Accurate statistical analysis in dynamic panel data models
Bun, M.J.G.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 4

Bias approximation in large T, small N panels

4.1 Introduction

In this chapter we analyse various least squares based estimation procedures for the higher-order dynamic panel data model with fixed individual effects and a non-scalar covariance matrix. Both the ordinary and generalised Least Squares Dummy Variables (LSDV) estimators are considered. The choice of the model and estimators is based mainly on the typical empirical study at hand, i.e. estimation of money demand functions in the area of the European Union (EU). The data are a cross-section of times series for 14 EU countries (Luxembourg is not included in the data) and the number of cross-section units \( N \) in the dataset is relatively small compared with the time dimension \( T \). In the simulation study of Chapter 2 we found that in the first-order dynamic panel data model with a scalar covariance matrix the bias of least squares based techniques is relatively small compared to instrumental variables based methods when \( T \) is larger than \( N \). Based on a mean squared error criterion, least squares methods are to be preferred in this case.

Notwithstanding the relatively good performance of least squares methods, they are still biased in dynamic models due to the inclusion of lagged dependent variable regressors and require \( T \) large for consistency. The simulations in Chapter 2 and other Monte Carlo experiments (Judson and Owen, 1999) show that these biases can be substantial especially for persistent models, i.e. stable dynamic models where the coefficient of the lagged dependent variable is close to one. As these models are commonly encountered in macro-economic applications like the empirical study on money demand, it is important to develop more accurate estimation procedures.
In this chapter we will develop alternative coefficient estimators by applying bias corrections to original estimators. Although the approach adopted here is comparable to that pursued in Chapters 2 and 3, there are several important differences. First, as the dimensions of the data are different from the typical small $T$, large $N$ panel, higher-order approximations will be developed from a large $T$ perspective while keeping $N$ fixed. Second, as $N$ is finite we assume fixed instead of random individual specific effects. Third, the model of Chapters 2 and 3 will be extended in several directions. These extensions are necessary to apply bias corrected estimators in applied research. In the empirical study on EU wide money demand, for example, it turns out that the first-order dynamic model is not general enough to capture all the dynamic features in the data. Regarding the disturbance covariance structure, when analysing time series for a group of countries both cross-sectional heteroscedasticity and interdependencies between countries are likely to be present. Hence, disturbance vectors of different cross-sectional units have different variances and may be correlated. To the extent that the covariance matrix of the disturbances is non-scalar, one should explicitly take this into account in any inference procedure exploiting the panel nature of the data.

Using asymptotic expansion techniques, Kiviet (1995, 1999) derives an approximation formula for the bias of the ordinary LSDV estimator in the first-order stable dynamic panel data model with normal disturbances and a scalar covariance matrix (see also Section 2.3 of this thesis). We use these and other results on bias approximation in higher-order dynamic regression models (Kiviet and Phillips, 1994) to develop bias expressions for the ordinary and generalised LSDV estimators in higher-order dynamic panel data models with general covariance structure.

Apart from developing more accurate estimation procedures for the so-called short-run parameters, estimation of other model parameters will be considered also. First, in the case of the money demand relationship the long-run effects are important for policymakers. Hence, a clear distinction is made between estimation of short- and long-run parameter vectors. The direct bias correction on long-run coefficients, proposed by Pesaran and Zhao (1999) in the context of the dynamic random coefficient model, is applicable here also. Second, in practice often particular linear restrictions are imposed on the parameters before estimation. For example, in empirical studies of money demand often long-run price or income homogeneity is imposed. Hence, we also develop bias approximation formulae for restricted estimators along the lines of Kiviet and Phillips (1994). Third, in the type of model analysed here, estimation of the variances of coefficient estimators by conventional asymptotic expressions can be dramatically inaccurate (Freedman and Peters, 1984; Beck and Katz, 1995). Hence, we make use of bootstrap procedures to
estimate standard errors.

The bias expressions developed in this chapter will be used to construct bias corrected estimators and they will be applied in the empirical study on money demand. Various authors have estimated a money demand function based on aggregated time series for the whole EU area and tested the stability of this function through time (Kremers and Lane, 1990; Monticelli and Papi, 1996; Fase and Winder, 1998). All those studies use time series techniques, but considering the EU countries as a cross-section one can possibly use panel data techniques. We will examine to what extent panel data techniques are a valid alternative for estimating EU wide money demand functions. Using aggregated time series it has been found that EU wide money demand is more stable than money demand in the individual member countries. An explanation for this fact is that when using aggregated data spillover effects between countries are internalised. As money demand specifications for individual countries typically do not contain variables measuring foreign developments, specification bias may cause them to be less stable than aggregated money demand. Exploiting the panel nature of the data we try to identify spillovers between countries by specifying a general disturbance covariance structure and including foreign variables in the empirical specification.

Section 4.2 gives an outline of the model. In evaluating the bias terms of the estimators, a detailed knowledge of the stochastic structure of the model is needed and this will be described in this section. In Section 4.3 bias expressions for ordinary and generalised LSDV estimators will be developed. In Section 4.4 we consider estimation under linear restrictions, while estimators of long-run parameters will be analysed in Section 4.5. In Section 4.6 estimation of asymptotic standard errors by using either analytical expressions or bootstrap procedures will be discussed. In Section 4.7 the estimation techniques will be applied to estimate EU wide money demand functions for M1, M2 and M3. The emphasis is on the plausibility of coefficient estimates and effectiveness of bias approximations. Section 4.8 concludes.

4.2 Model

In Chapters 2 and 3 we have considered the first-order dynamic panel data model\(^1\). Here, we extend the analysis to higher-order dynamic models, i.e.

\[
y_{it} = \sum_{p=1}^{P} \gamma_{p} y_{i,t-p} + \beta' x_{it} + \eta_{i} + \varepsilon_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T.
\]

\(^1\)We will use similar notation as in Section 2.2.
In this model the dependent variable $y_{it}$ is regressed on a $K \times 1$ vector of explanatory variables $x_{it}$ with parameter vector $\beta$, $P$ lagged values of the dependent variable and an individual specific constant. The explanatory variables in $x_{it}$ are assumed to be strictly exogenous, i.e.

$$E[x_{it}\varepsilon_{js}] = 0, \quad \forall i, j, t, s,$$  \tag{4.2}

and the individual effects $\eta_i$ are assumed fixed, but unknown. Note that both the univariate processes for $y_{it}$ and the elements of $x_{it}$ may contain unit roots. However, the relationship (4.1) between $y_{it}$ and $x_{it}$ is assumed to be stable. For $P = 1$ this implies $|\gamma_1| < 1$, but in higher-order models more complicated restrictions on the autoregressive coefficients are required for stability. Regarding the disturbances $\varepsilon_{it}$ it will be assumed throughout that they are normally distributed. Moreover, they are uncorrelated through time, but we allow for heteroscedasticity across cross-sectional units and non-zero contemporaneous cross-correlations, i.e.

$$E[\varepsilon_{it}] = 0, \quad \forall i, t,$$

$$E[\varepsilon_{it}\varepsilon_{jt}] = 0, \quad \forall i, j, t \neq s,$$

$$E[\varepsilon_{it}\varepsilon_{jt}] = \sigma_{ij}, \quad \forall i, j, t. \quad \tag{4.3}$$

Stacking the observations over time we get

$$y_i = \sum_{p=1}^{P} \gamma_p y_{i,-p} + X_i \beta + \eta_{i,T} + \varepsilon_i$$

$$= W_i \delta + \eta_{i,T} + \varepsilon_i, \quad \tag{4.4}$$

where $y_{i,-p} = (y_{i,1-p}, \ldots, y_{i,T-p})'$, $X_i = (x_{i1}, \ldots, x_{iT})'$, $\delta = (\gamma', \beta')'$, $\gamma = (\gamma_1, \ldots, \gamma_P)'$ and $W_i = [y_{i,-1} \ldots y_{i,-p} : X_i]$.

Like in Section 2.2 we decompose $y$ into a relevant random component, denoted by a tilde, and an irrelevant random plus deterministic components, denoted by a bar. The relevant random component is in some way related to the disturbance term $\varepsilon_{it}$, while the irrelevant component is not, i.e.

$$\tilde{y}_i = \sum_{p=1}^{P} \gamma_p \tilde{y}_{i,-p} + \varepsilon_i,$$

$$\bar{y}_i = \sum_{p=1}^{P} \gamma_p \bar{y}_{i,-p} + X_i \beta + \eta_{i,T}, \quad \tag{4.5}$$

where we use the assumption that we have fixed individual effects and only strictly exogenous explanatory variables, i.e. $\bar{X}_i = 0$ and $\tilde{y}_i = 0$. For the initial values we assume

$$\tilde{y}_{i,1-p} = 0, \quad \bar{y}_{i,1-p} = y_{i,1-p}, \quad \tag{4.6}$$
for \( p = 1, ..., P \), so we condition on \( p \) fixed starting values. Defining

\[
\Gamma_T = \left( I_T - \sum_{p=1}^{P} \gamma_p L_T^p \right)^{-1},
\]

we write for the relevant random components in (4.5)

\[
\tilde{y}_i = \Gamma_T \varepsilon_i, \quad i = 1, ..., N.
\]

To analyse the estimators in the next section we need a decomposition of the matrix \( A_T W_i \) into relevant and irrelevant random components. We may write \( A_T W_i = A_T \tilde{W}_i + A_T \tilde{\tilde{W}}_i \) with

\[
A_T \tilde{W}_i = \sum_{p=1}^{P} A_T L_T^p \Gamma_T \varepsilon_i e'_p,
\]

because \( \tilde{X}_i = O \) and \( A_T \tilde{y}_{i,-p} = A_T L_T^p \tilde{y}_i \) and where \( e_p \) is the \((P + K) \times 1\) unit vector with its \( p \)th element equal to one.

