Online Appendix

Validating China's output data using satellite observations

Stephen D. Morris*
Department of Economics
Bowdoin College

Junjie Zhang†
Duke Kunshan University and
Duke University

February 20, 2017

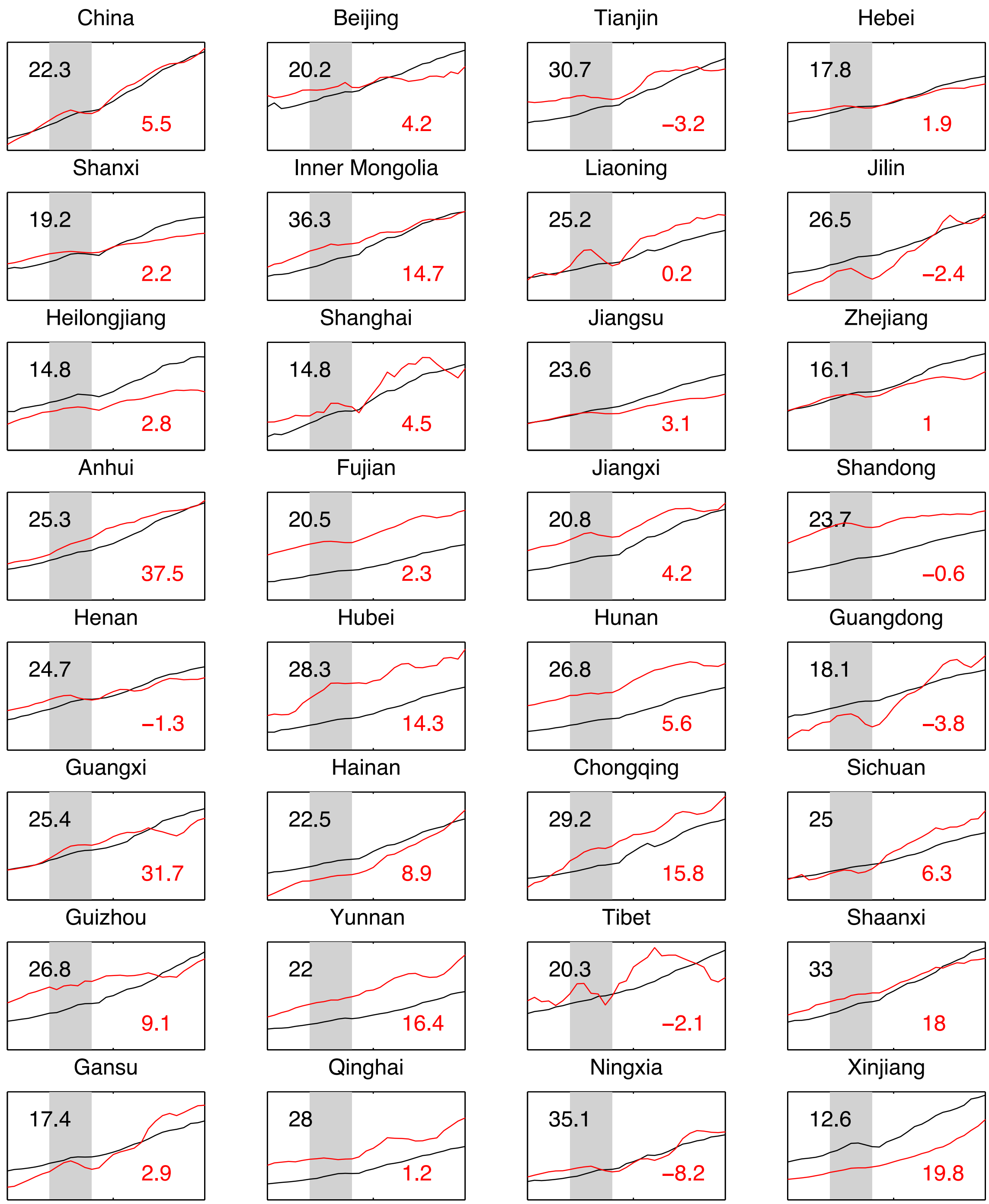
Revised: September 27, 2017

*Bowdoin College, Department of Economics, 9700 College Station, Brunswick, ME 04011. smorris@bowdoin.edu.

†Duke Kunshan University, Environmental Research Center; Duke University, Nicholas School of the Environment. No. 8 Duke Avenue, Kunshan, Jiangsu, China 215316. junjie.zhang@duke.edu.

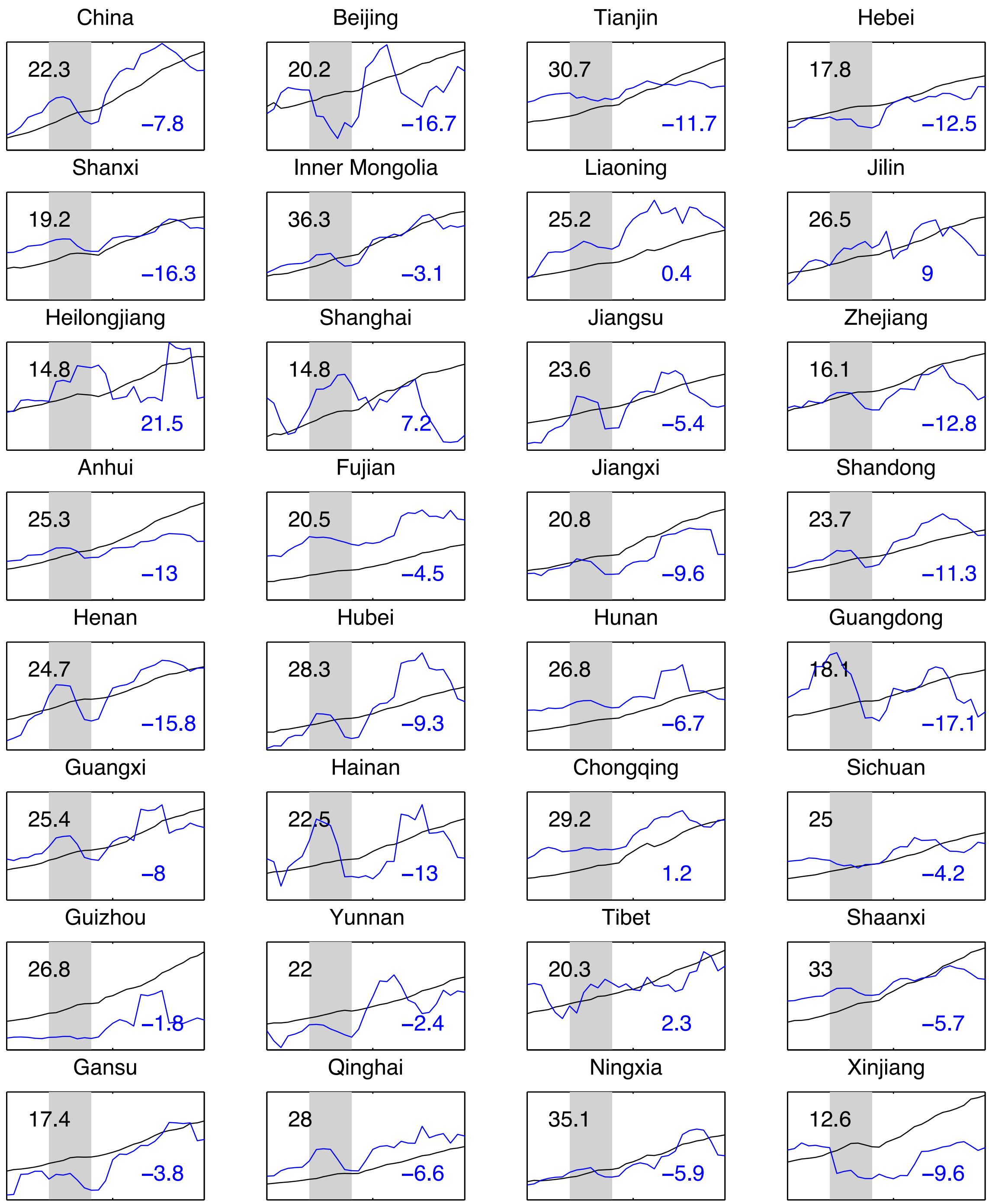
A Supplementary tables and figures

Figure A.1: Reported GDP vs. electricity generation, Great Recession period.



Notes: Units are equivalent to Figure 1 (a) in main paper (All-China: Top-left pane in this Figure). Scale varies by location. National level and $N = 31$ sub-national regions: Quarterly four-quarter rolling average, 2006-2013. Shaded: NBER recession dates (07Q4-09Q2). Black inset: GDP NBER dates % change. Red inset: Electricity NBER dates % change.

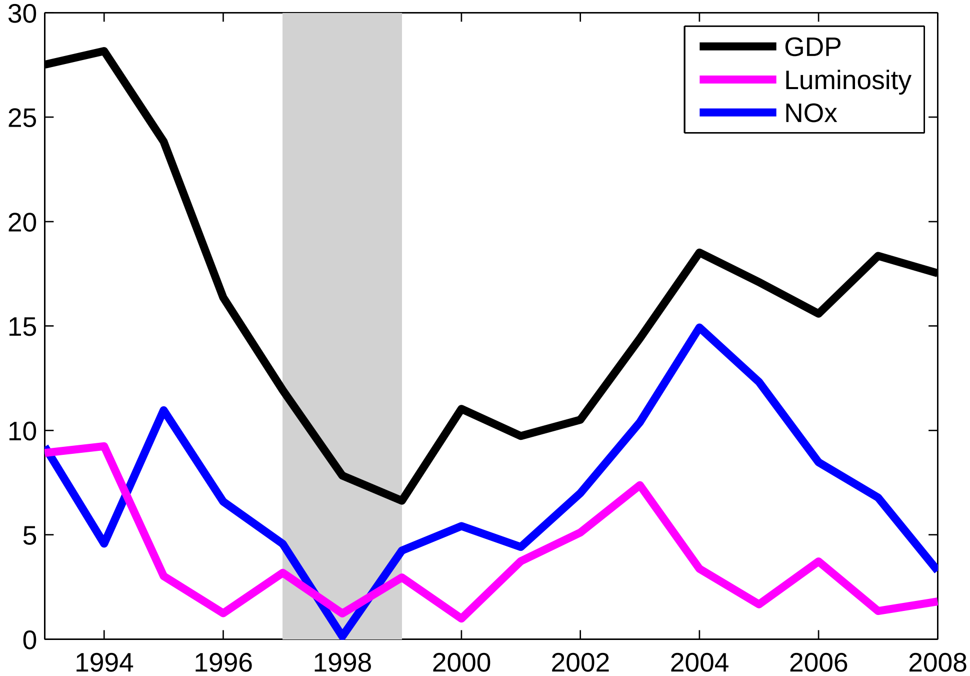
Figure A.2: Reported GDP vs. NO_x emissions, Great Recession period.



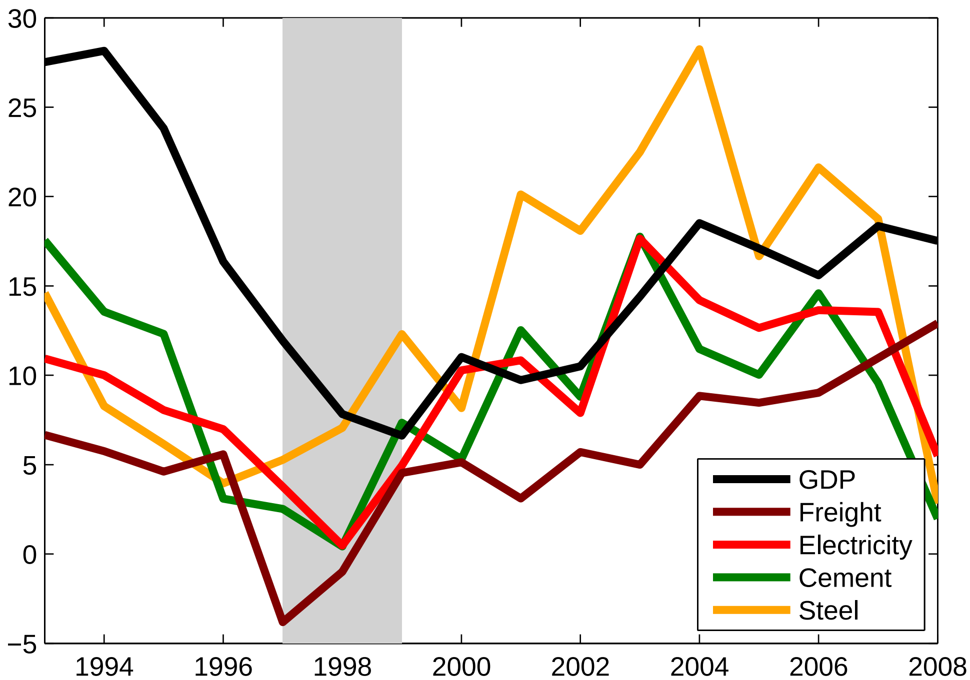
Notes: Units are equivalent to Figure 1 (b) in main paper (All-China: Top-left pane in this Figure). Scale varies by location. National level and $N = 31$ sub-national regions: Quarterly four-quarter rolling average, 2006-2013. Shaded: NBER recession dates (07Q4-09Q2). Black inset: GDP NBER dates % change. Blue inset: NO_x NBER dates % change.

Figure A.3: Annual % change, China: 1993-2008.

(a) Satellite-measured signals.



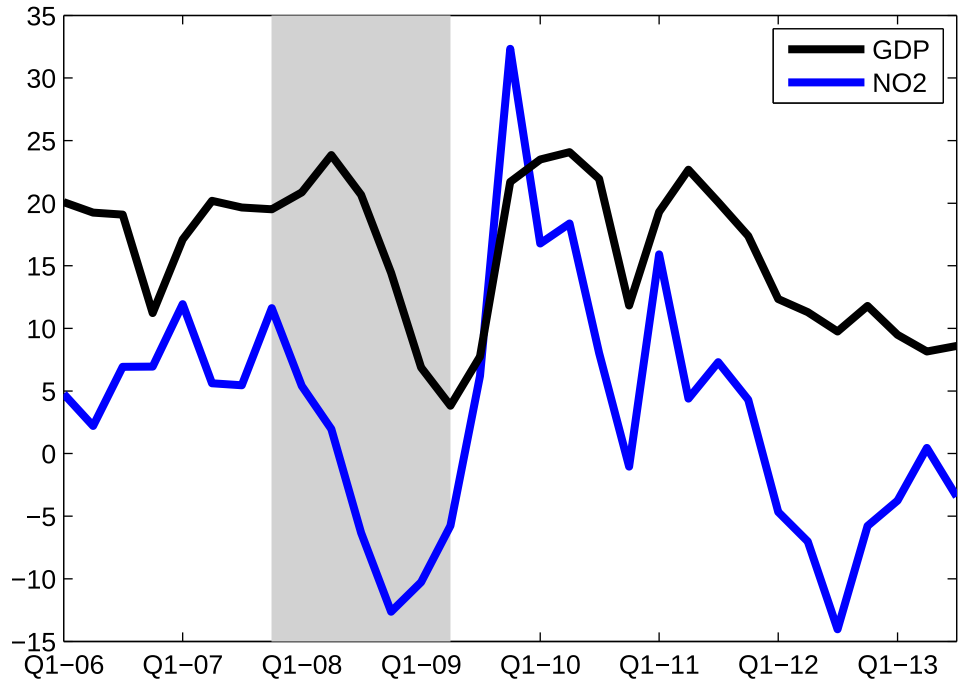
(b) Reported signals.



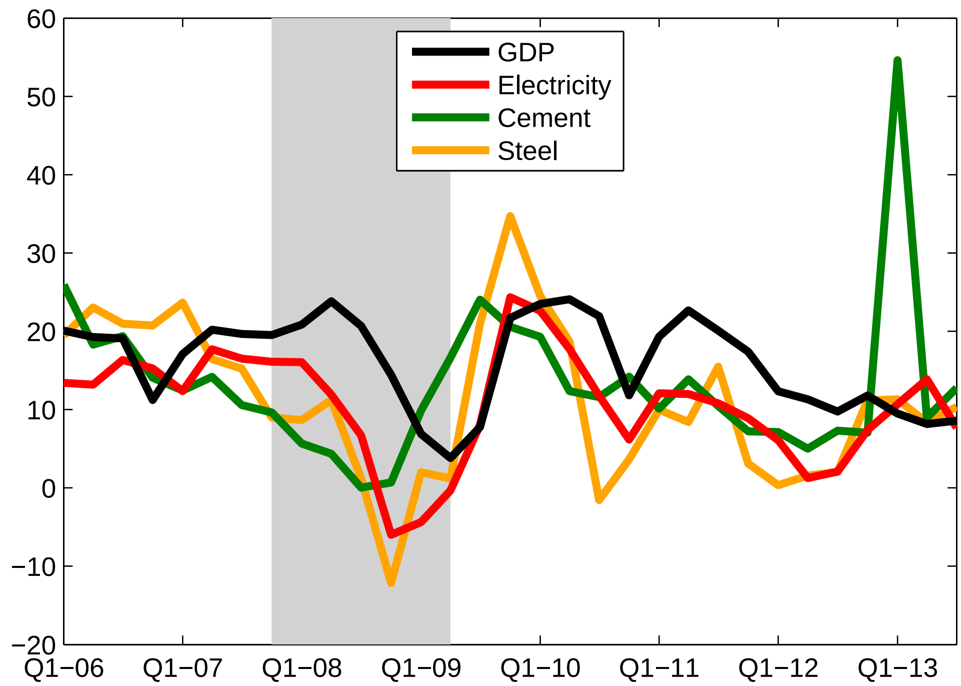
Notes: Shading: Asian Financial Crisis.

Figure A.4: Annualized quarterly % change, China: 2006 Q1 - 2013 Q3.

(a) Satellite-measured signals.

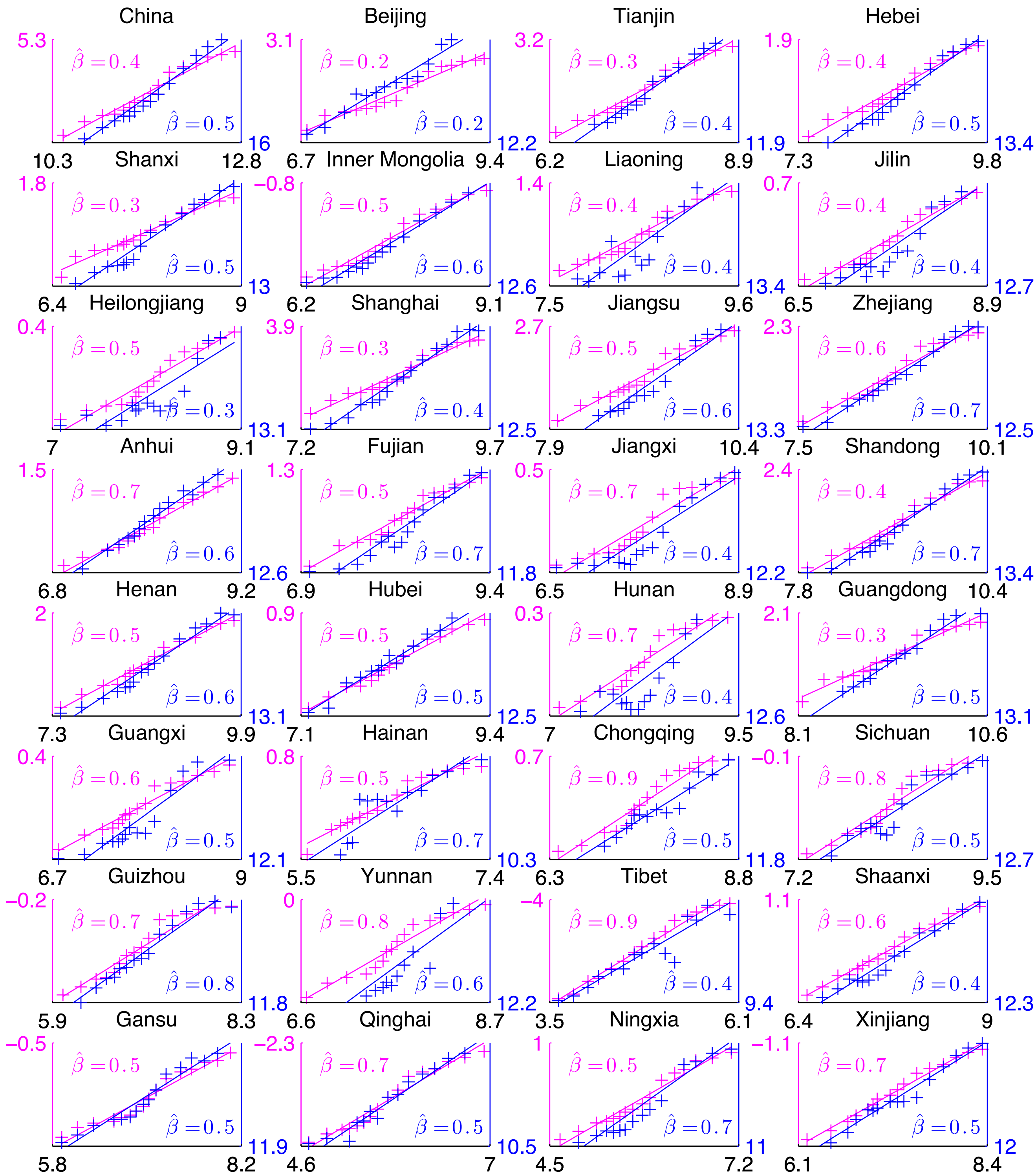


(b) Reported signals.



