Econometric Analysis of High Dimensional VARs Featuring a Dominant Unit

Alexander Chudik (1) Hashem Pesaran (2)
(1) CIMF and ECB
(2) Cambridge University, CIMF and USC

Prepared for presentation at First French Econometrics Confernece in Honour of Alian Monfort, Toulouse School of Economics, 14-15 December 2009

Outline

- Related literature
  - Weak vs strong cross section dependence
- IVAR Model with dominant unit
- Estimation and inference
- Monte Carlo experiments
- An empirical application

Motivation

- VAR models play a very important role in empirical macroeconomics and finance
- VAR theory is well developed when the number of variables N is small (6-7) and the number of time periods, T, is large
- For many empirical application such as global macroeconomic modelling, modelling of regions, firms, households, financial markets,... this theoretical framework is not appropriate
- Panel data models focus on small T and resolve the curse of dimensionality by dealing with homogeneous slope models with no cross section error dependence. See, panel VAR in Binder Hsiao and Pesaran (2005, ET)
- Our aim is to develop a theory for the analysis of VARs when both N and T are large

Related Literature: dynamics and cross section dependence

- Dealing with the **curse of dimensionality**:
  - (i) Data shrinkage (e.g. along the lines of index models or as in GVAR approach) and
  - (ii) Shrinkage of the parameter space (e.g. in form of Bayesian priors or spatial weights matrices).
- Dealing with **cross section dependence (CD)**:
  - (i) Spatial processes
  - (ii) Factor structures

- Spatial processes were pioneered by Whittle (1954), and later by Cliff and Ord (1973). Anselin(1988), Kelejian and Robinson (1995), Kelejian and Prucha (1999, 2004, 2007), and Lee (2002, 2004, 2007) have made important contributions in econometrics.
- Factor models were introduced by Hotelling (1933), applied in economics by Stone (1947), and recently extensively in finance and macroeconomics (Chamberlain and Rothschild 1983; Connor and Korajczyk, 1993; Kapetanios and Pesaran, 2007, Forni and Reichlin, 1998; Stock and Watson, 2002).
- Chudik, Pesaran and Tosetti (2009) distinguish between weak and strong, which will be useful in the analysis below.
- Chudik and Pesaran (2009) consider an IVAR model without a dominant unit. We extend this paper and allow for a dominant unit.

Weak and Strong Cross Section Dependence

- Understanding concepts of weak and strong CS dependence is important for the analysis of IVAR models.
- Let $\mathbf{z}_t = (z_{1t}, ..., z_{Nt})'$, with $E\left(\mathbf{z}_t \left| \mathcal{I}_{t-1}\right.\right) = \mathbf{0}$, $Var\left(\mathbf{z}_t \left| \mathcal{I}_{t-1}\right.\right) = \mathbf{\Sigma}_t$, where $\mathcal{I}_{t-1}$ is the information set at time $t-1$, and for each t where $\mathbf{\Sigma}_t$ has diagonal elements $0 < \sigma_{ii,t} \leq K$, for $i = 1, 2, ..., N$.
- Let $\mathbf{w}_t = (w_{1t}, ..., w_{Nt})'$ be a vector of weights satisfying the *granularity conditions*

$$\left\|\mathbf{w}_t\right\|_2 = O\left(N^{-\frac{1}{2}}\right), \qquad \frac{w_{jt}}{\left\|\mathbf{w}_t\right\|_2} = O\left(N^{-\frac{1}{2}}\right) \quad \text{for any } j \leq N \tag{1}$$

An obvious example is equal weights, $w_i = N^{-1}$.

- The process $\{z_{it}\}$ is *weakly cross sectionally dependent* (CWD) at a point in time t, if for *all* $\mathbf{w}_t$

$$\lim_{N\to\infty} Var(\mathbf{w}_t'\mathbf{z}_t \,|\mathcal{I}_{t-1}) = 0$$

The process $\{z_{it}\}$ is *cross sectionally strongly dependent* (CSD) at a point in time t, if there exists $\mathbf{w}_t$ such that

$$Var(\mathbf{w}_t'\mathbf{z}_t \,|\mathcal{I}_{t-1}) \geq K > 0$$

where K is a constant independent of N.

- The process $\{z_{it}\}$ is CSD at time $t \in \mathcal{T}$ if and only if

$$\lim_{N\to\infty} \frac{1}{N}\lambda_1\left(\mathbf{\Sigma}_t\right) = K > 0,$$

i.e. $\lambda_1\left(\mathbf{\Sigma}_t\right)$ increases to infinity at the rate N.

- If $\lambda_1\left(\mathbf{\Sigma}_t\right) = O(N^{1-\epsilon})$ for any $\epsilon > 0$, then

$$\lim_{N\to\infty} \left(\mathbf{w}_t'\mathbf{w}_t\right)\lambda_1\left(\mathbf{\Sigma}_t\right) = 0,$$

and the underlying process will be CWD. Hence, the bounded eigenvalue condition is *sufficient but not necessary* for CWD.