Stacking the observations across individuals too, one gets

\[
y = W \delta + S \eta + \varepsilon,
\]

where \( y \) and \( \varepsilon \) are \( NT \times 1 \) vectors, \( \eta = (\eta_1, ..., \eta_N)' \) is a \( N \times 1 \) vector and \( W = [W_1'; ..., W_N']' \)

and \( S = I_N \otimes I_T \) are \( NT \times (K + P) \) and \( NT \times N \) matrices respectively. The assumptions about \( \varepsilon \) can be written as

\[
\varepsilon \sim \mathcal{N}(0, \Omega),
\]

where \( \Omega = \Sigma \otimes I_T \) with \( \Sigma \) a \( N \times N \) matrix with typical element \( \sigma_{ij} \). For the relevant stochastic components in \( A W \) we find from (4.9)

\[
A \tilde{W} = \sum_{p=1}^{P} A L^p \Gamma \varepsilon e'_p
\]

\[
= \sum_{p=1}^{P} \Pi_p \varepsilon e'_p,
\]

where \( \Pi_p = A L^p \Gamma \).

Model (4.10) with (4.11) is a generalised normal regression model and in the next sections estimation of both short- and long-run coefficients will be considered. The elements of the parameter vector \( \delta \) are called short-run coefficients and \( \theta = \beta/(1 - \sum_{p=1}^{P} \gamma_p) \) is the vector of long-run effects.
4.3 Short-run coefficient estimators

4.3.1 LSDV estimator

The ordinary least squares estimator for $\delta$ in (4.10) is the familiar LSDV estimator

$$\hat{\delta}_{LSDV} = (W'AW)^{-1}W'Ay,$$

and its estimation error is

$$\hat{\delta}_{LSDV} - \delta = (W'AW)^{-1}W'A\varepsilon,$$

which depends in a non-linear way on the stochastic term $\varepsilon$. Defining $Q^{-1} = E[W'AW]$ and using a similar approach as in Section 2.3 (see also Kiviet, 1995, 1999), but now for fixed $N$, the first factor in (4.14) may be expanded as

$$(W'AW)^{-1} = Q - Q(W'AW - Q^{-1})Q + O_p(T^{-2}),$$

where the first two terms are $O(T^{-1})$ and $O_p(T^{-\frac{3}{2}})$ respectively. Hence, for the bias of the LSDV estimator we find

$$E \left[ \hat{\delta}_{LSDV} - \delta \right] = 2QE[W'A\varepsilon] - QE[W'AWQW'A\varepsilon] + o(T^{-1}),$$

where the first two terms in (4.16) are $O(T^{-1})$, as we shall see\(^2\). Note that the $O(T^{-1})$ contribution in (4.16) is encompassing its counterpart in the bias approximation (2.31). Specifically, it is similar to the sum of the $c_1(T^{-1})$ and $c_2(N^{-1}T^{-1})$ terms in (2.31). This is the result of using large $T$ approximations here keeping $N$ fixed.

In Appendix 4.A it is shown that the approximation for the bias in the LSDV estimator equals

$$E \left[ \hat{\delta}_{LSDV} - \delta \right] = B_{LSDV}(T^{-1}) + o(T^{-1}),$$

where

$$B_{LSDV}(T^{-1}) = \sum_p \text{tr}(\Pi_p \Omega)Qe_p - \sum_p Q\bar{W}'\Pi_p \Omega AWQe_p$$

$$-\sum_p \text{tr} [Q\bar{W}'\Pi_p \Omega AW] Qe_p$$

$$-2 \sum_p \sum_r \sum_s q_{rs} \text{tr}(\Omega \Pi'_p \Pi_r \Pi_s \Omega) Qe_p,$$

\(^2\)The notation $o(T^{-1})$ indicates that the corresponding term is of smaller order than $T^{-1}$.
4.3. Short-run coefficient estimators

with the indices \( p, r \) and \( s \) running from 1 to \( P \) and

\[
Q = \left[ \tilde{W}'A\tilde{W} + \sum_p \sum_r \text{tr}(\Pi_p'\Pi_r\Omega)e_pe_r' \right]^{-1},
\]

\[
q_{rs} = e'_rQe_s.
\]

Using this result an operational bias corrected estimator, denoted by LSDVc, can be constructed as

\[
\hat{\delta}_{LSDVc} = \hat{\delta}_{LSDV} - \hat{B}_{LSDV}(T^{-1}),
\]

(4.19)

using any consistent preliminary estimators for \( \delta \) and \( \Omega \) in \( \hat{B}_{LSDV}(T^{-1}) \), e.g based on LSDV residuals. The corrected LSDV estimator will be unbiased up to order \( O(T^{-1}) \), i.e.

\[
E[\hat{\delta}_{LSDVc} - \delta] = o(T^{-1}).
\]

(4.20)

4.3.2 (Feasible) Generalised LSDV estimator

The ordinary LSDV estimator in (4.13) does not take the covariance structure of \( \varepsilon \) into account. Hence, we analyse also the generalised LSDV estimator of \( \delta \), denoted by \( \hat{\delta}_{GLSDV} \). The GLSDV estimator of \( \delta \) is

\[
\hat{\delta}_{GLSDV} = (W'A\Omega^{-1}A)^{-1}W'A\Omega^{-1}A\gamma,
\]

(4.21)

and its estimation error is

\[
\hat{\delta}_{GLSDV} - \delta = (W'A\Omega^{-1}A)^{-1}W'A\Omega^{-1}A\varepsilon.
\]

(4.22)

Defining \( A^* = A\Omega^{-1}A \) and \( Q^{-1} = E[W'A^*W] \) the following expansion is valid under the assumptions made in Section 4.2, i.e.

\[
(W'A^*W)^{-1} = Q^* - Q^*(W'A^*W - Q^{-1})Q^* + O_p(T^{-2}).
\]

(4.23)

Hence, for the bias of the GLSDV estimator we find

\[
E[\hat{\delta}_{GLSDV} - \delta] = 2Q^*E[W'A^*\varepsilon] - Q^*E[W'A^*WQ^*W'A^*\varepsilon] + o(T^{-1}).
\]

(4.24)

In Appendix 4.A the following approximation is derived for the bias in the GLSDV estimator, i.e.

\[
E[\hat{\delta}_{GLSDV} - \delta] = B_{GLSDV}(T^{-1}) + o(T^{-1}),
\]

(4.25)
Chapter 4. Bias approximation in large T, small N panels

where

\[ B_{GLSDV}(T^{-1}) = \sum_p \text{tr}(\Pi_p)Q^*e_p - \sum_p Q^*\hat{W}'\Omega^{-1}\Pi_pA\hat{W}Q^*e_p \]

\[ - \sum_p \text{tr} (Q^*\hat{W}'\Omega^{-1}\Pi_pA\hat{W})Q^*e_p \]

\[ -2 \sum_p \sum_r \sum_s q^*_{rs}\text{tr}(\Pi_p^r\Pi_p^s)Q^*e_p, \]  \hspace{1cm} (4.26)

with

\[ Q^* = \left[ \hat{W}'A\Omega^{-1}A\hat{W} + \sum_p \sum_r \text{tr}(\Pi_p^r)\epsilon_p'\epsilon_r' \right]^{-1}, \]

\[ q^*_{rs} = \epsilon_r'Q^*e_s. \]

In practice the GLSDV estimator cannot be calculated because \( \Omega \) is unknown. We therefore analyse also the two-step feasible GLSDV estimator

\[ \hat{\delta}_{FGLSDV} = (W'A\hat{\Omega}^{-1}AW)^{-1}W'A\hat{\Omega}^{-1}Ay, \]  \hspace{1cm} (4.27)

where the covariance matrix \( \Omega \) is consistently estimated using the LSDV residuals, i.e.

\[ \hat{\Omega} = \hat{\Sigma} \otimes I_T, \]

\[ \hat{\sigma}_{ij} = \frac{(y_i - W_i\hat{\delta}_{LSDV})'A_T(y_j - W_j\hat{\delta}_{LSDV})}{T}. \]  \hspace{1cm} (4.28)

The estimation error of the FGLSDV estimator is

\[ \hat{\delta}_{FGLSDV} - \delta = (W'A\hat{\Omega}^{-1}AW)^{-1}W'A\hat{\Omega}^{-1}A\varepsilon, \]  \hspace{1cm} (4.30)

In Appendix 4.B it is shown that as long as a consistent estimator is used for \( \Omega \) the bias in the FGLSDV estimator is

\[ E \left[ \hat{\delta}_{FGLSDV} - \delta \right] = B_{FGLSDV}(T^{-1}) + o(T^{-1}), \]  \hspace{1cm} (4.31)

with

\[ B_{FGLSDV}(T^{-1}) = B_{GLSDV}(T^{-1}) + o(T^{-1}). \]  \hspace{1cm} (4.32)

In other words, the bias approximation to order \( O(T^{-1}) \) is equal for the FGLSDV and GLSDV estimators. Hence, we can construct an operational bias corrected estimator, denoted by FGLSDVc, as

\[ \hat{\delta}_{FGLSDVc} = \hat{\delta}_{FGLSDV} - \hat{B}_{FGLSDV}(T^{-1}), \]  \hspace{1cm} (4.33)

using any consistent preliminary estimators for \( \delta \) and \( \Omega \) in \( \hat{B}_{FGLSDV}(T^{-1}) \). The corrected FGLSDV estimator will be unbiased up to order \( O(T^{-1}) \), i.e.

\[ E \left[ \hat{\delta}_{FGLSDVc} - \delta \right] = o(T^{-1}). \]  \hspace{1cm} (4.34)
4.4 Estimation under linear restrictions

In practice, model (4.10) is often estimated in a restricted version for several reasons. First, economic theory may impose a priori certain values for some of the model parameters. For example, in money demand studies often a long-run income or price elasticity of one is imposed. Second, econometric modelling may lead, for example, to the exclusion of variables from the most general model to arrive at a more parsimoneous specification. In this section we will develop bias expressions for the restricted LSDV, GLSDV and FGLSDV short-run coefficient estimators along the lines of Kiviet and Phillips (1994).

Consider a set of \( J \) general linear restrictions on \( \delta \), i.e.

\[
R\delta = r, \tag{4.35}
\]

where \( R \) is \( J \times (K + P) \) and \( r = J \times 1 \). Note that no restrictions regarding the individual constants will be considered as they are typically filtered out before estimation.