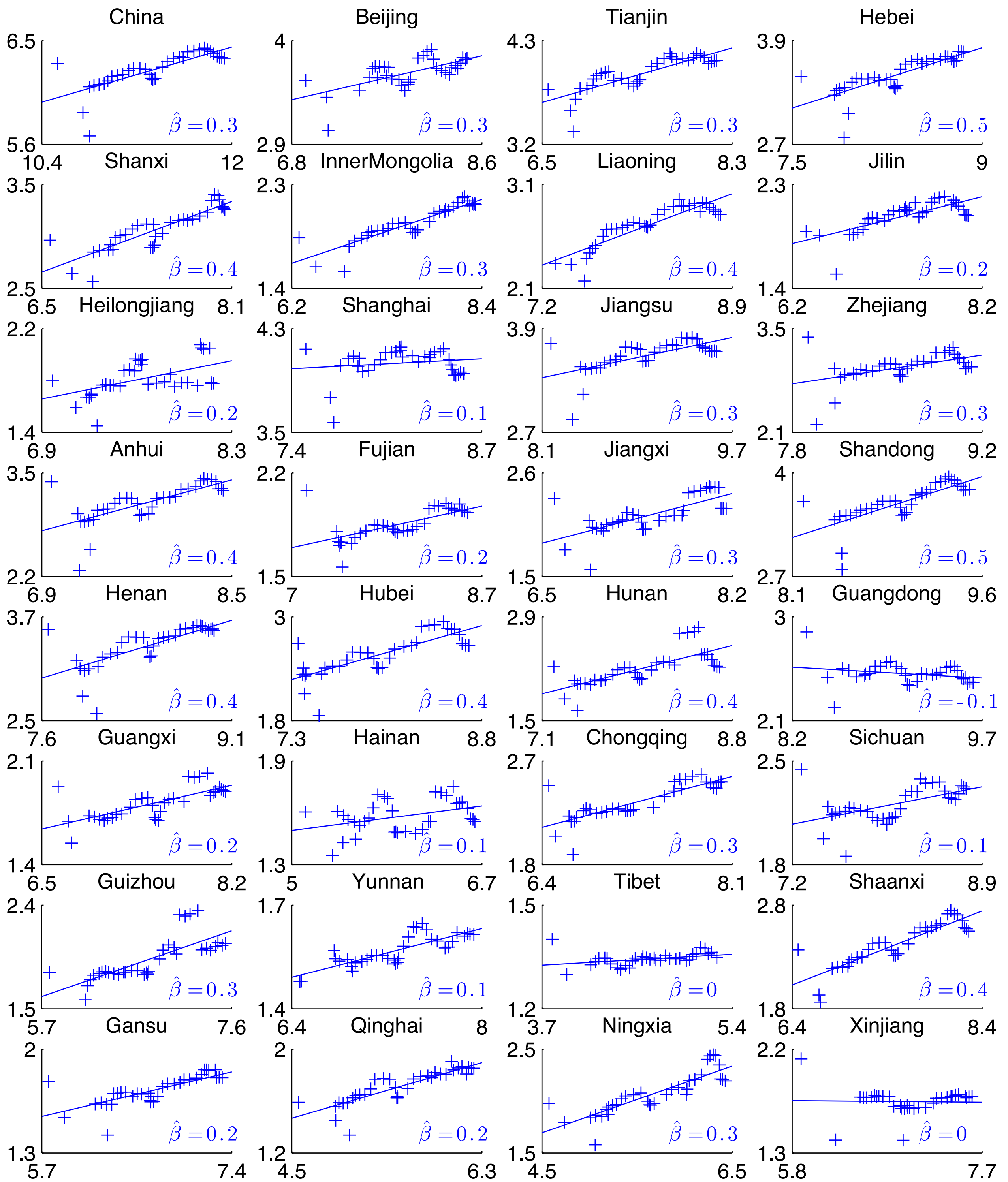
Notes: Shading: NBER recession dates.

Figure A.5: Annual elasticities by-region (1993-2008).



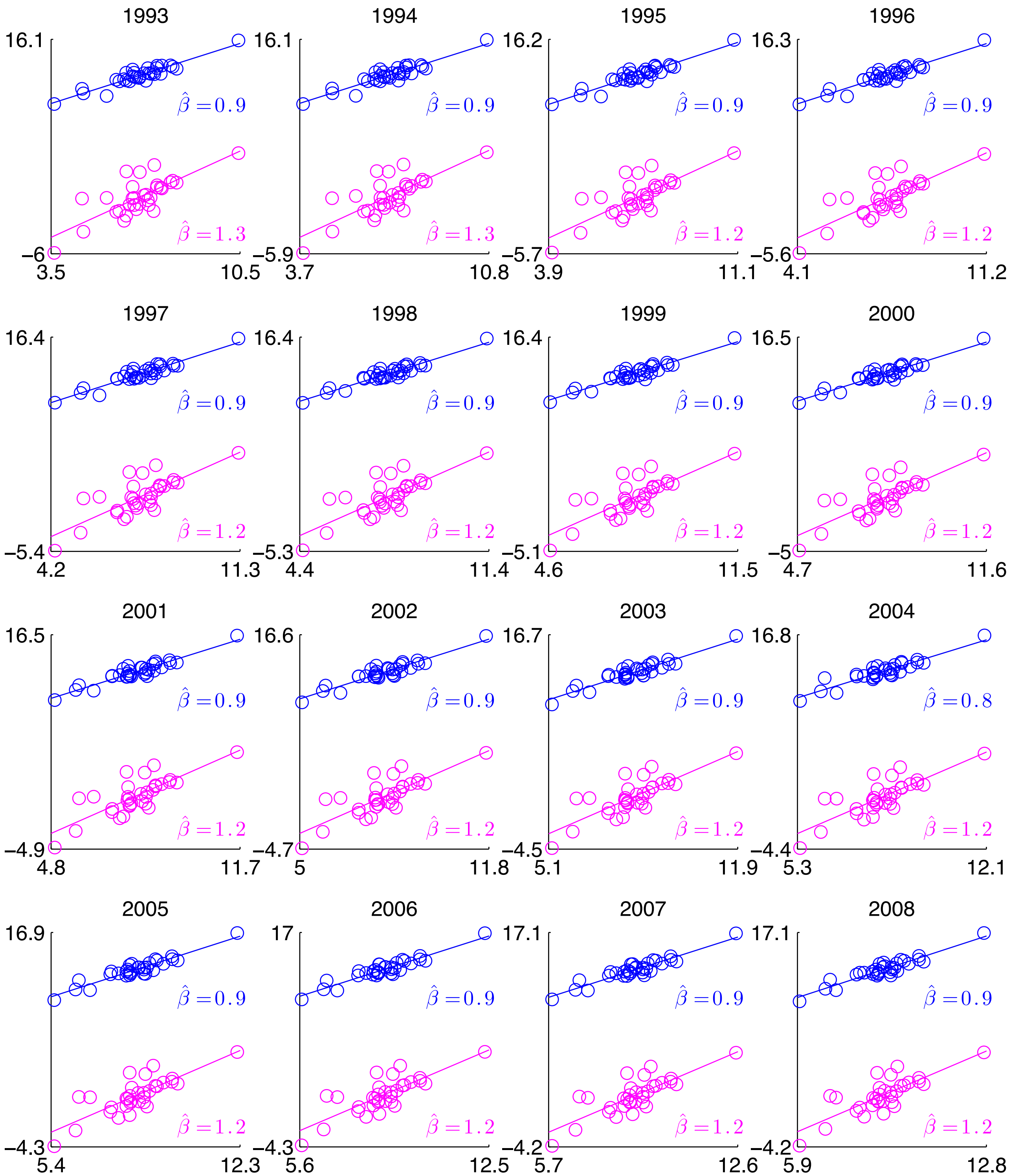
Notes: Pink: Luminosity. Blue: NO_x emissions. Across $T = 16$ quarters in each pane. Estimates are for regression model $\ln \text{Signal}_t = \beta \ln \text{Output}_t + \varepsilon_t$ for each n.

Figure A.6: NO_2 quarterly elasticities by-region.



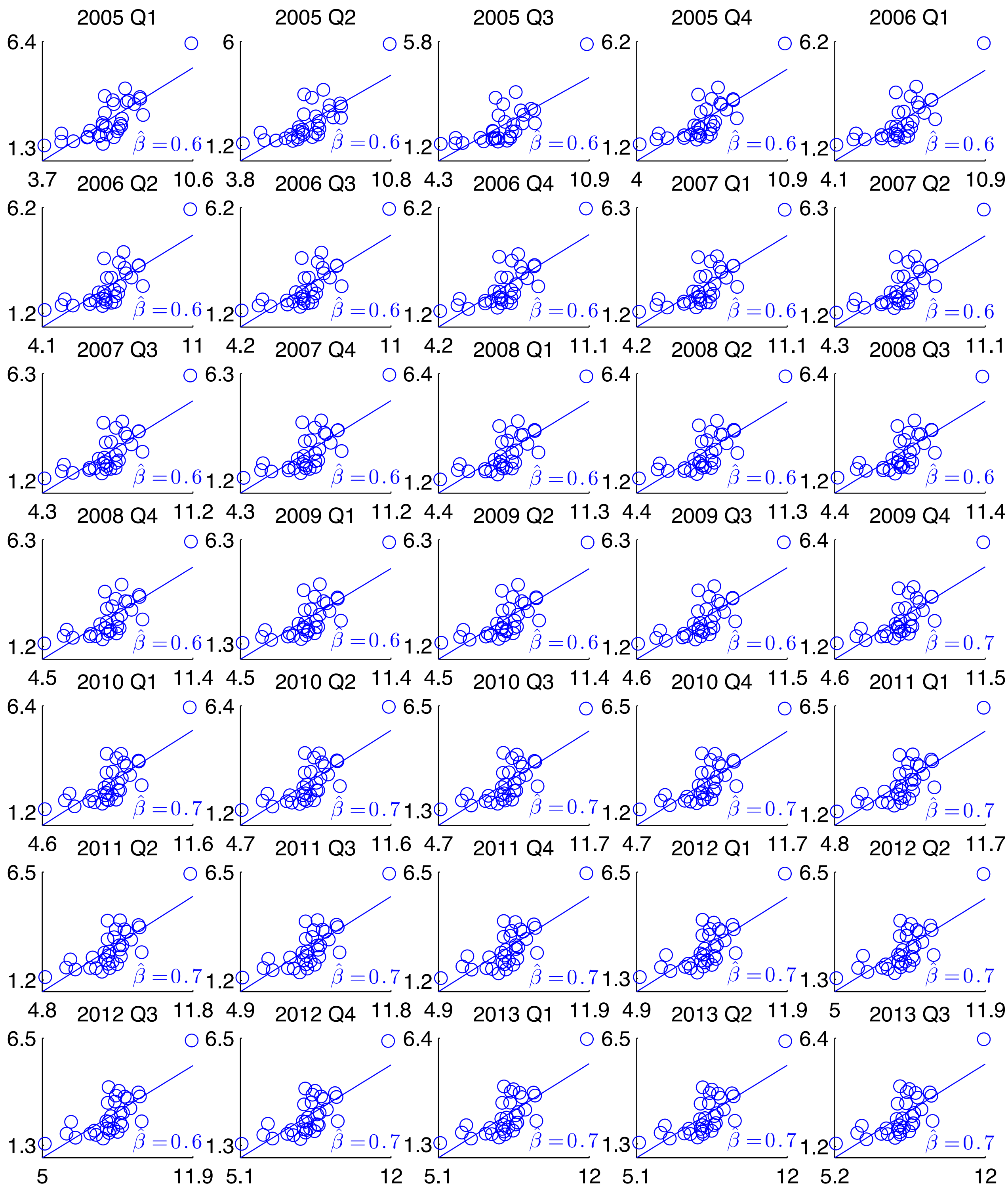
Notes: Across $T = 31$ quarters in each pane. Estimates are for regression model $\ln \text{Signal}_t = \beta \ln \text{Output}_t + \varepsilon_t$ for each n.

Figure A.7: Elasticities by-year.



Notes: Pink: Luminosity. Blue: NO_x emissions. Across $N = 31$ regions in each pane. Estimates are for regression model $\ln \text{Signal}_n = \beta \ln \text{Output}_n + \varepsilon_n$ for each t.

Figure A.8: NO_2 elasticities by-quarter.



Notes: Across $N = 31$ regions in each pane. Estimates are for regression model $\ln \text{Signal}_n = \beta \ln \text{Output}_n + \varepsilon_n$ for each t.

Table A.1: Correlations, annual sample (hundredths).

	GL	GN	GF	GE	GC	GS	LN	LF	LE	LC	LS	NF	NE	NC	NS
China	**53**	43	40	35	51	-5	18	-3	35	**73**	25	39	**72**	61	60
Beijing	45	**48**	-7	-26	40	45	38	5	-22	25	**43**	17	14	41	**54**
Tianjin	**45**	26	-19	6	-15	7	21	24	**39**	35	27	14	22	5	**37**
Hebei	**61**	44	31	21	34	-12	23	1	29	**55**	20	12	**62**	30	26
Shanxi	41	**43**	-9	26	16	17	21	-25	12	**36**	4	29	**72**	51	47
In. Mon.	37	**58**	40	55	32	28	**44**	8	23	22	26	42	**71**	69	40
Liaoning	15	25	**62**	-13	36	-4	37	30	42	**60**	27	53	52	**53**	17
Jilin	34	35	11	**47**	4	-52	-2	9	**59**	41	4	-1	**38**	32	8
Heil.	25	**32**	2	26	-38	-27	7	-48	1	4	**31**	-11	71	**29**	26
Shanghai	**59**	37	-41	9	12	39	33	-13	28	31	**40**	13	18	**36**	28
Jiangsu	**47**	37	29	41	45	18	3	0	33	71	**49**	49	**62**	40	15
Zhejiang	**57**	30	42	12	54	38	-1	38	23	**66**	22	48	53	**56**	53
Anhui	46	**48**	0	15	38	-14	32	29	**46**	29	-31	-1	14	**50**	-12
Fujian	**57**	-1	25	14	34	-38	-28	-12	26	**49**	-15	-3	**9**	-10	-19
Jiangxi	-9	25	**34**	16	9	-60	18	-5	**67**	62	30	15	**67**	59	14
Shandong	45	20	**46**	29	36	-24	38	22	32	**81**	17	-3	31	35	**58**
Henan	**47**	42	31	36	44	13	8	-2	19	**40**	19	-31	**71**	5	59
Hubei	31	**32**	-1	27	26	-9	**44**	-11	2	27	-8	-18	16	**22**	-13
Hunan	-20	37	47	28	**47**	-58	18	11	**43**	-6	0	52	53	**62**	5
Guangdong	59	29	-1	**63**	51	-10	-17	-7	**87**	59	28	8	-5	2	**12**
Guangxi	45	26	37	38	**71**	10	-4	-3	16	**42**	-1	-14	12	21	**52**
Hainan	**81**	-12	-13	58	23	41	-9	-27	**55**	34	45	**12**	-10	**12**	4
Chongqing	-5	9	24	**25**	6	-25	-15	-40	-20	**49**	-6	**8**	-33	-8	-25
Sichuan	0	23	**36**	19	16	7	**55**	6	48	23	21	9	**33**	10	15
Guizhou	-20	-6	-20	**15**	-12	-46	**46**	-26	24	24	9	-6	**17**	5	-21
Yunnan	26	-19	1	25	38	**40**	-4	**31**	25	13	3	15	15	**43**	32
Tibet	19	**31**	-23	9	-7	NA	4	-35	44	**49**	NA	-28	**-12**	-48	NA
Shaanxi	39	48	**75**	25	15	7	19	24	-3	9	**45**	**24**	11	0	15
Gansu	-7	**12**	3	-23	-41	-2	44	43	**51**	17	-18	-2	**37**	14	-5
Qinghai	**29**	-3	14	-6	-49	14	**30**	-41	3	7	-2	20	**38**	19	23
Ningxia	38	20	**50**	19	-20	-2	18	-7	**41**	26	10	-3	23	**33**	6
Xinjiang	43	35	31	**49**	-11	43	3	**60**	5	4	15	21	53	-16	**54**
Mean	**33**	27	18	23	18	-1	17	1	29	**36**	15	12	**33**	25	21
Max Count	9	8	6	5	2	1	5	2	9	10	5	3	12	10	6

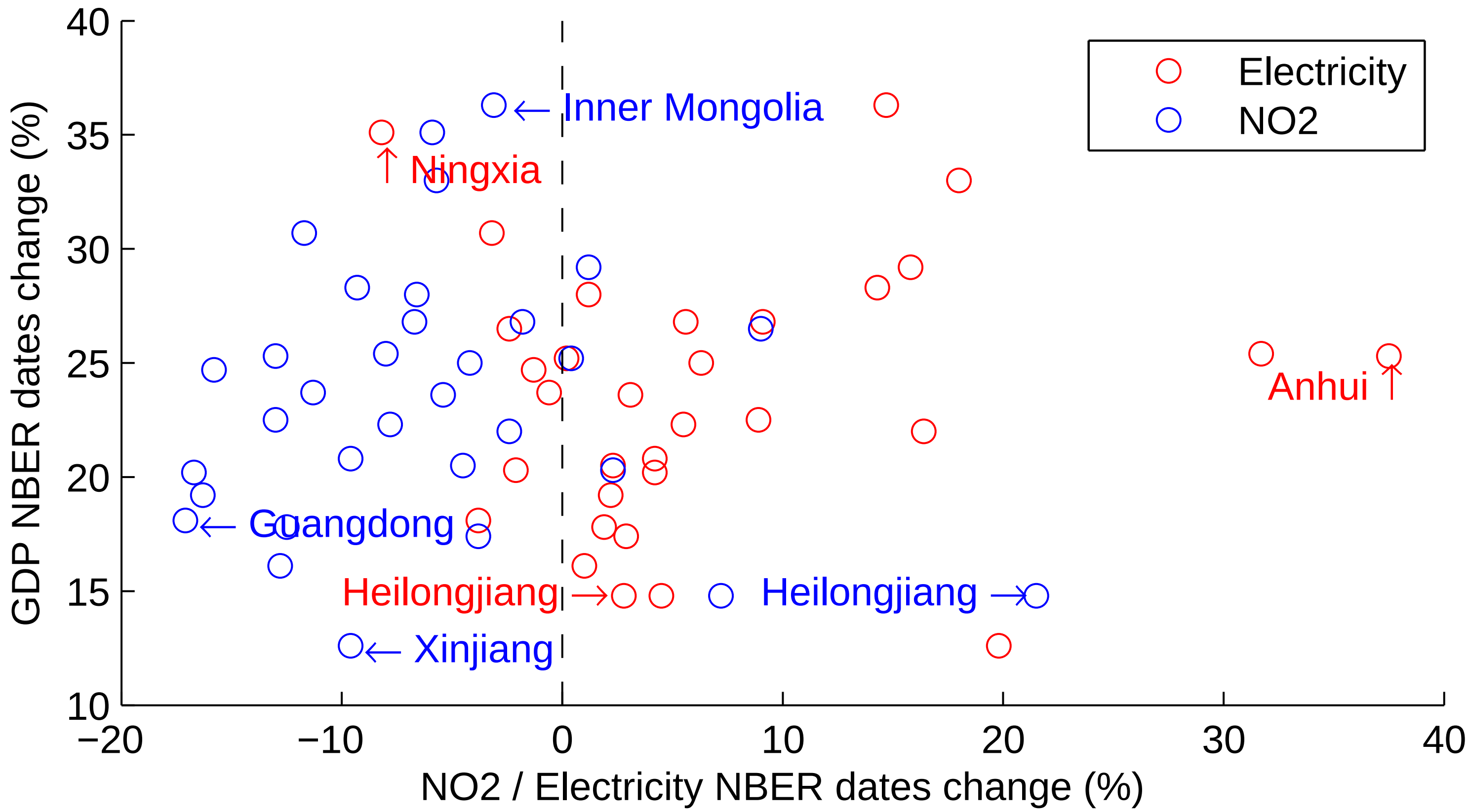
Notes: 1993-2008. G: GDP. L: Luminosity. N: NO_x. F: Freight. E: Electricity. C: Cement. S: Steel. "Max count" gives number of areas for which the leading G, L, or N correlation is the largest (bold).