- CWD and CSD can be defined equally with respect to any information set, such as $\mathcal{I}_{-M}$, for any fixed M, or as M tends to infinity (if the underlying process is stationary).

Stationary Infinite Dimensional VAR Model

Suppose that for each $N \in \mathbb{N}$, vector of N endogenous variables $\mathbf{x}_{(N),t} = \left(x_{(N),1t}, ..., x_{(N),Nt}\right)'$ is given by the following VAR model,

$$\mathbf{x}_{(N),t} = \mathbf{\Phi}_{(N)}\mathbf{x}_{(N),t-1} + \mathbf{u}_{(N),t}, \tag{2}$$

where $\mathbf{\Phi}_{(N)}$ is $N \times N$ matrix of coefficients, $\mathbf{u}_{(N),t}$ is $N \times 1$ vector of error terms given by

$$\mathbf{u}_{(N),t} = \delta_{(N)}u_{1t} + \mathbf{e}_{(N),t}, \text{ and } \ \mathbf{e}_{(N),t} = \mathbf{R}_{(N)}\boldsymbol{\varepsilon}_{(N),t}, \tag{3}$$

where $\boldsymbol{\varepsilon}_{(N),t} = (0, \varepsilon_{2t}, ..., \varepsilon_{Nt})'$, and the individual elements of the double index array $\{\varepsilon_{jt}, j \in \{2, 3, ...\}, t \in \mathbb{Z}\}$ are *IID*.

- The results that follow hold irrespective of whether the parameters $\phi_{(N),ij}$, $\delta_{(N),i}$, and $r_{(N),ij}$ are assumed to be varying with N or not.
- Note that even if individual coefficients do not change with N, covariance between individual units, $cov\left(x_{(N),it}, x_{(N),jt}\right)$ in general must be changing with N, unless matrices $\mathbf{\Phi}_{(N)}$ and $\mathbf{R}_{(N)}$ are lower triangular.
- Dependence on N is suppressed in the remainder of this presentation to simplify the notations, but it is understood that the parameters and the dimension of the random variables $\mathbf{x}_t$ and $\mathbf{u}_t$ vary with N, unless otherwise stated.

Dealing with the Curse of Dimensionality

- The equation for unit i in system (2) is

$$x_{it} = \sum_{j=1}^{N} \phi_{ij} x_{j,t-1} + u_{it}.$$

- We shall assume that the following absolute summability condition holds,

$$\sum_{j=1}^{N} \left| \phi_{ij} \right| < K \text{ for any } N \in \mathbb{N} \text{ and any } i \in \{1, .., N\}. \quad (4)$$

- This is not a restrictive condition, and is similar to the concept of absolutely summability in the time series literature, but the summing is over lagged values of cross section units as opposed to AR terms of unit i, or MA terms.

- Similar constraint is used in 'data mining' literature, in particular Lasso (Tibshirani, 1996) and Ridge regression shrinkage method. Using Lasso around 5-10% of the regression coefficients end up being exactly equal to zero.
- The Lasso minimizes residual sum of squares subject to $\sum_{j=1}^{N} \left|\phi_{ij}\right| \leq K$.
- Ridge regression minimizes residual sum of squares subject to $\sum_{j=1}^{N} \phi_{ij}^2 \leq K$.
- We do not choose the value for K, as it is necessary in shrinkage methods, but rather assume only its existence instead.

- For each i we divide the units into a finite number of 'neighbors', denoted by $\mathcal{N}_i$, that have fixed coefficients $\left|\phi_{ij}\right| < K$ for $j \in \mathcal{N}_i$, and nonneighbors $j \in \mathcal{N}_i^c \equiv \{1, .., N\} \setminus \mathcal{N}_i$ that have coefficients that are of order $1/N$:

$$
\begin{aligned}
x_{it} &= \underbrace{\sum_{j \in \mathcal{N}_i} \phi_{ij} x_{j,t-1}}_{Neighbors} + \underbrace{\sum_{j \in \mathcal{N}_i^c} \phi_{ij} x_{j,t-1}}_{Nonneighbors} + u_{it} \\
&= \mathring{\boldsymbol{\phi}}_i' \mathbf{x}_{t-1} + \tilde{\boldsymbol{\phi}}_i' \mathbf{x}_{t-1} + u_{it},
\end{aligned}
$$

$\mathring{\boldsymbol{\phi}}_i$ and $\tilde{\boldsymbol{\phi}}_i$ are obtained from $\boldsymbol{\phi}_i$ by replacing nonneighbors and neighbors coefficients with zeros.

