Supplementary Appendix to "Asymptotic Analysis of the Squared Estimation Error in Misspecified Factor Models"

Alexei Onatski*

July 7, 2013

Abstract

This note contains proofs of lemmas 2 and 3 in Onatski (2013), Asymptotic Analysis of the Squared Estimation Error in Misspecified Factor Models, where we refer to for definitions and notation. The note also shows that assumption A3 (iii) in Onatski (2013) holds for very wide classes of stationary processes $\{e_{\cdot t}, t \in \mathbb{Z}\}$.

1 Proof of Lemma 2.

Consider a decomposition

$$X'X/(nT) = M + M^{(1)}/\sqrt{T} + M^{(2)}/T, \tag{A1}$$

where

$$M = FD_nF'/T,\ M^{(1)} = (F\Lambda'e + e'\Lambda F')/(n\sqrt{T}),\ \text{and } M^{(2)} = e'e/n.$$

*Alexei Onatski: University of Cambridge, ao319@cam.ac.uk

We use Kato's (1980) theory to characterize the eigenvalues and eigenprojections of $X'X/(nT)$ as perturbations of those of M. Similar techniques were recently used in the analysis of the quasi maximum likelihood estimator in panel data models with interactive fixed effects by Moon and Weidner (2010). Let $R(z) = (M - zI_T)^{-1}$, $z \in \mathbb{C}$. Then, according to Kato (1980, p.78-79), for $1 \le j \le r$,

$$\mu_j(X'X)/(nT) = \mu_j(M) + \sum_{s=1}^{\infty} \mu_j^{(s)} T^{-s/2} = d_{jn} + \sum_{s=1}^{\infty} \mu_j^{(s)} T^{-s/2}, \quad \text{(A2)}$$

where

$$\mu_j^{(s)} = \sum_{\nu_1+\ldots+\nu_p=s} \frac{(-1)^p}{2\pi p \mathrm{i}} \operatorname{tr} \int_\Gamma M^{(\nu_1)} R(z) \ldots M^{(\nu_p)} R(z)\, \mathrm{d}z$$

with ν_k, $k = 1, \ldots, p$, taking on only values one or two; $\mathrm{i} \in \mathbb{C}$ being the imaginary unit; and Γ being the circle in $\mathbb{C}$ with center at d_{jn} and radius $r_j = \min_{i=0,1} \{d_{j+i-1,n} - d_{j+i,n}\}/2$. Here, we define $d_{0,n}$ as $+\infty$ and $d_{r+1,n}$ as 0.

As explained by Kato (1980, p.88), the series in (A2) are absolutely converging as long as $\sup_{z\in\Gamma} \sum_{i=1}^{2} T^{-i/2} \left\| M^{(i)} R(z) \right\| < 1$. By definition of Γ and $R(z)$, $\sup_{z\in\Gamma} \|R(z)\| = r_j^{-1}$. Therefore, a sufficient condition for the convergence is $\max_{i=1,2} \left\{ \left\| M^{(i)} \right\| \right\} < Tr_j/2$. We have

$$\begin{aligned} \left\| F\Lambda' e/(n\sqrt{T}) \right\| &\le \|\Lambda\| \|F\| \|e\| /(n\sqrt{T}) \\ &= \|\Lambda'\Lambda/n\|^{1/2} \|F'F/T\|^{1/2} \|e'e/n\|^{1/2} = d_{1n}^{1/2} \|e'e/n\|^{1/2}. \end{aligned}$$