Table A.2: Correlations, quarterly sample (hundredths).

	GN	GE	GC	GS	NE	NC	NS
China	62	**65**	-14	35	85	20	74
Beijing	**19**	-18	-1	3	19	43	**58**
Tianjin	24	**30**	0	28	45	36	**50**
Hebei	10	**50**	-3	10	**58**	53	28
Shanxi	58	**60**	-13	35	**71**	-4	69
Inner Mon.	49	**65**	24	4	**56**	33	-4
Liaoning	13	**42**	-13	6	41	-18	**58**
Jilin	7	**33**	-22	-1	**18**	-0	10
Heilongjiang	2	**48**	8	12	**10**	-5	-3
Shanghai	44	42	16	**66**	15	8	**26**
Jiangsu	57	**66**	-13	25	**60**	-8	33
Zhejiang	35	**40**	4	-0	**53**	4	4
Anhui	**45**	20	-34	1	22	-35	**27**
Fujian	**49**	34	1	23	**30**	4	17
Jiangxi	42	54	-27	**54**	**41**	-44	22
Shandong	45	**48**	-18	41	**56**	4	56
Henan	**32**	27	-9	13	**72**	20	43
Hubei	29	3	-31	**40**	24	-30	**55**
Hunan	**38**	48	-14	29	**24**	-17	19
Guangdong	13	39	18	**44**	**57**	20	36
Guangxi	15	**18**	-8	-12	**35**	-9	11
Hainan	24	35	**35**	NA	**18**	9	NA
Chongqing	**27**	-1	-16	4	22	44	**64**
Sichuan	-15	9	18	**20**	-3	**49**	15
Guizhou	14	-2	4	**29**	**0**	-7	-7
Yunnan	**3**	-18	-23	1	15	0	**48**
Tibet	-17	-4	**24**	NA	-13	**15**	NA
Shaanxi	**55**	34	-28	-25	**31**	-8	1
Gansu	23	**50**	-18	10	**47**	-33	14
Qinghai	**40**	27	-15	3	23	**28**	19
Ningxia	29	**38**	-27	NA	**81**	-41	NA
Xinjiang	-5	36	-8	**41**	**10**	-19	0
Mean	27	**32**	-6	17	**35**	3	26
Max Count	9	13	2	7	20	3	8

Notes: 2006Q1-2013Q3. G: GDP. L: N: NO_2. E: Electricity. C: Cement. S: Steel. "Max count" gives number of areas for which the leading G or N correlation is the largest (bold).

Figure A.9: China regional signal vs. GDP change: NBER recession dates (2007 Q4 - 2009 Q2).



B Output

This section justifies the assumption that y^*_{nt} has an AR(1) rule of motion (main paper Equation (2)). Consider a dynamic stochastic general equilibrium model not explicitly derived from microfoundations, but incorporating the main features of a wide class of models. It consists of a Taylor rule, linearized Euler equation, and Phillips curve.

$$r^*_{nt} = \psi_\pi \pi^*_{nt} + \psi_y y^*_{nt} + \varepsilon^r_{nt} \tag{B.1}$$

$$c^*_{nt} = E_t c^*_{nt+1} + h c^*_{nt-1} - (1/\tau)(r^*_{nt} - E_t \pi^*_{nt+1}) \tag{B.2}$$

$$\pi^*_{nt} = \beta E_t \pi^*_{nt+1} + \kappa y^*_{nt} \tag{B.3}$$

r^*_{nt} is the nominal interest rate, c^*_{nt} is nominal consumption, and π^*_{nt} is inflation. Each is stated in log difference from means. Stars denote they are unobservable values, comparable with y^*_{nt}. τ is the coefficient of relative risk aversion, h is habit formation, β is the discount factor, and κ is a function of the degree of price stickiness. ψ_π and ψ_y are indicative of monetary policy stance. $\varepsilon^r_{nt} \sim IWN(0,1)$. With no investment or government spending, we have the market-clearing condition $c^*_{nt} = y^*_{nt}$. Since c^*_{nt-1} is the only lagged variable in the system, the solution will have the form of a restricted singular VAR(1).

$$\begin{bmatrix} y^*_{nt} \\ r^*_{nt} \\ \pi^*_{nt} \end{bmatrix} = \begin{bmatrix} \rho_y \\ \rho_r \\ \rho_\pi \end{bmatrix} y^*_{nt-1} + \begin{bmatrix} \sigma_y \\ \sigma_r \\ \sigma_\pi \end{bmatrix} \varepsilon^r_{nt} \tag{B.4}$$

The VAR parameters are nonlinear functions of the underlying parameters of the model. This fact may be confirmed using the method of undetermined coefficients. Restricting the set of solutions to those for which $|\rho_y| \in (0,1)$, the first row of this VAR is simply Equation (2). But possibly many models have this same VAR solution. So, this particular structural model is not the unique underpinning of the assumption that y^*_{nt} follows an AR(1). Finally, if ε^y_{nt} is interpretable as a scaled monetary policy shock $\sigma_y \varepsilon^r_{nt}$, then the shocks should reasonably be correlated across areas n in a country. This is allowed by the assumptions.

C Representation

(Proof of Proposition 1) The model (2)-(6) may be stated in by-area n ABCD state space representation common of many macroeconomic models, for each signal i (Fernández-Villaverde et al., 2007).

$$
\begin{aligned}
\begin{bmatrix} y^*_{nt} \\ u^*_{nt} \end{bmatrix} &= \underbrace{\begin{bmatrix} \rho_y & 0 \\ 0 & \rho_u \end{bmatrix}}_{A} \begin{bmatrix} y^*_{nt-1} \\ u^*_{nt-1} \end{bmatrix} + \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{B} \varepsilon_{nt}(i) \\
\underbrace{\begin{bmatrix} y_{nt} \\ s_{nt}(i) \end{bmatrix}}_{Y_{nt}(i)} &= \underbrace{\begin{bmatrix} \rho_y & \rho_u \\ \beta(i)\rho_y & 0 \end{bmatrix}}_{C(i)} \begin{bmatrix} y^*_{nt-1} \\ u^*_{nt-1} \end{bmatrix} + \underbrace{\begin{bmatrix} 1 & 1 & 0 \\ \beta(i) & 0 & 1 \end{bmatrix}}_{D(i)} \underbrace{\begin{bmatrix} \varepsilon^y_{nt} \\ \varepsilon^u_{nt} \\ \varepsilon^s_{nt}(i) \end{bmatrix}}_{\varepsilon_{nt}(i)}
\end{aligned}
\tag{C.1}
$$

The admissible parameter space $\Theta(i)$ for $\theta(i)$, the structural parameters corresponding to any lone signal i, has been defined from economic fundamentals and benign normalizations.

$$\Theta(i): \ \beta(i) \neq 0, \ 0 < |\rho_y| < 1, \ 0 < |\rho_u| < 1, \ 0 < \sigma_y < \infty, \ 0 < \sigma_u < \infty, \ 0 < \sigma(i) < \infty$$

When $\theta(i) \in \Theta(i)$ the matrix $C(i)$ is upper-triangular and thus always invertible. So the observation equation implies $[y^*_{nt-1} \ u^*_{nt-1}]' = C(i)^{-1} Y_{nt}(i) - C^{-1}(i)D(i)\varepsilon_{nt}(i)$ and thus

$$Y_{nt}(i) = \begin{bmatrix} \rho_u & \psi(i) \\ 0 & \rho_y \end{bmatrix} Y_{nt-1}(i) + \begin{bmatrix} v^y_{nt}(i) \\ v^s_{nt}(i) \end{bmatrix} \tag{C.2}$$

$$\begin{bmatrix} v^y_{nt}(i) \\ v^s_{nt}(i) \end{bmatrix} = D(i)\varepsilon_{nt}(i) + \underbrace{\begin{bmatrix} 0 & 0 & -\psi(i) \\ 0 & 0 & -\rho_y \end{bmatrix}}_{C(B-AC^{-1}D)} \varepsilon_{nt-1}(i) = \underbrace{\begin{bmatrix} u^y_{nt}(i) \\ u^s_{nt}(i) \end{bmatrix}}_{D(i)\varepsilon_{nt}(i)} + \begin{bmatrix} 0 & -\psi(i)m(i) \\ 0 & -\rho_y m(i) \end{bmatrix} \begin{bmatrix} u^y_{nt-1}(i) \\ u^s_{nt-1}(i) \end{bmatrix} \tag{C.3}$$

$$\psi(i) = \lambda_i \frac{\rho_y - \rho_u}{\beta(i)} \qquad m(i) = \frac{\sigma(i)}{\sqrt{\beta(i)^2 \sigma_y^2 + \sigma^2(i)}}$$

This is VARMA(1,1) representation. See Morris (2016) for details on VARMA representations of ABCD models. Now consider the general case of $S \geq 1$ signals. Given (C .2),

$$\underbrace{\begin{bmatrix} y_{nt} \\ s_{nt}(1) \\ \vdots \\ s_{nt}(S) \end{bmatrix}}_{Y_{nt}} = \underbrace{\begin{bmatrix} \rho_u & \psi(1) & \dots & \psi(S) \\ 0 & \rho_y & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \rho_y \end{bmatrix}}_{P} \begin{bmatrix} y_{nt-1} \\ s_{nt-1}(1) \\ \vdots \\ s_{nt-1}(S) \end{bmatrix} + \underbrace{\begin{bmatrix} \sum_{i=1}^{S} \lambda_i v_{nt}^y(i) \\ v_{nt}^s(1) \\ \vdots \\ v_{nt}^s(S) \end{bmatrix}}_{V_{nt}} \tag{C.4}$$

for any weightings satisfying $\sum_{i=1}^{S} \lambda_i = 1$. Thus, $\lambda = [\lambda_1, \dots \lambda_S]'$ are nuisance parameters for all $S > 1$. They are only known when $S = 1$ (when $\lambda_1 = 1$). Using (C.3), the generalized error is

$$V_{nt} = \begin{bmatrix} \sum_{i=1}^{S} \lambda_i v_{nt}^y(i) \\ \underset{S\times 1}{v_{nt}^s} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & \underset{1\times S}{0'} \\ \underset{S\times 1}{\beta} & \underset{S\times 1}{0} & I_s \end{bmatrix}}_{\underset{(S+1)\times(S+2)}{M_0}} \varepsilon_{nt} + \underbrace{\begin{bmatrix} 0 & 0 & \underset{1\times S}{-\psi'} \\ \underset{S\times 1}{0} & \underset{S\times 1}{0} & -I_s \otimes \rho_y \end{bmatrix}}_{\underset{(S+1)\times(S+2)}{M_1}} \varepsilon_{nt-1} \tag{C.5}$$

for $\psi' = (\lambda_1\psi(1), \dots, \lambda_S\psi(S))$ and ε_{nt} the multiple signal generalization of $\varepsilon_{nt}(i)$.

$$\underset{(S+2)\times 1}{\varepsilon_{nt}} = \begin{bmatrix} \varepsilon_{nt}^y & \varepsilon_{nt}^u & \varepsilon_{nt}^s \end{bmatrix}'; \qquad \Sigma_\varepsilon = E(\varepsilon_{nt}\varepsilon_{nt}') = \begin{bmatrix} \sigma_y^2 & \cdot & \cdot \\ 0 & \sigma_u^2 & \cdot \\ 0 & 0 & \Sigma \end{bmatrix} \tag{C.6}$$

Next, using elementary matrix algebra operations, we have the equivalence $PY_{nt-1} = \text{vec}(PY_{nt-1}) = (Y_{nt-1}' \otimes I_{S+1})\text{vec}(P)$ where $\text{vec}(P) = \begin{bmatrix} \rho_u & 0_{1\times S} & K_{S,S+1}\text{vec}\left(\begin{bmatrix} \psi & I_s \otimes \rho_y \end{bmatrix}\right) \end{bmatrix}'$ with $K_{S,S+1}$ is the $S(S+1)$-dimensional square commutation matrix for which we have the exact equality $K_{S,S+1}\text{vec}\left(\begin{bmatrix}\psi & I_s \otimes \rho_y\end{bmatrix}\right) = \text{vec}(\begin{bmatrix}\psi & I_s \otimes \rho_y\end{bmatrix}')$ for ψ defined by (18) (Abadir and Magnus, 2005). Thus, $\text{vec}(P) = R\Psi$ for R the zero-one selection matrix,

$$\underset{(S+1)^2\times(S+2)}{R} = \begin{bmatrix} I_{S+1} & \underset{(S+1)\times(S+1)}{0} \\ \underset{S(S+1)\times(S+1)}{0} & K_{S,S+1}\begin{bmatrix} I_S & \underset{S\times 1}{0} \\ \underset{S^2\times S}{0} & \text{vec}(I_S)\end{bmatrix}\end{bmatrix} \times \begin{bmatrix} 0 & 1 & \underset{1\times S}{0} \\ \underset{S\times 1}{0} & \underset{S\times 1}{0} & \underset{S\times S}{0} \\ \underset{S\times 1}{0} & \underset{S\times 1}{0} & I_S \\ 1 & 0 & \underset{1\times S}{0}\end{bmatrix} \tag{C.7}$$

Inasmuch, (C.4) may be exactly restated in the form of (16) in the text. For future use, (C.4) additionally has the Wold decomposition

$$Y_{nt} = M_0\varepsilon_{nt} + \sum_{i=0}^{\infty} P^i\left(M_1 + PM_0\right)\varepsilon_{nt-i-1} \tag{C.8}$$

Finally, we must compute the variance-covariance matrix $E(V_{nt}V_{nt}') = \Omega$. Using (C.5),

$$\Omega = M_0\Sigma_\varepsilon M_0' + M_1\Sigma_\varepsilon M_1' \tag{C.9}$$

D Identification and estimation of reduced form

(Proof of Proposition 2) Here and henceforth we consider the case without a time decay in the signal-output relationship (Appendix H above). We wish to establish consistency and asymptotic normality of $\widehat{\Psi}$ (Equation (21)). The model (16) yields by-n representation $Y_n =$

$X_n\Psi + V_n$ with instruments Z_n for $Y_n = \begin{bmatrix} Y'_{n3} & \dots & Y'_{nT+2} \end{bmatrix}'$, $X_n = \begin{bmatrix} X'_{n3} & \dots & X'_{nT+2} \end{bmatrix}'$, $V_n = \begin{bmatrix} V'_{n3} & \dots & V'_{nT+2} \end{bmatrix}'$, and $Z_n = \begin{bmatrix} X'_{n2} & \dots & X'_{nT+1} \end{bmatrix}'$. However, since population means for the data are not known, in practice we require an initial within transformation to obtain small-sample analogues for each of these vectors.

$$\underbrace{\underbrace{\begin{bmatrix} \widetilde{Y}_{n3} \\ \vdots \\ \widetilde{Y}_{nT+2} \end{bmatrix}}_{\widetilde{Y}_n} - 1_T \otimes \overline{\overline{\widetilde{Y}}}_n}_{\widehat{Y}_n} = \underbrace{\left(\underbrace{\begin{bmatrix} \widetilde{X}_{n3} \\ \vdots \\ \widetilde{X}_{nT+2} \end{bmatrix}}_{\widetilde{X}_n} - 1_T \otimes \overline{\overline{\widetilde{X}}}_n \right)}_{\widehat{X}_n} \Psi + \underbrace{\underbrace{\begin{bmatrix} \widetilde{V}_{n3} \\ \vdots \\ \widetilde{V}_{nT+2} \end{bmatrix}}_{\widetilde{V}_n} - 1_T \otimes \overline{\overline{\widetilde{V}}}_n}_{\widehat{V}_n} \tag{D.1}$$

$$\widehat{Z}_n = \begin{bmatrix} \widetilde{X}'_{n2} & \dots & \widetilde{X}'_{nT+1} \end{bmatrix}' - 1_T \otimes \overline{\overline{\widetilde{Z}}}_n \quad \widetilde{Y}_{nt} = \begin{bmatrix} \widetilde{y}_{nt} \\ \widetilde{s}_{nt} \end{bmatrix} \quad \widetilde{X}_{nt} = \left(\begin{bmatrix} \widetilde{y}_{nt-1} \\ \widetilde{s}_{nt-1} \end{bmatrix}' \otimes I_{S+1} \right) R$$

$$\overline{\overline{\widetilde{Y}}}_n = \frac{1}{T} \sum_{t=3}^{T+2} \widetilde{Y}_{nt} \quad \overline{\overline{\widetilde{X}}}_n = \frac{1}{T} \sum_{t=3}^{T+2} \widetilde{X}_{nt} \quad \overline{\overline{\widetilde{Z}}}_n = \frac{1}{T} \sum_{t=2}^{T+1} \widetilde{X}_{nt}$$

for $1_T = [1, \dots, 1]'$ a $T \times 1$ vector of ones, $\widetilde{y}_{nt} = 100\Delta \ln Y_t$, $\widetilde{s}_{nt} = 100\Delta \ln S_t$ and $\{Y_{nt}\}$ and $\{S_{nt}\}$ raw output and signal data. (D.1) may be rewritten $Q_T\widetilde{Y}_n = Q_T\widetilde{X}_n\Psi + Q_T\widetilde{V}_n$ with instruments $Z_n = Q_T\widetilde{Z}_n$ for $Q_T = I_{(S+1)T} - (1/T)(1_T 1'_T \otimes I_{S+1})$ the $(S+1)T$ square within-groups transformation matrix with $Q'_T Q_T = Q_T$ (Arellano and Bover, 1995). So, an equivalent representation of (21) is

$$\widehat{\Psi} = \left[\sum_{n=1}^{N} \widehat{Z}_n' \widehat{X}_n\right]^{-1} \sum_{n=1}^{N} \widehat{Z}_n' \widehat{Y}_n = \left[\sum_{n=1}^{N} \widetilde{Z}_n' \underbrace{Q_T}_{Q_T' Q_T} \widetilde{X}_n\right]^{-1} \sum_{n=1}^{N} \widetilde{Z}_n' \underbrace{Q_T}_{Q_T' Q_T} \widetilde{Y}_n \tag{D.2}$$

Step 1. Before we may proceed, we must establish four auxiliary steps to the main results. First, we must find

$$E\left(\sum_{n=1}^{N} \widehat{Z}_n' \widehat{V}_n\right) = E \sum_{n=1}^{N} \widetilde{Z}_n' Q_T \widetilde{V}_n = N \underbrace{E(\widetilde{Z}_n' Q_T \widetilde{V}_n)}_{\text{(A)}} \tag{D.3}$$

This requires breaking down component (A). Note,

$$\text{(A)} \;:\; E(\widetilde{Z}_n' Q_T \widetilde{V}_n) = \underbrace{E(\widetilde{Z}_n' \widetilde{V}_n)}_{\text{(B)}} - \frac{1}{T} E(\underbrace{\widetilde{Z}_n'(1_T 1_T' \otimes I_{S+1})\widetilde{V}_n}_{\text{(C)}})$$

$$\text{(B)} \;:\; E(\widetilde{Z}_n' \widetilde{V}_n) = E \sum_{t=3}^{T+2} \widetilde{X}_{nt-1}' \widetilde{V}_{nt} = T E(\widetilde{X}_{nt-1}' \widetilde{V}_{nt}) = 0$$

$$\text{(C)} \;:\; \widetilde{Z}_n'(1_T 1_T' \otimes I_{S+1})\widetilde{V}_n = \underbrace{\left(\sum_{t=2}^{T+1} \widetilde{X}_{nt}'\right) \widetilde{V}_{n3}}_{\text{(D3)}} + \ldots + \underbrace{\left(\sum_{t=2}^{T+1} \widetilde{X}_{nt}'\right) \widetilde{V}_{nT+2}}_{(\text{D}T+2)}$$

where, $\forall\ \tau = 3, \ldots, T+2$ we have, utilizing the Wold representation (C.8), that

$$(\text{D}\tau) : \left(\sum_{t=2}^{T+1} \widetilde{X}_{nt}'\right) \widetilde{V}_{n\tau} = R' \sum_{t=2}^{T+1} (I_{S+1} \otimes M_0) \left(\widetilde{Y}_{nt-1} \otimes \varepsilon_{n\tau}\right) + (I_{S+1} \otimes M_1) \left(\widetilde{Y}_{nt-1} \otimes \varepsilon_{n\tau-1}\right)$$

$$E(\widetilde{Y}_{nt} \otimes \varepsilon_{n\tau}) = \begin{cases} 0 & \text{if } \tau > t \\ M_0 \text{vec}(\Sigma_\varepsilon) & \text{if } \tau = t \\ P^{t-\tau-1}(M_1 + PM_0)\text{vec}(\Sigma_\varepsilon) & \text{if } \tau < t \end{cases}$$

for $\Sigma_\varepsilon = E(\varepsilon_{nt}\varepsilon'_{nt})$. Thus, in expectation we have by backward induction,

$$\begin{aligned} E(\text{D}T+2) &= 0 \\ E(\text{D}T+1) &= R'(I_{S+1} \otimes M_1)M_0\text{vec}(\Sigma_\varepsilon) \\ E(\text{D}T) &= E(\text{D}T+1) \\ &+ R'[(I_{S+1} \otimes M_0)M_0 + (I_{S+1} \otimes M_1)(M_1 + PM_0)]\text{vec}(\Sigma_\varepsilon) \\ E(\text{D}T-q) &= E(\text{D}T-q+1) \\ &+ R'[(I_{S+1} \otimes M_0) + (I_{S+1} \otimes M_1)P]P^{q-1}(M_1 + PM_0)\text{vec}(\Sigma_\varepsilon) \\ &\forall \quad 1 \leq q \leq T-3 \end{aligned}$$

This implies that the expectation of (C) previous may be written

$$\begin{aligned} E[(C)] &= (T-1)R'(I_{S+1} \otimes M_1)M_0\text{vec}(\Sigma_\varepsilon) + (T-2)R'[(I_{S+1} \otimes M_0)M_0 \\ &+ (I_{S+1} \otimes M_1)(M_1 + PM_0)]\text{vec}(\Sigma_\varepsilon) \\ &+ R'\left[I_{S+1} \otimes M_0 + (I_{S+1} \otimes M_1)P\right] \times \left(\sum_{i=0}^{T-4}\sum_{j=0}^{i} P^j\right)(M_1 + PM_0)\text{vec}(\Sigma_\varepsilon) \end{aligned}$$

Because P is a triangular matrix with principal diagonal elements less than one, both P and $I_{S+1} - P$ are invertible. This latter point implies we have the geometric series

$\sum_{k=0}^{n-1} P^k = (I_{S+1} - P)^{-1}(I_{S+1} - P^n)$. When combined with the former point this yields

$$\sum_{i=0}^{T-4}\sum_{j=0}^{i} P^j = (T-3)(I_{S+1} - P)^{-1} - (I_{S+1} - P)^{-1}P^{-1}(I_{S+1} - P)^{-1}(I_{S+1} - P^{T-3})$$

And therefore in summary of the above points which began with (D.3),

$$E\left(\sum_{n=1}^{N} \widehat{Z}_n'\widehat{V}_n\right) = -N\left(\sum_{i=1}^{3} \frac{T-i}{T} C_i - \frac{1}{T} C_4(I_{S+1} - P^{T-3})C_5'\right) \tag{D.4}$$

For $\{C_i\}$ conformable constant matrices (cf. Alvarez and Arellano (2003) equation (16)).

