

Cointegration Vector Estimation by Panel DOLS and Long-run Money Demand*

NELSON C. MARK^{†,‡,§} and DONGGYU SUL[¶]

[†]*Ohio State University (e-mail: mark.1@osu.edu)*

[‡]*University of Notre Dame*

[§]*NBER*

[¶]*University of Auckland (e-mail: d.sul@auckland.ac.nz)*

Abstract

We study the panel dynamic ordinary least square (DOLS) estimator of a homogeneous cointegration vector for a balanced panel of N individuals observed over T time periods. Allowable heterogeneity across individuals include individual-specific time trends, individual-specific fixed effects and time-specific effects. The estimator is fully parametric, computationally convenient, and more precise than the single equation estimator. For fixed N as $T \rightarrow \infty$, the estimator converges to a function of Brownian motions and the Wald statistic for testing a set of s linear constraints has a limiting $\chi^2(s)$ distribution. The estimator also has a Gaussian sequential limit distribution that is obtained first by letting $T \rightarrow \infty$ and then letting $N \rightarrow \infty$. In a series of Monte-Carlo experiments, we find that the asymptotic distribution theory provides a reasonably close approximation to the exact finite sample distribution. We use panel DOLS to estimate coefficients of the long-run money demand function from a panel of 19 countries with annual observations that span from 1957 to 1996. The estimated income elasticity is 1.08 (asymptotic s.e. = 0.26) and the estimated interest rate semi-elasticity is -0.02 (asymptotic s.e. = 0.01).

*This paper was previously circulated under the title 'A Computationally Simple Cointegration Vector Estimator for Panel Data'. For valuable comments on earlier drafts, we thank Ronald Bewley, Roger Moon, Peter Phillips, seminar participants at Georgetown University, Ohio State University, the 2001 New Zealand Econometric study group meeting, the University of California at Santa Barbara, the University of Southern California, and an anonymous referee. The usual disclaimer applies.

I. Introduction

This paper considers the extension of the single equation dynamic ordinary least squares (DOLS) method of Saikkonen (1991) and Stock and Watson (1993) for estimating and testing hypotheses about a cointegrating vector to panel data. We call the estimator panel DOLS. We discuss its limit distribution and apply it to estimate the long-run money demand function using a panel data set of 19 countries with annual observations spanning from 1957 to 1996.

Panel DOLS is fully parametric and offers a computationally convenient alternative to the panel ‘fully modified’ OLS estimator proposed by Pedroni (1997) and Phillips and Moon (1999). Properties of panel DOLS, when there are fixed effects in the cointegrating regression, have been discussed by Kao and Chiang (2000). We take this to be the starting point for our analysis. In our environment, the cointegrating vector is homogeneous across individuals but we allow for individual heterogeneity through disparate short-run dynamics, individual-specific fixed effects and individual-specific time trends. Moreover, we permit a limited degree of cross-sectional dependence (CSD) through the presence of time-specific effects.

We present two limit distributions for panel DOLS. The first limit distribution is obtained for a fixed number of cross-sectional units N , letting $T \rightarrow \infty$. In this case, panel DOLS converges in distribution to a function of Brownian motions and the Wald statistic for testing a set of s linear constraints has a limiting $\chi^2(s)$ distribution. This limit theory seems well suited for many applied macroeconomic or international problems. Here, researchers often have available panel data sets of moderate N but much larger T . With the passage of time, these data sets will gain time-series observations but they are unlikely to acquire many more cross-sectional units.¹ We also obtain the sequential limit distribution by first letting $T \rightarrow \infty$ for fixed N , and then letting $N \rightarrow \infty$ as proposed by Phillips and Moon (1999). Here, panel DOLS has a limiting Gaussian distribution and as in the fixed N case, the Wald statistic has a limiting chi-square distribution. In the absence of linear trends in the cointegrating regression, the sequential limiting normality of the estimator is theoretically interesting but has less practical import because the limit distribution of the test statistics is identical to the $T \rightarrow \infty$ distribution with fixed N . However, when linear trends are present, the sequential limit theory produces considerable simplifications. Here, the estimator of the cointegration

¹For example, if the observational unit is a national economy, the total number of countries may fluctuate over time, but is unlikely to go to infinity. While the break-up of the Soviet Union created several new economies, the opposite trend is at work in Europe where the EMU may eventually combine to form a single economic unit. But beyond this, researchers typically choose to group countries into classes that share common characteristics such as income levels, stages of development or geography which often result in panels with 5–20 individuals.

vector and the time-trend slope coefficients remain correlated for fixed N as $T \rightarrow \infty$ but are asymptotically uncorrelated when $T \rightarrow \infty$ then $N \rightarrow \infty$.

As single equation cointegration vector estimators are super consistent, it is natural to ask what is to be gained by using the panel estimator. The answer is that super consistency means that convergence towards the asymptotic distribution occurs at rate T but it says nothing about the sampling variability of the estimator for a fixed value of T . In fact, the statistical properties of single-equation cointegration-vector estimators can be quite poor when applied to sample sizes associated with macroeconomic time series typically available to researchers (e.g. Inder, 1993; Stock and Watson, 1993). Moreover, even limited amounts of heterogeneity in the short-run dynamics across individuals can generate considerable disparities in single-equation DOLS estimates of the true homogeneous cointegration vector. In these situations, combining cross-sectional and time-series information in the form of a panel can provide much more precise point estimates of the cointegration vector with reasonably accurate asymptotic approximations to the exact sampling distribution. In a series of Monte-Carlo experiments, we study the small sample performance of panel DOLS and compare it with single-equation DOLS. Panel DOLS generally performs well under the short-run dynamic designs that we consider and can attain a striking improvement in estimation precision over that of single-equation DOLS with even a modest number of cross-sectional units.

We then apply panel DOLS to estimate the long-run demand for M1 money. The countries in our study are Austria, Australia, Belgium, Canada, Denmark, France, Finland, Germany, Iceland, Ireland, Japan, Norway, New Zealand, the Netherlands, Portugal, Spain, Switzerland, the United Kingdom, and the United States. Here, we build on the time-series contributions by Baba, Hendry and Starr (1992), Ball (1998), Hoffman, Rasche, and Tieslau (1995), Lucas (1988) and Stock and Watson (1993), and the cross-sectional studies by Mulligan and Sala-i-Martin (1992), and Mulligan (1997), most of which has focused on US data.²

The studies cited above report conflicting results along three dimensions. First, point estimates from time-series studies exhibit substantial dependence on the sample period of the data. Income elasticity estimates from post-WWII US data typically lie well below 1 – which implies the existence of economies of scale in money management – whereas estimates obtained from pre-WWII observations or those that combine pre- and postwar observations are generally close to 1. Using annual US data spanning from 1903 to 1987, Stock and Watson's (1993) DOLS estimate of the income elasticity is 0.97. When the sample spans from 1903 to 1945, their estimate is 0.89 but drops to 0.27 when

²Less recent cross-sectional studies include Meltzer (1963) and Gandolfi and Lothian (1976).

the data span from 1946 to 1987. Ball (2001) extends these data and obtains an estimate of 0.42 when the observations span from 1946 to 1996. Using annual observations from 1900 to 1958, Lucas's (1988) estimate of the M1 (permanent) income elasticity is 1.06 and his estimate of the (short-term) interest rate semi-elasticity is -0.07 . Using data spanning from 1958 to 1985, his income elasticity estimate drops to 0.21 and his interest semi-elasticity estimate is -0.01 . Secondly, there is tension generated by the large difference between the estimates from time-series studies and those from postwar cross-section studies. Mulligan and Sala-i-Martin's (1992) estimates from a 1989 cross-sectional data set from the Survey of Consumer Finances range between 0.82 and 1.37. Mulligan (1997) runs OLS on for a panel of 12,000 firms observed from 1961 to 1992 and obtains an income-elasticity estimate of 0.83. Thirdly, there is substantial cross-country variation even amongst economies of similar income levels and financial market development. In our analysis, single-equation DOLS with trend gives such disparate income elasticity estimates as -1.23 for New Zealand and 2.42 for Canada. The corresponding interest rate semi-elasticity estimates range from 0.02 for Ireland (which has the wrong sign) to -0.09 for the UK. When trends are omitted, the income elasticity estimates range from 0.13 for Belgium to 2.64 for Norway and the interest semi-elasticity estimates range from 0.02 for Ireland to -0.16 for Norway.

With only 40 annual observations, the heterogeneity that we observe in the point estimates may plausibly have been generated from an underlying data generating process with a homogeneous cointegration vector and heterogeneous short-run dynamics. When we include heterogeneous linear trends and estimate the cointegrating vector by panel DOLS, we obtain a point estimate of the income elasticity of 1.08 (asymptotic s.e. = 0.26) and a point estimate of the interest semi-elasticity of -0.02 (asymptotic s.e. = 0.01). Moreover, these estimates, which are more in line with those from cross-sectional studies on US data, are stable as the span of the time-series dimension is varied and are reasonably robust to the inclusion or omission of heterogeneous linear trends.

The remainder of the paper is organized as follows. The next section describes the representation of the non-stationary panel data and regularity conditions assumed in the paper. Section III describes the panel DOLS estimator and discusses its asymptotic properties. Section IV reports the results of a Monte-Carlo experiment to examine the small sample performance of the panel DOLS estimator and the accuracy of the asymptotic approximations. In section V we present our long-run money demand study and section VI concludes the paper. Proofs of propositions and supplementary results from the money demand study are given in an appendix which is available upon request from the authors.³

³Alternatively, the appendix can be downloaded from the <http://www.wcon.ohio-state.edu/Mark/nmark.htm>.

II. Representation of cointegrated observations in panel data

Consider a balanced panel of individuals indexed by $i = 1, \dots, N$ observed over time periods $t = 1, \dots, T$. Vectors are underlined and matrices appear in bold face. $\underline{W}(r)$ is a vector standard Brownian motion for $0 \leq r \leq 1$, and $[Tr]$ denotes the largest integer value of Tr for $0 \leq r \leq 1$. As is standard in the literature, we will drop the notational dependence on r and write $\int_0^1 \underline{W}(r) dr$ as $\int \underline{W}$ and $\int_0^1 \underline{W}(r) d\underline{W}(r)'$ as $\int \underline{W} d\underline{W}'$. Scaled vector Brownian motions are denoted by $\underline{B} = \underline{\Lambda} \underline{W}$ where $\underline{\Lambda}$ is a scaling matrix. For any matrix \mathbf{A} , $\|\mathbf{A}\|$ denotes the Euclidian norm, $[\text{Tr}(\mathbf{A}'\mathbf{A})]^{1/2}$.