- If it is possible to divide units in this way, then we no longer have dimensionality problem.
- Remark: This assumption could be relaxed without the loss of generality by assuming more general restrictions in form of spatial weights matrices as in Chudik and Pesaran (2009)
- Consider aggregate spatiotemporal impact of nonneighbors given by $\tilde{\boldsymbol{\phi}}_i' \mathbf{x}_{t-1}$.
- Notation: We use $\|.\|_\infty$ and $\|.\|_1$ to denote maximum absolute row and column sum matrix norms, respectively.
- Note that $\left\|\tilde{\boldsymbol{\phi}}_i\right\|_c = \sum_{j \in \mathcal{N}_i^c} \left|\phi_{ij}\right| < K$. Nonneighbors could have large aggregate impact on unit i.

- However, the euclidean norm $\left\|\tilde{\boldsymbol{\phi}}_i\right\| \leq \sqrt{\left\|\tilde{\boldsymbol{\phi}}_i\right\|_\infty \left\|\tilde{\boldsymbol{\phi}}_i\right\|_1} = \sqrt{O\left(\frac{1}{N}\right) O\left(1\right)} = O\left(N^{-\frac{1}{2}}\right)$ and it follows that $\tilde{\boldsymbol{\phi}}_i' \mathbf{x}_{t-1} \overset{q.m.}{\to} 0$ if and only if $\{x_{it}\}$ is weakly CS dependent.
- If $\{x_{it}\}$ is strongly CS dependent process then $\lim_{N\to\infty} Var\left(\tilde{\boldsymbol{\phi}}_i' \mathbf{x}_{t-1}\right)$ is not necessarily zero.
- Hence the aggregate spatiotemporal impact of nonneighbors is nonnegligible and generally important only in the case of strong CS dependence.
- This paper allows for dominant unit, which is a source of strong CS dependence.

Assumptions

- To simplify the exposition, we consider only the case where $\mathcal{N}_i = \{1, i\}$, that is only lagged dominant unit and own lagged coefficient are $O(1)$, and the remaining elements are nonneighbors:

Assumption 1: (*Influence of unit* 1 *on the rest of the system is unrestricted.*) There exists a constant $K < \infty$ (independent of N) such that $|\phi_{ii}| < K$, $|\phi_{i1}| < K$, and

$$\left\|\boldsymbol{\phi}_{-1,-i}\right\|_\infty < \frac{K}{N}, \text{ for any } i, N \in \mathbb{N},$$

where $\boldsymbol{\phi}_{-1,-i} = \left(0, \phi_{i2}, ..., \phi_{i,i-1}, 0, \phi_{i,i+1}, .., \phi_{iN}\right)'$.

Assumption 2 (Stationarity and bounded variances as $N \to \infty$) There exists a constant $0 < \rho < 1 - \epsilon$ (independent of N) such that for any $N \in \mathbb{N}$:

$$\lambda_1 \left(\mathbf{\Phi} \right) < \rho, \ \left\| \mathbf{\Phi}_{(-1)} \right\|_\infty < \rho, \ \left\| \mathbf{\Phi}_{(-1)} \right\| < \rho, \text{ and } \left\| \boldsymbol{\phi}_1 \right\|_\infty < \rho,$$

where $\mathbf{\Phi}_{(-1)}$ is $N \times N$ matrix constructed from matrix $\mathbf{\Phi}$ by replacing its first column with a zero vector, and $\boldsymbol{\phi}_1$ is the first column vector of matrix $\mathbf{\Phi}$.

Assumption 3 (Starting values) Available observations are $\mathbf{x}_0, \mathbf{x}_1, ..., \mathbf{x}_T$ with the starting values $\mathbf{x}_0 = \sum_{\ell=0}^{\infty} \mathbf{\Phi}^{\ell} \mathbf{u}(-\ell)$.
Assumption 4 (Errors) Vector of errors $\mathbf{u}_t$ is given by (3), $\delta_1 = 1$, $\|\delta\|_{\infty} = O(1)$, u_{1t} and the individual elements of the double index array $\{\varepsilon_{jt}, j \in \{2, 3, ...\}, t \in \mathbb{Z}\}$ are identically and independently distributed with mean 0, unit variances and finite fourth moments uniformly bounded in $j \in \mathbb{N}$. Furthermore, matrix $\mathbf{R}$ has bounded row and column matrix norms.