4.4.1 restricted LSDV estimator

The LSDV estimator of \( \delta \) under the restrictions (4.35) is

\[
\hat{\delta}_{LSDV,R} = \hat{\delta}_{LSDV} - (W'AW)^{-1}R' [R(W'AW)^{-1}R']^{-1} (R\hat{\delta}_{LSDV} - r). \tag{4.36}
\]

Below we shall derive a bias expression up to order \( O(T^{-1}) \) for the estimator in (4.36). The relation between the unrestricted and restricted LSDV estimators can be written as

\[
\hat{\delta}_{LSDV,R} - \delta = F_{(W)}(\hat{\delta}_{LSDV} - \delta), \tag{4.37}
\]

where \( F_{(W)} = I - (W'AW)^{-1}R' [R(W'AW)^{-1}R']^{-1} R \). Hence, the bias in the restricted LSDV estimator can be expressed as

\[
E \left[ \hat{\delta}_{LSDV,R} - \delta \right] = E \left[ F_{(W)}(W'AW)^{-1}W'A\varepsilon \right]. \tag{4.38}
\]

The evaluation of the right hand side of (4.39) is not straightforward because the elements of \( F_{(W)} \) are also stochastic. However, applying a similar reasoning as in the proof of Theorem 2 in Kiviet and Phillips (1994) it can be shown that

\[
E \left[ \hat{\delta}_{LSDV,R} - \delta \right] = -F_{(W)}Q E \left[ W'AWF_{(W)}QW'A\varepsilon \right] + o(T^{-1}), \tag{4.39}
\]

where \( F_{(W)} = I - QR' [RQR']^{-1} R \). The evaluation of the right hand side in (4.39) is now similar to earlier expectations and one gets

\[
E \left[ \hat{\delta}_{LSDV,R} - \delta \right] = B_{LSDV,R}(T^{-1}) + o(T^{-1}), \tag{4.40}
\]
Chapter 4. Bias approximation in large $T$, small $N$ panels

where

$$B_{LSDV,R}(T^{-1}) = \sum_p tr(\Pi_p \Omega) F(\bar{W}) Q e_p$$

(4.41)

$$- \sum_p F(\bar{W}) Q \bar{W}' \Pi_p \Omega A \bar{W} F(\bar{W}) Q e_p$$

$$- \sum_p tr [F(\bar{W}) Q \bar{W}' \Pi_p \Omega A \bar{W}] F(\bar{W}) Q e_p$$

$$- 2 \sum_p \sum_r \sum_s e_r^* F(\bar{W}) Q e_s tr (\Omega \Pi_p \Pi_s) F(\bar{W}) Q e_p.$$

4.4.2 restricted (Feasible) Generalised LSDV estimator

The GLSDV estimator of $\delta$ under the restrictions (4.35) is

$$\hat{\delta}_{GLSDV,R} = \hat{\delta}_{GLSDV} - (W' A \Omega^{-1} W)^{-1} R' [R(W' A \Omega^{-1} W)^{-1} R']^{-1} (R \hat{\delta}_{GLSDV} - \tau),$$

(4.42)

which can be rearranged as

$$\hat{\delta}_{GLSDV,R} - \delta = F(\bar{W}_\Omega) (\hat{\delta}_{GLSDV} - \delta),$$

(4.43)

where $F(\bar{W}) = I - (W' A \Omega^{-1} W)^{-1} R' [R(W' A \Omega^{-1} W)^{-1} R']^{-1} R$. The bias of the restricted GLSDV estimator is

$$E \left[ \hat{\delta}_{GLSDV,R} - \delta \right] = E \left[ F(\bar{W}_\Omega) (W' A \Omega^{-1} W)^{-1} W' A \Omega^{-1} \epsilon \right],$$

(4.44)

which can be written as

$$E \left[ \hat{\delta}_{GLSDV,R} - \delta \right] = - F(\bar{W}_\Omega) Q' E \left[ W' A \Omega^{-1} W F(\bar{W}_\Omega) Q' W' A \Omega^{-1} \epsilon \right] + o(T^{-1}),$$

(4.45)

where $F(\bar{W}_\Omega) = I - Q' R' [RQ' R']^{-1} R$. The evaluation of the right hand side is now similar to earlier derivations and one gets

$$E \left[ \hat{\delta}_{GLSDV,R} - \delta \right] = B_{GLSDV,R}(T^{-1}) + o(T^{-1}),$$

(4.46)

where

$$B_{GLSDV,R}(T^{-1}) = \sum_p tr(\Pi_p) F(\bar{W}_\Omega) Q^* e_p$$

(4.47)

$$ - \sum_p F(\bar{W}_\Omega) Q^* \bar{W}' \Omega^{-1} \Pi_p A \bar{W} F(\bar{W}_\Omega) Q^* e_p$$

$$ - \sum_p tr [F(\bar{W}_\Omega) Q^* \bar{W}' \Omega^{-1} \Pi_p A \bar{W}] F(\bar{W}_\Omega) Q^* e_p$$

$$ - 2 \sum_p \sum_r \sum_s e_r^* F(\bar{W}) Q^* e_s tr (\Omega \Pi_p \Pi_s) F(\bar{W}_\Omega) Q^* e_p.$$
4.5. Long-run coefficients

The restricted GLSDV estimator is not applicable because it depends on the unknown covariance matrix $\Omega$. Hence, we consider the feasible GLSDV estimator under restrictions, i.e.

$$\hat{\delta}_{FGLSDV,R} = \hat{\delta}_{FGLSDV} - (W'\hat{A}^{-1}W)^{-1}R' \left[R(W'\hat{A}^{-1}W)^{-1}R'\right]^{-1} (R\hat{\delta}_{FGLSDV} - r).$$

(4.48)

In Appendix 4.C it is shown that

$$E \left[\hat{\delta}_{FGLSDV,R}\right] = B_{FGLSDV,R}(T^{-1}) + o(T^{-1}),$$

(4.49)

with

$$B_{FGLSDV,R}(T^{-1}) = B_{GLSDV,R}(T^{-1}) + o(T^{-1}),$$

(4.50)

i.e. the bias approximation to order $O(T^{-1})$ is the same for the restricted FGLSDV and GLSDV estimators. Hence, a bias corrected restricted FGLSDV estimator can be constructed with the expression in (4.47), which is unbiased up to order $T^{-1}$.

4.5 Long-run coefficients

The short-run estimators in the previous sections can be used to construct estimators for the long-run effects $\theta$, which are defined by

$$\hat{\theta} = \hat{\beta}/(1 - \iota_P')\hat{\gamma},$$

(4.51)

where $\hat{\beta}$ and $\hat{\gamma}$ are any of the estimators considered before and $\iota_P$ is a $P \times 1$ vector of ones. If bias corrected short-run estimators like (4.19) or (4.33) are used to correct for bias in (4.51) the resulting long-run estimator is called "naive" by Pesaran and Zhao (1999), which analyse several estimators of the long-run coefficients in the context of the dynamic random coefficient model. The "naive" or indirect way of bias correction in (4.51) does not lead to an estimator unbiased to order $O(T^{-1})$. Note that the estimation error is $\hat{\delta} - \delta = O_p(T^{-\frac{1}{2}})$ irrespective of the estimator used (as long as it is $\sqrt{T}$ consistent).
Hence, in general we can write
\[
\hat{\theta} = \frac{\hat{\beta}}{(1 - \iota'_p \gamma) - \iota'_p (\hat{\gamma} - \gamma)}
\]

(4.52)
\[
= \frac{\hat{\beta}}{(1 - \iota'_p \gamma)} \left[ 1 - \iota'_p (\hat{\gamma} - \gamma) \right]^{-1}
\]
\[
= \frac{\hat{\beta}}{(1 - \iota'_p \gamma)} + \frac{\beta}{(1 - \iota'_p \gamma) (1 - \iota'_p \gamma)} \frac{\iota'_p (\hat{\gamma} - \gamma)}{(1 - \iota'_p \gamma)} + \frac{\beta}{(1 - \iota'_p \gamma)^2} \left( \iota'_p \hat{\gamma} - \iota'_p \gamma \right)^2 + O_p(T^{-\frac{1}{2}})
\]
\[
= \frac{\beta}{(1 - \iota'_p \gamma)} + \frac{1}{(1 - \iota'_p \gamma)^2} (\hat{\beta} - \beta) (\iota'_p \hat{\gamma} - \iota'_p \gamma) + \frac{\beta}{(1 - \iota'_p \gamma)^3} (\iota'_p \hat{\gamma} - \iota'_p \gamma)^2 + O_p(T^{-\frac{1}{2}}).
\]

Therefore, we find for the bias in estimating \( \theta \) the following
\[
E \left[ \hat{\theta} - \theta \right] = \frac{1}{(1 - \iota'_p \gamma)} E \left[ \hat{\beta} - \beta \right] + \frac{\beta}{(1 - \iota'_p \gamma)^2} E \left[ \iota'_p \hat{\gamma} - \iota'_p \gamma \right]
\]
(4.53)
\[
+ \frac{1}{(1 - \iota'_p \gamma)^2} E \left[ \hat{\beta} - \beta \right] (\iota'_p \hat{\gamma} - \iota'_p \gamma) + \frac{\beta}{(1 - \iota'_p \gamma)^3} E \left[ (\iota'_p \hat{\gamma} - \iota'_p \gamma)^2 \right] + o(T^{-1}),
\]
which is \( O(T^{-1}) \) because all explicit terms in (4.53) are in general non-zero and of order \( O(T^{-1}) \).

Pesaran and Zhao (1999) propose a direct way of bias correction. Using any original uncorrected estimator, i.e. (4.13), (4.27), (4.36) or (4.48), and rearranging (4.53) we find for the bias in the long-run coefficient vector \( \theta \)
\[
E \left[ \hat{\theta} - \theta \right] = B_\theta(T^{-1}) + o(T^{-1}),
\]
(4.54)
where
\[
B_\theta(T^{-1}) = \frac{1}{(1 - \iota'_p \gamma)^2} \left[ (1 - \iota'_p \gamma) (B_\beta + \theta \iota'_p B_\gamma) + \text{Cov} \left[ \hat{\beta}, \iota'_p \hat{\gamma} \right] + \theta \text{Var} \left[ \iota'_p \hat{\gamma} \right] \right],
\]
(4.55)
with \( B_\beta = E \left[ \hat{\beta} - \beta \right] \) and \( B_\gamma = E [\hat{\gamma} - \gamma] \). This can be used to construct corrected estimators of the long-run coefficients, which are unbiased up to order \( O(T^{-1}) \).

