

**SUPPLEMENTARY MATERIAL:**  
**GENERALIZED LEAST SQUARES ESTIMATION**  
**OF PANEL WITH COMMON SHOCKS**

**Paolo Zaffaroni\***

**Imperial College London**

This draft: 3 February 2010

**Abstract**

This paper considers GLS estimation of linear panel models when the innovation and the regressors can both contain a factor structure. A novel feature of this approach is that preliminary estimation of the latent factor structure is not necessary. Under a set of regularity conditions here provided, we establish consistency and asymptotic normality of the feasible GLS estimator as both the cross-section and time series dimensions diverge to infinity. Dependence, both cross-sectional and temporal, of the idiosyncratic innovation is permitted. Our results are presented separately for time regressions with unit-specific coefficients and for cross-section regressions with time-specific coefficients. Primitive conditions of our assumptions are established for Andrews (2005) and Pesaran (2006) regression models, as particular cases of our set up. Monte Carlo experiments corroborate our results.

Key Words: panel, factor model, heterogeneous coefficients, time-varying coefficients, estimation, generalized least squares.

---

\*I thank participants at the 2009 Cowles Foundation Econometrics Conference on "Handling Dependence: Temporal, Cross-sectional, and Spatial" and the 2009 NBER Summer Institute, Working Group on Empirical in Macroeconomics and Finance for comments. All errors are mine. Contact address: Imperial College Business School, Imperial College London, London SW7 2AZ United Kingdom, Telephone: +44 (0)20 7594 9186, email [p.zaffaroni@imperial.ac.uk](mailto:p.zaffaroni@imperial.ac.uk).

# 1 Applications in finance

Bai (2009) discusses in depth several examples of how a panel with constant regression coefficients and a common factor structure in both regressors and innovations can arise in finance, micro- and macro-econometrics. These examples are also valid here since panel models with constant regression coefficients represent a special case of the class of models considered in this paper. We now illustrate two examples taken from finance which specifically require heterogeneous (asset-specific) regression coefficients.

## 1.1 Capital asset pricing model

The traditional capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) implies that the  $i$ th asset return (in excess of the risk free rate) satisfies the linear factor model:

$$r_{it} - r_{ft-1} = \beta_i(r_{m,t} - r_{f,t-1}) + \varepsilon_{it}, \quad (1)$$

when risk-averse investors with a quadratic utility function trade  $N$  assets in a competitive market without transaction costs. Here  $r_{it}$  and  $r_{mt}$  are the return of asset  $i = 1, \dots, N$  and of the market portfolio, respectively,  $r_{ft-1}$  is the risk free rate and  $\beta_i = cov(r_{it}, r_{mt})/var(r_{mt})$ . The idiosyncratic innovation  $\varepsilon_{it}$ , orthogonal to  $r_{mt}$ , is typically assumed uncorrelated across asset or, at most, with a limited cross-correlation (see Chamberlain (1983)). Estimation and testing procedures for the CAPM are reviewed in detail in Campbell, Lo, and MacKinlay (1997, Chapter 5). It has been pointed out that inference concerning the CAPM could be seriously flawed since the market portfolio return  $r_{m,t}$  is unobserved (see Roll (1977)) and, in fact, proxies  $\hat{r}_{m,t}$  have to be necessarily considered. The sensitivity of the empirical results of using proxies has been investigated (see Stambaugh (1982), Shanken (1987) among others). Here we note that, as a consequence of using proxies, the model suitable to be considered for empirical analysis would then be

$$y_{it} = \beta_i x_{it} + u_{it}, \quad (2)$$

setting

$$\begin{aligned} y_{it} &= r_{it} - r_{ft-1}, \\ x_{it} &= x_t = \hat{r}_{m,t} - r_{f,t-1}, \\ u_{it} &= \beta_i(\hat{r}_{m,t} - r_{m,t}) + \varepsilon_{it} = \beta_i f_t + \varepsilon_{it}, \end{aligned}$$

where  $f_t = \hat{r}_{m,t} - r_{m,t}$  is a latent common factor since  $r_{m,t}$  is unobserved. Note that  $f_t$  and  $x_t$  will be, in general, cross-correlated. Model (2) represents a panel model with a common factor in both the regressors and the innovation.