Large N representation

- Under Assumptions $\lambda_1(\mathbf{\Phi}) < 1$ and Assumptions 3-4 (only), we have (for any $N \in \mathbb{N}$):

$$\begin{aligned} \mathbf{x}_t &= \mathbf{\Phi}_{-1}\mathbf{x}_{t-1} + \boldsymbol{\phi}_1 x_{1,t-1} + \boldsymbol{\delta} u_{1t} + \mathbf{e}_t, \\ &= \sum_{\ell=0}^{\infty} \mathbf{\Phi}_{-1}^{\ell} \boldsymbol{\phi}_1 x_{1,t-1-\ell} + \sum_{\ell=0}^{\infty} \mathbf{\Phi}_{-1}^{\ell} \boldsymbol{\delta} u_{1,t-1} + \boldsymbol{v}_t \end{aligned}$$

  where $\boldsymbol{v}_t = \sum_{\ell=0}^{\infty} \mathbf{\Phi}_{(-1)}^{\ell} \mathbf{R} \boldsymbol{\varepsilon}_{t-\ell}$ is CWD, in particular $\mathbf{w}'\boldsymbol{v}_t = O_p\left(N^{-\frac{1}{2}}\right)$.

- As a consequence, we have (for any i)

$$x_{it} = \mathbf{s}_i'\mathbf{x}_t = d_i(L)\, x_{1,t-1} + b_i(L)\, u_{1t} + v_{it}, \tag{5}$$

  where $d_i(L) = \mathbf{s}_i' \sum_{\ell=0}^{\infty} \mathbf{\Phi}_{-1}^{\ell} \boldsymbol{\phi}_1 L^{\ell}$, and $b_i(L) = \mathbf{s}_i' \sum_{\ell=0}^{\infty} \mathbf{\Phi}_{-1}^{\ell} \boldsymbol{\delta} L^{\ell}$.

- It follows that for $i = 1$,

$$\left(1 - d_1\left(L\right)L\right)x_{1t} = b_i\left(L\right)u_{1t} + \upsilon_{1t}, \tag{6}$$

  and for $i > 1$ it can be shown that

$$x_{it} = \phi_{ii}x_{i,t-1} + \beta_i\left(L\right)x_{1t} + e_{it} + O_p\left(N^{-\frac{1}{2}}\right), \tag{7}$$

  where in general the polynomial $\beta_i\left(L\right)$ is a function of all elements of $\mathbf{\Phi}$ and δ.
- Focus is on the consistent estimation of the unit-specific unknown coefficients ϕ_{ii} for $i > 1$ and also on the estimation of the impact of the dominant unit on the remaining units in the system, captured by $\beta_i\left(L\right)$.
- This paper does not deal with the pooled estimation of the mean coefficients over the cross section units.

- We allow for cross section dependence of innovations $\mathbf{e}_t = \mathbf{R}\varepsilon_t$, as characterized by matrix $\mathbf{R}$, but we ignore this matrix for estimation and therefore our estimators are not necessarily efficient in the presence of cross section dependence induced by the matrix $\mathbf{R}$.
- Based on the asymptotic representation of unit i, we consider the following auxiliary regression for $i > 1$:

$$\begin{aligned} x_{it} &= \phi_{ii} x_{i,t-1} + \sum_{\ell=0}^{k} \beta_{i\ell} x_{1,t-\ell} + \epsilon_{it} \\ &= \mathbf{g}_{it}' \boldsymbol{\pi}_i + \epsilon_{it}, \end{aligned} \tag{8}$$

where $\boldsymbol{\pi}_i = (\phi_{ii}, \beta_{i0}, ..., \beta_{ik})'$.

- Identification requires $\mathbf{C}_i = E\left(\mathbf{g}_{it}\mathbf{g}_{it}'\right)$ to be positive definite.

Theorems

Theorem 1 (Consistency) Under Assumptions 1-4, and invertibility of $\frac{1}{T-k}\sum_{t=k+1}^{T}\mathbf{g}_{it}\mathbf{g}_{it}'$, we have

$$\left\|\widehat{\boldsymbol{\pi}}_i-\boldsymbol{\pi}_i\right\|_{\infty}\xrightarrow{p}0\text{, for any }i>1,$$

as $N,T\xrightarrow{j}\infty$ at any order, and $k^2/T\rightarrow 0$ such that there exists constants $r_1, r_2>0$ satisfying $k>r_1T^{r_2}$.

- Remark: The number of regressors in the auxiliary regressions cannot go to infinity too fast, so that it is possible to satisfactorily estimate all coefficients, and not too slow, so that the omitted variable problem is asymptotically negligible.

Theorem 2. (Inference) Under assumptions of Theorem 1, for any sequence of $(k+1) \times 1$ dimensional vectors $\mathbf{a}_k$ such that $\|\mathbf{a}_k\| = 1$ and $\|\mathbf{a}_k\|_1 = O(1)$, and as $N, T \xrightarrow{j} \infty$, $T/N \to \varkappa < \infty$ ($\varkappa \geq 0$ is not necessarily nonzero), and $k^2/T \to 0$ such that there exists constants $r_1, r_2 > 0$ satisfying $k > r_1 T^{r_2}$, we have

$$\sqrt{T-k}\frac{1}{\sigma_i}\mathbf{a}_k' \mathbf{C}_i^{\frac{1}{2}} \left(\widehat{\boldsymbol{\pi}}_i - \boldsymbol{\pi}_i\right) \xrightarrow{d} N(0,1)\,, \text{ for } i > 1, \tag{9}$$

where $\widehat{\boldsymbol{\pi}}_i$ is LS estimator of $\boldsymbol{\pi}_i$ in regression (8), matrix $\mathbf{C}_i = E\left(\mathbf{g}_{it}\mathbf{g}_{it}'\right)$ can be consistently estimated by $\widehat{\mathbf{C}}_i = \frac{1}{T-k}\sum_{t=k+1}^{T} \mathbf{g}_{it}\mathbf{g}_{it}'$, and $\sigma_i^2 = Var\left(e_{it}\right)$.