### 4.6 Estimation of asymptotic standard errors

Standard errors of coefficient estimators can be estimated either by using asymptotic variance expressions following from limiting distributions or by applying bootstrap procedures. To save space we will discuss both approaches for the LSDV and FGLSDV
4.6. Estimation of asymptotic standard errors

estimators only, but similar results can be derived for bias corrected, restricted or long-run estimators. Let us define

\[ \Gamma_{W,AW} = \text{plim} \frac{1}{T} W'AW \]
\[ \Gamma_{W,A\Omega W} = \text{plim} \frac{1}{T} W'\Omega W \]
\[ \Gamma_{W,AW}^* = \text{plim} \frac{1}{T} W'\Omega^{-1}W. \]

From the usual asymptotic reasoning it follows that

\[ \sqrt{T} \left( \delta_{LSDV} - \delta \right) \xrightarrow{d} N \left[ 0, \text{plim} \frac{1}{T} W'AW \text{plim} \frac{1}{T} W'\Omega W \right], \quad (4.56) \]

and

\[ \sqrt{T} \left( \delta_{FGLSDV} - \delta \right) \xrightarrow{d} N \left[ 0, \text{plim} \frac{1}{T} W'AW \right]. \quad (4.57) \]

Furthermore, assuming

\[ \text{plim} \frac{1}{T} W'\Omega^{-1}W = 0 \]
\[ \text{plim} \frac{1}{T} W'\Omega^{-1}W = 0, \]

one gets

\[ \sqrt{T} \left( \delta_{FGLSDV} - \delta \right) \xrightarrow{d} N \left[ 0, \text{plim} \frac{1}{T} W'AW \right]. \quad (4.58) \]

Regarding restricted estimators or long-run estimators analogous limiting results can be derived. Note that bias corrected estimators have the same limiting behaviour as their uncorrected counterparts.

The limiting distributions above can be used to approximate the asymptotic distributions of the LSDV and FGLSDV estimators, i.e.

\[ \delta_{LSDV} \xrightarrow{T \to \infty} N \left[ \delta, V_{LSDV} \right], \]
\[ \delta_{FGLSDV} \xrightarrow{T \to \infty} N \left[ \delta, V_{FGLSDV} \right], \quad (4.59) \]

with

\[ V_{LSDV} = (W'AW)^{-1}W'\Omega W(W'AW)^{-1}, \]
\[ V_{FGLSDV} = (W'\Omega^{-1}W)^{-1}. \quad (4.60) \]

Hence, using a consistent estimator for \( \Omega \) the asymptotic variance matrix of the LSDV or FGLSDV estimator can be estimated consistently with the expressions in (4.60).
Various authors (Freedman and Peters, 1984; Beck and Katz, 1995) note that for the FGLSDV estimator the expression in (4.60) is very inaccurate in finite samples, i.e. true standard deviations are underestimated dramatically using conventional first-order asymptotic approximations. An alternative approach is using bootstrap procedures to estimate standard errors. We propose the following parametric resampling scheme, i.e.

1. Obtain the estimators \( \hat{\delta} \) and \( \hat{\Sigma} \) (LSDV, FGLSDV);
2. Take a random sample \( \varepsilon^{(b)} \sim N\left(0, \hat{\Sigma} \otimes I_T\right) \);
3. Calculate \( A_y^{(b)} = AW^{(b)} \hat{\delta} + A\varepsilon^{(b)} \);
4. Estimate the model with the resampled data \( (A_y^{(b)}, AW^{(b)}) \) giving the bootstrap estimator \( \hat{\delta}^{(b)} \).

Remark that due to the presence of lagged values of \( y \) in the regressor matrix \( W \) a recursive sampling scheme has to be used. Because the normality assumption is used in the derivation of the bias expressions we employ this assumption here and use a parametric bootstrap procedure contrary to Freedman and Peters (1984).

Considering the bootstrap LSDV and FGLSDV estimators, it can be shown that the following limiting results hold, i.e.

\[
\sqrt{T} \left( \hat{\delta}_{LSDV}^{(b)} - \delta \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma^{-1}_{WAW} \Sigma^{-1}_{WAW} \right),
\]

\[
\sqrt{T} \left( \hat{\delta}_{FGLSDV}^{(b)} - \delta \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma^{-1}_{WAW} \right).
\]

Hence, we can use the bootstrap estimator of the asymptotic variance matrix instead of the expressions in (4.60). Repeating the steps above \( B \) times, \( B \) realisations of \( \hat{\delta}^{(b)} \) are created and

\[
\hat{V}_{LSDV} = \frac{1}{B-1} \sum_{b=1}^{B} \left[ \hat{\delta}_{LSDV}^{(b)} - \frac{1}{B} \sum_{b=1}^{B} \hat{\delta}_{LSDV}^{(b)} \right] \left[ \hat{\delta}_{LSDV}^{(b)} - \frac{1}{B} \sum_{b=1}^{B} \hat{\delta}_{LSDV}^{(b)} \right] ',
\]

\[
\hat{V}_{FGLSDV} = \frac{1}{B-1} \sum_{b=1}^{B} \left[ \hat{\delta}_{FGLSDV}^{(b)} - \frac{1}{B} \sum_{b=1}^{B} \hat{\delta}_{FGLSDV}^{(b)} \right] \left[ \hat{\delta}_{FGLSDV}^{(b)} - \frac{1}{B} \sum_{b=1}^{B} \hat{\delta}_{FGLSDV}^{(b)} \right] ',
\]

are the bootstrap estimators of the asymptotic variance matrices of \( \hat{\delta}_{LSDV} \) and \( \hat{\delta}_{FGLSDV} \). The results in Freedman and Peters (1984) show that for the FGLSDV estimator the bootstrap variance estimator in (4.62) underestimates the true covariance matrix much less than the conventional expression in (4.60). Regarding bias corrected, restricted or long-run estimators similar resampling schemes, but now conditional on those estimators, can be exploited to calculate bootstrap standard errors.
4.7 The demand for money in the European Union

In this section the performance of the various estimators will be illustrated with an empirical application. Money demand in the European Union is analysed by panel data techniques. Earlier empirical studies (Kremers and Lane, 1990; Monticelli and Papi, 1996; Fase and Winder, 1998) use aggregated time series to estimate EU wide money demand functions, but considering the EU countries as a cross-section one can possibly use panel data techniques.

As compared with the aggregate time series approach the use of panel data techniques is different at least in three respects. First of all, it is not necessary to convert money stock and income measures for the different countries into one common currency as is the case for the aggregated time series approach. As long as a suitable conversion measure and functional form are chosen, the individual constants in the panel data model will absorb the effects of this conversion.

Second, as in dynamic panel data models the individual effects are typically filtered out before estimation, the cross-sectional dimension in the panel implies extra data to estimate the same number of unknown parameters. Hence, it seems possible to use fewer time observations than in the aggregate time series approach (see also Section 1.3). To the extent that one is primarily interested in a description of the very near past this is convenient, because especially the short-run parameters of the money demand relationship may not have been constant over the last few decades.

Third, an important specification issue in modeling money demand for a group of countries is interdependence between individual countries. Spillover effects between countries arise as a result of international integration. Also financial integration increases the elasticity of national money demand with respect to the return on foreign assets resulting in increased currency substitution. Hence, apart from domestic variables national money demand is likely to depend on foreign variables. The aggregated time series approach may overcome these specification problems because it internalises interdependencies between individual countries. However, it is clearly not capable of addressing the presence and importance of any spillover effects. In contrast, panel data models allowing for interdependencies between country specific disturbance terms are better suited for this purpose. Also, foreign variables can be incorporated as explanatory variables explicitly.

In this section we shall examine the possibility to estimate standard money demand functions for the European Union using panel data techniques. More in particular, the effectiveness of bias corrected estimators will be examined. Next, we will focus on spillovers between countries.
4.7.1 Standard money demand functions

The data used are from Vlaar and Schuberth (1998), which is an updated version of the data used in Fase and Winder (1998), and contain time series of several variables for Belgium (BE), Denmark (DK), Germany (GE), United Kingdom (UK), Finland (FIN), France (FR), Greece (GR), Ireland (IE), Italy (IT), The Netherlands (NL), Austria (AT), Portugal (PT), Spain (SP) and Sweden (SWE). Together with Luxembourg, which is not included in the data set, these 14 countries currently form the European Union. The time series of the variables have quarterly frequency, are not seasonally adjusted and are collected over the period 1970-1996. The variables in the dataset are M1, M2, M3, real GNP, GNP deflator, short- and long-term interest rates.

For each of the definitions of money stock specification (4.1) has been estimated using

\[ x_{it} = (\ln gnp_{it}, \ln gnp_{i,t-1}, rs_{it}, rs_{i,t-1}, rl_{it}, rl_{i,t-1}, ir_{it}, ir_{i,t-1}, s_{1,t}, s_{2,t}, s_{3,t})' \]  

(4.63)

and where the dependent variable \( y_{it} \) is the logarithm of real money stock, i.e. \( \ln(M1/P)_{it} \), \( \ln(M2/P)_{it} \) or \( \ln(M3/P)_{it} \). The explanatory variables in (4.63) are contemporaneous and one-period lagged values of the logarithm of real income (\( gnp \)), short- (\( rs \)) and long-term (\( rl \)) interest rates and the inflation rate (\( ir \)). To account for seasonal patterns a set of seasonal dummy variables (\( s_1, s_2 \) and \( s_3 \)) is included. Furthermore, lagged values of the dependent variable are incorporated to model autoregressive dynamic adjustments. Separate regressions for the individual countries, which are not reported here, suggest to include one lagged value for the M1 specification and to use two lagged values for the M2 and M3 specifications. Hence, the dimension of the parameter vector \( \delta \) is \( K + 1 \) for the M1 specification and \( K + 2 \) for M2 and M3 with \( K = 11 \).

In order to make valid inference with panel data techniques both parameter constancy through time and over countries should hold to some extent. To avoid parameter variability through time, we have chosen to analyse a relatively short time span, i.e. only the years after the German reunification in 1990 are considered and the sample period is 1991:I-1996:IV. As far as parameter constancy over countries is concerned, it is reasonable to assume that by taking a recent period the problem of parameter heterogeneity across countries is mitigated\(^3\). We are therefore confident to impose common slope vectors, but allow for individual specific effects.

