

University of Groningen

Measurement issues in political economy

Leertouwer, E.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2002

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Leertouwer, E. (2002). *Measurement issues in political economy*. s.n.

Copyright

Other than for strictly personal use, 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), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 2

Methodology

This thesis deals with applied research in the field of political economy. Before we present the results of this research, its econometric context needs to be discussed. The majority of the models described in the following chapters deal with issues concerning latent variables and measurement error. The first section of this methodological chapter provides a general introduction to latent variables models and the techniques to estimate them, while section 2.2 describes how to deal with underestimation of the coefficient of a latent variable in a regression model. In some of the models that are used in the thesis, dynamics play a role. Section 2.3 describes two tests for panel data models: a test whether or not the model is in fact dynamic, and a test for the presence of a unit root. These tests do not rely on asymptotic results.

2.1 Latent variables

Economic theory tries to describe the relationships between variables using mathematical models. To be able to apply these models empirically, a quantitative measure of these variables has to be available. Some economic variables are clearly defined and straightforward to measure, such as the consumer price level, the total value of exported goods or the number of unemployed people in the labour force of a country. Others, however, are more difficult to capture numerically. For mental constructs such as the independence of central banks, the level of corporatism or the amount of economic freedom of a country, different definitions and quantifications exist. These variables are called *latent* variables, since they are not directly observable. In order to use them in empirical models, we need to find observable related variables, also called *proxies* or *indicators*. This process creates *measurement error*, since the proxies are only approxima-

tions of the true unobservable phenomena. This section describes the basic concepts of latent variables models with measurement error and the techniques that can be used to estimate them. For a thorough discussion of the subject, see Wansbeek and Meijer (2000).

The standard linear multiple regression model is written as

$$y = \Xi\beta + \epsilon, \quad (2.1)$$

with y an observable N -vector, ϵ an unobservable N -vector of random disturbances, β a k -vector of unknown parameters and Ξ an $N \times k$ -matrix containing the regressors. The disturbances are assumed to be independently identically distributed (i.i.d.) with expectation zero and variance σ_ϵ^2 , and the regressors are assumed to be uncorrelated with the disturbances. Now, if there is measurement error in the regressors, the matrix X is observed instead of the unobservable Ξ :

$$X = \Xi + V. \quad (2.2)$$

Here, V is an $N \times k$ -matrix of measurement errors. Its rows are assumed to be i.i.d. with zero expectation and covariance matrix Ω , and are assumed to be uncorrelated with Ξ and ϵ .

Factor analysis The observable regressors X in equation (2.2) can also be assumed to be generated by the following model, known as the *multiple factor analysis* (MFA) model:

$$x_{ni} = \tau_i + \lambda'_i \xi_n + \delta_{ni}. \quad (2.3)$$

In model (2.3), there are k latent variables. The subscript i corresponds to the different observable regressors, and n to the observational units. Since in this thesis the factor analysis model is applied in a macroeconomic context, the observational units n are from now on denoted as countries. Then, x_{ni} denotes indicator i for country n and ξ_n is a k -vector containing the aspects of the unobservable concept (the *factors*) that the indicators are supposed to measure, for country n . The parameter τ_i captures the mean of indicator i , while λ_i is a k -vector of parameters (the *factor loadings*) that capture both the scale of indicator i and the strength of its relation to the factors. Further, δ_{ni} is a random measurement error, with mean zero and variance ψ_{ii} , often called the *unique variance*, and δ_{ni} and δ_{nj} are assumed uncorrelated for $i \neq j$. Both are assumed uncorrelated with the factors ξ_n .

The factor analysis model was originally developed in psychology to model the dependencies among different measures of intelligence (Spearman, 1904). The model is illustrated graphically by way of a path diagram in figure 2.1, for the simple case of one latent variable and three indicators. Drawing a path diagram obeys certain conventions. Circled variables denote latent variables, i.e. hypothetical constructs. Variables in square boxes denote observed variables, such as the different indicators of the latent variable. Variables that are not circled and not in square boxes denote error terms. An arrow denotes a causal dependency in the order indicated.

