E-Companion of "Asset Pricing with Spatial Interaction" by Steven Kou, Xianhua Peng, and Haowen Zhong

The E-Companion is organized as follows. Section EC.1 presents the mean-variance analysis with spatial interaction. Section EC.2 gives the proof of the S-CAPM theorem when there is a risk-free asset and when there is no risk free asset. The proof of the S-APT theorems is given in Section EC.3. Section EC.4 develops the econometric tools for implementing the S-APT Model. At last, Section EC.5 reports the spatial weight matrices, S&P credit ratings of eurozone countries and states of the USA, and parameter estimates of S-APT models in the empirical studies.

EC.1. Mean-Variance Analysis with Spatial Interaction

Assume that the n returns $\tilde{r} = (r_1, \ldots, r_{n_1}, r_{n_1+1}, \ldots, r_n)'$ satisfy the model (6), where the first n_1 are ordinary asset returns and the last n_2 are futures returns defined in (9). Then, the mean μ and covariance matrix Σ of $\tilde{r}$ are given by (8).

Consider the mean-variance problem faced by an investor in such a market. Let w be the initial wealth of the investor. Let $u = (u_1, \ldots, u_{n_1})'$ denote the vector of dollar-valued wealth invested in the first n_1 risky assets and $v = (v_1, \ldots, v_{n_2})'$ denote the vector of dollar-valued positions (i.e., the number of contracts times the futures price) of the investor on the n_2 futures contracts. Define the investor's portfolio weights as

$$\phi = (\phi_1, \ldots, \phi_n)' := \frac{1}{w}(u_1, \ldots, u_{n_1}, v_1, \ldots, v_{n_2})'.$$

Let r be the risk-free return. Then, the net return of the investor's portfolio is

$$\begin{aligned} r_p &= \frac{1}{w}\left(\sum_{i=1}^{n_1} u_i(1+r_i) + (w - \sum_{i=1}^{n_1} u_i)(1+r) + \sum_{j=1}^{n_2} \frac{v_j}{F_{j,0}}(F_{j,1} - F_{j,0})\right) - 1, \\ &= \frac{1}{w}\left(\sum_{i=1}^{n_1} u_i(r_i - r) + \sum_{j=1}^{n_2} v_j r_{n_1+j}\right) + r = \phi'(\tilde{r} - r 1_{n_1,n_2}) + r, \end{aligned} \tag{EC.1}$$

where $1_{n_1,n_2}$ is defined in (12). The mean and variance of r_p are given by

$$E[r_p] = h'\phi + r, \;\; Var(r_p) = \phi'\Sigma\phi, \;\text{where } h = \mu - r 1_{n_1,n_2}.$$

Let e denote the target mean portfolio return of the investor, then the mean-variance problem faced by the investor is

$$\min_{\phi} \frac{1}{2}\phi'\Sigma\phi \;\; s.t. \;\; h'\phi + r = e. \tag{EC.2}$$

Using Lagrange multiplier, we obtain the optimal solution to the problem:

$$\phi^* = (e - r)\frac{\Sigma^{-1}h}{h'\Sigma^{-1}h}. \tag{EC.3}$$

When there is no risk-free asset in the market, the net return of the investor's portfolio is

$$r_p = \frac{1}{w}\left(\sum_{i=1}^{n_1} u_i(1+r_i) + \sum_{j=1}^{n_2} \frac{v_j}{F_{j,0}}(F_{j,1}-F_{j,0})\right) - 1 = \phi'\tilde{r}.$$

Then, the mean-variance problem becomes

$$\min_{\phi} \frac{1}{2}\phi'\Sigma\phi \ s.t. \ \phi'\mu = e \text{ and } \phi' 1_{n_1,n_2} = 1, \tag{EC.4}$$

whose optimal solution can be shown to be

$$\phi^* = \psi + e\xi, \text{ where} \tag{EC.5}$$

$$\psi = \frac{1}{D}(B\Sigma^{-1}1_{n_1,n_2} - A\Sigma^{-1}\mu), \xi = \frac{1}{D}(F\Sigma^{-1}\mu - A\Sigma^{-1}1_{n_1,n_2}), \tag{EC.6}$$

$$A = \mu'\Sigma^{-1}1_{n_1,n_2}, B = \mu'\Sigma^{-1}\mu, F = 1'_{n_1,n_2}\Sigma^{-1}1_{n_1,n_2}, D = BF - A^2.$$

Because both μ and Σ are functions of ρ, the optimal portfolio weights ϕ^* and the efficient frontiers are affected by ρ. More specifically, (i) When there exists a risk-free return r, the efficient frontier is $e = r + \sigma\sqrt{H}$ and $H = (\mu - r1_{n_1,n_2})'\Sigma^{-1}(\mu - r1_{n_1,n_2}) = (\alpha - r1_{n_1,n_2})'V^{-1}(\alpha - r1_{n_1,n_2}) + 2r1'_{n_1,n_2}W'V^{-1}(\alpha - r1_{n_1,n_2})\rho + r^2 1'_{n_1,n_2}W'V^{-1}W1_{n_1,n_2}\rho^2$. Thus H is a quadratic function of ρ and the coefficient in front of ρ^2 is positive. (ii) When there is no risk-free asset, the effect of spatial interaction on efficient frontier is illustrated in Figure 1.

EC.2. Proof of the S-CAPM Theorems

EC.2.1. Proof of Theorem 1

LEMMA EC.1. *Let r_{mv} be any mean-variance efficient return other than the risk-free return. Then,*

(i) for any portfolio return y, it holds that

$$E[y] - r = \frac{Cov(y, r_{mv})}{Var(r_{mv})}(E[r_{mv}] - r); \tag{EC.7}$$

(ii) for the futures contracts, it holds that

$$E[F_{i,1}] - F_{i,0} = \frac{Cov(F_{i,1}, r_{mv})}{Var(r_{mv})}(E[r_{mv}] - r), i = 1, 2, \ldots, n_2. \tag{EC.8}$$

Proof. Let $e_{mv} := E[r_{mv}]$. Since r_{mv} is mean-variance efficient, it follows from (EC.3) that the portfolio weight of r_{mv} is given by

$$\phi_{mv} = (e_{mv} - r)\frac{\Sigma^{-1}h}{h'\Sigma^{-1}h}. \tag{EC.9}$$

For any portfolio return y with the portfolio weight ϕ, it follows from (EC.1) that

$$y = \phi'(\tilde{r} - r1_{n_1,n_2}) + r, E[y] - r = \phi'(\mu - r1_{n_1,n_2}) = \phi' h, \tag{EC.10}$$

and then

$$Cov(y, r_{mv}) = \phi' Cov(\tilde{r}) \phi_{mv} = (e_{mv} - r) \frac{\phi' h}{h' \Sigma^{-1} h}, \tag{EC.11}$$

$$Var(r_{mv}) = Cov(r_{mv}, r_{mv}) = (e_{mv} - r) \frac{\phi'_{mv} h}{h' \Sigma^{-1} h} = \frac{(e_{mv} - r)^2}{h' \Sigma^{-1} h}. \tag{EC.12}$$

It follows from (EC.10), (EC.11) and (EC.12) that

$$\frac{Cov(y, r_{mv})}{Var(r_{mv})} (E[r_{mv}] - r) = \phi' h = E[y] - r. \tag{EC.13}$$

In particular, consider a portfolio with all wealth w invested in the risk-free asset and a dollar-valued position u_i on the ith futures contracts. Then the portfolio return $y = u_i r_{n_1+i}/w + r$. Plugging y into (EC.13) leads to (EC.8). $\square$

Proof of Theorem 1. Suppose that there are J investors in the economy and w_j is the initial wealth of the jth investor. Suppose that each investor selects his/her investment portfolio by solving the mean-variance problem (EC.2) and the jth investor has a target mean return of e_j. Then, by (EC.3), the position of the jth investor is

$$\phi^j = (e_j - r) \frac{\Sigma^{-1} h}{h' \Sigma^{-1} h}, \text{ where } h = \mu - r 1_{n_1, n_2}, j = 1, 2, \ldots, J.$$

Let C_i be the market capitalization of asset i, $i = 1, 2, \ldots, n_1$, and $C_M = \sum_{i=1}^{n_1} C_i$ be the total market capitalization. In market equilibrium, since the aggregate of all positions on futures contracts and risk free asset is zero, it follows that

$$\sum_{j=1}^{J} w_j (e_j - r) \frac{\Sigma^{-1} h}{h' \Sigma^{-1} h} = (C_1, \ldots, C_{n_1}, 0, \ldots, 0)', \tag{EC.14}$$

$$\sum_{j=1}^{J} w_j \left(1 - (e_j - r) \frac{1'_{n_1, n_2} \Sigma^{-1} h}{h' \Sigma^{-1} h} \right) = 0, \tag{EC.15}$$

which leads to

$$\frac{\Sigma^{-1} h}{h' \Sigma^{-1} h} = \begin{pmatrix} g \\ 0 \end{pmatrix}, \text{ where } g = \frac{1}{\sum_{j=1}^{J} w_j (e_j - r)} (C_1, \ldots, C_{n_1})'; \text{ and } \sum_{j=1}^{J} w_j = C_M. \tag{EC.16}$$

Therefore, in equilibrium, each investor holds only the market portfolio and the risk-free asset, and no investor trades the futures contracts. Furthermore, let $e_M = \sum_{j=1}^{J} w_j e_j / C_M$. Then, (EC.14) and (EC.16) yield

$$(C_1, \ldots, C_{n_1}, 0, \ldots, 0)' = C_M (e_M - r) \frac{\Sigma^{-1} h}{h' \Sigma^{-1} h},$$

which shows that the market portfolio is mean-variance efficient. Then, applying Lemma EC.1 with r_{mv} being r_M leads to the conclusion of Theorem 1. $\square$

EC.2.2. The S-CAPM with Futures When There Is No Risk-free Asset

THEOREM EC.1. *(S-CAPM with Futures When There Is No Risk-free Asset) Suppose that there is no risk free asset and that the returns of $n = n_1 + n_2$ risky assets are generated by the model (6), of which the first n_1 are returns of ordinary assets and the others are defined in (9) which are "nominal returns" of futures contracts. Suppose that $n_1 > 0$. Let r_M be the return of market portfolio . If each investor holds mean-variance efficient portfolio, then in equilibrium, r_M is mean-variance efficient. Furthermore, if r_M is not the minimum-variance return, then there exists another mean-variance efficient return r_0 such that $Cov(r_M, r_0) = 0$, and it holds that*

(i) for the ordinary assets,

$$E[r_i] - E[r_0] = \frac{Cov(r_i, r_M)}{Var(r_M)}(E[r_M] - E[r_0]),\ i = 1, 2, \ldots, n_1;$$

(ii) for the futures contracts,

$$E[F_{i,1}] - F_{i,0} = \frac{Cov(F_{i,1}, r_M)}{Var(r_M)}(E[r_M] - E[r_0]),\ i = 1, 2, \ldots, n_2.$$

LEMMA EC.2. *Let r_{mv} be any mean-variance efficient return other than the minimum-variance return. Then, there exists another mean-variance efficient return r_0 such that $Cov(r_{mv}, r_0) = 0$. Furthermore,*

(i) for any portfolio return y, it holds that

$$E[y] - E[r_0] = \frac{Cov(y, r_{mv})}{Var(r_{mv})}(E[r_{mv}] - E[r_0]); \tag{EC.17}$$

(ii) for the futures contracts, it holds that

$$E[F_{i,1}] - F_{i,0} = \frac{Cov(F_{i,1}, r_{mv})}{Var(r_{mv})}(E[r_{mv}] - E[r_0]), i = 1, 2, \ldots, n_2. \tag{EC.18}$$

Proof. Let $e_{mv} := E[r_{mv}]$. By (EC.5), the portfolio weight of r_{mv} is given by $\phi_{mv} = \psi + e_{mv}\xi$, where ψ and ξ is given in (EC.6). For any portfolio return y with portfolio weight ϕ, it holds that $y = \phi'\tilde{r}$, $\phi' 1_{n_1,n_2} = 1$, and $E[y] = \phi'\mu$. Hence,

$$\begin{aligned} &Cov(y, r_{mv}) = \phi' Cov(\tilde{r})\phi_{mv} = \phi' Cov(\tilde{r})(\psi + e_{mv}\xi) \\ =&\frac{1}{D}\left(B\phi' 1_{n_1,n_2} - A\phi'\mu + (F\phi'\mu - A\phi' 1_{n_1,n_2})e_{mv}\right) \\ =&\frac{1}{D}(B - AE[y] + Fe_{mv}E[y] - Ae_{mv}). \end{aligned} \tag{EC.19}$$

In particular, letting $y = r_{mv}$ in the above equation leads to $Var(r_{mv}) = \frac{1}{D}(B + Fe_{mv}^2 - 2Ae_{mv})$. Since r_{mv} is not the minimum-variance return, $e_{mv} \neq \frac{A}{F}$. Let r_0 be a mean-variance efficient return with mean $E[r_0] = \frac{B - Ae_{mv}}{A - Fe_{mv}}$. It then follows from (EC.19) that $Cov(r_0, r_{mv}) = 0$. Since $E[r_{mv}] - E[r_0] = \frac{Fe_{mv}^2 + B - 2Ae_{mv}}{Fe_{mv} - A}$, it follows that

$$\begin{aligned}
&\frac{Cov(y, r_{mv})}{Var(r_{mv})}(E[r_{mv}] - E[r_0]) \\
=&\frac{\frac{1}{D}(B - AE[y] + Fe_{mv}E[y] - Ae_{mv})}{\frac{1}{D}(B + Fe_{mv}^2 - 2Ae_{mv})} \times \frac{Fe_{mv}^2 + B - 2Ae_{mv}}{Fe_{mv} - A} \\
=&\frac{(Fe_{mv} - A)E[y] + B - Ae_{mv}}{Fe_{mv} - A} = E[y] - E[r_0],
\end{aligned} \tag{EC.20}$$

which completes the proof of (EC.17). Letting $y = r_1$ in (EC.20), we have

$$E[r_1] - E[r_0] = \frac{Cov(r_1, r_{mv})}{Var(r_{mv})}(E[r_{mv}] - E[r_0]). \tag{EC.21}$$

