First-Differenced Inference for Panel Factor Series

Ekaterina Ipatova

Cass Business School, City University London

Lorenzo Trapani*

Cass Business School, City University London

September 21, 2012

Abstract

We complement existing inferential theory for panel factor models by deriving the asymptotics for the first differences of the estimated factors and common components obtained from a non-stationary panel factor model. As an application, we propose an estimator for the long run variance of the common components.

JEL Classification: C13, C23.

Keywords: Non-stationary panels, common factors, common components, first differences.

* *Corresponding author*: Centre for Econometric Analysis, Faculty of Finance, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ (U.K.). Tel. 00.44.(0)20.70405260 Fax. 00.44.(0)20.70408881, e-mail: L.Trapani@city.ac.uk

1 Introduction

Consider the non-stationary panel factor series

$$X_{it} = \lambda_i' F_t + e_{it}, \tag{1}$$

where $i = 1, ..., n$, $t = 1, ..., T$, F_t is a k-dimensional vector with DGP $F_t = F_{t-1} + \varepsilon_t$, and e_{it} is stationary. Bai (2004) develops the inferential theory for (1) - specifically, for F_t, λ_i, and for the non-stationary common component $C_{it} \equiv \lambda_i' F_t$. Alternatively, one may also consider the stationary, first-differenced model

$$x_{it} = \lambda_i' f_t + u_{it}, \tag{2}$$

where $x_{it} = \Delta X_{it}$ and $f_t = \Delta F_t$. In this case, estimators for λ_i, f_t and $c_{it} \equiv \lambda_i' f_t$ ($\hat{\lambda}_i$, $\hat{f}_t$ and $\hat{c}_{it}$ respectively) are provided by Bai (2003).

This note complements the existing inferential theory on (1) and (2), by studying estimation based on the first difference of the estimator of F_t, say $\hat{F}_t$, computed from (1). Indeed, instead of estimating f_t from (2), one could use $\tilde{f}_t = \hat{F}_t - \hat{F}_{t-1}$. Thence, using the either the estimated λ_i from (1), say $\hat{\lambda}_i$, or estimating λ_i from (2) using $\tilde{f}_t$, one can compute the first differenced estimator of c_{it} as $\tilde{c}_{it} \equiv \tilde{\lambda}_i' \tilde{f}_t$. Estimating f_t and c_{it} is useful for various purposes; in this paper we consider the estimation of the long run covariance matrices (henceforth, LRV) of F_t and C_{it}.

Some results have already been developed by Trapani (2012) in the context of bootstrapping nonstationary factor models. This note completes the inferential theory for the first-differenced estimators, reporting rates of convergence for: $\tilde{f}_t$; for the estimator of λ_i based on $\tilde{f}_t$, say $\tilde{\lambda}_i$; and for a weighted-sum-of-covariances estimator of the LRV of C_{it} based on $\tilde{f}_t$.

2 Results

All results are derived under the same assumptions as in Bai (2003, 2004), omitted for brevity. Henceforth, we define the $r \times r$ rotation matrix $H \equiv \left(\frac{\hat{F}'F}{T^2}\right)\left(\frac{\Lambda'\Lambda}{n}\right)$, where $F = [F_1, ..., F_T]'$ ($\hat{F}$ is defined similarly) and $\Lambda = [\lambda_1, ..., \lambda_n]'$. The number of factors, r, is assumed known.

We firstly report a Lemma containing rates of convergence for $\tilde{f}_t = \hat{F}_t - \hat{F}_{t-1}$.

Lemma 1 *As $(n, T) \to \infty$, it holds that*

$$\tilde{f}_t - H'f_t = O_p\left(\frac{1}{\sqrt{n}}\right) + O_p\left(\frac{1}{T^{3/2}}\right), \tag{3}$$

$$\max_{1 \leq t \leq T}\left\|\tilde{f}_t - H'f_t\right\| = O_p\left(\frac{1}{T}\right) + O_p\left(\sqrt{\frac{T}{n}}\right), \tag{4}$$

$$\frac{1}{T}\sum_{t=1}^{T}\left(\tilde{f}_t - H'f_t\right)u_{it} = O_p\left(\frac{1}{\sqrt{n}}\right) + O_p\left(\frac{1}{T^{3/2}}\right). \tag{5}$$

Under $\frac{n}{T^3} \to 0$, $\sqrt{n}\left(\tilde{f}_t - H'f_t\right) \xrightarrow{d} QN\left(0, \Upsilon_t\right)$, where Q is defined in Theorem 2 in Bai (2004, p. 148) and $\Upsilon_t \equiv \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n} E\left(\lambda_i\lambda_j' u_{it}u_{jt}\right)$.