From (2.3), it is clear that the mean and variance of ξ can be chosen arbitrarily, because a change in its mean or variance can be counteracted by changing the corresponding τ or λ accordingly, without changing the observed variables. Hence, it is customary to let ξ have mean zero and variance one. An introduction to this type of measurement model can be found in Bollen (1989, chapter 6) or Wansbeek and Meijer (2000, chapter 7).

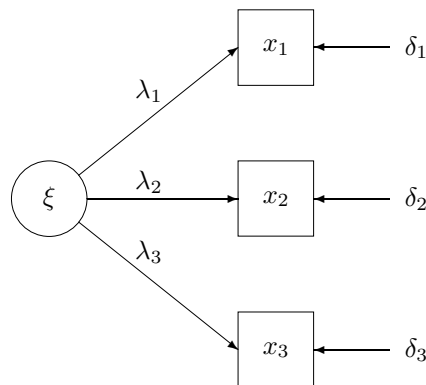


Figure 2.1: Path diagram of a factor analysis model with 1 factor and 3 indicators

Next, we discuss how the unknown parameters in model (2.3) can be estimated. For ease of exposition, we consider the simple factor analysis model with one factor, so $k = 1$. Estimation in MFA is in general a straightforward extension of this case. The parameters of the model, that is, the intercepts τ_i , the factor loadings λ_i , and the unique variances ψ_{ii} , are typically estimated from the means and covariance matrix of the indicators. From (2.3), the assumption that the errors are uncorrelated, and the imposed restriction that the mean of ξ is zero and its variance is one, it follows that the means of the indicators are

$\mu_i \equiv \mathbb{E}(x_i) = \tau_i$. Hence, a consistent estimator of τ_i is given by $\hat{\tau}_i \equiv \bar{x}_i$, the sample mean of the i -th indicator. The covariance of indicator i and indicator j is simply

$$\sigma_{ij} \equiv \mathbb{E}(x_i - \tau_i)(x_j - \tau_j) = \lambda_i \lambda_j, \quad (2.4)$$

where $i \neq j$. The variance of indicator i is

$$\sigma_{ii} \equiv \mathbb{E}(x_i - \tau_i)^2 = \lambda_i^2 + \psi_{ii}. \quad (2.5)$$

If there are at least three indicators, consistent estimators can be obtained by minimizing some sort of discrepancy between the sample variances and covariances and the theoretical variances and covariances as functions of the parameters, as given in (2.4) and (2.5). For example, (2.4) and (2.5) imply

$$\lambda_1 = \left(\frac{\sigma_{21}\sigma_{31}}{\sigma_{32}} \right)^{1/2}$$

and $\psi_{11} = \sigma_{11} - \lambda_1^2$. Hence, consistent estimators of λ_1 and ψ_{11} are given by

$$\hat{\lambda}_1 \equiv \left(\frac{s_{21}s_{31}}{s_{32}} \right)^{1/2}$$

and $\hat{\psi}_{11} = s_{11} - \hat{\lambda}_1^2$, where s_{ij} denotes the sample covariance between indicators i and j (provided that the expressions are nonnegative, cf. Dijkstra, 1992). However, when there are more than three indicators, generally more efficient estimators can be obtained that balance the discrepancies for different covariances optimally. Due to the small number of observational units in the empirical models, maximum likelihood procedures are used to estimate the unknown parameters throughout the rest of this thesis. Wansbeek and Meijer (2000) show that the estimation of the factor loadings and covariances in MFA comes down to solving a system of eigenequations. As will be shown later on, the corresponding eigenvalues are also used to determine the number of factors. The *reliability* of indicator i , denoted as r , is the squared correlation of the indicator and the unobservable concept ξ . As well as for the indicators, the reliability of the factor itself can also be estimated to assess the quality of the result.

In many cases, the scaling of the variables is arbitrary. For instance, there is no straightforward scale for central bank independence. In these cases, the model might be easier to interpret if the variables are rescaled such that they have variance 1. The corresponding solution of the FA model is called the *standardized solution*. The standardized solution is usually equivalent to the model

estimated on (a reparametrization of) the correlation instead of the covariance matrix. In the one-factor model, the factor loadings of the standardized solution are simply the correlations of the indicators with the factor.