Step 2. Next, we wish to find $\text{var}((TN)^{-1/2}\sum_{n=1}^{N} \widehat{Z}_n'\widehat{V}_n)$. To do so, recall that (C.4) gives the rule of motion $Y_{nt} = PY_{nt-1} + V_{nt}$ for the (unobservable) variables $\{Y_{nt}\}$ of which $\{\widehat{Y}_{nt}\}$ is a sample approximation. Note, were the actual series observable, one could write

$$\overline{Y}_{n(-2)} = \frac{1}{T}\sum_{t=1}^{T} Y_{nt} = \frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{Y}_{nt} - E(\widetilde{Y}_{nt})\right) = \overline{\widetilde{Y}}_{n(-2)} - E(\widetilde{Y}_{nt})$$

So, $Y_{nt} - \overline{Y}_{n(-2)} = (\widetilde{Y}_{nt} - E(\widetilde{Y}_{nt})) - (\overline{\widetilde{Y}}_{n(-2)} - E(\widetilde{Y}_{nt})) = \widetilde{Y}_{nt} - \overline{\widetilde{Y}}_{n(-2)}$ and thus,

$$\sum_{n=1}^{N} \widehat{Z}_n'\widehat{V}_n = \sum_{n=1}^{N} \widetilde{Z}_n' Q_T' Q_T \widetilde{V}_n = R'\sum_{n=1}^{N}\left(\begin{bmatrix}\widetilde{Y}_{n1} & \dots & \widetilde{Y}_{nT}\end{bmatrix}' \otimes I_{S+1}\right)' Q_T' Q_T \widetilde{V}_n$$

$$= R'\sum_{n=1}^{N}\left(\begin{bmatrix}Y_{n1}\dots Y_{nT}\end{bmatrix}' \otimes I_{S+1}\right)' Q_T V_n = \underbrace{R'\sum_{n=1}^{N}\sum_{t=3}^{T+2}(Y_{nt-2}\otimes V_{nt})}_{\text{(A)}} - \underbrace{TR'\sum_{n=1}^{N}(\overline{Y}_{n(-2)}\otimes\overline{V}_n)}_{\text{(B)}} \tag{D.5}$$

Using standard results for white noise VARMA processes (See Lutkepohl (2005)),

$$\text{(A)} \ : \ \text{var}\left((TN)^{-1/2}R'\sum_{n=1}^{N}\sum_{t=3}^{T+2}(Y_{nt-2}\otimes V_{nt})\right) = R'\text{var}\left(\sqrt{T}\sum_{t=3}^{T+2}(Y_{nt-2}\otimes V_{nt})\right)$$

$$\to R'\Sigma \text{ for } \Sigma = \Sigma_Y(0)\otimes\Omega(0) + \Sigma_Y(1)\otimes\Omega(1) + \Sigma_Y(1)'\otimes\Omega(1)'$$

$$\Omega(j) = E(V_{nt}V'_{nt-j}),\ \Sigma_Y(j) = E(Y_{nt}Y'_{nt-j}) \text{ and } \text{vec}(\Sigma_Y(0)) = (I_{(S+1)^2} - P\otimes P)^{-1}\text{vec}(\Omega)$$

A consistent estimator for Σ is

$$\widehat{\Sigma} = \widehat{\Sigma}_Y(0)\otimes\widehat{\Omega}(0) + \widehat{\Sigma}_Y(1)\otimes\widehat{\Omega}(1) + \widehat{\Sigma}_Y(1)'\otimes\widehat{\Omega}(1)') \tag{D.6}$$

$$\widehat{\Sigma}_Y(j) = ((T-j)N)^{-1}\sum_{n=2}^{N}\sum_{t=1}^{T}\widehat{Y}_{nt}\widehat{Y}'_{nt-j} \qquad \widehat{\Omega}(j) = ((T-j)N)^{-1}\sum_{n=2}^{N}\sum_{t=3}^{T+2}\widehat{V}_{nt}\widehat{V}'_{nt-j}$$

for $\widehat{V}_{nt} = \widehat{Y}_{nt} - \widehat{X}_{nt}\widehat{\Psi}$. Note, this estimator is consistent because the fourth moments of the observables are finite (See Hayashi (2000) pp. 123-4). Secondly, note that $\sqrt{T}(\overline{Y}_{n(-2)}, \overline{V}_n) \xrightarrow{d} (\zeta_n, \xi_n)$ for ζ_n and ξ_n jointly normally distributed random variables. Thus $\text{var}(T\overline{Y}_{n(-2)} \otimes \overline{V}_n) \to \text{var}(\zeta_n \otimes \xi_n)$, or, $\text{var}(\overline{Y}_{n(-2)} \otimes \overline{V}_n) \leq (1/T^2) \times$ constant and thus in other words, $\text{var}(\overline{Y}_{n(-2)} \otimes \overline{V}_n)$ is $O(T^{-2})$. Therefore,

$$\text{(B)} \ : \ \text{var}\left((TN)^{-1}T\sum_{n=1}^{N}(\overline{Y}_{n(-2)}\otimes\overline{V}_n)\right) = T\text{var}\left(\overline{Y}_{n(-2)}\otimes\overline{V}_n\right) = O(1/T)$$

Thus, the variance $\text{var}\left((TN)^{-1/2}\sum_{n=1}^{N} Z'_n V_n\right) \to$ constant as $T \to \infty$ regardless of the rate of growth of N (cf. Alvarez and Arellano equation (17)).

$$\text{var}\left((TN)^{-1/2}\sum_{n=1}^{N}\widehat{Z}_n'\widehat{V}_n\right) = R'\Sigma + O(1/T) \tag{D.7}$$

Step 3. Due to the previous arguments we have

$$\frac{1}{TN}\sum_{n=1}^{N}\widehat{Z}_n'\widehat{X}_n = \underbrace{\frac{1}{TN}\sum_{n=1}^{N}Z_n'X_n}_{\text{(A)}} - \underbrace{\frac{1}{T^2N}\sum_{n=1}^{N}Z_n'(1_T1_T'\otimes I_{S+1})X_n}_{\text{(B)}} \tag{D.8}$$

$$\sum_{n=1}^{N}Z_n'X_n = \sum_{n=1}^{N}\sum_{t=3}^{T}X_{nt-1}'X_{nt} = R'\left[\sum_{n=1}^{N}\sum_{t=3}^{T}(Y_{nt-2}\otimes I_{S+1})\left(Y_{nt-1}'\otimes I_{S+1}\right)\right]R$$

$$\text{(A)} \; : \; E\left(\frac{1}{TN}\sum_{n=1}^{N}Z_n'X_n\right) = R'\left[(\Sigma_Y P' + M_0\Sigma_\varepsilon M_1')\otimes I_{S+1}\right]R$$

because $E(Y_{nt}Y_{nt-1}') = P\Sigma_Y + M_1\Sigma_\varepsilon M_0'$ and

$$\sum_{n=1}^{N}Z_n'(1_T1_T'\otimes I_{S+1})X_n = \sum_{n=1}^{N}\left[\sum_{i=2}^{T+1}X_{ni}'\left(\sum_{j=3}^{T+2}X_{nj}\right)\right]$$

$$E\left[\sum_{i=2}^{T+1}X_{ni}'\left(\sum_{j=3}^{T+2}X_{nj}\right)\right] = TE(X_{nt-1}'X_{nt}) + s.o.t.$$

for $s.o.t.$ smaller order terms in the T-degree polynomial. This implies

$$\text{(B)} \; : \; E\left(\frac{1}{T^2N}\sum_{n=1}^{N}Z_n'(1_T1_T'\otimes I_{S+1})X_n\right) = O(T^{-1})$$

and thus in summary of the points which began with (D.8),

$$E\left[\frac{1}{TN}\sum_{n=1}^{N} Z_n' X_n\right] \rightarrow R'\left[(\Sigma_Y P' + M_0 \Sigma_\varepsilon M_1)\otimes I_{S+1}\right] R$$

Finally, note that (D.8) also implies

$$\text{var}\left(\frac{1}{TN}\sum_{n=1}^{N} Z_n' X_n\right) = \frac{1}{N^2}\text{var}\left(\frac{1}{T}\sum_{n=1}^{N} Z_n' X_n - \overline{Z}_n'\overline{X}_n\right) \rightarrow 0$$

for $\overline{Z}_n = (1/T)1_T' Z_n$ and $\overline{X}_n = (1/T)1_T' X_n$ since $\text{var}((1/T)\sum_{n=1}^{N} Z_n' X_n) = O(T^{-1})$ and additionally $\text{var}(\overline{Z}_n'\overline{X}_n) = O(T^{-2})$ for reasons similar to those previously discussed with respect to $\text{var}(\overline{Y}_{n(-2)}\otimes\overline{V}_n)$. Therefore, in summary of those points which began with (D.8) we have mean-squared convergence $\frac{1}{TN}\sum_{n=1}^{N} Z_n' X_n \overset{m.s.}{\rightarrow} R'\left[(\Sigma_Y P' + M_0\Sigma_\varepsilon M_1)\otimes I_{S+1}\right] R$, which itself always implies convergence in probability (cf. Alvarez and Arellano (2003) equation (18)).

$$\frac{1}{TN}\sum_{n=1}^{N} Z_n' X_n \overset{p}{\rightarrow} R'\left[(\Sigma_Y P' + M_0\Sigma_\varepsilon M_1)\otimes I_{S+1}\right] R \tag{D.9}$$

Step 4. Equation (D.4) implies that

$$\mu_{NT} = E\left((TN)^{-1/2}\sum_{n=1}^{N}\widehat{Z}_n'\widehat{V}_n\right) = -N^{1/2}\left(\sum_{i=1}^{3}\frac{T-i}{T^{3/2}}C_i - \frac{1}{T^{3/2}}C_4(I_{S+1}-P^{T-3})C_5'\right)$$

Therefore, using also (D.5) we have that

$$(TN)^{-1/2}\sum_{n=1}^{N}\widehat{Z}_n'\widehat{V}_n-\mu_{NT}=R'\sum_{n=1}^{N}\sum_{t=3}^{T+2}(Y_{nt-2}\otimes V_{nt})-R_{NT}$$

$$R_{NT}=(T/N)^{1/2}R'\sum_{n=1}^{N}(\overline{Y}_{n(-2)}\otimes\overline{V}_n)+\mu_{NT}$$

Evidently $\lim_{T\to\infty}R_{NT}=0$ which is sufficient to establish that R_{NT} is $o_p(1)$. Furthermore, due to the CLT for such processes we have

$$R'\sum_{n=1}^{N}\sum_{t=3}^{T+2}(Y_{nt-2}\otimes V_{nt})\overset{d}{\to}N\left(0,R'\Sigma R\right)$$

and so we have the following result analogous to Alvarez and Arellano (2003) equation (20).

$$(TN)^{-1/2}\sum_{n=1}^{N}\widehat{Z}_n'\widehat{V}_n-\mu_{NT}\overset{d}{\to}N\left(0,R'\Sigma R\right)\qquad\text{(D.10)}$$

Consistency. With the previous four steps completed, and in particular equations (D .4), (D.7), (D.9), and (D.10) in-hand, we may now establish the main results of consistency and asymptotic normality of $\widehat{\Psi}$. (D.4) implies $E((TN)^{-1}\sum_{n=1}^{N}\widehat{Z}_n'\widehat{V}_n)\to 0$ as $T\to\infty$ while (D.7) implies $\text{var}((TN)^{-1}\sum_{n=1}^{N}\widehat{Z}_n'\widehat{V}_n)=(TN)^{-1/2}(R'\Sigma+O(1/T))\to 0$. These two points mean that $(TN)^{-1}\sum_{n=1}^{N}\widehat{Z}_n'\widehat{V}_n\overset{m.s.}{\to}0$ which implies that $(TN)^{-1}\sum_{n=1}^{N}\widehat{Z}_n'\widehat{V}_n\overset{p}{\to}0$. So using also (D.9) we have

$$\widehat{\Psi} = \Psi_0 + \underbrace{\left[(TN)^{-1} \sum_{n=1}^{N} \widehat{Z}_n' \widehat{X}_n \right]^{-1}}_{\xrightarrow{p} \text{constant.}} \times \underbrace{(TN)^{-1} \sum_{n=1}^{N} \widehat{Z}_n' \widehat{V}_n}_{\xrightarrow{p} 0} \tag{D.11}$$

or in other words, $\widehat{\Psi} \xrightarrow{p} \Psi_0$ as $T \to \infty$ regardless of the asymptotic behavior of N, which is result (22) in Proposition 2.

Asymptotic Normality. Finally, (D.9) and (D.10) jointly imply by Cramér's theorem that

$$\left[(TN)^{-1} \sum_{n=1}^{N} \widehat{Z}_n' \widehat{X}_n \right]^{-1} \times \left[(TN)^{-1/2} \sum_{n=1}^{N} \widehat{Z}_n \widehat{V}_n - \mu_{NT} \right] \xrightarrow{d} N\left(0, \text{Avar}(\widehat{\Psi})\right)$$

where $\text{Avar}(\widehat{\Psi})$ is the explicit form of the asymptotic variance referred to in result (23) of Proposition 2.

$$\text{Avar}(\widehat{\Psi}) = E\left[\widehat{Z}_n' \widehat{X}_n \right]^{-1} R \Sigma R' E\left[\widehat{X}_n' \widehat{Z}_n \right]^{-1} \tag{D.12}$$

A consistent estimator $\widehat{\text{Avar}}(\widehat{\Psi})$ is therefore (24). The previous result also implies

$$\sqrt{TN}(\widehat{\Psi} - \Psi_0) - \left[(TN)^{-1} \sum_{n=1}^{N} \widehat{Z}_n' \widehat{X}_n \right]^{-1} \mu_{NT} \xrightarrow{d} N(0, \text{Avar}(\widehat{\Psi}))$$

$$\left[(TN)^{-1} \sum_{n=1}^{N} \widehat{Z}_n' \widehat{X}_n \right]^{-1} \mu_{NT} = \left[E\left((TN)^{-1} \sum_{n=1}^{N} \widehat{Z}_n' \widehat{X}_n \right) \right]^{-1} \mu_{NT} + R_{NT}^{\circ}$$

$$R^{\circ}_{NT} = \left[\left[(TN)^{-1} \sum_{n=1}^{N} \widehat{Z}'_n \widehat{X}_n \right]^{-1} - \left[E \left((TN)^{-1} \sum_{n=1}^{N} \widehat{Z}'_n \widehat{X}_n \right) \right]^{-1} \right] \mu_{NT}$$

Note, the previous results imply R°_{NT} is $o_p(1)$ as $T \to \infty$ regardless of the behavior of N. However, $E[(TN)^{-1} \sum_{n=1}^{N} \widehat{Z}'_n \widehat{X}_n]^{-1} \mu_{NT} = N^{-1/2} f(T)$ for $f(T)$ a function of T with $f(T) \to 0$ as $T \to \infty$ but $N^{-1/2} f(T) \to b(T)$ possibly not 0 if N is also increasing. This $b(T)$ is referred to in (23).

E Identification and estimation of structural parameters

E.1 Identification of structural parameters when $S = 1$

First let us establish identification for the case of $S = 1$ signal. In this case, there are no nuisance parameters, so the only pertinent restrictions on θ are those imposed by Ψ and Ω.

$$\pi = \underset{6\times 1}{\begin{bmatrix} \Psi' & \text{vech}(\Omega)' \end{bmatrix}'} \equiv \begin{bmatrix} \Psi_1 & \Psi_2 & \Psi_3 & \Omega_{11} & \Omega_{21} & \Omega_{22} \end{bmatrix}' = g(\underset{6\times 1}{\theta}) \qquad \text{(E.1)}$$

Ψ_i is the i-th row of Ψ and Ω_{ij} is the (i, j) entry of Ω. Establishing the global identifiability of structural parameters in nonlinear models is generally a nontrivial task (See Rothenberg (1971)). But despite the fact that the mapping g is nonlinear, it is evidently closed-form. Furthermore, there are equivalently 6 elements in both $[\Psi', \text{ vech}(\Omega)']'$ and θ. Thus, a sufficient condition for the global identifiability of θ is that the inverse g^{-1} exists. It does.

$$\theta = g^{-1}(\pi) : \begin{cases} \beta = (\Psi_3 - \Psi_1)/\Psi_2 \\ \rho_y = \Psi_3 \\ \rho_u = \Psi_1 \\ \sigma_y = \left| \sqrt{\frac{\Psi_2}{\Psi_3 - \Psi_1} \left(\Omega_{21} + \frac{\Psi_2 \Psi_3}{1+\Psi_3^2} \Omega_{22} \right) \Big/ \left(1 + \frac{\Psi_3}{1+\Psi_3^2} (\Psi_3 - \Psi_1) \right)} \right| \\ \sigma_s = \left| \sqrt{\frac{1}{1+\Psi_3^2} \left(\Omega_{22} - \left(\frac{\Psi_3 - \Psi_1}{\Psi_2} \right)^2 \sigma_y^2 \right)} \right| \\ \sigma_u = \left| \sqrt{\Omega_{11} - \sigma_y^2 - \Psi_2^2 \sigma_s^2} \right| \end{cases} \tag{E.2}$$

Therefore, θ is globally identifiable in any parameter space for which it is defined. The key to this conclusion is the VARMA representation (C.2). This is implied by the rule of motion (2) for y^*_{nt}, without which the observation matrix C is not invertible. Hence the rationale for (2): It provides the structure necessary for primitive identifiability, while remaining entirely consistent with macroeconomic theory.

E.2 Estimation of structural parameters when $S > 1$

We now consider the necessarily nonlinear problem of identification and estimation of θ when $S \geq 1$ given the estimators and restrictions $\widehat{\Psi}$, $\widehat{\Omega}$, and $\lambda' 1_S = 1$. The easiest way to begin any nonlinear identification exercise is to verify that necessary order conditions are satisfied. Specifically, in order to separately identify θ and λ, there must be at least as

Table E.1: Exact identification of structural θ and nuisance λ parameters.

