Supporting Information

The effect of economic growth, urbanization, and industrialization on fine particulate matter ($PM_{2.5}$) concentrations in China

Guangdong Li[1], Chuanglin Fang[1*], Shaojian Wang[2*], Siao Sun[1]

1. Key Laboratory of Regional Sustainable Development Modeling, Institute of Geographic Sciences and Natural Resources Research (IGSNRR), Chinese Academy of Sciences (CAS), 11A Datun Road, Chaoyang District, Beijing 100101, China

2. Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou, PR China

*Corresponding author: Chuanglin Fang, Shaojian Wang

E-mail: fangcl@igsnrr.ac.cn; wangshj8@mail.sysu.edu.cn.

TEL: +86-01064889101

Number of Pages: 31

Number of Figures: 11

Number of Text: 1

Number of Tables:11

Figure S1. Prefecture-level divisions in China for period of 1999-2011. (This figure was created by G. Li in ArcGIS10.2, http://desktop.arcgis.com/en/. The colors represent different regions in China. Prefecture-level division was the basic regional unit in China at the prefectural level. The administrative boundaries for the provinces, cities were obtained as GIS flies from the National Geomatics Center of China.)

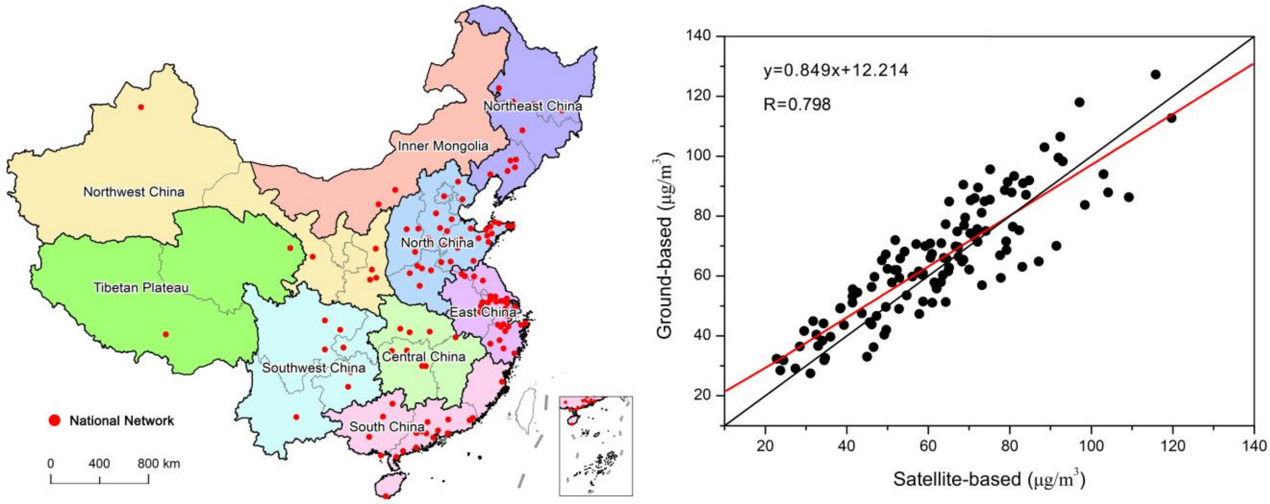

Figure S2. The validation result of satellite-derived estimates and ground-based measurements in China for $PM_{2.5}$. The map was created by G. Li in ArcGIS10.2, http://desktop.arcgis.com/en/. The scatter plot of satellite-derived estimates and ground-based measurements in China for $PM_{2.5}$ made with Microsoft Excel 2016. The satellite-derived $PM_{2.5}$ data drawn from the Atmospheric Composition Analysis Group (http://fizz.phys.dal.ca/~atmos/martin/?page_id=140). The ground-based $PM_{2.5}$ data gathered from the related literatures in China.

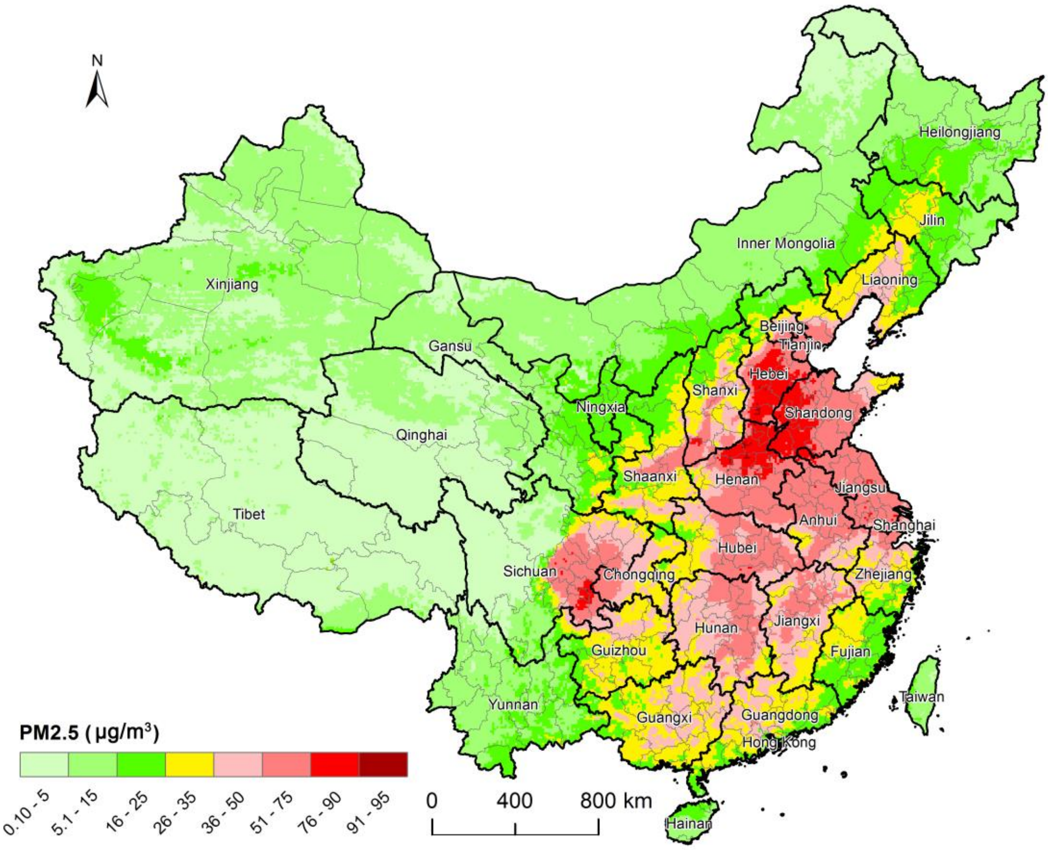

Figure S3. Spatial distribution of mean $PM_{2.5}$ concentrations in China from 1999 to 2011. (This figure was created by G. Li in ArcGIS10.2, http://desktop.arcgis.com/en/. The administrative boundaries for the provinces, cities were obtained as GIS flies from the National Geomatics Center of China.)

Table S1 Descriptive statistics for five panels used in this study for the period 1999-2011.

Panel	Variable	Obs.	Mean	Std. Dev.	Min	Max
C	$PM_{2.5}$ (μg/m^3)	4,381	34.26	21.91	0.38	98.75
	GDPPC (Yuan, RMB)	4,381	17448.04	17183.72	1179.79	161267.00
	IND (%)	4,381	44.95	12.80	8.60	90.09
	URBAN (%)	4,381	33.96	17.60	6.91	100.00
I1	$PM_{2.5}$ (μg/m^3)	3,198	33.36	22.08	0.38	98.75
	GDPPC (Yuan, RMB)	3,198	12858.23	11583.26	1179.79	130870.00
	IND (%)	3,198	42.18	12.42	8.60	82.02
	URBAN (%)	3,198	28.13	12.27	6.91	85.96
I2	$PM_{2.5}$ (μg/m^3)	728	38.56	20.39	7.09	89.99
	GDPPC (Yuan, RMB)	728	29739.09	23004.05	4495.82	161267.00
	IND (%)	728	55.88	8.73	29.89	90.09
	URBAN (%)	728	44.81	19.44	11.12	100.00
I3	$PM_{2.5}$ (μg/m^3)	455	33.68	22.25	3.26	95.01
	GDPPC (Yuan, RMB)	455	32042.17	22514.97	5329.98	160253.30
	IND (%)	455	46.92	11.08	15.45	81.79
	URBAN (%)	455	57.62	18.29	21.26	100.00
H	$PM_{2.5}$ (μg/m^3)	455	73.62	16.84	70.01	95.39
	GDPPC (Yuan, RMB)	455	21461.36	17512.43	2369.20	130870.00
	IND (%)	455	51.79	8.84	26.37	81.23
	URBAN (%)	455	32.45	11.93	14.07	65.18

Note: C is the total panel including all the prefecture-level divisions; I1 is agriculture-oriented panel including 246 prefecture-level divisions; I2 is industry-oriented panel including 56 prefecture-level divisions; I3 is service -oriented panel including 35 prefecture-level divisions; H is the heavily $PM_{2.5}$-polluted panel including 35 prefecture-level divisions.

Table S2 The mean changes of $PM_{2.5}$ concentrations, per capita GDP, urbanization and industrialization for five panels in 1999 and 2011.

Panel	Year	$PM_{2.5}$ ($\mu g/m^3$)	GDPPC (Yuan, RMB)	IND (%)	URBAN (%)
C	1999	25.23	7607.71	40.79	29.38
	2011	37.51	36343.93	49.71	38.75
I1	1999	24.78	5316.26	37.25	23.93
	2011	36.68	27852.18	48.33	32.78
I2	1999	27.51	13723.16	53.55	39.19
	2011	41.77	58202.37	56.93	50.16
I3	1999	24.78	13928.63	45.29	52.01
	2011	36.48	61055.32	47.87	62.48
H	1999	49.42	8557.32	47.30	27.46
	2011	85.97	44022.92	54.50	37.13
C		48.66	377.72	21.87	31.91
I1		48.05	423.91	29.76	37.00
I2	Change (%)	51.85	324.12	6.32	27.99
I3		47.21	338.34	5.70	20.14
H		73.96	414.45	15.23	35.21

Table S3 The list of prefecture-level divisions for four sub-panels.