Multi factor generalizations of (1), as for instance required by the arbitrage pricing theory of Ross (1976), whereby

$$r_{it} - r_{f,t-1} = \beta_{1i}z_{1t} + \beta_{2i}z_{2t} + \dots + \beta_{ki}z_{kt} + \varepsilon_{it}, \quad (3)$$

where  $z_{1t}, z_{2t}, \dots, z_{kt}$  represent  $k$  factors, could also be expressed as a panel model with common factors in the regressors and innovations, as (2), whenever one or more of the  $z_{jt}$  are unobserved and proxies are used. Some of the factors will be portfolio excess returns, such as  $r_{m,t} - r_{f,t-1}$ , other can be macroeconomic variables or sentiment indicators. For instance, inflation expectation is often considered to be an important factor although, being unobserved, would also give rise to a factor component in the innovation. Typically, each factor can be split into a predictable  $\lambda_{jt-1}$  and non-predictable component  $\tilde{z}_{jt}$  yielding  $z_{jt} = \lambda_{jt-1} + \tilde{z}_{jt}$ . Although this will give rise to a factor component in the innovation, it does not lead to the same problem associated with the use of proxies since  $\lambda_{jt-1}$  and  $\tilde{z}_{jt}$  are mutually orthogonal by construction.

## 1.2 Affine term structure models

In general terms, an affine term structure model with  $k$  state variables implies

$$p_{it} = \gamma_{0i} + \gamma'_i \mathbf{z}_t, \quad (4)$$

where  $p_{it}$  is the log price of a zero-coupon bond with maturity  $i$  traded in period  $t$  and  $\mathbf{z}_t = (z_{1t} \dots z_{kt})'$  is the vector of unobserved state variables with factor loadings  $\gamma_i$ . Model (4) is implied, under suitable simplifying conditions (see Campbell (1987, assumptions a) and b) in the appendix), by the intertemporal CAPM model of Merton (1973). It also follows that the expected excess return on long bonds over the short interest rate satisfies

$$E_t(r_{i,t+1} - y_{1t}) = \alpha_i + \delta'_i \mathbf{z}_t, \quad (5)$$

where  $r_{i,t+1} = p_{i-1,t+1} - p_{it}$  is the return on a  $i$ -period bond,  $y_{1t} = -p_{1t}$  is the short term interest rate and  $E_t(\cdot)$  denotes the expectation conditional on the all available information up to period  $t$ . Since the  $z_{jt}$  are unobserved,

proxies need to be used when fitting (5) to the data. Campbell (1987) and Stambaugh (1988), among others, assume that there is a vector of observed variables  $\mathbf{x}_t$  such that

$$E(z_{jt} | \mathbf{x}_t) = \theta'_j \mathbf{x}_t, j = 1, \dots, k. \quad (6)$$

In terms of choices for  $\mathbf{x}_t$ , the forward premium  $f_{it} = p_{i-1t} - p_{it} + p_{1t}$  and interest rate spreads  $-(j^{-1}p_{jt} - i^{-1}p_{it})$  have often been considered a natural predictors (see Campbell (1987) and Stambaugh (1988) among others). If (6) is assumed to hold exactly in population, the fitted model becomes, setting  $y_{it} = r_{i,t} - y_{1t-1}$ ,

$$y_{it+1} = \alpha_i + \beta'_i \mathbf{x}_t + \varepsilon_{it+1},$$

with the restriction  $\beta_i = \sum_{j=1}^k \delta_{ji} \theta_j$ , where the measurement error innovation  $\varepsilon_{it+1}$  satisfies  $E(\varepsilon_{it+1} | \mathbf{x}_t) = 0$ . However, if (6) is considered as an approximation or, alternatively, if the  $\mathbf{x}_t$  are affected by measurement errors (see Stambaugh (1988, section 3 and (17))), the empirical model is instead

$$y_{it+1} = \alpha_i + \beta'_i \mathbf{x}_t + u_{it+1},$$

where the innovation  $u_{it} = (\delta'_i \mathbf{z}_t - \beta'_i \mathbf{x}_t) + \varepsilon_{it+1}$  has a factor structure.