Now that the unknown parameters in (2.3) have been estimated, we would like to obtain values for the latent variable that we can use in empirical applications. These values are called *factor scores*, and the predictor $\hat{\xi}_n$ of ξ_n that has minimum mean squared error, under the restriction that the predictor is unbiased, is

$$\hat{\xi}_n = \lambda' \Sigma^{-1} (x - \tau)$$

for $k = 1$. This predictor is called the *Bartlett predictor*. Its expression for the MFA model follows analogously, see Wansbeek and Meijer (2000). If we rewrite the one-factor model as

$$\frac{x_{ni} - \tau_i}{\lambda_i} = \xi_n + \frac{1}{\lambda_i} \delta_{ni} = \xi_n + u_n,$$

we can write the Bartlett predictor as

$$\hat{\xi}_n = \lambda' \Sigma^{-1} \lambda (\xi_n + u_n) = \gamma \xi_n + v_n,$$

where $\gamma = \lambda' \Sigma^{-1} \lambda$ and $v_n = \lambda' \Sigma^{-1} \lambda u_n$.

Since the factor ξ is assumed to have expectation 0 and variance 1, and ξ and v_n are assumed to be uncorrelated, the variance of $\hat{\xi}_n$ is

$$\text{var}(\hat{\xi}_n) = \text{E}(\hat{\xi}_n^2) = \gamma^2 + \text{var}(v_n),$$

and consequently the estimated reliability \hat{r} is

$$\hat{r} = \frac{\gamma^2}{\text{var}(\hat{\xi}_n)} = \frac{\gamma^2}{\gamma^2 + \text{var}(v_n)}. \quad (2.6)$$

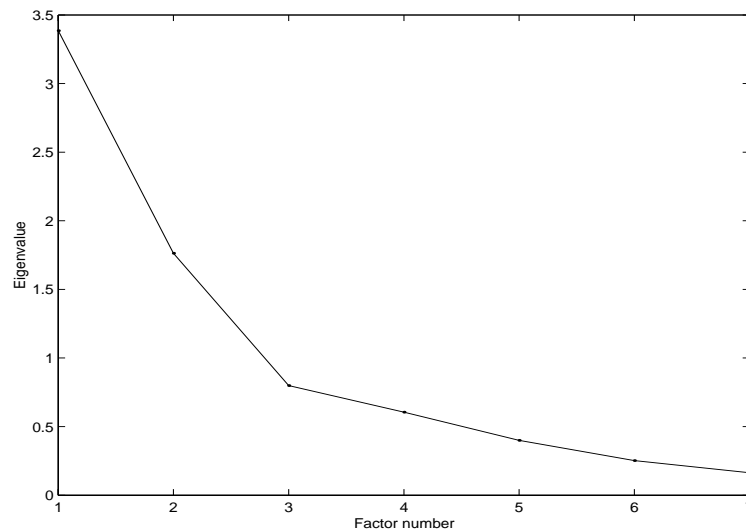
Finally, since the original factor ξ is assumed to have expectation 0 and variance 1, we would like the same to hold for our estimated counterpart. This means that $\hat{\xi}$ needs to be adjusted by a factor γ such that

$$\tilde{\xi}_n = \frac{1}{\gamma} \hat{\xi}_n = \xi_n + \frac{1}{\gamma} v_n = \xi_n + \tilde{v}_n. \quad (2.7)$$

The variance of \tilde{v} in (2.7) is called the *measurement error variance*, and is denoted as ϕ .