I1 panel (246)				**I2 panel (56)**	**I3 panel (35)**	**H panel (35)**
Qinhuangdao	Xinyu	Xiangxi	Puer	Shijiazhuang	Beijing	Tianjin
Xingtai	Yingtan	Guangzhou	Lincang	Tangshan	Tianjin	Shijiazhuang
Zhangjiakou	Ganzhou	Zhanjiang	Chuxiong	Handan	Taiyuan	Tangshan
Chengde	Ji'an	Maoming	Honghe	Baoding	Huhehaote	Handan
Cangzhou	Yichun	Zhaoqing	Wenshan	Langfang	Alxa League	Xingtai
Hengshui	Fuzhou	Meizhou	Xishuangbanna	Datong	Shenyang	Baoding
Changzhi	Shangrao	Shanwei	Dali	Yangquan	Dalian	Cangzhou
Jincheng	Jinan	Heyuan	Dehong	Baotou	Yanbian	Langfang
Shuozhou	Zibo	Yangjiang	Nujiang	Wuhai	Yichun	Hengshui
Jinzhong	Zaozhuang	Qingyuan	Diqing	Eerduosi	Great Khingan	Tai'an
Yuncheng	Dongying	Jieyang	Lhasa	Anshan	Shanghai	Hefei
Xinzhou	Yantai	Yunfu	Changdu	Fushun	Nanjing	Jinan
Linfen	Weifang	Nanning	Shannan	Benxi	Huaian	Zibo
Luliang	Jining	Liuzhou	Rikaze	Liaoyang	Yancheng	Zaozhuang
Chifeng	Tai'an	Guilin	Naqu	Panjin	Zhoushan	Dongying
Tongliao	Weihai	Wuzhou	Ali	Baishan	Hefei	Jining
Hulunbuer	Rizhao	Beihai	Linzhi	Jixi	Tongling	Laiwu
Bayannaoer	Laiwu	Fangchenggang	Tongchuan	Hegang	Fuzhou	Linyi
Wulanchabu	Linyi	Qinzhou	Baoji	Shuangyashan	Nanchang	Dezhou
Hinggan League	Dezhou	Guigang	Xianyang	Daqing	Qingdao	Liaocheng
Xilin Gol League	Liaocheng	Yulin	Weinan	Wuxi	Zhengzhou	Binzhou
Dandong	Binzhou	Baise	Yan'an	Xuzhou	Wuhan	Heze
Jinzhou	Heze	Hezhou	Hanzhong	Changzhou	Changsha	Zhengzhou
Yingkou	Kaifeng	Hechi	Yulin	Suzhou	Shenzhen	Kaifeng
Fuxin	Luoyang	Laibin	Ankang	Nantong	Hiakou	Luoyang
Tieling	Anyang	Chongzuo	Shangluo	Yangzhou	Sanya	Pingdingshan
Chaoyang	Hebi	Hainanzhixiaxian	Jinchang	Zhenjiang	Chengdu	Anyang
Huludao	Xinxiang	Chongqing	Baiyin	Taizhou	Guiyang	Jiaozuo
Changchun	Jiaozuo	Zigong	Tianshui	Hangzhou	Xi'an	Wuhan
Jilin	Puyang	Panzhihua	Wuwei	Ningbo	Lanzhou	Xuzhou
Siping	Xuchang	Luzhou	Zhangye	Wenzhou	Jiayuguan	Jingzhou
Liaoyuan	Luohe	Deyang	Pingliang	Jiaxing	Xining	Xi'an
Tonghua	Sanmenxia	Mianyang	Jiuquan	Huzhou	Yinchuan	Weinan
Songyuan	Nanyang	Guangyuan	Qingyang	Shaoxing	Shizuishan	Zigong
Baicheng	Shangqiu	Suining	Dingxi	Jinhua	Wulumuqi	Ma'anshan
Haerbin	Xinyang	Neijiang	Longnan	Taizhou		
Qiqiha'er	Zhoukou	Leshan	Linxia	Wuhu		
Jiamusi	Zhumadian	Nanchong	Gannan	Huainan		

Qitaihe	Huangshi	Meishan	Haidong	Ma'anshan
Mudanjiang	Shiyan	Yibin	Haibei	Huaibei
Heihe	Yichang	Guang'an	Huangnan	Xiamen
Suihua	Xiangyang	Dazhou	Hainan	Putian
Lianyungang	Ezhou	Ya'an	Guoluo	Quanzhou
Suqian	Jingmen	Bazhong	Yushu	Pingxiang
Quzhou	Xiaogan	Ziyang	Haixi	Jiujiang
Lishui	Jingzhou	Aba	Wuzhong	Pingdingshan
Bangbu	Huanggang	Ganzi	Guyuan	Shaoguan
Anqing	Xianning	Liangshan	Zhongwei	Zhuhai
Huangshan	Suizhou	Lupanshui	Turpan	Shantou
Chuzhou	Enshi	Zunyi	Hami(Kumul)	Foshan
Fuyang	Zhuzhou	Anshun	Changji	Jiangmen
Suzhou	Xiangtan	Tongren	Boertala	Huizhou
Lu'an	Hengyang	Qinxinan	Bayinguole	Dongwan
Bozhou	Shaoyang	Bijie	Aksu	Zhongshan
Chizhou	Yueyang	Qiandongnan	Kizilsu	Chaozhou
Xuanzhou	Changde	Qiannan	Kashi	Karamay
Sanming	Zhangjiajie	Kunming	Hetian	
Zhangzhou	Yiyang	Qujing	Kazak	
Nanping	Chenzhou	Yuxi	Tacheng	
Longyan	Yongzhou	Baoshan	Altay	
Ningde	Huaihua	Zhaotong		
Jingdezhen	Loudi	Lijiang		

Note: I1 is agriculture-oriented panel including 246 prefecture-level divisions; I2 is industry-oriented panel including 56 prefecture-level divisions; I3 is service -oriented panel including 35 prefecture-level divisions; H is the heavily $PM_{2.5}$-polluted panel including 35 prefecture-level divisions. The proportion of the primary, second and third industry in GDP was used to identify categories of prefecture-level divisions. The annual mean $PM_{2.5}$ concentrations >70 μg/m^3 (two times with the standard values of the IT-1 level of WHO) was utilized to group the heavily $PM_{2.5}$-polluted prefecture-level divisions.

Table S4 A list of baseline characteristics for the major prefecture-level divisions in China (Mean value in the period of 1999-2011 for provincial capitals and serious $PM_{2.5}$ polluted cities).

Code	Name	Annual mean $PM_{2.5}$ ($\mu g/m^3$)	Per capita GDP (Yuan)	Urbanization (%)	Industrialization (%)
1	Beijing	44.75	46732.93	74.05	30.34
2	Tianjin	70.16	41373.22	59.99	53.00
3	Shijiazhuang	75.67	21429.07	35.77	48.23
4	Tangshan	72.68	32833.51	31.29	55.46
5	Handan	73.14	15177.25	25.72	51.21
6	Xingtai	73.51	10977.20	20.63	55.05
7	Baoding	70.04	11663.30	23.01	48.13
8	Cangzhou	72.80	17313.19	20.95	50.80
9	Langfang	71.14	19266.36	22.67	53.98
10	Hengshui	78.95	12301.79	21.22	52.16
11	Taiyuan	34.94	27465.40	69.99	48.61
12	Yangquan	35.54	17699.27	49.62	59.07
13	Changzhi	37.91	14814.28	27.37	59.65
14	Huhehaote	17.00	33855.21	46.10	39.80
15	Shenyang	35.60	36815.92	64.68	46.96
16	Dalian	31.25	45381.19	55.63	48.50
17	Anshan	29.91	34336.21	52.30	55.29
18	Changchun	23.46	25870.51	43.89	47.75
19	Haerbin	18.30	21864.16	48.34	36.11
20	Mudanjiang	12.51	15034.79	55.89	39.28
21	Shanghai	50.81	57199.04	82.30	46.31
22	Nanjing	64.09	39534.81	70.58	48.14
23	Xuzhou	71.53	17655.75	32.07	49.39
24	Ningbo	32.06	47159.63	32.51	55.41
25	Hefei	72.33	23134.73	39.25	49.91
26	Wuhu	56.83	23260.87	43.47	57.56
27	Ma'anshan	72.05	31135.64	46.46	64.63
28	Fuzhou	19.68	27678.54	35.34	46.88
29	Xiamen	24.18	49063.15	61.97	53.36
30	Nanchang	51.61	25276.77	45.22	52.23
31	Jinan	71.35	34700.50	52.41	43.82
32	Qingdao	46.79	37753.99	53.81	50.37
33	Zibo	70.21	36654.66	45.56	61.79
34	Zaozhuang	71.79	20054.94	33.97	57.56
35	Dongying	71.91	67501.02	47.62	77.68
36	Weifang	54.77	19742.11	30.94	52.70
37	Jining	74.18	18152.87	29.81	50.43
38	Tai'an	73.64	19463.93	33.41	51.37

39	Laiwu	74.98	23553.63	37.54	59.05
40	Linyi	70.73	13788.39	24.70	49.68
41	Dezhou	76.72	17312.65	25.57	51.09
42	Liaocheng	77.21	15141.84	24.45	52.05
43	Binzhou	71.24	22077.00	25.08	54.72
44	Heze	76.56	7242.84	18.75	39.86
45	Zhengzhou	78.73	28352.77	40.37	52.54
46	Kaifeng	79.59	10693.12	21.12	39.81
47	Luoyang	72.12	19755.20	26.88	57.71
48	Pingdingshan	72.80	14228.62	25.97	58.89
49	Anyang	76.11	13139.85	21.19	55.12
50	Jiaozuo	77.98	19659.77	32.18	59.86
51	Sanmenxia	40.70	19955.33	29.12	59.99
52	Wuhan	71.36	32973.33	62.14	45.49
53	Yichang	36.75	20717.00	32.54	54.67
54	Jingzhou	70.98	8348.54	27.04	37.92
55	Changsha	46.63	32106.93	35.00	45.62
56	Zhuzhou	44.22	18184.70	27.59	51.54
57	Xiangtan	51.91	17621.67	29.45	46.83
58	Changde	49.89	14499.03	22.88	41.90
59	Zhangjiajie	35.99	9095.74	17.49	24.10
60	Shenzhen	22.38	71913.33	97.27	51.95
61	Guangzhou	31.12	60425.85	84.58	40.50
62	Nanning	30.22	14967.85	42.87	32.52
63	Liuzhou	33.74	18442.67	52.58	54.32
64	Guilin	33.88	12971.43	23.07	37.87
65	Hiakou	15.19	25099.00	57.02	27.50
66	Chengdu	58.20	24563.24	45.53	45.07
67	Zigong	74.09	12468.55	27.91	47.53
68	Nanchong	49.22	6949.85	18.75	35.06
69	Guiyang	29.19	16494.62	48.71	47.48
70	Kunming	12.88	21473.17	41.45	46.27
71	Lhasa	0.76	18798.18	36.37	26.15
72	Xi'an	72.89	20185.26	44.75	43.87
73	Tongchuan	33.31	11314.75	47.68	53.83
74	Baoji	28.36	13591.41	24.68	55.90
75	Xianyang	45.20	11303.62	20.89	45.81
76	Weinan	75.65	7517.76	23.00	43.61
77	Lanzhou	16.36	20061.67	59.45	49.28
78	Xining	7.04	14287.09	42.24	47.92
79	Yinchuan	16.12	23369.18	61.01	46.97
80	Wulumuqi	8.47	28358.91	81.66	38.62

Text S1

First, we introduce the basic definitions relating econometric methods used in this paper.

(1) In statistics, a unit root test tests whether a time series variable is non-stationary and possesses a unit root. The null hypothesis is generally defined as the presence of a unit root and the alternative hypothesis is either stationarity, trend stationarity or explosive root depending on the test used.

(2) If two or more series are individually integrated (in the time series sense) but some linear combination of them has a lower order of integration, then the series are said to be cointegrated.

(3) The Fully modified least squares (FMOLS) regression was designed to provide optimal estimates of cointegrating regressions. The method modifies least squares to account for serial correlation effects and for the endogeneity in the regressors that results from the existence of a cointegrating relationship.

(4) The Granger causality test is a statistical hypothesis test for determining whether one time series is useful in forecasting another and examining the direction of Granger causality between the variables.

(5) Impulse response was used to trace the effect of an exogenous shock or innovation in one of the variables on some or all of the other variables.

(6) The variance decomposition indicates the amount of information that each variable contributes to the other variables in the autoregression. It determines how much of the forecast error variance of each variable can be explained by exogenous shocks to the other variables.

In this study the first step is to test whether the variables contain a unit root. If the variables are stationary at the first difference, the second step is to test whether there is a long-run cointegrating relationship between PM2.5 concentrations, economic growth, urbanization, and industrialization. If the models are cointegrated, the third step is to estimate Eq. (1) using the Panel Fully Modified Least Squares (FMOLS) estimator. If a long-run relationship between the variables is found, the fourth step is to use a panel vector error correction model to infer both the short and long run bi-directional or un-directional causal relationships between the variables. The final step is to investigate the impact of an exogenous shock in one of the variables on some or all of the other variables, and test the amount of information that each variable contributes to the other variables using impulse response and variance decomposition, respectively.