The number of countries\(^4\) analysed is \( N = 13 \). For M1 one period is lost in constructing the lagged value of the dependent variable, so for this specification the first estimation period is 1991:II and \( T = 23 \). The estimation period for M2 and M3 is 1991:III-1996:IV.

\(^3\)We will explore slope heterogeneity across countries in more detail in Chapter 5.

\(^4\)We exclude Greece from the analysis because data on long-term interest rates are not available.
because one extra period is lost in constructing the two-period lagged value of the dependent variable and \( T = 22 \). In all tables bootstrap standard errors are given, because, as argued before, standard analytical variance expressions may be inaccurate here. The number of bootstrap replications used is 100.

For \( M1 \) the estimation results of both the short- and long-run coefficients are in Table 4.1. Regarding the short-run coefficients the bias corrected LSDV and FGLSDV estimators produce in general a higher autoregressive coefficient than the original estimators, while the bias correction in the other coefficients seems to be small. Considering the variance estimators the decrease in variance is apparent when using the FGLSDV estimator compared with the LSDV estimator. The bottom panel of the table with long-run coefficients reasserts these efficiency gains.

Selected estimation results for \( M2 \) and \( M3 \) are in the Tables 4.2 and 4.3. We show results for the FGLSDV estimator only. The first and second columns of these tables show unrestricted estimates, while the final two columns of each table present restricted estimates imposing a long-run income elasticity of one and some exclusion restrictions regarding insignificant dynamics\(^5\). For these money aggregates a two-period lagged value of the dependent variable has been included also, so the bias corrections according to (4.18) and (4.26) have been applied for \( P = 2 \). In general the results show no substantial differences between original and bias corrected estimates. The long-run coefficients presented in the bottom panels of Tables 4.2 and 4.3 are again plausible. Regarding \( M3 \) the long-run effect of inflation has the a priori expected negative sign and is significant.

We compare the general pattern of the long-run estimates with the (semi) elasticities found in earlier research based on the aggregated time series approach. Overviews of these results can be found in Fase and Winder (1993) and Monticelli and Papi (1996). If not restricted to one, the income elasticity is often found close to one for \( M1 \) and larger than one for both \( M2 \) and \( M3 \). In this study, the long-run estimates of the various corrected estimators reflect this pattern in general. As far as the interest rate semi-elasticities are concerned, in general they are close to the estimates found in earlier studies. The only exception is the effect of the long-term interest rate on real \( M1 \) and \( M2 \), which is found to be particularly strong as compared with other studies. Considering the inflation rate except for \( M3 \) no significant long-run effects have been found contrary to earlier studies.

\(^5\)Using a \( F \) test we cannot reject these restrictions (p-values are 0.47, 0.40 and 0.93 for the \( M1, M2 \) and \( M3 \) equations respectively)
## Table 4.1: Estimation results for M1

<table>
<thead>
<tr>
<th></th>
<th>LSDV</th>
<th>LSDV*</th>
<th>FGLSDV</th>
<th>FGLSDV*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Short-run estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\ln(M1/P)_{1,t-1})</td>
<td>0.79</td>
<td>0.87</td>
<td>0.79</td>
<td>0.83</td>
</tr>
<tr>
<td>(\ln gnp_{tt})</td>
<td>0.21</td>
<td>0.18</td>
<td>0.22</td>
<td>0.21</td>
</tr>
<tr>
<td>(\ln gnp_{t,t-1})</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>(rs_{it})</td>
<td>-0.53</td>
<td>-0.47</td>
<td>-0.33</td>
<td>-0.30</td>
</tr>
<tr>
<td>(rs_{i,t-1})</td>
<td>0.29</td>
<td>0.32</td>
<td>0.18</td>
<td>0.20</td>
</tr>
<tr>
<td>(rl_{it})</td>
<td>-0.58</td>
<td>-0.50</td>
<td>-0.59</td>
<td>-0.55</td>
</tr>
<tr>
<td>(rl_{i,t-1})</td>
<td>0.31</td>
<td>0.29</td>
<td>0.00</td>
<td>-0.01</td>
</tr>
<tr>
<td>(ir_{it})</td>
<td>-0.38</td>
<td>-0.35</td>
<td>-0.35</td>
<td>-0.34</td>
</tr>
<tr>
<td>(ir_{i,t-1})</td>
<td>0.34</td>
<td>0.33</td>
<td>0.35</td>
<td>0.35</td>
</tr>
</tbody>
</table>

| **Long-run estimates** |      |       |        |         |
| \(gnp\)               | 1.05 | 1.16  | 1.14   | 1.24    |
| \(rs\)                | -1.13| -1.16 | -0.72  | -0.60   |
| \(rl\)                | -1.28| -1.44 | -2.81  | -3.18   |
| \(ir\)                | -0.22| -0.22 | 0.00   | 0.07    |

\(N=13, T=23, P+K=12\)

Figures in parentheses are bootstrap standard errors
Table 4.2: Estimation results for $M2$

<table>
<thead>
<tr>
<th></th>
<th>$FGLSDV$</th>
<th>$FGLSDV_c$</th>
<th>$FGLSDV,R$</th>
<th>$FGLSDV_{c,R}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Short-run estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(M2/P)_{i,t-1}$</td>
<td>0.89</td>
<td>0.93</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>$\ln(M2/P)_{i,t-2}$</td>
<td>-0.07</td>
<td>-0.07</td>
<td>-0.06</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>$\ln gnp_{i,t}$</td>
<td>0.20</td>
<td>0.19</td>
<td>0.21</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>$\ln gnp_{i,t-1}$</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.05</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>$rs_{i,t}$</td>
<td>-0.06</td>
<td>-0.05</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.12)</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>$rs_{i,t-1}$</td>
<td>0.26</td>
<td>0.24</td>
<td>0.23</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.05)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>$rl_{i,t}$</td>
<td>-0.48</td>
<td>-0.48</td>
<td>-0.57</td>
<td>-0.54</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.17)</td>
<td>(0.10)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>$rl_{i,t-1}$</td>
<td>-0.08</td>
<td>-0.02</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.19)</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>$ir_{i,t}$</td>
<td>-0.38</td>
<td>-0.36</td>
<td>-0.37</td>
<td>-0.37</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.07)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$ir_{i,t-1}$</td>
<td>0.37</td>
<td>0.37</td>
<td>0.38</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.07)</td>
</tr>
<tr>
<td><strong>Long-run estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$gnp$</td>
<td>0.81</td>
<td>0.88</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.18)</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>$rs$</td>
<td>1.16</td>
<td>1.35</td>
<td>1.53</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.40)</td>
<td>(0.28)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>$rl$</td>
<td>-3.15</td>
<td>-3.55</td>
<td>-3.70</td>
<td>-3.94</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.65)</td>
<td>(0.68)</td>
<td>(0.71)</td>
</tr>
<tr>
<td>$ir$</td>
<td>-0.04</td>
<td>0.06</td>
<td>0.05</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.35)</td>
<td>(0.43)</td>
<td>(0.45)</td>
</tr>
</tbody>
</table>

N=13, T=22, P+K=13

Figures in parentheses are bootstrap standard errors
Chapter 4. Bias approximation in large $T$, small $N$ panels

Table 4.3: Estimation results for $M3$

<table>
<thead>
<tr>
<th></th>
<th>$\text{FGLSDV}$</th>
<th>$\text{FGLSDVc}$</th>
<th>$\text{FGLSDV R}$</th>
<th>$\text{FGLSDVc R}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short-run estimates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(M3/P)_{i,t-1}$</td>
<td>0.89</td>
<td>0.92</td>
<td>0.85</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.03)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>$\ln(M3/P)_{i,t-2}$</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>$\ln \text{gnpi}_{it}$</td>
<td>0.18</td>
<td>0.18</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$\ln \text{gnpi}_{i,t-1}$</td>
<td>-0.04</td>
<td>-0.04</td>
<td>-0.03</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$rs_{it}$</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>$rs_{i,t-1}$</td>
<td>0.20</td>
<td>0.20</td>
<td>0.19</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.05)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>$rt_{it}$</td>
<td>-0.37</td>
<td>-0.35</td>
<td>-0.36</td>
<td>-0.34</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.16)</td>
<td>(0.10)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>$rt_{i,t-1}$</td>
<td>0.05</td>
<td>0.06</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.17)</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>$ir_{it}$</td>
<td>-0.41</td>
<td>-0.39</td>
<td>-0.41</td>
<td>-0.40</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$ir_{i,t-1}$</td>
<td>0.31</td>
<td>0.30</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.07)</td>
</tr>
<tr>
<td></td>
<td>Long-run estimates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$gnp$</td>
<td>1.12</td>
<td>1.24</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.18)</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>$rs$</td>
<td>1.40</td>
<td>1.63</td>
<td>1.26</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>(0.45)</td>
<td>(0.47)</td>
<td>(0.24)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>$rl$</td>
<td>-2.46</td>
<td>-2.66</td>
<td>-2.43</td>
<td>-2.50</td>
</tr>
<tr>
<td></td>
<td>(0.63)</td>
<td>(0.66)</td>
<td>(0.48)</td>
<td>(0.49)</td>
</tr>
<tr>
<td>$ir$</td>
<td>-0.75</td>
<td>-0.82</td>
<td>-0.71</td>
<td>-0.72</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.42)</td>
<td>(0.38)</td>
<td>(0.39)</td>
</tr>
</tbody>
</table>

$N=13$, $T=22$, $P+K=13$

Figures in parentheses are bootstrap standard errors.
4.7. The demand for money in the European Union

4.7.2 Spillovers between countries

We will analyse possible spillover effects between countries in two ways. First, following Lane and Poloz (1992) and Angeloni et al. (1994) the covariance matrix of the disturbances is analysed more thoroughly to detect any interdependencies between countries. Second, we will focus on the importance of currency substitution and include both foreign income and foreign interest rates into the national money demand equations (Lane and Poloz, 1992).