Lemma 1 states that rates and uniform convergence of $\tilde{f}_t - H'f_t$ are the same as for $\hat{F}_t - H'F_t$ - see Lemma 2 in Bai (2004). This can also be compared with the results in Theorem 2 in Bai (2003), where it is shown that $\hat{f}_t - H_1'f_t = O_p\left(n^{-1/2}\right) + O_p\left(T^{-1}\right)$ - in general, the rotation matrices H and H_1 are different. Therefore, heuristically, $\tilde{f}_t$ should be a better estimator than $\hat{f}_t$ for the space spanned by f_t, especially when T is small. Lemma 1 is a complement, regarding the properties of $\tilde{f}_t$, to Lemma A.1 in Trapani (2012).

We now turn to presenting results on the estimation of the loadings λ_i. To this end, it is possible to use the estimator of λ_i from (1), say $\hat{\lambda}_i$. Bai (2004, p. 148-149) shows that $\hat{\lambda}_i$ is "superconsistent", viz. $\hat{\lambda}_i - H^{-1}\lambda_i = O_p\left(T^{-1}\right)$; also, the rate of convergence does not depend on n. Alternatively, it is possible to estimate loadings as $\tilde{\lambda}_i = \left[\sum_{t=1}^T \tilde{f}_t \tilde{f}_t'\right]^{-1} \left[\sum_{t=1}^T \tilde{f}_t x_{it}\right]$. Let $\Sigma_\varepsilon \equiv E\left(\varepsilon_t \varepsilon_t'\right) = E\left(f_t f_t'\right)$; it holds that:

Proposition 1 *As $(n,T) \to \infty$ it holds that $\tilde{\lambda}_i - H^{-1}\lambda_i = O_p\left(n^{-1}\right) + O_p\left(T^{-1/2}\right)$. Under $\frac{\sqrt{T}}{n} \to 0$, $\sqrt{T}\left(\tilde{\lambda}_i - H^{-1}\lambda_i\right) \xrightarrow{d} N\left(0, V_i\right)$ with $V_i = \left(H'\Sigma_\varepsilon H\right)^{-1}\left(H'\Phi_i H\right)\left(H\Sigma_\varepsilon H'\right)^{-1}$ and $\Phi_i = \lim_{T\to\infty} E\left(f_t f_s' u_{it} u_{is}\right)$.*

Proposition 1 states that the properties of $\tilde{\lambda}_i$ are (apart from the rotation matrix H) the same as in Theorem 2 in Bai (2003), where estimation of λ_i is based on using (2). This can be compared with $\hat{\lambda}_i$, whose convergence rate does not depend on n and it is faster in T.

Based on Lemma 1 and Proposition 1, consider the first-differenced estimator of the common components c_{it}, $\tilde{c}_{it} \equiv \hat{\lambda}_i' \tilde{f}_t = \hat{C}_{it} - \hat{C}_{it-1} = \hat{\lambda}_i'\left(\hat{F}_t - \hat{F}_{t-1}\right)$. By combining the results above, and using Lemma 3 in Bai (2004), we have $\tilde{c}_{it} - c_{it} = \hat{\lambda}_i'\tilde{f}_t - \lambda_i' f_t = \left(\hat{\lambda}_i - H^{-1}\lambda_i\right)'\tilde{f}_t + \left(\tilde{f}_t - H'f_t\right)' H^{-1}\lambda_i + \left(\hat{\lambda}_i - H^{-1}\lambda_i\right)'\left(\tilde{f}_t - H'f_t\right) = O_p\left(n^{-1/2}\right) + O_p\left(T^{-1}\right)$. Using Theorem 3 in Bai (2004) on the limiting distribution of $T\left(\hat{\lambda}_i - H^{-1}\lambda_i\right)$, the asymptotic distribution of $\tilde{c}_{it} - c_{it}$ has the same properties as in Theorem 4 in Bai (2004, p. 149).