Panel unit root tests.[1]

The precondition in the econometric analysis was to test the variables' stationariness. It is commonly recognized that panel unit root tests are much stronger than unit root tests that are based on a univariate time series or cross-sectional data. Our empirical work is based on a panel of 337 prefecture-level divisions in China from 1999 to 2011. This paper utilized two types of a panel unit root test, namely, Levin, Lin & Chu t (LLC)[2] and Im, Pesaran and Shin W-stat (IPS).[3]

The classification of our panel unit root tests is based on whether restrictions on the autoregressive process across cross-sections exist. In general, consider the following AR (1) process for the panel data:

$$y_{it} = \rho_i y_{it-1} + X_{it}\delta_i + \varepsilon_{it} \tag{1}$$

where i denotes the 1…337 prefecture-level divisions that were observed over period t (in this study, from 1999 to 2011). X_{it} represents the exogenous variables that are independently distributed for all t in the model, including any fixed effects (the fixed effect assumption is that an individual specific effect is correlated with the independent variables) or individual trends. The variables of ρ_i are the autoregressive coefficients, and the variables of ε_{it} are the error terms, which are assumed to be mutually independent idiosyncratic disturbances. If $|\rho_i| < 1$, then y_{it} is said to be weakly stationary. However, if $|\rho_i| = 1$, then y_{it} contains a unit root.[4]

The LLC and the IPS unit root tests are different because the LLC uses a common ADF regression (Eq. 1) for all prefecture-level division; specifically, the ρ_i is the same across prefecture-level division. The IPS, however, uses an individual unit root process; specifically, ρ_i can vary across prefecture-level division.

Both the LLC and IPS tests use the following specification of Augmented Dickey-Fuller (ADF):

$$\Delta y_{it} = \alpha_i y_{it-1} + \sum_{j=1}^{p_i} \beta_{ij} \Delta y_{it-j} + X'_{it}\delta + \varepsilon_{it} \tag{2}$$

where Δ is the first difference operator. We assume a common $\alpha_i = \rho_i - 1$ but allow the lag order for the difference terms, p_i, to vary across prefecture-level divisions, and Δy_{it-j} are the lag terms (for j=1,…,p_i).

Both panel unit root tests work under a null and alternative hypothesis. The LLC shows that the variables include a panel unit root, which indicates that the variables are not stationary, whereas the IPS indicates that the variables do not contain a panel unit root, which signifies that the variables are stationary.

Panel cointegration tests.

Panel cointegration tests are also well known for their high capacity compared with a normal time series cointegration.[5,6] Thus, a Pedroni panel cointegration test[7] was employed in this work to test whether a cointegration relation existed among the variables. Compared with traditional panel data models, the Pedroni panel cointegration test allows for heterogeneous intercepts and trend coefficients across cross-sections.[7] The Pedroni panel cointegration takes the following form:

$$y_{it} = \alpha_i + \delta_i t + \beta_{1i} x_{1i,t} + \beta_{2i} x_{2i,t} + \ldots + \beta_{Mi} x_{Mi,t} + e_{i,t} \tag{3}$$

where y and x are assumed to be integrated in order one. t is the number of observations over time, and M is the number of independent variables. It is assumed that the slope coefficients are $\beta_{1i},\ldots, \beta_{Mi}$. The parameters α_i and δ_i are the individual and trend effects, respectively, which may be set to zero if desired. Under the null hypothesis of no cointegration, the residual in the above equation ($e_{i,t}$) is integrated in order one. The null hypothesis of the Pedroni cointegration is that there is no cointegration. A rejection of the null hypothesis is based on several statistics. Pedroni[7] recommends two types of residual-based tests. In the first type, four tests are distributed as being standard normal asymptotically and are based on average test statistics for no cointegration in the time series across the cross-section for the within-group.[8] The tests that form this group are panel v-statistics, panel rho-statistic, panel PP-statistic (non-parametric) and panel ADF-statistic (parametric). For the second type, three tests, namely, group rho-statistics, group PP-statistics (non-parametric) and group ADF-statistics (parametric) are also distributed as being standard normal asymptotically but are based on the limits of piecewise numerator and denominator terms for the limiting distributions for the between-group.[8]

Panel Fully Modified Least Squares (FMOLS) estimates.

If the variables are cointegrated, the panel Fully Modified OLS (FMOLS) estimator can be used to estimate the single cointegration vector and to identify the direction and magnitude of the effects of the independent variables on the dependent variables based on equation (1). The FMOLS initially was proposed by Pedroni[9] and is a single equation estimator for cointegrated relations. The FMOLS utilizes errors to calculate the cumulative test volume. Specified estimates of the average long-term covariance of a one-sided long-term covariance matrix Λ_i and a long-term covariance matrix Ω_i enable the definition of the modified dependent variable ($\tilde{y}_{it}^{+}$) and serial correlation correction terms ($\hat{\lambda}_{12}^{+}$) in the long term:

$$\tilde{y}_{it}^{+} = \tilde{y}_{it} - \tilde{\omega}_{12}\hat{\Omega}_{22}^{-1}\hat{u}_2 \tag{4}$$

and

$$\hat{\lambda}_{12}^{+} = \hat{\lambda}_{12} - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}\hat{\Lambda}_{22} \tag{5}$$

where $\tilde{y}_{it}$ represents the corresponding data for the individual deterministic trends and $\tilde{\omega}_{12}$ is the long-term average variance of the residuals $\hat{u}_2$. The estimator of the pooled FMOLS can be formulated as:

$$\hat{\beta}_{FP}=\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{X}_{it}\tilde{X}_{it'}\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\tilde{X}_{it}\tilde{y}_{it}^{+}-\tilde{\lambda}_{12}^{+}\right) \tag{6}$$

where $\hat{\beta}_{FP}$ is the estimator of the pooled FMOLS and $\tilde{X}_{it}$ is the corresponding data purged of the individual deterministic trends. The pooled estimator is summed across prefecture-level divisions separately in the numerator and denominator.[10]

Panel Granger causality test.

The Granger causality test was used in this study to test the causal relationship between $PM_{2.5}$, economic growth, industrialization, and urbanization. If the models are cointegrated, the Granger causality that is based on the vector error-correction model (VECM) is utilized. The VECM Granger causality can indicate both the short-term causal relation that is based on the χ^2-Wald statistics and the long-term causal relation that is based on the error correction term *ETC* (-1). The VECM models can be formulized as follows:

$$(1\text{-}B)\begin{bmatrix}\Delta LNPM_{2.5it}\\ \Delta LNGDPPC_{it}\\ \Delta LNURBAN_{it}\\ \Delta LNIND_{it}\end{bmatrix}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\end{bmatrix}+\sum_{i=1}^{p}(1-B)\begin{bmatrix}\beta_{11i}&\beta_{12i}&\beta_{13i}&\beta_{14i}\\ \beta_{21i}&\beta_{22i}&\beta_{23i}&\beta_{24i}\\ \beta_{31i}&\beta_{32i}&\beta_{33i}&\beta_{34i}\\ \beta_{41i}&\beta_{42i}&\beta_{43i}&\beta_{44i}\end{bmatrix}\times\begin{bmatrix}\Delta LNPM_{2.5it-1}\\ \Delta LNGDPPC_{it-1}\\ \Delta LNURBAN_{it-1}\\ \Delta LNIND_{it-1}\end{bmatrix}+\begin{bmatrix}\varphi_1\\ \varphi_2\\ \varphi_3\\ \varphi_4\end{bmatrix}ECT_{it-1}+\begin{bmatrix}\varepsilon_{1it}\\ \varepsilon_{2it}\\ \varepsilon_{3it}\\ \varepsilon_{4it}\end{bmatrix} \tag{7}$$

where $LNPM_{2.5}$ is the natural logarithm of $PM_{2.5}$ concentrations, LNGDPPC is the natural logarithm of per capita GDP, LNURBAN is the natural logarithm of urbanization and LNIND is the natural logarithm of industrialization. The variable *t* indicates the time (1999 to 2011), *i* indicates the 337 prefecture-level divisions, (1-B) is the lag operator, Δ is the first difference operator, α's are the constant terms, β's and φ's are parameters to be estimated, ECT_{it-1} is the lagged error correction term that is obtained from the cointegration equation, and ε_{it} is the white noise error. The χ^2-Wald statistics for the lagged independent variables of the VECM represent the significance of the short-term causal effects. The *t*-statistics for the coefficients of the lagged error-correction term indicate the significance of the long-term causal effect. The lag length *p* is based on the Schwarz–Bayesian (SBC) and Akaike (AIC) Information Criteria.

However, if the models are not cointegrated, the Vector autoregressive (VAR) Granger causality model that identifies only the short-term causal relation is utilized. The VAR Granger causality model is as follows:

$$(1\text{-}B)\begin{bmatrix}\Delta LNPM_{2.5it}\\ \Delta LNGDPPC_{it}\\ \Delta LNURBAN_{it}\\ \Delta LNIND_{it}\end{bmatrix}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\end{bmatrix}+\sum_{i=1}^{p}(1-B)\begin{bmatrix}\beta_{11i} & \beta_{12i} & \beta_{13i} & \beta_{14i}\\ \beta_{21i} & \beta_{22i} & \beta_{23i} & \beta_{24i}\\ \beta_{31i} & \beta_{32i} & \beta_{33i} & \beta_{34i}\\ \beta_{41i} & \beta_{42i} & \beta_{43i} & \beta_{44i}\end{bmatrix}\times\begin{bmatrix}\Delta LNPM_{2.5it-1}\\ \Delta LNGDPPC_{it-1}\\ \Delta LNURBAN_{it-1}\\ \Delta LNIND_{it-1}\end{bmatrix}+\begin{bmatrix}\varepsilon_{1it}\\ \varepsilon_{2it}\\ \varepsilon_{3it}\\ \varepsilon_{4it}\end{bmatrix} \quad (8)$$

Impulse response and variance decomposition.

1) Impulse Response[11]

We utilized the impulse response function to identify the impacts of a shock to one variable on the other three variables. A shock to one variable not only directly impacts the variable itself but also is transmitted to all of the other variables by the dynamic structure of the model.

Lutkepohl[12] and Hamilton[13] both suggest that a VAR model is stable if all moduli of the companion matrix $\bar{A}$ are strictly less than one. The companion matrix can be shown as follows (see reference 11 for a detailed description):

$$\bar{A}=\begin{bmatrix}A_1 & A_2 & \cdots & A_p & A_{p-1}\\ I_k & O_k & \cdots & O_k & O_k\\ O_k & I_k & \cdots & O_k & O_k\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ O_k & O_k & \cdots & I_k & O_k\end{bmatrix} \quad (9)$$

Stability often means that the panel VAR is invertible and has an infinite-order vector moving-average (VMA) representation. The simple impulse-response function Φ_i may be calculated by rewriting the model as an infinite vector moving-average, where Φ_i are the VMA parameters.

$$\Phi_i=\begin{cases}I_k, & i=0\\ \sum_{j=1}^{i}\Phi_{i-j}A_j, & i=1,2,\ldots\end{cases} \quad (10)$$

However, the simple impulse response functions (IRFs) have no causal interpretation. Because the innovations of e_{it} (dependent variable-specific panel fixed-effects and idiosyncratic errors) are correlated contemporaneously, the shock to one variable is likely to be accompanied by shocks in the other variables as well. We assume that the matrix P exists such that $P'P=\sum$, and then P can be utilized to orthogonalize the innovations as $e_{it}P^{-1}$ and to convert the VMA parameters into the orthogonalized impulse-responses of $P\Phi_i$. The matrix of P effectively imposes identification restrictions on the system of dynamic equations. Sims[14] proposes that the Cholesky decomposition of

Σ imposes a recursive structure on a VAR. However, the decomposition is not unique but depends on the ordering of the variables in Σ.