Restriction	No. Elements	λ S	β S	ρ_y 1	ρ_u 1	σ_y 1	σ_u 1	σ $S(S+1)/2$	Total $n_\theta + S$
$\lambda' 1_S$	1	✓							
Ψ	$S+2$	✓	✓	✓	✓				
vech(Ω)	$(S+1)(S+2)/2$	✓	✓	✓	✓	✓	✓	✓	
Total	$n_\theta + S$								

much information in the $n_\theta + S$ vector of reduced form restrictions, π.

$$\underset{(n_\theta+S)\times 1}{\pi} = \begin{bmatrix} \lambda' 1_S & \Psi' & \text{vech}(\Omega)' \end{bmatrix}' \tag{E.3}$$

One way to cogently assess the dimension of either set of parameters is to enumerate them in a table. Such an analysis is presented in Table E.1. As described by this table, there are exactly as many structural and nuisance parameters as there are restrictions to be exploited for their identification. Therefore, the order condition is just satisfied for any $S \geq 1$.

When $S > 1$, π becomes too complicated to invert analytically. But Section E .3 establishes that one may compute analytically the square Jacobian,

$$\underset{(n_\theta+S)\times(n_\theta+S)}{J(\theta;\lambda)} = \frac{\partial \pi'}{\partial \begin{bmatrix} \theta' & \lambda' \end{bmatrix}} \tag{E.4}$$

Full rank of this matrix ensures point local identifiability of the structural parameters (Rothenberg (1971)). Furthermore, while π^{-1} is infeasible for $S > 1$, it is possible to compute an asymptotically efficient 2-step estimator on the basis of the Jacobian's inverse.

$$S > 1: \quad \begin{bmatrix} \widehat{\theta} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{\theta}^* \\ \widehat{\lambda}^* \end{bmatrix} + J(\widehat{\theta}^*, \widehat{\lambda}^*)^{-1} \left(\begin{bmatrix} 1 \\ \widehat{\Psi} \\ \text{vech}(\widehat{\Omega}) \end{bmatrix} - \begin{bmatrix} \widehat{\lambda}^{*\prime} 1_S \\ \Psi(\widehat{\theta}^*, \widehat{\lambda}^*) \\ \text{vech}(\Omega(\widehat{\theta}^*, \widehat{\lambda}^*)) \end{bmatrix} \right) \tag{E.5}$$

where $\widehat{\theta}^*$ and $\widehat{\lambda}^*$ are the signal-by-signal estimators for θ and λ constructed from each individual signal's consistent estimator (25). Section E.4 pedantically details the steps to be followed in constructing this two-step estimator and establishes its asymptotic equivalence to the infeasible efficient estimator based on inverting the mapping π directly.

E.3 Identification of structural parameters when $S > 1$

Now let us consider the case of $S > 1$ signals. In this case, the S nuisance parameters λ are not known and must be accounted for. It is not possible in this case to establish global identifiability, as was the case for $S = 1$ signal. The reason is that the mapping g becomes too complex. However, local identifiability may be established. The $(n_\theta + S) \times (n_\theta + S)$ Jacobian (E.4) may be written explicitly,

$$J = \begin{bmatrix} 0 & 1_S' \\ \partial \Psi / \partial \theta' & \partial \Psi / \partial \lambda' \\ \partial \text{vech}(\Omega) / \partial \theta' & \partial \text{vech}(\Omega) / \partial \lambda' \end{bmatrix} \qquad \underset{(S+2) \times n_\theta}{\partial \Psi / \partial \theta'} = \begin{bmatrix} \begin{bmatrix} \underset{1 \times S}{0} & \underset{1 \times S(S+1)/2}{0} & 1 & 0 & 0 & 0 \\ \underset{1 \times S}{0} & \underset{1 \times S(S+1)/2}{0} & 0 & 0 & 1 & 0 \end{bmatrix} \\ \partial \psi / \partial \theta' \end{bmatrix}$$

$$\underset{S\times n_\theta}{\partial\psi/\partial\theta'} = \left[\underset{S\times S}{\begin{bmatrix} \frac{-(\rho_y-\rho_u)\lambda_1}{\beta(1)^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{-(\rho_y-\rho_u)\lambda_S}{\beta(S)^2} \end{bmatrix}} \quad \underset{S\times S(S+1)/2}{0} \quad \underset{S\times 1}{\begin{bmatrix} \frac{\lambda_1}{\beta(1)} \\ \vdots \\ \frac{\lambda_S}{\beta(S)} \end{bmatrix}} \quad \underset{S\times 1}{0} \quad -\underset{S\times 1}{\begin{bmatrix} \frac{\lambda_1}{\beta(1)} \\ \vdots \\ \frac{\lambda_S}{\beta(S)} \end{bmatrix}} \quad \underset{S\times 1}{0}\right]$$

$$\underset{(S+2)\times S}{\partial\Psi/\partial\lambda'} = \begin{bmatrix} \underset{2\times S}{0} \\ \partial\psi/\partial\lambda' \end{bmatrix} \qquad \underset{S\times S}{\partial\psi/\partial\lambda'} = \begin{bmatrix} \frac{(\rho_y-\rho_u)}{\beta(1)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{(\rho_y-\rho_u)}{\beta(S)} \end{bmatrix}$$

$$\underset{\frac{(S+1)(S+2)}{2}\times n_\theta}{\partial \text{vech}(\Omega)/\partial\theta'} = \Delta_0\frac{\partial \text{vec}(M_0)}{\partial\theta'} + \Delta_1\frac{\partial \text{vec}(M_1)}{\partial\theta'} + \Delta_\varepsilon\frac{\partial \text{vech}(\Sigma_\varepsilon)}{\partial\theta'}$$

$$\underset{\frac{(S+1)(S+2)}{2}\times S}{\partial \text{vech}(\Omega)/\partial\lambda'} = \Delta_1\frac{\partial \text{vec}(M_1)}{\partial\lambda'}$$

$$\underset{\frac{(S+1)(S+2)}{2}\times(S+1)(S+2)}{\Delta_0} = D^+_{S+1}\left[(M_0\Sigma_\varepsilon)\otimes I_{S+1} + \left[I_{S+1}\otimes(M_0\Sigma_\varepsilon)\right]K_{S+1,S+2}\right]$$

$$\underset{\frac{(S+1)(S+2)}{2}\times(S+1)(S+2)}{\Delta_1} = D^+_{S+1}\left[(M_1\Sigma_\varepsilon)\otimes I_{S+1} + \left[I_{S+1}\otimes(M_1\Sigma_\varepsilon)\right]K_{S+1,S+2}\right]$$

$$\underset{\frac{(S+1)(S+2)}{2}\times\frac{(S+2)(S+3)}{2}}{\Delta_\varepsilon} = D^+_{S+1}\left[M_0\otimes M_0 + M_1\otimes M_1\right]D_{S+2}$$

$$\underset{(S+1)(S+2)\times n_\theta}{\frac{\partial \text{vec}(M_0)}{\partial \theta'}} = \begin{bmatrix} \underset{1\times S}{0} & \underset{1\times[S(S+1)/2+4]}{0} \\ I_S & 0 \\ \underset{(S+1)^2\times S}{0} & 0 \end{bmatrix}$$

$$\underset{(S+1)(S+2)\times n_\theta}{\frac{\partial \text{vec}(M_1)}{\partial \theta'}} = \begin{bmatrix} \underset{(2S+2)\times n_\theta}{0} \\ -\partial\psi/\partial\theta' \\ \begin{bmatrix} \underset{S^2\times S+\frac{S(S+1)}{2}}{0} & \text{vec}(I_S) & \underset{S^2\times 3}{0} \end{bmatrix} \end{bmatrix} \quad \underset{(S+1)(S+2)\times S}{\frac{\partial \text{vec}(M_1)}{\partial \lambda'}} = \begin{bmatrix} \underset{(2S+2)\times S}{0} \\ -\partial\psi/\partial\lambda' \\ \underset{S^2\times S}{0} \end{bmatrix}$$

$$\underset{(S+2)(S+3)/2\times n_\theta}{\frac{\partial \text{vech}(\Sigma_\varepsilon)}{\partial \theta'}} = \begin{bmatrix} \underset{1\times S}{0} & \underset{1\times S(S+1)/2}{0} & \underset{1\times 1}{0} & 2\sigma_y & \underset{1\times 1}{0} & \underset{1\times 1}{0} \\ \underset{(S+1)\times S}{0} & 0 & 0 & 0 & 0 & 0 \\ \underset{1\times S}{0} & 0 & 0 & 0 & 0 & 2\sigma_u \\ \underset{S\times S}{0} & 0 & 0 & 0 & 0 & 0 \\ \underset{S(S+1)/2\times S}{0} & \partial\text{vech}(\Sigma)/\partial\sigma' & 0 & 0 & 0 & 0 \end{bmatrix}$$

By definition, for $LL' = \Sigma$ the lower-left decomposition of Σ we have

$$\frac{\partial \text{vech}(\Sigma)}{\partial \sigma'} = \frac{\partial \text{vech}(LL')}{\partial \text{vech}(L)} = \underset{S(S+1)/2\times S(S+1)/2}{D_S^+ \left[L \otimes I_S + (I_S \otimes L)K_{S,S}\right] D_S}$$

We say θ are locally identifiable at a point (θ_0, λ_0) if $J(\theta_0, \lambda_0)$ is full column rank.

E.4 2-step estimation when $S > 1$.

Computing an estimator for the the structural parameters when $S > 1$ proceeds in three steps. We now describe these steps, and provide a proof for efficiency of the 2-step estimator.

1. Obtain a consistent first stage estimator $\widehat{\theta}^*$ using signal-by-signal estimates.

   (a) Obtain estimates for each signal-by-signal structural parameter using (25).

   (b) Obtain estimates $\widehat{\Psi}$ and $\widehat{\Omega}$ using the multiple signals under consideration jointly.

   (c) Individual signal estimates do not yield estimators for the off-diagonal elements of Σ. However, note that from (C.5) that we have $E(v_{nt}^s v_{nt}^{s'}) = \sigma_y^2 \beta\beta' + (1+\rho_y^2)\Sigma$. So a consistent estimator for Σ is $\widehat{\Sigma}^* = (1+\widehat{\rho}_y^2)^{-1}\left[\widehat{E}(v_{nt}^s v_{nt}^{s'}) - \widehat{\sigma}_y^{*2}\widehat{\beta}^*\widehat{\beta}^{*\prime}\right]$ where $\widehat{\beta}^* = [\beta(1)^* \;\ldots\; \beta(S)^*]'$ is comprised of estimates from each signal individually and $\widehat{E}(v_{nt}^s v_{nt}^{s'})$ is the lower right block of $\widehat{\Omega}$.

   (d) $\widehat{\theta}^* = (\widehat{\beta}^*, (\text{vech}(\widehat{\Sigma}^*))', \widehat{\rho}_y^*, \widehat{\sigma}_y^*, \widehat{\rho}_u^*, \widehat{\sigma}_u^*)'$ where $(\widehat{\rho}_y^*, \widehat{\sigma}_y^*, \widehat{\rho}_u^*, \widehat{\sigma}_u^*)$ come from the individual signal estimates for any signal, or an average, which remains consistent.

2. $\widehat{\lambda}^* = \text{diag}(\widehat{\beta}^*)\frac{1}{\widehat{\rho}_y^* - \widehat{\rho}_u^*}\widehat{\psi}$ is the signal-by-signal estimator for λ coming from (18).

3. We are looking for the infeasible estimator $[\widetilde{\theta}' \; \widetilde{\lambda}']'$ with the property $\widehat{\pi} = \pi([\widetilde{\theta}' \; \widetilde{\lambda}']')$ for $\widehat{\pi} = [1 \; \widehat{\Psi}' \; (\text{vech}(\widehat{\Omega}))']'$ the multiple signal reduced form estimates (with restrictions $\widehat{\lambda}' 1_S = 1$). A first order expansion of this infeasible estimator about the first stage estimators $[\widehat{\theta}^{*\prime} \; \widehat{\lambda}^{*\prime}]'$ yields $[\widetilde{\theta}' \; \widetilde{\lambda}']' \approx [\widehat{\theta}' \; \widehat{\lambda}']'$, the proposed estimator given by (E.5). The claim is that this estimator, restated here for convenience, is efficient. Note, the

premise is that θ is locally identifiable, which is sufficient for J to be invertible.

$$\begin{bmatrix} \widetilde{\theta} \\ \widetilde{\lambda} \end{bmatrix} \approx \begin{bmatrix} \widehat{\theta} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{\theta}^* \\ \widehat{\lambda}^* \end{bmatrix} + J(\widehat{\theta}^*, \widehat{\lambda}^*)^{-1} \left(\begin{bmatrix} 1 \\ \widehat{\Psi} \\ \text{vech}(\widehat{\Omega}) \end{bmatrix} - \begin{bmatrix} \widehat{\lambda}^{*\prime} 1_S \\ \Psi(\theta^*, \lambda^*) \\ \text{vech}(\Omega(\theta^*, \lambda^*)) \end{bmatrix} \right)$$

To establish efficiency of this two step estimator, first subtract the population values $[\theta_0' \ \lambda_0']'$ from both sides of the previous equation and premultiply by $\sqrt{TN}$.

$$\begin{aligned} \sqrt{TN} \left(\begin{bmatrix} \widehat{\theta} \\ \widehat{\lambda} \end{bmatrix} - \begin{bmatrix} \theta_0 \\ \lambda_0 \end{bmatrix} \right) &= \sqrt{TN} \left(\begin{bmatrix} \widehat{\theta}^* \\ \widehat{\lambda}^* \end{bmatrix} - \begin{bmatrix} \theta_0 \\ \lambda_0 \end{bmatrix} \right) \\ &+ J(\widehat{\theta}^*, \widehat{\lambda}^*)^{-1} \sqrt{TN} \left(\begin{bmatrix} 1 \\ \widehat{\Psi} \\ \text{vech}(\widehat{\Omega}) \end{bmatrix} - \begin{bmatrix} \widehat{\lambda}^{*\prime} 1_S \\ \Psi(\theta^*, \lambda^*) \\ \text{vech}(\Omega(\theta^*, \lambda^*)) \end{bmatrix} \right) \end{aligned}$$

We of course also have that

$$\begin{bmatrix} \widehat{\lambda}^{*\prime} 1_S \\ \Psi(\theta^*, \lambda^*) \\ \text{vech}(\Omega(\theta^*, \lambda^*)) \end{bmatrix} \approx \begin{bmatrix} \lambda_0' 1_S \\ \Psi(\theta_0, \lambda_0) \\ \text{vech}(\Omega(\theta_0, \lambda_0)) \end{bmatrix} + J(\theta_0, \lambda_0) \left(\begin{bmatrix} \widehat{\theta}^* \\ \widehat{\lambda}^* \end{bmatrix} - \begin{bmatrix} \theta_0 \\ \lambda_0 \end{bmatrix} \right)$$

Therefore, using the last two points,

$$\sqrt{TN}\left(\begin{bmatrix}\widehat{\theta}\\ \widehat{\lambda}\end{bmatrix}-\begin{bmatrix}\theta_0\\ \lambda_0\end{bmatrix}\right)=\sqrt{TN}\left(\begin{bmatrix}\widehat{\theta}^*\\ \widehat{\lambda}^*\end{bmatrix}-\begin{bmatrix}\theta_0\\ \lambda_0\end{bmatrix}\right)+J(\widehat{\theta}^*,\widehat{\lambda}^*)^{-1}\sqrt{TN}\left(\begin{bmatrix}1\\ \widehat{\Psi}\\ \text{vech}(\widehat{\Omega})\end{bmatrix}\right.$$
$$\left.-\left(\begin{bmatrix}\lambda_0'1_S\\ \Psi(\theta_0,\lambda_0)\\ \text{vech}(\Omega(\theta_0,\lambda_0))\end{bmatrix}+J(\theta_0,\lambda_0)\left(\begin{bmatrix}\widehat{\theta}^*\\ \widehat{\lambda}^*\end{bmatrix}-\begin{bmatrix}\theta_0\\ \lambda_0\end{bmatrix}\right)\right)\right)$$

This can be reorganized as,

$$\sqrt{TN}\left(\begin{bmatrix}\widehat{\theta}\\ \widehat{\lambda}\end{bmatrix}-\begin{bmatrix}\theta_0\\ \lambda_0\end{bmatrix}\right)=\underbrace{\left(I_{n_\theta+S}-J(\widehat{\theta}^*,\widehat{\lambda}^*)^{-1}J(\theta_0,\lambda_0)\right)}_{o_p(1)}\times\underbrace{\sqrt{TN}\left(\begin{bmatrix}\widehat{\theta}^*\\ \widehat{\lambda}^*\end{bmatrix}-\begin{bmatrix}\theta_0\\ \lambda_0\end{bmatrix}\right)}_{O_p(1)}$$
$$+\underbrace{J(\widehat{\theta}^*,\widehat{\lambda}^*)^{-1}\sqrt{TN}\left(\underbrace{\begin{bmatrix}1\\ \widehat{\Psi}\\ \text{vech}(\widehat{\Omega})\end{bmatrix}}_{\pi([\widetilde{\theta}'\ \widetilde{\lambda}']')}-\underbrace{\begin{bmatrix}\lambda_0'1_S\\ \Psi(\theta_0,\lambda_0)\\ \text{vech}(\Omega(\theta_0,\lambda_0))\end{bmatrix}}_{\pi([\theta_0'\ \lambda_0']')}\right)}_{\sqrt{TN}([\widetilde{\theta}'\ \widetilde{\lambda}']'-[\theta_0'\ \lambda_0']')+o_p(1)} \quad \text{(E.6)}$$

because

$$\pi([\widetilde{\theta}'\ \widetilde{\lambda}']')-\pi([\theta_0'\ \lambda_0']')=J(\theta_0,\lambda_0)([\widetilde{\theta}'\ \widetilde{\lambda}']'-[\theta_0'\ \lambda_0']')$$

$$\Longrightarrow J(\widehat{\theta}^*, \widehat{\lambda}^*)^{-1} J(\theta_0, \lambda_0)\sqrt{TN}([\widetilde{\theta}' \ \widetilde{\lambda}']' - [\theta_0' \ \lambda_0']') = \sqrt{TN}([\widetilde{\theta}' \ \widetilde{\lambda}']' - [\theta_0' \ \lambda_0']') + o_p(1)$$

Thus, (E.6) may be simplified to

$$\sqrt{TN}\left(\begin{bmatrix}\widehat{\theta}\\ \widehat{\lambda}\end{bmatrix} - \begin{bmatrix}\theta_0\\ \lambda_0\end{bmatrix}\right) = \sqrt{TN}\left(\begin{bmatrix}\widetilde{\theta}\\ \widetilde{\lambda}\end{bmatrix} - \begin{bmatrix}\theta_0\\ \lambda_0\end{bmatrix}\right) + o_p(1)$$

In other words, the 2 step feasible estimator is asymptotically equivalent to the infeasible estimator.

F Bootstrap

F.1 Bias

Bias of the estimator is written $b(\widehat{\theta}) = E(\widehat{\theta} - \theta_0)$ with θ_0 the population value: $\widehat{\theta} = \theta_0 + b(\widehat{\theta})$. A bootstrapped estimator for this bias is $b(\widehat{\theta}) \approx \frac{1}{B}\sum_{b=1}^{B} \widehat{\theta}^{(b)} - \widehat{\theta}$, where $\widehat{\theta}^{(b)}$ is the b-th draw from the distribution of $\widehat{\theta}$ and B is the number of draws. A bias-corrected estimator $\widehat{\theta}^\star$ is

$$\widehat{\theta} = \theta_0 + b(\widehat{\theta}) \approx \theta_0 + \frac{1}{B}\sum_{b=1}^{B} \widehat{\theta}^{(b)} - \widehat{\theta} \Longrightarrow \widehat{\theta}^\star = 2\widehat{\theta} - \frac{1}{B}\sum_{b=1}^{B} \widehat{\theta}^{(b)} \tag{F.1}$$

Pseudocode to obtain the required B draws follows.

1. Obtain $\widehat{\theta}$ from the data set $\{\{Y_{nt}\}_{t=1}^T\}_{n=1}^N$.
2. Obtain residuals $\{\{\widehat{U}_{nt}\}_{t=2}^T\}_{n=1}^N$ from the fitted values generated by $\widehat{\theta}$ for each $n = 1, \ldots, N$ and $t = 2, \ldots, T$ as follows:

(a) Assume $[0\ 1]\widehat{U}_{n1} = 0$ for each n. $M\widehat{U}_{n1} = 0_{2\times 1}$ because $M = [0_{2\times 1}, [m_{12}\ m_{22}]']$.