## 2 Mathematical Appendix

For random matrices  $\mathbf{A}$  non-singular of dimension  $m_1 \times m_1$ ,  $\mathbf{B}$  of dimension  $m_1 \times m_2$ ,  $\mathbf{C}$  non-singular of dimension  $m_2 \times m_2$ ,  $\mathbf{D}$  of dimension  $m_1 \times m_3$ , with  $m_1 \geq m_2$  and any  $m_3$ , we present the well-known Sherman-Morrison-Woodbury formula, followed by two lemmas. In particular, the proof of Lemma 1 is basically reproducing the proof of Lemma B in Pesaran and Zafaroni (2009) and it is here repeated for easy reference. Note that throughout the paper we will refer to the lemmas without reference to the matrixes  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$  when there is no risk of ambiguity.

**Sherman-Morrison-Woodbury formula.**

$$(\mathbf{BCB}' + \mathbf{A})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{A}^{-1} \text{ a.s.}$$

**Lemma 1**( $\mathbf{A}, \mathbf{B}, \mathbf{C}, m_1$ ). *Set*

$$\mathbf{E} = \mathbf{BCB}' + \mathbf{A} \text{ a.s.}$$

Let  $\mathbf{G}$  a random positive definitive matrix such that as  $m_1 \rightarrow \infty$ :

$$\frac{\mathbf{B}'\mathbf{A}^{-1}\mathbf{B}}{m_1} \rightarrow_{a.s.} \mathbf{G} \text{ non-singular.} \quad (7)$$

Then

$$\mathbf{E}^{-1}\mathbf{B} = \mathbf{A}^{-1}\mathbf{B}\left(\frac{\mathbf{C}^{-1}}{m_1} + \frac{\mathbf{B}'\mathbf{A}^{-1}\mathbf{B}}{m_1}\right)^{-1}\frac{\mathbf{C}^{-1}}{m_1} \quad (8)$$

and, denoting by  $\mathbf{e}_{m_1}^{(i)}$  the  $i$ -th column of the identity matrix  $\mathbf{I}_{m_1}$ , then for any  $1 \leq i \leq m_1$ :

$$\mathbf{e}_{m_1}^{(i)'}\mathbf{E}^{-1}\mathbf{b}^{(j)} \rightarrow_p 0, \quad 1 \leq j \leq m_2, \quad \text{as } m_1 \rightarrow \infty, \quad (9)$$

where  $\mathbf{b}^{(j)} = \mathbf{B}\mathbf{e}_{m_2}^{(j)}$  is the  $j$ th column of  $\mathbf{B}$ .

When (7) and

$$\frac{\mathbf{B}'\mathbf{A}^{-1}\mathbf{A}^{-1}\mathbf{B}}{m_1} \rightarrow_{a.s.} \mathbf{L} \geq 0, \quad (10)$$

for some a.s. finite random positive semi-definite matrix  $\mathbf{L}$ , then

$$\|\mathbf{E}^{-1}\mathbf{B}\|^2 = O_p(m_1^{-1}) \text{ as } m_1 \rightarrow \infty. \quad (11)$$

**Proof:** This follows precisely the proof of Pesaran and Zaffaroni (2009, Lemma B). We start from the Sherman-Morrison-Woodbury formula, rewritten as

$$\mathbf{E}^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\frac{\mathbf{C}^{-1}}{m_1} + \frac{\mathbf{B}'\mathbf{A}^{-1}\mathbf{B}}{m_1}\right)^{-1}\frac{\mathbf{B}'\mathbf{A}^{-1}}{m_1}. \quad (12)$$

Post-multiplying both sides by  $\mathbf{B}$  and simple manipulations yields (8). Pre-multiplying both sides by  $\mathbf{e}_{m_1}^{(j) \prime}$ , post-multiplying both sides by  $\mathbf{e}_{m_2}^{(j)}$  and then letting  $m_1 \rightarrow \infty$  yields (9).

We deal with (11) more explicitly. Since  $\mathbf{B}\mathbf{e}_{m_2}^{(j)} = \mathbf{b}^{(j)}$

$$\begin{aligned} & (m_1^{-1}\mathbf{C}^{-1} + m_1^{-1}\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}m_1^{-1}\mathbf{B}'\mathbf{A}^{-1}\mathbf{b}^{(j)} - \mathbf{e}_{m_2}^{(j)} \\ &= (m_1^{-1}\mathbf{C}^{-1} + m_1^{-1}\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}m_1^{-1}\mathbf{B}'\mathbf{A}^{-1}\mathbf{b}^{(j)} - (m_1^{-1}\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}m_1^{-1}\mathbf{B}'\mathbf{A}^{-1}\mathbf{b}^{(j)} \\ &= m_1^{-1} \left[ -(m_1^{-1}\mathbf{C}^{-1} + m_1^{-1}\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}^{-1}(m_1^{-1}\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}m_1^{-1}\mathbf{B}'\mathbf{A}^{-1}\mathbf{b}^{(j)} \right] \\ &= m_1^{-1}\mathbf{g}^{(j)}, \end{aligned}$$