Monte Carlo Simulations

Monte Carlo Design

- We consider factor augmented IVAR model featuring dominant unit,

$$\left(\mathbf{x}_t - \gamma f_t\right) = \mathbf{\Phi}\left(\mathbf{x}_{t-1} - \gamma f_{t-1}\right) + \mathbf{u}_t,$$

$$\mathbf{u}_t = \delta u_{1t} + \mathbf{e}_t, \text{ and } \mathbf{e}_t = \mathbf{R}\varepsilon_t.$$

- $f_t = 0.9 f_{t-1} + \varepsilon_{ft}$, and $\varepsilon_{ft} \sim IIDN\left(0, 1 - 0.9^2\right)$.
- We need to generate coefficient matrix $\mathbf{\Phi}$, errors $\mathbf{u}_t$, and factor loadings γ.

Construction of Coefficient Matrix

- Neighbor for unit i is unit 1 (dominant unit), unit $i+1$, and unit i itself.
- We generate $O_p(N^{-1})$ random variables λ_{ij} such that $\sum_{j=1}^{N} \lambda_{ij} = 1$ and $\lambda_1 = \lambda_i = \lambda_{i+1} = 0$ first.
- Matrix $\mathbf{\Phi}$ is then constructed as follows.
  - (Dominant unit $i=1$) $\phi_{11} = 0.7$, $\phi_{12} = 0.1$, and $\phi_{1j} = \alpha_1 \lambda_{1j}$ for $j = 3, .., N$, with $\alpha_1 = 0.1$.
  - (Unit $i=2$) $\phi_{21} = 0.1$, $\phi_{22} = 0.5$, $\phi_{23} = 0.1$, and $\phi_{2j} = \alpha_2 \lambda_{2j}$ for $j = 4, .., N$, with $\alpha_2 = 0.1$.
  - (Remaining units $i > 2$) $\phi_{ii} \sim IIDU\,(0.3, 0.5)$, $\phi_{i1} \sim IIDU\,(0, 0.1)$, $\phi_{i,i+1} \sim IIDU\,(-0.2, 0.2)$, $\phi_{ij} = \alpha_i \lambda_{ij}$ for $j \notin \{1, i, i+1\}$, where $\alpha_i \sim IIDU\,(0.05, 0.15)$.

Construction of Reduced Form Errors and Factor loadings

- $\mathbf{u}_t = \delta u_{1t} + \mathbf{e}_t$, where $u_{1t} \sim IIDN\left(0, 0.15\right)$, we set $\delta_1 = 1$, $\delta_2 = 0.1$ and generate $\delta_i \sim IIDU\left(0, 0.3\right)$ for $i = 3, .., N$.
- We set $e_{1t} = 0$ and $\left\{e_{2t}, ..., e_{Nt}\right\}$, are generated from a stationary bilateral Spatial Autoregressive Model (SAR) in order to show that our estimators are invariant to the weak cross section dependence of innovations:

$$e_{it} = \frac{a_e}{2}\left(e_{i-1,t} + e_{i+1,t}\right) + \vartheta_{eit},$$

where $\vartheta_{eit} \sim IIDN\left(0, \sigma^2_{\vartheta e}\right)$, and SAR parameter $a_e = 0.4$ (this is CWD, see Pesaran and Tosetti, 2009)
- Two sets of factor loadings are considered, $\boldsymbol{\gamma} = \mathbf{0}$ (no unobserved common factor) and $\boldsymbol{\gamma} \neq \mathbf{0}$. Under the latter we set $\gamma_1 = 1$, $\gamma_2 = -0.5$ and generate $\gamma_i \sim IIDN\left(1, 1\right)$ for $i = 3, .., N$.

