SUPPLEMENTARY APPENDIX

The Impact of Macro-Prudential Policies on Chinese Housing Markets: A Panel VAR-X Approach

Yiming Lin*

November 8, 2020

*Department of Economics, North Carolina State University, e-mail: ylin23@ncsu.edu. The supplementary appendix is downloadable at https://yiminglin.wordpress.ncsu.edu/research. I extend my gratitude to my dissertation advisor Jim Nason for his encouragement and guidance. I appreciate Douglas Pearce, Giuseppe Fiori, and Xiaoyong Zheng for their comments and support. I also thank Hanming Fang for providing the Chinese house price indexes. I am also grateful to Bogdan Nikiforov of the Department of Agricultural and Resource Economics for providing computing support. Fabio Canova, Tao Zha, and colleagues at NC State provided helpful guidance and advice.

A China's Economic Regions

The panel VAR is estimated on a sample of seven economic regions. I follow the structure proposed by Li and Hou (2003). Their definition of eight economic regions of China are the northeast (NE), north coast (NC), east coast (EC), south coast (SC), middle reaches of the Yellow River (YEL), middle reaches of the Yangtze River (YNG), southwest and northwest regions. I take southwest and northwest regions as the west (WST) region.

Figure 14 displays a map of China's economic regions and provinces.

The economic regions are divided by their economic development conditions. Table 9 shows the differences of the regional GDP in levels in 2015, as well as their main industries.

Table 9: China's Economic Regions

Economic Regions		GDP in 2015 (CNY billion)**	Main Industries
NE		5781.58	Equipment manufacturing, agriculture
NC		13236.12	High-tech R&D
EC		13812.63	Manufacturing
SC		10249.51	High-end consumer goods, manufacturing, high-tech manufacturing
YEL		8562.20	Energy development, steel and nonferrous metals, dairy industry
YNG		9718.18	Agricultural products deep processing, auto industry
WST	**Southwest**	8669.52	Heavy chemical
	Northwest	2247.03	Agriculture

* This table reports the regional GDP in 2015 and the main industries of each region.
** Data source: National Bureau of Statistics of China.

B Data

The data series used to estimate the panel VAR are real GDP growth ($\Delta lnGDP$), real loan growth ($\Delta lnLoan$), real house price growth ($\Delta lnHP$), and a macro-prudential policy index change (MPP). Table 10 lists the raw data used for construction a new regional data set.

Figure 14: Map of China with Economic Regions

Map source: `http://www.51pptmoban.com/sucai/1834.html`

Note: This is a map of China with provinces and the seven economic regions.

Table 10: Original Data Description

Data	Level	Frequency	Seasonal Adjustment	Source
				$\Delta lnGDP$
GDP (year to date)	Provincial	Quarterly	No	The National Bureau of Statistics of China (NBSC); see http://data.stats.gov.cn/easyquery.htm?cn=E0102
Consumer price index: all items for China	National	Quarterly	No	The Federal Reserve Bank of St. Louis (FRED); see https://fred.stlouisfed.org/series/CHNCPIALLQINMEI
M1	National	Monthly	Yes	The Federal Reserve Bank of St. Louis; see https://fred.stlouisfed.org/series/MANMM101CNM189S
M2	National	Monthly	Yes	The Federal Reserve Bank of St. Louis; see https://fred.stlouisfed.org/series/MABMM201CNM189S
M3	National	Monthly	Yes	The Federal Reserve Bank of St. Louis; see https://fred.stlouisfed.org/series/MABMM301CNM189S
Consumer price index: all items for China	National	Monthly	No	The Federal Reserve Bank of St. Louis; see https://fred.stlouisfed.org/series/CHNCPIALLMINMEI
3-month or 90-day rates and yields growth rate(treasury yields)	National	Monthly	No	The Federal Reserve Bank of St. Louis; see https://fred.stlouisfed.org/series/IR3TTS01CNM156N
Ratio of exports to imports (extoim)	National	Monthly	Yes	The Federal Reserve Bank of St. Louis; see https://fred.stlouisfed.org/series/XTEITT01CNM156S
GDP: ratio to trend (output gap)	National	Monthly	Yes	The Federal Reserve Bank of St. Louis; see https://fred.stlouisfed.org/series/CHNLORSGPRTSTSAM
Purchasing managers' index (PMI)	National	Monthly	No	The National Bureau of Statistics of China; see http://data.stats.gov.cn/easyquery.htm?cn=A01&zb=A0B01&sj=201812
				$\Delta lnLoan$
Total loan	Provincial	Monthly	No	The CEIC Data China Premium Database; see https://www.ceicdata.com/en/products/china-economic-database?gclid=EAIaIQobChMIw4C7lbH13wIVR-DICh0MwweZEAAYASACEgKblfD_BwE
Consumer price index: all items for China	National	Monthly	No	The Federal Reserve Bank of St. Louis; see https://fred.stlouisfed.org/series/CHNCPIALLMINMEI
				$\Delta lnHP$
House price index	City	Monthly	No	Fang et al. (2015)
Floor space of residential building sold	Provincial	Yearly	NA	The National Bureau of Statistics of China; see http://data.stats.gov.cn/easyquery.htm?cn=E0103
Floor space of residential building sold	City	Yearly	NA	The National Bureau of Statistics of China; see http://data.stats.gov.cn/easyquery.htm?cn=E0105
				MPP
Macro-prudential policy index	National	Monthly	NA	Kuttner and Shim (2016) and Shim et.al (2013); see https://www.bis.org/publ/qtrpdf/r_qt1309i_appendix.xls.
Macro-prudential policy information	National	Monthly	NA	The People's Bank of China (PBC); see http://www.pbc.gov.cn/en/3688229/index.html

* This table lists all original data series and their descriptions that this paper used to construct the variables in the panel VAR. See section 3 for details of constructing data used in this paper.

Table 11: Unconditional Summary Statistics of Data

Region	Mean	Std. Dev.	Min	Max	ACF(1)
Annualized Monthly Real Regional GDP Growth					
SC	3.6689	9.5741	-17.6314	61.1569	-0.0620
EC	3.5932	8.4169	-20.2101	62.3281	-0.0105
NC	3.7879	7.5360	-14.0262	46.3597	0.0217
NE	3.8723	12.1025	-44.9617	54.0952	0.0038
YEL	4.4079	10.8833	-47.8408	40.6148	-0.1291
YNG	4.3660	5.6393	-10.5519	21.2731	-0.2851
WST	4.4121	5.8202	-12.2692	29.8053	0.1087
Annualized Monthly Real Regional Loan Growth					
SC	13.1914	12.6672	-10.0111	66.6924	0.3525
EC	13.5411	12.7628	-12.7848	70.9776	0.2174
NC	12.6686	12.2815	-11.9484	70.5350	0.3606
NE	10.8443	18.1566	-92.0116	75.8968	0.4736
YEL	13.0495	12.5163	-12.3339	71.8613	0.3307
YNG	14.1284	14.0785	-20.4195	83.8916	0.3155
WST	15.3928	11.7908	-7.7965	77.0989	0.3881
Annualized Monthly Real Regional House Price Growth					
SC	8.5730	12.8833	-23.6568	43.6519	0.6902
EC	8.4365	12.1879	-29.9459	48.2168	0.6057
NC	9.1216	10.4930	-17.6652	33.6659	0.4335
NE	7.6606	14.7828	-55.8072	44.8842	0.2020
YEL	8.6938	13.2054	-31.1343	49.5390	0.1704
YNG	9.4990	10.0180	-15.4383	33.0771	0.6433
WST	9.8932	11.5576	-18.3001	45.0076	0.6192
12-Month Average Macro-Prudential Policy Index					
	-0.4328	0.5515	0.5000	1.5000	0.9737
12-Month Average Maximum Loan-to-Value Ratio Change					
	-0.3681	0.6503	-2.0833	0.5000	0.9364

* This table summarizes unconditional statistics of annualized monthly real GDP growth, real loan growth, real house price growth, and policy indexes constructed in this paper for each of the seven regions from 2005m02 to 2013m03.

B.1 Real GDP Growth ($\Delta lnGDP$)

The following steps generate the annualized monthly real regional GDP growth. First, I obtain provincial quarterly accumulated (year-to-date) nominal GDP from the National Bureau of Statistics of China (NBSC), and unwind the accumulated data to obtain flow data. Then I aggregate provincial flow GDP into seven regional quarterly nominal GDP series (see Figure 14 in appendix A for regional division). I deflate the seven nominal output series by using a not seasonally adjusted quarterly consumer price index (CPI): all items for China. Quarterly real regional GDP series are then deseasonalized using X-12-ARIMA[1]. Then temporal disaggregation methods are applied to obtain monthly data. The methods in used are Chow-Lin (CL), Chow-Lin AR(1) (CLAR1) and Santos Silva-Cardoso (SSC). Appendix C describes the disaggregation methods. Finally, I compute the monthly growth rates using the disaggregated data and multiply by 1200 to produce annualized monthly seasonally adjusted, real, regional GDP growth rates.

The construction of regional annualized monthly real GDP growth is summarized as follows.