where it is easy to see that  $\mathbf{g}^{(j)} \rightarrow_p -\mathbf{G}^{-1}\mathbf{C}^{-1}\mathbf{e}_{m_2}^{(j)}$ . Therefore, substituting the latter expression into (12) yields  $\mathbf{E}^{-1}\mathbf{b}^{(j)} = \mathbf{A}^{-1}\mathbf{b}^{(j)} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{e}_{m_2}^{(j)} + m_1^{-1}\mathbf{g}^{(j)}) = -m_1^{-1}\mathbf{A}^{-1}\mathbf{B}\mathbf{g}^{(j)}$  and thus

$$\begin{aligned} \|\mathbf{E}^{-1}\mathbf{b}^{(j)}\|^2 &= m_1^{-1}\mathbf{g}^{(j) \prime} (m_1^{-1}\mathbf{B}'\mathbf{A}^{-1}\mathbf{A}^{-1}\mathbf{B}) \mathbf{g}^{(j)} \\ &= O_p(m_1^{-1} \mathbf{e}_{m_2}^{(j) \prime} \mathbf{C}^{-1} \mathbf{G}^{-1} \mathbf{L} \mathbf{G}^{-1} \mathbf{C}^{-1} \mathbf{e}_{m_2}^{(j)}). \end{aligned}$$

At last (11) simply follows from

$$\|\mathbf{E}^{-1}\mathbf{B}\|^2 \leq \sum_{j=1}^{m_2} \|\mathbf{E}^{-1}\mathbf{b}^{(j)}\|^2. \quad \square$$

**Lemma 2**( $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, m_1$ ).

Set

$$\mathbf{E} = \mathbf{BCB}' + \mathbf{A} \quad a.s.$$

When (7) and  $\mathbf{D}'\mathbf{A}^{-1}\mathbf{B} = O_p(m_1^{-1} \iota_{m_3} \iota'_{m_2})$  then

$$\mathbf{D}'\mathbf{E}^{-1}\mathbf{B} = O_p(\iota_{m_3} \iota'_{m_2}) \quad \text{as } m_1 \rightarrow \infty. \quad (13)$$

When (7) and  $\mathbf{D}'\mathbf{A}^{-1}\mathbf{B} = O_p(m_1^{-\frac{1}{2}} \iota_{m_3} \iota'_{m_2})$  then

$$\mathbf{D}'\mathbf{E}^{-1}\mathbf{B} = O_p(m_1^{-\frac{1}{2}} \iota_{m_3} \iota'_{m_2}) \quad \text{as } m_1 \rightarrow \infty. \quad (14)$$

**Proof:** By (12)

$$\mathbf{D}'\mathbf{E}^{-1}\mathbf{B} = \mathbf{D}'\mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}^{-1}.$$

and (13) and (14) easily follows along the lines of the proof of Lemma 1.  $\square$

**Proof of Theorem 2.** This follows the proof to Theorem 1.

(i) All the limits below hold as  $N \rightarrow \infty$ . The results follow since

$$\hat{\beta}_t^{OLS} - \beta_{t0} = (\mathbf{X}'_t\mathbf{X}_t)^{-1}\mathbf{X}'_t\mathbf{u}_t,$$

can be written as

$$\hat{\beta}_t^{OLS} - \beta_{t0} - \hat{\gamma}_t^{OLS} = N^{-\frac{1}{2}}\left(\frac{\mathbf{X}'_t\mathbf{X}_t}{N}\right)^{-1}N^{-\frac{1}{2}}\mathbf{X}'_t\varepsilon_t,$$

setting  $\hat{\gamma}_t^{OLS} = (\mathbf{X}'_t\mathbf{X}_t)^{-1}\mathbf{X}'_t\mathbf{B}\mathbf{f}_t$ . The result then easily follows by our assumptions.