Related issues The MFA model can be used in two different situations. In exploratory factor analysis (EFA), the analysis is purely exploratory and does not use subject matter theory to restrict the model parameters. Confirmatory factor analysis (CFA), however, takes subject matter theory as its point of departure and uses restrictions on the factor loadings and the covariance matrix of the factors. The EFA model, by contrast, is estimated on the basis of the correlation matrix. If we want to perform EFA, we have to choose the number of factors k to be used in the analysis. There are two commonly applied rules to do this, which are both based on the eigenvalues of the correlation matrix of the standardized variables. The first rule, which is also called the *Kaiser rule*, says that relevant factors correspond to eigenvalues larger than 1. The second rule uses the *scree plot*, which plots the number of factors against the eigenvalues. It states that the number of factors to select is the number of eigenvalues before a ‘kink’ that is often found in the scree plot. An example is shown in figure 2.2

Figure 2.2: Scree plot of a factor analysis



There are two eigenvalues larger than 1 in figure 2.2. Also, the number of factors before the kink in the plot is two, so using either rule leads to the same conclusion: we should select two factors in this case. If the two rules have different outcomes, the number of factors corresponding to the solution that is interpreted more easily is selected.

It may happen that the solution of a MFA is difficult to interpret. In that case, we can make use of the fact that the matrix of factor loadings is not identified: it can be multiplied with any orthonormal matrix without affecting the distri-

bution of the indicators. Put differently, the factor loadings matrix is open to *rotation*, yielding a solution that may be easier to interpret because the matrix has a simpler structure. Ideally, each indicator is correlated with as few factors as possible. Several rotation methods can be applied: the one most frequently used in practice is varimax rotation. Another rotation method that is used in this thesis is direct oblimin rotation, which minimizes the correlation between columns of the factor loadings matrix. For a more detailed discussion of rotation and rotation methods see Wansbeek and Meijer (2000, pp. 167–169).

Finally, there are a number of criteria available to judge the fit of a factor analysis model. Of these criteria, we use the following two in this thesis: the χ^2 -statistic, which compares the proposed restricted model to an unrestricted alternative (the *saturated model*), and the so-called *comparative fit index* (CFI) which considers the proposed model compared to a highly restrictive *null model*. In the null model, all factor loadings are restricted to zero. The CFI is an assessment of model fit that is especially valuable in small samples, which are encountered in most of the macroeconomic models in this thesis. An elaborate discussion of these and other model fit measures is found in chapter 10 of Wansbeek and Meijer.

Principal components analysis As we have mentioned above, the factor analysis model imposes a specific structure on the covariance matrix, implying assumptions that may not be satisfied in practical applications. As an alternative, one may drop the assumptions and try to find a Ξ and λ in (2.1) so that the resulting errors are small in some sense or other. This is the idea behind a data analysis method called *principal components analysis* (PCA). Here, the columns of the matrix Ξ are the principal components of y , which are uniquely obtained as a linear combination of the observed variables. If the number of indicators is not too small, the solutions of PCA and FA are quite similar. An application of PCA is found in chapter 7, where the measurement of economic freedom is discussed. More on PCA is found in Wansbeek and Meijer (2000).

2.2 Measurement error in a single regressor: the CALS estimator

When a latent variable is used as a regressor in a regression model, we are confronted with the problem that it can only be imperfectly measured. It is well known that a neglect of this problem leads to inconsistent estimation results. In particular, the coefficient of the latent variable will be underestimated. This phe-

nomenon is quite widespread (as discussed earlier, many economic variables are theoretical ideals that allow no direct measurement) and is equally widely ignored. It may explain the experience, well known to applied researchers, that regression coefficient estimates are often disappointingly low.

If the variance of the differences between true and observed values of the regressor is known, it is possible to adapt the results and come up with consistent estimates. Meijer and Wansbeek (2000) and Wansbeek and Meijer (2000, section 5.2) describe how to do it, using the general approach due to Kapteyn and Wansbeek (1984) known as the consistent adjusted least squares (CALs) estimator.

Although the principle is easy to understand, the attention paid to it is modest. The main reason for this is that the condition of a known variance is usually not met in practice. However, if we know the reliability of the construct that has been derived from the unobservable phenomenon by factor analysis techniques, the adaptation to adjust the underestimation is carried out easily. Here, knowing the reliability should be interpreted in the sense of a consistent estimator being available. In the previous section, we have seen that the reliability of the factor can be estimated from the factor model. Consequently, CALs estimators can be computed.

Summarizing, our aim is to provide consistent estimation results and get rid of the underestimation of the coefficient of the latent variable. Then, the term we are interested in is the measurement error variance ϕ . From equation (2.7) in section 2.1 it follows that

$$\phi = \text{var}(\tilde{v}_n) = \frac{\text{var}(v_n)}{\gamma^2} = \frac{1}{r} - 1.$$

Since an estimate of r is available, the measurement error variance ϕ follows.