Specification of Auxiliary Regressions

- We consider three different augmentations:
  - (i) by dominant unit, neighbor $i+1$ and arithmetic cross section averages $\{x_{1t}, \overline{x}_t, x_{i+1,t-1}, x_{1,t-1}, \overline{x}_{t-1}\}$,
  - (ii) augmentation by dominant unit and neighbor $i+1$, $x_{i+1,t-1} \cup \{x_{1,t-\ell}\}_{\ell=0}^{p}$ with p being the largest integer smaller than $T^{1/3}/2$, and
  - (iii) augmentation by neighbor $i+1$ and arithmetic cross section averages, $x_{i+1,t-1} \cup \{\overline{x}_{t-\ell}\}_{\ell=0}^{p}$ with the number of lags p chosen to grow with T in the same way.
- Auxiliary regression for unit $i = 2$ corresponding to augmentation of form (i) is:

$$x_{2t} = c_2 + \phi_{22} x_{2,t-1} + \phi_{23} x_{3,t-1} + \sum_{\ell=0}^{1} b_{1\ell} x_{1,t-\ell} + \sum_{\ell=0}^{1} b_{2\ell} \overline{x}_{t-\ell} + \epsilon_{2t}.$$

Table 1a: RMSE of $\widehat{\phi}_{22}$ in the experiments zero factor loadings ($\gamma = 0$).

	Auxiliary regressions are augmented by neighbor $x_{3,t-1}$ and:								
	$\{x_{1,t-\ell}\}_{\ell=0}^{p(T)}$			$\{\overline{x}_{t-\ell}\}_{\ell=0}^{p(T)}$			$\{x_{1,t-\ell}, \overline{x}_{t-\ell}\}_{\ell=0}^{1}$		
N\T	50	100	200	50	100	200	50	100	200
25	15.90	10.45	7.12	15.74	10.64	7.26	17.00	11.08	7.47
50	16.07	10.69	6.84	15.88	10.60	6.79	16.94	11.29	7.00
75	16.18	10.77	7.02	15.97	10.69	6.96	17.20	11.23	7.16
100	16.26	10.70	7.13	16.20	10.70	7.10	17.29	11.16	7.29
200	15.74	10.66	7.07	15.69	10.66	7.08	16.60	11.12	7.26

Table 1b: Size ($\times$100) (5% level, $H_0 : \phi_{22} = 0.50$) in experiments with zero factor loadings ($\gamma = 0$).

	Auxiliary regressions are augmented by neighbor $x_{3,t-1}$ and:								
	$\{x_{1,t-\ell}\}_{\ell=0}^{p(T)}$			$\{\bar{x}_{t-\ell}\}_{\ell=0}^{p(T)}$			$\{x_{1,t-\ell}, \bar{x}_{t-\ell}\}_{\ell=0}^{1}$		
N\T	50	100	200	50	100	200	50	100	200
25	7.30	6.35	6.20	7.15	6.90	6.50	8.45	7.35	6.65
50	7.15	6.60	5.25	7.45	6.80	5.40	8.80	7.90	5.50
75	7.55	6.80	5.75	7.45	6.50	5.95	8.70	7.65	6.20
100	7.40	6.30	5.90	7.85	6.75	6.45	8.85	6.90	6.50
200	6.60	6.35	5.95	6.60	6.45	5.95	7.55	6.95	6.45

Table 1c: Power ($\times$100) (5% level, $H_1 : \phi_{22} = 0.70$) in experiments with zero factor loadings ($\gamma = 0$).

	Auxiliary regressions are augmented by neighbor $x_{3,t-1}$ and:								
	$\{x_{1,t-\ell}\}_{\ell=0}^{p(T)}$			$\{\overline{x}_{t-\ell}\}_{\ell=0}^{p(T)}$			$\{x_{1,t-\ell}, \overline{x}_{t-\ell}\}_{\ell=0}^{1}$		
N\T	50	100	200	50	100	200	50	100	200
25	47.85	67.45	93.40	49.70	70.45	95.20	51.95	71.70	94.90
50	46.95	67.90	92.80	46.50	69.65	94.10	48.60	69.40	94.10
75	48.00	70.05	92.90	48.05	70.55	93.60	49.05	71.35	93.20
100	48.05	69.80	91.85	48.00	70.40	92.75	49.40	71.60	92.80
200	48.40	69.05	92.05	47.15	69.25	92.25	49.90	69.85	92.25

Table 2a: RMSE of $\widehat{\phi}_{22}$ in the experiments with nonzero factor loadings ($\gamma \neq 0$).

	Auxiliary regressions are augmented by neighbor $x_{3,t-1}$ and:								
	$\{x_{1,t-\ell}\}_{\ell=0}^{p(T)}$			$\{\bar{x}_{t-\ell}\}_{\ell=0}^{p(T)}$			$\{x_{1,t-\ell}, \bar{x}_{t-\ell}\}_{\ell=0}^{1}$		
N\T	50	100	200	50	100	200	50	100	200
25	14.38	12.59	12.82	14.40	11.88	11.09	15.48	10.18	7.45
50	14.31	12.70	12.78	14.57	11.44	10.39	16.35	10.54	6.77
75	14.02	12.58	12.61	14.55	10.93	10.16	17.18	10.27	6.86
100	14.38	12.54	12.81	14.29	11.00	10.05	17.45	10.97	6.79
200	14.22	12.83	12.91	14.58	11.13	10.25	17.87	11.08	7.08

Table 2b: Size ($\times$100) (5% level, $H_0 : \phi_{22} = 0.50$) in experiments with nonzero factor loadings ($\gamma \neq 0$).