(b) Given $M\widehat{U}_{n1} = 0_{2\times 1}$, then $\widehat{U}_{n2} = \widehat{V}_{n2} = Y_{n2} - \widehat{P}Y_{n1}$ for each n.

(c) For all $t = 3, \ldots, T$ and each n, $\widehat{U}_{nt} = \widehat{V}_{nt} - M(\widehat{\theta}_s)\widehat{U}_{nt-1}$ for $\widehat{V}_{nt} = Y_{nt} - \widehat{P}Y_{nt-1}$.

3. Take as given $\{\{\widehat{U}_{nt}\}_{t=2}^{T}\}_{n=1}^{N}$ from the last step and draw with replacement as follows:

(a) Define the set of fitted vectors $\{\widehat{U}_t\}_{t=2}^{T}$ for $\widehat{U}_t = [\widehat{U}_{1t}' \ \ldots \widehat{U}_{Nt}']'$ from $\{\{\widehat{U}_{nt}\}_{t=2}^{T}\}_{n=1}^{N}$.

(b) Draw with replacement from $\{\widehat{U}_t\}_{t=2}^{T}$ $T-1$ times. One must draw from joint set of fitted values across regions n because U_{nt} may be correlated across n. Label these draws $\{\widehat{U}_t^{(1)}\}_{t=2}^{T}$.

(c) Repeat the previous step $B-1$ times, resulting in the set $\{\{\widehat{U}_t^{(b)}\}_{t=2}^{T}\}_{b=1}^{B}$.

(d) The draws from the last step may now also be written $\{\{\{\widehat{U}_{nt}^{(b)}\}_{t=2}^{T}\}_{n=1}^{N}\}_{b=1}^{B}$.

4. Take as given the draws $\{\{\{\widehat{U}_{nt}^{(b)}\}_{t=2}^{T}\}_{n=1}^{N}\}_{b=1}^{B}$ from the last step and generate draws from the observables for each $n = 1, \ldots, N$ and $b = 1, \ldots, B$ as follows:

(a) Set $\widehat{Y}_{n1}^{(b)} = Y_{n1}$.

(b) Compute $\widehat{Y}_{n2}^{(b)} = \widehat{P}\widehat{Y}_{n1}^{(b)} + \widehat{U}_{n2}^{(b)}$.

(c) Compute $\widehat{Y}_{nt}^{(b)} = \widehat{P}\widehat{Y}_{nt-1}^{(b)} + \widehat{U}_{nt}^{(b)} + M(\widehat{\theta})\widehat{U}_{nt-1}^{(b)}\ \forall\ t = 3, \ldots, T$.

(d) The resulting set of draws from the observables is written $\{\{\{\widehat{Y}_{nt}^{(b)}\}_{t=1}^{T}\}_{n=1}^{N}\}_{b=1}^{B}$.

5. For each $b = 1, \ldots B$, Obtain the estimator $\widehat{\theta}$ from each synthetic data set $\{\{\widehat{Y}_{nt}^{(b)}\}_{t=1}^{T}\}_{n=1}^{N}$. This results in the set of B draws $\{\widehat{\theta}_s^{(b)}\}_{b=1}^{B}$ from which to estimate the bias of $\widehat{\theta}$.

F .2 Confidence intervals

Take as given the draws $\{\widehat{\theta}^{(b)}\}_{b=1}^{B}$ from the above estimation of bias. Recalling that θ has n_θ elements, write $\widehat{\theta} = (\widehat{\theta}_1, \ldots, \widehat{\theta}_{n_\theta})'$ and $\widehat{\theta}^{(b)} = (\widehat{\theta}_1^{(b)}, \ldots, \widehat{\theta}_{n_\theta}^{(b)})'$ for scalar structural parameter estimators $\{\widehat{\theta}_i\}$ and b-th draw $\{\widehat{\theta}_i^{(b)}\}$. Define the corresponding bias-corrected estimator $\widehat{\theta}_i^\star = 2\widehat{\theta}_i - \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_i^{(b)}$ and bias-corrected bootstrap draws

$$\widehat{\theta}_i^{\star(b)} = 2\left(2\widehat{\theta}_i^{(b)} - \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_i^{(b)}\right) - \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_i^{(b)}$$

Note that this also implies

$$\widehat{\theta}_i^{\star(b)} - \widehat{\theta}_i^\star = \left(4\widehat{\theta}_i^{(b)} - 3\frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_i^{(b)}\right) - \left(2\widehat{\theta}_i^{(b)} - \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_i^{(b)}\right) = 2\left(\widehat{\theta}_i^{(b)} - \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_i^{(b)}\right)$$

Then, for each i define the set

$$\mathbb{S}_i = \left\{\widehat{\theta}_i^{\star(1)} - \widehat{\theta}_i^\star, \ldots, \widehat{\theta}_i^{(B)} - \widehat{\theta}_i^\star\right\} = \left\{2\left(\widehat{\theta}_i^{(1)} - \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_i^{(b)}\right), \ldots, 2\left(\widehat{\theta}_i^{(B)} - \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_i^{(b)}\right)\right\}$$

So, for q_i is the quantile function of $\mathbb{S}_i$, the $1-\alpha$ bootstrap confidence interval of $\widehat{\theta}_i^\star$ is

$$C_{1-\alpha}(\widehat{\theta}_i^\star) = \left[\widehat{\theta}_i^\star - q_i(1-\alpha/2), \quad \widehat{\theta}_i^\star - q_i(\alpha/2)\right]$$

The confidence interval $C_{1-\alpha}$ makes no assumptions about the distribution of U_{nt} in order to be correctly sized. See Hansen (2015) p. 233. A similar method is followed to compute intervals for $\widehat{\phi}$ using the draws from $\widehat{\theta}$'s distribution as inputs to (9).

F.3 Confidence bands

Take as given the draws $\{\{\{\widehat{Y}_{nt}^{(b)}\}_{t=1}^{T}\}_{n=1}^{N}\}_{b=1}^{B}$ and $\{\widehat{\theta}^{(b)}\}_{b=1}^{B}$ from the above estimation of bias. Label $\widehat{Y}_{nt}^{(b)} = [\widehat{y}_{nt}^{(b)}\ \widehat{s}_{nt}^{(b)\prime}]'$, collect $\{\widehat{\beta}^{(b)}\}_{b=1}^{B}$ the first elements of $\{\widehat{\theta}^{(b)}\}_{b=1}^{B}$, and compute $\{\widehat{\phi}^{(b)}\}_{b=1}^{B}$ using each $\widehat{\theta}^{(b)}$ and (9). For each $t = 3, \ldots, T$, $n = 1, \ldots, N$, and $b = 1, \ldots, B$

$$\widehat{x}_{nt}^{(b)} = (1 - \widehat{\phi}^{(b)})\widehat{y}_{nt}^{(b)} + \widehat{\phi}^{(b)}\widehat{\beta}^{(b)-1}\widehat{s}_{nt}^{(b)}$$

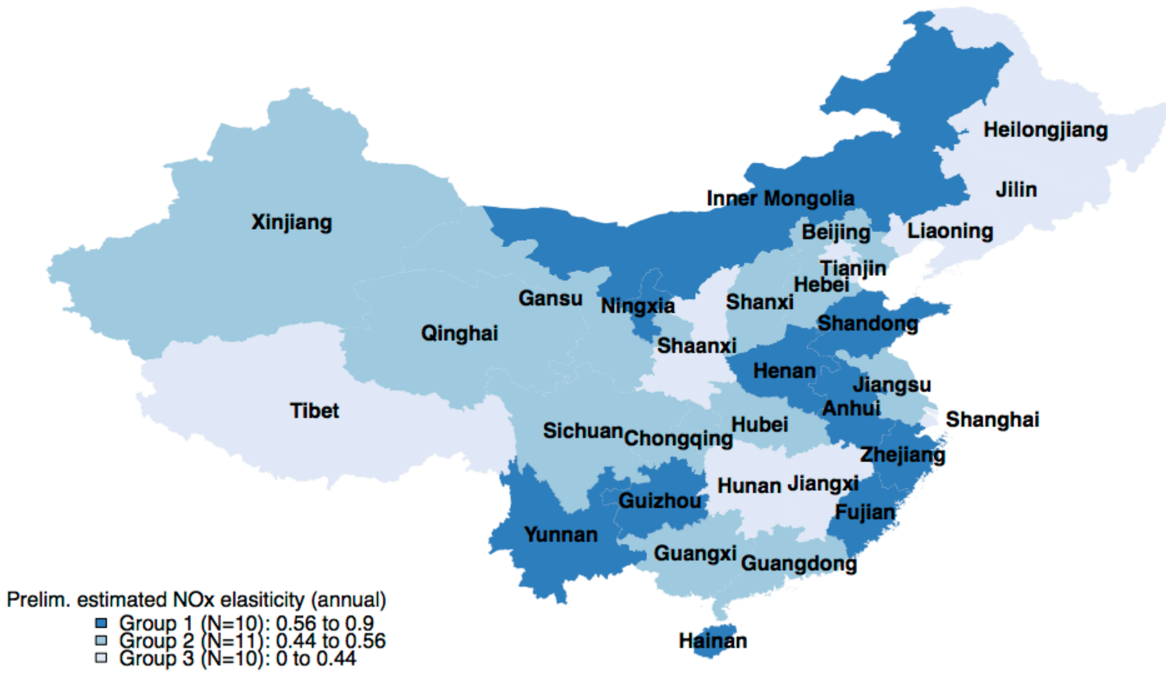
Define the deviations $\mathbb{S}$ and confidence interval $C_{1-\alpha}$ for $\widehat{x}_{nt}$ as in the above confidence interval computation for each n and t using $\widehat{x}_{nt}$ and $\{\widehat{x}_{nt}^{(b)}\}_{b=1}^{B}$.

G Results

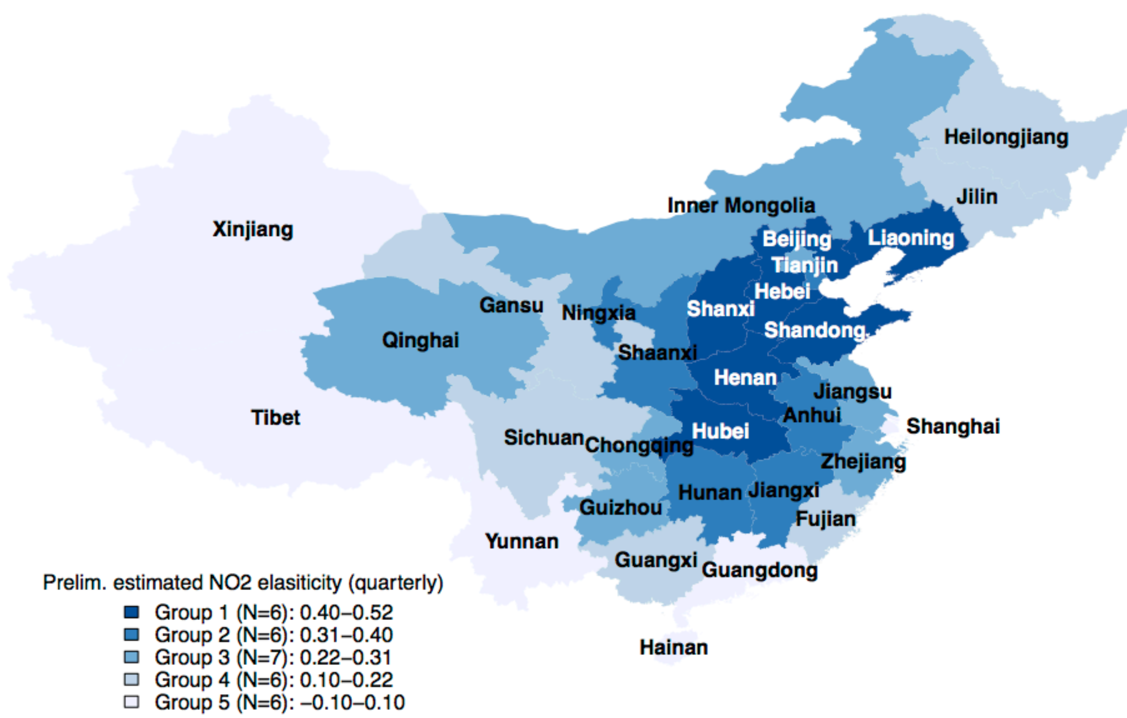
Beginning on the next page, find supplementary figures and tables to the main paper's empirical results.

Figure G.1: Estimation groupings.

(a) Annual sample groupings.



(b) Quarterly sample groupings.



Notes: "Preliminary estimated NO_2 elasticity" refers to the estimates $\widehat{\beta}$ computed using NO_2 columns data in Figure A.5 for the annual sample, and Figure A.6 for the quarterly sample.

Table G.1: Structural parameter estimates by signal: Annual sample, Groups 1 and 2.

Group 1 ($N = 10$).

	L	N	F	E	C
β	-2.24	0.40	-23.64	-2.48	55.01*
	(-13.88,17.16)	(-10.05,12.63)	(-24.79,-22.29)	(-8.01,8.92)	(17.4,108.88)
σ	7.07*	11.90*	9.53*	7.23*	-2.49
	(0.89,10.89)	(4.27,18.26)	(0.57,20.15)	(2.74,12.11)	(-16.39,10.71)
ρ_y	-0.45	1.03	-1.37*	0.97*	-0.37
	(-17.57,17.33)	(-11.98,12.36)	(-1.40,-1.35)	(0.16,1.58)	(-2.55,2.32)
σ_y	1.22	2.62	1.23*	3.38	-0.17
	(-4.26,4.49)	(-4.70,7.05)	(0.25,2.15)	(-0.08,6.55)	(-4.40,2.37)
ρ_u	0.65	0.58	0.41	0.57	0.50
	(-1.66.4.48)	(-0.98,2.27)	(-1.53,2.71)	(-0.83,3.25)	(-1.62,2.62)
σ_u	0.01	38.43*	2.49	2.23	3.45*
	(-3.52,6.15)	(32.80,47.45)	(-0.28,5.19)	(-1.49,7.17)	(0.22,7.74)

Group 2 ($N = 11$).

	L	N	F	E	C
β	-2.96	-0.63	-2.54*	-3.04	0.46*
	(-20.63,12.04)	(-13.81,9.04)	(-2.54,-2.53)	(-47.34,35.82)	(0.39,0.53)
σ	6.06	0.65	3.58*	10.60*	0.26
	(-0.23,10.56)	(-0.03,12.82)	(1.57,5.96)	(1.96,18.99)	(-3.22,3.48)
ρ_y	-1.42	-0.60	3.50*	0.26	-0.95*
	(-30.35,17.93)	(-7.33,6.24)	(3.49,3.50)	(-18.54,20.14)	(-1.31,-0.67)
σ_y	1.76	8.38*	13.50*	4.47	27.90*
	(-3.42,5.43)	(3.24,12.70)	(6.34,17.41)	(-0.01,7.44)	(18.48,37.96)
ρ_u	0.36	0.33	0.73	0.50	1.01*
	(-3.87,4.36)	(-2.40,2.70)	(-0.29,1.36)	(-1.21,2.37)	(0.79,1.32)
σ_u	1.20	5.46*	2.57	3.79	35.34*
	(-3.01,6.78)	(0.47,12.68)	(-1.68,7.45)	(-8.00,2.04)	(22.71,56.81)

Notes: $T = 16$. *Significiant at 95% confidence level (confidence interval). L: Luminosity. N: NO_x emissions. F: Freight volume. E: Electricity generation. C: Cement production. See Figure 8 (a) for annual Group definitions.

Table G.2: Structural parameter estimates by signal: Annual sample, Groups 3 and Pooled.

Group 3 ($N = 10$).

	L	N	F	E	C
β	-6.03	3.08	-1.48	-0.40	5.29*
	(-18.20,11.54)	(-53.80,56.00)	(-15.36,18.91)	(-5.23,8.23)	(5.06,5.44)
σ	8.68*	18.11*	20.46*	8.30*	4.91*
	(1.83,13.44)	(3.29,27.28)	(12.83,30.88)	(1.39,13.85)	(1.06,6.87)
ρ_y	-1.36	12.56	-1.00	-1.03	-1.02*
	(-28.00,15.45)	(-24.22,33.40)	(-35.79,14.40)	(-15.92,21.68)	(-1.22,-1.16)
σ_y	2.43	-0.13	3.24	10.15*	3.81*
	(-3.71,6.75)	(-4.93,2.64)	(-5.76,8.99)	(4.76,21.68)	(-1.22,-1.16)
ρ_u	0.37	0.55	0.30	0.45	0.54
	(-3.89,4.29)	(-2.39,3.47)	(-5.58,8.59)	(-2.31,4.09)	(-0.55,1.58)
σ_u	3.61	3.82	3.48	3.99	5.95*
	(-0.22,8.72)	(-0.47,9.78)	(-6.98,13.05)	(-0.45,9.22)	(4.26,7.57)

Pooled ($N = 31$).

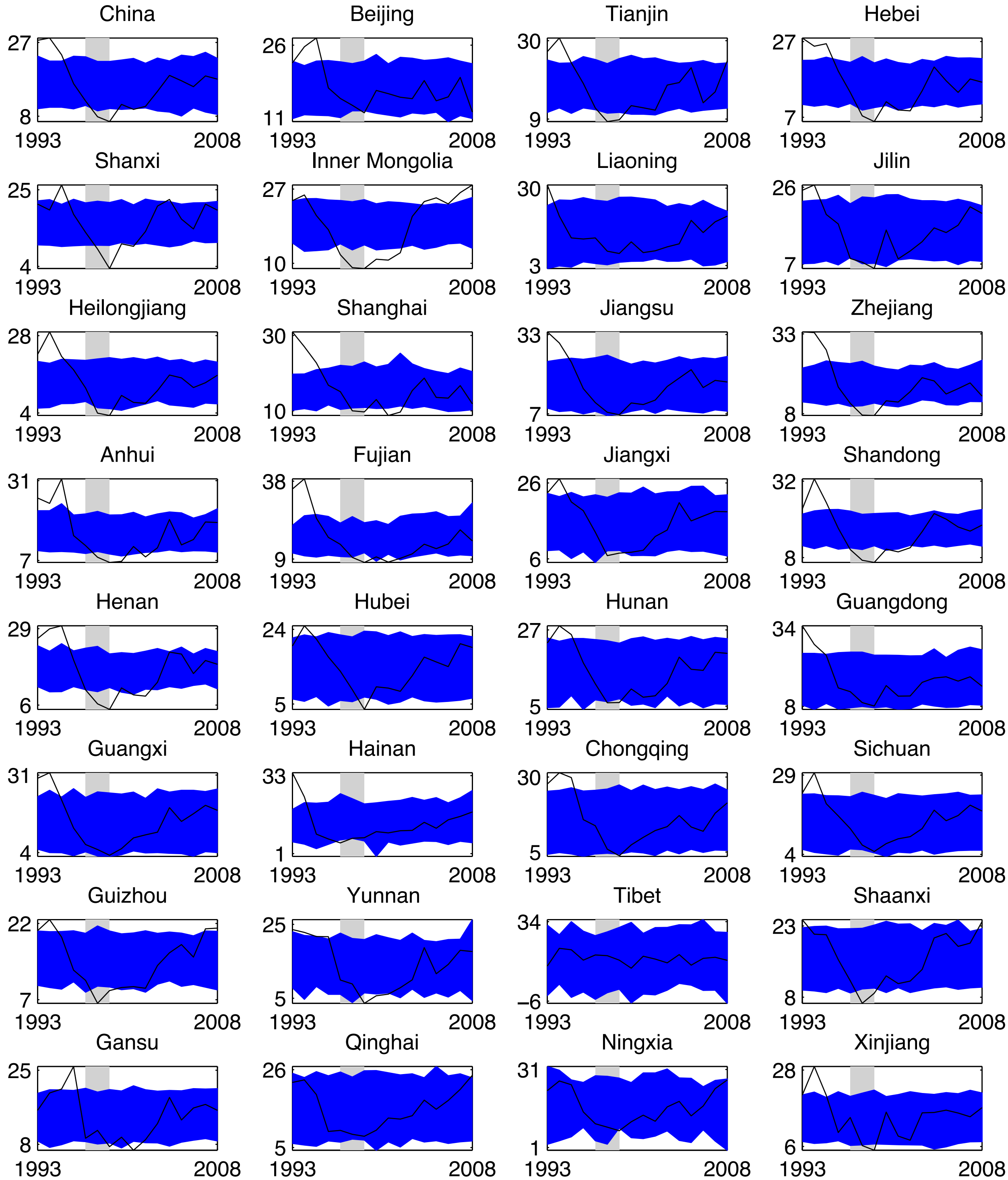
	L	N	F	E	C
β	-4.14	1.07	0.01	-3.02	5.54
	(-19.47,9.03)	(-39.81,38.66)	(-37.74,43.09)	(-22.91,18.76)	(-0.08,10.26)
σ	6.93*	14.50*	23.07*	12.02*	1.34
	(0.21,11.44)	(4.37,21.88)	(5.72,34.33)	(4.02,18.68)	(-7.04,8.74)
ρ_y	-1.05	0.61	-0.84	-0.29	-0.62
	(-20.01,14.75)	(-23.83,26.80)	(-20.81,21.74)	(-24.19,18.29)	(-1.43,0.03)
σ_y	0.47	1.63	2.23	2.09	2.46
	(-5.02,4.10)	(-23.82,26.80)	(-6.46,6.28)	(-3.58,5.78)	(-0.11,4.44)
ρ_u	0.58	0.59	0.40	0.61	0.50
	(-2.02,5.05)	(-1.31,2.70)	(-4.01,8.67)	(-1.55,3.57)	(-0.92,2.14)
σ_u	1.91	9.87*	9.43*	-3.91	4.39*
	(-2.24,8.04)	(5.76,16.29)	(0.62,20.20)	(-7.69,3.34)	(1.55,8.07)

Notes: $T = 16$. * Significiant at 95% confidence level (confidence interval). L: Luminosity. N: NO_x emissions. F: Freight volume. E: Electricity generation. C: Cement production. See Figure 8 (a) for annual Group definitions.