1. Obtain accumulated quarterly provincial (year-to-date) nominal GDP. The series is not seasonally adjusted. Next, I unwind the accumulated provincial nominal GDP data to create flows of quarterly provincial nominal GDP. The flows of quarterly provincial nominal GDPs evolve from the accumulated data as

$$GDP_t = GDP_t^A - GDP_{t-1}^A, \tag{B.1}$$

   where GDP_t stands for GDP quarter t, and GDP_t^A is accumulated GDP at quarter t. The initial condition is $GDP_1 = GDP_1^A$.

2. Sum provincial GDP_t up to produce regional GDP_{it}, where $i = 1, ..., 7$ represents the seven regions.

[1]X-12-ARIMA is a program that is officially used by the United States Census Bureau for seasonal adjustment.

Table 12: Candidate Regressor Combinations

Regressor I	M1, PMI, treasury yields, CPI, extoim, outputgap
Regressor II	M2, PMI, treasury yields, CPI, extoim, outputgap
Regressor III	M3, PMI, treasury yields, CPI, extoim, outputgap
Regressor IV	PMI, treasury yields, CPI, extoim, outputgap
Regressor V	PMI, treasury yields, extoim, outputgap

[*] This table lists candidate regressor combinations in temporal disaggregating of quarterly GDP to monthly GDP.

3. Deflate regional GDP_{it} by the quarterly consumer price index CPI_t, which includes all items for China. The CPI is not seasonally adjusted.

4. Deseasonalize regional real GDP_{it} using X-12-ARIMA.

5. Take the logarithms of real GDP_{it} to get $lnGDP_{it}$.

6. Use Chow-Lin (CL) estimator with a drift to estimate monthly regional real GDP.[2] Appendix C describes the CL estimator and another two estimators I tried.

   Regressors are needed in temporal disaggragation as indicated in appendix C. The regressors can be different for different regions. I select the following series as candidate regressors: logM1, logM2, logM3, CPI, 3-Month or 90-day Rates and Yields growth rate (treasury yields), ratio of exports to imports (extoim), GDP: ratio to trend (output gap), and purchasing managers' index (PMI). These regressors are seasonally adjusted. Candidate regressor combinations are listed in table 12. Regressor combinations used in each region and are estimation results are reported in table 13.

7. Compute the annualized monthly growth rate of $lnGDP_{it}$. The equation is

$$\Delta lnGDP_{it} = (lnGDP_{it} - lnGDP_{it-1}) * 1,200, \tag{B.2}$$

   where $\Delta lnGDP_{it}$ is annualized monthly growth rate of real GDP (deseasonalized) in the i-th region.

[2] I also tried Chow-Lin AR(1) (CLAR1) and Santos Silva-Cardoso (SSC) estimators. CL estimators gives the highest correlations between the quarterly data and the estimated monthly data.

Table 13: Regional Monthly $lnGDP$

Region	Frequency**	Mean	Standard Deviation	Min	Max	Corr(y, $\hat{y}$)
NE Regressor I	**LF*** **	2.2596	0.1033	2.0911	2.4121	NA
	HF	2.2596	0.1023	2.0895	2.4181	0.9932
NC Regressor I	**LF**	2.5353	0.0925	2.3566	2.6661	NA
	HF	2.5353	0.0916	2.3531	2.6705	0.9964
EC Regressor I	**LF**	2.5465	0.0890	2.3737	2.6698	NA
	HF	2.5465	0.0881	2.3709	2.6726	0.9958
SC Regressor I	**LF**	2.4403	0.0899	2.2670	2.5732	NA
	HF	2.4403	0.0891	2.2634	2.5755	0.9945
YEL Regressor I	**LF**	2.3797	0.1115	2.1692	2.5300	NA
	HF	2.3797	0.1105	2.1648	2.5342	0.9960
YNG Regressor III	**LF**	2.3761	0.1167	2.1924	2.5504	NA
	HF	2.3761	0.1157	2.1897	2.5644	0.9989
WST Regressor II	**LF**	2.4141	0.1167	2.1924	2.5504	NA
	HF	2.4141	0.1101	2.2189	2.5847	0.9985

[*] This table displays the result and performance of temporal disaggregation of regional GDP by using CL method with a drift for each region, with showing the regressor combinations used. See table 12 for candidate regressor combinations.
[**] Frequency includes low frequency (LF): quarterly, and high frequency (HF): monthly.
[***] All quarterly statistics are averaged out to monthly (1/3 of the quarterly).

B.2 Real Loan Growth ($\Delta lnLoan$)

The raw data of constructing annualized monthly real regional loan growth ($\Delta lnLoan$) is monthly provincial total loan (for all sectors, not seasonally adjusted) obtained from the CEIC Data China Premium Database. Total loans include short-term loans, medium and long-term loans, designated loans, bill financing, and other loans of all financial institutions in China (i.e., the People's Bank of China, deposit financial institutions, trust and investment corporations, financial leasing corporations, and automobile financial corporations). I deflate the raw data using not seasonally adjusted monthly CPI, and deseasonalize using X-12-ARIMA. Then I aggregate these provincial-level data to regional-level, and compute annualized monthly real regional loan growth rate.

The construction of annualized monthly regional real loan growth is summarized as follows.

1. Obtain provincial monthly total loan (not seasonally adjusted).
2. Deflate provincial monthly total loan by national monthly CPI: all items for China (not seasonally adjusted).
3. Deseasonalize provincial monthly real loan by X-12-ARIMA.
4. Compute regional monthly real loan by summing provincial data up by groups of region.
5. Take logarithms of regional monthly real loan. Compute regional annualized monthly real loan growth rate,

$$\Delta lnLoan_{it} = (lnLoan_{it} - lnLoan_{it-1}) * 1,200, \tag{B.3}$$

   where t denotes month, and i denotes the i-th region.

B.3 Real House Price Growth ($\Delta lnHP$)

The construction of regional annualized monthly real house price growth is summarized as below.

1. Obtain 120 city-level monthly house price indexes computed by Fang et al. (2016). It is common to group Chinese cities into three tiers. The 120 cities in Fang et al. (2016) includes all first and second tier cities, and some third tier cities. I pick the 98 cities with complete series. Table 14 lists the provinces in each region, and cities that have complete series of house price index. All the first and second tier cities are complete in the sample period.

2. Compute monthly regional house price indexes by taking weighted average of monthly provincial-level house price indexes. The provincial-level indexes is the weighted average of city-level indexes.

    (a) Calculate weights of cities and provinces. The weights computed can be find in table 15 and 16.

        i. Annually provincial and city-level floor space of residential building sold (sq) are used to compute the weights. Due to availabilities, the data are annually and only includes provincial-level, first and second tier cities, but not third tier cities.

        The weights of the three tiers of cities in each region are

$$weight_{i1} = sq_{i1}/sq_i, \tag{B.4}$$

$$weight_{i2} = sq_{i2}/sq_i, \tag{B.5}$$

$$weight_{i3} = 1 - weight_{i1} - weight_{i2}. \tag{B.6}$$

        where $weight_{ih}$ denotes the weight of the hth tier city in the ith region, sq_i is the sq of the i-th region, and sq_{ih} denotes the sum of sq of the hth tier city in the i-th region.

Table 14: List of Economics Regions

Economic Regions		Provinces	Cities
NE		Helongjiang	Harbin (2)*
		Jilin	Changchun (2), Songyuan
		Liaoning	Shenyang (2), Dalian (2), Anshan, Dandong, Yingkou, Tieling, Huludao
NC		Beijing (M)**	Beijing (1)
		Tianjin (M)	Tianjin (2)
		Hebei	Shijiazhuang (2), Tangshan, Qinhuangdao, Xingtai, Baoding, Zhangjiakou, Langfang
		Shandong	Jinan (2), Qingdao (2), Zaozhuang, Rizhao
EC		Shanghai (M)	Shanghai (1)
		Jiangsu	Nanjing (2), Wuxi, Xuzhou, Changzhou, Suzhou, Nantong, Lianyungang, Huai'an, Yancheng, Yangzhou, Zhenjiang, Suqian
		Zhejiang	Hangzhou (2), Ningbo (2), Wenzhou, Jiaxing, Huzhou, Shaoxing, Jinhua, Taizhou
SC		Fujian	Fuzhou (2), Xiamen (2), Quanzhou, Zhangzhou, Ningde
		Guangdong	Guangzhou (1), Shenzhen (1), Shantou, Foshan, Jiangmen, Zhaoqing, Huizhou, Shanwei, Heyuan, Qingyuan, Dongguan, Zhongshan, Jieyang
		Hainan	Haikou (2)
YEL		Shaanxi	Xi'an (2)
		Shanxi	
		Henan	Zhengzhou (2), Kaifeng, Luoyang, Xinxiang, Xuchang, Luohe, Nanyang, Zhumadian
		Inner Mongolia	Hohhot (2), Baotou
YNG		Hubei	
		Hunan	Changsha (2)
		Jiangxi	Nanchang (2), Jingdezhen, Jiujiang, Yichun, Shangrao, Fuzhou
		Anhui	Hefei (2), Wuhu, Bengbu, Anqing, Chuzhou, Xuancheng
WST	Southwest	Yunnan	Kunming (2)
		Sichuan	Chengdu (2), Luzhou, Deyang, Mianyang, Leshan, Nanchong
		Chongqing (M)	Chongqing (2)
		Guizhou	
		Guangxi	Nanning (2)
	Northwest	Gansu	
		Xinjiang	Urumq (2)
		Qinghai	Xining (2)
		Ningxia	
		Tibet	

* Number after city shows tier of the city. Cities without number are third tier cities. The first tier cities are the most economically developed cities and with the largest population. The second tier cities include 36 cities that are selected by the National Bureau of Statistics of China due to their importance economically, and exclude 4 first tier cities, and cities with incomplete series such as Yinchuan, Lanzhou, Lhasa, Guiyang, Wuhan and Taiyuan. All other cities in the sample are classified as third tier cities.

