Appendix A. The structural PVAR model

The structural PVAR model analyzing the dynamics among structure (*S*), pollution (*P*), income (*E*) and health (*H*) is

$$\boldsymbol{A}\begin{pmatrix} S_{it} \\ P_{it} \\ E_{it} \\ H_{it} \end{pmatrix} = \sum_{l=1}^{k} \begin{pmatrix} \phi_{l,11} & \phi_{l,12} & \phi_{l,13} & \phi_{l,14} \\ \phi_{l,21} & \phi_{l,22} & \phi_{l,23} & \phi_{l,24} \\ \phi_{l,31} & \phi_{l,32} & \phi_{l,33} & \phi_{l,34} \\ \phi_{l,41} & \phi_{l,42} & \phi_{l,43} & \phi_{l,44} \end{pmatrix} \begin{pmatrix} S_{it-l} \\ P_{it-l} \\ E_{it-l} \\ H_{it-l} \end{pmatrix} + \begin{pmatrix} \alpha_i^S \\ \alpha_i^P \\ \alpha_i^E \\ \alpha_i^H \end{pmatrix} + \begin{pmatrix} u_{it}^S \\ u_{it}^P \\ u_{it}^E \\ u_{it}^H \end{pmatrix} \quad \text{(A.1)}$$

Define $\boldsymbol{y_{it}} = (S_{it}, P_{it}, E_{it}, H_{it})'$, using matrix, Eq. (A.1) can be rewritten as

$$\boldsymbol{A}\boldsymbol{y_{it}} = \sum_{l=1}^{k} \boldsymbol{\Phi_l}\boldsymbol{y_{it-l}} + \boldsymbol{\alpha_i} + \boldsymbol{u_{it}} \quad \text{(A.2)}$$

The 4×4 matrix $\boldsymbol{A} = (a_{ij})$ reflects the contemporaneous relation of the variables in $\boldsymbol{y_{it}}$, and $\boldsymbol{\Phi_l} = (\phi_{l,ij})$ are the coefficients of the lagged endogenous variables. Because both contemporaneous and long-run relationships are accounted for, we assume that the structural shock $\boldsymbol{u_{it}}$ is homoskedastic, that is

$$\boldsymbol{u_{it}} \sim (\boldsymbol{0}, \boldsymbol{\Sigma}), \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma_S^2} & & & \\ & \boldsymbol{\sigma_P^2} & & \\ & & \boldsymbol{\sigma_E^2} & \\ & & & \boldsymbol{\sigma_H^2} \end{pmatrix}$$

To estimate eq. (A.2), transform Eq. (A.1) into the reduced form

$$\boldsymbol{y_{it}} = \sum_{l=1}^{k} \Pi_l \boldsymbol{y_{it-l}} + \boldsymbol{\mu_i} + \boldsymbol{\varepsilon_{it}},\ \boldsymbol{\varepsilon_{it}} \sim (\boldsymbol{0}, \boldsymbol{\Omega}) \quad \text{(A.3)}$$

where

$$\Pi_l = \boldsymbol{A^{-1}\Phi_l}, \boldsymbol{\mu_i} = \boldsymbol{A^{-1}\alpha_i}, \boldsymbol{\varepsilon_{it}} = \boldsymbol{A^{-1}u_{it}}.$$

Since matrix A is of full rank, the covariance matrix Ω is no longer diagonal. In fact, we have

$$\boldsymbol{\Omega} = \boldsymbol{A^{-1}\Sigma(A^{-1})'} \quad \text{(A.4)}$$

To recover the structural parameters in $\boldsymbol{A}$ and $\boldsymbol{\Sigma}$, we must impose additional restrictions a priori. Note that the left-hand side of Eq. (A.4) contains 10 coefficients, because $\boldsymbol{\Omega}$ is symmetric. But on the right-hand side of (A.2), the number of unknowns in $\boldsymbol{A}$ and $\boldsymbol{\Sigma}$ are 16 and 4, respectively. Therefore, to pin down $\boldsymbol{A}$ and $\boldsymbol{\Sigma}$, we need 16 + 4 − 10 = 10 more restrictions.

Applying Cholesky decomposition, we impose a lower triangular structure on $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

To finally identify $\boldsymbol{A}$ and $\boldsymbol{\Sigma}$, we should either set $a_{11} = a_{22} = a_{33} = a_{44} = 1$, or $\sigma_S = \sigma_P = \sigma_E = \sigma_H = 1$. Based on these restrictions, the PVAR model is transformed into a recursive system.