Letting $y = r_1 + r_{n_1+i}$, which corresponds to $\phi = (1, 0, \ldots, 0, 0, \ldots, 0, 1, \ldots, 0, 0)'$, in (EC.20), we have

$$E[r_1] + E[r_{n_1+i}] - E[r_0] = \frac{Cov(r_1 + r_{n_1+i}, r_{mv})}{Var(r_{mv})}(E[r_{mv}] - E[r_0]). \tag{EC.22}$$

Then, (EC.18) follows from subtracting (EC.21) from (EC.22). □

Proof of Theorem EC.1. When there is no risk-free asset, the mean-variance problem faced by an investor is given by (EC.4), and the optimal portfolio weight is given by (EC.5). Suppose that there are J investors in the economy and let w_j and e_j be the initial wealth and target mean return of the jth investor. Then, the position of the jth investor is

$$\phi^j = \psi + e_j \xi, \;\; j = 1, 2, \ldots, J.$$

Let C_i be the market capitalization of asset i, $i = 1, 2, \ldots, n_1$, and $C_M = \sum_{i=1}^{n_1} C_i$ be the total market capitalization. In market equilibrium, since the aggregate of all positions of futures contracts is zero, it follows that

$$\sum_{j=1}^{J} w_j \phi^j = \left(\sum_{j=1}^{J} w_j\right)\psi + \left(\sum_{j=1}^{J} w_j e_j\right)\xi = (C_1, \ldots, C_{n_1}, 0, \ldots, 0)'. \tag{EC.23}$$

By (EC.23), $\sum_{j=1}^{J} w_j 1'_{n_1,n_2}\phi^j = 1'_{n_1,n_2}(C_1, \ldots, C_{n_1}, 0, \ldots, 0)' = C_M$, which in combination with $1'_{n_1,n_2}\phi^j = 1$ leads to $\sum_{j=1}^{J} w_j = C_M$. Let $e_M = \sum_{j=1}^{J} w_j e_j / C_M$. Then, it follows from (EC.23) that

$$\psi + e_M \xi = \frac{1}{C_M}(C_1, \ldots, C_{n_1}, 0, \ldots, 0)',$$

which shows that the market portfolio is mean-variance efficient with a target mean return e_M. Then, the conclusion in Theorem EC.1 follows by applying Lemma EC.2 with the market portfolio return r_M being r_{mv}. □

EC.3. Proof of the S-APT Theorems

EC.3.1. Proof of Theorem 2

Proof. Fix any $\delta > 0$. Without loss of generality, assume that $|\bar{\alpha}_j^{(n)}| > \delta,\ j = 1, \ldots, N(n, \delta)$. We rewrite (17) as

$$(I_n - \rho^{(n)} W^{(n)})(\tilde{r}^{(n)} - r 1_{n_1,n_2}) = \bar{\alpha}^{(n)} + B^{(n)} \tilde{F} + \tilde{\epsilon}^{(n)}.$$

Let η_j be the jth column of I_n. For $1 \leq j \leq N(n, \delta)$, if $\bar{\alpha}_j^{(n)} > \delta$, consider the zero-cost portfolio with payoff $\eta_j'(I_n - \rho^{(n)} W^{(n)})(\tilde{r}^{(n)} - r 1_{n_1,n_2}) - \eta_j' B^{(n)} \tilde{F}$, which by definition is equal to $\bar{\alpha}_j^{(n)} + \epsilon_j^{(n)}$, a random variable with mean $\bar{\alpha}_j^{(n)} > \delta$ and variance not exceeding $\bar{\sigma}^2$; if $\bar{\alpha}_j^{(n)} < -\delta$, one can construct another zero-cost portfolio with payoff $-\bar{\alpha}_j^{(n)} - \epsilon_j^{(n)}$ by taking opposite positions of the previous portfolio. In this way, $N(n, \delta)$ such portfolios can be constructed. Since the components of $\tilde{\epsilon}^{(n)}$ are uncorrelated with each other, a portfolio with equal weights in these $N(n, \delta)$ portfolios has a payoff with mean greater than δ and variance less than $\bar{\sigma}^2 / N(n, \delta)$. If there exists a subsequence $\{m_1, m_2, \ldots\}$ such that $N(m_k, \delta)$ grows unboundedly as k goes to infinity, then the corresponding sequence of portfolios will have payoffs with means greater than δ and diminishing variances, constituting an asymptotic arbitrage opportunity. Therefore, if no asymptotic arbitrage opportunities exist, then there exists a number N_δ not depending on n such that $N(n, \delta) < N_\delta$ for all n. Since δ can be arbitrarily small, we conclude that $\bar{\alpha}^{(n)} \approx 0$. □

EC.3.2. Proof of Theorem 3

Proof. Since $I_n - \rho^{(n)} W^{(n)}$ is invertible, the model (19) can be written as

$$\tilde{r}^{(n)} = Q^{(n)} \alpha^{(n)} + Q^{(n)} B^{(n)} \tilde{f} + Q^{(n)} \tilde{\epsilon}^{(n)}, \text{ where } Q^{(n)} = (I_n - \rho^{(n)} W^{(n)})^{-1}. \tag{EC.24}$$

For the sake of notational simplicity, the superscript (n) will be dropped when the meaning is clear. Let

$$\hat{\alpha} = Q\alpha, \hat{B} = QB, \tilde{\varepsilon} = Q\tilde{\epsilon}, \Omega = QVQ'. \tag{EC.25}$$

Since Ω is positive definite, it can be factored as $\Omega = CC'$ where C is a nonsingular matrix. $C^{-1}\hat{\alpha}$ can be orthogonally projected into the space spanned by $C^{-1} 1_{n_1,n_2}$ and the columns of $C^{-1}\hat{B}$:

$$C^{-1}\hat{\alpha} = C^{-1} 1_{n_1,n_2} \lambda_0 + C^{-1} \hat{B} \lambda + u,$$

where u satisfies the orthogonality condition:

$$0 = \hat{B}'(C')^{-1} u, \tag{EC.26}$$

$$0 = 1_{n_1,n_2}'(C')^{-1} u. \tag{EC.27}$$

Define the pricing errors $v := \hat{\alpha} - 1_{n_1,n_2}\lambda_0 - \hat{B}\lambda = Cu$. Then, by (EC.26) and (EC.27), we have

$$0 = \hat{B}'(C')^{-1}u = \hat{B}'(C')^{-1}C^{-1}v = \hat{B}'\Omega^{-1}v, \tag{EC.28}$$

$$0 = 1'_{n_1,n_2}(C')^{-1}u = 1'_{n_1,n_2}(C')^{-1}C^{-1}v = 1'_{n_1,n_2}\Omega^{-1}v. \tag{EC.29}$$

There are two cases: (i) there exists $N > 0$ such that $v^{(n)} = 0$ for all $n > N$; and (ii) for any $N > 0$, there exists $n > N$ such that $v^n \neq 0$.

In case (i), since $U^{(n)} = (Q^{(n)})^{-1}v^{(n)} = 0$ for all $n > N$, it follows that (28) holds for $A := \max\{(U^{(n)})'(V^{(n)})^{-1}U^{(n)} \mid n = 1, \ldots, N\}$.

In case (ii), since $U^{(n)} = (Q^{(n)})^{-1}v^{(n)} = 0$ for any n such that $v^{(n)} = 0$, we will only consider those index n such that $v^{(n)} \neq 0$ in the following. Consider the portfolio $h = \Omega^{-1}v(v'\Omega^{-1}v)^{-1}$. Since $v \neq 0$, $v'\Omega^{-1}v > 0$, $h \neq 0$, and $U = Q^{-1}v \neq 0$. By (EC.29), h is a zero-cost portfolio. By (EC.28), the payoff of the zero-cost portfolio is

$$\begin{aligned} h'\tilde{r} &= h'(Q\alpha + QB\tilde{f} + Q\tilde{\epsilon}) = h'(\hat{\alpha} + \hat{B}\tilde{f} + \tilde{\varepsilon}) = h'\hat{\alpha} + (v'\Omega^{-1}v)^{-1}v'\Omega^{-1}\hat{B}\tilde{f} + h'\tilde{\varepsilon} \\ &= h'\hat{\alpha} + h'\tilde{\varepsilon}, \end{aligned}$$

whose expectation and variance are

$$\begin{aligned} E[h'\tilde{r}] &= h'\hat{\alpha} = (v'\Omega^{-1}v)^{-1}v'\Omega^{-1}(1_{n_1,n_2}\lambda_0 + \hat{B}\lambda + v) = 1, \\ Var(h'\tilde{r}) &= h'\Omega h = (v'\Omega^{-1}v)^{-2}v'\Omega^{-1}\Omega\Omega^{-1}v = (v'\Omega^{-1}v)^{-1} \\ &= [(\hat{\alpha} - 1_{n_1,n_2}\lambda_0 - \hat{B}\lambda)'(Q')^{-1}V^{-1}(Q)^{-1}(\hat{\alpha} - 1_{n_1,n_2}\lambda_0 - \hat{B}\lambda)]^{-1} \\ &= [(\alpha - (I_n - \rho W)1_{n_1,n_2}\lambda_0 - B\lambda)'V^{-1}(\alpha - (I_n - \rho W)1_{n_1,n_2}\lambda_0 - B\lambda)]^{-1} \\ &= (U'V^{-1}U)^{-1}. \end{aligned}$$

Therefore, if (28) is violated, the variance of $h'\tilde{r}$ vanishes along some subsequence, which constitutes an asymptotic arbitrage opportunity.

The proof for the case when there exists a risk-free return r is almost a copy of the above. Let $Q, \hat{\alpha}, \hat{B}, \tilde{\varepsilon}$, and Ω be defined in (EC.24) and (EC.25). Since Ω is positive definite, it can be factored as $\Omega = CC'$ where C is a nonsingular matrix. Project $C^{-1}(\hat{\alpha} - r1_{n_1,n_2})$ onto the space spanned by the columns of $C^{-1}\hat{B}$:

$$C^{-1}(\hat{\alpha} - r1_{n_1,n_2}) = C^{-1}\hat{B}\lambda + u.$$

Define the pricing errors $v := \hat{\alpha} - r1_{n_1,n_2} - \hat{B}\lambda = Cu$. Then, by orthogonality, we have

$$0 = \hat{B}'(C')^{-1}u = \hat{B}'(C')^{-1}C^{-1}v = \hat{B}'\Omega^{-1}v. \tag{EC.30}$$

Let $U^{(n)}$ be defined in (29) with $\lambda_0^{(n)}$ replaced by r. There are two cases: (i) there exists $N>0$ such that $v^{(n)}=0$ for all $n>N$; and (ii) for any $N>0$, there exists $n>N$ such that $v^n \neq 0$.

In case (i), since $U^{(n)}=(Q^{(n)})^{-1}v^{(n)}=0$ for all $n>N$, it follows that (28) (with $\lambda_0^{(n)}$ replaced by r) holds for $A:=\max\{(U^{(n)})'(V^{(n)})^{-1}U^{(n)} \mid n=1,\dots,N\}$.

In case (ii), since $U^{(n)}=(Q^{(n)})^{-1}v^{(n)}=0$ for any n such that $v^{(n)}=0$, we will only consider those index n such that $v^{(n)} \neq 0$ in the following. Consider the zero-cost portfolio which has dollar-valued positions $h=(v'\Omega^{-1}v)^{-1}\Omega^{-1}v$ in the n_1 risky assets and the n_2 futures contracts and the position $-h'1_{n_1,n_2}$ in the risk-free asset. Since $v\neq 0$, $v'\Omega^{-1}v>0$, $h\neq 0$, and $U=Q^{-1}v\neq 0$. By (EC.25) and (EC.30), the payoff of the portfolio is

$$\begin{aligned}
&h'(1_{n_1,n_2}+\tilde{r})-h'1_{n_1,n_2}(1+r)=h'(\tilde{r}-r1_{n_1,n_2})=h'(Q\alpha+QB\tilde{f}+Q\tilde{\epsilon}-r1_{n_1,n_2})\\
&=h'(\hat{\alpha}+\hat{B}\tilde{f}+\tilde{\varepsilon}-r1_{n_1,n_2})=h'(\hat{\alpha}-r1_{n_1,n_2})+(v'\Omega^{-1}v)^{-1}v'\Omega^{-1}\hat{B}\tilde{f}+h'\tilde{\varepsilon}\\
&=h'(\hat{\alpha}-r1_{n_1,n_2})+h'\tilde{\varepsilon},
\end{aligned}$$

whose mean and variance are

$$\begin{aligned}
&E[h'(\tilde{r}-r1_{n_1,n_2})]=h'(\hat{\alpha}-r1_{n_1,n_2})=(v'\Omega^{-1}v)^{-1}v'\Omega^{-1}(\hat{B}\lambda+v)=1,\\
&Var(h'(\tilde{r}-r1_{n_1,n_2}))=Var(h'\tilde{\varepsilon})=h'\Omega h=(v'\Omega^{-1}v)^{-2}v'\Omega^{-1}\Omega\Omega^{-1}v\\
&=(v'\Omega^{-1}v)^{-1}=[(\hat{\alpha}-r1_{n_1,n_2}-\hat{B}\lambda)'(Q')^{-1}V^{-1}(Q)^{-1}(\hat{\alpha}-r1_{n_1,n_2}-\hat{B}\lambda)]^{-1}\\
&=[(\alpha-(I_n-\rho W)1_{n_1,n_2}r-B\lambda)'V^{-1}(\alpha-(I_n-\rho W)1_{n_1,n_2}r-B\lambda)]^{-1}\\
&=[U'V^{-1}U]^{-1}.
\end{aligned}$$

Therefore, if (28) (with $\lambda_0^{(n)}$ replaced by r) is violated, then the variance of $h'(\tilde{r}-r1_{n_1,n_2})$ vanishes along some subsequence, and an asymptotic arbitrage opportunity exists. The proof is thus completed. □

EC.4. Econometric Tools for Implementing the S-APT Model (34)

Let $\theta_0=(\rho_0,\underline{b}_0',\sigma_0^2)'$ be the true model parameters that lie in the interior of the parameter space Θ defined by:

$$\Theta:=[\zeta,\gamma]\times[-\delta_b,\delta_b]^{n\times(K+1)}\times[\delta_s^{-1},\delta_s], \tag{EC.31}$$

where $\zeta<0<\gamma, \delta_b>0, \delta_s>0$ are constants and $I_n-\rho W$ is invertible for $\rho\in[\zeta,\gamma]$.[18]

We first make the following mild assumptions regarding $\tilde{g}_t$, which will be needed for proving Proposition EC.1 and Proposition EC.2:

Assumption EC.1. *We assume that $\tilde{g}_t$ satisfies the following mild technical conditions:*

[18] See footnote 11 for more details on the specification of the interval $[\zeta,\gamma]$.