Figure G.2: Annual luminosity 95% confidence bands.

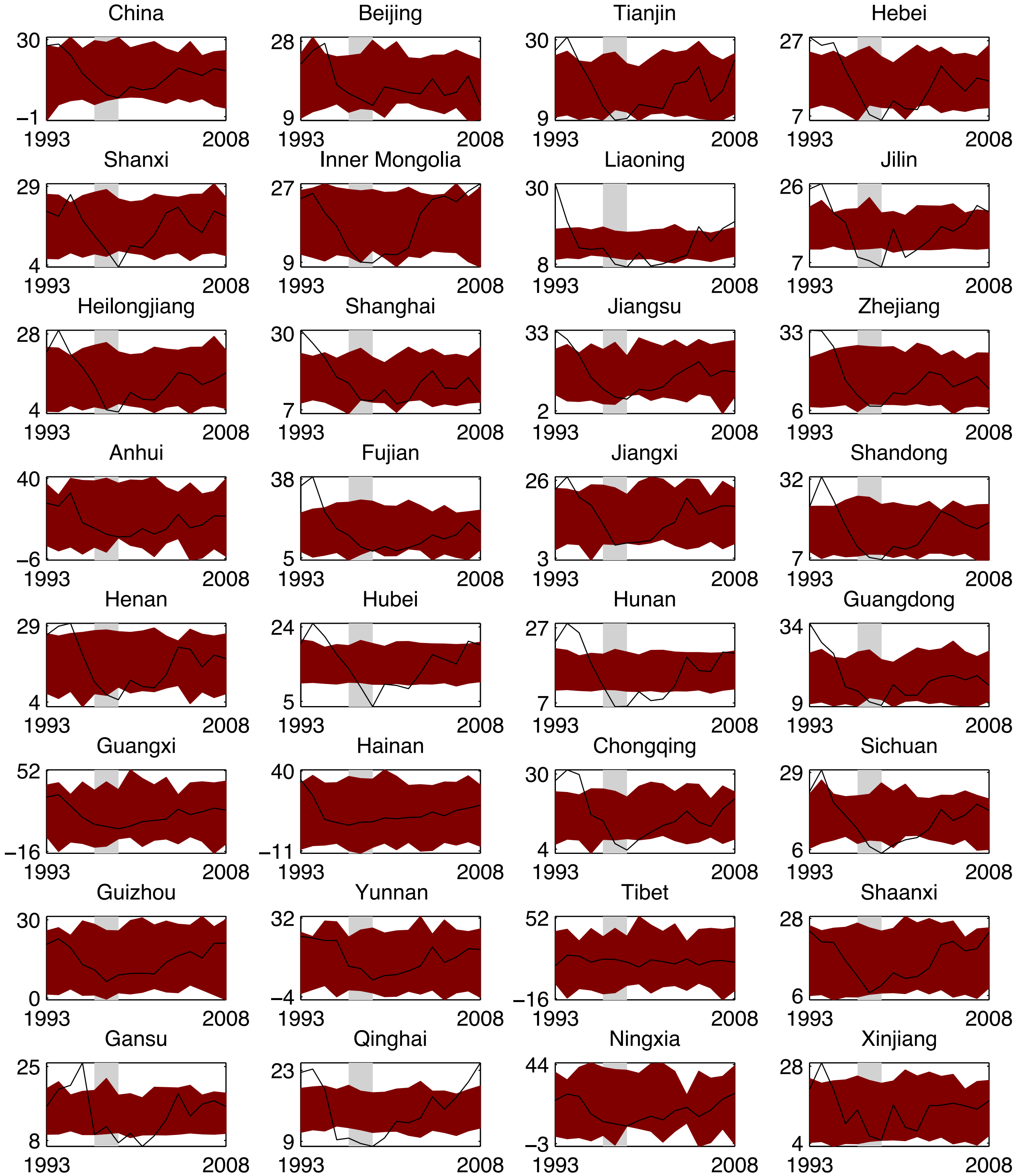
Notes: Black line: Annual reported % change, regional output. Gray shading: 1997-1999 Asian Financial Crisis period. Pink shading: confidence bands. Confidence bands are computed by bootstrap using each respective Group 1-3 estimates. China computed using "Pooled" estimates.

Figure G.3: 95% Confidence bands: Annual sample, NO_x emissions.



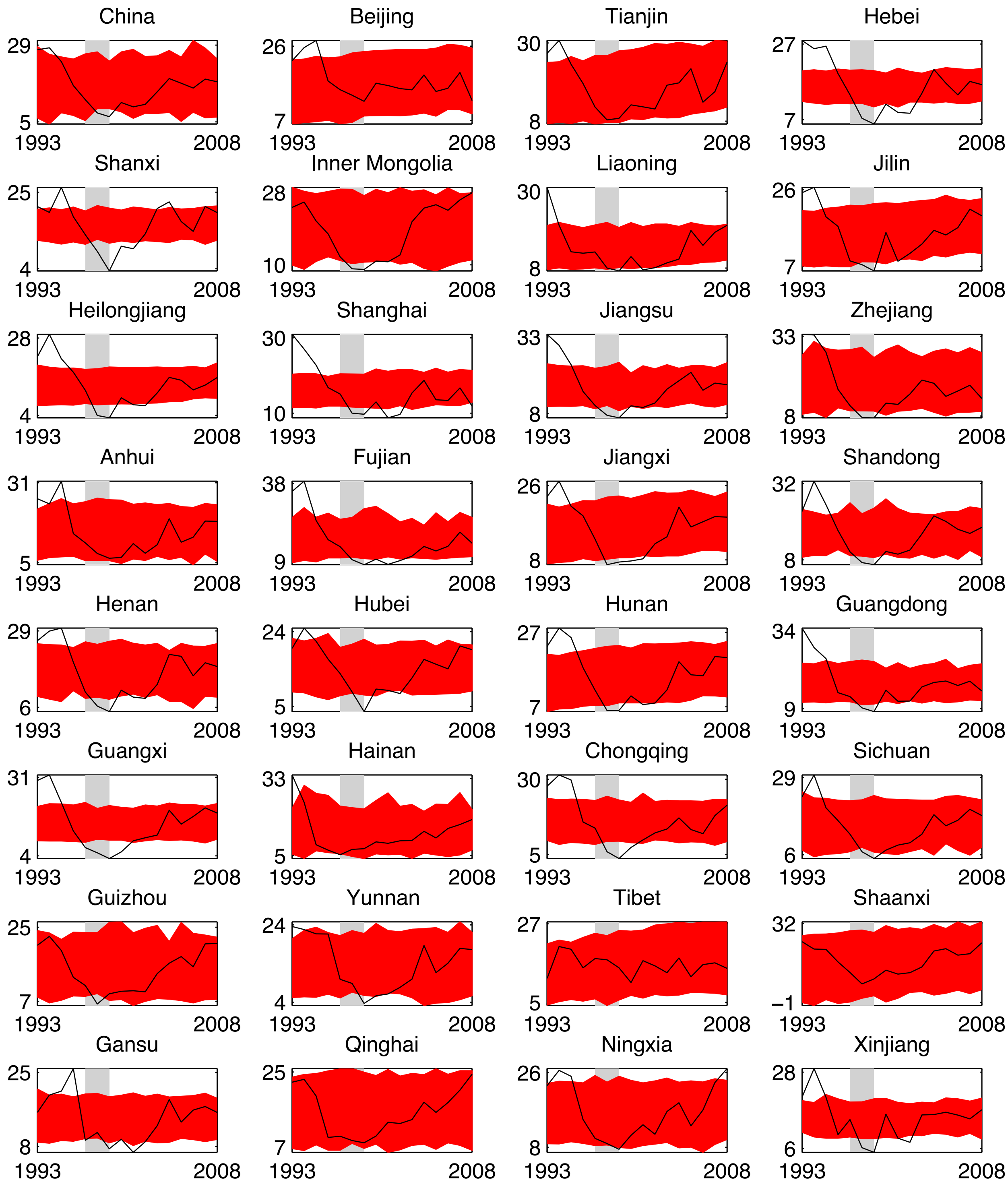
Notes: 1993-2008. Black line is reported annualized percentage change in GDP.

Figure G.4: 95% Confidence bands: Annual sample, freight traffic.



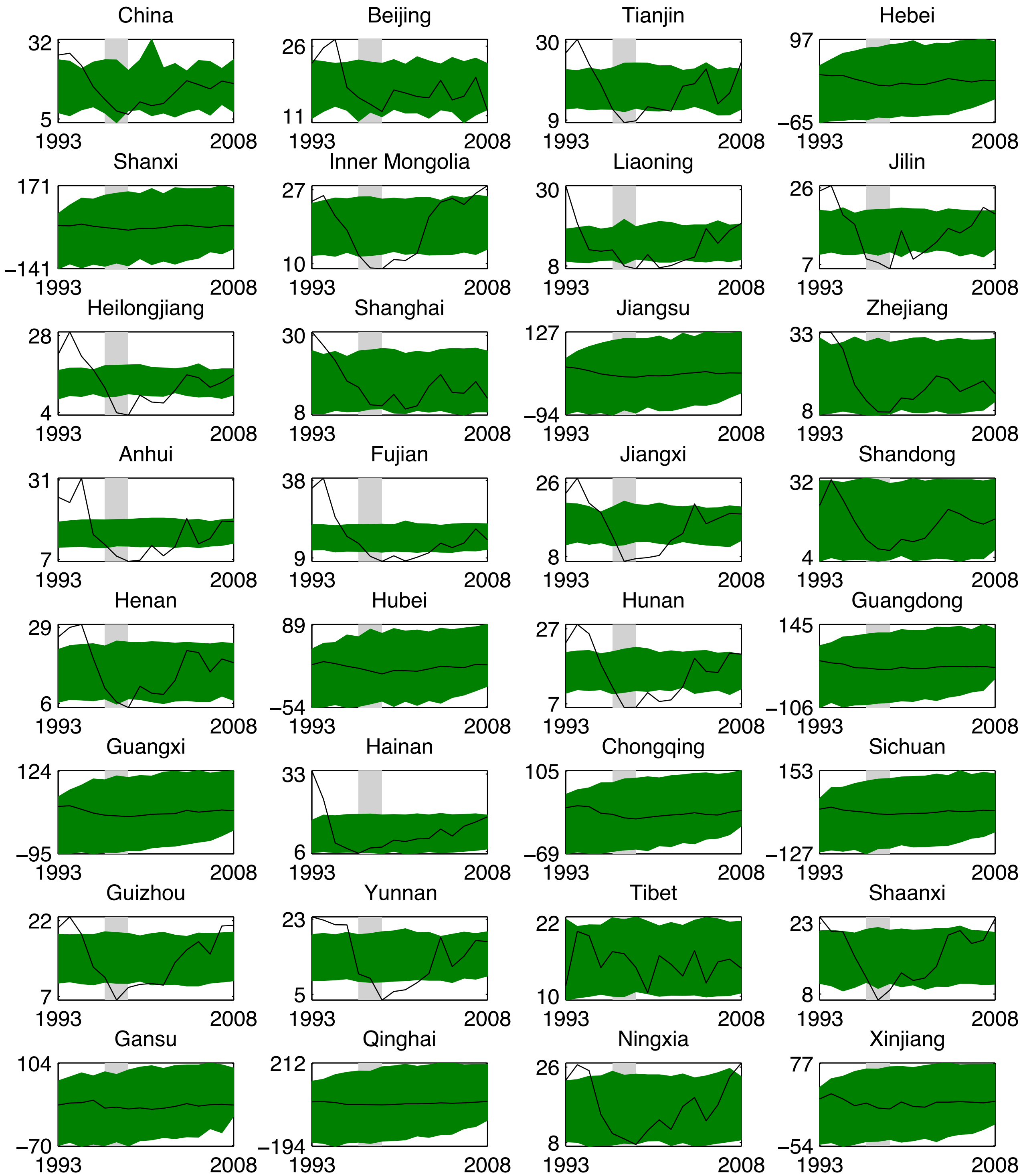
Notes: 1993-2008. Black line is reported annualized percentage change in GDP.

Figure G.5: 95% Confidence bands: Annual sample, electricity generation.



Notes: 1993-2008. Black line is reported annualized percentage change in GDP.

Figure G.6: 95% Confidence bands: Annual sample, cement production.



Notes: 1993-2008. Black line is reported annualized percentage change in GDP.

Table G.3: Structural parameter estimates by signal: Quarterly sample, Groups 1-4.

Group 1 ($N = 6$).

	N	E	C
β	1.09	5.99	-26.30
	(-2.31,2.72)	(-12.97,29.92)	(-182.22,151.59)
σ	7.87*	10.32	52.74
	(1.50,13.27)	(-1.98,22.50)	(-34.38,108.36)
ρ_y	0.81*	0.57	-0.41
	(0.15,1.48)	(-12.75,13.39)	(-16.30,14.91)
σ_y	10.30*	1.68	2.28
	(5.16,17.74)	(-4.75,6.49)	(-5.78,6.43)
ρ_u	-0.06	0.16	0.33
	(-2.36,1.85)	(-4.75,10.12)	(-7.91,10.56)
σ_u	1.92	9.02*	9.47*
	(-5.38,10.48)	(4.60,17.74)	(4.94,18.72)

Group 2 ($N = 6$).

	N	E	C
β	5.49	-2.93	-0.13
	(-24.40,41.96)	(-11.08,9.61)	(-54.54,43.70)
σ	53.89*	17.25*	23.34
	(10.89,74.35)	(3.92,27.17)	(-5.19,44.84)
ρ_y	-0.10	2.30	3.09
	(-32.89,24.06)	(-14.97,23.83)	(-5.19,44.84)
σ_y	3.58	3.71	4.94
	(-6.55,10.13)	(-2.04,9.76)	(-0.58,8.99)
ρ_u	0.61	0.45	0.49
	(-0.84,2.24)	(-1.22,2.71)	(-0.96,2.15)
σ_u	15.73*	10.71*	5.32*
	(6.52,25.69)	(6.35,17.92)	(0.74,13.26)

Group 3 ($N = 7$).

	N	E	C
β	-0.11	3.08	2.11
	(-4.38,6.57)	(-11.99,18.60)	(-79.19,76.15)
σ	20.49*	9.70	46.27
	(2.70,36.51)	(-1.40,24.41)	(-0.59,74.01)
ρ_y	0.76	0.42	0.32
	(-7.90,11.64)	(-7.25,6.23)	(-23.88,31.51)
σ_y	28.30*	2.94	6.99
	(16.56,41.52)	(-5.98,10,00)	(-3.12,13.29)
ρ_u	-0.10	0.36	0.35
	(-3.77,5.68)	(-1.86,3.78)	(-2.91,4.93)
σ_u	10.37	10.52*	8.13*
	(-0.39,27.01)	(3.47,22.51)	(1.00,20.11)

Group 4 ($N = 6$).

	N	E	C
β	-0.13	2.70	-1.67
	(-5.08,9.07)	(-5.37,15.86)	(-41.29,36.54)
σ	25.87	14.97	48.02*
	(-0.53,37.22)	(-2.81,20.98)	(6.77,65.49)
ρ_y	0.67	1.04	-0.66
	(-7.15,12.82)	(-16.61,16.34)	(-17.56,19.07)
σ_y	22.82*	4.03	2.16
	(11.78,39.62)	(-3.90,10.85)	(-4.64,12.50)
ρ_u	-0.18	-0.16	0.40
	(-5.56,5.47)	(-5.54,7.21)	(-3.94,5.47)
σ_u	10.34*	9.54*	9.41*
	(0.81,23.79)	(3.77,21.07)	(3.39,17.69)

Notes: $T = 31$. * Significiant at 95% confidence level (CI). N: NO_2 columns. E: Electricity generation. C: Cement production. See Figure 8 (b) for quarterly Group definitions.

Table G.4: Structural parameter estimates by signal: Quarterly sample, Groups 5 and Pooled.

Group 5 ($N = 6$).

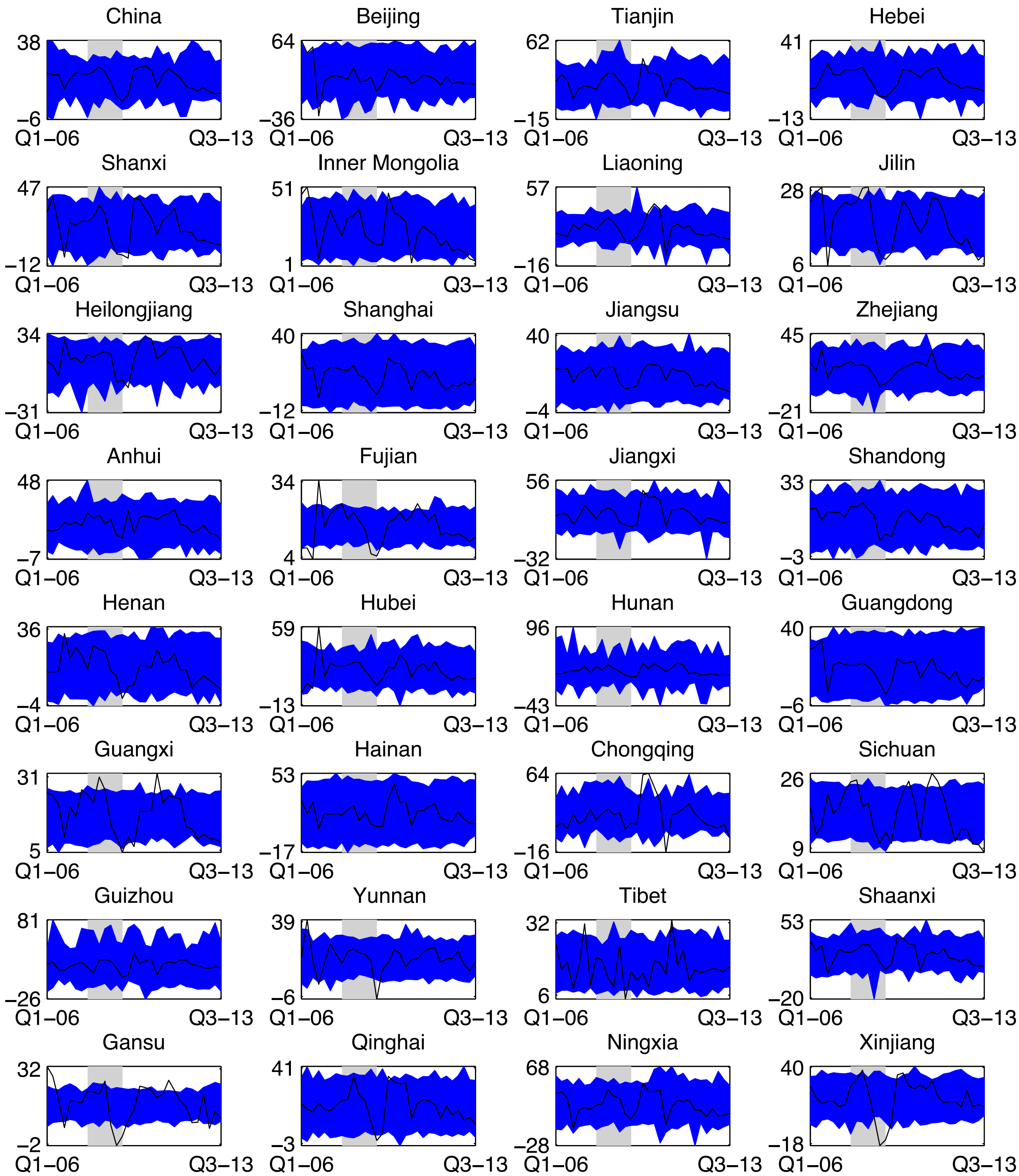
	N	E	C
β	-1.04	-0.75	13.27
	(-8.43,12.43)	(-4.88,9.98)	(-37.99,54.77)
σ	23.11*	13.33*	21.74
	(3.22,33.88)	(2.85,16.66)	(-5.24,68.59)
ρ_y	0.69	0.15	0.71
	(-16.55,15.05)	(-17.77,19.29)	(-2.88,2.32)
σ_y	11.48*	12.94*	1.04
	(3.48,36.41)	(5.92,9.81)	(-6.19,11.40)
ρ_u	-0.03	0.07	0.46
	(-9.55,9.26)	(-5.47,4.44)	(-0.82,2.02)
σ_u	8.26*	6.95*	9.15*
	(1.67,20.39)	(0.61,20.81)	(3.70,17.42)

Pooled ($N = 31$).

