix available at <http://www.cemmap.ac.uk/publications.php>.)

<sup>8</sup>Note that Assumption 3.2 is a sufficient condition for the relevance of the instruments, but nothing is known about the necessity of this assumption. The LS-MD estimator may also give consistent parameter estimates in some situations where the assumption is violated.

where

$$\begin{aligned}
& C^{(2)}(Z_l, u) \\
&= \text{tr} \left( u \mathbb{M}_{f^0} u' \mathbb{M}_{\lambda^0} Z_l f^0 (f^{0'} f^0)^{-1} (\lambda^{0'} \lambda^0)^{-1} \lambda^{0'} \right) + \text{tr} \left( u' \mathbb{M}_{\lambda^0} u \mathbb{M}_{f^0} Z_l' \lambda^0 (\lambda^{0'} \lambda^0)^{-1} (f^{0'} f^0)^{-1} f^{0'} \right) \\
&\quad + \text{tr} \left( u' \mathbb{M}_{\lambda^0} Z_l \mathbb{M}_{f^0} u' \lambda^0 (\lambda^{0'} \lambda^0)^{-1} (f^{0'} f^0)^{-1} f^{0'} \right).
\end{aligned}$$

MSW show that under Assumption 3.1, as  $N, T \rightarrow \infty$  with  $\frac{N}{T} \rightarrow \kappa^2$ , where  $0 < \kappa < \infty$ , we can approximate

$$\sqrt{NT}(\hat{\alpha} - \alpha_0) = (GWG')^{-1} GW \left[ \frac{1}{\sqrt{NT}} (z^{\lambda f})' u + c^{(2)} \right] + o_p(1). \quad (5)$$

Notice that as  $N, T \rightarrow \infty$  with  $\frac{N}{T} \rightarrow \kappa^2$ , where  $0 < \kappa < \infty$ , under Assumptions 3.1 we can show that

$$\frac{1}{\sqrt{NT}} (z^{\lambda f})' u \Rightarrow N(-\kappa b, \Omega), \quad \text{and} \quad c^{(2)} \rightarrow_p 0, \quad (6)$$

where

$$\begin{aligned}
b &= (b_1, \dots, b_L)', \quad b_l = \text{plim}_{N, T \rightarrow \infty} \frac{1}{N} \text{tr} \left[ P_{f^0} \left[ E(\epsilon' \tilde{\epsilon}_{-l-1}) + E\left((\eta - \alpha_0 \eta_{-1})' \eta_{-l-1}\right) \right] \right], \\
\tilde{\epsilon}_{-l} &= [\tilde{\epsilon}_{it-l}], \quad \tilde{\epsilon}_{it-l} = \sum_{s=0}^{\infty} \alpha_0^s \epsilon_{it-l-s}, \\
\Omega &= \left( \frac{\sigma_\epsilon^2}{1 - \alpha_0^2} + \sigma_\eta^2 \right) \left\{ \frac{\sigma_\epsilon^2}{1 - \alpha_0^2} \begin{bmatrix} 1 & \dots & \alpha_0^{L-1} \\ \vdots & \ddots & \vdots \\ \alpha_0^{L-1} & \dots & 1 \end{bmatrix} + \sigma_\eta^2 I_L \right\}.
\end{aligned}$$

Combining (5) and (6), we have the following theorem.

**Theorem 3.2** *Suppose that Assumptions 3.1 hold. As  $N, T \rightarrow \infty$  with  $\frac{N}{T} \rightarrow \kappa^2$  and  $0 < \kappa < \infty$ , we have*

$$\sqrt{NT}(\hat{\alpha} - \alpha_0) \Rightarrow \mathcal{N} \left( -\kappa (GWG')^{-1} GWb, (GWG')^{-1} GW\Omega WG' (GWG')^{-1} \right).$$

Notice that the bias  $b$  in the limit distribution is due to the incidental parameters  $\lambda_i^0 f_t^0$  and the lagged dependent variables as IVs, which is similar to the bias in Moon and Weidner (2010). This bias can be consistently estimated and is correctable, for details we



refer to Moon, Shum, and Weidner (2011).