(i) $E[\|\tilde{g}_t\|^2] < \infty$, *where* $\|\tilde{g}_t\|^2 := \sum_{i=1}^{K} g_{it}^2$.

(ii) There exists an open set $\mathbb{A} \subset \mathbb{R}^K$ *such that* $P(\tilde{g}_t \in \mathbb{A}) > 0$ *and the distribution of* $\tilde{g}_t$ *restricted on* $\mathbb{A}$ *has a strictly positive density.*

We also assume that the spatial weight matrix W satisfies the following standard conditions used in the spatial econometrics literature:

Assumption EC.2. *W is non-negative; $W \neq 0$; and the diagonal elements of W are all equal to zero.*

EC.4.1. Identifiability of Model Parameters

Let $l(\tilde{y}_t \mid \tilde{g}_t, \theta)$ be defined in (38). We recall the following definition of identifiability; see, e.g., Neway and McFadden (1994, Lemma 2.2).

Definition EC.1. θ_0 is identifiable if for any $\theta \neq \theta_0$ and $\theta \in \Theta$ it holds that $P(l(\tilde{y}_t \mid \tilde{g}_t, \theta) \neq l(\tilde{y}_t \mid \tilde{g}_t, \theta_0)) > 0$.

It turns out that the identifiability of θ_0 depends largely on the property of the spatial weight matrix W. In particular, we need the following definition for W:

Definition EC.2. The spatial weight matrix W is regular if there exist no $c_1 > 0$ and $c_2 \geq 0$ such that

$$\sum_{k=1}^{n} W_{ki}^2 = c_1, \ \forall i = 1, \ldots, n, \tag{EC.32}$$

$$\sum_{k=1}^{n} W_{ki} W_{kj} = c_2(W_{ij} + W_{ji}), \ \forall 1 \leq i < j \leq n. \tag{EC.33}$$

If W is not regular, then the pair of constants (c_1, c_2) that satisfy (EC.32) and (EC.33) are unique. Indeed, by Assumption EC.2, there exists $i < j$ such that $W_{ij} + W_{ji} > 0$; hence, it follows from (EC.33) that c_2 is uniquely determined by W. Apparently, c_1 is also unique. We have the following proposition regarding the identifiability of θ_0.

Proposition EC.1. *Suppose Assumption EC.1 and Assumption EC.2 hold. Then, any θ_0 is identifiable if W is regular. More generally, a particular θ_0 is identifiable if and only if W satisfies one of the following conditions:*

(i) W is regular.

(ii) W is not regular and corresponds to the unique pair (c_1, c_2) in (EC.32) and (EC.33), and one of the following conditions holds:

$$\rho_0 = -\frac{c_2}{c_1}; \tag{EC.34}$$

$$\rho_0 \neq -\frac{c_2}{c_1} \ \textit{and} \ \frac{1 - c_2 \rho_0}{c_2 + c_1 \rho_0} = \rho_0; \tag{EC.35}$$

$$\rho_0 \neq -\frac{c_2}{c_1} \text{ and } \frac{1-c_2\rho_0}{c_2+c_1\rho_0} \neq \rho_0 \text{ and } \theta_* := (\rho_*, \bar{\alpha}_*, B_*, \sigma_*^2) \notin \Theta, \tag{EC.36}$$

where $\rho_* := \frac{1-c_2\rho_0}{c_2+c_1\rho_0}$*,* $\sigma_*^2 := \sigma_0^2 \frac{c_2^2+c_1}{(c_2+c_1\rho_0)^2}$*,* $\bar{\alpha}_* := \frac{\sigma_*^2}{\sigma_0^2}(I_n - \rho_* W')^{-1}(I_n - \rho_0 W')\bar{\alpha}_0$*, and* $B_* := \frac{\sigma_*^2}{\sigma_0^2}(I_n - \rho_* W')^{-1}(I_n - \rho_0 W')B_0$.

Proof. θ_0 is not identifiable if and only if there exists $\theta \in \Theta$ and $\theta \neq \theta_0$ such that

$$P(l(\tilde{y}_t \mid \tilde{g}_t; \theta) = l(\tilde{y}_t \mid \tilde{g}_t; \theta_0)) = 1. \tag{EC.37}$$

It follows from the part (ii) of Assumption EC.1 that $P((\tilde{y}_t', \tilde{g}_t')' \in \mathbb{R}^n \times \mathbb{A}) > 0$ and the joint distribution of $(\tilde{y}_t', \tilde{g}_t')'$ has a strictly positive density on $\mathbb{R}^n \times \mathbb{A}$. Therefore, (EC.37) implies that

$$l(\tilde{y} \mid \tilde{g}; \theta) = l(\tilde{y} \mid \tilde{g}; \theta_0), \ \forall (\tilde{y}, \tilde{g}) \in \mathbb{R}^n \times \mathbb{A}.$$

By (38), we have

$$l(\tilde{y} \mid \tilde{g}; \theta) = \tilde{y}' C_1(\theta) \tilde{y} + \tilde{g}' C_2(\theta) \tilde{g} + \tilde{g}' C_3(\theta) \tilde{y} + C_4(\theta)' \tilde{y} + C_5(\theta)' \tilde{g} + C_6(\theta), \tag{EC.38}$$

where

$$C_1(\theta) = -\frac{1}{2\sigma^2}(I_n - \rho W')(I_n - \rho W)$$

$$C_2(\theta) = -\frac{1}{2\sigma^2} B' B \tag{EC.39}$$

$$C_3(\theta) = \frac{1}{\sigma^2} B'(I_n - \rho W) \tag{EC.40}$$

$$C_4(\theta) = \frac{1}{\sigma^2}(I_n - \rho W')\bar{\alpha} \tag{EC.41}$$

$$C_5(\theta) = -\frac{1}{\sigma^2} B' \bar{\alpha} \tag{EC.42}$$

$$C_6(\theta) = -\frac{n}{2}\log(2\pi\sigma^2) + \frac{1}{2}\log(\det((I_n - \rho W')(I_n - \rho W))) - \frac{1}{2\sigma^2}\bar{\alpha}'\bar{\alpha}. \tag{EC.43}$$

By the equality of partial derivatives of $l(\tilde{y} \mid \tilde{g}, \theta)$ and $l(\tilde{y} \mid \tilde{g}, \theta_0)$ on $\mathbb{R}^n \times \mathbb{A}$, we obtain that

$$C_i(\theta) = C_i(\theta_0), \ i = 1, 2, \ldots, 6. \tag{EC.44}$$

Since $I_n - \rho W$ is invertible, (EC.40) and (EC.41) imply that

$$B' = \sigma^2 C_3(\theta)(I_n - \rho W)^{-1}, \ \bar{\alpha} = \sigma^2 (I_n - \rho W')^{-1} C_4(\theta). \tag{EC.45}$$

Then, plugging (EC.45) into (EC.39), (EC.42), and (EC.43), we obtain

$$C_2(\theta) = \frac{1}{4} C_3(\theta) C_1(\theta)^{-1} C_3(\theta)',$$

$$C_5(\theta) = \frac{1}{2} C_3(\theta) C_1(\theta)^{-1} C_4(\theta),$$

$$C_6(\theta) = -\log((2\pi)^{\frac{n}{2}} \det(-2C_1(\theta))^{-\frac{1}{2}}) + \frac{1}{4} C_4(\theta)' C_1(\theta)^{-1} C_4(\theta).$$

Hence, (EC.44) is equivalent to

$$C_i(\theta) = C_i(\theta_0), \ i = 1, 3, 4. \tag{EC.46}$$

Now we are ready to prove the proposition. We will first show the sufficiency.

(i) Suppose that W is regular. Suppose for the sake of contradiction that there exists $\theta \neq \theta_0$ such that (EC.37) holds. Then, (EC.46) holds. $C_1(\theta) = C_1(\theta_0)$ is equivalent to that

$$(\frac{1}{\sigma_0^2} - \frac{1}{\sigma^2}) I_n - (\frac{\rho_0}{\sigma_0^2} - \frac{\rho}{\sigma^2})(W' + W) + (\frac{\rho_0^2}{\sigma_0^2} - \frac{\rho^2}{\sigma^2}) W'W = 0. \tag{EC.47}$$

Considering the (i,i)-element of the matrix on the left and noting $W_{ii} = 0$, we obtain

$$\frac{1}{\sigma_0^2} - \frac{1}{\sigma^2} + (\frac{\rho_0^2}{\sigma_0^2} - \frac{\rho^2}{\sigma^2}) \sum_{k=1}^{n} W_{ki}^2 = 0, \ i = 1, 2, \ldots, n. \tag{EC.48}$$

For $i < j$, considering the (i,j)-element of the matrices on both sides of (EC.47), we obtain

$$-(\frac{\rho_0}{\sigma_0^2} - \frac{\rho}{\sigma^2})(W_{ij} + W_{ji}) + (\frac{\rho_0^2}{\sigma_0^2} - \frac{\rho^2}{\sigma^2}) \sum_{k=1}^{n} W_{ki} W_{kj} = 0, \ \forall 1 \leq i < j \leq n. \tag{EC.49}$$

Suppose that $\frac{\rho_0^2}{\sigma_0^2} - \frac{\rho^2}{\sigma^2} = 0$, then (EC.48) implies $\sigma = \sigma_0$, which, together with (EC.49) and that there exists $i \neq j$ such that $W_{ij} + W_{ji} > 0$ (by Assumption EC.2), implies that $\rho = \rho_0$. Then, since $I_n - \rho W$ is invertible, $C_3(\theta) = C_3(\theta_0)$ and (EC.40) imply $B = B_0$; $C_4(\theta) = C_4(\theta_0)$ and (EC.41) imply that $\bar{\alpha} = \bar{\alpha}_0$. Therefore, we have shown that $\theta = \theta_0$, but this contradicts to the assumption that $\theta \neq \theta_0$. Hence, $\frac{\rho_0^2}{\sigma_0^2} \neq \frac{\rho^2}{\sigma^2}$. Suppose without generality that $\frac{\rho_0^2}{\sigma_0^2} > \frac{\rho^2}{\sigma^2}$. Since (EC.48) and that there exists i such that $\sum_{k=1}^{n} W_{ki}^2 > 0$ (by Assumption EC.2), it follows that $\sigma_0 > \sigma$. Hence, (EC.48) and (EC.49) imply that

$$\sum_{k=1}^{n} W_{ki}^2 = c_1, \ i = 1, 2, \ldots, n; \text{ and } \sum_{k=1}^{n} W_{ki} W_{kj} = c_2 (W_{ij} + W_{ji}), \ \forall 1 \leq i < j \leq n;$$

where

$$c_1 = -\frac{\frac{1}{\sigma_0^2} - \frac{1}{\sigma^2}}{\frac{\rho_0^2}{\sigma_0^2} - \frac{\rho^2}{\sigma^2}} > 0 \text{ and } c_2 = \frac{\frac{\rho_0}{\sigma_0^2} - \frac{\rho}{\sigma^2}}{\frac{\rho_0^2}{\sigma_0^2} - \frac{\rho^2}{\sigma^2}} \geq 0, \tag{EC.50}$$

which contradicts to the assumption that W is regular.

(ii) Suppose that W is not regular and corresponds to $c_1 > 0$ and $c_2 \geq 0$ in (EC.32) and (EC.33). Suppose for the sake of contradiction that there exists $\theta \neq \theta_0$ such that (EC.37) holds. Then, by the same argument in case (i) and by the uniqueness of (c_1, c_2), $\sigma \neq \sigma_0$ and (ρ, σ) must satisfy the equations in (EC.50) and hence must be a solution to the following two equations

$$c_1 (\frac{\rho_0^2}{\sigma_0^2} - \frac{\rho^2}{\sigma^2}) = -\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2} \text{ and } c_2 (\frac{\rho_0^2}{\sigma_0^2} - \frac{\rho^2}{\sigma^2}) = \frac{\rho_0}{\sigma_0^2} - \frac{\rho}{\sigma^2}.$$

It can be shown by simple algebra that the above two equations are equivalent to

$$(c_2+c_1\rho_0)\rho^2-(c_1\rho_0^2+1)\rho-(c_2\rho_0^2-\rho_0)=0 \text{ and } \sigma^2=\frac{1+c_1\rho^2}{1+c_1\rho_0^2}\sigma_0^2. \tag{EC.51}$$

If $c_2+c_1\rho_0=0$, then the above system of equations has a unique solution (ρ_0,σ_0); otherwise, the above equations have two solutions (ρ_0,σ_0) and $(\frac{1-c_2\rho_0}{c_2+c_1\rho_0},\sigma_0\sqrt{\frac{c_2^2+c_1}{(c_2+c_1\rho_0)^2}})$.