(ii) All the limits below hold as  $N \rightarrow \infty$ . The result follows since

$$\begin{aligned} \hat{\beta}_t^{UGLS} - \beta_{t0} &= (\mathbf{X}'_t\mathcal{S}_t^{-1}\mathbf{X}_t)^{-1}\mathbf{X}'_t\mathcal{S}_t^{-1}(\mathbf{B}\mathbf{f}_t + \varepsilon_t) \\ &= \hat{\gamma}_t^{UGLS} + N^{-\frac{1}{2}}\left(\frac{\mathbf{X}'_t\mathcal{S}_t^{-1}\mathbf{X}_t}{N}\right)^{-1}\frac{\mathbf{X}'_t\mathcal{S}_t^{-1}\varepsilon_t}{N^{\frac{1}{2}}}, \end{aligned}$$

where by Lemma 1( $\mathbf{H}_t, \mathbf{B}, \mathcal{F}_t, N$ ) and the SMW formula

$$\hat{\gamma}_t^{UGLS} = (\mathbf{X}'_t\mathcal{S}_t^{-1}\mathbf{X}_t)^{-1}\mathbf{X}'_t\mathbf{H}_t^{-1}\mathbf{B}(\mathcal{F}_t^{-1} + \mathbf{B}'\mathbf{H}_t^{-1}\mathbf{B})^{-1}\mathcal{F}_t^{-1}\mathbf{f}_t = O_p(N^{-1}),$$

together with

$$\left(\frac{\mathbf{X}'_t\mathcal{S}_t^{-1}\mathbf{X}_t}{N}\right)^{-1} = \left(\boldsymbol{\Sigma}_{tX'H^{-1}X} - \boldsymbol{\Sigma}_{tX'H^{-1}B}\boldsymbol{\Sigma}_{tB'H^{-1}B}^{-1}\boldsymbol{\Sigma}'_{tX'H^{-1}B}\right)^{-1}(1 + o_p(1)),$$

and

$$\frac{\mathbf{X}'_t\mathcal{S}_t^{-1}\varepsilon_t}{N^{\frac{1}{2}}} \rightarrow_d \left(\boldsymbol{\Sigma}_{tX'H^{-1}X} - \boldsymbol{\Sigma}_{tX'H^{-1}B}\boldsymbol{\Sigma}_{tB'H^{-1}B}^{-1}\boldsymbol{\Sigma}'_{tX'H^{-1}B}\right)^{\frac{1}{2}}\zeta_{kt}.$$

(iii) All the limits below hold as  $(N, T) \rightarrow \infty$ . We must assume  $T \geq N$  and, with no loss of generality, that there are no observed common factors implying that  $\hat{\boldsymbol{\Sigma}}_N = T^{-1}\sum_{t=1}^T \hat{u}_t\hat{u}'_t$  where

$$\hat{\mathbf{u}}_t = (\mathbf{I}_N - \mathbf{M}_t)\mathbf{u}_t,$$

setting

$$\mathbf{M}_t = \mathbf{X}_t(\mathbf{X}'_t\mathbf{X}_t)^{-1}\mathbf{X}'_t.$$

Then

$$\begin{aligned} \hat{\mathbf{u}}_t\hat{\mathbf{u}}'_t &= \underbrace{\mathbf{B}\mathbf{f}_t\mathbf{f}'_t\mathbf{B}'}_I + \underbrace{(\mathbf{I}_N - \mathbf{M}_t)\varepsilon_t\varepsilon'_t(\mathbf{I}_N - \mathbf{M}_t)}_{II} \\ &+ \underbrace{\mathbf{M}_t\mathbf{B}\mathbf{f}_t\mathbf{f}'_t\mathbf{B}'\mathbf{M}_t}_{III} + \underbrace{(\mathbf{B}\mathbf{f}_t\varepsilon'_t(\mathbf{I}_N - \mathbf{M}_t) + (\mathbf{I}_N - \mathbf{M}_t)\varepsilon_t\mathbf{b}'_t\mathbf{B}')}_{IV} \\ &- \underbrace{((\mathbf{I}_N - \mathbf{M}_t)\varepsilon_t\mathbf{f}'_t\mathbf{B}'\mathbf{M}_t + \mathbf{M}_t\mathbf{B}\mathbf{f}_t\varepsilon'_t(\mathbf{I}_N - \mathbf{M}_t))}_{V} - \underbrace{(\mathbf{M}_t\mathbf{B}\mathbf{f}_t\mathbf{f}'_t\mathbf{B}' + \mathbf{B}\mathbf{f}_t\mathbf{f}'_t\mathbf{B}'\mathbf{M}_t)}_{VI}. \end{aligned}$$