** Label with M means it is a provincial-level municipality.

Table 15: Weights of Each Tier of City in Coastal Regions

Year	SC_1 **	SC_2	SC_3	EC_1	EC_2	EC_3	NC_1	NC_2	NC_3
2005	0.33	0.19	0.48	0.28	0.19	0.5	0.32	0.29	0.39 3
2006	0.28	0.16	0.56	0.24	0.19	0.57	0.24	0.28	0.48
2007	0.21	0.14	0.65	0.23	0.20	0.57	0.18	0.29	0.53
2008	0.22	0.11	0.67	0.21	0.17	0.62	0.11	0.27	0.62
2009	0.21	0.13	0.66	0.18	0.18	0.64	0.15	0.26	0.59
2010	0.16	0.10	0.74	0.12	0.15	0.73	0.08	0.22	0.70
2011	0.15	0.10	0.75	0.14	0.15	0.72	0.06	0.22	0.72
2012	0.15	0.14	0.71	0.12	0.18	0.70	0.10	0.24	0.67
2013	0.14	0.14	0.72	0.12	0.17	0.71	0.08	0.24	0.68

*

* This tables shows the weights of tiers of cities and provinces in calculating monthly regional house price indexes in coastal regions.

** The subscripts label tiers of city.

Table 16: Weights of Each Tier of City in Interior Regions

Year	NE_2 **	NE_3	YEL_2	YEL_3	YNG_2	YNG_3	WST_2	WST_3
2005	0.55	0.45	0.30	0.70	0.24	0.76	0.45	0.55
2006	0.54	0.46	0.29	0.71	0.22	0.78	0.47	0.53
2007	0.53	0.47	0.26	0.74	0.24	0.76	0.50	0.50
2008	0.48	0.52	0.23	0.77	0.24	0.76	0.45	0.55
2009	0.45	0.55	0.28	0.72	0.26	0.74	0.47	0.53
2010	0.40	0.60	0.30	0.70	0.22	0.78	0.43	0.57
2011	0.38	0.62	0.27	0.73	0.20	0.80	0.42	0.58
2012	0.38	0.62	0.26	0.74	0.21	0.79	0.42	0.58
2013	0.39	0.61	0.24	0.76	0.21	0.79	0.40	0.60

*

* This tables shows the weights of tiers of cities and provinces in calculating monthly regional house price indexes in interior regions.

** The subscripts label tiers of city.

The weights showed a declining trend in all regions for first and second tier cities, and a inclining trend for third tier cities.

(b) Compute city-tier-level house price indexes in each region, which is denoted as HP_{ih}, by taking averages of city-level house price indexes in the hth tier in region i.

(c) Regional house price indexes, HP^i, is computed by taking weighted averages of HP_{ih}. It can be written as the following equation,

$$HP_i = \left(\sum_{h=1}^{3} (HP_{ih} * weight_{ih}) \right) * 100. \tag{B.7}$$

3. Deflate monthly regional house price indexes by national monthly CPI: all items for China (not seasonally adjusted).

4. Deseasonalize monthly real house price by X-12-ARIMA.

5. Take log-difference of HP_i, and multiply by 1,200 to compute the annualized monthly real regional house price growth.

$$\Delta lnHP_{it} = (lnHP_{it} - lnHP_{it-1}) * 1,200, \tag{B.8}$$

where t indicate month, and $\Delta lnHP_{it}$ is annualized monthly real house price growth rate in the tth month in region i (seasonally adjusted).

B.4 Policy Indexes

The policy indexes are macro-prudential policy index (MPP) and maximum loan-to-value ratio change (ΔLTV). The raw data of MPP comes from Shim et al. (2013), and the raw data of ΔLTV comes from Alam et al. (2019). I take the 12-month moving averages of the change of the raw data.

C Disaggregation Methods

The frequency of China's provincial GDP provided by the NBSC is quarterly. The problem of constructing monthly GDP from January 2005 to March 2013, given quarterly GDP, are dealt with Chow-Lin, Chow-Lin AR(1) and Santos Silva-Cardoso approaches. The approaches are processes by a Matlab program *Temporal Disaggregation* developed by Quilis (2019). This section briefly introduces the application of the 3 approaches.

C.1 Chow-Lin (CL) Method

The Chow-Lin method is used to temporally disaggregate data from a lower to a higher frequency. Chow and Lin (1971) use a multiple linear regression model to estimate high-frequency dependent variables from low-frequency data, using related high-frequency series as regressors.

The monthly series to be estimated is written in a regression model

$$y = X\beta + u, \tag{C.1}$$

where y is $3n \times 1$ given n observations of the quarterly series y^*, X is a $3n \times p$ explanatory matrix composed of p variables x_1, x_2,..., x_p, and u is a vector of residuals with covariance matrix V. An intercept is tested. If there an intercept is included, $X = [1_{3n\times1}, x_1, x_2, ..., x_p]$.

Converting quarterly GDP to monthly GDP is a distribution problem, which means the sum of the estimated monthly GDP should equal the observed quarterly GDP. For distribution problems, define the $n \times 3n$ aggregation matrix as

$$C = \begin{bmatrix} 1_{3\times1} & 0_{3\times1} & \cdots & 0_{3\times1} \\ 0_{3\times1} & 1_{3\times1} & \cdots & 0_{3\times1} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{3\times1} & 0_{3\times1} & \cdots & 1_{3\times1} \end{bmatrix}, \tag{C.2}$$

then quarterly series y^* can be expressed as

$$y^* = Cy = CX\beta + Cu, \tag{C.3}$$

where Cu is the regression residuals of the quarterly series with covariance matrix CVC'.

Estimates of the parameters in equation (C.1) can be obtained from the regression of quarterly series, as shown in equation (C.3). The best linear estimate of the regression coefficients $\hat{\beta}$ is obtained by generalized least squares (GLS),

$$\hat{\beta} = [(CX)'(CVC')^{-1}CX]^{-1}(CX)'(CVC')^{-1}y^*. \tag{C.4}$$

and the best linear unbiased estimator $\hat{y}$ is

$$\hat{y} = X\hat{\beta} + VC'(CVC')^{-1}(y^* - CX\hat{\beta}), \tag{C.5}$$

The covariance matrix V of the residuals in equation (C.1) needs to be estimated. The Chow-Lin method assumes that the residuals are homoskedastic and no correlations, so $V = I_{3n}\sigma^2$, where σ^2 is the variance of each residual. This structure distributes the quarterly residuals uniformly to each month. However, this method may lead to discontinuities in monthly series between contiguous quarters.

C.2 Chow-Lin AR(1) (CLAR1) Method

Chow-Lin AR(1) regression has the same structure as Chow-Lin method, the only difference is the assumption of the covariance matrix V. It estimates the covariance matrix V assuming the monthly residuals in equation (C.1) follow AR(1) process,

$$u_t = \rho u_{t-1} + \epsilon_t, \tag{C.6}$$

where $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. When $\rho = 0$, CLAR1 will be reduced to CL method as described in C.1; and when $\rho = 1$, as assumed by Fernandez (1981), the residuals follows a random walk. In the practice of this paper, the optimal ρ is obtained by weighted least squares estimator.

C.3 Santos Silva-Cardoso (SSC) Method

An extension to the Chow-Lin method proposed by Silva and Cardoso (2001) introduces dynamic feature into the regression model by adding lagged dependent variables as regressors. The dynamic model can be written as

$$y_t = \zeta y_{t-1} + X_t\beta + u_t, \tag{C.7}$$

where X_t is a vector of dependent variables and the autoregression coefficients *zeta* should be between -1 and 1. u_t is assumed to be white noise, which means the residuals are not serially correlated.