(ii.1) Suppose that (EC.34) or (EC.35) holds, then the two equations in (EC.51) have a unique solution (ρ_0,σ_0), and hence the two equations in (EC.50) do not have a solution $(\rho,\sigma)\neq(\rho_0,\sigma_0)$, which leads to a contradiction.

(ii.2) Suppose that (EC.36) holds, then $(\frac{1-c_2\rho_0}{c_2+c_1\rho_0},\sigma_0\sqrt{\frac{c_2^2+c_1}{(c_2+c_1\rho_0)^2}})$ is the unique solution to (EC.50); hence, $\rho=\frac{1-c_2\rho_0}{c_2+c_1\rho_0}=\rho_*$ and $\sigma=\sigma_0\sqrt{\frac{c_2^2+c_1}{(c_2+c_1\rho_0)^2}}=\sigma_*$. Since $C_3(\theta)=C_3(\theta_0)$ and $C_4(\theta)=C_4(\theta_0)$, it follows that $\bar{\alpha}=\bar{\alpha}_*$ and $B=B_*$. Hence, $\theta=\theta_*$. However, $\theta\in\Theta$ but $\theta_*\notin\Theta$, which constitutes a contradiction.

Therefore, we have completed the proof of sufficiency. We will then show the necessity. Suppose for the sake of contradiction that W does not satisfy any of the conditions specified. Then, W is not regular, and it corresponds to a unique pair of $c_1>0$ and $c_2\geq 0$, and $\rho_0\neq -\frac{c_2}{c_1}$, and $\frac{1-c_2\rho_0}{c_2+c_1\rho_0}\neq\rho_0$, and $\theta_*\in\Theta$. Then, by the definition of θ_*, it holds that $\theta_*\neq\theta_0$ and $C_i(\theta_*)=C_i(\theta_0)$ for $i=1,3$, and 4, which further implies that $C_i(\theta_*)=C_i(\theta_0)$ for $i=1,2,\ldots,6$. Therefore, $l(\tilde{y}_t\mid\tilde{g}_t,\theta_*)=l(\tilde{y}_t\mid\tilde{g}_t,\theta_0)$, but this contradicts to that θ_0 is identifiable. □

Proposition EC.1 is equivalent to the following statement: A particular θ_0 is not identifiable if and only if W is not regular and it corresponds to the unique pair (c_1,c_2) with $c_1>0$ and $c_2\geq 0$ in (EC.32) and (EC.33), and $\rho_0\neq -\frac{c_2}{c_1}$, and $\rho_0\neq\frac{1-c_2\rho_0}{c_2+c_1\rho_0}$, and $\theta_*\in\Theta$.

EC.4.2. Asymptotic and Small Sample Properties of the Conditional MLE

We first prove the following proposition, which will be used to show the asymptotic properties of the conditional MLE.

Proposition EC.2. *Suppose Assumption EC.1 and Assumption EC.2 hold and θ_0 is identifiable. Define*

$$Q_0(\theta):=E[l(\tilde{y}_t\mid\tilde{g}_t,\theta)],\ \hat{Q}_T(\theta):=\frac{1}{T}\sum_{t=1}^{T}l(\tilde{y}_t\mid\tilde{g}_t,\theta). \tag{EC.52}$$

Then, $\hat{Q}_T(\theta)$ is twice continuously differentiable on the interior of Θ. Define

$$s(\tilde{y}_t,\tilde{g}_t;\theta):=\frac{\partial l(\tilde{y}_t\mid\tilde{g}_t,\theta)}{\partial\theta} \text{ and } H(\tilde{y}_t,\tilde{g}_t;\theta):=\frac{\partial^2 l(\tilde{y}_t\mid\tilde{g}_t,\theta)}{\partial\theta\partial\theta'}. \tag{EC.53}$$

Then, the following statements hold:

(i) $Q_0(\theta)$ is uniquely maximized at θ_0.

(ii) $\sup_{\theta\in\Theta}|\hat{Q}_T(\theta)-Q_0(\theta)|\xrightarrow{p}0$, *as* $T\to\infty$.

(iii) $Q_0(\theta)$ *is continuous on* Θ.

(iv) $E[s(\tilde{y}_t,\tilde{g}_t;\theta_0)]=0$.

(v) $H(\tilde{y}_t,\tilde{g}_t;\theta)$ *is equal to*

$$\begin{bmatrix} -\operatorname{tr}(W(I_n-\rho W)^{-1}W(I_n-\rho W)^{-1})-\frac{1}{\sigma^2}\tilde{y}_t'W'W\tilde{y}_t & -\frac{1}{\sigma^2}\tilde{y}_t'W'X_t & -\frac{1}{\sigma^4}\tilde{y}_t'W'\tilde{\xi}_t \\ -\frac{1}{\sigma^2}X_t'W\tilde{y}_t & -\frac{1}{\sigma^2}X_t'X_t & -\frac{1}{\sigma^4}X_t'\tilde{\xi}_t \\ -\frac{1}{\sigma^4}\tilde{\xi}_t'W\tilde{y}_t & -\frac{1}{\sigma^4}\tilde{\xi}_t'X_t & \frac{n}{2\sigma^4}-\frac{1}{\sigma^6}\tilde{\xi}_t'\tilde{\xi}_t \end{bmatrix}, \tag{EC.54}$$

where $\tilde{\xi}_t:=(I_n-\rho W)\tilde{y}_t-X_t\underline{b}$, *and* $\operatorname{tr}(\cdot)$ *denotes the matrix trace.*

(vi) $-E[H(\tilde{y}_t,\tilde{g}_t;\theta_0)]=E[s(\tilde{y}_t,\tilde{g}_t;\theta_0)s(\tilde{y}_t,\tilde{g}_t;\theta_0)']$.

(vii) $E[H(\tilde{y}_t,\tilde{g}_t;\theta_0)]$ *is invertible.*

(viii) *There is a neighborhood* $\mathcal{N}$ *of* θ_0 *such that* $E[\sup_{\theta\in\mathcal{N}}\|H(\tilde{y}_t,\tilde{g}_t;\theta)\|]<\infty$.

Proof. Since $\det((I_n-\rho W')(I_n-\rho W))$ is equal to a polynomial of ρ, it follows from (EC.38) that $\hat{Q}_T(\theta)$ is twice continuously differentiable on the interior of Θ. The proof of part (i) to (vi) is as follows.

(i) Let $f(\tilde{y}_t\mid\tilde{g}_t,\theta)$ denote the conditional density. It follows from the model (34) and part (i) of Assumption EC.1 that $E[\|\tilde{y}_t\|^2]<\infty$. Hence, (EC.38) implies that for any $\theta\in\Theta$, $E[|l(\tilde{y}_t\mid\tilde{g}_t,\theta)|]<\infty$. For any $\theta\neq\theta_0$, define $g(\tilde{y}_t,\tilde{g}_t):=\frac{f(\tilde{y}_t|\tilde{g}_t,\theta)}{f(\tilde{y}_t|\tilde{g}_t,\theta_0)}$. Since θ_0 is identifiable, it follows that $P(g(\tilde{y}_t,\tilde{g}_t)\neq 1)>0$. Therefore, it follows from the strict Jensen's inequality that

$$E[l(\tilde{y}_t\mid\tilde{g}_t,\theta_0)-l(\tilde{y}_t\mid\tilde{g}_t,\theta)]=E[-\log g(\tilde{y}_t,\tilde{g}_t)]>-\log E[g(\tilde{y}_t,\tilde{g}_t)]. \tag{EC.55}$$

Since

$$E[g(\tilde{y}_t,\tilde{g}_t)\mid\tilde{g}_t]=\int\frac{f(\tilde{y}_t\mid\tilde{g}_t,\theta)}{f(\tilde{y}_t\mid\tilde{g}_t,\theta_0)}f(\tilde{y}_t\mid\tilde{g}_t,\theta_0)d\tilde{y}_t=\int f(\tilde{y}_t\mid\tilde{g}_t,\theta)d\tilde{y}_t=1,$$

it follows that $E[g(\tilde{y}_t,\tilde{g}_t)]=1$, which in combination with (EC.55) implies that $Q_0(\theta)$ has a unique maximizer θ_0.

(ii) and (iii). We first show that

$$E[\sup_{\theta\in\Theta}|l(\tilde{y}_t\mid\tilde{g}_t;\theta)|]<\infty. \tag{EC.56}$$

By (EC.38), $l(\tilde{y}_t\mid\tilde{g}_t;\theta)=\sum_{i,j}a_{ij}(\theta)y_{it}y_{jt}+\sum_{i,j}b_{ij}(\theta)g_{it}g_{jt}+\sum_{i,j}c_{ij}(\theta)y_{it}g_{jt}+\sum_{i=1}^n d_i(\theta)y_{it}$ $+\sum_{i=1}^K e_i(\theta)g_{it}+C_6(\theta)$, where $a_{ij}(\cdot)$, $b_{ij}(\cdot)$, $c_{ij}(\cdot)$, $d_i(\cdot)$, $e_i(\cdot)$, and $C_6(\cdot)$ are all continuous functions. Since Θ is compact, it follows that $E[\sup_{\theta\in\Theta}|a_{ij}(\theta)y_{it}y_{jt}|]=E[|y_{it}y_{jt}|]\sup_{\theta\in\Theta}|a_{ij}(\theta)|<\infty$. Similarly, the expectation of the supremum (with respect to θ) of the absolute value of each term in the summation for $l(\tilde{y}_t\mid\tilde{g}_t;\theta)$ is finite; therefore, $E[\sup_{\theta\in\Theta}|l(\tilde{y}_t\mid\tilde{g}_t;\theta)|]<\infty$. Then, since $l(\tilde{y}_t\mid\tilde{g}_t,\theta)$

is continuous at every $\theta \in \Theta$, it follows from Lemma 2.4 in Neway and McFadden (1994, p. 2129) that (ii) and (iii) hold.

(iv) Define $\tilde{\xi}_t := \tilde{y}_t - \rho W \tilde{y}_t - X_t \underline{b}$. By Jacobi's formula of matrix calculus,

$$\frac{d}{d\rho}\det((I_n - \rho W')(I_n - \rho W)) = -2\det((I_n - \rho W')(I_n - \rho W))\,\mathrm{tr}((I_n - \rho W)^{-1}W),$$

where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. Hence,

$$\frac{\partial l(\tilde{y}_t \mid \tilde{g}_t, \theta)}{\partial \rho} = -\,\mathrm{tr}((I_n - \rho W)^{-1}W) + \frac{1}{\sigma^2}\tilde{\xi}_t' W \tilde{y}_t. \tag{EC.57}$$

By simple algebra, we have

$$\frac{\partial l(\tilde{y}_t \mid \tilde{g}_t, \theta)}{\partial \underline{b}} = \frac{1}{\sigma^2} X_t' \tilde{\xi}_t,\ \frac{\partial l(\tilde{y}_t \mid \tilde{g}_t, \theta)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\tilde{\xi}_t'\tilde{\xi}_t. \tag{EC.58}$$

Hence,

$$\begin{aligned}
E\left[\frac{\partial l(\tilde{y}_t \mid \tilde{g}_t, \theta_0)}{\partial \rho}\right] &= -\,\mathrm{tr}((I_n - \rho_0 W)^{-1}W) + \frac{1}{\sigma_0^2}E[\tilde{\xi}_t' W (I_n - \rho_0 W)^{-1}(\bar{\alpha}_0 + B_0 \tilde{g}_t + \tilde{\xi}_t)] \\
&= -\,\mathrm{tr}((I_n - \rho_0 W)^{-1}W) + \frac{1}{\sigma_0^2}E[\tilde{\xi}_t' W (I_n - \rho_0 W)^{-1}\tilde{\xi}_t] \\
&= -\,\mathrm{tr}((I_n - \rho_0 W)^{-1}W) + \frac{1}{\sigma_0^2}E[\mathrm{tr}(W (I_n - \rho_0 W)^{-1}\tilde{\xi}_t\tilde{\xi}_t')] \\
&= -\,\mathrm{tr}((I_n - \rho_0 W)^{-1}W) + \frac{1}{\sigma_0^2}\sigma_0^2\,\mathrm{tr}(W (I_n - \rho_0 W)^{-1}) = 0.
\end{aligned}$$

Also, by (EC.58),

$$E\left[\frac{\partial l(\tilde{y}_t \mid \tilde{g}_t, \theta_0)}{\partial \underline{b}}\right] = E\left[\frac{1}{\sigma_0^2}X_t'\tilde{\xi}_t\right] = 0,\ E\left[\frac{\partial l(\tilde{y}_t \mid \tilde{g}_t, \theta_0)}{\partial \sigma^2}\right] = E\left[-\frac{n}{2}\frac{1}{\sigma_0^2} + \frac{1}{2\sigma_0^4}\tilde{\xi}_t'\tilde{\xi}_t\right] = 0,$$

which completes the proof of part (iv).

(v) For brevity of notation, we use $l(\theta)$ to denote $l(\tilde{y}_t \mid \tilde{g}_t, \theta)$ in the sequel. By (EC.57) and (EC.58), we have

$$\frac{\partial^2 l(\theta)}{\partial \rho^2} = -\,\mathrm{tr}(W(I_n - \rho W)^{-1}W(I_n - \rho W)^{-1}) - \frac{1}{\sigma^2}\tilde{y}_t' W' W \tilde{y}_t, \tag{EC.59}$$

$$\frac{\partial^2 l(\theta)}{\partial \rho \partial \underline{b}} = -\frac{1}{\sigma^2}X_t' W \tilde{y}_t, \frac{\partial^2 l(\theta)}{\partial \rho \partial \sigma^2} = -\frac{1}{\sigma^4}\tilde{\xi}_t' W \tilde{y}_t, \frac{\partial^2 l(\theta)}{\partial \underline{b} \partial \underline{b}'} = -\frac{1}{\sigma^2}X_t' X_t, \tag{EC.60}$$

$$\frac{\partial^2 l(\theta)}{\partial \sigma^2 \partial \underline{b}} = -\frac{1}{\sigma^4}X_t'\tilde{\xi}_t, \frac{\partial^2 l(\theta)}{\partial \sigma^2 \partial \sigma^2} = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\tilde{\xi}_t'\tilde{\xi}_t, \tag{EC.61}$$

which completes the proof.