Confidence intervals for the impulse-response function may be derived analytically based on the asymptotic distribution of the panel VAR parameters and the cross-equation error variance-covariance matrix. Alternatively, the confidence interval may likewise be estimated using a Monte Carlo simulation and bootstrap resampling methods (for example, see Lutkepohl[12] for details that are applied to time-series VAR).

2) Forecast-error variance decomposition[11]

The Granger-causality tests show only the existence of causality among the variables. To determine the importance of the causal effect of one variable on another and estimate how each variable responds to the changes in the other variables, we utilized a variance decomposition analysis by applying the Cholesky decomposition technique in VECM.[12]

The h-step ahead forecast-error can be formulized as:

$$Y_{it+h} - E[Y_{it+h}] = \sum_{i=0}^{h-1} e_{i(t+h-i)} \Phi_i \tag{11}$$

where Y_{it+h} is the observed vector at time $t+h$ and $E[Y_{it+h}]$ is the h-step ahead predicted vector made at time t. Similar to impulse-response functions, we orthogonalized the shocks by employing matrix P to isolate each variable's contribution to the forecast-error variance. The orthogonalized shocks of $e_{it}P^{-1}$ have a covariance matrix of I_k, which enables the straightforward decomposition of the forecast-error variance. Specifically, the contribution of a variable m to the h-step ahead forecast-error variance of variable n may be computed as:

$$\sum_{i=0}^{h-1} \theta_{mn}^2 = \sum_{i=1}^{h-1} \left(i_n' P \Phi_i i_m \right)^2 \tag{12}$$

where i_s is the s-th column of I_k. In practice, the shares are often normalized relative to the h-step ahead forecast-error variance of variable *n:*

$$\sum_{i=0}^{h-1} \theta_n^2 = \sum_{i=1}^{h-1} i_n' \Phi_i' \Sigma \Phi_i i_n \tag{13}$$

Similar to impulse-response functions, confidence intervals may be derived analytically or estimated using various resampling techniques.

Table S5 Panel unit root test results of variables for four sub-panels.

Panel	Variable	Level		First difference	
		Intercept	Intercept and trend	Intercept	Intercept and trend
I1		Levin, Lin & Chu t			
	$LNPM_{2.5}$	-29.3778^{***}	-12.4353^{***}	-24.2788^{***}	-37.8992^{***}
	LNGDPPC	9.81137	-31.4240^{***}	-24.9052^{***}	-16.2063^{***}
	LNING	-41.2730	-21.7901^{***}	-15.6273^{***}	-24.4725^{***}
	LNURBAN	-18.7918^{***}	-56.8121^{***}	-45.8896^{***}	-45.5637^{***}
		Im, Pesaran and Shin W-stat			
	$LNPM_{2.5}$	-14.5734	7.2350	-16.7463^{***}	-20.9570^{***}
	LNGDPPC	32.0693	-12.4816^{***}	-12.5586^{***}	-1.7384^{***}
	LNING	-9.8633	-3.6300^{***}	-10.4152^{***}	-10.8882^{***}
	LNURBAN	14.4419	-12.1242^{***}	-15.6272^{***}	-15.3581^{***}
I2		Levin, Lin & Chu t			
	$LNPM_{2.5}$	-16.7356^{***}	-3.2596^{***}	-7.8100^{***}	-16.5209^{***}
	LNGDPPC	-1.9300^{***}	-10.5141^{***}	-14.8711^{***}	-7.7821^{***}
	LNING	-28.5468^{***}	-5.9551^{***}	-3.1916^{***}	-13.9121^{***}
	LNURBAN	-14.8550^{***}	-20.5201^{***}	-8.6026^{***}	-9.4325^{***}
		Im, Pesaran and Shin W-stat			
	$LNPM_{2.5}$	-8.8888^{***}	6.3084	-3.8607^{***}	-9.0067^{***}
	LNGDPPC	9.4432	-5.1393^{***}	-10.1468^{***}	-1.8264^{***}
	LNING	-11.4571^{***}	1.6378	-2.1020^{***}	-7.2185^{***}
	LNURBAN	-3.2058^{***}	-7.9188^{***}	-4.0697^{***}	-4.1683^{***}
I3		Levin, Lin & Chu t			
	$LNPM_{2.5}$	-15.2012^{***}	-2.9799^{***}	-9.2142^{***}	-13.1852^{***}
	LNGDPPC	-0.4019	-11.6466^{***}	-22.4365^{***}	-19.1971^{***}
	LNING	-19.4891^{***}	-5.8497^{***}	-3.6098^{***}	-10.2955^{***}
	LNURBAN	-2693.0000^{***}	-3462.4200^{***}	-3055.3400^{***}	-12.2496^{***}
		Im, Pesaran and Shin W-stat			
	$LNPM_{2.5}$	-9.0082^{***}	3.3917	-5.4802^{***}	-7.4945^{***}
	LNGDPPC	7.0881	-5.4087^{***}	-9.7737^{***}	-3.8168^{***}
	LNING	-6.1772^{***}	-0.2585	-2.1999^{***}	-4.5975^{***}
	LNURBAN	-535.9400^{***}	-190.1580^{***}	-555.7420^{***}	-4.5029^{***}
H		Levin, Lin & Chu t			
	$LNPM_{2.5}$	-12.8666^{***}	-3.0703^{***}	-10.6617^{***}	-17.7616^{***}
	LNGDPPC	-8.4986^{***}	-5.9312^{***}	-8.4435^{***}	-8.0632^{***}
	LNING	-21.6561^{***}	-10.8614^{***}	-5.5660^{***}	-9.8972^{***}
	LNURBAN	-8.3829^{***}	-10.5774^{***}	-4.0015^{***}	-6.5278^{***}
		Im, Pesaran and Shin W-stat			
	$LNPM_{2.5}$	-6.9944^{***}	3.9292	-5.7752^{***}	-8.1370^{***}
	LNGDPPC	2.0160	-3.2514^{***}	-5.5888^{***}	-6.5167^{***}
	LNING	-11.3915^{***}	-0.7224	-2.5705^{***}	-5.7045^{***}
	LNURBAN	-1.2806^{***}	-1.6206^{***}	-2.2794^{***}	-3.1815^{***}

Notes: The panel unit root tests with intercept and trend are carried independently; the optimal lag lengths are obtained automatically with the Schwarz information criteria (SIC).

Table S6 Pedroni cointegration test results for four sub-panels.

Panel	*Alternative hypothesis: common AR coefs. (within-dimension)*	Statistic	Prob.
I1	Panel v-Statistic	0.8585	0.1953
	Panel rho-Statistic	1.6349	0.9490
	Panel PP-Statistic	-23.4958	0.0000
	Panel ADF-Statistic	-23.3785	0.0000
	Panel v-Statistic (Weighted statistic)	-4.7415	1.0000
	Panel rho-Statistic (Weighted statistic)	7.5469	0.0046
	Panel PP-Statistic (Weighted statistic)	-23.4115	0.0000
	Panel ADF-Statistic (Weighted statistic)	-22.7677	0.0000
	Alternative hypothesis: individual AR coefs. (between-dimension)		
	Group rho-Statistic	8.9545	1.0000
	Group PP-Statistic	-37.0879	0.0000
	Group ADF-Statistic	-26.8270	0.0000
I2	*Alternative hypothesis: common AR coefs. (within-dimension)*	Statistic	Prob.
	Panel v-Statistic	-0.9160	0.8202
	Panel rho-Statistic	0.9243	0.8223
	Panel PP-Statistic	-13.2019	0.0000
	Panel ADF-Statistic	-14.0104	0.0000
	Panel v-Statistic (Weighted statistic)	-2.6867	0.9964
	Panel rho-Statistic (Weighted statistic)	8.0338	0.0024
	Panel PP-Statistic (Weighted statistic)	-14.2050	0.0000
	Panel ADF-Statistic (Weighted statistic)	-15.1890	0.0000
	Alternative hypothesis: individual AR coefs. (between-dimension)		
	Group rho-Statistic	3.6528	0.9999
	Group PP-Statistic	-20.3997	0.0000
	Group ADF-Statistic	-18.4817	0.0000
I3	*Alternative hypothesis: common AR coefs. (within-dimension)*	Statistic	Prob.
	Panel v-Statistic	-0.8342	0.7979
	Panel rho-Statistic	1.5020	0.9335
	Panel PP-Statistic	-7.7953	0.0000
	Panel ADF-Statistic	-9.8030	0.0000
	Panel v-Statistic (Weighted statistic)	-2.5484	0.9946
	Panel rho-Statistic (Weighted statistic)	6.3139	0.0056
	Panel PP-Statistic (Weighted statistic)	-6.8599	0.0000
	Panel ADF-Statistic (Weighted statistic)	-8.6895	0.0000
	Alternative hypothesis: individual AR coefs. (between-dimension)		
	Group rho-Statistic	3.9825	1.0000
	Group PP-Statistic	-13.2066	0.0000
	Group ADF-Statistic	-11.0367	0.0000
H	*Alternative hypothesis: common AR coefs. (within-dimension)*	Statistic	Prob.
	Panel v-Statistic	-1.5556	0.9401
	Panel rho-Statistic	1.5563	0.9402
	Panel PP-Statistic	-9.6981	0.0000
	Panel ADF-Statistic	-9.9505	0.0000
	Panel v-Statistic (Weighted statistic)	-2.8160	0.9976
	Panel rho-Statistic (Weighted statistic)	5.9055	0.0016
	Panel PP-Statistic (Weighted statistic)	-9.3155	0.0000
	Panel ADF-Statistic (Weighted statistic)	-10.4871	0.0000
	Alternative hypothesis: individual AR coefs. (between-dimension)		
	Group rho-Statistic	4.2853	1.0000
	Group PP-Statistic	-15.7159	0.0000
	Group ADF-Statistic	-12.0924	0.0000

Note: we use the automatic selection based on Schwarz to choose the optimal lag length.