Therefore, by A3 (iii), $\left\| M^{(1)} \right\| = O_{\mathrm{P}}(1)$ and $\left\| M^{(2)} \right\| = O_{\mathrm{P}}(1)$. In particular, for any $\varepsilon > 0$ and any sequence $\{n_T, T = 1, 2, ..\}$ such that $n_T, T \to_c \infty$, there exists $\bar{T} > 0$ such that $\Pr\left(\max_{i=1,2} \left\{ \left\| M^{(i)} \right\| \right\} < \bar{T} r_j/2 \right) > 1 - \varepsilon$ for all $T > \bar{T}$. That is, with probability larger than $1 - \varepsilon$, the convergence in (A2) takes place for all $T > \bar{T}$. Furthermore, by Kato's (1980, p.89) formula (3.6), with the same probability, for all $T > \bar{T}$,

$$\left|\mu_j\left(X'X/(nT)\right) - d_{jn} - \mu_j^{(1)}/T^{1/2} - \mu_j^{(2)}/T\right| \leq \frac{r_j T^{-3/2}}{\bar{T}^{-1}\left(\bar{T}^{-1/2} - T^{-1/2}\right)},$$

which implies that

$$\mu_j\left(X'X/(nT)\right) = d_{jn} + \mu_j^{(1)}/T^{1/2} + \mu_j^{(2)}/T + o_{\mathrm{P}}(1/T). \tag{A3}$$

Let $P_j = F_{\cdot j}\left(F_{\cdot j}'F_{\cdot j}\right)^{-1}F_{\cdot j}' = F_{\cdot j}F_{\cdot j}'/T$ be the eigenprojection corresponding to the j-th eigenvalue of M, $\mu_j(M) = d_{jn}$, and let P_0 be the projection on the subspace of $\mathbb{R}^T$ orthogonal to all columns of F. Kato (1980, p.79) gives the following explicit formulae for $\mu_j^{(1)}$ and $\mu_j^{(2)}$:

$$\mu_j^{(1)} = \operatorname{tr}\left[M^{(1)}P_j\right] \text{ and } \mu_j^{(2)} = \operatorname{tr}\left[M^{(2)}P_j - M^{(1)}S_jM^{(1)}P_j\right], \tag{A4}$$

where

$$S_j = \sum\nolimits_{k\neq j,k=1}^{r} P_k/\left(d_{kn} - d_{jn}\right) - P_0/d_{jn}. \tag{A5}$$

Using (A4) and the definition of $M^{(1)}$, we have

$$\mu_j^{(1)} = 2\operatorname{tr}\left[F\Lambda'eF_{\cdot j}F_{\cdot j}'\right]/\left(nT^{3/2}\right) = 2\Lambda_{\cdot j}'eF_{\cdot j}/\left(nT^{1/2}\right) = 2\sqrt{d_{jn}}e^{(j,j)}/\sqrt{n}. \tag{A6}$$

Further, straightforward algebra that employs (A5), shows that

$$\begin{aligned}
\operatorname{tr}\left[M^{(1)}S_jM^{(1)}P_j\right] &= \sum_{k\neq j,k=1}^{r}\frac{\left(\sqrt{d_{jn}}e^{(j,k)} + \sqrt{d_{kn}}e^{(k,j)}\right)^2}{n\left(d_{kn} - d_{jn}\right)} - \Lambda_{\cdot j}'eP_0e'\Lambda_{\cdot j}/\left(d_{jn}n^2\right) \\
&= \sum_{k\neq j,k=1}^{r}\frac{\left(\sqrt{d_{jn}}e^{(j,k)} + \sqrt{d_{kn}}e^{(k,j)}\right)^2}{n\left(d_{kn} - d_{jn}\right)} - \Lambda_{\cdot j}'ee'\Lambda_{\cdot j}/\left(d_{jn}n^2\right) \\
&\quad + \sum_{k=1}^{r}\left(e^{(j,k)}\right)^2/n.
\end{aligned}$$

By A3 (ii), $e^{(j,k)} = O_{\mathrm{P}}(1)$ and $e^{(k,j)} = O_{\mathrm{P}}(1)$. Therefore, recalling that n and T are of the same order when $n, T \to_c \infty$, we get

$$\operatorname{tr}\left[M^{(1)} S_j M^{(1)} P_j\right] = -\Lambda'_{\cdot j} e e' \Lambda_{\cdot j} / \left(d_{jn} n^2\right) + O_{\mathrm{P}}\left(1/T\right).$$

Since $\operatorname{tr}\left[M^{(2)} P_j\right] = F'_{\cdot j} e' e F_{\cdot j} / nT$, we obtain

$$\mu_j^{(2)} = F'_{\cdot j} e' e F_{\cdot j} / nT + \Lambda'_{\cdot j} e e' \Lambda_{\cdot j} / \left(d_{jn} n^2\right) + O_{\mathrm{P}}\left(1/T\right). \tag{A7}$$

Equalities (A3), (A6), and (A7) imply the lemma.