To correct for the downward bias, we define the matrix $A_N \equiv \frac{1}{N}X'X$ and scalar $\alpha \equiv e_1'A_N^{-1}e_1$. Here, X is the $N \times k$ matrix of observed variables of equation (2.2) and e_1 is an N -vector with first element equal to 1 and zeros otherwise. Further, let $\theta = \frac{1}{1-\phi\alpha}$. Then, if b_1 is the estimated coefficient of the latent variable using OLS, the CALs estimator that corrects for underestimation in case of measurement error is

$$\hat{\beta}_1 = \theta b_1.$$

Now, we can also derive the t -value that corresponds to a measurement error of size ϕ . The t -statistic in case of no measurement error is

$$t_0 = \frac{b_1 \sqrt{N}}{\sqrt{s_\epsilon^2 \alpha}}.$$

Then, the t -statistic corresponding to a measurement error of size ϕ is given by

$$t_\phi = \frac{t_0}{\sqrt{1 + \frac{2}{N}(\theta - 1)^2 t_0^2}}.$$

Meijer and Wansbeek show that, while the coefficient estimate increases when the CALS estimator is used, the t -value decreases. Consequently, for large values of the measurement error variance ϕ , the coefficient estimate becomes insignificant.

2.3 An exact test for dynamic panel data models

In this section, we shortly step away from models containing latent variables to discuss econometric models that include dynamics. In particular, we are interested in panel data models that contain a lagged dependent variable among the regressors. We use the simple model specification

$$y = \gamma y_{-1} + X\beta + Z\alpha + u, \quad (2.8)$$

where y and y_{-1} are $NT \times 1$ -vectors. For ease of exposition, we let the regressor matrix X be an $NT \times 1$ -vector, with β its corresponding unknown parameter. The procedure outlined below can easily be generalized to the case of k regressors. The term $Z\alpha$ denotes the individual specific effects, with Z an $NT \times N$ -matrix defined as $Z = \iota_T \otimes I_N$ and α an $N \times 1$ -parameter vector. Here, ι_T is a vector consisting of T ones, and I_N is the identity matrix of order N . Finally, u is an $NT \times 1$ -disturbance vector with variance matrix $\sigma^2 I$. Then, u/σ has expectation 0 and variance 1 and, under normality, does not depend on unknown parameters. Our aim is to derive an exact test for different values of γ . To do this, we use a test described in Van den Doel and Kiviet (1995) that does not rely on asymptotics.

Since the inclusion of a lagged dependent variable complicates the estimation procedure of a panel data model, it is useful to test first whether the model is dynamic. In this case, the null hypothesis is $H_0 : \gamma = 0$. Next, we adapt this test to test for unit roots, with null hypothesis $H_0 : \gamma = 1$. In chapter 4, these tests are applied in an empirical, macroeconomic context.

2.3.1 Testing for dynamics

To test for $\gamma = 0$ in model (2.8), we might use the fixed effects estimator of γ ,

$$\hat{\gamma}_{FE} = \frac{y'_{-1} M_{XZ} y}{y'_{-1} M_{XZ} y_{-1}},$$

where the matrix $M_{XZ} = I_{NT} - (X, Z)\{(X, Z)'(X, Z)\}^{-1}(X, Z)'$ is the projector orthogonal to (X, Z) . However, basing a test on $\hat{\gamma}_{FE}$ has the disadvantage that its distribution depends on unknown parameters. If we estimate γ from a different regression in an augmented model, this dependence can be avoided and the exact distribution of the resulting estimator of γ can be computed. In order to do so, define lag operators

$$B_0 = \begin{bmatrix} 0 & 0 \\ I_{T-1} & 0 \end{bmatrix},$$

and

$$B = B_0 \otimes I_N.$$

Our aim is to estimate γ in the augmented regression model

$$\begin{aligned} y &= \gamma y_{-1} + X\beta + BX\beta^* + Z\alpha + BZ\alpha^* + u \\ &= \gamma y_{-1} + W\theta + u, \end{aligned}$$