To avoid conditioning on the initial observations, as shown by Salazer et al. (1994) and Gregoir (1995), Silva and Cardoso (2001) transform the model into a recursive form

$$y_t = \left(\sum_{k=0}^{t-1} \zeta^k X_{t-k}\right)\beta + \zeta^t y_0 + \left(\sum_{k=0}^{t-1} \zeta^k u_{t-k}\right), \tag{C.8}$$

thus we can find

$$y_0 = \left(\sum_{k=0}^{\infty} \zeta^k X_{-k}\right)\beta + \left(\sum_{k=0}^{\infty} \zeta^k u_{-k}\right). \tag{C.9}$$

From the equation above, define $\eta = E(y_0|X_0, X_{-1}, ...)$ as the truncation remainder, $(\sum_{i=0}^{\infty} \zeta^k X_{-k})\beta$, which gives initial condition for y. Given equation (C.9), equation (C.8) can rewritten as

$$y_t = \left(\sum_{k=0}^{t-1} \zeta^k X_{t-k}\right)\beta + \zeta^t \eta + \left(\sum_{k=0}^{\infty} \zeta^k u_{t-k}\right), \tag{C.10}$$

it can be expressed as

$$y_t = X(\zeta)_t\beta + \zeta^t\eta + \gamma_t, \tag{C.11}$$

where $X(\zeta)_t = \sum_{k=0}^{t-1} \zeta^k X_{t-k}$, which means the weighted sum of past and present values of regressors, and $\gamma_t = k\gamma_t + u_t$ is the AR(1) process error term. SSC further express it in vector form

$$y = \chi(\zeta)\tilde{\beta} + \gamma. \tag{C.12}$$

Then estimators $\hat{\tilde{\beta}}$ and $\hat{y}$ can be obtained as in Chow-Lin method, see equation (C.4) and equation (C.5) in appendix C.1, where V is a diagonal matrix with its element $\omega_{ij} = \beta^{|i=j|}$.

D TVP Panel VAR

I employ a panel vector autoregression (PVAR) model with time-varying parameters (TVP) and stochastic volatility to model the interactions among $\Delta lnGDP$, $\Delta lnLoan$, $\Delta lnHP$ and national-level macro-prudential policy variable in seven economic regions. Estimating the panel VAR is a difficult computational problem. The problem results from the dimensionality of the panel VAR. The solution is to map the panel VAR into a state space model.

D.1 A TVP panel VAR

This subsection sets up reduced-form VARs for each region, and combines them into a reduced-form panel VAR.

D.1.1 The Panel VAR Framework

Let y_{it} denotes a vector of monthly data of real GDP growth, $\Delta lnGDP_{it}$, real loan growth, $\Delta lnLoan_{it}$, and real house price growth, $\Delta lnHP_{it}$ in economic region of China i, $i = 1, ..., 7$, at time t. The regions are the NE, NC, EC, SC, YEL, YNG and WST, as introduced in appendix B. The data runs from 2005m02 to 2013m03, $T = 97$ and $N = 7$. The reduced-form vector autoregression model, with time-varying parameters

A_{it} and C_{it}, for each region can be written as:

$$y_{it} = Intcpt_{it} + A_{it}Y_{t-1} + C_{it}D_{t-1} + e_{it}, \tag{D.1}$$

where $Y_t = (y_{1t}^{'}, ..., y_{7t}^{'})^{'}$ stacks the vectors of the three endogenous variables of the seven regions, D is a vector of macro-prudential policy indexes or maximum loan-to-value ratio change common to all regions, A_{it} is a 3×21 coefficient matrix, C_{it} is a 3×1 matrix, e_{it} is a 3×1 vector of random disturbances for each region, and $e_t \sim N(0, \Sigma_{iit})$ where Σ_{iit} denotes the variance-covariance matrix of the errors of y_{it}. The VAR of equation (D.1) shows that y_{it} depends on lagged variables of every region.

The panel VAR below stacks each of the $i = 1, ..., 7$ VARs of equation (D.1) to form the reduced form panel VAR,

$$Y_t = Intcpt_t + A_tY_{t-1} + C_tD_{t-1} + e_t, \tag{D.2}$$

where $e_t = (e_{1t}^{'}, ..., e_{7t}^{'})' \sim N(0_{21\times1}, \Sigma_{t(21\times21)})$ and Σ_t is the time-varying covariance matrix of the errors. Denote the covariance matrix between the errors in the VARs of any 2 regions as $\Sigma_{ijt} = cov_t(e_{it}, e_{jt})$, and $\Sigma_{ijt(3\times3)}$ is ith-jth block of Σ_t.

D.1.2 Features of Panel VAR: An Example with 2 regions

The panel VAR framework of equation (D.2) allows for dynamic interdependencies, static interdependencies, cross section heterogeneities, and dynamic heterogeneity.

Consider a panel VAR with only 2 regions. The data vectors of region 1 and 2 are denoted y_{1t} and y_{2t}.

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} Intcpt_{1t} \\ Intcpt_{2t} \end{bmatrix} + \begin{bmatrix} A_{11,t} & A_{12,t} \\ A_{21,t} & A_{22,t} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \tag{D.3}$$

where $\left[\begin{smallmatrix} e_{1t} \\ e_{2t} \end{smallmatrix}\right] \sim N\left(0, \left[\begin{smallmatrix} \Sigma_{11,t} & \Sigma_{12,t} \\ \Sigma_{21,t} & \Sigma_{22,t} \end{smallmatrix}\right]\right)$.

Dynamic Interdependency

A dynamic interdependency between region i to region j occurs when $A_{ijt} \neq 0$. For example, if $A_{21t} \neq 0$, there are dynamic interdependencies between these regions.

Static Interdependency

Static interdependencies depend on $\Sigma_{ijt} \neq 0$. If $\Sigma_{12t} \neq 0$, there are static interdependancies between these regions, which means the errors are correlated across regions in the same period.

Cross Section Heterogeneity

Cross section heterogeneities exist when the coefficients differ across regions, $A_{iit} \neq A_{jjt}$. If $A_{iit} = A_{jjt}$, the region responses are homogeneous. When $A_{11t} \neq A_{22t}$, there is cross sectional heterogeneity.

Dynamic Heterogeneity

Dynamic heterogeneities occurs when the coefficients and errors vary across time. If dynamic heterogeneity holds in region 1, $A_{11t} \neq A_{11s}$ and $\Sigma_{11t} \neq \Sigma_{11s}$, where date $s \neq t$.

D.1.3 Dimensionality of Panel VAR

The dimension of the panel VAR is large. There are 21 dependent variables in the panel VAR. Without restrictions, a total of 714 parameters need to be estimated at every date t. The panel VAR spreads out the over 700 parameters across $441(= 21 \times 21)$ impact parameters, $21(= 3 \times 7)$ parameters on the exogenous variable, $21(= 3 \times 7)$ parameters on the intercept, and $231(= \frac{21\times21+21}{2})$ covariance parameters.

D.2 State Space Structure

The large dimension makes it very challenging computationally. To reduce the dimensionality of the problem, Canova and Ciccarelli (2009) propose a factor structure of the coefficients. The factor structure transforms the reduced-form panel VAR into a state space structural panel VAR. The structural panel VAR is estimated by state space methods.

Factorizing the impact parameters of the panel VAR reduces the dimensionality, and the factors can be explained economically. The economic interpretation is described below.

Dynamic factor models have state space representation. Bayesian MCMC algorithms exist to estimate the panel VAR in state space form. I use a Metropolis within Gibbs algorithm developed by Dieppe, Legrand, and van Roye (2016). The algorithm is described in appendix E.

D.2.1 Stacked-Form Panel VAR

For factorizing the impacts parameters, a stacked-form coefficient matrix is needed.

First, construct each region's stacked-form VAR from equation (D.1). The matrices A_{it} and C_{it} are mapped into $(23 \times 3) \times 1$ vector $\delta_{it} = (vec(A_{it})', Intcpt_t', vec(C_{it})')$. Stack independent variables Y_{t-1} and D_{t-1} into a 23×1 vector $X_t = (Y_{t-1}', I', D_{t-1}')'$. Then the stacked-form VAR for each region is:

$$y_{it} = (I_3 \otimes X_t')\delta_{it} + e_{it} \tag{D.4}$$

Next, stack the region VARs in equation (D.4) into a panel VAR. Define $\delta_t = (\delta_{1t}', ..., \delta_{7t}')'$ as the integrated impact matrix. The stacked-form panel VAR is:

$$Y_t = W_t \delta_t + e_t, \tag{D.5}$$

where $W_t = I_{21} \otimes X_t'$ is a $21 \times (21 \times 23)$ matrix. The error term e_t is defined in equation (D.2).

Illustration in Matrix Form: An Example with 2 Regions

The equation above can be expanded into matrix form. Consider an example with only 2 regions, $Y_t = (y_{1t}', y_{2t}')'$ where $y_{it} = (y_{i1t}, y_{i2t})'$ is a vector of only 2 variables, and there is 1 predetermined variable.

$$\begin{bmatrix} y_{11t} \\ y_{12t} \\ y_{21t} \\ y_{22t} \end{bmatrix}_{4\times 1} = \begin{bmatrix} [X_t']_{1\times 5} & O_{1\times 5} & O_{1\times 5} & O_{1\times 5} \\ O_{1\times 5} & [X_t']_{1\times 5} & O_{1\times 5} & O_{1\times 5} \\ O_{1\times 5} & O_{1\times 5} & [X_t']_{1\times 5} & O_{1\times 5} \\ O_{1\times 5} & O_{1\times 5} & O_{1\times} & [X_t']_{1\times 5} \end{bmatrix}_{4\times(4\times 5)} \begin{bmatrix} \begin{bmatrix} A_{11,11t} \\ A_{11,12t} \\ A_{11,21t} \\ A_{11,22t} \\ C_{11t} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} A_{22,11t} \\ A_{22,12t} \\ A_{22,21t} \\ A_{22,22t} \\ C_{22t} \end{bmatrix} \end{bmatrix}_{(4\times 5)\times 1} + e_t \quad \text{(D.6)}$$

where the first two subscripts of As are the indexes of dependent region and variable,and the latter two are the indexes of independent region and variable. The vector of the impact parameters is δ_t.