For  $I$

$$\mathbf{B}(T^{-1}\sum_{t=1}^T\mathbf{f}_t\mathbf{f}'_t)\mathbf{B}' = \mathbf{B}\mathcal{F}\mathbf{B}'(1 + o_p(1)).$$

For  $II$

$$\begin{aligned} T^{-1}\sum_{t=1}^T(\mathbf{I}_N - \mathbf{M}_t)\varepsilon_t\varepsilon'_t(\mathbf{I}_N - \mathbf{M}_t) &= T^{-1}\sum_{t=1}^T\varepsilon_t\varepsilon'_t \\ &+ \left( N^{-2}T^{-1}\sum_{t=1}^T\mathbf{X}_t\Sigma_{tX'X}^{-1}\mathbf{X}'_t\varepsilon_t\varepsilon'_t\mathbf{X}_t\Sigma_{tX'X}^{-1}\mathbf{X}'_t \right) (1 + o_p(1)) \\ &- \left( N^{-1}T^{-1}\sum_{t=1}^T(\mathbf{X}_t\Sigma_{tX'X}^{-1}\mathbf{X}'_t\varepsilon_t\varepsilon'_t + \varepsilon_t\varepsilon'_t\mathbf{X}_t\Sigma_{tX'X}^{-1}\mathbf{X}'_t) \right) (1 + o_p(1)), \end{aligned}$$

yielding

$$T^{-1}\sum_{t=1}^T(\mathbf{I}_N - \mathbf{M}_t)\varepsilon_t\varepsilon'_t(\mathbf{I}_N - \mathbf{M}_t) = \mathcal{H}_{N+N^{-1}}(\mathcal{A}_{3N} - \mathcal{A}_{4N} - \mathcal{A}'_{4N})(1 + o_p(1)).$$

For  $III$

$$T^{-1}\sum_{t=1}^T\mathbf{M}_t\mathbf{B}\mathbf{f}_t\mathbf{f}'_t\mathbf{B}'\mathbf{M}_t = \mathcal{A}_{1N}(1 + o_p(1)).$$

For  $IV$

$$T^{-\frac{1}{2}}\sum_{t=1}^T\mathbf{B}\mathbf{f}_t\varepsilon'_t = \mathbf{B}\mathcal{C}_{1N}(1 + o_p(1)), \quad (NT)^{-\frac{1}{2}}\sum_{t=1}^T\mathbf{B}\mathbf{f}_t\varepsilon'_t\mathbf{M}_t = \mathbf{B}\mathcal{C}_{3N}(1 + o_p(1)),$$



and combining the above results yields

$$T^{-1} \sum_{t=1}^T \mathbf{B} \mathbf{f}_t \varepsilon_t' (\mathbf{I}_N - \mathbf{M}_t) = T^{-\frac{1}{2}} \mathbf{B} \left( \mathcal{C}_{1N} - N^{-\frac{1}{2}} \mathcal{C}_{3N} \right) (1 + o_p(1)).$$

For  $V$  using the same arguments one gets

$$T^{-1} \sum_{t=1}^T \mathbf{M}_t \mathbf{B} \mathbf{f}_t \varepsilon_t' (\mathbf{I}_N - \mathbf{M}_t) = T^{-\frac{1}{2}} \left( \mathcal{C}_{2N} - N^{-\frac{1}{2}} \mathcal{C}_{4N} \right) (1 + o_p(1)).$$

For  $VI$

$$T^{-1} \sum_{t=1}^T \mathbf{M}_t \mathbf{B} \mathbf{f}_t \mathbf{f}_t' \mathbf{B}' = \mathcal{A}_{2N} (1 + o_p(1)).$$