We will be working with double indexed sums. In some instances – to deal with individual-specific fixed effects or common-time effects – these sums will involve transformations of the original observations. To handle such situations, we generically denote the sample cross-moment matrix averaged over N and T^2 as \mathbf{M}_{NT} and let the precise definition depend upon the particular model under consideration. Also, we generically denote the limit of the moment matrix as $T \rightarrow \infty$ for any given N by \mathbf{M}_N . As $N \rightarrow \infty$, \mathbf{M}_N need not converge to a constant and we denote this limit as $\overline{\mathbf{M}}_N$. Similarly, our generic notation for the sample cross-product vector between the regressors and the equilibrium error is \underline{m}_{NT} , and the limit for fixed N as $T \rightarrow \infty$ is \underline{m}_N .

As the model we study allows for individual specific effects, perhaps it would be more accurate to call the estimator dynamic least squares dummy variable (LSDV). However, in the interests of simplicity, we will refer to the estimator as panel DOLS.

(i) Triangular representation

Let $\{(y_{it}, \underline{x}'_{it})'\}$ be a $(k + 1)$ dimensional vector of observations where y_{it} is a scalar and \underline{x}_{it} is a k -dimensional vector. Observations on each individual i obey the triangular representation

$$y_{it} = \alpha_i + \lambda_i t + \theta_t + \underline{\gamma}' \underline{x}_{it} + u_{it}^\dagger, \tag{1}$$

$$\Delta \underline{x}_{it} = \underline{v}_{it}, \tag{2}$$

where $(1, -\underline{\gamma}')$ is a cointegrating vector between y_{it} and \underline{x}_{it} that is identical across individuals. The composite equilibrium error $y_{it} - \underline{\gamma}' \underline{x}_{it}$ is potentially comprised of an individual-specific effect α_i , an individual-specific linear trend $\lambda_i t$, and a common time-specific factor θ_t . The remaining idiosyncratic error u_{it}^\dagger is independent across i but possibly dependent across t . An alternative representation for equation (2) allows \underline{x}_{it} to have an individual-specific vector of drift terms and for the trend in equation (1) to be induced by this drift. With

some minor modifications, the results of this paper continue to hold in this alternative environment.⁴

In addition to individual-specific fixed effects and linear trends, potentially disparate short-run dynamics of the covariance stationary error process $\{w_{it}^\dagger\} = \{(u_{it}^\dagger, v_{it}^\dagger)'\}$ introduces an additional source of heterogeneous behaviour across individuals. The underlying error dynamics are given in Assumption 1.

Assumption 1. (Error Dynamics.) $\{w_{it}^\dagger\}$ is independent across $i = 1, \dots, N$, and has the moving average representation

$$w_{it}^\dagger = \Psi_i^\dagger(L)\epsilon_{it}^\dagger, \tag{3}$$

where $\{\epsilon_{it}^\dagger\} \sim i.i.d.$ with $E(\epsilon_{it}^\dagger) = \mathbf{0}$, $E(\epsilon_{it}^\dagger \epsilon_{it}^{\dagger 0}) = \mathbf{I}$, $E\|\epsilon_{it}^\dagger\|^4 < \infty$, $\Psi_i^\dagger(L) = \sum_{j=0}^\infty \Psi_{ij}^\dagger L^j$ is a $(k + 1) \times (k + 1)$ dimensional matrix lag polynomial in the lag operator L , where $\sum_{j=0}^\infty j |\psi_{ij}^{\dagger mn}| < \infty$, and $\psi_{ij}^{\dagger mn}$ is the m, n -th element of the matrix Ψ_{ij}^\dagger .

Our assumption that w_{it}^\dagger is independent across individuals ($E[w_{it}^\dagger w_{jt-k}^\dagger] = 0$, $i \neq j$, $-\infty \leq k \leq \infty$) follows the recent econometrics literature on non-stationary panel data (e.g. Phillips and Moon, 1999a,b; Kao and Chiang, 2000; Pedroni, 1997). Unlike these authors, we assume that the coefficients in the $\Psi_i^\dagger(L)$ polynomial are fixed for a given i although they can differ across individuals.

Let W_j be a $k + 1$ -dimensional standard Brownian motion. By Assumption 1, it follows that for each $i = 1, \dots, N$, $\{w_{it}^\dagger\}$ obeys the functional central limit theorem

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[T]} w_{it}^\dagger \xrightarrow{D} \Psi_i^\dagger(1) W_i \equiv B_i^\dagger, \tag{4}$$

as $T \rightarrow \infty$, where $B_i^\dagger = (B_{ui}^\dagger, B_{vi}^\dagger)'$ is a scaled mixed Brownian motion and

$$\Omega_i^\dagger = E[B_i^\dagger(1)B_i^\dagger(1)'] = \begin{bmatrix} \Omega_{uu,i}^\dagger & \Omega_{vu,i}^{\dagger 0} \\ \Omega_{vu,i}^\dagger & \Omega_{vv,i}^\dagger \end{bmatrix} = \Psi_i^\dagger(1)\Psi_i^{\dagger 0}(1) = \Gamma_{0,i}^\dagger + \sum_{j=1}^\infty (\Gamma_{j,i}^\dagger + \Gamma_{j,i}^{\dagger 0}),$$

$$\Gamma_{j,i}^\dagger = E(w_{it}^\dagger w_{it-j}^{\dagger 0}) = E \begin{bmatrix} u_{it}^\dagger u_{it-j}^\dagger & u_{it}^\dagger v_{it-j}^\dagger \\ v_{it}^\dagger u_{it}^\dagger & v_{it}^\dagger v_{it-j}^\dagger \end{bmatrix} = \begin{bmatrix} \Gamma_{uu,j,i}^\dagger & \Gamma_{vu,j,i}^{\dagger 0} \\ \Gamma_{vu,j,i}^\dagger & \Gamma_{vv,j,i}^\dagger \end{bmatrix}.$$

⁴The trend in equation (1) can be induced by a drift in $\{y_{it}\}$. If instead, there is a drift in $\{x_{it}\}$ but none in $\{y_{it}\}$, where $\Delta x_{it} = a_i + v_{it}$ with x_{i0} given, then repeated substitution gives $x_{it} = x_{i0} + a_i t + \xi_{it}$ where $\xi_{it} = \sum_{j=1}^t v_{ij}$ is a driftless vector $I(1)$ process. The cointegrating regression becomes $y_{it} = \gamma_0' x_{i0} + \gamma_1' a_i t + \gamma_2' \xi_{it} + u_{it}^\dagger$. In this case, the ensuing analysis is to be performed using the statistical properties of ξ_{it} . In computations, one can follow the recommendations of Phillips and Moon (2000) to obtain an estimate of ξ_{it} first by estimating the drift $a_{iT} = \frac{1}{T} \sum_{t=1}^T \Delta x_{it} = \frac{1}{T} (x_{iT} - x_{i1})$ and then using $\hat{\xi}_{it} = x_{it} - x_{i0} - ta_{iT}$.

The issues involved in panel cointegration vector estimation and testing parallels that in the single equation environment. For a single equation, OLS is a consistent estimator of the cointegrating vector but its asymptotic distribution depends on the long-run covariance between u_{it}^\dagger and \underline{v}_{it} . This nuisance parameter dependency invalidates standard hypothesis testing in the OLS framework without modifications. DOLS, dynamic generalized least square (GLS), and fully modified OLS are examples of such modifications. Similarly, in panel data, Phillips and Moon (1999) and Pedroni (1997) show that for fixed N , the pooled OLS estimator is a consistent estimator of the cointegrating vector as $T \rightarrow \infty$ and can be used in a first pass in getting point estimates. In panel data, however, the problems of second-order asymptotic bias and nuisance parameter dependence are compounded and are potentially more serious in the sense that the bias accumulates with the size of the cross-section. In particular, if $\underline{\gamma}_{NT}^{\text{OLS}}$ is the OLS estimator for the pooled cross-section time-series data, one cannot rule out the possibility that $\sqrt{NT}(\underline{\gamma}_{NT}^{\text{OLS}} - \underline{\gamma})$ diverges as $T \rightarrow \infty$ and then $N \rightarrow \infty$. It follows that the distribution for a Wald statistic for testing linear restrictions becomes even less useful as the cross-sectional dimension of the panel grows as it too can diverge.

III. Panel DOLS

We consider the panel DOLS estimator of the vector of slope coefficients $\underline{\gamma}$ and its limit distribution for various subcases of the model in equations (1) and (2). We take the model with individual-specific effects as our starting point.⁵ Section (i) discusses the baseline fixed-effects model. Once we obtain the limit distribution for this baseline case, the limit theory for more general versions of the model with heterogeneous linear trends and common time effects follow in an analogous manner. In section (ii) we add heterogeneous trends to the fixed-effects model, and section (iii) considers the model with fixed-effects, trends, and common-time effects.

(i) Fixed effects

In applied work, the researcher will almost always need to include individual-specific constants in the regression. To handle this situation, we

⁵Kao and Chiang (2000) derive the sequential limit distribution ($T \rightarrow \infty$, $N \rightarrow \infty$) for panel DOLS in a model with individual-specific effects. They do not consider the fixed the $T \rightarrow \infty$ limit theory with fixed N , nor do they allow for time trends or time-specific effects.

begin by setting $\lambda_i = 0, \theta_t = 0$ for all i and t in equation (1), which we write as

$$y_{it} = \alpha_i + \underline{\gamma}'x_{it} + u_{it}^\dagger. \tag{5}$$

Assume that u_{it}^\dagger is correlated with at most p_i leads and lags of $\underline{v}_{it} = \Delta x_{it}$. To control for this endogeneity, project u_{it}^\dagger onto these p_i leads and lags

$$u_{it}^\dagger = \sum_{s=-p_i}^{p_i} \delta'_{i,s} v_{it-s} + u_{it} = \sum_{s=-p_i}^{p_i} \delta'_{i,s} \Delta x_{it-s} + u_{it} = \underline{\delta}'_i z_{it} + u_{it}, \tag{6}$$

where $\underline{\delta}_{i,s}$ is a $k \times 1$ vector of projection coefficients, $\underline{\delta}_i = (\delta'_{i,-p_i}, \dots, \delta'_{i,0}, \dots, \delta'_{i,p_i})'$ is a $(2p_i + 1)k$ -dimensional vector and $\underline{z}_{it} = (\Delta x'_{it-p_i}, \dots, \Delta x'_{it}, \dots, \Delta x'_{it+p_i})'$ is a $(2p_i + 1)k$ -dimensional vector of leads and lags of the first differences of the variables x_{it} . The projection error u_{it} is, by construction, orthogonal to all leads and lags of \underline{v}_{it} . It follows from Assumption 1 that because $\underline{w}_{it}^\dagger$ is independent across i we project u_{it}^\dagger only onto leads and lags of Δx_{it} for individual i and not onto leads and lags of the other individuals $(\Delta x_{jt}, j \neq i)$.