(vi) Let

$$C := W(I_n - \rho_0 W)^{-1} \tag{EC.62}$$

and let C_{ij} be the (i,j) element of C. Then by (EC.59), we have

$$\begin{aligned}
&E[\frac{\partial^2 l(\theta_0)}{\partial \rho^2} \mid \tilde{g}_t] \\
&= -\operatorname{tr}(C^2) - \frac{1}{\sigma_0^2} E[((I_n - \rho_0 W)^{-1}(X_t \underline{b}_0 + \tilde{\epsilon}_t))' W' W (I_n - \rho_0 W)^{-1}(X_t \underline{b}_0 + \tilde{\epsilon}_t) \mid \tilde{g}_t] \\
&= -\operatorname{tr}(C^2) - \frac{1}{\sigma_0^2} \underline{b}_0' X_t' C' C X_t \underline{b}_0 - \frac{1}{\sigma_0^2} E[\tilde{\epsilon}_t' C' C \tilde{\epsilon}_t] \\
&= -\operatorname{tr}(C^2) - \frac{1}{\sigma_0^2} \underline{b}_0' X_t' C' C X_t \underline{b}_0 - \operatorname{tr}(C'C).
\end{aligned} \tag{EC.63}$$

By (EC.57) and simple algebra,

$$\begin{aligned}
&E[\left(\frac{\partial l(\theta_0)}{\partial \rho}\right)^2 \mid \tilde{g}_t] = E[(-\operatorname{tr}(C) + \frac{1}{\sigma_0^2} \tilde{\epsilon}_t' W \tilde{y}_t)^2 \mid \tilde{g}_t] \\
&= E[(-\operatorname{tr}(C) + \frac{1}{\sigma_0^2} \tilde{\epsilon}_t' C (X_t \underline{b}_0 + \tilde{\epsilon}_t))^2 \mid \tilde{g}_t] \\
&= E[\operatorname{tr}(C)^2 - \frac{2\operatorname{tr}(C)}{\sigma_0^2} \tilde{\epsilon}_t' C (X_t \underline{b}_0 + \tilde{\epsilon}_t) \mid \tilde{g}_t] + E[\frac{1}{\sigma_0^4} \tilde{\epsilon}_t' C (X_t \underline{b}_0 + \tilde{\epsilon}_t)(X_t \underline{b}_0 + \tilde{\epsilon}_t)' C' \tilde{\epsilon}_t \mid \tilde{g}_t] \\
&= -\operatorname{tr}(C)^2 + \frac{1}{\sigma_0^2} \underline{b}_0' X_t' C' C X_t \underline{b}_0 + \frac{2}{\sigma_0^4} \underline{b}_0' X_t' C' E[\tilde{\epsilon}_t \tilde{\epsilon}_t' C \tilde{\epsilon}_t] + \frac{1}{\sigma_0^4} E[(\tilde{\epsilon}_t' C \tilde{\epsilon}_t)^2] \\
&= -\operatorname{tr}(C)^2 + \frac{1}{\sigma_0^2} \underline{b}_0' X_t' C' C X_t \underline{b}_0 + \frac{1}{\sigma_0^4} E[(\tilde{\epsilon}_t' C \tilde{\epsilon}_t)^2] \\
&= -\operatorname{tr}(C)^2 + \frac{1}{\sigma_0^2} \underline{b}_0' X_t' C' C X_t \underline{b}_0 + \frac{1}{\sigma_0^4} E[(\sum_{i=1}^n \sum_{j=1}^n C_{ij} \epsilon_{it} \epsilon_{jt})^2] \\
&= -(\sum_{i=1}^n C_{ii})^2 + \frac{1}{\sigma_0^2} \underline{b}_0' X_t' C' C X_t \underline{b}_0 + \frac{1}{\sigma_0^4} E[(\sum_{i=1}^n C_{ii} \epsilon_{it}^2 + \sum_{i<j} (C_{ij} + C_{ji}) \epsilon_{it} \epsilon_{jt})^2] \\
&= 2\sum_{i=1}^n C_{ii}^2 + \sum_{i \neq j} C_{ij}^2 + 2\sum_{i<j} C_{ij} C_{ji} + \frac{1}{\sigma_0^2} \underline{b}_0' X_t' C' C X_t \underline{b}_0 \\
&= -E[\frac{\partial^2 l(\theta_0)}{\partial \rho^2} \mid \tilde{g}_t]. \text{ (by (EC.63))}
\end{aligned}$$

By (EC.60),

$$E[\frac{\partial^2 l(\theta_0)}{\partial \rho \partial \underline{b}} \mid \tilde{g}_t] = -\frac{1}{\sigma_0^2} E[X_t' W (I_n - \rho_0 W)^{-1} (X_t \underline{b}_0 + \tilde{\epsilon}_t) \mid \tilde{g}_t] = -\frac{1}{\sigma_0^2} X_t' C X_t \underline{b}_0. \tag{EC.64}$$

By (EC.57) and (EC.58),

$$\begin{aligned}
&E[\frac{\partial l(\theta_0)}{\partial \rho} \frac{\partial l(\theta_0)}{\partial \underline{b}} \mid \tilde{g}_t] = -\frac{\operatorname{tr}(C)}{\sigma_0^2} E[X_t' \tilde{\epsilon}_t \mid \tilde{g}_t] + \frac{1}{\sigma_0^4} E[\tilde{\epsilon}_t' W \tilde{y}_t X_t' \tilde{\epsilon}_t \mid \tilde{g}_t] \\
&= \frac{1}{\sigma_0^4} E[\tilde{\epsilon}_t' C (X_t \underline{b}_0 + \tilde{\epsilon}_t) X_t' \tilde{\epsilon}_t \mid \tilde{g}_t] = \frac{1}{\sigma_0^4} E[X_t' \tilde{\epsilon}_t \tilde{\epsilon}_t' C X_t \underline{b}_0 \mid \tilde{g}_t] + \frac{1}{\sigma_0^4} E[\tilde{\epsilon}_t' C \tilde{\epsilon}_t X_t' \tilde{\epsilon}_t \mid \tilde{g}_t] \\
&= \frac{1}{\sigma_0^2} X_t' C X_t \underline{b}_0 = -E[\frac{\partial^2 l(\theta_0)}{\partial \rho \partial \underline{b}} \mid \tilde{g}_t]. \text{ (by (EC.64))}
\end{aligned}$$

By (EC.60),

$$E[\frac{\partial^2 l(\theta_0)}{\partial \rho \partial \sigma^2} \mid \tilde{g}_t] = -\frac{1}{\sigma_0^4} E[\tilde{\epsilon}_t' C (X_t \underline{b}_0 + \tilde{\epsilon}_t) \mid \tilde{g}_t] = -\frac{1}{\sigma_0^4} E[\tilde{\epsilon}_t' C \tilde{\epsilon}_t] = -\frac{1}{\sigma_0^4} \operatorname{tr}(E[\tilde{\epsilon}_t \tilde{\epsilon}_t' C])$$

$$= -\frac{1}{\sigma_0^2}\operatorname{tr}(C). \tag{EC.65}$$

By (EC.57) and (EC.58),

$$\begin{aligned}
&E[\frac{\partial l(\theta_0)}{\partial\rho}\frac{\partial l(\theta_0)}{\partial\sigma^2}\mid\tilde{g}_t]\\
&=\frac{n\operatorname{tr}(C)}{2\sigma_0^2}-\frac{\operatorname{tr}(C)}{2\sigma_0^4}E[\tilde{\epsilon}_t'\tilde{\epsilon}_t]-\frac{n}{2\sigma_0^4}E[\tilde{\epsilon}_t'W\tilde{y}_t\mid\tilde{g}_t]+\frac{1}{2\sigma_0^6}E[\tilde{\epsilon}_t'W\tilde{y}_t\tilde{\epsilon}_t'\tilde{\epsilon}_t\mid\tilde{g}_t]\\
&=-\frac{n}{2\sigma_0^4}E[\tilde{\epsilon}_t'C\tilde{\epsilon}_t]+\frac{1}{2\sigma_0^6}E[\tilde{\epsilon}_t'C(X_t\underline{b}_0+\tilde{\epsilon}_t)\tilde{\epsilon}_t'\tilde{\epsilon}_t\mid\tilde{g}_t]=-\frac{n}{2\sigma_0^2}\operatorname{tr}(C)+\frac{1}{2\sigma_0^6}\operatorname{tr}(CE[\tilde{\epsilon}_t\tilde{\epsilon}_t'\tilde{\epsilon}_t\tilde{\epsilon}_t'])\\
&=-\frac{n}{2\sigma_0^2}\operatorname{tr}(C)+\frac{1}{2\sigma_0^6}(n+2)\sigma_0^4\operatorname{tr}(C)=\frac{1}{\sigma_0^2}\operatorname{tr}(C)=-E[\frac{\partial^2 l(\theta_0)}{\partial\rho\partial\sigma^2}\mid\tilde{g}_t].\ \text{(by (EC.65))}
\end{aligned}\tag{EC.66}$$

By (EC.58),

$$E[\frac{\partial l(\theta_0)}{\partial\underline{b}}\frac{\partial l(\theta_0)}{\partial\underline{b}'}\mid\tilde{g}_t]=\frac{1}{\sigma_0^4}E[X_t'\tilde{\epsilon}_t\tilde{\epsilon}_t'X_t\mid\tilde{g}_t]=\frac{1}{\sigma_0^2}X_t'X_t=-E[\frac{\partial^2 l(\theta_0)}{\partial\underline{b}\partial\underline{b}'}\mid\tilde{g}_t],$$

where the last equality follows from (EC.60). By (EC.58),

$$\begin{aligned}
&E[\frac{\partial l(\theta_0)}{\partial\sigma^2}\frac{\partial l(\theta_0)}{\partial\underline{b}}\mid\tilde{g}_t]=-\frac{n}{2\sigma_0^4}E[X_t'\tilde{\epsilon}_t\mid\tilde{g}_t]+\frac{1}{2\sigma_0^6}E[\tilde{\epsilon}_t'\tilde{\epsilon}_tX_t'\tilde{\epsilon}_t\mid\tilde{g}_t]\\
&=-\frac{n}{2\sigma_0^4}X_t'E[\tilde{\epsilon}_t\mid\tilde{g}_t]+\frac{1}{2\sigma_0^6}E[X_t'\tilde{\epsilon}_t\tilde{\epsilon}_t'\tilde{\epsilon}_t\mid\tilde{g}_t]=\frac{1}{2\sigma_0^6}X_t'E[\tilde{\epsilon}_t\tilde{\epsilon}_t'\tilde{\epsilon}_t]\\
&=0=-E[\frac{\partial^2 l(\theta_0)}{\partial\sigma^2\partial\underline{b}}\mid\tilde{g}_t],
\end{aligned}$$

where the last equality follows from (EC.61). At last, by (EC.58),

$$\begin{aligned}
&E[\left(\frac{\partial l(\theta_0)}{\partial\sigma^2}\right)^2\mid\tilde{g}_t]=E[\left(-\frac{n}{2\sigma_0^2}+\frac{1}{2\sigma_0^4}\tilde{\epsilon}_t'\tilde{\epsilon}_t\right)^2\mid\tilde{g}_t]\\
&=E[\frac{n^2}{4\sigma_0^4}+\frac{1}{4\sigma_0^8}\tilde{\epsilon}_t'\tilde{\epsilon}_t\tilde{\epsilon}_t'\tilde{\epsilon}_t-\frac{n}{2\sigma_0^6}\tilde{\epsilon}_t'\tilde{\epsilon}_t]=\frac{n}{2\sigma_0^4}=-E[\frac{\partial^2 l(\theta_0)}{\partial\sigma^2\partial\sigma^2}\mid\tilde{g}_t],
\end{aligned}$$

where the last equality follows from (EC.61). Hence, we have shown that $-E[H(\tilde{y}_t,\tilde{g}_t;\theta_0)\mid\tilde{g}_t]=E[s(\tilde{y}_t,\tilde{g}_t;\theta_0)s(\tilde{y}_t,\tilde{g}_t;\theta_0)'\mid\tilde{g}_t]$, which completes the proof of (vi).