where $W = (X, BX, Z, BZ)$ and θ contains all parameters except γ . The estimator for γ in this model is

$$\hat{\gamma}_{FE}^* = \frac{y'_{-1} M_W y}{y'_{-1} M_W y_{-1}}, \quad (2.9)$$

which can also be written as

$$\hat{\gamma}_{FE}^* = \frac{u' B' M_W u}{u' B' M_W B u}. \quad (2.10)$$

Under $H_0 : \gamma = 0$, the distribution of this estimator does no longer depend on unknown parameters. Moreover, it is easily computed, since we can draw a large number of normally distributed vectors u and simulate the distribution of $\hat{\gamma}_{FE}^*$ using (2.10).

The augmented regression comes down to adding BX , which contains the regressors X lagged by one period, to the regression, along with BZ . Since

$BZ = Z - e_1 \otimes I_N$, adding this term means an additional transformation to the original data: adding BZ is equivalent to ignoring the first N observations of X , BX and Z . For notational convenience, we still denote the transformed data as X , BX and Z , which are now two vectors of order $N(T - 1) \times 1$ and one matrix of order $N(T - 1) \times N$, respectively. The matrix W is redefined as $W = (X, BX, Z)$.

In order to compute $M_W u$ and $M_W B u$ in (2.10), the matrix $W'W$ needs to be inverted. Since this matrix can become quite large, it is useful to make the computation more efficient by avoiding direct inversion of the matrix. This is done as follows. To compute $M_W u$, let $R = (X, BX)$ and consider

$$u = R\delta + Z\eta + \epsilon = W(\delta', \eta')' + \epsilon.$$

The unknown parameters δ and η are estimated by

$$(\hat{\delta}', \hat{\eta}')' = (W'W)^{-1}W'u,$$

hence the estimate of u is

$$\hat{u} = W(\hat{\delta}', \hat{\eta}')' = W(W'W)^{-1}W'u.$$

Then

$$M_W u = (I_{N(T-1)} - W(W'W)^{-1}W')u = u - \hat{u}.$$

Using the Frisch-Waugh theorem, see Wansbeek and Meijer (2000, p.352), the parameters δ and η can be estimated separately by

$$\hat{\delta} = (R' M_Z R)^{-1} R' M_Z u$$

and

$$\hat{\eta} = (Z'Z)^{-1} Z'(u - R\hat{\delta}).$$

To start with $\hat{\delta}$, note that

$$M_Z = (I_{N(T-1)} - Z(Z'Z)^{-1}Z') = I_{N(T-1)} - \frac{1}{T-1}(\iota_{T-1}\iota'_{T-1} \otimes I_N).$$

To write this in a more convenient form, use $\text{vec}(ABC) = (C' \otimes A)\text{vec } B$, see Wansbeek and Meijer (2000, p.350). If we let $u = \text{vec } U$, where the matrix U is of order $N \times (T - 1)$ with elements u_{it} , then

$$M_Z u = u - \frac{1}{T-1} \text{vec}(U \iota_{T-1} \iota'_{T-1}).$$

The computation of $M_Z R$ follows by analogy:

$$M_Z R = (X, BX) - \frac{1}{T-1} (\iota_{T-1} \iota'_{T-1} \otimes I_N).$$

Let $X = \text{vec } \tilde{X}$ and $BX = \text{vec } \widetilde{BX}$ with \tilde{X} and \widetilde{BX} matrices of order $N \times (T-1)$, then

$$M_Z R = (X, BX) - \frac{1}{T-1} \{ \text{vec}(\tilde{X} \iota_{T-1} \iota'_{T-1}), \text{vec}(\widetilde{BX} \iota_{T-1} \iota'_{T-1}) \},$$

where $\text{vec}(\tilde{X} \iota_{T-1} \iota'_{T-1})$ is simply a vector containing the column sums of \tilde{X} stacked $T-1$ times. Then, $\hat{\delta} = (\hat{\delta}_1, \hat{\delta}_2)'$ is obtained by regressing $M_Z u$ on the columns of $M_Z R$, and

$$\begin{aligned} \hat{\eta} &= (Z'Z)^{-1} Z'(u - R\hat{\delta}) = \frac{1}{T-1} Z'(u - R\hat{\delta}) \\ &= \frac{1}{T-1} (\iota'_{T-1} \otimes I_N)(u - R\hat{\delta}). \end{aligned}$$