	N	E	C
β	-0.12	0.92	-5.59
	(-4.77,7.04)	(-6.64,12.97)	(-156.23,116.90)
σ	23.61*	14.09*	40.29
	(0.06,41.51)	(1.61,26.66)	(-11.30,73.95)
ρ_y	0.48	0.94	0.80
	(-11.29,9.67)	(-13.19,16.54)	(-18.88,23.23)
σ_y	28.08*	5.66	2.85
	(17.61,40.33)	(-2.11,12.27)	(-5.01,7.22)
ρ_u	-0.29	0.30	0.24
	(-3.08,2.90)	(-5.15,5.62)	(-2.75,3.80)
σ_u	10.60	6.87*	8.89*
	(-1.11,27.62)	(1.25,14.97)	(3.89,18.65)

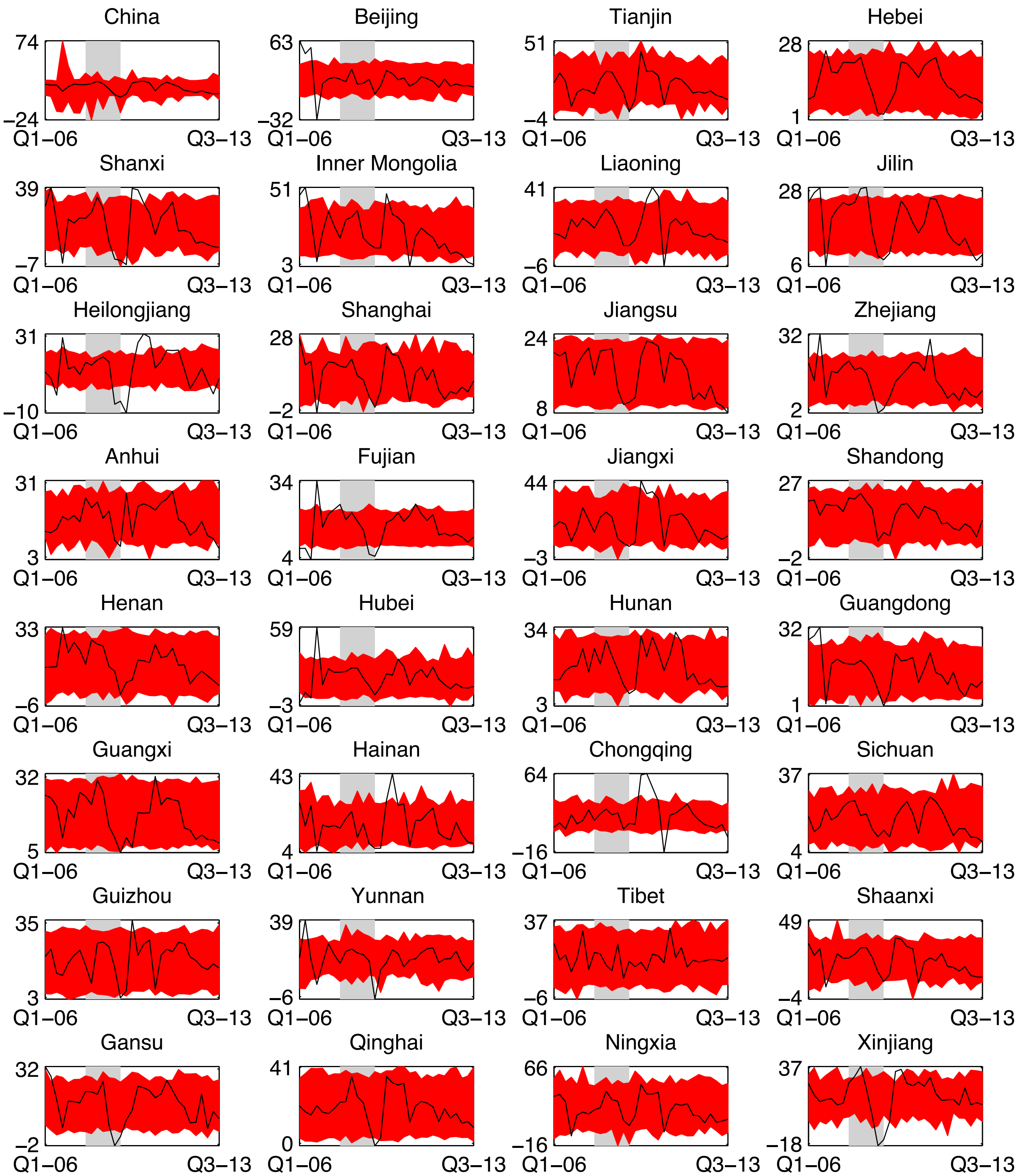
Notes: $T = 31$. * Significiant at 95% confidence level (CI). N: NO_2 columns. E: Electricity generation. C: Cement production. See Figure 8 (b) for quarterly Group definitions.

Figure G.7: Quarterly NO_2 95% confidence bands.



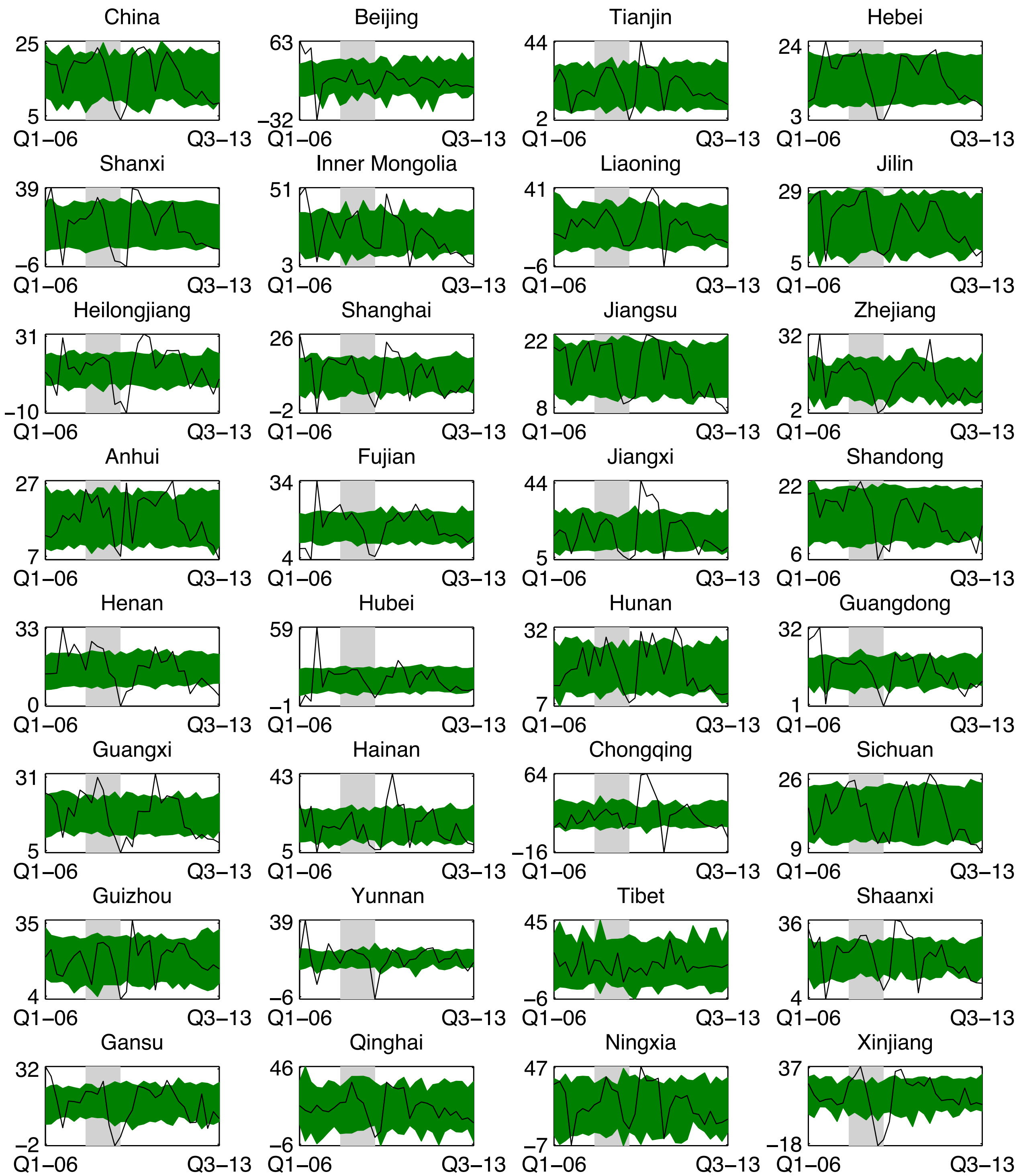
Notes: 2006Q1-2013Q3. Black line is reported annualized percentage change in GDP.

Figure G.8: 95% Confidence bands: Quarterly sample, electricity generation.



Notes: 2006Q1-2013Q3. Black line is reported annualized percentage change in GDP.

Figure G.9: 95% Confidence bands: Quarterly sample, cement production.



Notes: 2006Q1-2013Q3. Black line is reported annualized percentage change in GDP.

H Time decay in signal-output relationship

Above in Appendix C, we derived the representation of the model with a constant elasticity of the signal with respect to output. Alternatively, say the relationship between signal and output is now subject to a time decay as a result of technical progress, Equation (26) in the main text. This relationship replaces main paper Equation (3), while the other relationships in Equations (2)-(6) which encapsulate the model remain the same. Then dropping (i) subscripts for clarity, the ABCD representation (C.1) above for the model becomes,

$$\begin{aligned}
\begin{bmatrix} y^*_{nt} \\ u^*_{nt} \\ t \end{bmatrix} &= \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{f} + \underbrace{\begin{bmatrix} \rho_y & 0 & 0 \\ 0 & \rho_u & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} y^*_{nt-1} \\ u^*_{nt-1} \\ t-1 \end{bmatrix} + \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{B} \varepsilon_{nt} \\
\underbrace{\begin{bmatrix} y_{nt} \\ s_{nt} \\ t \end{bmatrix}}_{Y_{nt}} &= \underbrace{\begin{bmatrix} 0 \\ b(\bar{t}-1) \\ 1 \end{bmatrix}}_{g} + \underbrace{\begin{bmatrix} \rho_y & \rho_u & 0 \\ \beta\rho_y & 0 & -b \\ 0 & 0 & 1 \end{bmatrix}}_{C} \begin{bmatrix} y^*_{nt-1} \\ u^*_{nt-1} \\ t-1 \end{bmatrix} + \underbrace{\begin{bmatrix} 1 & 1 & 0 \\ \beta & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} \varepsilon^y_{nt} \\ \varepsilon^u_{nt} \\ \varepsilon^s_{nt} \end{bmatrix}}_{\varepsilon_{nt}}
\end{aligned} \tag{H.1}$$

The new variable in this model is t, a time trend. So, following steps exactly similar to Appendix C, the observation equation implies the states are $\begin{bmatrix} y^*_{nt-1} & u^*_{nt-1} & t-1 \end{bmatrix}' = -C^{-1}g + C^{-1}Y_{nt} - C^{-1}D\varepsilon_{nt}$. Plugging this into the state equation and rearranging yields,

$$\underbrace{\begin{bmatrix} y_{nt} \\ s_{nt} \\ t \end{bmatrix}}_{Y_{nt}} = \underbrace{\begin{bmatrix} -b\bar{t}\psi \\ -b[1-(1-\rho_y)\bar{t}] \\ 1 \end{bmatrix}}_{C[f+(I_3-A)C^{-1}g]} + \underbrace{\begin{bmatrix} \rho_u & \psi & b\psi \\ 0 & \rho_y & -b(1-\rho_y) \\ 0 & 0 & 1 \end{bmatrix}}_{CAC^{-1}} \underbrace{\begin{bmatrix} y_{nt-1} \\ s_{nt-1} \\ t-1 \end{bmatrix}}_{Y_{nt-1}} + \begin{bmatrix} v^y_{nt} \\ v^s_{nt} \\ 0 \end{bmatrix} \tag{H.2}$$

for the MA(1) error and parameters,

$$\begin{bmatrix} v_{nt}^y \\ v_{nt}^s \\ 0 \end{bmatrix} = D\varepsilon_{nt} + \underbrace{\begin{bmatrix} 0 & 0 & -\psi \\ 0 & 0 & -\rho_y \\ 0 & 0 & 0 \end{bmatrix}}_{C(B-AC^{-1}D)} \varepsilon_{nt-1} = \underbrace{\begin{bmatrix} u_{nt}^y \\ u_{nt}^s \\ 0 \end{bmatrix}}_{D\varepsilon_{nt}} + \begin{bmatrix} 0 & -\psi m & 0 \\ 0 & -\rho_y m & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_{nt-1}^y \\ u_{nt-1}^s \\ 0 \end{bmatrix} \tag{H.3}$$

$$\psi = \frac{\rho_y - \rho_u}{\beta} \qquad m = \frac{\sigma}{\sqrt{\beta^2\sigma_y^2 + \sigma^2}}$$

with the iid errors $u_{nt}^y = \varepsilon_{nt}^y + \varepsilon_{nt}^u \sim N(0, \sigma_y^2 + \sigma_u^2)$ and $u_{nt}^s = \beta\varepsilon_{nt}^y + \varepsilon_{nt}^s \sim N(0, \beta^2\sigma_y^2 + \sigma^2)$. Note, there is no need to account for a nuisance parameter λ since we are only dealing with one signal in this instance and so $\lambda = 1$ always. First-differencing yields,

$$\begin{bmatrix} \Delta y_{nt} \\ \Delta s_{nt} \end{bmatrix} = \begin{bmatrix} b\psi \\ -b(1-\rho_y) \end{bmatrix} + \begin{bmatrix} \rho_u & \psi \\ 0 & \rho_y \end{bmatrix} \begin{bmatrix} \Delta y_{nt-1} \\ \Delta s_{nt-1} \end{bmatrix} + \begin{bmatrix} \Delta v_{nt}^y \\ \Delta v_{nt}^s \end{bmatrix} \tag{H.4}$$

for $\Delta v_{nt}^y = u_{nt}^y - u_{nt-1}^y - \psi m u_{nt-1}^s + \psi m u_{nt-2}^s$ and $\Delta v_{nt}^s = u_{nt}^s - u_{nt-1}^s - \rho_y m u_{nt-1}^s + \rho_y m u_{nt-2}^s$ each MA(2) errors. The two rows of (H.4) correspond to (27) and (28) in the main text. To test the null hypothesis $H_0 : b \equiv b_0/b_2 = 0$ using the first row, called "Test 1" in the text, we may make use of the Wald statistic

$$W = (T-4)N(\widehat{b}_0/\widehat{b}_2)^2 \widehat{A}\widehat{\text{Avar}} \left(\begin{bmatrix} \widehat{b}_0 & \widehat{b}_1 & \widehat{b}_2 \end{bmatrix}' \right)^{-1} \widehat{A}' \xrightarrow{d} \chi^2(1) \tag{H.5}$$

$$\widehat{A} = \begin{bmatrix} 1/\widehat{b}_2 & 0 & -\widehat{b}_0/\widehat{b}_2^2 \end{bmatrix}$$

$$\begin{bmatrix}\widehat{b}_0 & \widehat{b}_1 & \widehat{b}_2\end{bmatrix}' = \left(\sum_{n=1}^{N}\sum_{t=5}^{T} x_{nt-3}x'_{nt-1}\right)^{-1} \sum_{n=1}^{N}\sum_{t=5}^{T} x_{nt-3}\Delta y_{nt}$$

$$x'_{nt} = \begin{bmatrix}1 & \Delta y_{nt} & \Delta s_{nt}\end{bmatrix}$$

$$\text{Avar}\left(\begin{bmatrix}\widehat{b}_0 & \widehat{b}_1 & \widehat{b}_2\end{bmatrix}'\right) = E\left[\left(x_{nt-3}\Delta v^y_{nt}\right)\left(x_{nt-3}\Delta v^y_{nt}\right)'\right]$$

$$\widehat{\text{Avar}}\left(\begin{bmatrix}\widehat{b}_0 & \widehat{b}_1 & \widehat{b}_2\end{bmatrix}'\right) = \widehat{\Sigma}_x(0)\widehat{\sigma}_v(0) + 2\widehat{\Sigma}_x(1)\widehat{\sigma}^2_v(1) + 2\widehat{\Sigma}_x(2)\widehat{\sigma}^2_v(2)$$

$$\widehat{\Sigma}_x(j) = (N(T-j))^{-1}\sum_{n=1}^{N}\sum_{t=1+j}^{T} x_{nt}x'_{nt-j}$$

$$\widehat{\sigma}^2_v(j) = (N(T-j))^{-1}\sum_{n=1}^{N}\sum_{t=1+j}^{T} \widehat{v}^y_{nt}\widehat{v}^y_{nt-j}$$

$$\widehat{v}^y_{nt} = \Delta y_{nt} - \begin{bmatrix}\widehat{b}_0 & \widehat{b}_1 & \widehat{b}_2\end{bmatrix} x_{nt-1}$$

To test the null hypothesis $H_0 : b \equiv b_0/(b_1 - 1) = 0$ using the second row, called "Test 2" in the text, we may make use of the Wald statistic,

$$W = (T-4)N(\widehat{b}_0/(\widehat{b}_1 - 1))^2 \widehat{A}\widehat{\text{Avar}}\left(\begin{bmatrix}\widehat{b}_0 & \widehat{b}_1\end{bmatrix}'\right)^{-1} \widehat{A}' \xrightarrow{d} \chi^2(1) \tag{H.6}$$

$$\widehat{A} = \begin{bmatrix}1/(\widehat{b}_1 - 1) & -\widehat{b}_0/(\widehat{b}_1 - 1)^2\end{bmatrix}$$

$$\begin{bmatrix}\widehat{b}_0 & \widehat{b}_1\end{bmatrix}' = \left(\sum_{n=1}^{N}\sum_{t=5}^{T} x_{nt-3}x'_{nt-1}\right)^{-1} \sum_{n=1}^{N}\sum_{t=5}^{T} x_{nt-3}\Delta s_{nt}$$

$$x'_{nt} = \begin{bmatrix}1 & \Delta s_{nt}\end{bmatrix}$$

$$\text{Avar}\left(\begin{bmatrix}\widehat{b}_0 & \widehat{b}_1\end{bmatrix}'\right) = E\left[\left(x_{nt-3}\Delta v_{nt}^s\right)\left(x_{nt-3}\Delta v_{nt}^s\right)'\right]$$

$$\widehat{\text{Avar}}\left(\begin{bmatrix}\widehat{b}_0 & \widehat{b}_1\end{bmatrix}'\right) = \widehat{\Sigma}_x(0)\widehat{\sigma}_v^2(0) + 2\widehat{\Sigma}_x(1)\widehat{\sigma}_v^2(1) + 2\widehat{\Sigma}_x(2)\widehat{\sigma}_v^2(2)$$

and $\widehat{\Sigma}_x$ and $\widehat{\sigma}_v^2$ defined similar to with respect to Test 1, but using $\widehat{v}_{nt}^s$.

References

Abadir, K. M. and J. R. Magnus (2005). *Matrix algebra.* Cambridge University Press.

Alvarez, J. and M. Arellano (2003). The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica*, 1121–1159.

Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics 68*(1), 29–51.

Fernández-Villaverde, J., J. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2007). ABCs (and Ds) of Understanding VARs. *American Economic Review 97*(3), 1021–1026.

Hansen, B. E. (2015). *Econometrics.* Unpublished manuscript.

Hayashi, F. (2000). *Econometrics.* Princeton University Press.

Lutkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis.* Springer.

Morris, S. D. (2016). VARMA representation of DSGE models. *Economics Letters 138.*

Rothenberg, T. (1971). Identification in parametric models. *Econometrica 39*(3), 577–91.