(vii) Suppose for the sake of contradiction that $E[H(\tilde{y}_t,\tilde{g}_t;\theta_0)]$ is not invertible, then there exists $a=(a_1,a_2',a_3)'\in\mathbb{R}^{2+n(K+1)}$, $a\neq 0$, such that

$$0=a'E[H(\tilde{y}_t,\tilde{g}_t;\theta_0)]a=-E[(a's(\tilde{y}_t,\tilde{g}_t;\theta_0))^2],$$

where the last equality follows from (vi). This implies that $a's(\tilde{y}_t,\tilde{g}_t;\theta_0)=0, a.s.$ Denote $a_2=(v_1,u_{11},u_{12},\ldots,u_{1K},\ldots,v_n,u_{n1},u_{n2},\ldots,u_{nK})'$, $v=(v_1,\ldots,v_n)'$, $U=(u_{ij})$. Let C be defined in (EC.62). Then, we have

$$0=a's(\tilde{y}_t,\tilde{g}_t;\theta_0)=-a_1\operatorname{tr}(C)-\frac{na_3}{2\sigma_0^2}+\frac{a_1}{\sigma_0^2}\tilde{\epsilon}_t'C\bar{\alpha}+\frac{a_1}{\sigma_0^2}\tilde{\epsilon}_t'CB_0\tilde{g}_t+\frac{a_1}{\sigma_0^2}\tilde{\epsilon}_t'C\tilde{\epsilon}_t$$

$$+\frac{1}{\sigma_0^2}v'\tilde{\epsilon}_t+\frac{1}{\sigma_0^2}\tilde{\epsilon}_t'U\tilde{g}_t+\frac{a_3}{2\sigma_0^4}\tilde{\epsilon}_t'\tilde{\epsilon}_t,\ a.s. \tag{EC.67}$$

It follows from (EC.67) and part (ii) of Assumption EC.1 that

$$\begin{aligned}0=&-a_1\operatorname{tr}(C)-\frac{na_3}{2\sigma_0^2}+\frac{a_1}{\sigma_0^2}\tilde{\epsilon}'C\bar{\alpha}_0+\frac{a_1}{\sigma_0^2}\tilde{\epsilon}'CB\tilde{g}+\frac{a_1}{\sigma_0^2}\tilde{\epsilon}'C\tilde{\epsilon}\\&+\frac{1}{\sigma_0^2}v'\tilde{\epsilon}+\frac{1}{\sigma_0^2}\tilde{\epsilon}'U\tilde{g}+\frac{a_3}{2\sigma_0^4}\tilde{\epsilon}'\tilde{\epsilon},\ \text{for any } (\tilde{\epsilon}',\tilde{g}')'\in\mathbb{R}^n\times\mathbb{A}.\end{aligned} \tag{EC.68}$$

By taking partial derivatives with respect to $(\tilde{\epsilon}',\tilde{g}')'$ on both sides of (EC.68), we obtain that

$$-a_1\operatorname{tr}(C)-\frac{na_3}{2\sigma_0^2}=0,\ \frac{a_1}{\sigma_0^2}C\bar{\alpha}_0+\frac{1}{\sigma_0^2}v=0,\ \frac{a_1}{\sigma_0^2}CB_0+\frac{1}{\sigma_0^2}U=0, \tag{EC.69}$$

$$\frac{a_1}{\sigma_0^2}C+\frac{a_3}{2\sigma_0^4}I_n=0. \tag{EC.70}$$

There are two cases:

Case 1: $a_1=0$. Then, it follows from (EC.69) that $a_3=0$, $v=0$, and $U=0$, which contradict to $a\neq 0$.

Case 2: $a_1\neq 0$. Then, it follows from (EC.70) that $C=-\frac{a_3}{2\sigma_0^2a_1}I_n$, which in combination with (EC.62) implies that $W(1-\frac{a_3\rho_0}{2\sigma_0^2a_1})=-\frac{a_3}{2\sigma_0^2a_1}I_n$. If $a_3=0$, then $W=0$, which contradicts to Assumption EC.2; if $a_3\neq 0$, then $1-\frac{a_3\rho_0}{2\sigma_0^2a_1}\neq 0$, and $W=\frac{-a_3}{2\sigma_0^2a_1-a_3\rho_0}I_n$, which contradicts to that the diagonal elements of W are zero (Assumption EC.2). Hence, $E[H(\tilde{y}_t,\tilde{g}_t;\theta_0)]$ is invertible.

(viii) Let $\mathcal{N}$ be any neighborhood of θ_0 that lies in the interior of Θ. We have

$$\begin{aligned}E[\sup_{\theta\in\mathcal{N}}\|H(\tilde{y}_t,\tilde{g}_t;\theta)\|]\leq{}& E[\sup_{\theta\in\mathcal{N}}|\frac{\partial^2 l(\theta)}{\partial\rho^2}|]+E[\sup_{\theta\in\mathcal{N}}\|\frac{\partial^2 l(\theta)}{\partial\rho\partial\underline{b}}\|]+E[\sup_{\theta\in\mathcal{N}}|\frac{\partial^2 l(\theta)}{\partial\rho\partial\sigma_2}|]\\&+E[\sup_{\theta\in\mathcal{N}}\|\frac{\partial^2 l(\theta)}{\partial\underline{b}\partial\underline{b}'}\|]+E[\sup_{\theta\in\mathcal{N}}\|\frac{\partial^2 l(\theta)}{\partial\sigma^2\partial\underline{b}}\|]+E[\sup_{\theta\in\mathcal{N}}|\frac{\partial^2 l(\theta)}{\partial\sigma^2\partial\sigma^2}|].\end{aligned} \tag{EC.71}$$

We only need to show that each term on the right side of (EC.71) is finite. Let $C(\rho):=W(I_n-\rho W)^{-1}$, then $C(\rho)$ is continuous. By (EC.59), (EC.60), and (EC.61), we have

$$E[\sup_{\theta\in\mathcal{N}}|\frac{\partial^2 l(\theta)}{\partial\rho^2}|]\leq\sup_{\theta\in\mathcal{N}}|\operatorname{tr}(C(\rho)^2)|+E[\tilde{y}_t'W'W\tilde{y}_t]\sup_{\theta\in\mathcal{N}}\frac{1}{\sigma^2}, \tag{EC.72}$$

$$E[\sup_{\theta\in\mathcal{N}}\|\frac{\partial^2 l(\theta)}{\partial\rho\partial\underline{b}}\|]\leq E[\|X_t'W\tilde{y}_t\|]\sup_{\theta\in\mathcal{N}}\frac{1}{\sigma^2}, \tag{EC.73}$$

$$\begin{aligned}&E[\sup_{\theta\in\mathcal{N}}|\frac{\partial^2 l(\theta)}{\partial\rho\partial\sigma^2}|]=E[\sup_{\theta\in\mathcal{N}}|\frac{1}{\sigma^4}\tilde{\xi}_t'W\tilde{y}_t|]\\&\leq E[\sup_{\theta\in\mathcal{N}}|\frac{1}{\sigma^4}\tilde{y}_t'W\tilde{y}_t|]+E[\sup_{\theta\in\mathcal{N}}|\frac{\rho}{\sigma^4}\tilde{y}_t'W'W\tilde{y}_t|]+E[\sup_{\theta\in\mathcal{N}}|\frac{1}{\sigma^4}\bar{\alpha}'W\tilde{y}_t|]+E[\sup_{\theta\in\mathcal{N}}|\frac{1}{\sigma^4}\tilde{g}_t'B'W\tilde{y}_t|]\\&\leq E[|\tilde{y}_t'W\tilde{y}_t|]\sup_{\theta\in\mathcal{N}}\frac{1}{\sigma^4}+E[\tilde{y}_t'W'W\tilde{y}_t]\sup_{\theta\in\mathcal{N}}\frac{|\rho|}{\sigma^4}+E[\|W\tilde{y}_t\|]\sup_{\theta\in\mathcal{N}}\frac{1}{\sigma^4}\|\bar{\alpha}\|\\&\quad+E[\|\tilde{g}_t\|^2]\sup_{\theta\in\mathcal{N}}\frac{1}{2\sigma^4}\|B\|^2+E[\|W\tilde{y}_t\|^2]\sup_{\theta\in\mathcal{N}}\frac{1}{2\sigma^4},\end{aligned} \tag{EC.74}$$

$$E[\sup_{\theta\in\mathcal{N}}\|\frac{\partial^2 l(\theta)}{\partial\underline{b}\partial\underline{b}'}\|]\leq E[\|X_t'X_t\|]\sup_{\theta\in\mathcal{N}}\frac{1}{\sigma^2}, \tag{EC.75}$$

$$\begin{aligned}
&E[\sup_{\theta\in\mathcal{N}}\|\frac{\partial^2 l(\theta)}{\partial\sigma^2\partial\underline{b}}\|]=E[\sup_{\theta\in\mathcal{N}}\|\frac{1}{\sigma^4}X_t'\tilde{\xi}_t\|]\\
\leq\, & E[\sup_{\theta\in\mathcal{N}}\|\frac{1}{\sigma^4}X_t'\tilde{y}_t\|]+E[\sup_{\theta\in\mathcal{N}}\|\frac{\rho}{\sigma^4}X_t'W\tilde{y}_t\|]+E[\sup_{\theta\in\mathcal{N}}\|\frac{1}{\sigma^4}X_t'\bar{\alpha}\|]+E[\sup_{\theta\in\mathcal{N}}\|\frac{1}{\sigma^4}X_t'B\tilde{g}_t\|]\\
\leq\, & E[\|X_t'\tilde{y}_t\|]\sup_{\theta\in\mathcal{N}}\frac{1}{\sigma^4}+E[\|X_t'W\tilde{y}_t\|]\sup_{\theta\in\mathcal{N}}\frac{|\rho|}{\sigma^4}+E[\|X_t\|]\sup_{\theta\in\mathcal{N}}\frac{\|\bar{\alpha}\|}{\sigma^4}+E[\|X_t'\|\|\tilde{g}_t\|]\sup_{\theta\in\mathcal{N}}\frac{\|B\|}{\sigma^4}\\
\leq\, & E[\|X_t'\tilde{y}_t\|]\sup_{\theta\in\mathcal{N}}\frac{1}{\sigma^4}+E[\|X_t'W\tilde{y}_t\|]\sup_{\theta\in\mathcal{N}}\frac{|\rho|}{\sigma^4}+E[\|X_t\|]\sup_{\theta\in\mathcal{N}}\frac{\|\bar{\alpha}\|}{\sigma^4}\\
&+\frac{1}{2}(E[\|X_t'\|^2]+E[\|\tilde{g}_t\|^2])\sup_{\theta\in\mathcal{N}}\frac{\|B\|}{\sigma^4},
\end{aligned} \tag{EC.76}$$

$$\begin{aligned}
&E[\sup_{\theta\in\mathcal{N}}|\frac{\partial^2 l(\theta)}{\partial\sigma^2\partial\sigma^2}|]=E[\sup_{\theta\in\mathcal{N}}|\frac{n}{2\sigma^4}-\frac{1}{\sigma^6}\tilde{\xi}_t'\tilde{\xi}_t|]\\
=\, & E[\sup_{\theta\in\mathcal{N}}|\frac{n}{2\sigma^4}-\frac{1}{\sigma^6}\|(I_n-\rho W)\tilde{y}_t-\bar{\alpha}-B\tilde{g}_t\|^2|]\\
\leq\, & E[\sup_{\theta\in\mathcal{N}}(\frac{n}{2\sigma^4}+\frac{16}{\sigma^6}(\|\tilde{y}_t\|^2+\rho^2\|W\tilde{y}_t\|^2+\|\bar{\alpha}\|^2+\|B\|^2\|\tilde{g}_t\|^2))]\\
\leq\, & \sup_{\theta\in\mathcal{N}}(\frac{n}{2\sigma^4}+\frac{16\|\bar{\alpha}\|^2}{\sigma^6})+E[\|\tilde{y}_t\|^2]\sup_{\theta\in\mathcal{N}}\frac{16}{\sigma^6}+E[\|W\tilde{y}_t\|^2]\sup_{\theta\in\mathcal{N}}\frac{16\rho^2}{\sigma^6}+E[\|\tilde{g}_t\|^2]\sup_{\theta\in\mathcal{N}}\frac{16\|B\|^2}{\sigma^6}.
\end{aligned} \tag{EC.77}$$

Since Θ is compact, all the supremums on the right-hand side of (EC.72)-(EC.77) are finite. Furthermore, by part (i) of Assumption EC.1, $\tilde{g}_t$ and hence $\tilde{y}_t$ have finite second moments. Thus, all the expectations on the right-hand side of (EC.72)-(EC.77) are finite. Therefore, each term on the right-hand side of (EC.71) is finite, which completes the proof. □

THEOREM EC.2. *(Asymptotic properties of the conditional MLE) Suppose the conditions in Proposition EC.2 hold. Then, the conditional MLE $\hat{\theta}:=(\hat{\rho},\hat{\underline{b}},\widehat{\sigma^2})$ has consistency and asymptotic normality:*

$$(i)\ \hat{\theta}\xrightarrow{p}\theta_0,\ \ as\ T\to\infty, \tag{EC.78}$$

$$(ii)\ \sqrt{T}(\hat{\theta}-\theta_0)\xrightarrow{d}N(0,-E[H(\tilde{y}_t,\tilde{g}_t;\theta_0)]^{-1}),\ \ as\ T\to\infty, \tag{EC.79}$$

$$(iii)\ \frac{1}{T}\sum_{t=1}^{T}H(\tilde{y}_t,\tilde{g}_t;\hat{\theta})\xrightarrow{p}E[H(\tilde{y}_t,\tilde{g}_t;\theta_0)],\ \ as\ T\to\infty, \tag{EC.80}$$

where $H(\tilde{y}_t,\tilde{g}_t;\theta)$ is equal to (EC.54).

Proof. By (i), (ii), and (iii) of Proposition EC.2 and the compactness of Θ, it follows from Theorem 2.1 in Neway and McFadden (1994) that (EC.78) holds.

We will show the asymptotic normality (EC.79) by applying Proposition 7.9 in Hayashi (2000, p. 475). The consistency of $\hat{\theta}$ has been proved in above. The condition (1) of Proposition 7.9 holds by the assumption that θ_0 lies in the interior of Θ. The conditions (2), (3), (4), and (5) of Proposition

7.9 follow from Proposition EC.2 of this paper. Hence, all the conditions of Proposition 7.9 hold and its conclusion implies (EC.79).