Summarizing:

$$\begin{aligned} \hat{\boldsymbol{\Sigma}}_N &= [\mathbf{B} \mathcal{F} \mathbf{B}' + \mathcal{H}_N + \mathcal{A}_{1N} - \mathcal{A}_{2N} - \mathcal{A}'_{2N} + N^{-1} (\mathcal{A}_{3N} - \mathcal{A}_{4N} - \mathcal{A}'_{4N}) + \mathcal{D}_{T,N}] (1 + o_p(1)) \\ &= [\mathbf{B} \mathcal{I}_{1N} \mathbf{B}' + \mathcal{I}_{2N} + \mathcal{D}_{T,N}] (1 + o_p(1)) = [\boldsymbol{\Sigma}_N + \mathcal{D}_{T,N}] (1 + o_p(1)) \end{aligned}$$

setting

$$\mathcal{D}_{T,N} = T^{-\frac{1}{2}} (\mathbf{B} \mathcal{C}_{1N} + \mathcal{C}'_{1N} \mathbf{B}' + \mathcal{C}_{2N} + \mathcal{C}'_{2N}) - (NT)^{-\frac{1}{2}} (\mathbf{B} \mathcal{C}_{3N} + \mathcal{C}'_{3N} \mathbf{B}' + \mathcal{C}_{4N} + \mathcal{C}'_{4N}).$$

The following decomposition holds

$$\begin{aligned} (\hat{\beta}_t^{GLS} - \beta_{t0}) &= ((\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t)^{-1} + (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t)^{-1} (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathcal{D}_{T,N} \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t) (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t)^{-1}) \\ &\quad \times (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} (\mathbf{B} \mathbf{f}_t + \varepsilon_t) + \mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathcal{D}_{T,N} \boldsymbol{\Sigma}_N^{-1} (\mathbf{B} \mathbf{f}_t + \varepsilon_t)) (1 + o_p(1)) \\ &= (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t)^{-1} \mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} (\mathbf{B} \mathbf{f}_t + \varepsilon_t) (1 + o_p(1)) \\ &+ (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t)^{-1} \mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathcal{D}_{T,N} \boldsymbol{\Sigma}_N^{-1} (\mathbf{B} \mathbf{f}_t + \varepsilon_t) (1 + o_p(1)) \\ &+ (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t)^{-1} (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathcal{D}_{T,N} \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t) (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t)^{-1} \mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} (\mathbf{B} \mathbf{f}_t + \varepsilon_t) (1 + o_p(1)) \\ &+ (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t)^{-1} (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathcal{D}_{T,N} \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t) (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t)^{-1} \mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathcal{D}_{T,N} \boldsymbol{\Sigma}_N^{-1} (\mathbf{B} \mathbf{f}_t + \varepsilon_t) (1 + o_p(1)). \end{aligned}$$

Following precisely the same steps of part (ii) but replacing  $\mathbf{S}_t$ ,  $\mathcal{F}_t$ ,  $\mathbf{H}_t$  by  $\boldsymbol{\Sigma}_N$ ,  $\mathcal{I}_{1N}$ ,  $\mathcal{I}_{2N}$  respectively, and using Lemma 1( $\mathcal{I}_{2N}$ ,  $\mathbf{B}$ ,  $\mathcal{I}_{1N}$ ,  $N$ )

$$N^{\frac{1}{2}} (\mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} \mathbf{X}_t)^{-1} \mathbf{X}_t' \boldsymbol{\Sigma}_N^{-1} (\mathbf{B} \mathbf{b}_t + \varepsilon_t) \rightarrow_d (\mathcal{V}_t^{GLS})^{\frac{1}{2}} \zeta_{kt} \text{ as } N \rightarrow \infty$$

where  $\mathcal{V}_t^{GLS} = (\mathcal{M}_t^{GLS})^{-1} \mathcal{N}_t^{GLS} (\mathcal{M}_t^{GLS})^{-1}$  with

$$\begin{aligned} \mathcal{N}_t^{GLS} &= \boldsymbol{\Sigma}_{tX'I_2^{-1} \mathcal{H} I_2^{-1} X}^{-1} + \boldsymbol{\Sigma}_{tX'I_2^{-1} B} \boldsymbol{\Sigma}_{tB'I_2^{-1} B}^{-1} \boldsymbol{\Sigma}_{tB'I_2^{-1} \mathcal{H} I_2^{-1} B} \boldsymbol{\Sigma}_{tB'I_2^{-1} B}^{-1} \boldsymbol{\Sigma}'_{tX'I_2^{-1} B} \\ &- \left( \boldsymbol{\Sigma}_{tX'I_2^{-1} \mathcal{H} I_2^{-1} B} \boldsymbol{\Sigma}_{tB'I_2^{-1} B}^{-1} \boldsymbol{\Sigma}'_{tX'I_2^{-1} B} + \boldsymbol{\Sigma}_{tX'I_2^{-1} B} \boldsymbol{\Sigma}_{tB'I_2^{-1} B}^{-1} \boldsymbol{\Sigma}'_{tX'I_2^{-1} \mathcal{H} I_2^{-1} B} \right) \end{aligned}$$