## 4 Discussion

Our LS-MD estimation can be used for more sophisticated cases. We briefly discuss how to extend our simple model.

1. **Inclusion of covariates:** The LS-MD estimation procedure can be easily extended to include a model with other exogenous regressors, say  $X_{it}$ . For example, in the first step one can regress  $Y_{it} - \alpha Y_{it-1}$  on  $X_{it}$ ,  $Z_{it}$  with interactive fixed effects  $\lambda_i f_t$  for fixed  $\alpha$ . In the second step, minimize  $\hat{\gamma}(\alpha)' W_{NT}^\gamma \hat{\gamma}(\alpha)$  w.r.t.  $\alpha$ .
2. **Heteroskedastic error:** Until now, we assume that the errors  $\epsilon_{it}$  and  $\eta_{it}$  are homoskedastic for simplicity. If the errors are heteroskedastic, then the limit of  $c^{(2)}$  is non-zero and contributes additional bias terms to the limit distribution of  $\hat{\alpha}$ . These biases are correctable (see e.g. Bai, 2009, and Moon and Weidner, 2010), and can be interpreted as mis-specification bias, since the least squares step in the LS-MD procedure can only correspond to a correctly specified likelihood maximization under homoscedasticity.
3. **Non-classical measurement error:** Measurement error so far is assumed to be classical. In many applications, however, measurement error can be correlated with the unobserved latent variable and the covariates. Our estimation method is still valid under more general measurement error models. For example, suppose that people tend to report income,  $Y_{it}$  proportionally to  $Y_{it}^*$  as

$$Y_{it} = \gamma_{0i} + \gamma_{1it} Y_{it}^* + v_{it}, \quad (7)$$

where  $v_{it}$  is an unobserved error. Note that the measurement error in model (7) is non-classical since the measurement error,  $\eta_{it} = Y_{it} - Y_{it}^*$ , could be correlated with  $Y_{it}^*$  and the mean of the measurement error is not necessarily zero.<sup>9</sup> Model (7) is a modified version of a linear measurement error model that allows for a heterogeneous relationship between  $Y_{it}$  and  $Y_{it}^*$  across cross-section and over time.<sup>10</sup> When the

---

<sup>9</sup>A special case of model (7) is  $\gamma_{0it} = 0$  and  $\gamma_{1it} = 1$ , in which case  $v_{it}$  is classical.

<sup>10</sup>Bollinger and Chandra (2005) and Kim and Solon (2005) develop a model allowing for a constant linear relationship between  $Y_{it}$  and  $Y_{it}^*$ , based on the evidence in surveyed income; i.e., those who earn higher than

coefficient  $\gamma_{1it}$  is random satisfying  $\gamma_{1it} = \gamma_1 + w_{it}$ , where  $\{w_{it}\}$  and  $\{v_{it}\}$  are *iid* across  $i$  and over  $t$  with zero mean, and  $\{w_{it}\}, \{v_{it}\}, \{\epsilon_{it}\}$  are independent of each other, then we have the following dynamic equation with two factors (or one factor and a time invariant fixed effect) as

$$Y_{it} = \alpha Y_{it-1} + \delta_i' h_t + U_{it},$$

where  $\delta_i = [\gamma_1 \lambda_i, (1 - \alpha) \gamma_{0i}]'$ ,  $h_t = [f_t, 1]'$  and

$$U_{it} = \gamma_1 \epsilon_{it} + v_{it} - \alpha v_{it-1} + Y_{it}^* w_{it} - \alpha Y_{it-1}^* w_{it-1}. \quad (9)$$

Note that the composite error  $U_{it}$  in (9) has serial dependence structure similar to an  $MA(1)$  process, and we can use  $Z_{it} = (Z_{1,it}, \dots, Z_{L,it})' = (Y_{it-2}, \dots, Y_{it-1-L})'$  as IVs for the LS-MD estimator.