At last, we will show that (EC.80) holds. Define $\hat{H}(\theta) := \frac{1}{T}\sum_{t=1}^{T} H(\tilde{y}_t, \tilde{g}_t; \theta)$. Let $\mathcal{N}$ be a neighborhood such that (viii) in Proposition EC.2 holds and let $\Theta_0 \subset \mathcal{N}$ be a compact set that contains θ_0. Then, it follows from (viii) in Proposition EC.2 and Lemma 2.4 in Neway and McFadden (1994, p. 2129) that $H(\theta) := E[H(\tilde{y}_t, \tilde{g}_t; \theta)]$ is continuous and

$$\sup_{\theta\in\Theta_0}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 l(\tilde{y}_t \mid \tilde{g}_t, \theta)}{\partial\theta\partial\theta'} - H(\theta)\right\| \xrightarrow{p} 0, \text{ as } T\to\infty. \tag{EC.81}$$

Since $\hat{\theta} \xrightarrow{p} \theta_0$ and $H(\theta)$ is continuous, it follows that $H(\hat{\theta}) \xrightarrow{p} H(\theta_0)$. For any $\varepsilon > 0$,

$$\begin{aligned}
&P(\|\hat{H}(\hat{\theta}) - H(\theta_0)\| > \varepsilon) \le P(\|\hat{H}(\hat{\theta}) - H(\hat{\theta})\| + \|H(\hat{\theta}) - H(\theta_0)\| > \varepsilon)\\
&\le P(\|\hat{H}(\hat{\theta}) - H(\hat{\theta})\| > \frac{\varepsilon}{2}, \hat{\theta}\in\Theta_0) + P(\|\hat{H}(\hat{\theta}) - H(\hat{\theta})\| > \frac{\varepsilon}{2}, \hat{\theta}\notin\Theta_0)\\
&\quad + P(\|H(\hat{\theta}) - H(\theta_0)\| > \frac{\varepsilon}{2})\\
&\le P(\sup_{\theta\in\Theta_0}\|\hat{H}(\theta) - H(\theta)\| > \frac{\varepsilon}{2}) + P(\hat{\theta}\notin\Theta_0) + P(\|H(\hat{\theta}) - H(\theta_0)\| > \frac{\varepsilon}{2})\\
&\to 0, \text{ as } T\to\infty,
\end{aligned}$$

where the limit follows from (EC.81), $\hat{\theta} \xrightarrow{p} \theta_0$, and $H(\hat{\theta}) \xrightarrow{p} H(\theta_0)$. □

The asymptotic properties of the conditional MLE of S-APT model (34) are obtained by letting $T\to\infty$ and keeping n fixed; in contrast, those of the SAR model are obtained by letting $n\to\infty$. As a result, the conditional MLE of the S-APT model has a $\sqrt{T}$-rate of convergence as long as W satisfies the identifiability condition specified in Proposition EC.1, but those of the SAR may not have the desired $\sqrt{n}$-rate of convergence when W is not sparse enough; see Lee (2004).[19]

We also investigate the finite-sample performance of the estimators using 2000 data sets simulated from the model (34). In all the simulation studies, we use the locations of twenty major cities in the United States as asset locations and W is defined by the method of Delaunay triangularization, which is commonly adopted in spatial econometrics literature.[20] It is easy to check that the specified matrix W is regular;[21] hence, by Proposition EC.1, the true parameter is identifiable.

We specify $\bar{\alpha}_0 = 0$ and $\sigma_0^2 = 0.5$. An i.i.d. draw of 20 samples from $N(0,1)$ is fixed as the elements of B_0. $\{\tilde{g}_t : t = 1,\ldots,131\}$ is generated as one realization of 131 i.i.d. random variables with distribution $N(0.5, 0.5)$. For the fixed B_0 and $\{\tilde{g}_t : t = 1,\ldots,131\}$, 2000 i.i.d. samples of $\{\tilde{\epsilon}_t : t = 1,\ldots,131\}$

[19] More precisely, Lee (2004) shows that when each asset can be influenced by many neighbors, various components of the estimators may have different rates of convergence.

[20] The twenty cities correspond to the twenty MSAs that have S&P/Case-Shiller home price indices; their locations are specified by their geographic coordinates. See Lesage and Pace (2009, Chap. 4.11) for the details of the method of Delaunay triangularization. We use the program fdelw2 in the Spatial Statistics Toolbox (Pace (2003)) to compute W by this method.

[21] This is because the sum of square of different columns of W are not equal.

are then simulated and $\{\tilde{y}_t : t = 1, \ldots, 131\}$ are then computed from (34). Then, the conditional MLE $\hat{\rho}$ is obtained from each of the simulated data sets.

Table EC.1 shows the mean and standard deviation of the conditional MLE $\hat{\rho}$ for the 2000 simulated data sets for different values of $\rho_0 = 0.2$, 0.4, 0.6, and 0.8, respectively. Figure EC.1 shows the histogram of 2000 estimates $\hat{\rho}$ for different ρ_0, which seems to indicate that $\hat{\rho}$ has an asymptotic normal distribution with mean ρ_0.

ρ_0	0.2	0.4	0.6	0.8
mean of $\hat{\rho}$	0.199	0.399	0.599	0.799
(theoretical) asymptotic standard deviation of $\hat{\rho}$	0.028	0.025	0.019	0.011
empirical standard deviation of $\hat{\rho}$	0.031	0.029	0.020	0.013

Table EC.1 The mean and standard deviation of $\hat{\rho}$. The asymptotic standard deviations are estimated from the sample average of Hessian matrix (see Eq. (EC.54) and (EC.80)).

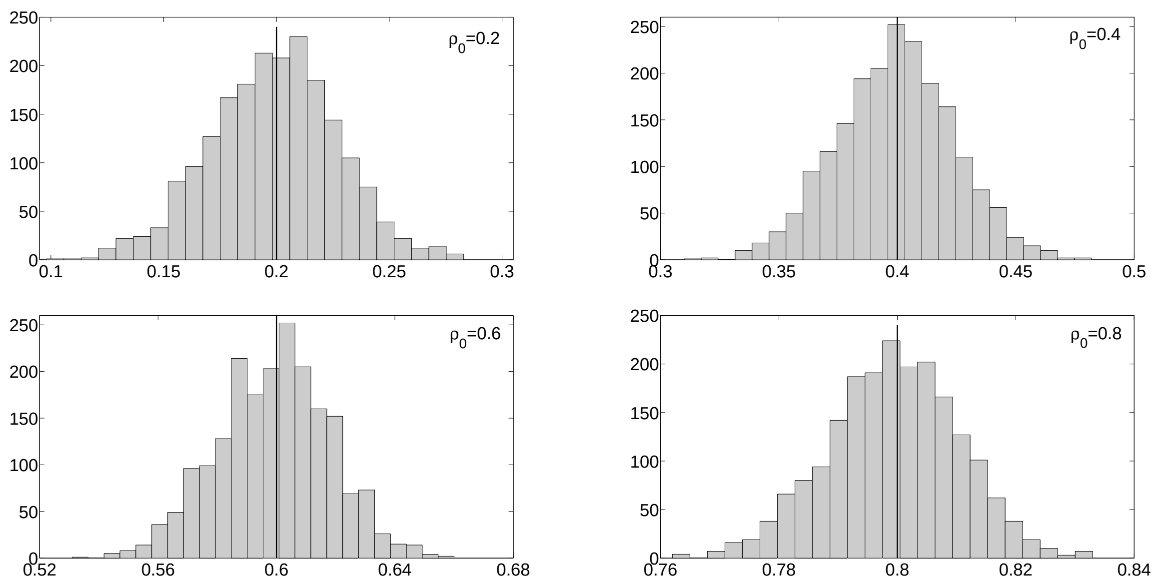

Figure EC.1 Histogram of the conditional MLE $\hat{\rho}$ for 2000 data sets simulated from the model (34) for different values of ρ_0 with $n = 20$, $K = 1$, and $T = 131$.

EC.4.3. Simulation Studies for Testing S-APT and Estimating Adjusted R^2

Using the likelihood ratio test statistic defined in (41), we test the zero-intercept constraint of S-APT for 10000 simulated data sets at the confidence level of 95%.[22] The size of the test is 5.91%,

[22] The test data include the 8000 data sets used for Table EC.1 and additional 2000 data sets simulated in the same way with $\rho_0 = 0.5$.

which is slightly higher than the theoretical value of 5%. This may result from small sample bias, as discussed in Campbell, Lo, and MacKinlay (1996, Chap. 5.4).

To show the effectiveness of the adjusted R^2, two data sets are simulated according to the same model specification as that in Table EC.1, except that ρ_0 is fixed at 0.5, and two values of σ_0^2 (0.01 and 0.5) are used, respectively, for the two data sets, which correspond to the two cases of high and low adjusted R^2. In the simulation, the factor realization $\{\tilde{g}_t, t=1,\dots,T\}$ is first simulated and fixed. Then, for each chosen value of σ_0^2, the residuals $\{\tilde{\epsilon}_t, t=1,\dots,T\}$ are simulated and the realized returns $\{\tilde{y}_t, t=1,\dots,T\}$ are calculated according to the model (34). For each simulated data set, we calculate the conditional MLE estimate θ^* under the constraint $\bar{\alpha}_0=0$ and obtain the fitted residual series $\{\hat{\epsilon}_t=\tilde{y}_t-\rho^*W\tilde{y}_t-B^*\tilde{g}_t : t=1,\dots,T\}$, where ρ^* and B^* are the conditional MLE. The sample adjusted R^2 of y_i is computed and compared to the theoretical adjusted R^2 of y_i. Table EC.2 shows that the sample adjusted R^2 and the theoretical adjusted R^2 align well.

	$\sigma_0^2=0.01$									
	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9	r_{10}
theoretical adjusted R^2	0.7763	0.9818	0.0910	0.4064	0.9654	0.9730	0.9690	0.3210	0.8415	0.3110
sample adjusted R^2	0.7748	0.9817	0.0848	0.4023	0.9652	0.9728	0.9688	0.3162	0.8404	0.3062
	r_{11}	r_{12}	r_{13}	r_{14}	r_{15}	r_{16}	r_{17}	r_{18}	r_{19}	r_{20}
theoretical adjusted R^2	0.1833	0.9203	0.8952	0.9906	0.6367	0.6634	0.9672	0.0451	0.2236	0.9394
sample adjusted R^2	0.1777	0.9197	0.8945	0.9905	0.6342	0.6611	0.9670	0.0385	0.2183	0.9390
	$\sigma_0^2=0.5$									
	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9	r_{10}
theoretical adjusted R^2	0.2328	0.4256	0.0007	0.1308	0.4691	0.3521	0.4452	0.0859	0.0567	0.0002
sample adjusted R^2	0.2294	0.4230	-0.0037	0.1269	0.4667	0.3491	0.4427	0.0817	0.0525	-0.0043
	r_{11}	r_{12}	r_{13}	r_{14}	r_{15}	r_{16}	r_{17}	r_{18}	r_{19}	r_{20}
theoretical adjusted R^2	0.1003	0.2366	0.1816	0.6552	0.1691	0.0235	0.4598	-0.0705	0.0390	0.1301
sample adjusted R^2	0.0961	0.2331	0.1779	0.6536	0.1654	0.0191	0.4574	-0.0754	0.0347	0.1262

Table EC.2 Simulation study of the sample adjusted R^2. We use the same model specification as that for Table EC.1, except that ρ_0 is fixed at 0.5 and two values of σ_0^2 (0.01 and 0.5) are used, respectively, for the two data sets. For each data set, the conditional MLE of parameters is estimated for the model (34) under the constraint $\bar{\alpha}_0=0$ and then the sample adjusted R^2 for each element of $\tilde{r}$ is calculated and compared to its theoretical counterpart. It appears that the sample adjusted R^2 and its theoretical counterpart align well.

EC.5. Spatial Weight Matrices, S&P Credit Ratings and Parameter Estimates of S-APT models

The spatial weight matrix W that is used in Section 6 and calculated using driving distance is

$$\begin{pmatrix} 0.000 & 0.080 & 0.082 & 0.116 & 0.210 & 0.083 & 0.069 & 0.127 & 0.125 & 0.049 & 0.059 \\ 0.041 & 0.000 & 0.036 & 0.236 & 0.095 & 0.026 & 0.075 & 0.050 & 0.359 & 0.036 & 0.046 \\ 0.132 & 0.113 & 0.000 & 0.147 & 0.142 & 0.067 & 0.077 & 0.080 & 0.121 & 0.057 & 0.064 \\ 0.074 & 0.295 & 0.058 & 0.000 & 0.087 & 0.031 & 0.086 & 0.064 & 0.182 & 0.052 & 0.072 \\ 0.176 & 0.157 & 0.074 & 0.114 & 0.000 & 0.051 & 0.070 & 0.079 & 0.183 & 0.043 & 0.052 \\ 0.146 & 0.090 & 0.073 & 0.086 & 0.107 & 0.000 & 0.066 & 0.198 & 0.088 & 0.067 & 0.078 \\ 0.084 & 0.179 & 0.059 & 0.164 & 0.102 & 0.046 & 0.000 & 0.074 & 0.154 & 0.063 & 0.075 \\ 0.148 & 0.114 & 0.058 & 0.117 & 0.110 & 0.131 & 0.071 & 0.000 & 0.100 & 0.066 & 0.085 \\ 0.069 & 0.388 & 0.041 & 0.157 & 0.120 & 0.027 & 0.070 & 0.047 & 0.000 & 0.035 & 0.045 \\ 0.068 & 0.097 & 0.048 & 0.113 & 0.071 & 0.052 & 0.071 & 0.078 & 0.088 & 0.000 & 0.315 \\ 0.069 & 0.107 & 0.046 & 0.132 & 0.072 & 0.052 & 0.073 & 0.085 & 0.095 & 0.268 & 0.000 \end{pmatrix} \tag{EC.82}$$