To see the impacts of shocks, we transform Eq. (A.1) to

$$\boldsymbol{A}\left(\boldsymbol{I} - \sum_{l=1}^{k} \boldsymbol{\Pi_l L^l}\right)\boldsymbol{y_{it}} = \boldsymbol{\alpha_i} + \boldsymbol{u_{it}} \quad \text{(A.5)}$$

where L is the lag operator. Let $\boldsymbol{\Pi(L)} = \boldsymbol{I} - \sum_{\boldsymbol{l=1}}^{\boldsymbol{k}} \boldsymbol{\Pi_l L^l}$, then when $k = 1$ just like in our empirical study

$$\boldsymbol{\Pi(L)} = \boldsymbol{I} - \boldsymbol{\Pi_1 L}$$

Multiply both sides of Eq. (A.5) by $\boldsymbol{\Pi(L)^{-1} A^{-1}}$, we obtain a VMA model

$$\boldsymbol{y_{it}} = \boldsymbol{\beta_i} + \sum_{p=0}^{\infty} \Theta_p \, \boldsymbol{u_{it-p}} \tag{A.6}$$

where $\boldsymbol{\beta_i} = \boldsymbol{\Pi(L)^{-1} A^{-1} \alpha_i}$ and $\Theta_p = \boldsymbol{\Pi}_1^{\mathrm{p}} \boldsymbol{A^{-1}}$.

Appendix B. The spatial distribution of the cities in our sample

Table A. 1 List of Sample Cities

Group	Cities
1.East	Beijing, Tianjin, Shanghai, Nanjing, Wuxi, Xuzhou, Changzhou, Suzhou, Nantong, Lianyungang, Yangzhou, Zhenjiang, Hangzhou, Ningbo, Wenzhou, Jiaxing, Huzhou, Shaoxing, Taizhou, Fuzhou, Xiamen, Quanzhou, Guangzhou, Shaoguan, Shenzhen, Zhuhai, Shantou, Foshan, Zhanjiang, Zhong-shan
2.Central	Shijiazhuang, Tangshan, Qinhuangdao, Handan, Baoding, Taiyuan, Datong, Changzhi, Linfen, Huhehaote, Baotou, Chifeng, Shenyang, Dalian, Anshan, Fushun, Benxi, Jinzhou, Changchun, Jilin, Harbin, Qiqihar, Daqing, Mudanjiang, Hefei, Wuhu, Ma'anshan, Nanchang, Jiujiang, Jinan, Qingdao, Zibo, Zaozhuang, Yantai, Weifang, Jining, Tai'an, Weihai, Rizhao, Zhengzhou, Kaifeng, Luoyang, Pingdingshan, Anyang, Jiaozuo, Sanmenxia, Wuhan, Yichang, Jingzhou, Changsha, Zhuzhou, Xian-gtan, Yueyang, Changde
3.West	Nanning, Liuzhou, Guilin, Chongqing, Chengdu, Zigong, Luzhou, Deyang, Mianyang, Yibin, Guiyang, Zunyi, Kunming, Qujing, Yuxi, Xi'an, Baoji, Xianyang, Weinan, Yan'an, Lanzhou, Xining, Yinchuan, Urumqi

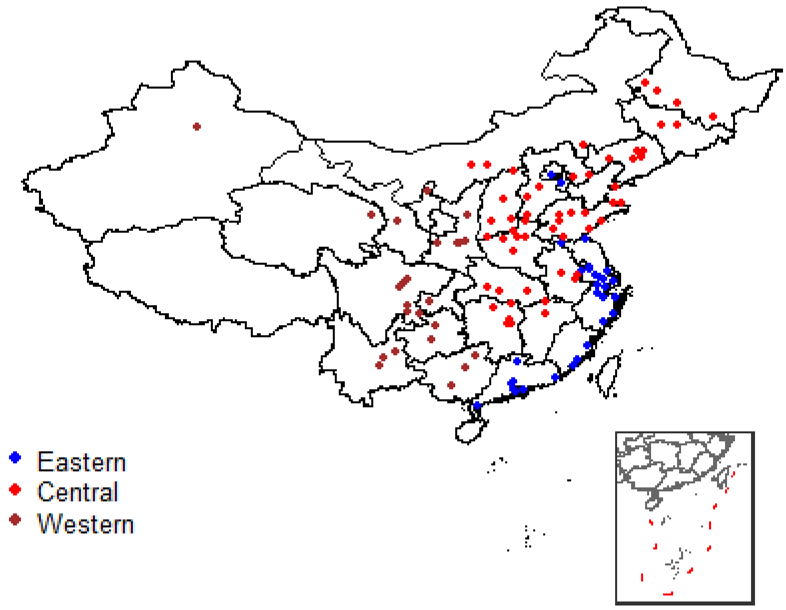

Fig A.1 The Distribution of the Cities

Notes: Our sample consists of 108 cities, of which 30 are in the east, 54 in the central China, and 24 in the west.