The above matrix form equation consists of 4 equations. For example, the first line is the equation for y_{11t}, the first variable in the first region,

$$y_{11t} = \begin{bmatrix} y_{11,t-1} \\ y_{12,t-1} \\ y_{21,t-1} \\ y_{22,t-1} \\ D_{t-1} \end{bmatrix}' \begin{bmatrix} A_{11,11t} \\ A_{11,12t} \\ A_{11,21t} \\ A_{11,22t} \\ C_{11t} \end{bmatrix} + e_{1t}. \tag{D.7}$$

D.2.2 Factorization of the Impact Matrix

The stacked-form impact matrix δ_t, which is defined in equation (D.4) and equation (D.5), is mapped into a linear combination of structural factors,

$$\delta_t = \Xi_1\theta_{1t} + \Xi_2\theta_{2t} + \Xi_3\theta_{3t} + \Xi_4\theta_{4t}, \tag{D.8}$$

where θ_{1t} is a scalar captures impacts that commonly come from all variables in all regions in δ_t, θ_{2t} is a 7×1 vector of factors that capture impacts come from all variables each region, θ_{3t} is a 3×1 vector of factors that capture impacts come from a specific variable in all regions, and θ_{4t} is a 2×1 vector of factor that capture intercept and exogenous variable's impacts. The factors follow stochastic processes and their law of motion are described in appendix D.2.3.

Conformable matrices $\Xi_1, ..., \Xi_4$ are predetermined loadings on the factors with blocks of ones and zeros relevant to $\theta_1, ..., \theta_4$. These selector matrices are defined as

$$\Xi_1 \equiv \begin{bmatrix} i_1 \\ \vdots \\ i_1 \end{bmatrix}_{(21\times23)\times1}, \tag{D.9}$$

where $i_1 = ((1, ..., 1)_{1\times23})'$. Ξ_1 labels the place of the factor commonly comes from all regions and variables.

$$\Xi_2 \equiv \begin{bmatrix} i_2 & 0 & \cdots & 0 \\ i_2 & 0 & \cdots & 0 \\ i_2 & 0 & \cdots & 0 \\ 0 & i_3 & \cdots & 0 \\ 0 & i_3 & \cdots & 0 \\ 0 & i_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & i_8 \\ 0 & 0 & \cdots & i_8 \\ 0 & 0 & \cdots & i_8 \end{bmatrix}_{(21\times23)\times7}, \tag{D.10}$$

where i_k is a 23×1 matrix, $k = 2, ..., 8$, with the positions corresponding to the dependent variables of the $(k-1)$th region are ones, and others are zeros, e.g., $i_2 = ((1, 1, 1, 0, ..., 0)_{1\times23})'$. The $(k-1)$th column of Ξ_2 labels the place of the factor comes from variables in the $(k-1)$th region.

$$\Xi_3 \equiv \begin{bmatrix} i_9 & 0 & 0 \\ 0 & i_{10} & 0 \\ 0 & 0 & i_{11} \\ i_9 & 0 & 0 \\ 0 & i_{10} & 0 \\ 0 & 0 & i_{11} \\ \vdots & \ddots & \vdots \\ i_9 & 0 & 0 \\ 0 & i_{10} & 0 \\ 0 & 0 & i_{11} \end{bmatrix}_{(21\times23)\times3}, \tag{D.11}$$

where i_k is a 23×1 matrix, $k = 9, ..., 11$, with the positions corresponding to the $(k-8)$th dependent variable in every unit are ones, and others are zeros, e.g., $i_9 = ((1, 0, 0, 1, 0, 0, ..., 1, 0, 0, 0)_{1\times23})'$. The $(k-8)$th column of Ξ_3 labels the place of the factor comes from the $(k-8)$th variable in all regions.

$$\Xi_4 \equiv \begin{bmatrix} i_{12} & i_{13} \\ \vdots & \vdots \\ i_{12} & i_{13} \end{bmatrix}_{(21\times23)\times1}, \tag{D.12}$$

where $i_k = ((0, ...0, 1)_{1\times23})'$ for $k = 12$. The second column of Ξ_4 labels the place of the factor comes from the exogenous variable.

Then rewrite equation (D.8) in compact form:

$$\delta_t = \Xi\theta_t, \tag{D.13}$$

Take equation (D.13) back to equation (D.5) to get the observation equation of the state space model,

$$\begin{aligned} Y_t &= W_t \left(\sum_i^4 \Xi_i \theta_{it} \right) + e_t \\ &= W_t \Xi \theta_t + e_t \\ &= \chi_t \theta_t + e_t, \end{aligned} \tag{D.14}$$

where $\theta_t \equiv (\theta_{1t}^{'}, \theta_{2t}^{'}, \theta_{3t}^{'}, \theta_{4t}^{'})^{'}$, $\Xi \equiv (\Xi_1, \Xi_2, \Xi_3, \Xi_4)$, and $\chi \equiv W_t\Xi$.

Take the second line of Y_t into factors as an example:

$$\begin{aligned} y_{12t} = {} & (y_{11t-1} + y_{12t-1} + ... + y_{73t-1})\theta_{1t} \\ & + (y_{11t-1} + y_{12t-1} + y_{13t-1})\theta_{21t} \\ & + (y_{12t-1} + y_{22t-1} + y_{32t-1} + y_{42t-1} + y_{52t-1} + y_{62t-1} + y_{72t-1})\theta_{32t} \\ & + \theta_{41t} + (D_{t-1})\theta_{42t} + e_{12t}, \end{aligned} \tag{D.15}$$

The equation shows that the second variable in the first region depends on (i) the common component, (ii) the component common to region 1, (iii) the component common to variable 2, and (iv) the component comes from the intercept and the exogenous variable.

D.2.3 Law of Motion of the Factors

I assume the law of motion of the vector of factors θ_t, which is defined in equation (D.14), is in form of autoregressive process,

$$\theta_t = (1-\rho)\bar{\theta} + \rho\theta_{t-1} + \eta_t, \tag{D.16}$$

where $\bar{\theta}$ is the population mean of θ, and ρ determines the persistence of the factors. The error term $\eta_t \sim N(0, B)$, and

$$B = \begin{bmatrix} b_1 I_1 & 0 & 0 & 0 \\ 0 & b_2 I_7 & 0 & 0 \\ 0 & 0 & b_3 I_3 & 0 \\ 0 & 0 & 0 & b_4 I_1 \end{bmatrix} \tag{D.17}$$

where b_i, $i = 1, ..., 4$, are the variances of the factors and have the same dimensions as the factors. B is a block diagonal matrix making the errors of the factors independent.

D.2.4 Flexible Heteroskedasticity

As described in equation (D.2), the error term $e_t \sim N(0, \Sigma_t)$. The variance Σ_t evolve over time implies heteroskedasticity in the model. In this paper, Σ_t is decomposited into a homoskedastic and a heteroskedastic part, as described in Dieppe, Legrand, and van Roye (2016),

$$\Sigma_t = exp(\zeta_t)\tilde{\Sigma}, \tag{D.18}$$

where Σ is a 21×21 matrix captures homoskedasticities, and ζ_t is a dynamic coefficient follows a autoregressive process,

$$\zeta_t = \gamma \zeta_{t-1} + \nu_t, \tag{D.19}$$

and the variance of its disturbances ν_t is φ. The initial value of ζ_0 is set to 0, which means the $\Sigma_0 = \tilde{\Sigma}$, and it varies with time with nonzero γ and φ.

With this setup, restrictions on parameters decides whether the residual e_t is homoskedastic or heteroskedastic and the precise form of heteroskedasticity. The variance φ decides whether there are heteroskedasticity in this model: when $\varphi = 0$, the residual is homoskedastic. The parameter γ determines the form of heteroskedasticity: when $\gamma = 1$, ζ_t follows a random walk; when $\gamma = 0$, ζ_t depends on a white noise ν_t; when $0 < \gamma < 1$, the effects from previous periods gradually disappear.

D.2.5 Summarize in State Space Model

A state space model consists of the observation equation of Y_t, the states θ_t and $\zeta_t$that include exogenous law of motion of shocks.