The spatial weight matrix W that is used in Section 6 and calculated using geographic distance is

$$\begin{pmatrix} 0.000 & 0.116 & 0.074 & 0.103 & 0.203 & 0.083 & 0.063 & 0.139 & 0.113 & 0.046 & 0.059 \\ 0.065 & 0.000 & 0.036 & 0.226 & 0.092 & 0.029 & 0.077 & 0.051 & 0.344 & 0.035 & 0.045 \\ 0.129 & 0.113 & 0.000 & 0.097 & 0.168 & 0.075 & 0.092 & 0.084 & 0.124 & 0.055 & 0.063 \\ 0.074 & 0.291 & 0.040 & 0.000 & 0.087 & 0.037 & 0.098 & 0.069 & 0.178 & 0.053 & 0.073 \\ 0.184 & 0.149 & 0.087 & 0.110 & 0.000 & 0.054 & 0.073 & 0.082 & 0.168 & 0.042 & 0.052 \\ 0.149 & 0.091 & 0.077 & 0.091 & 0.106 & 0.000 & 0.067 & 0.182 & 0.088 & 0.067 & 0.081 \\ 0.075 & 0.163 & 0.062 & 0.162 & 0.096 & 0.044 & 0.000 & 0.067 & 0.167 & 0.077 & 0.087 \\ 0.166 & 0.108 & 0.058 & 0.115 & 0.107 & 0.121 & 0.067 & 0.000 & 0.098 & 0.068 & 0.093 \\ 0.070 & 0.376 & 0.043 & 0.152 & 0.113 & 0.030 & 0.086 & 0.050 & 0.000 & 0.035 & 0.044 \\ 0.067 & 0.090 & 0.046 & 0.107 & 0.067 & 0.054 & 0.095 & 0.083 & 0.083 & 0.000 & 0.307 \\ 0.072 & 0.099 & 0.044 & 0.124 & 0.070 & 0.055 & 0.090 & 0.096 & 0.088 & 0.260 & 0.000 \end{pmatrix}. \tag{EC.83}$$

The spatial weight matrix for 10 CSI indices futures that is used in Section 7 and calculated using driving distance is

$$\begin{pmatrix} 0.000 & 0.468 & 0.152 & 0.057 & 0.022 & 0.021 & 0.029 & 0.019 & 0.211 & 0.021 \\ 0.509 & 0.000 & 0.125 & 0.058 & 0.023 & 0.024 & 0.030 & 0.021 & 0.188 & 0.022 \\ 0.293 & 0.221 & 0.000 & 0.088 & 0.039 & 0.037 & 0.052 & 0.036 & 0.196 & 0.038 \\ 0.137 & 0.129 & 0.110 & 0.000 & 0.083 & 0.067 & 0.139 & 0.071 & 0.187 & 0.078 \\ 0.034 & 0.034 & 0.032 & 0.054 & 0.000 & 0.088 & 0.129 & 0.207 & 0.037 & 0.385 \\ 0.072 & 0.074 & 0.064 & 0.095 & 0.189 & 0.000 & 0.143 & 0.133 & 0.077 & 0.154 \\ 0.067 & 0.065 & 0.063 & 0.134 & 0.192 & 0.099 & 0.000 & 0.137 & 0.077 & 0.167 \\ 0.032 & 0.031 & 0.030 & 0.048 & 0.215 & 0.064 & 0.096 & 0.000 & 0.035 & 0.449 \\ 0.307 & 0.252 & 0.148 & 0.113 & 0.034 & 0.033 & 0.048 & 0.031 & 0.000 & 0.033 \\ 0.027 & 0.027 & 0.026 & 0.043 & 0.325 & 0.060 & 0.095 & 0.366 & 0.030 & 0.000 \end{pmatrix}. \tag{EC.84}$$

The spatial weight matrix W for 10 CSI indices futures that is used in Section 7 and calculated using geographic distance is

$$\begin{pmatrix} 0.000 & 0.456 & 0.147 & 0.061 & 0.022 & 0.022 & 0.029 & 0.020 & 0.222 & 0.021 \\ 0.488 & 0.000 & 0.119 & 0.065 & 0.024 & 0.024 & 0.031 & 0.021 & 0.205 & 0.022 \\ 0.272 & 0.206 & 0.000 & 0.099 & 0.039 & 0.036 & 0.051 & 0.035 & 0.226 & 0.037 \\ 0.140 & 0.139 & 0.122 & 0.000 & 0.078 & 0.067 & 0.126 & 0.066 & 0.191 & 0.071 \\ 0.034 & 0.035 & 0.032 & 0.053 & 0.000 & 0.086 & 0.133 & 0.201 & 0.038 & 0.388 \\ 0.072 & 0.074 & 0.065 & 0.098 & 0.183 & 0.000 & 0.142 & 0.134 & 0.077 & 0.155 \\ 0.067 & 0.068 & 0.063 & 0.128 & 0.197 & 0.099 & 0.000 & 0.138 & 0.077 & 0.164 \\ 0.032 & 0.033 & 0.031 & 0.048 & 0.213 & 0.067 & 0.099 & 0.000 & 0.035 & 0.442 \\ 0.298 & 0.258 & 0.163 & 0.113 & 0.033 & 0.031 & 0.045 & 0.029 & 0.000 & 0.031 \\ 0.028 & 0.028 & 0.026 & 0.042 & 0.332 & 0.062 & 0.095 & 0.357 & 0.030 & 0.000 \end{pmatrix}. \tag{EC.85}$$

The spatial weight matrix W for 9 CSI indices futures (excluding New York) that is used in Section 7 and calculated using driving distance is

$$\begin{pmatrix} 0.000 & 0.478 & 0.156 & 0.058 & 0.022 & 0.022 & 0.029 & 0.020 & 0.216 \\ 0.520 & 0.000 & 0.128 & 0.060 & 0.024 & 0.024 & 0.031 & 0.021 & 0.193 \\ 0.305 & 0.230 & 0.000 & 0.091 & 0.041 & 0.038 & 0.054 & 0.037 & 0.204 \\ 0.148 & 0.140 & 0.119 & 0.000 & 0.090 & 0.073 & 0.150 & 0.077 & 0.203 \\ 0.055 & 0.055 & 0.052 & 0.088 & 0.000 & 0.142 & 0.211 & 0.337 & 0.060 \\ 0.085 & 0.087 & 0.076 & 0.112 & 0.223 & 0.000 & 0.169 & 0.157 & 0.092 \\ 0.080 & 0.078 & 0.076 & 0.161 & 0.230 & 0.118 & 0.000 & 0.165 & 0.093 \\ 0.057 & 0.056 & 0.055 & 0.087 & 0.390 & 0.116 & 0.175 & 0.000 & 0.063 \\ 0.318 & 0.261 & 0.153 & 0.117 & 0.035 & 0.034 & 0.050 & 0.032 & 0.000 \end{pmatrix}. \tag{EC.86}$$

The spatial weight matrix W for 9 CSI indices futures (excluding New York) that is used in Section 7 and calculated using geographic distance is

$$\begin{pmatrix} 0.000 & 0.466 & 0.150 & 0.063 & 0.023 & 0.022 & 0.030 & 0.020 & 0.227 \\ 0.499 & 0.000 & 0.121 & 0.067 & 0.024 & 0.025 & 0.032 & 0.022 & 0.210 \\ 0.282 & 0.214 & 0.000 & 0.103 & 0.040 & 0.038 & 0.053 & 0.036 & 0.234 \\ 0.150 & 0.150 & 0.131 & 0.000 & 0.084 & 0.072 & 0.136 & 0.071 & 0.206 \\ 0.056 & 0.057 & 0.053 & 0.087 & 0.000 & 0.140 & 0.217 & 0.328 & 0.062 \\ 0.085 & 0.088 & 0.077 & 0.116 & 0.216 & 0.000 & 0.168 & 0.159 & 0.091 \\ 0.080 & 0.081 & 0.075 & 0.153 & 0.236 & 0.118 & 0.000 & 0.165 & 0.092 \\ 0.058 & 0.058 & 0.056 & 0.085 & 0.382 & 0.120 & 0.177 & 0.000 & 0.063 \\ 0.308 & 0.266 & 0.169 & 0.116 & 0.034 & 0.032 & 0.046 & 0.030 & 0.000 \end{pmatrix}. \tag{EC.87}$$

The S&P credit ratings of the 11 eurozone countries during 2001–2013 are shown in Table EC.3.

The S&P credit ratings of the states where the ten MSAs are located during 2006–2013 are shown in Table EC.4.

The parameter estimates of the intercepts α in unit of percent and the factor loading matrix B under the S-APT model 4s for the European stock indices returns (see Section 6.2), with cor-

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Austria	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AA+	AA+
Belgium	AA+	AA+	AA+	AA+	AA+	AA+	AA+	AA+	AA+	AA+	AA	AA	AA
Finland	AA+	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA
France	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AA+	AA+	AA+
Germany	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA
Greece	A	A	A+	A+	A	A	A	A	A-	BBB+/BB+	BB/B/CCC/CC	SD/CCC	B-
Ireland	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AA+/AA	AA-/A	A-/BBB+	BBB+
Italy	AA	AA	AA	AA	AA-	AA-	A+	A+	A+	A	BBB+	BBB+	BBB
Netherlands	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AA+
Portugal	AA	AA	AA	AA	AA-	AA-	AA-	AA-	A+	A-	BBB-	BB	BB
Spain	AA+	AA+	AA+	AA+	AAA	AAA	AAA	AAA	AA+	AA	AA-	A/BBB+/BBB-	BBB-

Table EC.3 The S&P credit ratings of the 11 eurozone countries during 2001–2013.

	2013	2012	2011	2010	2009	2008	2007	2006
Florida	AAA	AAA	AAA	AAA	AAA	AAA	AAA	AAA
Nevada	AA	AA	AA	AA+	AA+	AA+	AA+	AA+
Massachusetts	AA+	AA+	AA	AA	AA	AA	AA	AA
New York	AA	AA	AA	AA	AA	AA	AA	AA
Colorado	AA	AA	AA	AA	AA	AA	AA	AA-
Illinois	A-	A+	A+	A+	A+	A	A	A
California	A	A-	A-	A-	A	A+	A+	A+

Table EC.4 The S&P credit ratings of the states where the ten MSAs are located during 2006–2013.

responding standard errors (standard deviations of the estimates) in parentheses, are reported in Eq. (EC.88). The estimate of variance of residuals $\widehat{\sigma^2}$ is 9.4577 (0.0261), in unit of $percent^2$.

$$(\hat{\alpha}, \hat{B}) = \left(\begin{array}{cccccc} & \hat{\alpha} & MKV & SMB & MoM & Credit \\ Austria & 0.2897(0.0213) & 1.1757(0.0058) & 0.6209(0.0104) & 0.2897(0.0050) & -0.3342(0.0058) \\ Belgium & 0.0503(0.0058) & 0.9506(0.0213) & 0.1039(0.0058) & 0.0503(0.0103) & -0.1684(0.0050) \\ Finland & -0.1803(0.0050) & 0.9902(0.0058) & -0.3077(0.0213) & -0.1803(0.0059) & -0.0389(0.0103) \\ France & -0.0327(0.0103) & 0.9782(0.0050) & -0.3702(0.0058) & -0.0327(0.0213) & 0.0162(0.0058) \\ Germany & -0.1289(0.0058) & 1.0007(0.0103) & -0.1241(0.0050) & -0.1289(0.0058) & 0.5561(0.0213) \\ Greece & -0.2231(0.0213) & 1.2267(0.0059) & 0.1149(0.0103) & -0.2231(0.0050) & -0.7719(0.0059) \\ Ireland & -0.0732(0.0059) & 0.8797(0.0213) & 0.3734(0.0058) & -0.0732(0.0103) & -0.1513(0.0050) \\ Italy & -0.1449(0.0050) & 1.0554(0.0058) & -0.4576(0.0213) & -0.1449(0.0058) & -0.3635(0.0103) \\ Netherland & 0.0119(0.0.0103) & 1.0220(0.0050) & -0.0929(0.0058) & 0.0119(0.0213) & 0.1051(0.0059) \\ Portugal & -0.0680(0.0059) & 0.8422(0.0103) & 0.0433(0.0050) & -0.0680(0.0059) & -0.4295(0.0213) \\ Spain & -0.1118(0.0213) & 1.0084(0.0058) & -0.6441(0.0103) & -0.1118(0.0050) & -0.5120(0.0058) \end{array}\right) \tag{EC.88}$$

The parameter estimates of the intercepts α in unit of percent and the factor loading matrix B under the S-APT model for the S&P/Case-Shiller home price indices futures returns (see Section 7.2), with corresponding standard errors (standard deviations of the estimates) in parentheses, are reported in Eq. (EC.89). The estimate of variance of residuals $\widehat{\sigma^2}$ is 1.3720 (0.0067), in unit of

$percent^2$.

$$(\hat{\alpha}, \hat{B}) = \begin{pmatrix} & \hat{\alpha} & CS10f & CS10fTr & Credit \\ LosAngeles & -0.2099(0.0140) & 0.4724(0.0463) & -0.2099(0.0425) & 0.2170(0.0036) \\ SanDiego & -0.1624(0.0036) & 0.0612(0.0140) & -0.1624(0.0464) & 0.1809(0.0425) \\ SanFrancisco & 0.0833(0.0425) & 0.4398(0.0036) & 0.0833(0.0140) & 0.1916(0.0463) \\ Denver & 0.2355(0.0463) & 0.1939(0.0425) & 0.2335(0.0036) & 0.0570(0.0140) \\ WashingtonDC & 0.1887(0.0140) & 0.0883(0.0464) & 0.1887(0.0425) & 0.1595(0.0036) \\ Miami & -0.3032(0.0036) & 0.7042(0.0140) & -0.3032(0.0464) & -0.0735(0.0425) \\ Chicago & -0.7817(0.0425) & 1.7215(0.0036) & -0.7817(0.0140) & -0.1247(0.0464) \\ Boston & 0.1948(0.0464) & 0.1619(0.0425) & 0.1948(0.0036) & 0.0378(0.0140) \\ LasVegas & -0.0673(0.0140) & 0.4620(0.0463) & -0.0673(0.0425) & -0.3199(0.0036) \\ NewYork & -0.2263(0.0036) & 0.5321(0.0140) & -0.2263(0.0464) & -0.0247(0.0425) \end{pmatrix} \tag{EC.89}$$