Substituting the projection representation for u_{it} in equation (6) into equation (5) yields

$$y_{it} = \alpha_i + \underline{\gamma}'x_{it} + \underline{\delta}'_i z_{it} + u_{it}. \tag{7}$$

The projection defines the new covariance stationary vector process, $\underline{w}_{it} = (u_{it}, \underline{v}'_{it})'$ where for each i ,

$$\underline{w}_{it} = \Psi_i(L)\epsilon_{it}, \quad \Psi_i(L) = \begin{bmatrix} \Psi_{uu,i}(L) & \underline{0}' \\ \underline{0} & \Psi_{vv,i}(L) \end{bmatrix},$$

and \underline{w}_{it} obeys the functional central limit theorem

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \underline{w}_{it} \xrightarrow{D} \underline{B}_i = \Psi_i(1)\underline{W}_i,$$

where $\underline{B}_i = (B_{ui}, B'_{vi})'$, B_{ui} and B_{vi} are independent, and

$$\Omega_i = E[\underline{B}_i(1)\underline{B}_i(1)'] = \begin{bmatrix} \Psi_{uu,i}(1)^2 & \underline{0}' \\ \underline{0} & \Psi_{vv,i}(1)\Psi_{vv,i}(1)' \end{bmatrix} = \begin{bmatrix} \Omega_{uu,i} & \underline{0}' \\ \underline{0} & \Omega_{vv,i} \end{bmatrix}.$$

Taking the time-series average of equation (7) gives

$$\frac{1}{T} \sum_{t=1}^T y_{it} = \alpha_i + \underline{\gamma}' \frac{1}{T} \sum_{t=1}^T x_{it} + \delta'_i \frac{1}{T} \sum_{t=1}^T z_{it} + \frac{1}{T} \sum_{t=1}^T u_{it}. \tag{8}$$

Subtracting equation (8) from equation (7) eliminates α_i and gives

$$\tilde{y}_{it} = \underline{\gamma}' \tilde{\underline{x}}_{it} + \underline{\delta}'_i \tilde{\underline{z}}_{it} + \tilde{u}_{it}, \tag{9}$$

where a ‘tilde’ denotes the deviation of an observation from its time-series average,

$$\tilde{y}_{it} = y_{it} - \frac{1}{T} \sum_{t=1}^T y_{it},$$

$$\tilde{\underline{x}}_{it} = \underline{x}_{it} - \frac{1}{T} \sum_{t=1}^T \underline{x}_{it},$$

$$\tilde{\underline{z}}_{it} = \underline{z}_{it} - \frac{1}{T} \sum_{t=1}^T \underline{z}_{it},$$

$$\tilde{u}_{it} = u_{it} - \frac{1}{T} \sum_{t=1}^T u_{it}.$$

To set up the estimation problem, let $\tilde{\underline{q}}_{it}$ be the $2k(1 + \sum_{i=1}^N p_i)$ dimensional vector whose first k elements are $\tilde{\underline{x}}_{it}$, elements $k(1 + \sum_{j=1}^{i-1} (2p_j + 1)) + 1$ to $k(1 + \sum_{j=1}^i (2p_j + 1))$ are $\tilde{\underline{z}}_{it}$ and 0s elsewhere. That is,

$$\begin{aligned} \tilde{\underline{q}}_{1t} &= (\tilde{\underline{x}}'_{1t} \quad \tilde{\underline{z}}'_{1t} \quad \underline{0}' \quad \dots \quad \underline{0}')' \\ \tilde{\underline{q}}_{2t} &= (\tilde{\underline{x}}'_{2t} \quad \underline{0}' \quad \tilde{\underline{z}}'_{2t} \quad \dots \quad \underline{0}')' \\ &\vdots \\ \tilde{\underline{q}}_{Nt} &= (\tilde{\underline{x}}'_{Nt} \quad \underline{0}' \quad \underline{0}' \quad \dots \quad \tilde{\underline{z}}'_{Nt})' \end{aligned}$$

Let the grand coefficient vector be $\underline{\beta} = (\underline{\gamma}', \underline{\delta}'_1, \dots, \underline{\delta}'_N)'$ and the compact form of the regression be $\tilde{y}_{it} = \underline{\beta}' \tilde{\underline{q}}_{it} + \tilde{u}_{it}$. The panel DOLS estimator for the fixed-effects model is $\underline{\beta}_{NT}$, where

$$(\underline{\beta}_{NT} - \underline{\beta}) = \left[\sum_{i=1}^N \sum_{t=1}^T \tilde{\underline{q}}_{it} \tilde{\underline{q}}'_{it} \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T \tilde{\underline{q}}_{it} \tilde{u}_{it} \right]. \tag{10}$$

We exploit the fact that the limiting behaviour of the regression error \tilde{u}_{it} , is identical to that of u_{it} . Some algebra reveals that

$$\frac{1}{T} \sum_{t=1}^T \tilde{\underline{x}}_{it} \tilde{u}_{it} = \frac{1}{T} \sum_{t=1}^T \tilde{\underline{x}}_{it} u_{it} \xrightarrow{D} \sqrt{\Omega_{uu,i}} \int \tilde{\underline{B}}_{vi} dW_{ui} \quad \text{where } \tilde{\underline{B}}_{vi} = \underline{B}_{vi} - \int \underline{B}_{vi}.$$

We also have

$$\frac{1}{\sqrt{T}} \sum_{i=1}^T \tilde{u}_{it} = \frac{1}{\sqrt{T}} \sum_{i=1}^T u_{it} - (1/T) \frac{1}{\sqrt{T}} \sum_{i=1}^T u_{it} \xrightarrow{D} B_{ui}(1) = \sqrt{\Omega_{uu,i}} W_{ui}(1),$$

so we are able to use estimated values of \tilde{u}_{it} to obtain a consistent estimate of $\Omega_{uu,i}$. The $T \rightarrow \infty$ limit theory with fixed N for panel DOLS with individual-specific fixed effects is given in

Proposition 1. (Fixed N , $T \rightarrow \infty$ with fixed effects.) Let $\tilde{B}_{vi} = B_{vi} - \int B_{vi}$. For the panel DOLS estimator (10), for fixed N as $T \rightarrow \infty$,

- a. $T(\gamma_{NT} - \gamma)$ and $\sqrt{T}(\hat{\delta}_i - \delta_i)$ are independent for each i .
- b. $\sqrt{NT}(\gamma_{NT} - \gamma) \xrightarrow{D} \mathbf{M}_N^{-1} \underline{m}_N$, where $\mathbf{M}_N = \frac{1}{N} \sum_{i=1}^N \int \tilde{B}_{vi} \tilde{B}'_{vi}$, and $\underline{m}_N = \frac{1}{N} \sum_{i=1}^N \sqrt{\Omega_{uu,i}} \left[\int \tilde{B}_{vi} dW_{ui} \right]$.
- c. $[\sqrt{NT} \mathbf{R}(\gamma_{NT} - \gamma)]' [\mathbf{R} \mathbf{D}_N \mathbf{R}']^{-1} [\sqrt{NT} \mathbf{R}(\gamma_{NT} - \gamma)] \xrightarrow{D} \chi^2(s)$, where \mathbf{R} is an $s \times k$ restriction matrix, $\mathbf{D}_N = \mathbf{M}_N^{-1} \mathbf{V}_N \mathbf{M}_N^{-1}$, $\hat{\mathbf{M}}_N = \frac{1}{N} \sum_{i=1}^N \int \tilde{B}_{vi} \tilde{B}'_{vi}$, and $\mathbf{V}_N = \frac{1}{N} \sum_{i=1}^N \Omega_{uu,i} \int \tilde{B}_{vi} \tilde{B}'_{vi}$.
- d. $\hat{\mathbf{D}}_{NT} - \mathbf{D}_N \xrightarrow{p} \mathbf{0}$, where $\hat{\mathbf{D}}_{NT} = \mathbf{M}_{NT}^{-1} \hat{\mathbf{V}}_{NT} \mathbf{M}_{NT}^{-1}$, $\mathbf{M}_{NT} = \left[\frac{1}{N} \sum_{i=1}^N \left(\frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \tilde{x}'_{it} \right) \right]$, $\hat{\mathbf{V}}_{NT} = \frac{1}{N} \sum_{i=1}^N \hat{\Omega}_{uu,i} \left(\frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \tilde{x}'_{it} \right)$, and $\hat{\Omega}_{uu,i}$ is a consistent estimator of $\Omega_{uu,i}$.

The asymptotic independence of $T(\gamma_{NT} - \gamma)$ and $\sqrt{T}(\hat{\delta}_i - \delta_i)$ as $T \rightarrow \infty$ follows for the same reasons as in the single-equation environment and because $T(\gamma_{NT} - \gamma)$ converges in distribution to a mixed normal random vector, the limiting chi-square distribution of the quadratic form in part (c) of the proposition also follows by the standard argument. The asymptotic covariance matrix \mathbf{D}_N can be consistently estimated by $\hat{\mathbf{D}}_{NT}$ and it follows that under the null hypotheses $\mathbf{R}\gamma = \underline{r}$, the Wald statistic

$$[\sqrt{NT}(\mathbf{R}\gamma_{NT} - \underline{r})]' [\mathbf{R} \hat{\mathbf{D}}_{NT} \mathbf{R}']^{-1} [\sqrt{NT}(\mathbf{R}\gamma_{NT} - \underline{r})] \xrightarrow{D} \chi^2(s) \tag{11}$$

as $T \rightarrow \infty$ for any given N .

The sequential limit distribution of γ_{NT} is obtained by showing that the sequence $\left\{ \int \tilde{B}_{vi} \tilde{B}'_{vi} \right\}_{i=1}^{\infty}$ obeys a law of large numbers for independent but heterogeneously distributed observations and that the sequence $\left\{ \sqrt{\Omega_{uu,i}} \int \tilde{B}_{vi} dW_{ui} \right\}_{i=1}^{\infty}$ obeys a central limit theorem for independent but heterogeneously distributed observations. The sequential limit theory for panel DOLS is given in Proposition 2.

Proposition 2. (Sequential limit distribution, fixed effects.) For the panel DOLS estimator (10), as $T \rightarrow \infty$ and then $N \rightarrow \infty$,

- a. $\mathbf{C}_N^{-1/2} \sqrt{NT}(\underline{y}_{NT} - \underline{\gamma}) \overset{A}{\sim} N(0, \mathbf{I}_k)$, where $\mathbf{C}_N = (\mathbf{C}_N^{-1/2})(\mathbf{C}_N^{-1/2})' = \overline{\mathbf{M}}_N^{-1} \overline{\mathbf{V}}_N \overline{\mathbf{M}}_N^{-1}$, $\overline{\mathbf{M}}_N = \frac{1}{6N} \sum_{i=1}^N \boldsymbol{\Omega}_{vv,i}$, and $\overline{\mathbf{V}}_N = \frac{1}{6N} \sum_{i=1}^N \boldsymbol{\Omega}_{uu,i} \boldsymbol{\Omega}_{vv,i}$.
- b. $\hat{\mathbf{D}}_{NT} - \mathbf{C}_N \xrightarrow{P} \mathbf{0}$, where $\hat{\mathbf{D}}_{NT}$ is defined in Proposition 1(d).