The observation equation is derived in equation (D.14)

$$Y_t = \chi_t \theta_t + e_t. \tag{D.20}$$

The law of motion of factors θ_t and the heteroskedastic part ζ_t, as defined in equation (D.16) and equation (D.19), form a system of state equations,

$$\theta_t = (1-\rho)\bar{\theta} + \rho\theta_{t-1} + \eta_t, \tag{D.21}$$

$$\zeta_t = \gamma\zeta_{t-1} + \nu_t, \tag{D.22}$$

The aboved equations described the state space model, and the disturbances are

$$e_t \sim N(0, exp(\zeta_t)\tilde{\Sigma}), \tag{D.23}$$

$$\nu_t \sim N(0, \varphi), \tag{D.24}$$

$$\eta_t \sim N(0, B_t). \tag{D.25}$$

E Bayesian Inference

This section discuss the Bayesian inference used in sampling the state space model mapped from the panel VAR. This section is organized by the following order: section 2 sets priors for parameters of interest and hyperparameters; section E.3 describes the distribution of the observed data, and steps for sampling from Metropolis within Gibbs algorithm.

E.1 Parameters of Interest

Priors of parameters of interest and likelihood of observed data are needed for obtaining posterior of the parameters.

For the state space model described above, the parameters of interest are factors $\theta = \{\theta_t\}_{t=1}^{T=97}$, heteroskedastic component in variance, $\zeta = \{\zeta_t\}_{t=1}^{T=97}$, residual variance, φ, factor variance, $b = \{b_i\}_{i=1}^{4}$ and homoskedasiticity component in variance, $\tilde{\Sigma}$. The joint posterior distribution of the parameters can be obtained by

$$\pi(\theta, \zeta, \varphi, b, \tilde{\Sigma}|Y) \propto L(Y|\theta, \zeta, \varphi, b, \tilde{\Sigma})p(\theta, \zeta, \varphi, b, \tilde{\Sigma}), \tag{E.1}$$

where $\pi(\cdot|\cdot)$ is the joint posterior distribution, and Y is the a vector of the observed data $\{Y_t\}_{t=1}^{T=97}$, $L(\cdot|\cdot)$ denotes likelihood of data, and $p(\cdot|\cdot)$ denote the joint prior distribution.

Since θ is conditional on b, and ζ is conditional on φ, equation (E.1) can be decomposed into

$$\begin{aligned} \pi(\theta, \zeta, \varphi, b, \tilde{\Sigma}|Y) &\propto L(Y|\theta, \zeta, \varphi, b, \tilde{\Sigma})p(\theta|b)p(b)p(\zeta|\varphi)p(\varphi)p(\tilde{\Sigma}) \\ &= L(Y|\theta, \zeta, \tilde{\Sigma})p(\theta|b)p(b)p(\zeta|\varphi)p(\varphi)p(\tilde{\Sigma}). \end{aligned} \tag{E.2}$$

The equation shows that prior distributions of factors θ, dynamic coefficient ζ, residual variance φ, factor variance b and homoskedasiticity component $\tilde{\Sigma}$, and the conditional distribution of the observed data $L(Y|\theta, \zeta, \tilde{\Sigma})$ are the basis of obtaining posterior distribution.

E.2 Priors

This subsection introduces priors for factors θ, heteroskedastic component ζ, residual variance φ, factor variance b and homoskedasticity component $\tilde{\Sigma}$.

E.2.1 Factors θ

Factors θ are mapped into a sparse matrix following Dieppe, Legrand, and van Roye (2016), instead of Kalman filter in Canova and Ciccarelli (2009). The sparse matrix approach is proposed by Chan and Jeliazkov (2009), and it is less computational demanding and more efficient while dealing with large state vectors. This approach is used to map the set of the high-density parameters into a low-density simple joint formulation.

This approach begins by the law of motion of factors θ, defined in equation (D.16), can be mapped into simultaneous equation form:

$$H\Theta = \tilde{\Theta} + \eta, \tag{E.3}$$

where

$$H = \begin{bmatrix} I_{13} & 0 & 0 & \cdots & 0 \\ -\rho I_{13} & I_{13} & 0 & \cdots & 0 \\ 0 & -\rho I_{13} & I_{13} & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\rho I_{13} & I_{13} \end{bmatrix} \tag{E.4}$$

is a transition matrix across time, and ρ is the persistence of the factors in equation (D.16). There are $T \times T$ ($T = 97$) blocks in H. The reason for the identity matrix in dimension of 13×13 is because there 13 factors. Θ is a compact matrix for θ_t,

$$\Theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_T \end{bmatrix}, \tag{E.5}$$

and

$$\tilde{\Theta} = \begin{bmatrix} (1-\rho)\bar{\theta}+\rho\theta_0 \\ (1-\rho)\bar{\theta} \\ \vdots \\ (1-\rho)\bar{\theta} \end{bmatrix} \tag{E.6}$$

is a $T\times 1$ matrix, and θ_0 and $\bar{\theta}$ are the initial value, and population mean obtained by OLS estimation of equation (D.14), as $\chi = W_t\xi$ can be taken as independent variables that are observable. As defined, $\theta_t = (\theta'_{1t}, \theta'_{2t}, \theta'_{3t}, \theta'_{4t})'$ and η is a vector of the disturbances in factors over time that defined in equation (D.16), $\{\eta_t\}_{t=1}^{T=97}$, where the prior

$$\eta \sim N(0, \tilde{B}), \tag{E.7}$$

the prior covariance matrix is $\tilde{B} \equiv I_T \otimes B$, where B is block diagonal and it is the covariance of η_t as defined in equation (D.17). $\tilde{B}$ shows that the variances of the factors are not correlated across time.

Take the first line of the matrix-form equation (E.3) as an example:

$$\theta_1 = (1-\rho)\bar{\theta} + \rho\theta_0 + \eta_1. \tag{E.8}$$

Thus, combining the simultaneous equations above, and equation (D.16), the prior distribution of the vector of factors across time follows $\Theta \sim N(\Theta_0, B_0)$, where $\Theta_0 = H^{-1}\tilde{\Theta}$ and $B_0 = H^{-1}\tilde{B}(H^{-1})'$.

E.2.2 Dynamic Coefficient ζ

The law of motion of ζ_t can also be mapped into a simultaneous equation form by the sparse matrix approach described above,

$$KZ = v, \tag{E.9}$$

where

$$K = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\gamma & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & -\gamma & 1 \end{bmatrix} \tag{E.10}$$

is a transition matrix across time, with γ defined in equation (D.19) as the persistence of ζ_t. There are $T \times T$ ($T = 97$) blocks in K.

The vector Z is the stacked-form ζ across time, $Z = (\zeta_1, \zeta_2, ..., \zeta_T)$, with the initial value $\zeta_0 = 0$. The disturbances v is a vector of the errors of the heteroskedastic part that defined in equation (D.19), $v = (\nu_1, \nu_2, ..., \nu_T)$, where the prior

$$v \sim N(0, \Phi), \tag{E.11}$$

the prior covariances matrix is $\Phi = \varphi I_T$, where φ is the covariance of v_t as defined in equation (D.19).

Take the first line of the matrix-form equation (E.9) as an example:

$$\zeta_1 = \nu_1, \tag{E.12}$$

as the initial value ζ_0 is set to 0, and the second line is

$$\zeta_2 = \gamma\zeta_1 + \nu_2. \tag{E.13}$$

Combine the simultaneous equations of ζ and the distribution of the errors $\nu = \{\nu\}_{t=1}^{T=97} \sim N(0, \varphi I_{97})$, the prior distribution of Z follows

$$Z \sim N(0, \varphi(K^{'}K)^{-1}). \tag{E.14}$$

This stacked-from distribution will be used in deriving the conditional posterior distribution of φ. But prior distribution for ζ_t are needed for deriving the conditional posterior of itself. Since ζ_t depends on previous period, the prior distribution of Z can be written as

$$p(Z|\varphi) = \prod_{t=1}^{T=97} p(\zeta_t|\zeta_{t-1}, \varphi). \tag{E.15}$$

Then from equation (D.19), the prior distribution of individual (ζ_t)s follows

$$\zeta_t \sim N(0, \nu_t), \tag{E.16}$$

and

$$p(\zeta_t|\zeta_{t-1}, \varphi) \propto \varphi^{-\frac{1}{2}} exp(-\frac{1}{2}\frac{(\zeta_t - \gamma\zeta_{t-1})^2}{\varphi}). \tag{E.17}$$

E.2.3 Residual Variance φ, Factor Variance b and Homoskedasticity Component $\tilde{\Sigma}$

The priors of the residual variance, as defined in equation (D.19), and the factor variance, as defined in equation (D.17), are inverse gamma distributions,

$$\varphi \sim IG\left(\frac{\alpha_0}{2}, \frac{\delta_0}{2}\right), \tag{E.18}$$

$$b_i \sim IG\left(\frac{a_0}{2}, \frac{c_0}{2}\right) \quad for \quad i = 1, ..., 4. \tag{E.19}$$

The prior for the variance homoskedastic part $\tilde{\Sigma}$, which is defined in equation (D.18), is a Jeffreys prior,

$$p(\tilde{\Sigma}) \propto |\tilde{\Sigma}|^{(21+1)/2}, \tag{E.20}$$

where the power of $|\tilde{\Sigma}|$ comes from the size of the covariance matrix. Jeffreys prior is non-informative as it is invariant to a change of the variable.

E.2.4 Parameterizing Priors

The priors need to be parameterized are the inverse gamma shape and scale of residual variance φ in equation (D.19), factor variance b_i in equation (D.17), the factor persistence ρ in equation (D.16), and the residual persistence γ in equation (D.19).