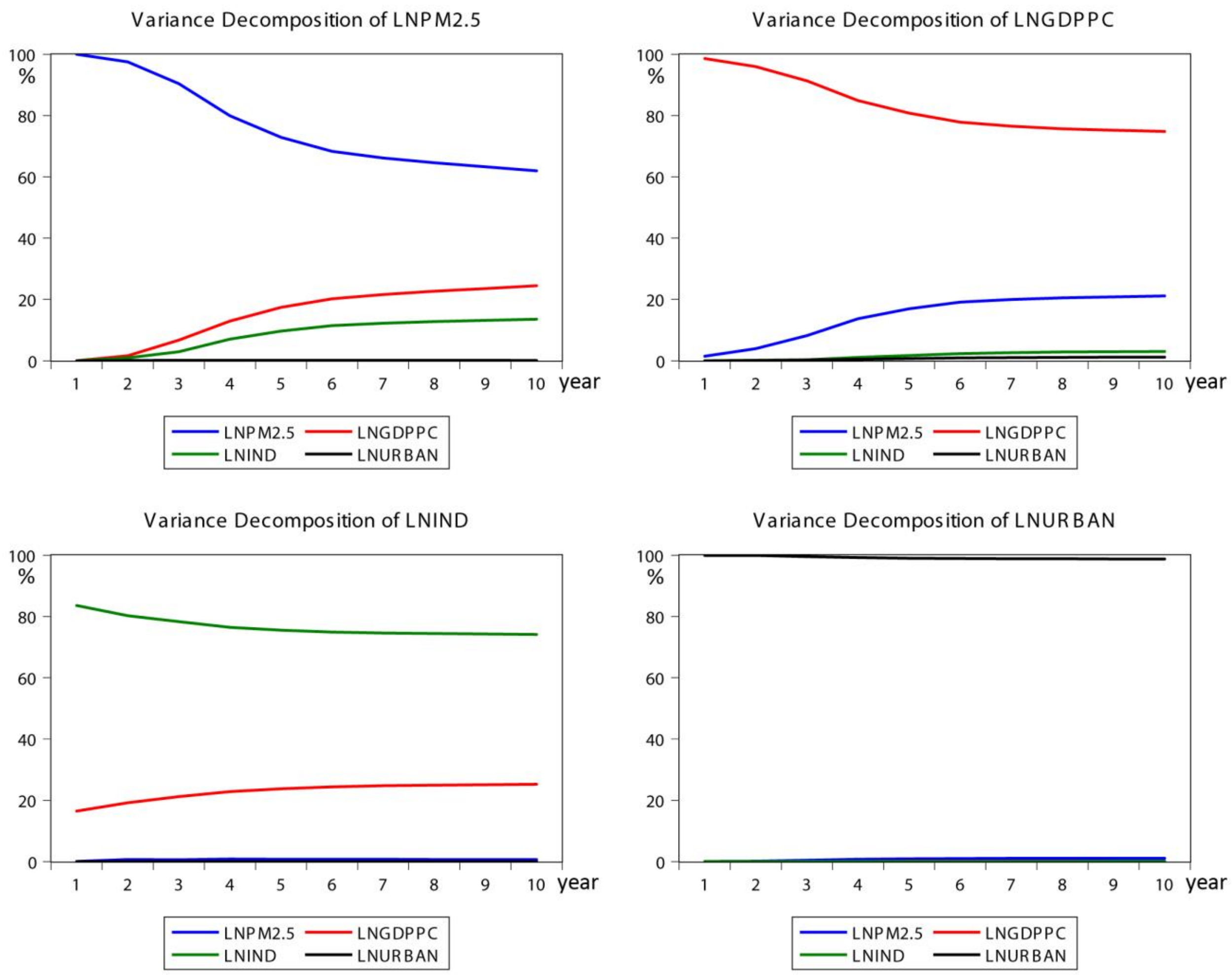

Figure S4. Variance decomposition of four variables for panel I1.

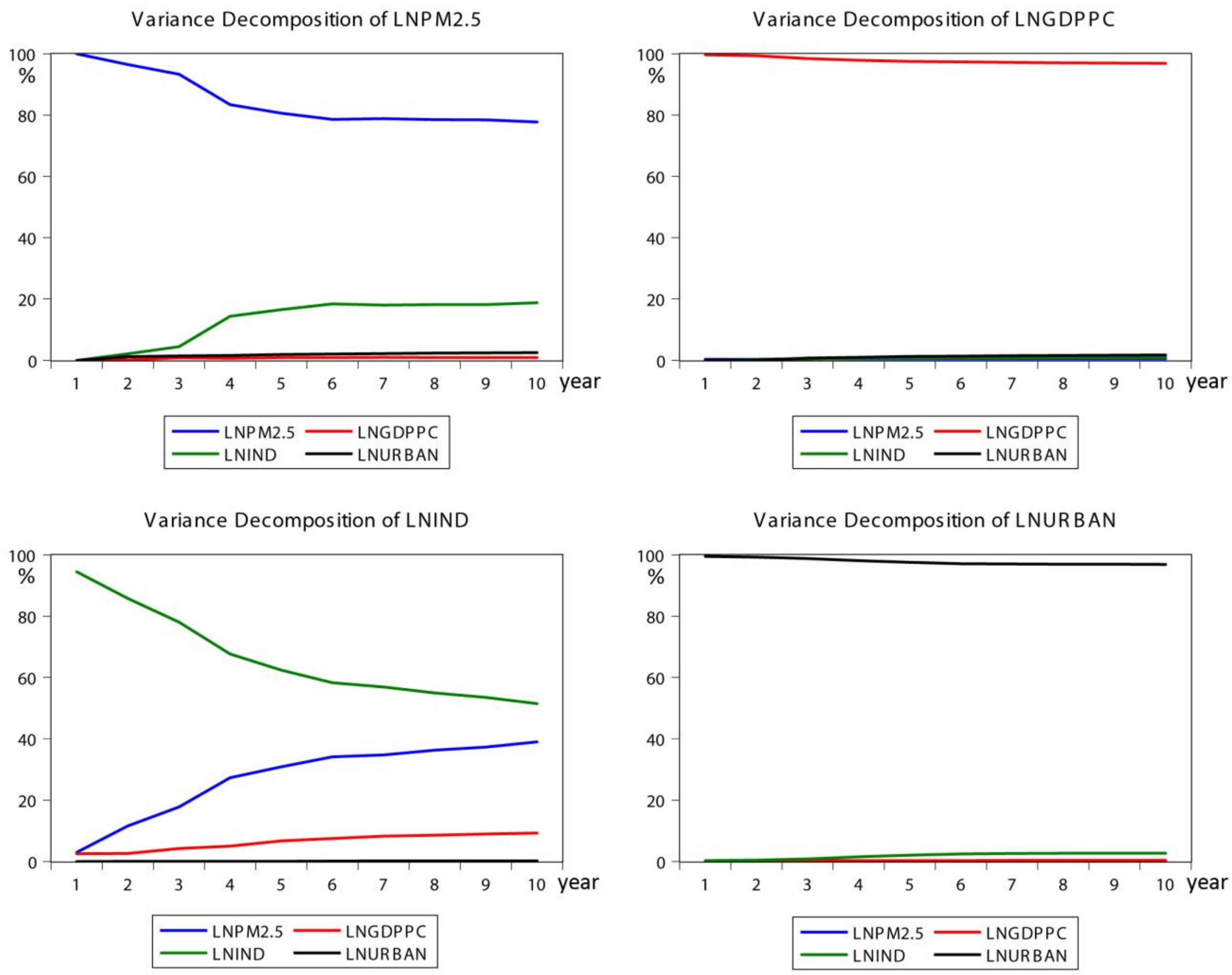

Figure S5. Variance decomposition of four variables for panel I2.

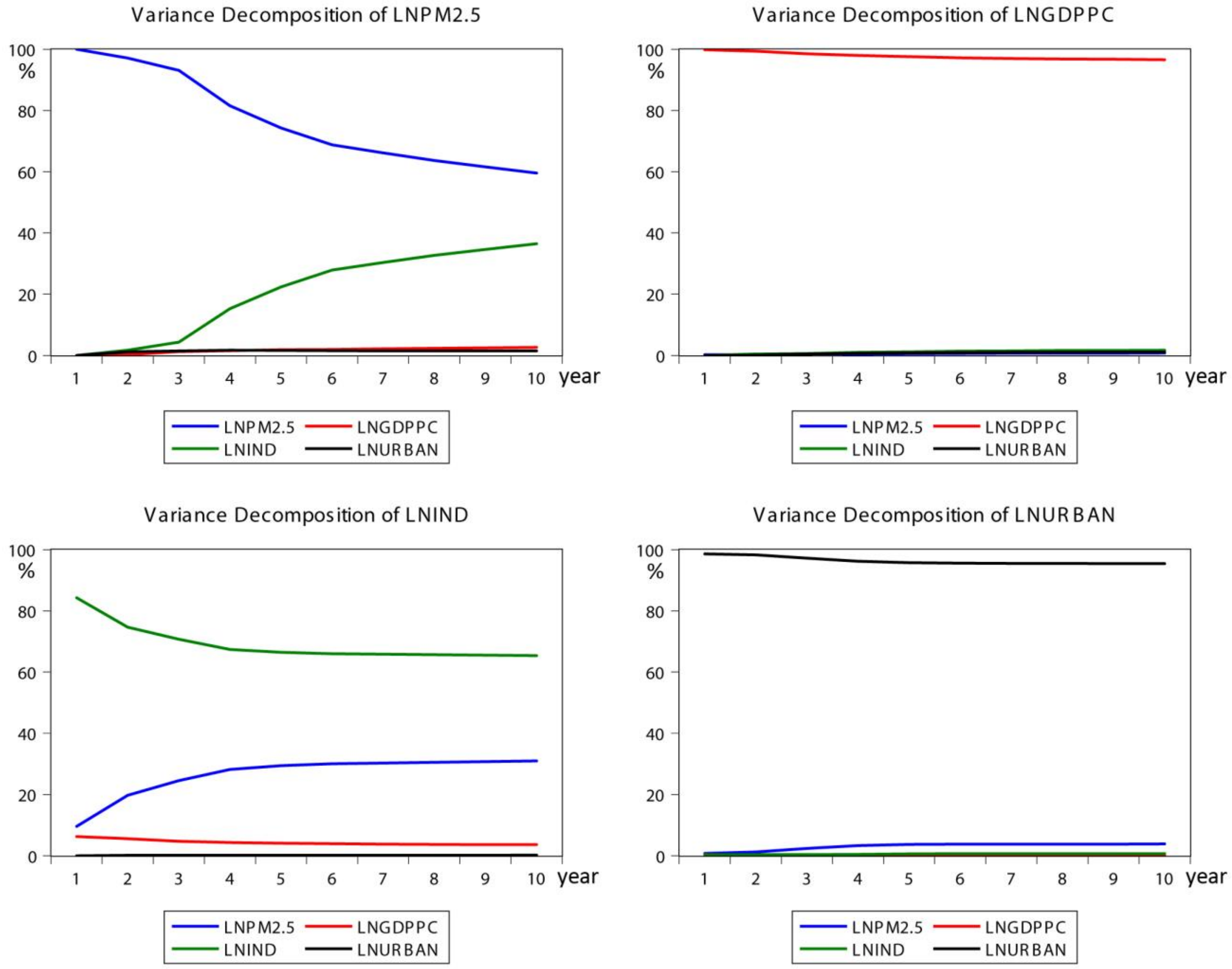

Figure S6. Variance decomposition of four variables for panel I3.