	Auxiliary regressions are augmented by neighbor $x_{3,t-1}$ and:								
	$\{x_{1,t-\ell}\}_{\ell=0}^{p(T)}$			$\{\overline{x}_{t-\ell}\}_{\ell=0}^{p(T)}$			$\{x_{1,t-\ell}, \overline{x}_{t-\ell}\}_{\ell=0}^{1}$		
N\T	50	100	200	50	100	200	50	100	200
25	9.15	22.90	49.20	8.30	17.30	35.20	7.20	7.10	10.60
50	9.40	21.70	49.20	7.55	14.10	29.60	7.45	6.35	6.25
75	8.70	23.20	48.15	6.90	12.40	28.30	9.50	5.60	6.15
100	9.00	22.70	48.65	6.40	12.95	26.75	9.10	7.20	5.00
200	9.55	23.40	49.85	7.40	14.00	28.80	10.65	6.50	5.85

Table 2c: Power ($\times 100$) (5% level, $H_1 : \phi_{22} = 0.70$) in experiments with nonzero factor loadings ($\gamma \neq 0$).

	Auxiliary regressions are augmented by neighbor $x_{3,t-1}$ and:								
	$\{x_{1,t-\ell}\}_{\ell=0}^{p(T)}$			$\{\overline{x}_{t-\ell}\}_{\ell=0}^{p(T)}$			$\{x_{1,t-\ell}, \overline{x}_{t-\ell}\}_{\ell=0}^{1}$		
N\T	50	100	200	50	100	200	50	100	200
25	25.00	26.80	33.50	25.75	30.70	43.60	39.25	51.20	75.85
50	24.25	25.80	34.40	27.25	33.05	48.40	44.70	61.70	88.25
75	24.15	25.95	33.35	29.15	33.25	51.70	49.40	66.80	90.00
100	24.85	25.55	34.00	27.50	35.55	50.40	49.85	67.85	91.50
200	24.00	26.45	34.30	28.25	34.10	49.00	52.80	69.20	91.65

Modelling Returns on Equity Futures

- We model weekly returns on equity futures across the globe, where we assume that S&P is dominant unit.
- $N = 26$, and $T = 308$ observations of weekly returns (06 January 2003 - 24 November 2008)
- The equity series refer to futures contracts downloaded from Datastream. Daily returns were calculated allowing for contract rollovers. Weekly returns were calculated from daily returns by summing working days (Monday to Friday).

- It is assumed that $\mathbf{x}_t$ (vector of returns) is given by the following VAR model,

$$\mathbf{x}_t = \mathbf{d} + \mathbf{\Phi}\mathbf{x}_{t-1} + \mathbf{u}_t,$$

and $\mathbf{u}_t = \delta u_{1t} + \mathbf{e}_t$,where $\mathbf{e}_t = \mathbf{R}\varepsilon_t$ is CWD.

- The neighbors of unit i are S&P (unit 1), unit i itself, and the following spatial weighted average:

$$\overline{x}_{wit} = \sum_{j=2}^{N} \frac{w_{ij}}{1 - w_{i1}} x_{jt},$$

where w_{ij} are financial weights constructed according to average of assets and liabilities holdings for equities during 2001-2007 period ($w_{ii} = 0$). Source: IMF CPIS database.

Financial weights matrix (selected countries)

	US	UK	France	Germany	NL	Japan
US	-	20.9%	7.4%	6.6%	8.3%	15.0%
UK	45.4%	-	7.2%	6.1%	5.6%	8.9%
France	31.3%	14.1%	-	13.2%	6.7%	4.9%
Germany	32.3%	13.8%	15.2%	-	7.0%	3.8%
NL	44.0%	13.7%	8.4%	7.6%	-	4.2%
Japan	56.7%	15.5%	4.3%	2.9%	3.0%	-

- The following *conditional* models are estimated

$$x_{it} = c_i + \sum_{\ell=1}^{k_{oi}} \phi_{ii,\ell} x_{i,t-1} + \sum_{\ell=0}^{k_{di}} \beta_{i\ell} x_{1,t-\ell} + \sum_{\ell=1}^{k_{si}} h_{i\ell} \overline{x}_{wi,t-\ell} + \epsilon_{it},,$$

for $i = 2, .., N$, and for $i = 1$ (S&P), we estimate the following *marginal* model,

$$x_{1t} = c_1 + \sum_{\ell=1}^{k_{o1}} a_\ell x_{1,t-\ell} + \sum_{\ell=1}^{k_{s1}} b_\ell \overline{x}_{w1,t-\ell} + \epsilon_{1t},$$

- Truncation lags were chosen according to SBC criterion with the maximum lag set to 4.