The inverse gamma shape and scale of residual variance φ and factor variance b_i are set to be 10000 and 1, which is non-informative as the distribution is flat, so that zero mean and the homoskedastic condition are assumed.

The priors for the parameters and hyperparameters are summarized in table 2 in section 2.2.2.

E.3 Metropolis within Gibbs Algorithm

From equation (D.14), and $e_t \sim N(0, exp(\zeta_t)\tilde{\Sigma})$, the likelihood of data is

$$L(Y|\theta, \tilde{\Sigma}, \zeta) = \prod_{t=1}^{T=97} f(y_t|\theta_t, \tilde{\Sigma}, \zeta_t)$$
$$\propto |\tilde{\Sigma}|^{\frac{97}{2}} exp(-\frac{1}{2}\sum_{t=1}^{T}\{exp(-\zeta_t)(Y_t - \chi_t\theta_t)'\tilde{\Sigma}^{-1}(Y_t - \chi_t\theta_t) + 21 \times \zeta_t\}). \tag{E.21}$$

Given the priors and likelihood, the posterior distributions of parameters of interest can be obtained by Metropolis-within-Gibbs algorithm. The algorithm is described in the following steps.

1. Initialize the starting values for $\theta^{(0)} = \{\theta_t^{(0)}\}_{t=1}^{97}$, $b^{(0)} = \{b_i^{(0)}\}_{i=1}^{4}$, $\tilde{\Sigma}^{(0)}$, $\zeta^{(0)} = \{\zeta_t^{(0)}\}_{t=1}^{97}$ and $\varphi^{(0)}$. The starting value of factors θ_t are set to $\bar{\theta}$ for all periods which are obtained from OLS estimation of equation (D.14), as Y_t and $\chi_t = W_t\xi$ are observable. The starting value of the variance of factors b_i is set to 10^5, which means the distribution of factors are diffuse. The homoskedasticity component $\tilde{\Sigma}^{(0)}$ is set to $\frac{1}{97}\sum_{i=1}^{97} e_t e_t'$, given the initial value of θ_t. The starting value of the heteroskedasticity component ζ_t is set to 0 for all periods, which assumes no heteroskedasticity. The the variance of the disturbances of the dynamic coefficient $\varphi^{(0)}$ is set to 0.001.

2. At each iteration n, repeat the following steps:

   (a) Draw the homoskedasticity component $\tilde{\Sigma}^n$, as defined in equation (D.18): Draw $\tilde{\Sigma}^n$ from $\pi(\tilde{\Sigma}^{(n)}|y, \theta^{(n-1)}, b^{(n-1)}, \zeta^{(n-1)}, \varphi^{(n-1)}) \sim IW(\bar{S}^{(n)}, T)$, where the scale $\bar{S}^{(n)} = \sum_{t=1}^{T}(y_t - \chi_t^{(n-1)}\theta_t^{(n-1)})exp(-\zeta_t^{(n-1)})(\sum_{t=1}^{T}(y_t - \chi_t^{(n-1)}\theta_t^{(n-1)})'$, and $T = 97$ degrees of freedom.

   (b) Draw the heteroskedasticity component $\zeta^{(n)}$, as defined in equation (D.18):

Using equation (E.21) and equation (E.14), the posterior of ζ is

$$\pi(\zeta^{(n)}|y,\theta^{(n-1)},b^{(n-1)},\zeta^{(n-1)},\varphi^{(n-1)})$$
$$\propto exp(-\frac{1}{2}[\sum_{t=1}^{T}\{exp(-\zeta_t^{(n-1)})(y_t-\chi_t\theta_t^{(n-1)})'\tilde{\Sigma}^{(n)(-1)}(y_t-\chi_t\theta_t^{(n-1)}) \tag{E.22}$$
$$+21\times\zeta_t^{(n-1)}\}+Z^{(n-1)'}\Phi(G'G)^{-1}Z^{(n-1)}]).$$

Since the likelihood above doesn't belongs to any known distribution, so direct sampling like Gibbs sampler is not applicable. In this case, Metropolis algorithm can be used to obtain $\zeta_t^{(n)}$ instead of equation (E.22). The Metropolis algorithm is described below.

i. Define the transition kernel of ζ_t as

$$\zeta_t^{(n)}=\zeta_t^{(n-1)}+\omega \tag{E.23}$$

where $\omega\sim N(0,\psi I_{97})$ with $\psi=10,000$ chosen to balance the variance of the draw and the acceptance probability.

ii. In each iteration, obtain a candidate from $\zeta_t^n\sim N(\bar{\zeta},\bar{\varphi})$, where $\bar{\varphi}=\frac{\varphi}{1+\gamma^2}$ and $\bar{\zeta}=\bar{\varphi}\frac{\gamma(\zeta_{t-1}+\zeta_{t+1})}{\varphi}$.

iii. In each iteration, compute the acceptance probability

$$\alpha(\zeta^{(n-1)},\zeta^{(n)})$$
$$=\frac{\pi(\zeta^{(n)})}{\pi(\zeta^{(n-1)})}$$
$$=exp(-\frac{1}{2}\sum_{t=1}^{97}(y_t-\chi_t\theta_t^{(n-1)})'\tilde{\Sigma}^{(n)(-1)}(y_t-\chi_t\theta_t^{(n-1)})\{exp(-\zeta_t^{(n)})-exp(-\zeta_t^{(n-1)})\})$$
$$\times exp(-\frac{21}{2}\sum_{t=1}^{97}\{\zeta_t^{(n)}-\zeta_t^{(n-1)}\})$$
$$\times exp(-\frac{1}{2}\{(\zeta_t^{(n)})'(\Phi(G'G)^{-1})^{-1}\zeta_t^{(n)}-(\zeta_t^{(n-1)})'(\Phi(G'G)^{-1})^{-1}\zeta_t^{(n-1)}\}). \tag{E.24}$$

iv. In each iteration, draw a random number k, where $k \sim U(0,1)$.

v. The acceptance law is: if $k \leq \alpha(\zeta^{(n-1)}, \zeta^{(n)})$, then the draw is accepted and $\zeta^{(n)} = \hat{\zeta}$; if not, the draw is rejected and keep the last draw $\zeta^{(n-1)}$.

vi. Repeat from step 2(b)i until iterations end.

(c) Draw the variance of the heteroskedasticity component, $\varphi^{(n)}$, as defined in equation (D.19):

The posterior of φ is

$$\begin{aligned} &\pi(\varphi^{(n)}|y, \theta^{(n-1)}, b^{(n-1)}, \zeta^{(n)}, \varphi^{(n-1)}) \\ &\propto \varphi^{-\frac{97+\alpha_0}{2}-1} exp(-\frac{Z^{(n)\prime} G' G Z^{(n)} + \delta_0}{2\varphi^{(n-1)}}), \end{aligned} \tag{E.25}$$

which is the kernel of an inverse Gamma distribution, and φ can be drawn from $IG(\frac{97+\alpha_0}{2}, \frac{Z^{(n)\prime} G' G Z^{(n)} + \delta_0}{2})$.

(d) Draw the variance of the factors, b_i^n, as defined in equation (D.17):

The posterior of b_i is

$$\begin{aligned} &\pi(b_i^{(n)}|y, \theta^{(n-1)}, b^{(n-1)}, \zeta^{(n)}, \varphi^{(n)}) \\ &\propto b_i^{-\frac{97d_i}{2}-1} exp(-\frac{\sum_{t=1}^{97}(\theta_{i,t}^{(n-1)} - \theta_{i,t-1}^{(n-1)})'(\theta_{i,t}^{(n-1)} - \theta_{i,t-1}^{(n-1)}) + c_0}{2b_i}), \end{aligned} \tag{E.26}$$

which is the kernel of an inverse Gamma distribution, where d_i denote dimensions of b_i, and b_i can be drawn from $IG(\frac{97d_i+a_0}{2}, \frac{\sum_{t=1}^{97}(\theta_{i,t}^{(n-1)} - \theta_{i,t-1}^{(n-1)})'(\theta_{i,t}^{(n-1)} - \theta_{i,t-1}^{(n-1)}) + c_0}{2})$.

(e) Recover the variance Σ^n, as defined in equation (D.1):

Draw Σ^n from the newly drawn $\tilde{\Sigma}^{(n)}$ and $\zeta^{(n)}$.

$$\Sigma_t^{(n)} = exp(\zeta_t^{(n)})\tilde{\Sigma}^{(n)}. \tag{E.27}$$

(f) Draw the factors θ^n, as defined in equation (D.8):

The posterior of Θ is

$$\begin{aligned}&\pi(\Theta^{(n)}|y,\theta^{(n-1)},b^{(n)},\zeta^{(n)},\varphi^{(n)})\\&\propto exp(-\frac{1}{2}(\Theta^{(n-1)}-\bar{\Theta}^{(n)})'(\bar{B}^{(n)})^{-1}(\Theta^{(n-1)}-\bar{\Theta}^{(n)}))\end{aligned} \tag{E.28}$$

which is a kernel of a multivariate normal distribution with

$$\bar{B}^{(n)}=(\xi'(\Sigma^n)^{-1}\xi+B_0^{-1})^{-1}, \tag{E.29}$$

and

$$\bar{\Theta}^{(n)}=\bar{B}^{(n)}(\xi'(\Sigma^n)^{-1}y+B_0^{-1}\Theta_0). \tag{E.30}$$

Thus Θ can be drawn from $N(\bar{\Theta},\bar{B})$.