The results in Lemma 1 and Proposition 1 can be combined in order to estimate the LRV of F_t and C_{it}. Let Σ_F be the LRV of F_t, and define similarly the LRV of C_{it} as Σ_C. A rotation of Σ_F can be estimated as

$$\hat{\Sigma}_F = \hat{\gamma}_0^F + \sum_{j=1}^{h}\left(1 - \frac{j}{h+1}\right)\left(\hat{\gamma}_j^F + \hat{\gamma}_j^{F\prime}\right),$$

where h is a bandwidth parameter and $\hat{\gamma}_j^F \equiv T^{-1}\sum_{t=j+1}^T \tilde{f}_t \tilde{f}_{t-j}'$. Of course, $\hat{\Sigma}_F$ does not estimate Σ_F consistently due to rotational indeterminacy; it can be expected that $\left\|\hat{\Sigma}_F - H'\Sigma_F H\right\| = o_p(1)$. Similarly, Σ_C can be estimated either as $\hat{\Sigma}_C = \hat{\lambda}_i'\hat{\Sigma}_F\hat{\lambda}_i$, or as

$\tilde{\Sigma}_C = \tilde{\lambda}_i' \hat{\Sigma}_F \tilde{\lambda}_i$. By virtue of Proposition 1, $\hat{\Sigma}_C$ should be better, and we focus our attention on it.

Theorem 1 *Assume that $\sum_{j=0}^{\infty} j^s \left|\gamma_j^F\right| < \infty$. It holds that*

$$\left\|\hat{\Sigma}_C - \Sigma_C\right\| = O_p\left(\frac{h}{\sqrt{T}}\right) + O_p\left(\frac{h}{n}\right) + O_p\left(\frac{1}{h}\right). \quad (6)$$

Theorem 1 contains rates of convergence for $\hat{\Sigma}_C$, which is consistent provided that $h \to \infty$ and $h/\min\left\{n, \sqrt{T}\right\} \to 0$. This also gives a selection rule for h; the choice of the bandwidth that maximizes the speed of convergence is $h^* = O\left(\min\left\{T^{1/4}, n^{1/2}\right\}\right)$.

We point out that $\hat{\Sigma}_C$ is not the only possible estimator for Σ_C. One could consider estimating a rotation of Σ_F using $\hat{f}_t$ calculated from (2). Given that H differs depending on whether (1) or (2) is used, in this case it is necessary to employ the estimated loadings from model (2), which have the same properties as $\tilde{\lambda}_i$ in Proposition 1. Based on this, and on Lemma 1, it can be expected that this estimator does not converge as fast as $\hat{\Sigma}_C$. Similarly, it is possible to estimate Σ_C using the x_{it}s directly. Theoretically, this estimator should work, since e_{it} is stationary, although this may introduce some noise in the estimation of Σ_C.

Proofs

Proof of Lemma 1. See the online material.