The expression for $\hat{\eta}$ can be computed more easily by noting that

$$u - R\hat{\delta} = u - \hat{\delta}_1 X - \hat{\delta}_2 BX = \text{vec}(U - \hat{\delta}_1 \tilde{X} - \hat{\delta}_2 \widetilde{BX})$$

and, using $\text{vec}(ABC) = (C' \otimes A)\text{vec } B$ again,

$$\hat{\eta} = \frac{1}{T-1} \text{vec}\{(U - \hat{\delta}_1 \tilde{X} - \hat{\delta}_2 \widetilde{BX}) \iota_{T-1}\}.$$

Now, all that is left to obtain γ_{FE}^* in (2.10) is the term $M_W B u$. This can be done by analogy to the computation of $M_W u$, using

$$B u = W(\delta'_B, \eta'_B)' + \epsilon,$$

and estimating δ_B and η_B as above, yielding $\widehat{B}u = W(\hat{\delta}'_B, \hat{\eta}'_B)'$. Then

$$\hat{\gamma}_{FE}^* = \frac{u' B'(u - \hat{u})}{u' B'(Bu - \widehat{B}u)},$$

which concludes the computation of the test statistic.

2.3.2 Testing for unit roots

The procedure described for testing dynamics can easily be adjusted to derive an exact test for a unit root. Let

$$\Gamma = \begin{bmatrix} 1 & & 0 \\ \vdots & \ddots & \\ 1 & \cdots & 1 \end{bmatrix} \otimes I_N.$$

Then, the aim is to estimate γ in the augmented regression model

$$\begin{aligned} y &= \gamma y_{-1} + X\beta + B\Gamma X\beta^* + Z\alpha + B\Gamma Z\alpha^* + u \\ &= \gamma y_{-1} + W\theta + u, \end{aligned}$$

where $W = (X, B\Gamma X, Z, B\Gamma Z)$ and θ contains all parameters except γ . The estimator for γ , under the null hypothesis $H_0 : \gamma = 1$, is

$$\hat{\gamma}_{FE}^{**} = \frac{y'_{-1} M_W y}{y'_{-1} M_W y_{-1}}, \quad (2.11)$$

which can also be written as

$$\hat{\gamma}_{FE}^{**} = 1 + \frac{u' \Gamma' B' M_W u}{u' \Gamma' B' M_W B \Gamma u}. \quad (2.12)$$

This estimator has a distribution that does not depend on unknown parameters. It is noted that $B\Gamma = \Gamma - I_{NT}$, so the augmented regression is computed using $W = (X, \Gamma X - X, Z, \Gamma Z - Z)$.

As in the previous subsection, the terms $M_W u$ and $M_W B \Gamma u$ in (2.12) can be computed more efficiently by avoiding direct inversion of the matrix $W'W$, using $R = (X, \Gamma X - X)$ and $Q = (Z, \Gamma Z - Z)$. Further, let

$$u = R\delta + Q\eta + \epsilon = W(\delta', \eta')' + \epsilon,$$

and

$$\Gamma u = R\delta_\Gamma + Q\eta_\Gamma + \epsilon = W(\delta'_\Gamma, \eta'_\Gamma)' + \epsilon,$$

to get $\hat{u} = W(\hat{\delta}', \hat{\eta}')'$ and $\hat{\Gamma}u = W(\hat{\delta}'_\Gamma, \hat{\eta}'_\Gamma)'$. Then equation (2.12) can be written as

$$\hat{\gamma}_{FE}^{**} = 1 + \frac{(u' \Gamma' - u')(u - \hat{u})}{(u' \Gamma' - u')(\Gamma u - \hat{\Gamma}u)(u - \hat{u})},$$

which concludes the computation of $\hat{\gamma}_{FE}^{**}$.