The likelihood ratio (LR) test for the null hypothesis of zero cross-correlations, i.e. $\Sigma$ is a diagonal matrix $\Sigma_d$, is

$$LR = T \left( \ln |\hat{\Sigma}_d| - \ln |\hat{\Sigma}| \right),$$

(4.64)

where $\Sigma_d$ is estimated under the restrictions. Under the null hypothesis the statistic $LR$ is asymptotically $\chi^2$ distributed with $\frac{1}{2}N(N - 1)$ degrees of freedom. The values of $LR$ for the money demand equations estimated in the previous subsection are 498.33, 517.71 and 597.06 for $M1$, $M2$ and $M3$ respectively, clearly rejecting the null hypothesis. Hence, we conclude that interdependencies between countries are important.

Next, we include foreign variables in the specifications to quantify the effects of portfolio diversification on domestic money demand. We include weighted averages of foreign income and foreign long-term interest rates of other EU countries, i.e.

$$g_{ni} = \sum_{j=1}^{N} c_j g_{nj} - c_i g_{ni},$$

$$r_{ti} = \sum_{j=1}^{N} w_{jt} r_{jt} - w_{it} r_{it},$$

$$w_{it} = \frac{c_i g_{ni}}{\sum_{j=1}^{N} c_j g_{nj}}$$

where the constant $c_i$ is the nominal exchange rate of country $i$ against the DMark in 1985. The weights are depending on national incomes of the individual countries denoted in a single currency. For each of the definitions of money stock similar specifications as before are estimated using

$$\ln g_{np}, r_s, r_l, (r_{li} - r_l), \ln g_{np}^{\prime}, s_1, s_2, s_3,$$

(4.65)

as explanatory variables.

The estimation results are in Table 4.4. To save space only the implied long-run effects of the corrected FGLSDV estimator are given, the pattern of the short-run estimates and
the differences between estimators are similar to the results in the previous subsection. The estimation results show that both foreign income and foreign long-term interest rates do influence domestic money demand. If foreign long-term interest rates increase relative to local rates domestic money demand drops except for $M_1$, which shows no significant effect. Regarding foreign income a significant long-run effect is found, which is negative for $M_1$ and positive for $M_2$ and $M_3$.

### 4.8 Concluding remarks

In this chapter we have analysed various least squares based estimation procedures for the dynamic panel data model with fixed individual effects and a non-scalar covariance matrix. Because of the typical dimensions of the panel at hand, which is dominated by its time dimension, least squared based methods are used instead of instrumental variables techniques. The latter are commonly used in the typical small $T$, large $N$ panel.

Despite its relatively good performance in this type of panel, least squares estimators are biased in dynamic models and the bias may be substantial in finite samples. Hence, approximation formulae for the bias of the various estimators are developed up to order $O(T^{-1})$ using results of Kiviet (1995, 1999) and related work on bias approximation. The resulting bias expressions are then used to construct bias corrected estimators. From the bias approximations it is seen that operating from a large $N$, large $T$ perspective (like in Chapters 2 and 3) in a panel with $N$ fixed is inappropriate as it would omit terms that cannot be ignored from a large $T$ perspective. Also falsely assuming a scalar covariance matrix will lead to corrected estimators, which still contain a bias term of order $O(T^{-1})$.
4.8. Concluding remarks

This latter result underlines the importance of taking into account the true covariance structure of the disturbances.

With panel data techniques money demand functions for $M_1$, $M_2$ and $M_3$ have been estimated for the EU area as a whole. As far as we know, until now only aggregate time series studies have been undertaken in this area. Regarding estimation bias the empirical results are ambiguous. As is shown by the simulation study in Chapter 2, and to some extent by the estimation results for $M_1$ in this chapter, in some cases bias terms can be substantial in this type of data reasserting the importance of more refined estimation techniques. However, the majority of the estimated bias terms are rather small indicating that small sample bias is not much of a problem here. Regarding efficiency we find that the efficiency gains of exploiting the heteroscedasticity and cross-correlation patterns between countries are sometimes considerable.

The empirical results show that panel data estimators produce plausible long-run effects commonly found in other empirical studies on money demand. As such, the panel data approach is a valuable alternative to the aggregate time series approach. Moreover, exploiting the cross-sectional variation in the panel we are able to identify and estimate any interdependencies between countries. Significant spillover effects between EU countries are found as the cross-correlations in the disturbance covariance matrix are significantly different from zero. Also foreign income and interest rates turn out to have explanatory power for national money demand, asserting the dependence of domestic money demand on foreign developments.
4.A Bias in the ordinary and generalised LSDV estimators

In this appendix the expressions (4.18) and (4.26) will derived. Using the decomposition of \( W \) into irrelevant and relevant stochastic parts, i.e. \( W = \bar{W} + \bar{W} \), and exploiting the normality of \( \varepsilon \) we have

\[
E[\bar{W}'A\varepsilon] = E[\bar{W}'A\varepsilon] + E[\bar{W}'A\varepsilon] = \sum_{p=1}^{P} tr(\Pi_p\Omega)e_p, \tag{4.A.1}
\]

using \( A\bar{W} = \sum_{p=1}^{P} \Pi_p\varepsilon' e_p \). For \( Q^{-1} \) we write

\[
Q^{-1} = E[W'AW] = \bar{W}'A\bar{W} + E[\bar{W}'A\bar{W}] = \bar{W}'A\bar{W} + \sum_{p} \sum_{r} tr(\Pi_p\Pi_r\Omega)e_p e_r. \tag{4.A.2}
\]

Omitting terms with zero moments we also have

\[
E[W'AWQW'A\varepsilon] = \bar{W}'A\bar{W}'QE[\bar{W}'A\varepsilon] + \bar{W}'E[A\bar{W}Q\bar{W}'A\varepsilon] + E[\bar{W}'A\bar{W}Q\bar{W}'A\varepsilon] = \bar{W}'A\bar{W}'Q \sum_{p} tr(\Pi_p\Omega)e_p + \sum_{p} \bar{W}'\Pi_p\Omega A\bar{W} Q e_p + \sum_{p} tr[Q\bar{W}'\Pi_p\Omega A\bar{W}] e_p + \sum_{p} \sum_{r} \sum_{s} q_{rs} [tr(\Pi'_p\Pi_r\Omega)tr(\Pi'_s\Omega) + 2tr(\Omega\Pi'_p\Pi_r\Pi'_s\Omega)] e_p = \sum_{p} tr(\Pi_p\Omega)e_p + \sum_{p} \bar{W}'\Pi_p\Omega A\bar{W} Q e_p + \sum_{p} tr[Q\bar{W}'\Pi_p\Omega A\bar{W}] e_p + 2\sum_{p} \sum_{r} \sum_{s} q_{rs} tr(\Omega\Pi'_p\Pi_r\Pi'_s\Omega)e_p, \tag{4.A.3}
\]

where \( q_{rs} = e'_r Q e_s \) and we have used \( \bar{W}'A\bar{W} = Q^{-1} - \sum_s \sum_r tr(\Pi'_s\Pi_r\Omega)e_s e_r \). Using (4.A.1) and (4.A.3) in (4.16) the result in (4.18) readily follows.

For the bias in the GLSDV estimator the expectations in (4.24) have to be evaluated.
4.B. Bias in the feasible generalised LSDV estimator

Now

\[ E[W^*A^*\epsilon] = E[\tilde{W}'A\Omega^{-1}A\epsilon] + E[\tilde{W}'A\Omega^{-1}A\epsilon] \]
\[ = \sum_{p=1}^{P} E[\epsilon^*\Omega^{-1}\Pi_p\epsilon] \epsilon_p \]
\[ = \sum_{p=1}^{P} tr(\Pi_p)\epsilon_p, \quad (4.A.4) \]

and for \( Q^{-1} \) we write

\[ Q^{-1} = \tilde{W}'A\Omega^{-1}A\tilde{W} + E[\tilde{W}'A\Omega^{-1}A\tilde{W}] \]
\[ = W'A\Omega^{-1}AW + \sum_{p} \sum_{r} tr(\Pi_p\Pi_r)\epsilon_p\epsilon'_r. \quad (4.A.5) \]

Also we write omitting terms with zero moments

\[ E[W^*A^*WQ^*W^*A^*\epsilon] = \tilde{W}'A^*\tilde{W}Q^*E[\tilde{W}'A^*\epsilon] + \tilde{W}'E[A^*\tilde{W}Q^*\tilde{W}^*A^*\epsilon] \]
\[ + E[\tilde{W}'A^*\tilde{W}Q^*\tilde{W}^*A^*\epsilon] + E[\tilde{W}'A^*\tilde{W}Q^*\tilde{W}^*A^*\epsilon] \]
\[ = \tilde{W}'A^*\tilde{W}Q^*\sum_{p} tr(\Pi_p)\epsilon_p \]
\[ + \sum_{p} \tilde{W}'\Omega^{-1}\Pi_pAWQ^*\epsilon_p + \sum_{p} tr(Q^*\tilde{W}'\Omega^{-1}\Pi_pAW) \epsilon_p \]
\[ + \sum_{p} \sum_{r} \sum_{s} q^*_{rs} [tr(\Pi_p\Pi_r)tr(\Pi_s) + 2tr(\Pi_p\Pi_r\Pi_s)] \epsilon_p \]
\[ = \sum_{p} tr(\Pi_p)\epsilon_p + \sum_{p} \tilde{W}'\Omega^{-1}\Pi_pAWQ^*\epsilon_p \]
\[ + \sum_{p} tr(Q^*\tilde{W}'\Omega^{-1}\Pi_pAW) \epsilon_p \]
\[ + 2\sum_{p} \sum_{r} \sum_{s} q^*_{rs} tr(\Pi_p\Pi_s)\epsilon_p, \quad (4.A.6) \]

where \( q^*_{rs} = \epsilon^*_rQ^*e_s \) and we have used \( \tilde{W}'A^*\tilde{W} = Q^{-1} - \sum_{s} \sum_{r} tr(\Pi_s\Pi_r)e_s\epsilon'_r \). Hence, inserting (4.A.4) and (4.A.6) in (4.24) the bias expression in (4.26) follows.