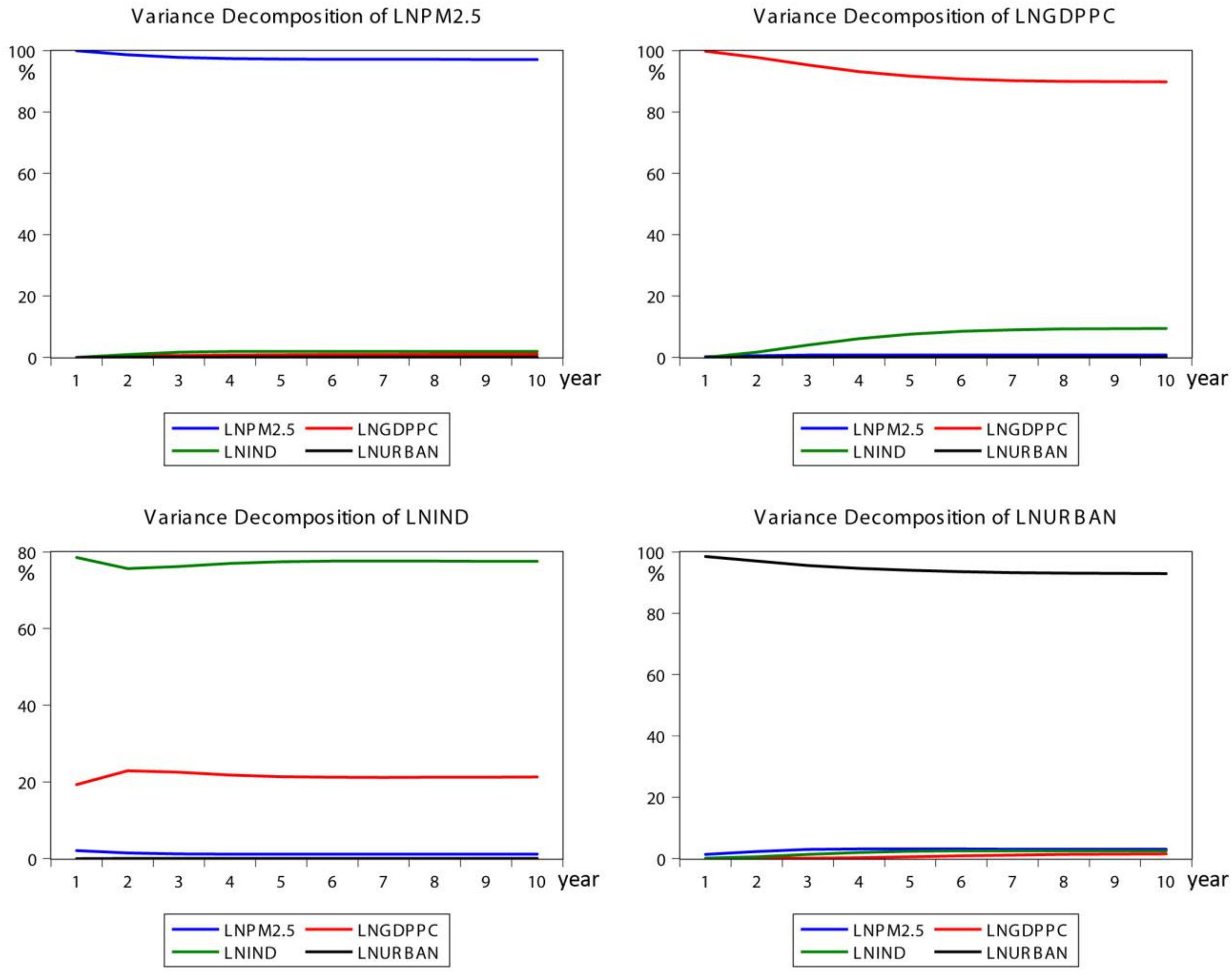

Figure S7. Variance decomposition of four variables for panel H.

Table S7 Variance decomposition of four variables for panel C.

Period	S.E.	LNPM2.5	LNGDPPC	LNIND	LNURBAN
Variance Decomposition of LNPM2.5:					
1	0.0705	100.0000	0.0000	0.0000	0.0000
2	0.0734	97.7341	0.7282	1.4936	0.0441
3	0.0787	91.6850	4.1771	4.0432	0.0947
4	0.0845	80.5531	7.8911	11.4735	0.0822
5	0.0908	73.2269	10.8909	15.8108	0.0715
6	0.0968	68.0199	12.6875	19.2276	0.0650
7	0.1024	65.4714	13.8019	20.6581	0.0687
8	0.1074	63.4230	14.6574	21.8444	0.0752
9	0.1121	61.7868	15.4113	22.7204	0.0815
10	0.1165	60.1198	16.1254	23.6700	0.0847
Variance Decomposition of LNGDPPC:					
1	0.0456	0.6330	99.3670	0.0000	0.0000
2	0.0608	1.5575	98.1640	0.2654	0.0130
3	0.0779	3.1635	96.0435	0.5133	0.2797
4	0.0908	5.3141	92.8924	1.3259	0.4677
5	0.1027	6.4290	90.9898	1.9422	0.6389
6	0.1131	7.1409	89.6055	2.5040	0.7497
7	0.1227	7.3610	89.0217	2.7931	0.8242
8	0.1315	7.5006	88.6406	2.9868	0.8720
9	0.1398	7.5721	88.4267	3.0947	0.9064
10	0.1476	7.6684	88.2099	3.1882	0.9335
Variance Decomposition of LNIND:					
1	0.0371	0.8370	12.7311	86.4319	0.0000
2	0.0492	4.0694	15.1092	80.8097	0.0117
3	0.0611	5.1507	16.9557	77.8860	0.0076
4	0.0704	7.2101	18.8296	73.9532	0.0071
5	0.0786	7.8251	20.1155	72.0529	0.0065
6	0.0859	8.3021	21.0373	70.6545	0.0061
7	0.0927	8.3659	21.5768	70.0521	0.0053
8	0.0991	8.4512	21.9477	69.5964	0.0047
9	0.1050	8.4889	22.2049	69.3020	0.0041
10	0.1107	8.5657	22.4248	69.0058	0.0037
Variance Decomposition of LNURBAN:					
1	0.0126	0.0200	0.0008	0.0132	99.9659
2	0.0190	0.0388	0.0008	0.0067	99.9537
3	0.0234	0.1633	0.0058	0.0051	99.8258
4	0.0270	0.2109	0.0104	0.0062	99.7724
5	0.0302	0.2383	0.0132	0.0083	99.7402
6	0.0331	0.2442	0.0144	0.0086	99.7329
7	0.0358	0.2470	0.0148	0.0084	99.7298
8	0.0383	0.2482	0.0150	0.0080	99.7289
9	0.0406	0.2506	0.0152	0.0077	99.7266
10	0.0428	0.2531	0.0153	0.0075	99.7241

Table S8 Variance decomposition of four variables for panel I1.

Period	S.E.	LNPM2.5	LNGDPPC	LNIND	LNURBAN
Variance Decomposition of LNPM2.5:					
1	0.0718	100.0000	0.0000	0.0000	0.0000
2	0.0749	97.4874	1.6261	0.8848	0.0017
3	0.0818	90.3335	6.6657	2.9220	0.0788
4	0.0879	79.9338	12.9313	7.0432	0.0916
5	0.0951	72.8512	17.4028	9.6450	0.1010
6	0.1017	68.2929	20.1700	11.4304	0.1067
7	0.1080	66.1192	21.6056	12.1792	0.0960
8	0.1136	64.5451	22.6567	12.7114	0.0868
9	0.1188	63.2690	23.5364	13.1153	0.0794
10	0.1237	61.9537	24.4178	13.5553	0.0732
Variance Decomposition of LNGDPPC:					
1	0.0433	1.4217	98.5783	0.0000	0.0000
2	0.0570	3.9380	95.9014	0.1436	0.0170
3	0.0727	8.2015	91.3175	0.2859	0.1951
4	0.0844	13.6779	84.8490	1.0550	0.4180
5	0.0951	16.9233	80.7283	1.6905	0.6580
6	0.1042	19.0741	77.7787	2.2948	0.8524
7	0.1125	19.9429	76.4634	2.6136	0.9802
8	0.1202	20.4723	75.6550	2.8136	1.0591
9	0.1275	20.7658	75.2085	2.9181	1.1076
10	0.1344	21.0816	74.7716	3.0031	1.1438
Variance Decomposition of LNIND:					
1	0.0373	0.0133	16.4461	83.5406	0.0000
2	0.0501	0.6316	19.1474	80.2204	0.0007
3	0.0621	0.5647	21.1954	78.2358	0.0041
4	0.0718	0.7806	22.7940	76.4217	0.0036
5	0.0805	0.7511	23.7792	75.4655	0.0041
6	0.0883	0.7423	24.3873	74.8660	0.0044
7	0.0955	0.6987	24.7187	74.5772	0.0054
8	0.1022	0.6733	24.9373	74.3830	0.0065
9	0.1085	0.6519	25.0961	74.2444	0.0075
10	0.1145	0.6411	25.2371	74.1137	0.0082
Variance Decomposition of LNURBAN:					
1	0.0123	0.0262	0.0157	0.0560	99.9021
2	0.0186	0.1009	0.0142	0.0296	99.8552
3	0.0229	0.4456	0.0238	0.0235	99.5071
4	0.0264	0.7088	0.0537	0.0358	99.2017
5	0.0295	0.8941	0.0825	0.0547	98.9688
6	0.0323	0.9804	0.1027	0.0682	98.8487
7	0.0348	1.0187	0.1125	0.0748	98.7940
8	0.0372	1.0354	0.1170	0.0772	98.7704
9	0.0394	1.0485	0.1192	0.0781	98.7542
10	0.0415	1.0623	0.1212	0.0788	98.7378

Table S9 Variance decomposition of four variables for panel I2.

Period	S.E.	LNPM2.5	LNGDPPC	LNIND	LNURBAN
Variance Decomposition of LNPM2.5:					
1	0.0643	100.0000	0.0000	0.0000	0.0000
2	0.0704	96.4897	0.2059	2.0918	1.2127
3	0.0771	93.3320	0.7936	4.4567	1.4177
4	0.0867	83.3808	0.6822	14.3323	1.6047
5	0.0944	80.6287	0.9196	16.5759	1.8759
6	0.1018	78.6075	0.8871	18.4215	2.0839
7	0.1081	78.8485	0.9347	17.9737	2.2431
8	0.1139	78.5423	0.9089	18.1790	2.3699
9	0.1193	78.4115	0.9284	18.1978	2.4623
10	0.1246	77.7453	0.9239	18.7921	2.5387
Variance Decomposition of LNGDPPC:					
1	0.0503	0.3274	99.6726	0.0000	0.0000
2	0.0654	0.3661	99.3399	0.2506	0.0435
3	0.0837	0.5105	98.3972	0.3909	0.7014
4	0.0970	0.5186	97.8488	0.6491	0.9836
5	0.1098	0.5108	97.5115	0.7363	1.2414
6	0.1209	0.5051	97.2937	0.8125	1.3887
7	0.1313	0.4951	97.1628	0.8315	1.5107
8	0.1408	0.4914	97.0546	0.8589	1.5952
9	0.1498	0.4884	96.9735	0.8735	1.6646
10	0.1582	0.4880	96.8989	0.8954	1.7176
Variance Decomposition of LNIND:					
1	0.0349	2.9605	2.5539	94.4856	0.0000
2	0.0387	11.5931	2.5939	85.8120	0.0010
3	0.0433	17.7424	4.2347	78.0222	0.0008
4	0.0465	27.3110	4.9960	67.6533	0.0397
5	0.0490	30.8431	6.7139	62.3728	0.0702
6	0.0511	34.0932	7.4511	58.3183	0.1374
7	0.0536	34.7085	8.2224	56.8982	0.1709
8	0.0558	36.2834	8.5463	54.9706	0.1998
9	0.0582	37.3046	8.9690	53.5151	0.2113
10	0.0602	39.0281	9.2810	51.4652	0.2258
Variance Decomposition of LNURBAN:					
1	0.0148	0.0959	0.1557	0.2684	99.4800
2	0.0218	0.1446	0.1535	0.4388	99.2631
3	0.0272	0.1621	0.2810	0.7923	98.7646
4	0.0317	0.1193	0.3007	1.4999	98.0801
5	0.0357	0.0953	0.3055	2.0740	97.5253
6	0.0393	0.0842	0.3089	2.4747	97.1321
7	0.0426	0.0843	0.3137	2.6031	96.9989
8	0.0456	0.0856	0.3218	2.6514	96.9412
9	0.0485	0.0856	0.3282	2.6644	96.9218
10	0.0512	0.0829	0.3338	2.7028	96.8805

Table S10 Variance decomposition of four variables for panel I3.