Controlling for fixed effects results in a shrinkage of the sequential limit asymptotic variance, compared to when there are no fixed effects. In the case without fixed effects where $\alpha_i = 0$ for all i , $\overline{\mathbf{M}}_N = \frac{1}{2N} \sum_{i=1}^N \boldsymbol{\Omega}_{vv,i}$ and $\overline{\mathbf{V}}_N = \frac{1}{2N} \sum_{i=1}^N \boldsymbol{\Omega}_{uu,i} \boldsymbol{\Omega}_{vv,i}$.

(ii) Fixed effects and heterogeneous trends

We now admit both individual-specific fixed effects and heterogeneous time trends into the specification. Upon substitution of the projection representation for the equilibrium error (6) into (1) (with $\theta_t = 0$ for all t) we have,

$$y_{it} = \alpha_i + \lambda_i t + \underline{\gamma}' \underline{x}_{it} + \underline{\delta}'_i \underline{z}_{it} + u_{it}. \tag{12}$$

The time-series average of equation (12) yields

$$\frac{1}{T} \sum_{t=1}^T y_{it} = \alpha_i + \lambda_i \left(\frac{T+1}{2} \right) + \underline{\gamma}' \frac{1}{T} \sum_{t=1}^T \underline{x}_{it} + \underline{\delta}'_i \frac{1}{T} \sum_{t=1}^T \underline{z}_{it} + \frac{1}{T} \sum_{t=1}^T u_{it}, \tag{13}$$

where we use the fact that $\frac{1}{T} \sum_{t=1}^T t = (T+1)/2$. To control for the fixed-effects subtract equation (13) from equation (12) to get

$$\tilde{y}_{it} = \lambda_i \tilde{t} + \underline{\gamma}' \tilde{\underline{x}}_{it} + \underline{\delta}'_i \tilde{\underline{z}}_{it} + \tilde{u}_{it}, \tag{14}$$

where again we use a ‘tilde’ to denote the deviation of an observation from its time-series average,

$$\begin{aligned} \tilde{y}_{it} &= y_{it} - \frac{1}{T} \sum_{t=1}^T y_{it}, \\ \tilde{\underline{x}}_{it} &= \underline{x}_{it} - \frac{1}{T} \sum_{t=1}^T \underline{x}_{it}, \\ \tilde{\underline{z}}_{it} &= \underline{z}_{it} - \frac{1}{T} \sum_{t=1}^T \underline{z}_{it}, \\ \tilde{u}_{it} &= u_{it} - \frac{1}{T} \sum_{t=1}^T u_{it}, \\ \tilde{t} &= t - \frac{T+1}{2}. \end{aligned}$$

To set up panel DOLS, let $\underline{\lambda}_N = (\lambda_1, \lambda_2, \dots, \lambda_N)'$ be the vector of trend-slope coefficients, $\beta = (\underline{\gamma}', \underline{\lambda}'_N, \underline{\delta}'_1, \dots, \underline{\delta}'_N)'$ be the grand coefficient vector, and define

$$\begin{aligned} \tilde{\mathbf{q}}'_{1t} &= (\tilde{\mathbf{x}}'_{1t} \quad \tilde{t} \quad 0 \quad \dots \quad 0 \quad \underline{z}'_{1t} \quad \underline{0}' \quad \dots \quad \underline{0}')' \\ \tilde{\mathbf{q}}'_{2t} &= (\tilde{\mathbf{x}}'_{2t} \quad 0 \quad \tilde{t} \quad \dots \quad 0 \quad \underline{0}' \quad \underline{z}'_{2t} \quad \dots \quad \underline{0}')' \\ &\vdots \\ \tilde{\mathbf{q}}'_{Nt} &= (\tilde{\mathbf{x}}'_{Nt} \quad 0 \quad 0 \quad \dots \quad \tilde{t} \quad \underline{0}' \quad \underline{0}' \quad \dots \quad \underline{z}'_{Nt})' \end{aligned} \tag{15}$$

Then the panel DOLS estimator of $\underline{\beta}$ is:

$$\underline{\beta}_{NT} = \left[\sum_{i=1}^N \sum_{t=1}^T \tilde{\mathbf{q}}_{it} \tilde{\mathbf{q}}'_{it} \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T \tilde{\mathbf{q}}_{it} \tilde{y}_{it} \right]. \tag{16}$$

For fixed N , as $T \rightarrow \infty$, $\sqrt{T}(\hat{\underline{\delta}}_i - \underline{\delta}_i)$ is independent of $T(\underline{\gamma}_{NT} - \underline{\gamma})$ and $T^{3/2}(\hat{\underline{\lambda}}_N - \underline{\lambda}_N)$ for the standard reasons but $T(\underline{\gamma}_{NT} - \underline{\gamma})$ and $T^{3/2}(\hat{\underline{\lambda}}_N - \underline{\lambda}_N)$ remain correlated. The $T \rightarrow \infty$ limit theory with fixed N for the fixed-effects model with trend is given in Proposition 3.

Proposition 3. (Fixed N , $T \rightarrow \infty$, fixed effects and trends.) Let $\tilde{\underline{\mathbf{B}}}_{vi} = \underline{\mathbf{B}}_{vi} - \int \underline{\mathbf{B}}_{vi}$. For the panel DOLS estimator (16), for fixed N as $T \rightarrow \infty$,

a. $\sqrt{T}(\hat{\underline{\delta}}_i - \underline{\delta}_i)$ is independent of $T(\underline{\gamma}_{NT} - \underline{\gamma})$ and $T^{3/2}(\hat{\underline{\lambda}}_N - \underline{\lambda}_N)$ for each i .

b. $\begin{bmatrix} T(\underline{\gamma}_{NT} - \underline{\gamma}) \\ T^{3/2}(\hat{\underline{\lambda}}_N - \underline{\lambda}_N) \end{bmatrix} \xrightarrow{D} \mathbf{M}_N^{-1} \underline{m}_N$, where $\mathbf{M}_N = \begin{bmatrix} \mathbf{M}_{11,N} & \mathbf{M}'_{21,N} \\ \mathbf{M}_{21,N} & \mathbf{M}_{22,N} \end{bmatrix}$,

$$\begin{aligned} \mathbf{M}_{11,N} &= \frac{1}{N} \sum_{i=1}^N \int \tilde{\underline{\mathbf{B}}}_{vi} \tilde{\underline{\mathbf{B}}}'_{vi}, \quad \mathbf{M}_{22,N} = \frac{1}{12} \mathbf{I}_N, \\ \mathbf{M}'_{21,N} &= \left[\frac{1}{\sqrt{N}} \left(\int r \tilde{\underline{\mathbf{B}}}_{v1} - \frac{1}{2} \int \tilde{\underline{\mathbf{B}}}_{v1} \right), \dots, \frac{1}{\sqrt{N}} \left(\int r \tilde{\underline{\mathbf{B}}}_{vN} - \frac{1}{2} \int \tilde{\underline{\mathbf{B}}}_{vN} \right) \right], \text{ and} \\ \underline{m}_N &= \left[\begin{array}{c} \frac{1}{\sqrt{N}} \sum_{i=1}^N \sqrt{\Omega_{uu,i}} \int \tilde{\underline{\mathbf{B}}}_{vi} dW_{ui} \\ \left[\sqrt{\Omega_{uu,1}} \left[\int r dW_{u1} - \frac{W_{u1}(1)}{2} \right], \dots, \sqrt{\Omega_{uu,N}} \left[\int r dW_{uN} - \frac{W_{uN}(1)}{2} \right] \right]' \end{array} \right]. \end{aligned}$$

When $T \rightarrow \infty$ then $N \rightarrow \infty$, the panel DOLS estimator of the trend-slope coefficients and the cointegration vector are independent which results in considerable simplification. The sequential limit theory for panel DOLS in this case is given in Proposition 4.

Proposition 4. (Sequential limits, fixed effects and trends.) For the panel DOLS estimator (16), as $T \rightarrow \infty$ then $N \rightarrow \infty$,

a. $\sqrt{NT}(\underline{\gamma}_{NT} - \underline{\gamma})$ and $T^{3/2}(\hat{\underline{\lambda}}_N - \underline{\lambda}_N)$ are independent.

- b. $\mathbf{C}_N^{-1/2} \sqrt{NT}(\underline{\gamma}_{NT} - \underline{\gamma}) \overset{A}{\sim} N(\mathbf{0}, \mathbf{I}_k)$, where $\mathbf{C}_N = (\mathbf{C}_N^{1/2})(\mathbf{C}_N^{1/2})' = \overline{\mathbf{M}}_{11,N}^{-1} \times \overline{\mathbf{V}}_{11,N} \overline{\mathbf{M}}_{11,N}^{-1}$, $\overline{\mathbf{M}}_{11,N} = \frac{1}{6N} \sum_{i=1}^N \mathbf{\Omega}_{vv,i}$, and $\overline{\mathbf{V}}_{11,N} = \frac{1}{6N} \sum_{i=1}^N \mathbf{\Omega}_{uu,i} \mathbf{\Omega}_{vv,i}$.
- c. $\hat{\mathbf{D}}_{NT} - \mathbf{C}_N \xrightarrow{p} \mathbf{0}$, where $\hat{\mathbf{D}}_{NT} = \mathbf{M}_{NT}^{-1} \hat{\mathbf{V}}_{NT} \mathbf{M}_{NT}^{-1}$, $\mathbf{M}_{NT} = \frac{1}{NT^2} \sum_{i=1}^N \sum_{t=1}^T \tilde{\mathbf{x}}_{it} \tilde{\mathbf{x}}'_{it}$, $\hat{\mathbf{V}}_{NT} = \frac{1}{NT^2} \sum_{i=1}^N \sum_{t=1}^T \hat{\mathbf{\Omega}}_{uu,i} \tilde{\mathbf{x}}_{it} \tilde{\mathbf{x}}'_{it}$, and $\hat{\mathbf{\Omega}}_{uu,i}$ is a consistent estimator of $\mathbf{\Omega}_{uu,i}$

Notice that the sequential limit distribution for $\sqrt{NT}(\underline{\gamma}_{NT} - \underline{\gamma})$ is identical to that obtained in Proposition 2 in the absence of trends. Construction of a Wald test under the sequential limit theory can proceed as in section (i).

(iii) Fixed effects, heterogeneous trends, and common time effects

The asymptotic distribution theory that we employ requires that observations are independent across individuals but in applications, one typically encounters some degree of CSD. In this section we take up the complete model (1) which allows us to model a limited form of CSD in which the equilibrium error for each individual is driven in part by θ_t .

Begin by substituting the projection representation for u_{it}^\ddagger into equation (1) to get

$$y_{it} = \alpha_i + \lambda_{it} + \theta_t + \underline{\gamma}' \underline{x}_{it} + \underline{\delta}'_j \underline{z}_{it} + u_{it}. \tag{17}$$

Controlling for the common time effect requires an analysis of the cross-sectional average of the observations. Because we admit heterogeneity in the projection coefficients $\underline{\delta}_j$ across i , the resulting cross-sectional averages will involve sums such as $\sum_{j=1}^N \underline{\delta}'_j \underline{z}_{jt}$ which complicates estimation of the $\underline{\delta}_j$ coefficients. The estimation problem can be simplified by proceeding sequentially and addressing the endogeneity correction separately from cointegration vector estimation.