2 Proof of Lemma 3

Consider decomposition (A1). According to Kato (1980, p.68),

$$\hat{P}_j = P_j + \sum_{s=1}^{\infty} P_j^{(s)} / T^{s/2}, \tag{A8}$$

where

$$P_j^{(s)} = -\sum\nolimits_{\nu_1 + \ldots + \nu_p = s} \frac{(-1)^p}{2\pi i} \int_\Gamma R\left(z\right) M^{(\nu_1)} R\left(z\right) M^{(\nu_2)} \ldots M^{(\nu_p)} R\left(z\right) \mathrm{d}z \tag{A9}$$

with ν_k, $k = 1, ..., p$, $R(z)$, and Γ defined as in the proof of Lemma 2. As in that proof, for any $\varepsilon > 0$ and any sequence $\{n_T, T = 1, 2, ..\}$ such that $n_T, T \to_c \infty$, let $\bar{T}$ be such that $\Pr\left(\max_{i=1,2}\left\{\left\|M^{(i)}\right\|\right\} < \bar{T} r_j / 2\right) > 1 - \varepsilon$ for all $T > \bar{T}$. Then, since $\sup_{z \in \Gamma}\left|R\left(z\right)\right| = 1/r_j$, with probability larger than $1 - \varepsilon$,

$$\left\|P_j^{(s)}\right\| \leq \sum\nolimits_{\nu_1 + \ldots + \nu_p = s} \frac{1}{2\pi} \int_\Gamma \left(1/r_j\right)^{p+1} \left(\bar{T} r_j / 2\right)^p \left|\mathrm{d}z\right| = \sum\nolimits_{\nu_1 + \ldots + \nu_p = s} \left(\bar{T}/2\right)^p.$$

Since ν_i may only be equal to one or two, there are no more than 2^s summands in the latter sum. Therefore, with probability larger than $1 - \varepsilon$, $\left\|P_j^{(s)}\right\| \leq$

$\left(2\bar{T}\right)^s$ for all $s = 1, 2, ...$ and all $T > \bar{T}$. Hence, by (A8), with probability larger than $1 - \varepsilon$, for all $T > \left(2\bar{T}\right)^2$,

$$\left\| \hat{P}_j - P_j - P_j^{(1)}/\sqrt{T} - P_j^{(2)}/T \right\| \leq \left(2\bar{T}/\sqrt{T}\right)^3 / \left(1 - 2\bar{T}/\sqrt{T}\right),$$

which implies that

$$\hat{P}_j = P_j + P_j^{(1)}/\sqrt{T} + P_j^{(2)}/T + o_{\mathrm{P}}\left(1/T\right), \tag{A10}$$

where $To_{\mathrm{P}}\left(1/T\right)$ converges to zero in probability in spectral norm.

Kato (1980, p.77) gives the following explicit formulae for $P_j^{(1)}$ and $P_j^{(2)}$:

$$P_j^{(1)} = -P_j M^{(1)} S_j - S_j M^{(1)} P_j, \text{ and} \tag{A11}$$

$$\begin{aligned} P_j^{(2)} &= -P_j M^{(2)} S_j - S_j M^{(2)} P_j + P_j M^{(1)} S_j M^{(1)} S_j \\ &+ S_j M^{(1)} P_j M^{(1)} S_j + S_j M^{(1)} S_j M^{(1)} P_j - P_j M^{(1)} P_j M^{(1)} S_j^2 \\ &- P_j M^{(1)} S_j^2 M^{(1)} P_j - S_j^2 M^{(1)} P_j M^{(1)} P_j. \end{aligned} \tag{A12}$$