Proof of Proposition 1. Let $\delta_{nT} \equiv \min\{\sqrt{n}, T\}$. By definition, $\tilde{\lambda}_i - H^{-1}\lambda_i = \left(\sum_{t=1}^T \tilde{f}_t \tilde{f}_t'\right)^{-1} \times \left[\sum_{t=1}^T H' f_t u_{it} + \sum_{t=1}^T \tilde{f}_t' \left(\tilde{f}_t - H' f_t\right)\lambda_i + \sum_{t=1}^T \left(\tilde{f}_t - H' f_t\right) u_{it}\right] = \left(\sum_{t=1}^T \tilde{f}_t \tilde{f}_t'\right)^{-1} (I + II + III)$. Consider the denominator. By Lemma A.1 in Trapani (2012), $\sum_{t=1}^T \left\|\tilde{f}_t - H' f_t\right\|^2 = O_p\left(T\delta_{nT}^{-2}\right)$ and $\sum_{t=1}^T \left(\tilde{f}_t - H' f_t\right)' f_t = O_p\left(\sqrt{T}\delta_{nT}^{-1}\right) + O_p\left(\frac{\sqrt{T}}{n}\right)$. Hence, $\sum_{t=1}^T \tilde{f}_t \tilde{f}_t' = H' \sum_{t=1}^T f_t f_t' H + o_p(T) = O_p(T)$. As regards the numerator, $I = O_p\left(\sqrt{T}\right)$ by a CLT. Using the same arguments as for the denominator, $II = O_p\left(\sqrt{T}\delta_{nT}^{-1}\right) + O_p\left(\frac{\sqrt{T}}{n}\right)$. Hence, $\tilde{\lambda}_i - H^{-1}\lambda_i = O_p\left(T^{-1/2}\right) + O_p\left(n^{-1}\right)$. Finally, $III = O_p\left(n^{-1/2}\right) + O_p\left(T^{-3/2}\right)$ using (5). The limiting distribution follows from noting that, when $\frac{\sqrt{T}}{n} \to 0$, the dominating $O_p\left(T^{-1/2}\right)$ term is $\left(H' \sum_{t=1}^{T} f_t f_t' H\right)^{-1} \left(\sum_{t=1}^{T} H' f_t u_{it}\right)$.

Proof of Theorem 1. We omit H for simplicity when this does not cause ambiguity. We start by showing that $\left\| \hat{\Sigma}_F - H' \Sigma_F H \right\| = O_p\left(\frac{h}{\sqrt{T}}\right) + O_p\left(\frac{h}{n}\right) + O_p\left(\frac{1}{h}\right)$. By definition, $\Sigma_F = \gamma_0^F + \sum_{j=1}^{\infty} \left(\gamma_j^F + \gamma_j^{F\prime}\right)$, whence

$$
\begin{aligned}
\hat{\Sigma}_F - \Sigma_F &= \left(\hat{\gamma}_0^F - \gamma_0^F\right) + \sum_{j=1}^{h} \left(1 - \frac{j}{h+1}\right) \left[\left(\hat{\gamma}_j^F + \hat{\gamma}_j^{F\prime}\right) - \left(\gamma_j^F + \gamma_j^{F\prime}\right)\right] \\
&\quad - \sum_{j=1}^{h} \left(\frac{j}{h+1}\right) \left(\gamma_j^F + \gamma_j^{F\prime}\right) - \sum_{j=h+1}^{\infty} \left(\gamma_j^F + \gamma_j^{F\prime}\right) \\
&= I - II - III.
\end{aligned}
$$

Consider I. We have $\hat{\gamma}_0^F - \gamma_0^F = T^{-1} \sum_{t=j+1}^{T} \tilde{f}_t \tilde{f}_t' - \gamma_0^F = \left(T^{-1} \sum_{t=j+1}^{T} f_t f_t' - \gamma_0^F\right) - T^{-1} \sum_{t=j+1}^{T} \left(\tilde{f}_t - f_t\right) f_t' - T^{-1} \sum_{t=j+1}^{T} f_t \left(\tilde{f}_t - f_t\right)' + T^{-1} \sum_{t=j+1}^{T} \left(\tilde{f}_t - f_t\right) \left(\tilde{f}_t - f_t\right)' = I_a + I_b + I_b' + I_c$. The CLT yields $I_a = O_p\left(T^{-1/2}\right)$; as far as I_b and I_c are concerned, Lemma A.1 in Trapani (2012) entails that they are both $O_p\left(n^{-1}\right) + O_p\left(T^{-2}\right)$. The same holds for $\hat{\gamma}_j^F - \gamma_j^F$; putting all together $I = O_p\left(hT^{-1/2}\right) + O_p\left(hn^{-1}\right)$. Standard arguments yield $II = O\left(h^{-1}\right)$ and $III = o\left(h^{-s}\right)$. The Theorem follows from $\hat{\lambda}_i - H^{-1}\lambda_i = O_p\left(T^{-1}\right)$.

References

Bai, J., 2003. Inferential theory for structural models of large dimensions. Econometrica. 71, 135-171.

Bai, J., 2004. Estimating cross-section common stochastic trends in nonstationary panel data. Journal of Econometrics 122, 137-183

Trapani, L., 2012. On bootstrapping panel factor series. Journal of Econometrics, forthcoming.