To do this, let y_{it}^\ddagger be the error from projecting each element of y_{it} onto $\underline{n}_{it} = (1, t, \underline{z}'_{it})'$ and $\underline{x}_{it}^\ddagger = \underline{x}_{it} - \mathbf{\Phi}_i \underline{n}_{it}$ be the vector of projection errors from projecting each element of \underline{x}_{it} onto \underline{n}_{it} , where $\mathbf{\Phi}_i$ is a $(k + 2) \times p_i$ matrix of projection coefficients. Substituting the projection representations for y_{it} and \underline{x}_{it} into equation (17) gives

$$y_{it}^\ddagger = \underline{\gamma}' \underline{x}_{it}^\ddagger + \theta_t + u_{it}. \tag{18}$$

We now work with equation (18) as for the purposes of estimating and drawing inference about $\underline{\gamma}$ it is equivalent to equation (17). Now take the cross-sectional average of equation (18) to get

$$\frac{1}{N} \sum_{j=1}^N y_{jt}^{\ddagger} = \gamma' \left[\frac{1}{N} \sum_{j=1}^N \underline{x}_{jt}^{\ddagger} \right] + \theta_t + \frac{1}{N} \sum_{j=1}^N u_{jt}. \tag{19}$$

Subtracting equation (19) from equation (18) eliminates the common time effect giving

$$y_{it}^{\ddagger*} = \gamma' \underline{x}_{it}^{\ddagger*} + u_{it}^*, \tag{20}$$

where an ‘asterisk’ denotes the deviation of an observation from its cross-sectional average. That is,

$$\begin{aligned} y_{it}^{\ddagger*} &= y_{it}^{\ddagger} - \frac{1}{N} \sum_{j=1}^N y_{jt}^{\ddagger}, \\ \underline{x}_{it}^{\ddagger*} &= \underline{x}_{it}^{\ddagger} - \frac{1}{N} \sum_{j=1}^N \underline{x}_{jt}^{\ddagger}, \\ u_{it}^* &= u_{it} - \frac{1}{N} \sum_{j=1}^N u_{jt}. \end{aligned}$$

The panel DOLS estimator of γ is

$$\underline{\gamma}_{NT} = \left[\sum_{i=1}^N \sum_{t=1}^T \underline{x}_{it}^{\ddagger*} \underline{x}_{it}^{\ddagger*0} \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T \underline{x}_{it}^{\ddagger*} y_{it}^{\ddagger*} \right]. \tag{21}$$

As in the case of the fixed-effects model with linear trends, the panel DOLS estimator of the grand coefficient vector converges to a mixed normal random vector but $T(\underline{\gamma}_{NT} - \gamma)$ and $T^{3/2}(\hat{\underline{\lambda}}_N - \underline{\lambda}_N)$ are asymptotically correlated for fixed N as $T \rightarrow \infty$. We omit a statement of the limit theory for this case. As $T \rightarrow \infty$ then $N \rightarrow \infty$, however, $T(\underline{\gamma}_{NT} - \gamma)$ and $T^{3/2}(\hat{\underline{\lambda}}_N - \underline{\lambda}_N)$ are independent and the limit theory for this case is given in Proposition 5.

Proposition 5. (Sequential limit distribution.) For the panel DOLS estimator (21), as $T \rightarrow \infty$ then $N \rightarrow \infty$,

- a. $\sqrt{NT}(\underline{\gamma}_{NT} - \gamma)$ and $T^{3/2}(\hat{\underline{\lambda}}_N - \underline{\lambda}_N)$ are independent.
- b. $\mathbf{C}_N^{-1/2} \sqrt{NT}(\underline{\gamma}_{NT} - \gamma) \overset{A}{\rightsquigarrow} N(0, \mathbf{I}_K)$, where $\mathbf{C}_N = (\mathbf{C}_N^{1/2})(\mathbf{C}_N^{1/2})' = \bar{\mathbf{M}}_{11,N}^{-1} \bar{\mathbf{V}}_{11,N} \bar{\mathbf{M}}_{11,N}^{-1}$, $\bar{\mathbf{M}}_{11,N} = \frac{1}{6N} \sum_{i=1}^N \Omega_{vv,i}$, and $\bar{\mathbf{V}}_{11,N} = \frac{1}{6N} \sum_{i=1}^N \Omega_{uu,i} \Omega_{vv,i}$.
- c. $\hat{\mathbf{D}}_{NT} - \mathbf{C}_N \xrightarrow{P} \mathbf{0}$ where $\hat{\mathbf{D}}_{NT} = \mathbf{M}_{11,NT}^{-1} \hat{\mathbf{V}}_{11,NT} \mathbf{M}_{11,NT}^{-1}$, $\mathbf{M}_{11,NT} = \frac{1}{N} \sum_{i=1}^N \left[\frac{1}{T^2} \sum_{t=1}^T \underline{x}_{it}^{\ddagger*} \underline{x}_{it}^{\ddagger*'} \right]$, $\hat{\mathbf{V}}_{11,NT} = \frac{1}{N} \sum_{i=1}^N \sqrt{\hat{\Omega}_{uu,i}} \left[\frac{1}{T^2} \sum_{t=1}^T \underline{x}_{it}^{\ddagger*} \underline{x}_{it}^{\ddagger*'} \right]$, and $\hat{\Omega}_{uu,i}$ is a consistent estimator of $\Omega_{uu,i}$.

Notice that the limit distribution of Proposition 5 is identical to the sequential limit distribution of Proposition 4. Controlling for fixed effects again produces a shrinkage of the asymptotic variance, while controlling for the common time effect requires taking the deviation from the cross-sectional average. These cross-sectional transformations have no effect on the sequential asymptotic variance of the estimator.

If the modifications to OLS are successful in removing the correlation between the equilibrium error u_{it}^\dagger and the leads and lags of Δx_{jt} for $j = 1, \dots, N$ but the time-specific effects do not fully account for CSD, then the residual cross-sectional correlation in the projection error u_{it} changes only the formula for the asymptotic standard errors.⁶ This is a feasible estimation strategy for a small to moderate N . But if there remains a correlation between the equilibrium error and the leads and lags of *other* equation Δx_{jt} , $i \neq j$, then panel DOLS exhibits the same sort of second-order asymptotic bias as pooled OLS as discussed in section II. For small to moderate N , a feasible solution to this problem is to include leads and lags of Δx_{jt} , $j = 1, \dots, N$ in the projection (6).

We close this section by noting that for large N , modelling CSD in panel data is itself an active area of research and one that has shown itself to be a thorny problem. What one seeks in this case is a simple parametric structure that does an adequate job of capturing the long-run covariance structure. Bai and Ng (2001), Moon and Perron (2002), and Phillips and Sul (2002) study models in which the error terms in dynamic panel data regressions have a factor structure, but the implications for such factor models have not been studied in the panel cointegration context.

IV. Monte-Carlo experiments

In this section, we present some Monte-Carlo experiments to investigate some small sample properties of panel DOLS and to compare them with single-equation DOLS in the presence of individual fixed effects and CSD.⁷ Our data generating process (DGP) includes two regressors in the cointegrating relation and is given by

⁶In this case, the asymptotic variance of panel DOLS is consistently estimated by $(\sum_{t=1}^T \mathbf{X}_t \mathbf{X}_t')^{-1} (\sum_{t=1}^T \mathbf{X}_t \boldsymbol{\Omega}_{uu,T} \mathbf{X}_t') (\sum_{t=1}^T \mathbf{X}_t \mathbf{X}_t')^{-1}$ where $\mathbf{X}_t = (x_{1t}, \dots, x_{Nt})$ and $\boldsymbol{\Omega}_{uu,T}$ is a consistent estimator of the long-run covariance matrix of u_{it} , $i = 1, \dots, N$.

⁷Kao and Chiang (2000) compared the small-sample performance of panel DOLS and panel fully modified OLS with fixed effects in the case of a single regressor. They found that panel dynamic OLS performed much better than panel fully modified OLS in removing finite sample bias so we do not include panel fully modified OLS in the comparison.

$$\begin{aligned}
 y_{it} &= \alpha_i + \gamma_1 x_{1,it} + \gamma_2 x_{2,it} + \eta_{it}, \\
 \Delta x_{1,it} &= a_i + v_{1,it}, \\
 \Delta x_{2,it} &= v_{2,it}.
 \end{aligned}$$

Letting $\underline{w}_{it} = (\eta_{it}, v_{1,it}, v_{2,it})'$, $\underline{\epsilon}_{it} = (\epsilon_{1,it}, \epsilon_{2,it}, \epsilon_{3,it})'$, and $\underline{e}_{it} = (e_{1,it}, e_{2,it}, e_{3,it})'$, the short-run dynamics are given by

$$\begin{aligned}
 \underline{w}_{it} &= \mathbf{A}_i \underline{w}_{it-1} + \underline{\epsilon}_{it}, \\
 \underline{\epsilon}_{it} &= \sqrt{\phi} \underline{\theta}_t + \sqrt{1 - \phi} \underline{e}_{it},
 \end{aligned}$$

where for $j = 1, 2, 3$, $i = 1, \dots, N$, $e_{j,it} \stackrel{i.i.d.}{\sim} N(0, \sigma_{ji}^2)$, $\underline{\theta}_t = (\theta_{1t}, \theta_{2t}, \theta_{3t})'$, $\theta_{jt} \stackrel{i.i.d.}{\sim} N(0, \sigma_{\theta j}^2)$.

We designed the DGP to provide a connection to the empirical work on money demand of the next section, where the regressors are real income (which has a drift), and the nominal interest rate (which does not). Accordingly, we induce a trend into the cointegrating relation through the drift term a_i for the first regressor $x_{1,it}$ and specify the second regressor $x_{2,it}$ to be a driftless $I(1)$ process. In addition, we model the equilibrium error $\alpha_i + \eta_{it}$, to admit a more general form of CSD than the common time effect model considered in the previous section.⁸ This single-factor model of the short-run innovations is of the type considered by Phillips and Sul (2002). CSD in the equilibrium errors is induced by θ_{1t} , while θ_{2t} and θ_{3t} induce cross-sectional endogeneity between $x_{j,kt}$ and u_{it}^\dagger , $j \neq i$, $k = 1, 2$. These features were not explicitly accounted for in the theoretical analysis, but may be encountered in empirical work. In the presence of heterogeneous CSD, subtracting off the cross-sectional average does not completely eliminate CSD. Our interest here is in evaluating the seriousness of the resulting distortions. The degree of CSD is modulated by the size of ϕ .