Using (A10)-(1) and the fact that $P_j S_j = 0$, we obtain, for $j \leq r$,

$$\mathrm{tr}\left[P_j \hat{P}_j\right] = 1 - \mathrm{tr}\left[P_j M^{(1)} S_j^2 M^{(1)} P_j\right]/T + o_{\mathrm{P}}\left(1/T\right).$$

From the latter formula and definition (A5) of S_j, we have

$$\begin{aligned} \mathrm{tr}\left[P_j \hat{P}_j\right] &= 1 - \sum_{k \neq j, k=1}^{r} \frac{\left(\sqrt{d_{kn}} e^{(k,j)} + \sqrt{d_{jn}} e^{(j,k)}\right)^2}{\left(d_{kn} - d_{jn}\right)^2 nT} \\ &\quad - \Lambda'_{\cdot j} e P_0 e' \Lambda_{\cdot j} / (d_{jn}^2 n^2 T) + o_{\mathrm{P}}\left(1/T\right). \end{aligned}$$

Assumption A3 (ii) implies that the second summand on the right hand side of the above equation is $o_{\mathrm{P}}(1/T)$, and hence,

$$\operatorname{tr}\left[P_j\hat{P}_j\right] = 1 - \Lambda'_{\cdot j} e P_0 e' \Lambda_{\cdot j}/(d_{jn}^2 n^2 T) + o_{\mathrm{P}}(1/T).$$

Since $P_0 = I_T - \sum_{k=1}^{r} P_k$, we have

$$\operatorname{tr}\left[P_j\hat{P}_j\right] = 1 - \Lambda'_{\cdot j} e e' \Lambda_{\cdot j}/(d_{jn}^2 n^2 T) + \sum_{k=1}^{r} \Lambda'_{\cdot j} e P_k e' \Lambda_{\cdot j}/(d_{jn}^2 T n^2) + o_{\mathrm{P}}(1/T).$$

Noting that $\Lambda'_{\cdot j} e P_k e' \Lambda_{\cdot j} = d_{jn} n \left(e^{(j,k)}\right)^2$ and using A3 (ii) one more time, we get

$$\operatorname{tr}\left[P_j\hat{P}_j\right] = 1 - \Lambda'_{\cdot j} e e' \Lambda_{\cdot j}/(d_{jn}^2 n^2 T) + o_{\mathrm{P}}(1/T). \tag{A13}$$

For $k \neq j$, using (A10)-(1) and (A5), we have

$$\operatorname{tr}\left[P_k\hat{P}_j\right] = \frac{1}{nT}\left(\sqrt{d_{jn}}e^{(j,k)} + \sqrt{d_{kn}}e^{(k,j)}\right)^2 / \left(d_{kn} - d_{jn}\right)^2 + o_{\mathrm{P}}(1/T).$$

By A3 (ii), the first term in the above sum is $o_{\mathrm{P}}(T^{-1})$, and thus, for $k \neq j$,

$$\operatorname{tr}\left[P_k\hat{P}_j\right] = o_{\mathrm{P}}\left(T^{-1}\right). \tag{A14}$$

Lemma 3 (ii) follows from the symmetry of our model with respect to interchanging temporal and cross-sectional dimensions. The symmetry holds up to different normalizations of $\Lambda'\Lambda$ and $F'F$, which explains the "extra d_{jn}" in the denominator of the formula for $\operatorname{tr}\left[P_k\hat{P}_j\right]$ relative to that for $\operatorname{tr}\left[Q_k\hat{Q}_j\right]$.

3 Primitive conditions for A3 (iii)

Proposition 1 *Let $\varepsilon_t = (\varepsilon_{1t}, ..., \varepsilon_{nt})'$, where ε_{it} with $i \in \mathbb{N}$ and $t \in \mathbb{Z}$ are independent zero mean random variables with uniformly bounded fourth mo-*

ments. Assumption A3 (iii) is satisfied for $e = [e_1, ..., e_T]$ *with*

$$e_t = \sum\nolimits_{j=0}^{\infty} \Psi_{nj} \varepsilon_{t-j},$$

where Ψ_{nj} *are* $n \times n$ *matrices such that* $\sum_{j=0}^{\infty} j \|\Psi_{nj}\|^2 < M$, *and* $\sum_{j=0}^{\infty} \|\Psi_{nj}\| < M$ *for an* $M < \infty$ *that does not depend on* n.