## References

- [1] Bai, J. (2009): Panel Data Models with Interactive Fixed Effects, *Econometrica*, 77, 1229-1279.
- [2] Bollinger, C and A. Chandra (2005): Introgenic specification error: a cautionary tale of cleaning data, *Journal of Labor Economics* 23(2), 235-257.
- [3] Bound, J, C. C. Brown, and N. Mathiowetz (2001): Measurement Error in Survey Data. *Handbook of Labor Economics*, 5, 3705-3843.
- [4] Chernozhukov, V. and C. Hansen (2006): Instrumental Quantile Regression Inference for Structural and Treatment Effect Models, *Journal of Econometrics*, 132, 491-525.
- [5] Chernozhukov, V. and C. Hansen (2008): Instrumental Variable Quantile Regression: A Robust Inference Approach, *Journal of Econometrics*, 142, 379-398.
- [6] Hahn, J. and G. Kuersteiner (2002): Asymptotically Unbiased Inference of a Dynamic Panel Model with Fixed Effects When Both  $n$  and  $T$  Are Large, *Econometrica*, 70, 1639-1657.

---

average tend to report their earning less, while those who earn lower than average tend to report higher. See also Bound, Brown and Mathiowetz (2001).

- [7] Harding, M. and C. Lamarche (2011): Least Squares Estimation of a Panel Data Model with Multifactor Error Structure and Endogenous Covariates, *Economic Letters*, 111, 197–199.
- [8] Kim, B and G. Solon (2005): Implications of mean-reverting measurement error for longitudinal studies of wages and employment. *Review of Economics and Statistics* 87 (1), 193-196.
- [9] Moon, H.R. and M. Shum, M. Weidner (2011): Interactive Fixed Effects in the BLP Random Coefficients Demand Model, *Working Paper*, Univ. of Southern California.
- [10] Moon, H.R. and M. Weidner (2010): Dynamic Panel Linear Regression Models with Interactive Fixed Effects, *Working Paper*, Univ. of Southern California.

## 5 Supplementary Appendix (Not for Publication)

### 5.1 Proof of Consistency

We show that Assumptions 3.1 and 3.2 in the current model are sufficient for Assumption 1 of Moon, Shum and Weidner (2011) with  $\delta(\alpha)$  in MSW replaced by  $Y - \alpha Y_{-1}$ , and  $X_k$  in MSW replaced by 0.

- Assumption 1(i) holds since uniformly in  $\alpha$  outside of any neighborhood of  $\alpha_0$  we have

$$\frac{\|\delta(\alpha) - \delta(\alpha_0)\|_F}{\|\alpha - \alpha_0\|} = \|Y_{-1}\|_F = \sqrt{\sum_{i=1}^N \sum_{t=1}^T Y_{it-1}^2} = O_p(\sqrt{NT}).$$

Also, it follows that

$$\|Z_l\|_F = \sqrt{\sum_{i=1}^N \sum_{t=1}^T Y_{it-1-l}^2} = O_p(\sqrt{NT}).$$

- Assumption 1(ii) is satisfied because  $\|U\| = \|\epsilon + \eta - \alpha_0 \eta_{-1}\| \leq \|\epsilon\| + \|\eta\| + |\alpha_0| \|\eta_{-1}\| = O_p(\sqrt{\max(J, T)})$  because  $\{\epsilon_{it}\}, \{\eta_{it}\} \sim iid$  with mean zero and finite moments higher than 4 (See Moon and Weidner (2010)).
- Assumption 1(iii)

$$\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T Y_{it-1-l} U_{it} = o_p(1)$$

follows for  $l \geq 1$  since  $E(Y_{it-1-l} U_{it}) = 0$  if  $l \geq 1$ .

- Assumption 1(iv) follows since any (nontrivial) linear combinations of  $Z'_l$ s have rank higher than two under Assumption 3.1.
- Assumption 1(v) holds by Assumption 3.2 with  $\Delta \xi_{\alpha, \beta} = -(\alpha - \alpha_0) y_{-1}$ .
- Assumption 1(vi) holds by Assumption 3.1 (iii).

### 5.2 Asymptotic Normality

- Assumptions 2 and 3 in MSW follow immediately under Assumption 3.1.