Period	S.E.	LNPM2.5	LNGDPPC	LNIND	LNURBAN
Variance Decomposition of LNPM2.5:					
1	0.0770	100.0000	0.0000	0.0000	0.0000
2	0.0801	97.0768	0.1846	1.6508	1.0878
3	0.0831	93.0731	1.1812	4.3159	1.4299
4	0.0890	81.5207	1.5254	15.2591	1.6949
5	0.0958	74.2321	1.7979	22.3663	1.6037
6	0.1031	68.7320	1.9215	27.8298	1.5167
7	0.1094	66.1046	2.0997	30.3511	1.4446
8	0.1149	63.6457	2.2854	32.6459	1.4230
9	0.1198	61.5598	2.4632	34.5663	1.4107
10	0.1246	59.5133	2.6044	36.4801	1.4022
Variance Decomposition of LNGDPPC:					
1	0.0481	0.2187	99.7813	0.0000	0.0000
2	0.0662	0.2079	99.3309	0.3945	0.0667
3	0.0842	0.4696	98.5028	0.5901	0.4375
4	0.0986	0.4438	97.9428	0.9431	0.6703
5	0.1116	0.5216	97.5225	1.1558	0.8001
6	0.1233	0.5635	97.1930	1.3573	0.8862
7	0.1339	0.6276	96.9644	1.4666	0.9414
8	0.1438	0.6646	96.7954	1.5552	0.9848
9	0.1530	0.6947	96.6749	1.6133	1.0172
10	0.1616	0.7131	96.5783	1.6649	1.0437
Variance Decomposition of LNIND:					
1	0.0373	9.5517	6.2307	84.2176	0.0000
2	0.0541	19.6818	5.5519	74.6527	0.1136
3	0.0697	24.5131	4.6970	70.6761	0.1138
4	0.0815	28.1677	4.3270	67.3735	0.1319
5	0.0913	29.3606	4.0676	66.4359	0.1360
6	0.1000	30.0136	3.9109	65.9337	0.1419
7	0.1081	30.2255	3.7900	65.8395	0.1450
8	0.1156	30.4955	3.6993	65.6561	0.1492
9	0.1228	30.7284	3.6263	65.4931	0.1522
10	0.1296	30.9638	3.5683	65.3129	0.1550
Variance Decomposition of LNURBAN:					
1	0.0105	0.7210	0.3825	0.2973	98.5992
2	0.0158	1.2350	0.1797	0.3123	98.2730
3	0.0196	2.3824	0.1388	0.3163	97.1624
4	0.0228	3.2713	0.1288	0.4240	96.1760
5	0.0256	3.6458	0.1205	0.5441	95.6896
6	0.0281	3.7296	0.1144	0.6243	95.5318
7	0.0304	3.7361	0.1107	0.6596	95.4937
8	0.0326	3.7551	0.1088	0.6740	95.4622
9	0.0346	3.7950	0.1077	0.6831	95.4143
10	0.0365	3.8391	0.1067	0.6929	95.3612

Table S11 Variance decomposition of four variables for panel H.

Period	S.E.	LNPM2.5	LNGDPPC	LNIND	LNURBAN
Variance Decomposition of LNPM2.5:					
1	0.0706	100.0000	0.0000	0.0000	0.0000
2	0.0711	98.6471	0.4912	0.8593	0.0024
3	0.0716	97.7552	0.5654	1.6771	0.0024
4	0.0717	97.4044	0.7072	1.8859	0.0025
5	0.0717	97.2667	0.8075	1.9234	0.0025
6	0.0718	97.1972	0.8741	1.9263	0.0025
7	0.0718	97.1585	0.9133	1.9255	0.0026
8	0.0718	97.1370	0.9338	1.9264	0.0028
9	0.0718	97.1254	0.9435	1.9283	0.0029
10	0.0718	97.1194	0.9476	1.9300	0.0030
Variance Decomposition of LNGDPPC:					
1	0.0251	0.1856	99.8144	0.0000	0.0000
2	0.0318	0.5091	97.7773	1.6996	0.0141
3	0.0351	0.7286	95.2854	3.9715	0.0146
4	0.0367	0.7442	93.1688	6.0553	0.0317
5	0.0375	0.7272	91.6589	7.5600	0.0538
6	0.0379	0.7148	90.7302	8.4833	0.0718
7	0.0380	0.7090	90.2233	8.9843	0.0834
8	0.0381	0.7071	89.9740	9.2290	0.0899
9	0.0381	0.7070	89.8629	9.3370	0.0931
10	0.0381	0.7073	89.8183	9.3800	0.0944
Variance Decomposition of LNIND:					
1	0.0166	2.0964	19.3289	78.5747	0.0000
2	0.0208	1.4296	22.8700	75.6598	0.0406
3	0.0224	1.2298	22.5150	76.2089	0.0463
4	0.0231	1.1629	21.8050	76.9881	0.0439
5	0.0233	1.1443	21.3711	77.4411	0.0435
6	0.0234	1.1430	21.2059	77.6072	0.0439
7	0.0235	1.1457	21.1827	77.6276	0.0441
8	0.0235	1.1481	21.2084	77.5995	0.0441
9	0.0235	1.1494	21.2392	77.5674	0.0441
10	0.0235	1.1500	21.2614	77.5446	0.0441
Variance Decomposition of LNURBAN:					
1	0.0179	1.3742	0.0275	0.1121	98.4863
2	0.0219	2.3173	0.0893	0.5607	97.0327
3	0.0233	2.9905	0.0960	1.3287	95.5847
4	0.0238	3.1264	0.2566	1.9793	94.6377
5	0.0240	3.1193	0.5552	2.3423	93.9832
6	0.0241	3.0996	0.8790	2.4809	93.5405
7	0.0242	3.0864	1.1464	2.5126	93.2547
8	0.0242	3.0789	1.3327	2.5119	93.0765
9	0.0242	3.0749	1.4477	2.5085	92.9689
10	0.0242	3.0727	1.5120	2.5090	92.9064

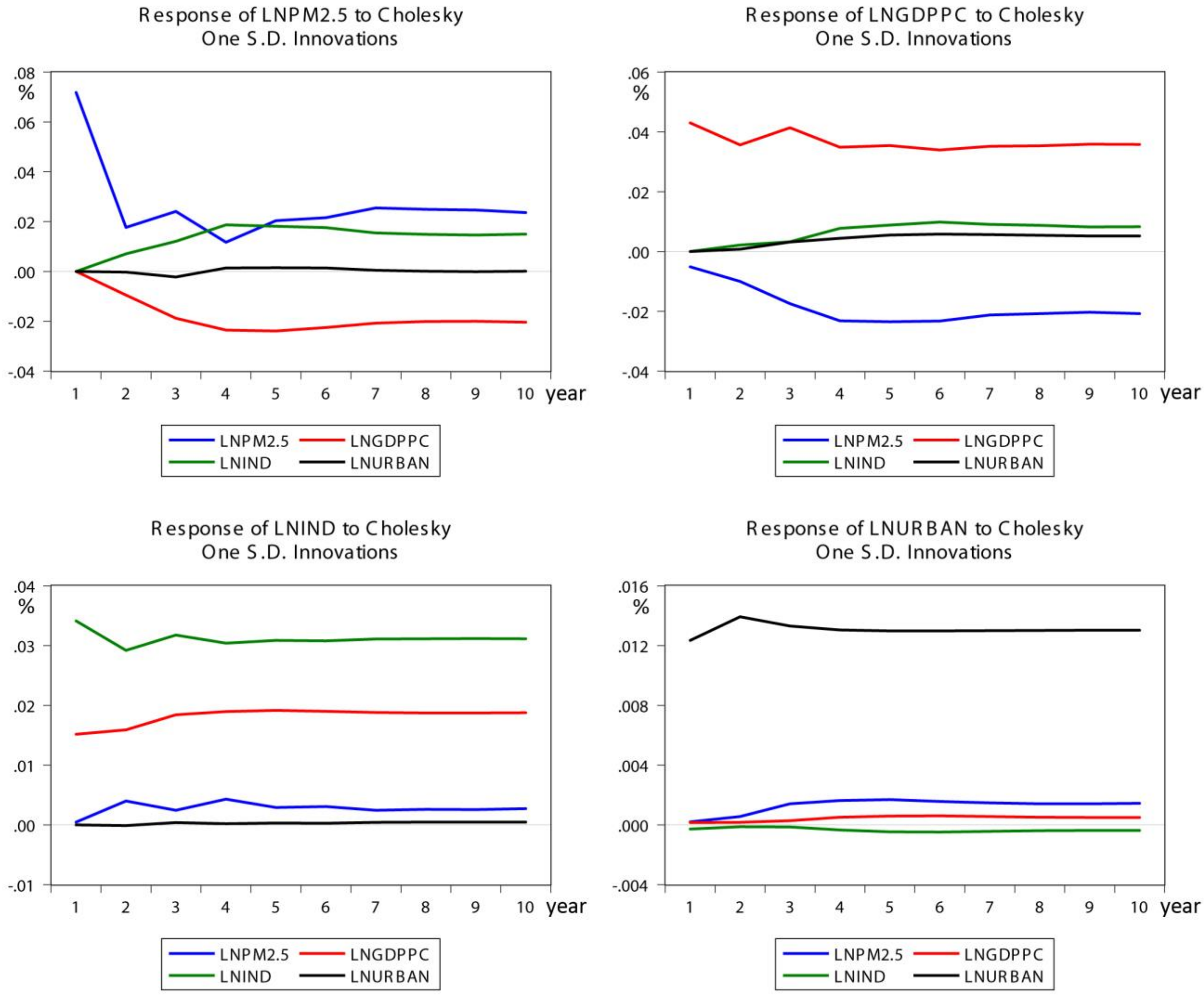

Figure S8. Responses of variables to one S.D. innovations for panel I1.

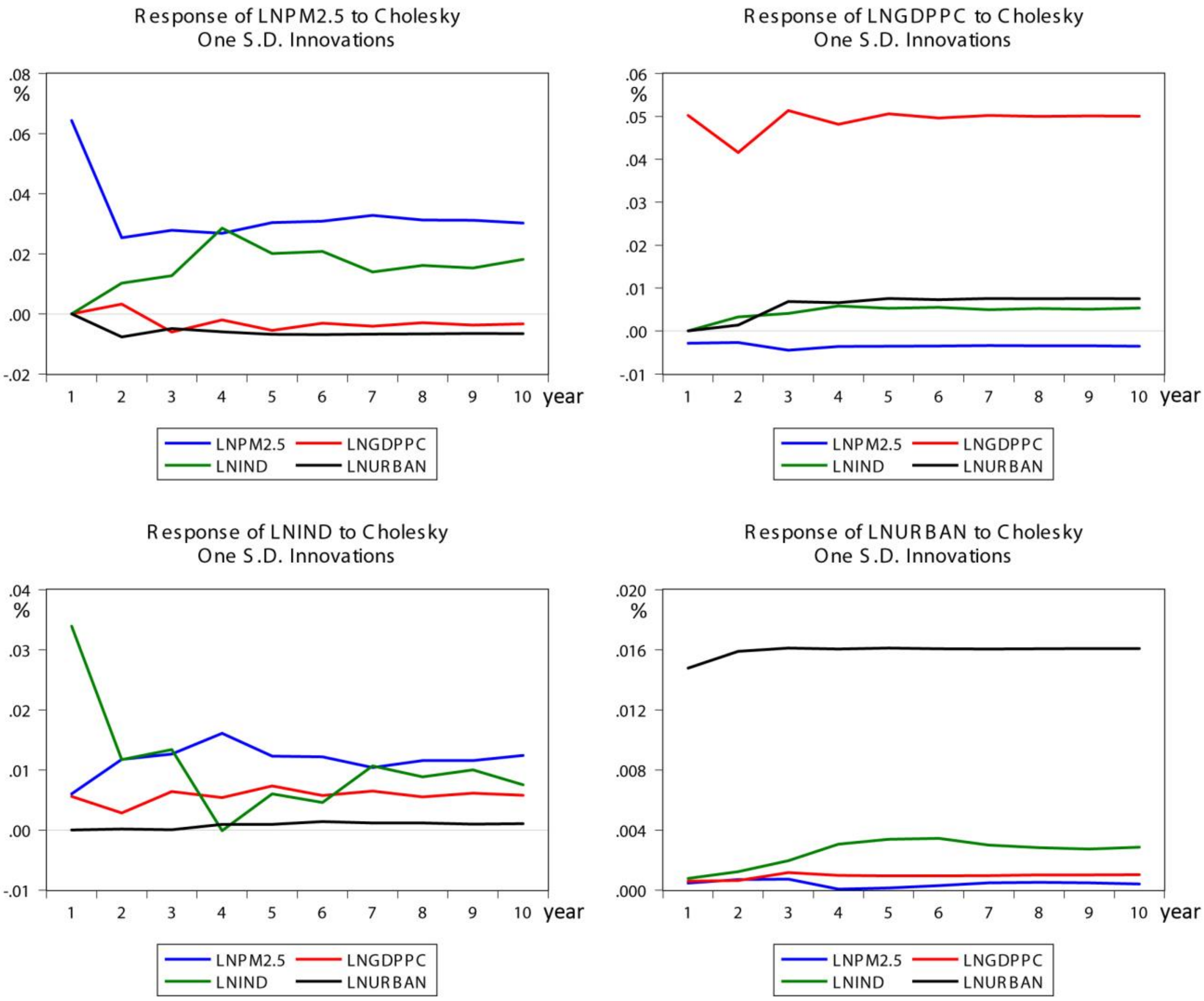

Figure S9. Responses of variables to one S.D. innovations for panel I2.