Proof: Our proof is similar to Moon and Weidner's (2010a) proof of their example (ii). We have

$$\left(\mu_1 \left(ee'/T\right)\right)^{1/2} = \|e\| / \sqrt{T} \leq \sum\nolimits_{j=0}^{T} \|\Psi_{nj}\| \|\varepsilon_{-j}\| / \sqrt{T} + \|r_{n,T}\|,$$

where $\varepsilon_{-j} = [\varepsilon_{1-j}, ..., \varepsilon_{T-j}]$ and $r_{n,T} = \sum_{j=T+1}^{\infty} \Psi_{nj} \varepsilon_{-j} / \sqrt{T}$. Obviously, for any $j = 0, ..., T$, $\|\varepsilon_{-j}\| \leq \|\varepsilon\|$, where $\varepsilon = [\varepsilon_{1-T}, ..., \varepsilon_T]$. As explained by Moon and Weidner (2010a), $\|\varepsilon\| / \sqrt{T} = O_{\mathrm{P}}(1)$. Therefore,

$$\left(\mu_1 \left(ee'/T\right)\right)^{1/2} \leq O_{\mathrm{P}}(1) \sum\nolimits_{j=0}^{T} \|\Psi_{nj}\| + \|r_{n,T}\| = O_{\mathrm{P}}(1) + \|r_{n,T}\|. \quad \text{(A15)}$$

Next, since the fourth moments of ε_{it} are uniformly bounded, and $E\varepsilon_{it}^2 \leq \left(E\varepsilon_{it}^4\right)^{1/2}$, the second moments of ε_{it} are uniformly bounded too. Let us denote the uniform bound on the second moments of ε_{it} as B. We have

$$\begin{aligned} E\|r_{n,T}\|^2 &\leq \sum_{i=1}^{n} \sum_{t=1}^{T} E\left(\left(r_{n,T}\right)_{it}^2\right) = \frac{1}{T} \sum_{i=1}^{n} \sum_{t=1}^{T} E\left(\sum_{j=T+1}^{\infty} \sum_{s=1}^{n} \left(\Psi_{nj}\right)_{is} \varepsilon_{s,t-j}\right)^2 \\ &\leq B \sum_{j=T+1}^{\infty} \|\Psi_{nj}\|_F^2 \leq \frac{B}{T} \sum_{j=T+1}^{\infty} j \|\Psi_{nj}\|_F^2, \end{aligned}$$

where $\|M\|_F$ denotes the Frobenius norm of matrix M. Since $\|\Psi_{nj}\|_F^2 \leq n \|\Psi_{nj}\|^2$ (see Horn and Johnson (1985), p. 314), we have

$$E\|r_{n,T}\|^2 \leq \frac{Bn}{T} \sum_{j=T+1}^{\infty} j \|\Psi_{nj}\|^2 = o(1).$$

Hence, $\|r_{n,T}\|^2 = o_{\mathrm{P}}(1)$, and $\|r_{n,T}\| = o_{\mathrm{P}}(1)$ too. Combining this with (A15), we obtain $\mu_1(ee'/T) = O_{\mathrm{P}}(1)$.

References

[1] Horn, R.A., and C.R. Johnson (1985), *Matrix Analysis*, Cambridge University Press.

[2] Kato, T. (1980) Perturbation Theory for Linear Operators. Second Edition. Springer-Verlag. Berlin, Heidelberg, New York.

[3] Moon, H.R., and M. Weidner (2010), "Dynamic Linear Panel Regression Models with Interactive Fixed Effects", manuscript, UCL.

[4] Moon, H.R., and M. Weidner (2010a), A Supplementary Appendix to "Dynamic Linear Panel Regression Models with Interactive Fixed Effects", manuscript, UCL.

[5] Onatski, A. (2013), "Asymptotic Analysis of the Squared Estimation Error in Misspecified Factor Models", manuscript, University of Cambridge.