The true value of the cointegration vector is $(\gamma_1, \gamma_2) = (1.0, 0.1)$. For each individual i , the values of a_i , \mathbf{A}_i , and σ_{ji} are first obtained by a draw from the uniform distribution then held fixed throughout the experiment. The persistence in the short-run dynamics are controlled by varying the support of the uniform distribution from which the elements of \mathbf{A}_i are drawn. We consider three levels of persistence and three alternative degrees of CSD. Persistence levels can be low ($A_{11,i} \sim U_{[0.3,0.5]}$), medium ($A_{11,i} \sim U_{[0.5,0.7]}$), or

⁸In the common time effect specification, the cross-sectional correlation between individuals i and j is identical for all i, j . This homogeneous CSD is obtained here by setting $A_{11,i}$ to be identical across i . That allowing for heterogeneity in $A_{11,i}$ results in heterogeneous CSD can be seen in the case of an AR(1) where $A_{11,i} = \rho_i$ and all other elements of \mathbf{A}_i are set to 0. Then it can be shown that $\text{Corr}(\eta_{it}, \eta_{jt}) = c_{ij} = b_{ij}^c \frac{\sqrt{1-\rho_i^2} \sqrt{1-\rho_j^2}}{1-\rho_i \rho_j}$, where $b_{ij} = \frac{E(\epsilon_{1,it} \epsilon_{1,jt})}{\{E(\epsilon_{1,it}^2)E(\epsilon_{1,jt}^2)\}^{\frac{1}{2}}} = \frac{\phi \sigma_{\theta 1}^2}{\sqrt{(\phi \sigma_{\theta 1}^2 + (1-\phi)\sigma_i^2)(\phi \sigma_{\theta 1}^2 + (1-\phi)\sigma_j^2)}}$.

high ($A_{11,i} \sim U_{[0.7,0.9]}$), and degrees of CSD are either none ($\phi = 0$), low ($\phi = 0.3$) or high ($\phi = 0.7$). Assignment of the remaining parameter values are determined by, $A_{21,i} \sim U_{[-0.05,0.05]}$, $A_{12,i} \sim U_{[-0.05,0.05]}$, $A_{23,i} \sim U_{[-0.05,0.05]}$, $A_{22,i} \sim U_{[0.0,0.4]}$, $A_{33,i} \sim U_{[0.0,0.04]}$, $a_i \sim U_{[23,53]} \times 10^{-3}$, $\sigma_{1i}^2 \sim U_{[1,33]} \times 10^{-3}$, $\sigma_{2i}^2 \sim U_{[0.25,1.34]} \times 10^{-3}$, $\sigma_{3i}^2 \sim U_{[2.3,57]}$, and $\sigma_{0j}^2 = (1/N) \sum_{i=1}^N \sigma_{ji}^2$. The long-run variance $\Omega_{uu,i}$ is estimated by the prewhitened quadratic spectral (QSPW) method suggested by Sul, Phillips and Choi (2002). Each experiment consists of 5,000 random samples of $T = 40, 100, \text{ or } 200$ observations on $N = 10$ or 20 individuals. The number of leads and lags of Δx_{it} included are 2 ($T = 40$), 3 ($T = 100$), and 4 ($T = 200$). We organize the experiments according to the following three cases.

Case 1. (No CSD, Variable Persistence). Setting $\phi = 0$ yields no CSD. Persistence levels are low, medium and high.

Case 2. (Homogeneous CSD and Persistence). Setting $A_{11,i} = A_{11,j}$ yields the homogeneous CSD as in the common time effect specification. We consider high and low CSD and low, medium and high levels of persistence.

Case 3. (Heterogeneous CSD and Persistence). Allowing $A_{11,i} \neq A_{11,j}$ yields heterogeneous CSD. We consider high and low CSD and low, medium and high levels of persistence.

We begin with the effective size of nominal 5% and 10% sized tests of the hypothesis $H_0: \gamma_1 = 1$ and $H_0: \gamma_2 = 0.1$. To provide a point of comparison, Table 1 displays the effective size of (single-equation) DOLS tests. Table 2 presents the panel DOLS size results for Case 1. Under low and medium levels of persistence, the tests are reasonably sized. Size accuracy is seen to

TABLE 1
Effective size of DOLS tests

<i>T</i>	<i>Persistence</i>	<i>H</i> ₀ : $\gamma_1 = 1$		<i>H</i> ₀ : $\gamma_2 = 0.1$	
		5%	10%	5%	10%
40	Low	0.097	0.152	0.093	0.148
	Medium	0.118	0.179	0.114	0.174
	High	0.179	0.241	0.175	0.239
100	Low	0.100	0.155	0.093	0.151
	Medium	0.110	0.170	0.104	0.163
	High	0.184	0.250	0.182	0.248
200	Low	0.067	0.128	0.067	0.124
	Medium	0.071	0.132	0.074	0.127
	High	0.117	0.181	0.114	0.180

TABLE 2
Effective size of panel DOLS tests. Case 1: No CSD, variable persistence

<i>T</i>	<i>Persistence</i>	<i>H</i> ₀ : $\gamma_1 = 1$				<i>H</i> ₀ : $\gamma_2 = 0.1$			
		<i>N</i> = 10		<i>N</i> = 20		<i>N</i> = 10		<i>N</i> = 20	
		5%	10%	5%	10%	5%	10%	5%	10%
40	Low	0.092	0.154	0.087	0.142	0.096	0.156	0.082	0.142
	Medium	0.076	0.133	0.065	0.126	0.103	0.169	0.104	0.167
	High	0.056	0.098	0.039	0.074	0.110	0.180	0.155	0.237
100	Low	0.072	0.122	0.059	0.109	0.071	0.126	0.064	0.118
	Medium	0.062	0.111	0.054	0.106	0.080	0.136	0.072	0.127
	High	0.051	0.093	0.042	0.091	0.113	0.184	0.159	0.250
200	Low	0.060	0.105	0.060	0.110	0.058	0.113	0.064	0.115
	Medium	0.059	0.102	0.058	0.108	0.062	0.121	0.067	0.117
	High	0.045	0.097	0.044	0.085	0.090	0.167	0.147	0.228

improve with increasing sample size both in the time series as well as in the cross-sectional dimensions. Under high levels of persistence, the test for γ_1 remains reasonably sized but the test for γ_2 becomes slightly mis-sized. This mis-sizing worsens somewhat as the cross-section increases (e.g. for $T = 100$, the 5% test has size of 16% for $N = 10$ and 25% for $N = 20$). In comparison with DOLS, the test for γ_1 is better sized whereas the test for γ_2 is roughly equivalent.

Table 3 reports the effective size of panel DOLS tests under Case 2. For a low degree of CSD, the size of the test for γ_1 improves with persistence and is accurate when the level of persistence is high.⁹ The length of the time series is relatively unimportant. Similar results are obtained for the test on γ_2 under low CSD. Under high CSD, there is some mis-sizing of the test on γ_1 , which is comparable with the size of the DOLS test. For the test on γ_2 , size accuracy improves with the size of the cross-section and overall size distortion is modest.

Effective size performance of panel DOLS tests under Case 3, shown in Table 4, is very similar to that under Case 2. Subtracting off the cross-sectional average works reasonably well as a control for the heterogeneous CSD considered here.

Table 5 reports quantiles of $\hat{\gamma}_1$ from DOLS and panel DOLS under Case 3. Here, it is seen that dramatic precision gains over single-equation DOLS can be attained in small samples. For $T = 40$, $N = 10$ under high persistence and high CSD, the inter-95 percentile range for DOLS is (-0.304; 2.495) while for

⁹This is largely a feature of Sul, Phillips and Choi's (2002) QSPW estimator of the long-run variance which works well under high persistence.

TABLE 3
Effective size of panel DOLS tests. Case 2: homogeneous CSD

<i>T</i>	<i>CSD</i>	<i>Persistence</i>	<i>H</i> ₀ : $\gamma_1 = 1$				<i>H</i> ₀ : $\gamma_2 = 0.1$			
			<i>N</i> = 10		<i>N</i> = 20		<i>N</i> = 10		<i>N</i> = 20	
			5%	10%	5%	10%	5%	10%	5%	10%
40	Low	Low	0.103	0.160	0.111	0.182	0.106	0.165	0.091	0.157
		Medium	0.087	0.147	0.088	0.150	0.127	0.201	0.083	0.142
		High	0.069	0.113	0.051	0.092	0.165	0.245	0.069	0.117
100	Low	Low	0.067	0.128	0.079	0.136	0.073	0.129	0.065	0.117
		Medium	0.063	0.118	0.073	0.124	0.084	0.143	0.059	0.113
		High	0.050	0.094	0.053	0.098	0.140	0.227	0.067	0.129
200	Low	Low	0.068	0.131	0.063	0.115	0.074	0.128	0.050	0.098
		Medium	0.068	0.122	0.057	0.110	0.075	0.127	0.045	0.090
		High	0.047	0.098	0.046	0.097	0.120	0.204	0.053	0.103
40	High	Low	0.105	0.167	0.161	0.234	0.168	0.240	0.199	0.267
		Medium	0.096	0.152	0.131	0.210	0.167	0.246	0.151	0.217
		High	0.065	0.112	0.082	0.139	0.155	0.230	0.088	0.135
100	High	Low	0.094	0.155	0.133	0.206	0.100	0.161	0.124	0.195
		Medium	0.087	0.147	0.123	0.186	0.113	0.179	0.113	0.178
		High	0.064	0.115	0.107	0.173	0.150	0.241	0.098	0.161
200	High	Low	0.102	0.164	0.130	0.198	0.089	0.145	0.091	0.146
		Medium	0.105	0.163	0.128	0.200	0.090	0.152	0.083	0.144
		High	0.087	0.147	0.124	0.191	0.134	0.216	0.080	0.144

panel DOLS is (0.883; 1.152). Precision gains continue to accrue when $T = 200$. For $N = 10$, under high persistence and high CSD, the panel DOLS inter-95 percentile range of (0.979; 1.028) whereas for DOLS it is (0.89; 1.122). Precision advantages are also seen to accrue from enlarging the cross-sectional dimension. Under high persistence and high CSD, $T = 40$, the inter-95 percentile range shrinks from (0.883; 1.152) for $N = 10$ to (0.924; 1.102) for $N = 20$.

Table 6 displays analogous quantile information for $\hat{\gamma}_2$. Here, the benefits from the cross-section dimension are largely obtained with $N = 10$. For $T = 40$, under high persistence and CSD, the inter-95 percentile range of panel DOLS is (0.096; 0.110), which is an improvement over the (0.054; 0.170) range for DOLS.