4.B Bias in the feasible generalised LSDV estimator

We give in this appendix a proof of (4.32), i.e. the bias approximations of the GLSDV and FGLSDV estimators are the same up to order \( O(T^{-1}) \). The estimation error of the
Chapter 4. Bias approximation in large $T$, small $N$ panels

FGLSDV estimator (4.30) consists of two factors. The first factor in (4.30) can be expressed as

$$W' A(\bar{\Sigma}^{-1} \otimes I_T) AW = W' A(\Sigma^{-1} \otimes I_T) AW + W' A((\bar{\Sigma}^{-1} - \Sigma^{-1}) \otimes I_T) AW,$$  

(4.3.1)

with

$$W' A(\Sigma^{-1} \otimes I_T) AW = Q^{*-1} + \bar{W}' A(\Sigma^{-1} \otimes I_T) A\bar{W} + \bar{W}' A(\Sigma^{-1} \otimes I_T) A\bar{W}$$

$$+ \left( \bar{W}' A(\Sigma^{-1} \otimes I_T) A\bar{W} - E \left[ \bar{W}' A(\Sigma^{-1} \otimes I_T) A\bar{W} \right] \right)$$

$$= Q^{*-1} + A_1 + A_2 + A_3,$$  

(4.3.2)

where $Q^{*-1} = O(T)$ and $A_1$, $A_2$ and $A_3$ are $O_p(T^{\frac{1}{2}})$. For the second term in (4.3.1) we will need the following

$$\hat{\Sigma} = \hat{E}' A_T \hat{E} / T + O_p(T^{-1}),$$  

(4.3.3)

where $\hat{E} = (\varepsilon_1, ..., \varepsilon_N)$ and $\varepsilon_i$ is the $T \times 1$ disturbance vector belonging to individual $i$. A proof of a similar result is given in Kiviet et al. (1995). Now $\hat{\Sigma} - \Sigma$ can be replaced by $\hat{E}' A_T \hat{E} / T - \Sigma$ without changing the order of the approximation, i.e

$$\hat{E}' A_T \hat{E} / T = \Sigma + O_p(T^{-\frac{1}{2}}),$$  

(4.3.4)

and we may write

$$\hat{\Sigma}^{-1} = \Sigma^{-1} - \Sigma^{-1} \left[ \frac{\hat{E}' A_T \hat{E}}{T} - \Sigma \right] \Sigma^{-1} + O_p(T^{-1}).$$  

(4.3.5)

Exploiting (4.3.5) the second term in (4.3.1) becomes

$$W' A((\bar{\Sigma}^{-1} - \Sigma^{-1}) \otimes I_T) AW = -W' A \left( \Sigma^{-1} \left[ \frac{\hat{E}' A_T \hat{E}}{T} - \Sigma \right] \Sigma^{-1} \otimes I_T \right) AW + O_p(1)$$

$$= -\bar{W}' A \left( \Sigma^{-1} \left[ \frac{\hat{E}' A_T \hat{E}}{T} - \Sigma \right] \Sigma^{-1} \otimes I_T \right) A\bar{W}$$

$$- \bar{W}' A \left( \Sigma^{-1} \left[ \frac{\hat{E}' A_T \hat{E}}{T} - \Sigma \right] \Sigma^{-1} \otimes I_T \right) A\bar{W} + O_p(1)$$

$$= A_4 + A_5 + O_p(1),$$  

(4.3.6)

where $A_4$ and $A_5$ are $O_p(T^{\frac{1}{2}})$. Hence, from (4.3.1), (4.3.2) and (4.3.6) we find

$$W' A(\bar{\Sigma}^{-1} \otimes I_T) AW = Q^{*-1} + \sum_{i=1}^{5} A_i + O_p(1)$$

$$= \left( I + \left( \sum_{i=1}^{5} A_i + O_p(1) \right) Q^* \right) Q^{*-1},$$  

(4.3.7)
and
\[
(W'\Sigma^{-1} \otimes I_T)AW = Q^* \left( I + \sum_{i=1}^{5} A_i + O_p(1) \right) Q^* \]
\[
= Q^* - Q^* \left( \sum_{i=1}^{5} A_i + O_p(1) \right) Q^* + O_p(T^{-2})
\]
\[
= Q^* - Q^* \left( \sum_{i=1}^{5} A_i \right) Q^* + O_p(T^{-2}). \quad (4.B.8)
\]
The second factor in (4.30) can be written as
\[
W'\Sigma^{-1} \otimes I_T)A\varepsilon = W'\Sigma^{-1} \otimes I_T)A\varepsilon + W'\Sigma^{-1} \otimes I_T)A\varepsilon
\]
\[
= \tilde{W}'\Sigma^{-1} \otimes I_T)A\varepsilon + \tilde{W}'\Sigma^{-1} \otimes I_T)A\varepsilon
\]
\[
= \tilde{W}'A \left( \Sigma^{-1} \left[ \frac{E'\hat{A}_T \hat{E}}{T} - \Sigma \right] \otimes I_T \right) A\varepsilon
\]
\[
= \tilde{W}'A \left( \Sigma^{-1} \left[ \frac{E'\hat{A}_T \hat{E}}{T} - \Sigma \right] \otimes I_T \right) A\varepsilon + O_p(T^{-\frac{1}{2}})
\]
where \(A_6, A_7\) are \(O_p(T^{\frac{1}{2}})\) and \(A_8\) and \(A_9\) are \(O_p(1)\). From (4.B.8) and (4.B.9) the estimation error of the FGLSDV estimator is
\[
\hat{\delta}_{FGLSDV} - \delta = Q^* \left( \sum_{i=6}^{9} A_i \right)
\]
\[
- Q^* \left( \sum_{i=1}^{5} A_i \right) Q^* \left( \sum_{i=6}^{9} A_i \right) + O_p(T^{-\frac{1}{2}}). \quad (4.B.10)
\]
using the fact that \(A_8\) and \(A_9\) are \(O_p(1)\) and \(Q^*\) is \(O(T^{-1})\). Evaluating the expectation of the estimation error in (4.B.10) we got many terms. Noting that
\[
E \left[ \hat{\delta}_{GLSDV} - \delta \right] = Q^*E [A_6 + A_7] - Q^*E [(A_1 + A_2 + A_3)Q^* (A_6 + A_7)] + o(T^{-1}),
\]
we have the following
\[
E \left[ \hat{\delta}_{FGLSDV} - \hat{\delta}_{GLSDV} \right] = Q^*E [A_8 + A_9] - Q^*E [(A_4 + A_5)Q^* (A_6 + A_7)] + o(T^{-1}). \quad (4.B.11)
\]
We have to evaluate the expectations of the six remaining terms on the right hand side in (4.B.12). It is easily seen that \(E [A_8] = 0\) and \(E [A_4Q^* A_6] = 0\). We will sketch
Chapter 4. Bias approximation in large $T$, small $N$ panels

the proof that the other four expectations are all of order $O(T^{-1})$, so premultiplied by $Q^*$ their contribution is $o(T^{-1})$. In the following all summations run from 1 to $N$ except the index $p$, which runs from 1 to $P$. Consider first

$$E[A_9] = -E \left[ \hat{W}'A \left( \Sigma^{-1} \left[ \frac{\hat{E}'A_T \hat{E}}{T} - \Sigma \right] \Sigma^{-1} \otimes I_T \right) A \varepsilon \right]$$

$$= -E \left[ \hat{W}'A \left( \Sigma^{-1} \frac{\hat{E}'A_T \hat{E}}{T} \Sigma^{-1} \otimes I_T \right) A \varepsilon \right]$$

$$+ E \left[ \hat{W}'A (\Sigma^{-1} \otimes I_T) A \varepsilon \right]. \quad (4.B.13)$$

Defining $\xi_j$ as the $j$th column of $\Sigma^{-1}$ and $\sigma^{ij}$ as its $ij$th element we write for the first term in (4.B.13)

$$E \left[ \hat{W}'A \left( \Sigma^{-1} \frac{\hat{E}'A_T \hat{E}}{T} \Sigma^{-1} \otimes I_T \right) A \varepsilon \right] = \sum_i \sum_j E \left[ \xi_i \frac{\hat{E}'A_T \hat{E}}{T} \xi_j \hat{W}'_i A_T \varepsilon_j \right]. \quad (4.B.14)$$

Evaluating a particular term in (4.B.14) we got

$$E \left[ \xi_i \frac{\hat{E}'A_T \hat{E}}{T} \xi_j \hat{W}'_i A_T \varepsilon_j \right] = \sum_r \sum_s \sum_r \sum_s \sum_p \frac{\sigma^{ri} \sigma^{sj} \xi_i A_T \varepsilon_i \xi_j \hat{W}'_i A_T \varepsilon_j}{T}$$

$$= \sum_r \sum_s \sum_p \frac{\sigma^{ri} \sigma^{sj} \xi_i A_T \varepsilon_i \xi_j \hat{W}'_i A_T \varepsilon_j}{T}$$

$$= \frac{1}{T} \sum_r \sum_s \sum_p \sigma^{ri} \sigma^{sj} \xi_i A_T \varepsilon_i \xi_j \hat{W}'_i A_T \varepsilon_j$$

$$= \frac{1}{T} \sum_r \sum_s \sum_p \sigma^{ri} \sigma^{sj} \xi_i A_T \varepsilon_i \xi_j \hat{W}'_i A_T \varepsilon_j$$

$$= \frac{1}{T} \sum_r \sum_s \sum_p \sigma^{ri} \sigma^{sj} \xi_i A_T \varepsilon_i \xi_j \hat{W}'_i A_T \varepsilon_j + O(T^{-1})$$

$$= E \left[ \xi_i \xi_j \hat{W}'_i A_T \varepsilon_j \right] + O(T^{-1}). \quad (4.B.15)$$

where $\Pi_{pT} = A_T L_{pT}^T \Gamma_T = O(1)$. Hence, substituting (4.B.15) into (4.B.14)

$$E \left[ \hat{W}'A \left( \Sigma^{-1} \frac{\hat{E}'A_T \hat{E}}{T} \Sigma^{-1} \otimes I_T \right) A \varepsilon \right] = E \left[ \hat{W}'A (\Sigma^{-1} \otimes I_T) A \varepsilon \right] + O(T^{-1}), \quad (4.B.16)$$

and now it is easily seen from (4.B.13) and (4.B.16) that $E[A_9]$ is of order $O(T^{-1})$. 