Table 3 in section 2.2.2 summarizes posterior distributions of the parameters of interest.

F Empirical Results

F.1 Structural Shocks

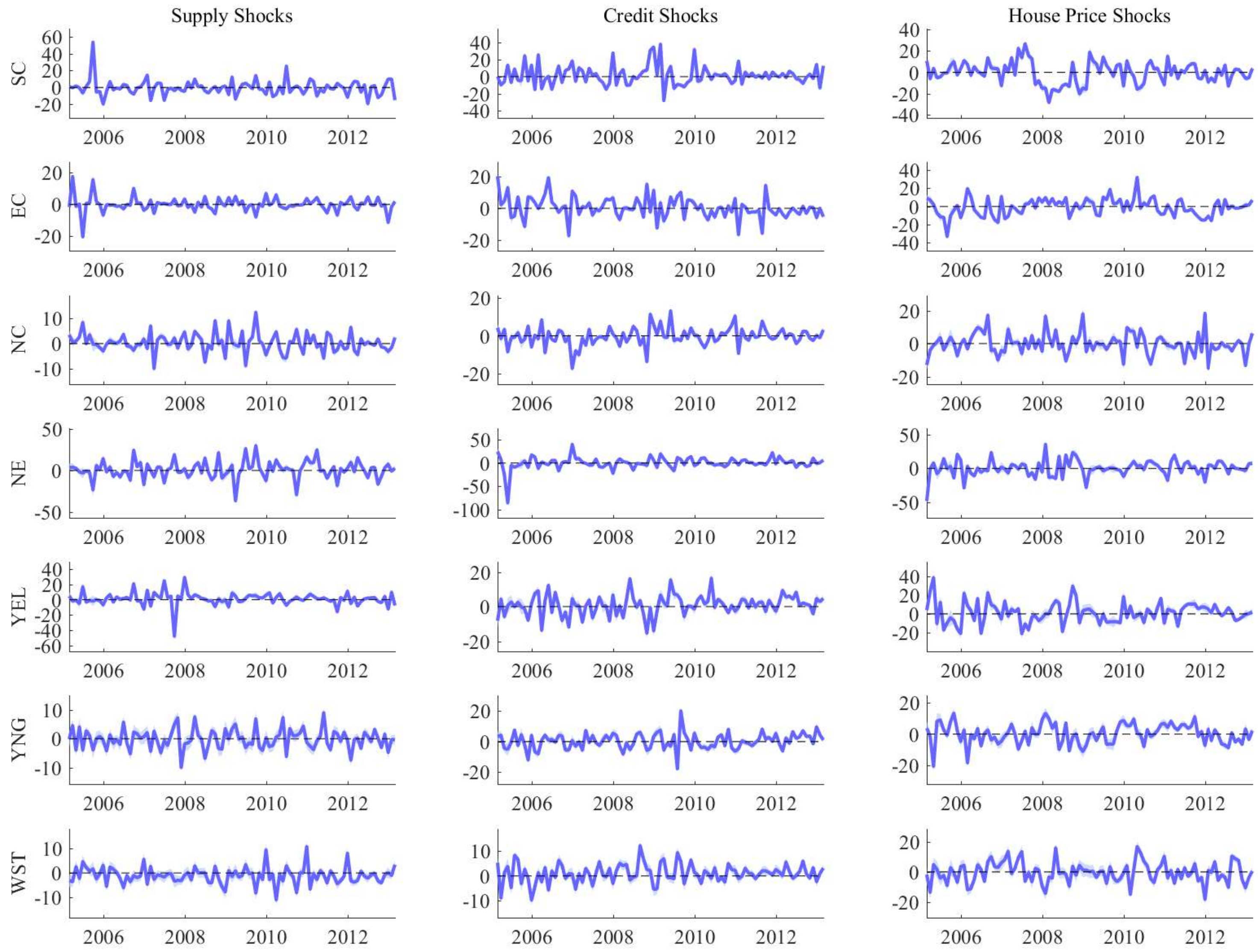
This subsection reports structural shocks. The structural shocks are supply shocks, credit supply shocks, and housing demand shocks from each region. They are mutually independent, and identified by recursive decomposition with the ordering as in section 2.3. Table 17 summarizes statistics of the structural shocks in each region during the sample period. Figure 15 plots the structural shocks over the sample period. The structural shocks of $\Delta lnGDP$ in the SC and EC regions are more volatile at the beginning of the sample, while the NC, NE and YEL regions exhibit the most volatility around the financial crisis. The structural credit shock from the NE region is stable except in 2005. The coastal regions shows a similar fluctuation in credit shocks, as they are less volatile during 2007-2009, and more volatile immediately after the crisis.

Table 17: Structural Shocks

Supply Shocks					
Region	**Mean**	**Std. Dev.**	**Min**	**Max**	**ACF(1)**
SC	-0.2921	9.1750	-20.0719	53.5273	-0.0681
EC	-0.0075	4.4942	-20.4063	17.5653	-0.0524
NC	0.3437	3.6759	-9.8930	12.3652	-0.1337
NE	0.8892	10.5582	-36.3460	29.7045	-0.0127
YEL	0.9959	8.4517	-47.8936	29.0607	-0.1334
YNG	-0.0924	3.4089	-9.9596	8.9837	-0.1419
WST	-1.0159	3.7755	-11.3667	11.0457	-0.1152
Credit supply shocks					
Region	**Mean**	**Std. Dev**	**Min**	**Max**	**ACF(1)**
SC	1.6460	11.5344	-28.1976	38.1590	-0.1784
EC	0.4701	6.6159	-17.1231	20.0154	-0.1053
NC	-0.2933	4.7862	-17.4283	13.3305	-0.0660
NE	0.3381	13.1017	-85.8850	40.1585	0.2024
YEL	1.0032	5.8490	-15.4132	16.4362	0.0646
YNG	0.3379	5.2026	-17.8307	20.0757	-0.0552
WST	1.0765	3.7408	-9.8058	12.2953	0.0199
Housing demand shocks					
Region	**Mean**	**Std. Dev**	**Min**	**Max**	**ACF(1)**
SC	0.3842	10.1248	-28.3985	26.4946	0.4013
EC	-0.7124	9.5555	-33.1004	32.1288	0.3116
NC	-0.0287	6.5093	-15.0279	18.3486	-0.0239
NE	0.5009	11.6469	-47.7767	35.5847	-0.0376
YEL	1.2162	10.9082	-21.0922	38.6016	0.1585
YNG	0.1673	6.3926	-20.7323	13.4027	0.2410
WST	-0.3146	6.8780	-18.2993	16.7764	0.1748

* This table summarizes unconditional statistics of the structural supply, credit supply, and housing demand shocks in each region during the sample.

Figure 15: Structural Shocks



Note: This figure displays structural supply, credit supply, and housing demand shocks from each region over the sample with 68% uncertainty bands.

F.2 House Price Spillovers

This subsection provides supplementary information and results for section 4.1.

Table 18: Comparison of Migrant Population in Five Cities

Year	Percentage of Immigrants to Permanent Population (%)				
	Beijing (1)**	**Shanghai (1)**	**Shenzhen (1)**	**Tianjin (2)**	**Shantou (3)**
2005	23.23		78.02	9.94	
2006	25.20		77.40	11.73	
2007	27.61		76.72	13.98	
2008	30.55		76.10	17.61	
2009	33.02	37.08	75.73	20.22	
2010	35.92	39.00	75.80	24.20	
2011	36.77	39.85	74.40	26.44	
2012	37.39	40.34	72.73	26.72	
2013	37.96	41.00	70.79	31.81	
Year	**Population Density (ppl/sq.km)**				
	Beijing (1)	**Shanghai (1)**	**Shenzhen (1)**	**Tianjin (2)**	**Shantou (3)**
2005	937	2981		887	2380
2006	976	3098		914	2399
2007	1021	3255		948	2426
2008	1079	3376		1000	2454
2009	1133	3486	4475	1044	2474
2010	1196	3632	5208	1105	2539
2011	1230	3702	5256	1152	2565
2012	1261	3754		1202	2581
2013	1289	3809	5323	1252	2616

* The top panel of the table summarizes the percentage of immigrants to permanent population in Beijing, Shanghai, Shenzhen, Tianjin, and Shantou from 2005 to 2013. Permanent population are residents with local *Hukou*. Immigrants lack local *Hukou*. The bottom panel display the population density in these cities. The data are provided by the National Bureau of Statistics in China.

** Beijing, Shanghai, and Shenzhen are first-tier cities in the NC, EC, and SC regions. Tianjin is a second-tier municipality adjacent to Beijing. Shantou is a third-tier city that is in Guangdong province, as are Shenzhen and Guangzhou (one of the four first-tier cities but not listed in the table due to data availability).