We summarize the Monte-Carlo results with four general observations. First, for given T , the empirical size of the panel DOLS t -tests worsens slightly when N is increased from 10 to 20. Secondly, size distortion, while not particularly severe at $T = 40$ is reasonably small at $T = 200$. Thirdly, subtracting the cross-sectional average to control for CSD works reasonably

TABLE 4
Effective size of panel DOLS tests. Case 3: heterogeneous CSD

<i>T</i>	<i>CSD</i>	<i>Persistence</i>	<i>H</i> ₀ : $\gamma_1 = 1$				<i>H</i> ₀ : $\gamma_2 = 0.1$			
			<i>N</i> = 10		<i>N</i> = 20		<i>N</i> = 10		<i>N</i> = 20	
			5%	10%	5%	10%	5%	10%	5%	10%
40	Low	Low	0.104	0.161	0.101	0.166	0.098	0.159	0.091	0.156
		Medium	0.083	0.134	0.082	0.140	0.109	0.172	0.096	0.155
		High	0.059	0.104	0.050	0.085	0.131	0.198	0.124	0.198
100	Low	Low	0.075	0.134	0.082	0.140	0.071	0.128	0.068	0.123
		Medium	0.069	0.121	0.076	0.131	0.085	0.144	0.077	0.132
		High	0.051	0.099	0.051	0.099	0.112	0.187	0.138	0.229
200	Low	Low	0.042	0.092	0.061	0.113	0.040	0.088	0.055	0.105
		Medium	0.040	0.082	0.057	0.112	0.046	0.088	0.056	0.104
		High	0.040	0.076	0.042	0.091	0.066	0.131	0.110	0.190
40	High	Low	0.114	0.172	0.147	0.217	0.152	0.221	0.206	0.283
		Medium	0.094	0.142	0.118	0.190	0.147	0.215	0.167	0.234
		High	0.058	0.098	0.053	0.094	0.142	0.213	0.089	0.147
100	High	Low	0.098	0.165	0.138	0.202	0.098	0.165	0.137	0.207
		Medium	0.088	0.141	0.131	0.199	0.108	0.173	0.131	0.191
		High	0.061	0.110	0.082	0.138	0.146	0.225	0.095	0.159
200	High	Low	0.090	0.150	0.128	0.199	0.068	0.119	0.102	0.164
		Medium	0.088	0.147	0.145	0.207	0.074	0.131	0.102	0.166
		High	0.070	0.124	0.119	0.190	0.108	0.188	0.087	0.145

well even in the presence of heterogenous CSD. Fourthly, panel DOLS is much more precise than single-equation DOLS.

V. Long-run money demand

We now employ panel DOLS to estimate coefficients of the long-run M1 demand function. Economists have long been interested in obtaining precise estimates of money demand for at least two reasons. First, knowing the income elasticity of money demand helps in determining the rate of monetary expansion that is consistent with long-run price level stability. Secondly, knowing the interest elasticity of money demand aids in calculating the area under the demand curve and to assess the welfare costs of long-run inflation (Bailey, 1956). Additionally, because a stable money-demand function is a building block of the IS-LM model, economists have historically been interested in knowing how well this particular aspect of the model performed. While this motive has become less important in the era of dynamic general equilibrium models, Lucas (1988) shows that such a

TABLE 5
 Quantiles for first regressor slope (Case 3): panel DOLS and DOLS

CSD	Persistence	DOLS (N = 10)			Panel DOLS (N = 10)			Panel DOLS (N = 20)		
		2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
<i>T</i> = 40										
None	Low	0.460	0.999	1.512	0.952	1.001	1.048	0.976	1.000	1.025
	Med	0.225	0.999	1.769	0.934	1.001	1.070	0.967	1.002	1.038
	High	-0.192	1.010	2.420	0.890	1.007	1.137	0.937	1.008	1.078
Low	Low	0.456	0.999	1.529	0.951	1.000	1.052	0.973	1.000	1.029
	Med	0.206	0.996	1.782	0.931	1.001	1.075	0.963	1.003	1.044
	High	-0.259	1.006	2.435	0.882	1.010	1.150	0.936	1.011	1.090
High	Low	0.432	0.999	1.536	0.956	1.000	1.047	0.971	1.000	1.031
	Med	0.181	0.996	1.783	0.936	1.001	1.073	0.960	1.003	1.049
	High	-0.304	1.004	2.495	0.883	1.011	1.152	0.924	1.012	1.102
<i>T</i> = 100										
None	Low	0.906	1.000	1.094	0.988	1.000	1.012	0.994	1.000	1.006
	Med	0.852	1.000	1.146	0.982	1.000	1.018	0.991	1.000	1.010
	High	0.692	1.002	1.347	0.966	1.003	1.039	0.981	1.002	1.022
Low	Low	0.904	1.000	1.095	0.985	1.000	1.016	0.991	1.000	1.009
	Med	0.850	1.000	1.150	0.978	1.000	1.023	0.988	1.001	1.014
	High	0.682	1.001	1.369	0.961	1.005	1.052	0.978	1.004	1.030
High	Low	0.901	1.000	1.099	0.986	1.000	1.014	0.991	1.000	1.009
	Med	0.847	0.999	1.154	0.978	1.000	1.022	0.987	1.001	1.015
	High	0.666	1.000	1.378	0.957	1.005	1.056	0.974	1.004	1.036
<i>T</i> = 200										
None	Low	0.970	1.000	1.029	0.996	1.000	1.004	0.998	1.000	1.002
	Med	0.955	1.000	1.045	0.994	1.000	1.006	0.997	1.000	1.004
	High	0.898	1.000	1.107	0.987	1.001	1.014	0.993	1.000	1.008
Low	Low	0.970	1.000	1.030	0.993	1.000	1.007	0.996	1.000	1.004
	Med	0.955	1.000	1.047	0.990	1.000	1.010	0.995	1.000	1.006
	High	0.895	1.000	1.117	0.983	1.002	1.024	0.991	1.001	1.012
High	Low	0.969	1.000	1.031	0.994	1.000	1.006	0.996	1.000	1.004
	Med	0.952	1.000	1.048	0.990	1.000	1.010	0.994	1.000	1.007
	High	0.890	1.001	1.122	0.979	1.002	1.028	0.987	1.002	1.018

neoclassical model with a cash in advance constraint generates a standard money demand function.

We follow Stock and Watson (1993), Ball (2001), and Hoffman *et al.* (1995) and approach long-run money demand as a cointegrating relationship. Our analysis suggests that instability exhibited by time-series estimates from the literature do not reflect underlying shifts in behavioural relationships but instead indicate inherent difficulties associated with estimation using

TABLE 6
Quantiles for second regressor slope (Case 3): panel DOLS and DOLS

CSD	Persistence	DOLS ($N = 10$)			Panel DOLS ($N = 10$)			Panel DOLS ($N = 20$)		
		2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
<i>T = 40</i>										
None	Low	0.072	0.100	0.129	0.098	0.100	0.102	0.099	0.100	0.101
	Med	0.065	0.101	0.142	0.098	0.100	0.103	0.098	0.100	0.101
	High	0.050	0.105	0.172	0.097	0.102	0.107	0.094	0.098	0.101
Low	Low	0.075	0.100	0.127	0.098	0.100	0.102	0.099	0.100	0.101
	Med	0.068	0.101	0.140	0.098	0.101	0.103	0.098	0.100	0.102
	High	0.052	0.105	0.172	0.097	0.102	0.108	0.094	0.098	0.102
High	Low	0.076	0.100	0.126	0.098	0.100	0.102	0.098	0.100	0.102
	Med	0.069	0.101	0.139	0.097	0.101	0.104	0.097	0.100	0.103
	High	0.054	0.105	0.170	0.096	0.103	0.110	0.092	0.098	0.104
<i>T = 100</i>										
None	Low	0.093	0.100	0.107	0.099	0.100	0.101	0.100	0.100	0.100
	Med	0.091	0.100	0.111	0.099	0.100	0.101	0.099	0.100	0.100
	High	0.085	0.102	0.128	0.099	0.101	0.102	0.098	0.099	0.101
Low	Low	0.094	0.100	0.106	0.099	0.100	0.101	0.100	0.100	0.100
	Med	0.092	0.100	0.110	0.099	0.100	0.101	0.099	0.100	0.101
	High	0.086	0.102	0.125	0.099	0.101	0.103	0.098	0.099	0.101
High	Low	0.094	0.100	0.106	0.099	0.100	0.101	0.099	0.100	0.101
	Med	0.092	0.100	0.110	0.099	0.100	0.101	0.099	0.100	0.101
	High	0.087	0.102	0.124	0.099	0.101	0.104	0.097	0.099	0.101
<i>T = 200</i>										
None	Low	0.097	0.100	0.103	0.100	0.100	0.100	0.100	0.100	0.100
	Med	0.096	0.100	0.105	0.100	0.100	0.100	0.100	0.100	0.100
	High	0.094	0.101	0.113	0.100	0.100	0.101	0.099	0.100	0.100
Low	Low	0.102	0.100	0.098	0.100	0.100	0.100	0.100	0.100	0.100
	Med	0.103	0.100	0.096	0.100	0.100	0.100	0.100	0.100	0.100
	High	0.105	0.099	0.089	0.100	0.100	0.099	0.101	0.100	0.100
High	Low	0.102	0.100	0.098	0.100	0.100	0.100	0.100	0.100	0.100
	Med	0.103	0.100	0.096	0.100	0.100	0.099	0.100	0.100	0.100
	High	0.105	0.099	0.090	0.101	0.099	0.098	0.101	0.100	0.099

relatively short sample spans in environments with persistent short-run dynamics. Combining observations across countries allows us to obtain relatively sharp and stable estimates of money demand elasticities and the panel cointegration approach seems well suited to take up King's (1988) suggestion to extend the money-demand analysis beyond the US. In his words, 'the results of such investigations would provide us with sharper estimates of the long run values of Friedman's (1956) "numerical constants

of monetary behavior” when we approach the difficult problem of the short run demand for money”.

The equation that we estimate is,

$$\ln \left[\frac{M_{it}}{P_{it}} \right] = \alpha_i + \lambda_i t + \theta_t + \gamma_y \ln Y_{it} + \gamma_r R_{it} + u_{it}^\dagger \quad (22)$$

for $i = 1, \dots, 19$, where M_{it} is an M1 measure of money, P_{it} is the price level, Y_{it} is real GDP, and R_{it} is a nominal short-term interest rate. Data definitions and sources are available in the unpublished appendix. In addition to country specific effects, α_i , we allow for possibly heterogeneous linear trends and common time effects. These trends are included to capture changes in the financial technology that affects money demand independent of income and the opportunity cost of holding money.

(i) Pre-testing: cointegration and homogeneity restrictions

Panel DOLS estimation of equation (22) requires that the equilibrium errors are stationary and that the cointegrating vectors for each country must be identical. To investigate the stationarity of the equilibrium errors, we employ Pedroni's (1999) panel- t test. This results in the rejection of the null hypothesis of no cointegration at the 0.1% level whether heterogeneous linear trends and common time effects are included or not.¹⁰

Next, we conduct a Wald test of the homogeneity restrictions on the cointegrating vector. When trends are omitted from the regression, the evidence against homogeneity is mixed. The asymptotic test rejects the restrictions in this case, but in some unreported Monte-Carlo experiments, we found moderate size distortion in the Wald test for sample sizes of $N = 19$ and $T = 40$. Using a size adjustment from these experiments, the homogeneity restrictions on income is rejected at the 5% level but not for the interest rate ($P = 0.30$). However, when we impose homogeneity on the interest rate slope, the test for slope homogeneity on income is not rejected ($P = 0.70$). When heterogeneous linear trends are included, the evidence supporting homogeneity strengthens. Here, we obtain a P -value of 0.63 for the test of homogeneity on the income coefficient and a P -value of 0.22 for the test of homogeneity on the interest rate coefficient.

