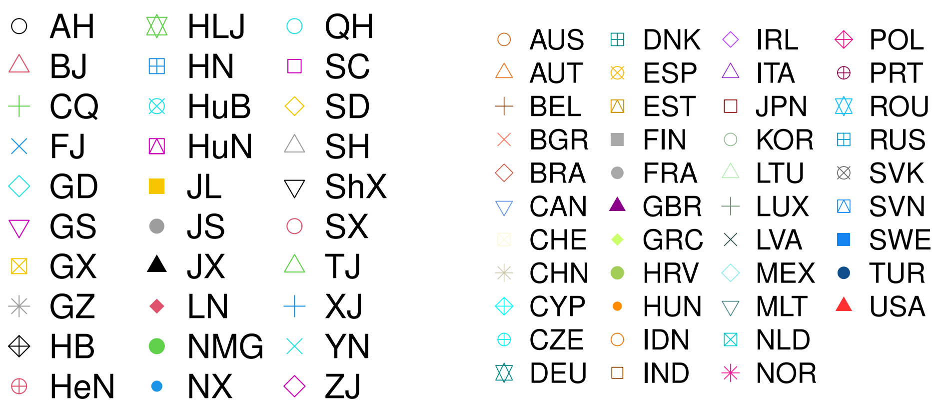
Appendices

A Abbreviations & Legends

Table 2: Abbreviations of Province Names.

Province	Abbreviation	Province	Abbreviation
Beijing	BJ	Henan	HeN
Tianjin	TJ	Hubei	HuB
Hebei	HB	Hunan	HuN
Shanxi	SX	Guangdong	GD
InnerMongolia	NMG	Guangxi	GX
Liaoning	LN	Hainan	HN
Jilin	JL	Chongqing	CQ
Heilongjiang	HLJ	Sichuan	SC
Shanghai	SH	Guizhou	GZ
Jiangsu	JS	Yunnan	YN
Zhejiang	ZJ	Shaanxi	ShX
Anhui	AH	Gansu	GS
Fujian	FJ	Qinghai	QH
Jiangxi	JX	Ningxia	NX
Shandong	SD	Xinjiang	XJ

Notes. This table shows the abbreviations of province names used in this work.

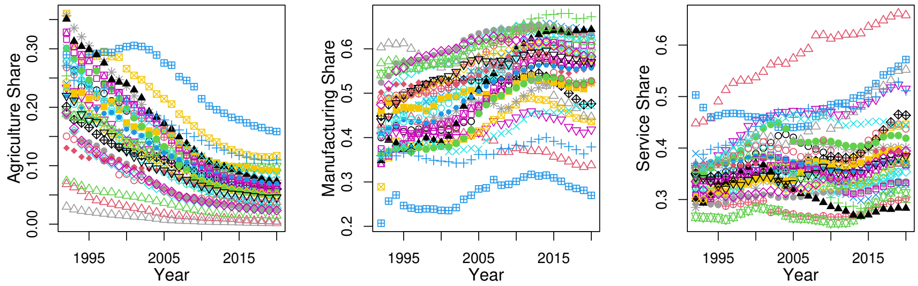


(a) Legend for China Provinces

(b) Legend for Countries

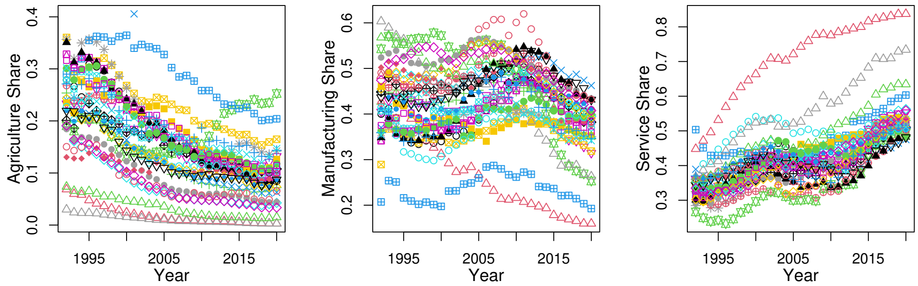
Figure 14: Legends. These figures show the legends used in this paper. The left panel shows the legend for selected Chinese provinces. The right panel shows the legend for selected countries.