- Assumptions 4 and 5 in MSW follow since in this paper  $\delta(\alpha) = Y - \alpha Y_{-1}$  is linear in  $\alpha$  and by the conditions in Assumption 3.1.

### 5.3 Sufficient Conditions for Assumption 3.2

In matrix notation we can write (3) as

$$Y = \alpha^0 Y_{-1} + \lambda^0 f^{0'} + U,$$

where  $Y_{-1} = [Y_{i,t-1}]_{i=1,\dots,N;t=1,\dots,T}$ , and we define  $Y_{-2}$  analogously. By recursively applying the model we find

$$Y = \lambda^0 F^{0'} + E + Y^{\text{init}},$$

where  $F$  is the  $T \times 1$  vector with entries  $F_t = \sum_{\tau=0}^{t-1} \alpha_0^\tau f_{t-\tau}^0$ , and  $E$  and  $Y^{\text{init}}$  are the  $T \times N$  matrices with entries  $E_{it} = \eta_{it} + \sum_{\tau=0}^{t-1} \alpha_0^\tau \epsilon_{t-\tau}$ , and  $Y_{it}^{\text{init}} = \alpha_0^t Y_{i0}$ . We denote lagged versions of  $F^0$  and  $E$  by  $F_{-1}^0$  and  $E_{-1}$ , etc.

In the following we assume  $L = 1$ . In that case Assumption 3.2 is satisfied if

$$\frac{(\text{plim}_{N,T \rightarrow \infty} \frac{1}{NT} y'_{-1} z)^2}{\text{plim}_{N,T \rightarrow \infty} \frac{1}{NT} z' z} - \text{plim}_{N,T \rightarrow \infty} \left( \max_{\lambda} \frac{1}{NT} y'_{-1} \mathbb{P}_{I_T \otimes \tilde{\lambda}} y_{-1} \right) > 0, \quad (10)$$

where  $\tilde{\lambda} = (\lambda^0, \lambda)$ . If  $f_t$  is mean zero and independent across  $t$  we have

$$\begin{aligned} \frac{1}{NT} y'_{-1} z &= \frac{1}{NT} \text{Tr}(Y_{-1} Y_{-2}) \\ &= \frac{1}{NT} \text{Tr}(E_{-1} E_{-2}) + \frac{1}{NT} \|\lambda\|^2 (F'_{-1} F_{-2}) + o_p(1) \\ &= \frac{\alpha_0}{1 - \alpha_0^2} (\sigma_\epsilon^2 + \Sigma_\lambda \Sigma_f) + o_p(1), \\ \frac{1}{NT} z' z &= \frac{1}{NT} \text{Tr}(Y_{-1} Y_{-1}) \\ &= \frac{1}{1 - \alpha_0^2} (\sigma_\epsilon^2 + \Sigma_\lambda \Sigma_f) + \sigma_\eta^2 + o_p(1), \end{aligned}$$

and

$$\begin{aligned}
\max_{\lambda} \frac{1}{NT} y'_{-1} \mathbb{P}_{I_T \otimes \tilde{\lambda}} y_{-1} &= \max_{\lambda} \frac{1}{NT} \text{Tr}(Y'_{-1} P_{\tilde{\lambda}} Y_{-1}) \\
&= \max_{\lambda} \frac{1}{NT} \text{Tr}(F_{-1} \lambda' P_{\tilde{\lambda}} \lambda F'_{-1}) + o_p(1) \\
&= \frac{1}{NT} \|\lambda\|^2 \|F_{-1}\|^2 + o_p(1) \\
&= \frac{1}{1 - \alpha_0^2} \Sigma_{\lambda} \Sigma_f + o_p(1).
\end{aligned}$$

Plugging these results into condition (10) immediately yields condition (4). Here, we have assumed only one instrument ( $L = 1$ ), but in the special case under consideration ( $f_t$  independent across  $t$ ) there is actually no additional explanatory power for  $Y_{it}$  contained in  $Y_{i,t-2}$ , i.e the probability limit of  $\frac{1}{NT} y'_{-1} \mathbb{P}_z y_{-1}$  is the same for all  $L \geq 1$ .