and

$$\mathcal{M}_t^{GLS} = \Sigma_{tX'T_2^{-1}X} - \Sigma_{tX'T_2^{-1}B} \Sigma_{tB'T_2^{-1}B}^{-1} \Sigma_{tX'T_2^{-1}B}'.$$

For the second and third term after the second equality sign,

$$(\mathbf{X}_t' \Sigma_N^{-1} \mathbf{X}_t)^{-1} \mathbf{X}_t' \Sigma_N^{-1} \mathcal{D}_{T,N} \Sigma_N^{-1} (\mathbf{B} \mathbf{f}_t + \varepsilon_t) = O_p(T^{-\frac{1}{2}} N^{-1} (N^a + N^{b-\frac{1}{2}})),$$

$$(\mathbf{X}_t' \Sigma_N^{-1} \mathbf{X}_t)^{-1} (\mathbf{X}_t' \Sigma_N^{-1} \mathcal{D}_{T,N} \Sigma_N^{-1} \mathbf{X}_t) (\mathbf{X}_t' \Sigma_N^{-1} \mathbf{X}_t)^{-1} \mathbf{X}_t' \Sigma_N^{-1} (\mathbf{B} \mathbf{f}_t + \varepsilon_t) = O_p(T^{-\frac{1}{2}} N^{-\frac{3}{2}} (N^c + N^{d-\frac{1}{2}})),$$

whereas the fourth can be easily seen to be  $O_p\left(N^{-2} T^{-1} (N^c + N^{d-\frac{1}{2}}) (N^a + N^{b-\frac{1}{2}})\right)$ .

□

## References

- ANDREWS, D. (2005): “Cross-section regression with common shocks,” *Econometrica*, 73, 1551–1585.
- BAI, J. (2009): “Panel data models with interactive fixed effects,” *Econometrica*, 77, 1229–1279.
- CAMPBELL, J., A. LO, AND A. MACKINLAY (1997): *The econometrics of financial markets*. Princeton (NJ): Princeton University Press.
- CAMPBELL, J. Y. (1987): “Stock returns and the term structure,” *Journal of Financial Economics*, 18, 373–399.
- CHAMBERLAIN, G. (1983): “Funds, factors, and diversification in arbitrage pricing models,” *Econometrica*, 51, 1305–1323.
- LINTNER, J. (1965): “The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets,” *Review of Economics and Statistics*, 47, 13–37.
- MERTON, R. C. (1973): “An Intertemporal Capital Asset Pricing Model,” *Econometrica*, 41, 867–87.
- PESARAN, M. H. (2006): “Estimation and inference in large heterogeneous panels with a multifactor error structure,” *Econometrica*, 74, 967–1012.
- PESARAN, M. H., AND P. ZAFFARONI (2009): “Optimality and diversifiability of mean variance and arbitrage pricing portfolios,” Cambridge University and Imperial College London, preprint.

- ROLL, R. (1977): “A critique of the asset pricing theory’s tests Part I: On past and potential testability of the theory,” *Journal of Financial Economics*, 4, 129–176.
- ROSS, S. (1976): “The arbitrage theory of capital asset pricing,” *Journal of Economic Theory*, 13, 341–360.
- SHANKEN, J. (1987): “Multivariate proxies and asset pricing relations: Living with the Roll critique,” *Journal of Financial Economics*, 18, 91–110.
- SHARPE, W. (1964): “Capital asset prices: a theory of market equilibrium under conditions of risk,” *Journal of Finance*, 1919, 425–442.
- STAMBAUGH, R. F. (1982): “On the exclusion of assets from tests of the two-parameter model: a sensitivity analysis,” *Journal of Financial Economics*, 10, 237–268.
- STAMBAUGH, R. F. (1988): “The information in forward rates : Implications for models of the term structure,” *Journal of Financial Economics*, 21, 41–70.