¹⁰We also confirmed these cointegration test results by using Im, Pesaran and Shin (1997) and Maddala and Wu (1999) panel unit root tests under the assumption that the cointegrating vector is known to be $(1, -1.0, 0.05)$. The justification for using these values is that 1.0 is a typical value of the income elasticity estimated in the literature while a common estimate of the interest rate semi elasticity -0.05 .

TABLE 7

Single-equation and panel dynamic OLS estimates of long-run money demand

Country	No trend		With trend		Trend
	$\hat{\gamma}_y$ (s.e.)	$\hat{\gamma}_R$ (s.e.)	$\hat{\gamma}_y$ (s.e.)	$\hat{\gamma}_R$ (s.e.)	
Austria	0.901 (0.139)	-0.009 (0.029)	1.552 (0.349)	-0.037 (0.026)	-0.001
Belgium	0.134 (0.218)	0.009 (0.039)	1.183 (0.444)	-0.033 (0.026)	0.002
Denmark	1.460 (0.170)	-0.043 (0.009)	0.684 (0.321)	-0.036 (0.006)	0.009
Finland	1.019 (0.634)	-0.006 (0.114)	-0.740 (0.881)	0.009 (0.011)	-0.005
France	0.677 (0.213)	0.010 (0.020)	0.842 (0.699)	0.004 (0.031)	0.003
Germany	1.548 (0.033)	-0.019 (0.008)	1.691 (0.197)	-0.023 (0.009)	0.003
Iceland	0.594 (0.161)	-0.010 (0.005)	-0.451 (1.093)	-0.004 (0.008)	0.004
Ireland	0.507 (0.169)	0.022 (0.022)	1.670 (2.805)	0.015 (0.029)	0.005
Netherlands	1.112 (0.111)	-0.045 (0.020)	0.309 (0.415)	-0.011 (0.022)	0.003
Norway	2.641 (0.450)	-0.160 (0.046)	-0.676 (2.154)	-0.092 (0.060)	0.013
Portugal	0.517 (0.136)	-0.037 (0.017)	1.624 (0.379)	-0.043 (0.011)	0.010
Spain	1.203 (0.091)	-0.030 (0.008)	1.203 (0.190)	-0.030 (0.009)	0.003
Switzerland	1.020 (0.208)	-0.062 (0.021)	1.447 (0.482)	-0.053 (0.021)	0.011
UK	1.738 (0.097)	-0.089 (0.008)	2.128 (0.726)	-0.089 (0.008)	0.016
Japan	0.889 (0.599)	0.009 (0.200)	1.798 (0.415)	-0.076 (0.061)	0.016
Australia	0.926 (0.136)	-0.043 (0.012)	0.068 (0.329)	-0.048 (0.007)	0.002
New Zealand	1.349 (0.539)	-0.076 (0.026)	-1.233 (1.149)	-0.084 (0.018)	-0.001
Canada	1.245 (0.219)	-0.057 (0.024)	2.420 (0.903)	-0.078 (0.024)	-0.013
US	0.428 (0.074)	-0.035 (0.008)	1.022 (0.417)	-0.039 (0.007)	-0.001
Panel	0.860 (0.092)	-0.020 (0.007)	1.079 (0.264)	-0.022 (0.006)	—
Panel ^a	0.820 (0.105)	-0.017 (0.005)	0.986 (0.336)	-0.016 (0.005)	—

Note: ^aControls for common time effect.**(ii) Comparison between single-equation and panel DOLS estimates**

Our panel DOLS estimates use two leads and two lags of $\Delta \ln Y_{it}$ and ΔR_{it} in the regressions. Point estimates and asymptotic standard errors are reported in Table 7.

Single-equation DOLS estimates are seen to display such cross-sectional variability that they are difficult to interpret. In DOLS regressions without trend, the income elasticities are all positive, ranging from 0.134 (Belgium) to a whopping 2.64 (Norway), but the interest semi-elasticity has the wrong sign for Belgium, France, Ireland, and Japan. When a trend is included in the regression, income elasticity estimates are negative for Finland, Iceland, Norway, and New Zealand, and interest semi-elasticity estimates are positive for Finland, France, and Iceland. If we maintain an underlying belief that the financial systems and transactions technologies across modern economies are essentially similar, the cross-sectional variability in these estimates must reflect the inherent difficulty of obtaining good estimates rather than evidence of disparate economic behavior.

Panel DOLS estimates are shown at the bottom of Table 7. When the panel regression omits trends, we estimate 0.86 (asymptotic s.e. = 0.09) and the interest semi-elasticity to be -0.02 (asymptotic s.e. = 0.01). When we include heterogeneous trends, we estimate the income elasticity to be 1.08 (asymptotic s.e. = 0.26) and the interest semi-elasticity to be -0.02 (asymptotic s.e. = 0.01). Results obtained from controlling for CSD are very similar.

To further illustrate the problem of estimation instability in the time dimension, we constructed recursive single-equation DOLS coefficient estimates for the US, UK, France, and Japan and panel DOLS for all 19 countries. Recursive DOLS estimates from 1979 to 1995 for both the income elasticity and interest semi-elasticity exhibit substantially more variability than the recursive panel DOLS estimates and in several instances even change sign. These results are also contained in the unpublished appendix.

VI. Conclusions

Heterogeneity and persistence in short-run dynamics can create substantial variability in single-equation cointegration vector point estimates. The result is that these estimators can be quite sensitive to the particular time span of the observations as well as to the particular individual being studied. This small sample fragility can be encountered in spite of the superconsistency of these estimators.

In these environments, panel DOLS can provide much more precise estimates. Panel DOLS is straightforward to compute and relevant test statistics have standard asymptotic distributions. The asymptotic distributions were found to provide reasonably close approximations to the exact sampling distributions in small samples.

We applied the panel DOLS method to estimate the long-run money demand function using a panel of 19 countries with annual data from 1957 to 1996. The estimates in which we have the most confidence are an income elasticity near 1 and an interest rate semi-elasticity of -0.02 .

Final Manuscript received: April 2003

References

- Baba, Y., Hendry, D. F. and Starr, R. M. (1992). 'The demand for M1 in the U.S.A., 1960–1988', *Review of Economic Studies*, Vol. 59, pp. 25–61.
- Bai, J. and Ng, S. (2001). 'Determining the number of factors in approximate factor models', *Econometrica*, Vol. 70, pp. 191–221.
- Bailey, M. J. (1956). 'The welfare cost of inflationary finance', *Journal of Political Economy*, Vol. 64, pp. 93–110.

- Ball, L. (2001). 'Another look at long-run money demand', *Journal of Monetary Economics*, Vol. 47, pp. 37–44.
- Friedman, M. (1956). 'The quantity theory of money – a restatement', in Friedman M. (ed.), *Studies in the Quantity Theory of Money*, University of Chicago Press, Chicago.
- Gandolfi, A. E. and Lothian, J. R. (1976). 'The demand for money from the great depression to the present', *American Economic Review*, Vol. 66, pp. 46–51.
- Hoffman, D. L., Rasche, R. H. and Tieslau, M. A. (1995). 'The stability of long-run money demand in five industrial countries', *Journal of Monetary Economics*, Vol. 35, pp. 317–339.
- Im, K. S., Pesaran, M. H. and Shin, Y. (1997). *Testing for Unit Roots in Heterogeneous Panels*, Discussion Paper, University of Cambridge, Cambridge.
- Inder, B. (1993). 'Estimating long-run relationships in economics: a comparison of different approaches', *Journal of Econometrics*, Vol. 57, pp. 53–68.
- Kao, C. and Chiang, M.-H. (2000). 'On the estimation and inference of a cointegrated regression in panel data', in Baltagi B. H. (ed.), *Advances in Econometrics: Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Vol. 15, pp. 179–222.
- King, R. G. (1988). *Money Demand in the United States: A Quantitative Review: A Comment*, Carnegie-Rochester Conference Series on Public Policy, Vol. 29, pp. 169–172.
- Lucas, R. E. Jr (1988). *Money Demand in the United States: A Quantitative Review*, Carnegie-Rochester Conference Series on Public Policy, Vol. 29, pp. 137–168.
- Maddala, G. S. and Wu, S. (1999). 'A comparative study of unit root tests with panel data and a new simple test', *Oxford Bulletin of Economics and Statistics*, Vol. 61, pp. 631–652.
- Meltzer, A. H. (1963). 'The demand for money: a cross-section study of business firms', *Quarterly Journal of Economics*, Vol. 77, pp. 405–422.
- Moon, H. R. and Perron, P. (2002). 'Testing for a unit root in panels with dynamic factors', University of Southern California, mimeo.
- Mulligan, C. B. (1997). 'Scale economies, the value of time, and the demand for money: longitudinal evidence from firms', *Journal of Political Economy*, Vol. 105, pp. 1061–1079.
- Mulligan, C. B. and Sala-i-Martin, X. (1992). 'U.S. money demand: surprising cross-sectional estimates', *Brookings Papers on Economic Activity*, No. 2, pp. 285–329.
- Pedroni, P. (1997). 'Fully modified OLS for heterogeneous cointegrated panels and the case of purchasing power parity', Department of Economics, Indiana University, mimeo.
- Pedroni, P. (1999). 'Critical values for cointegration tests in heterogeneous panels with multiple regressors', *Oxford Bulletin of Economics and Statistics*, Vol. 61, pp. 653–670.
- Phillips, P. C. B. and Moon, H. R. (1999). 'Linear regression limit theory for nonstationary panel data', *Econometrica*, Vol. 67, pp. 1057–1111.
- Phillips, P. C. B. and Moon, H. R. (2000). 'Nonstationary panel data analysis: an overview of some recent development', *Econometric Reviews*, Vol. 19, pp. 263–286.
- Phillips, P. C. B. and Sul, D. (2002). 'Dynamic panel estimation and homogeneity testing under cross section dependence', Yale University, mimeo.
- Saikkonen, P. (1991). 'Asymptotically efficient estimation of cointegration regressions', *Econometric Theory*, Vol. 7, pp. 1–21.
- Stock, J. H. and Watson, M. W. (1993). 'A simple estimator of cointegrating vectors in higher order integrated systems', *Econometrica*, Vol. 61, pp. 783–820.
- Sul, D., Phillips, P. C. B. and Choi, C.-Y. (2002). 'Prewhitening Bias in HAC Estimation', Yale University, mimeo.