Next consider

\[ E[A_4 Q^* A_T] = -E \left[ \tilde{W}' A \left( \Sigma^{-1} \left[ \frac{\tilde{E}' A_T \tilde{E}}{T} - \Sigma \right] \Sigma^{-1} \otimes I_T \right) A \tilde{W} Q^* \tilde{W}' A (\Sigma^{-1} \otimes I_T) A \epsilon \right] \]

\[ = -E \left[ \tilde{W}' A \left( \Sigma^{-1} \left[ \frac{\tilde{E}' A_T \tilde{E}}{T} \right] \Sigma^{-1} \otimes I_T \right) A \tilde{W} Q^* \tilde{W}' A (\Sigma^{-1} \otimes I_T) A \epsilon \right] \]

\[ + \tilde{W}' A (\Sigma^{-1} \otimes I_T) A \tilde{W} Q^* E \left[ \tilde{W}' A (\Sigma^{-1} \otimes I_T) A \epsilon \right]. \quad (4.B.17) \]

For the first term in (4.B.17) we write

\[ \tilde{W}' A \left( \Sigma^{-1} \left[ \frac{\tilde{E}' A_T \tilde{E}}{T} \right] \Sigma^{-1} \otimes I_T \right) A \tilde{W} Q^* \tilde{W}' A (\Sigma^{-1} \otimes I_T) A \epsilon \]

\[ = \sum_i \sum_j \sum_k \sum_l \xi_i \frac{\tilde{E}' A_T \tilde{E}}{T} \xi_j \tilde{W}'_i A_T \tilde{W}_j Q^* \sigma^{kl} \tilde{W}'_k A_T \epsilon_l, \quad (4.B.18) \]

with

\[ \xi_i \frac{\tilde{E}' A_T \tilde{E}}{T} \xi_j \tilde{W}'_i A_T \tilde{W}_j Q^* \sigma^{kl} \tilde{W}'_k A_T \epsilon_l \]

\[ = \sum_p \sum_r \sum_{\sigma^{ij}} \sigma^{ij} \xi_i \frac{\tilde{E}' A_T \tilde{E}}{T} \xi_j \tilde{W}'_i A_T \tilde{W}_j Q^* \sigma^{kl} \tilde{W}'_k A_T \epsilon_l \epsilon_p. \quad (4.B.19) \]

The expectation of a particular term in (4.B.19) is

\[ E \left[ \sigma^{ij} \xi_i \frac{\tilde{E}' A_T \tilde{E}}{T} \xi_j \tilde{W}'_i A_T \tilde{W}_j Q^* \sigma^{kl} \tilde{W}'_k A_T \epsilon_l \epsilon_p \right] \]

\[ = \frac{1}{T} \sigma^{ij} \sigma^{kl} \epsilon_l \epsilon_p (tr(A_T)tr(\Pi_T) + 2tr(\Pi_T)) \tilde{W}'_i A_T \tilde{W}_j Q^* \epsilon_p \]

\[ = \sigma^{ij} \sigma^{kl} \epsilon_l \epsilon_p tr(\Pi_T) \tilde{W}'_i A_T \tilde{W}_j Q^* \epsilon_p + O(T^{-1}). \quad (4.B.20) \]

Using this result and follow the same steps back it follows that \( E[A_4 Q^* A_T] \) is of order \( O(T^{-1}) \). In the same fashion the expectations of the remaining two terms \( A_5 Q^* A_6 \) and \( A_8 Q^* A_T \) can be shown to be of order \( O(T^{-1}) \) too.

Having derived the order of magnitude of the expectations of the several terms on the right hand side of (4.B.12) and noting that \( Q^* \) is of order \( O(T^{-1}) \), it is straightforward to see that

\[ E\left[ \hat{\delta}_{FGLSDV} - \delta \right] - E\left[ \hat{\delta}_{GLSDV} - \delta \right] = o(T^{-1}), \quad (4.B.21) \]

Hence, the result in (4.32) readily follows or, in other words, the magnitude of the bias upto order \( O(T^{-1}) \) is the same for the GLSDV and FGLSDV estimators.
Chapter 4. Bias approximation in large T, small N panels

4.C Bias in the restricted feasible generalised LSDV estimator

In this appendix a proof of (4.50) is given, i.e. the bias approximations of the restricted GLSDV and FGLSDV estimators are the same up to order $O(T^{-1})$. We can write for the restricted FGLSDV estimator in (4.48) the following

$$\hat{\delta}_{\text{FGLSDV},R} - \delta = F_{(W^l)}(\hat{\delta}_{\text{FGLSDV}} - \delta), \quad (4.1)$$

where $F_{(W^l)} = I - (W'A\hat{\Omega}^{-1}W)^{-1}R'\left[R(W'A\hat{\Omega}^{-1}W)^{-1}R'\right]^{-1}R$. Defining $C = (W'A\Omega^{-1}W)^{-1}$ we have

$$W'A\hat{\Omega}^{-1}W = W'A\Omega^{-1}W + W'A(\hat{\Omega}^{-1} - \Omega^{-1})W$$
$$\left[I + W'A(\hat{\Omega}^{-1} - \Omega^{-1})WC\right]C^{-1}, \quad (4.2)$$

where $C$ and $W'A(\hat{\Omega}^{-1} - \Omega^{-1})W$ are of order $O_{p}(T^{-1})$ and $O_{p}(T^{1/2})$ respectively. Hence, we may expand

$$\left(W'A\hat{\Omega}^{-1}W\right)^{-1} = C\left[I - W'A(\hat{\Omega}^{-1} - \Omega^{-1})WC\right] + O_{p}(T^{-2}). \quad (4.3)$$

As a result of this, we have

$$R\left(W'A\hat{\Omega}^{-1}W\right)^{-1}R' = RCR' - RCR'W'A(\hat{\Omega}^{-1} - \Omega^{-1})WC$$
$$+ O_{p}(T^{-2}), \quad (4.4)$$

and

$$\left[R\left(W'A\hat{\Omega}^{-1}W\right)^{-1}R'\right]^{-1} = [RCR']^{-1} + [RCR']^{-1} \times$$
$$RCW'A(\hat{\Omega}^{-1} - \Omega^{-1})WC$$
$$+ O_{p}(1). \quad (4.5)$$

Using the expansions in (4.3) and (4.5) we can write for $F_{W^l}$ the following

$$F_{(W^l)} = I - \left[C - CW'A(\hat{\Omega}^{-1} - \Omega^{-1})WC + O_{p}(T^{-2})\right] \times R'(RCR')^{-1}$$
$$+ (RCR')^{-1}RCW'A(\hat{\Omega}^{-1} - \Omega^{-1})WC$$
$$R' (RCR')^{-1} + O_{p}(1)\}R$$
$$= F_{W} + F_{W}CW'A(\hat{\Omega}^{-1} - \Omega^{-1})WC$$
$$R' (RCR')^{-1} R + O_{p}(T^{-1}). \quad (4.6)$$
For the second factor in the estimation error (4.3.1) we can write

$$\hat{\delta}_{FGLSDV} - \delta = CW'\Omega^{-1}\epsilon + O_p(T^{-1})$$  

(4.3.7)

so we have for the estimation error in (4.3.1) the following

$$F_{(W\tilde{\Omega})}(\hat{\delta}_{FGLSDV} - \delta) = F_{W\tilde{\Omega}}(\hat{\delta}_{GLSDV} - \delta) 
+ F_{W\tilde{\Omega}}CW'A(\hat{\Omega}^{-1} - \Omega^{-1})WCR'(RCR')^{-1}R \times 
CW'\Omega^{-1}\epsilon + O_p(T^{-3})$$  

(4.3.8)

so we can write for the bias in the restricted FGLSDV estimator

$$E\left[\hat{\delta}_{FGLSDV,R} - \delta\right] = E\left[\hat{\delta}_{GLSDV,R} - \delta\right]$$

$$+ E[F_{W\tilde{\Omega}}CW'A(\hat{\Omega}^{-1} - \Omega^{-1})WCR'(RCR')^{-1}R \times 
CW'\Omega^{-1}\epsilon] + o(T^{-1}).$$  

(4.3.9)

We will now show that the second term in (4.3.9) is actually $o(T^{-1})$, which establishes the result of (4.50).

From Appendix 4.B we have

$$C = Q^* + O_p(T^{-3/2}),$$

$$W'A(\hat{\Omega}^{-1} - \Omega^{-1})W = A_4 + A_5 + O_p(1),$$

$$W'A\Omega^{-1}\epsilon = A_6 + A_7 + O_p(1),$$

so

$$[RCR']^{-1} = [RQ^*R']^{-1} + O_p(T^{1/2}),$$  

(4.3.10)

and

$$F_{W\tilde{\Omega}} = I - Q^*R'[RQ^*R']^{-1}R + O_p(T^{-1/2}).$$  

(4.3.11)

Hence, we can write

$$F_{W\tilde{\Omega}}CW'A(\hat{\Omega}^{-1} - \Omega^{-1})WCR'(RCR')^{-1}RCW'A\Omega^{-1}\epsilon$$

$$= \left(I - Q^*R'[RQ^*R']^{-1}R\right)Q^*(A_4 + A_5)Q^* \times 
R'(RQ^*R')^{-1}RQ^*(A_6 + A_7) + O_p(T^{-3})$$

$$= F_{Q^*Q^*}(A_4 + A_5)(I + F_{Q^*})Q^*(A_6 + A_7) + O_p(T^{-3}),$$  

(4.3.12)
Chapter 4. Bias approximation in large $T$, small $N$ panels

with $F_{Q^*} = I - Q^*R'[RR^*]^{-1}R$. Using similar derivations as for $E[A_4Q^*A_6]$ in Appendix 4.B it can be derived that

$$E[A_4(I + F_{Q^*})Q^*A_6] = O(T^{-1})$$
$$E[A_4(I + F_{Q^*})Q^*A_7] = O(T^{-1})$$
$$E[A_5(I + F_{Q^*})Q^*A_6] = O(T^{-1})$$
$$E[A_5(I + F_{Q^*})Q^*A_7] = O(T^{-1}),$$

Noting that $F_{Q^*}Q^*$ is $O(T^{-1})$ the second term in (4.9) is $o(T^{-1})$ and we have

$$E\left[\hat{\delta}_{FGLSDV,R} - \delta\right] = E\left[\hat{\delta}_{GLSDV,R} - \delta\right] + o(T^{-1}), \quad (4.13)$$

which establishes the result in (4.50).