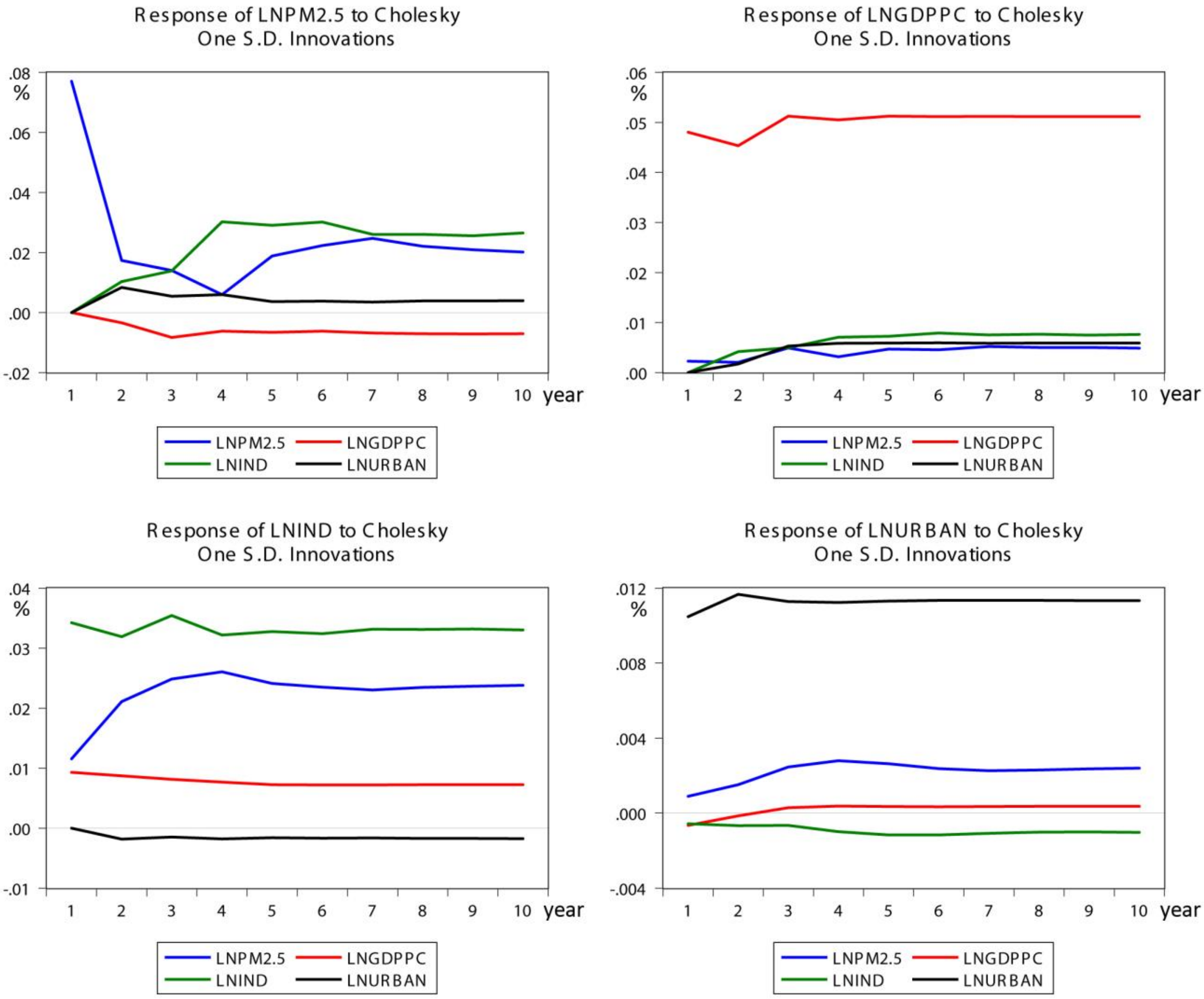

Figure S10. Responses of variables to one S.D. innovations for panel I3.

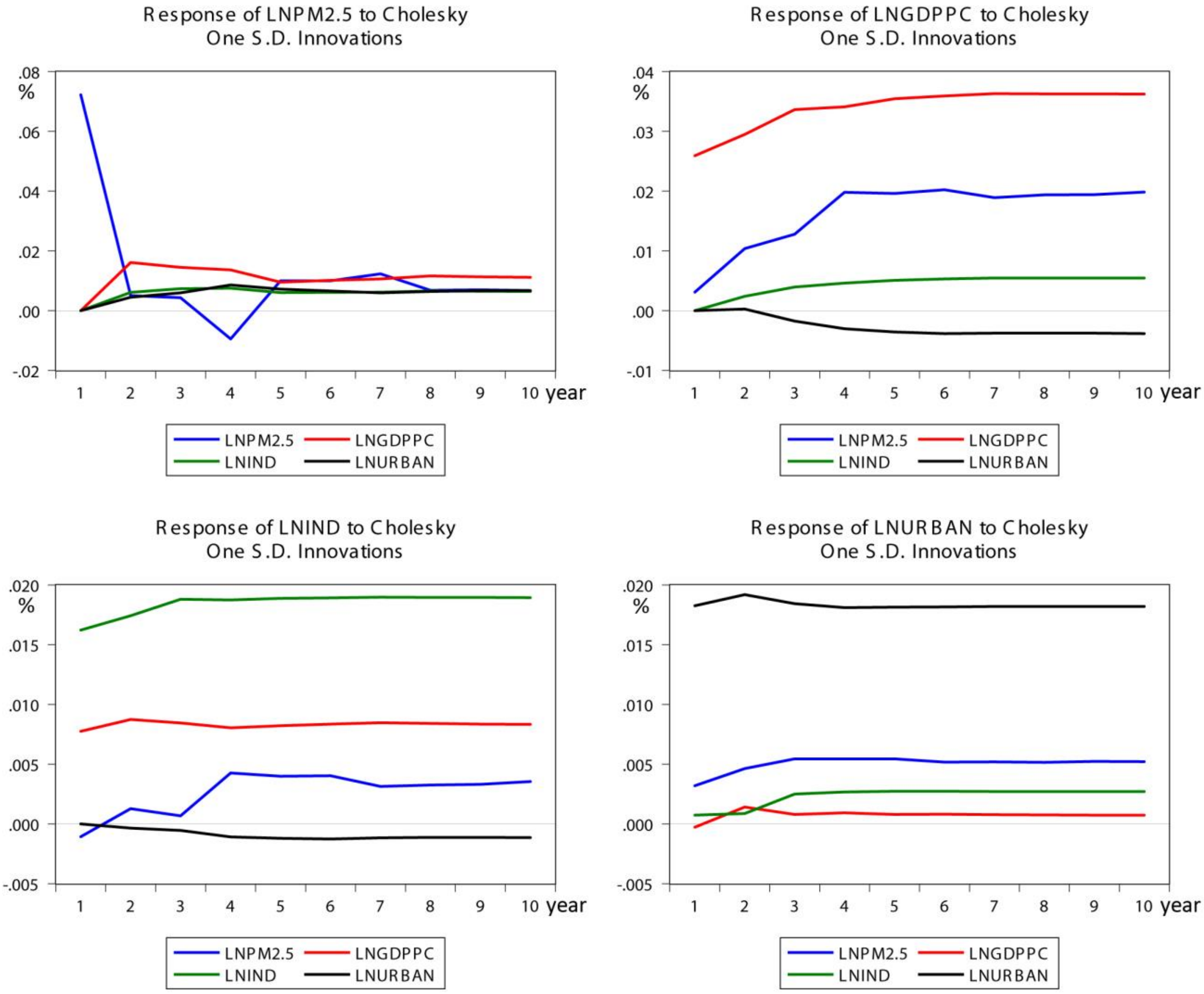

Figure S11. Responses of variables to one S.D. innovations for panel H.

References:

1 IHS Global Inc. EViews 9 User's Guide II. 2015, p558-563, www.eviews.com.

2 Levin, A., Lin, C. F. & Chu, C. S. J. Unit root tests in panel data: asymptotic and finite-sample properties. *J Econometrics* **108**, 1-24, doi:10.1016/s0304-4076(01)00098-7 (2002).

3 Im, K. S., Pesaran, M. H. & Shin, Y. Testing for unit roots in heterogeneous panels. *J Econometrics* **115**, 53-74, doi:10.1016/s0304-4076(03)00092-7 (2003).

4 Phillips, P. C. & Perron, P. Testing for a unit root in time series regression. *Biometrika* **75**, 335-346 (1988).

5 Pedroni, P. Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. Economet Theor 20, 597-625 (2004).

6 Al-mulali U. Factors affecting CO_2 emission in the Middle East: A panel data analysis. Energy 2012; 4(1):564–9.

7 Pedroni P. Panel cointegration: asymptotic and finite sample properties of pooled time series tests, with an application to the PPP hypothesis: new results. Working paper, Indiana University, 1999.

8 Li F, Dong SC, Li X, Liang QX, Yang WZ. Energy consumption-economic growth relationship and carbon dioxide emissions in China. Energy Policy 2011; 39(2):568-74.
9 Pedroni P. Fully modified OLS for heterogeneous cointegrated panels. Advanced Econometrics 2000; 15: 93–130.
10 Al-Mulali, U.; Ozturk, I., The effect of energy consumption, urbanization, trade openness, industrial output, and the political stability on the environmental degradation in the MENA (Middle East and North African) region. Energy 2015, 84, 382-389.
11 Abrigo, M. R.; Love, I., Estimation of panel vector autoregression in Stata: A package of programs. *manuscript, Febr 2015 available on http://paneldataconference2015. ceu. hu/Program/Michael-Abrigo. pdf* **2015**.
12 Lütkepohl, H. New introduction to multiple time series analysis. Springer Science & Business Media, 2005.
13 Hamilton, J. D., Time Series Analysis. Princeton University Press. Princeton, NJ 1994.
14 Sims, C.A., Macroeconomics and reality. Econometrica, 1980, 48(1), 1-48.