- Formal tests of global dominance in IVAR models are yet to be developed.
- Nevertheless, the assumption that x_{1t} is weakly exogenous in the equation x_{it}, $i = 2, .., N$, can be tested using the procedure advanced by Wu (1973) and Hausman (1978).
- Wu's approach is to test the statistical significance of the S&P residuals $\widehat{\epsilon}_{1t}$ in the equation for remaining units. This test is asymptotically equivalent to using Hausman's procedure.

	SP Cont.		SP Lag		Own Lag		Spatial Lag			
Unit	Coef	t	Coef.	t	Coef.	t	Coef.	t	$\bar{R}^2$	Wu-H.
Dominant Unit										
S&P	-	-	-	-	0.563	3.6	-0.543	-4.3	0.067	-
Advanced European Countries										
AEX	1.086	24.3	0.441	4.5	-0.538	-4.0	0.469	2.4	0.695	1.44
BEL	0.946	21.8	0.718	6.4	-0.046	-0.5	-0.241	-1.8	0.665	2.00
CAC	0.959	28.0	0.328	4.5	-0.274	-2.5	0.027	0.2	0.749	-0.23
DAX	1.089	26.9	0.461	4.3	-0.294	-3.8	0.061	0.6	0.753	0.91
FTSE	0.833	27.4	0.292	3.2	-0.181	-2.4	-0.062	-0.8	0.751	0.54
FOX	0.900	20.7	0.612	4.7	-0.209	-2.2	-0.208	-1.8	0.638	0.04
GRX	0.880	13.8	0.616	3.6	-0.095	-1.6	0.225	1.6	0.461	-1.61
IBE	0.857	20.9	0.433	4.1	0.034	0.4	-0.303	-2.8	0.651	2.59
KFX	0.837	18.6	0.603	4.9	-0.114	-1.7	-0.111	-0.9	0.590	-2.53
MIB	0.852	25.6	0.326	4.5	-0.328	-2.7	0.185	1.4	0.722	0.33
OBX	0.928	15.5	0.649	5.1	0.049	0.8	-0.334	-2.2	0.510	-1.90
OMX	0.969	23.2	0.337	3.8	-0.190	-2.5	0.052	0.5	0.656	-0.54
PSI	0.619	14.9	0.325	3.7	0.036	0.6	0.036	0.4	0.504	-1.16
SMI	0.855	22.8	0.211	2.7	-0.268	-3.4	0.038	0.4	0.685	-0.38

	SP Cont.		SP Lag		Own Lag		Spatial Lag			
Unit	Coef	t	Coef	t	Coef	t	Coef	t	$\overline{R}^2$	Wu-H.
Other Advanced Countries										
TSX	0.751	21.9	0.161	2.3	-0.095	-1.5	0.054	0.7	0.633	0.16
NK	0.919	17.4	0.323	2.8	-0.108	-1.7	0.107	0.8	0.539	-0.42
ASX	0.600	16.5	0.284	3.7	-0.115	-1.8	0.019	0.2	0.503	-1.84
Latin America										
BRX	1.210	18.7	0.327	1.8	-0.109	-1.9	0.107	0.7	0.571	-0.07
Emerging Europe										
HUX	0.831	13.3	0.399	3.0	-0.018	-0.2	0.193	1.2	0.458	1.14
POX	0.898	13.1	0.180	1.2	-0.023	-0.4	0.100	0.7	0.370	-1.96
SAX	0.756	12.9	0.217	2.2	-0.152	-2.5	-0.031	-0.3	0.361	-1.92
Emerging Asia										
HKX	0.707	11.7	0.713	4.5	-0.206	-3.2	0.003	0.0	0.391	-2.30
KOX	0.768	11.2	0.050	0.4	-0.551	-5.7	1.047	5.9	0.409	-1.17
SIX	0.718	14.3	1.127	7.2	-0.386	-4.0	-0.122	-0.9	0.493	1.12
TWX	0.717	10.4	0.772	4.2	-0.380	-4.2	0.007	0.0	0.358	0.70

Conclusion

- This paper considered the problem of estimation of high dimensional VARs featuring a dominant unit.
- We showed that the asymptotic normality of the cross section augmented least squares estimator continues to hold (once the individual auxiliary regressions are correctly specified).
- How to correctly specify the individual regressions is an important topic, and the correct specification depends on the assumption about the presence of dominant units, observed and unobserved common factors and the (local) spatiotemporal neighborhood effects.
- The framework developed here can be applied to model spatio-temporal dependence. See "Spatial and Temporal Diffusion of House Prices in the UK" by Holly, Pesaran and Yamagata (2009, forthcoming).