B More Figures and Tables



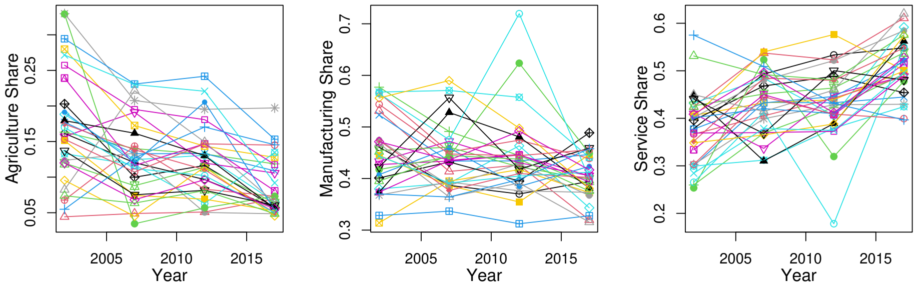
(a) Agricultural Shares (b) Manufacturing Shares (c) Service Shares

Figure 15: Real production VA Shares. The figures plot the real production VA shares of the primary, secondary, and tertiary sectors in the years from 1992 to 2020. Each style of point denotes a province according to the legend presented in Figure 14 in Appendix A. Xizang, Taiwan, Hong Kong, and Macao are absent in the figure lacking data. The data come from the NBS.



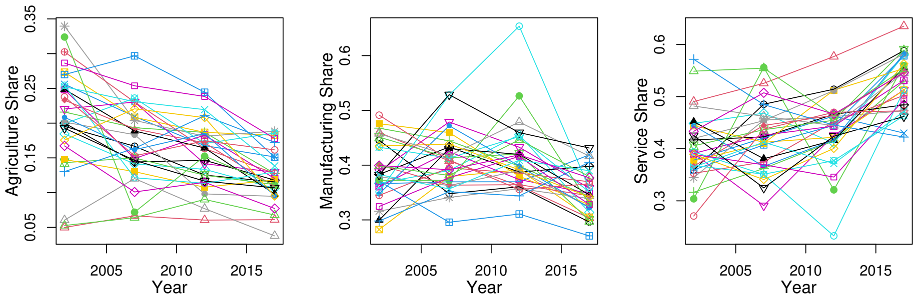
(a) Agricultural Shares (b) Manufacturing Shares (c) Service Shares

Figure 16: Nominal production VA Shares: This figure plots the nominal production VA shares of the primary, secondary, and tertiary sectors in the years from 1992 to 2020. Each style of point denotes a province according to the legend presented in Figure 14 in Appendix A. Xizang, Taiwan, Hong Kong, and Macao are absent in the figure lacking data. The data come from the NBS.



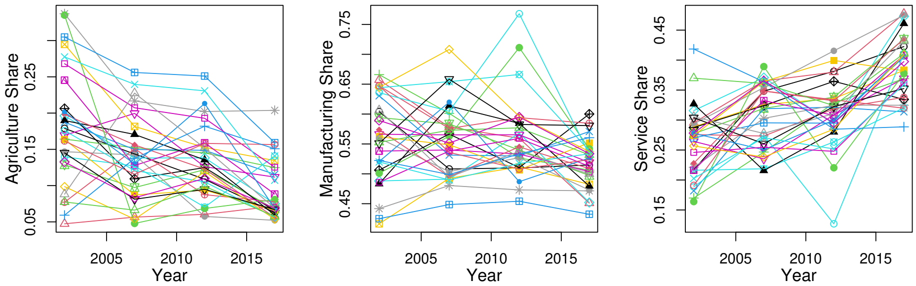
(a) Agricultural Shares (b) Manufacturing Shares (c) Service Shares

Figure 17: Final Consumption Shares: This figure plots the shares of consumption on final goods/services from the primary, secondary, and tertiary sectors in 2002, 2007, 2012, and 2017. The FC shares are calculated based on China's provincial I-O tables. Each style of point denotes a province according to the legend presented in Figure 14 in Appendix A. Xizang, Taiwan, Hong Kong, and Macao are absent in the figure lacking data.



(a) Agricultural Shares (b) Manufacturing Shares (c) Service Shares

Figure 18: Consumption VA Shares: This figure plots the shares of consumption on VA from the primary, secondary, and tertiary sectors in 2002, 2007, 2012, and 2017. The consumption VA shares are calculated based on China's provincial I-O tables. Each style of point denotes a province according to the legend presented in Figure 14 in Appendix A. Xizang, Taiwan, Hong Kong, and Macao are absent in the figure lacking data.



(a) Agricultural Shares (b) Manufacturing Shares (c) Service Shares

Figure 19: Provincial Consumption Shares on Food, Goods, and *CS* VA: This figure plots the shares of consumption on farm-gate products (food), factory-gate (goods), and *CS* VA in the years 2002, 2007, 2012, and 2017. The consumption shares are calculated based on China's provincial I-O tables. Each style of point denotes a province according to the legend presented in Figure 14 in Appendix A. Xizang, Taiwan, Hong Kong, and Macao are absent in the figure lacking data.

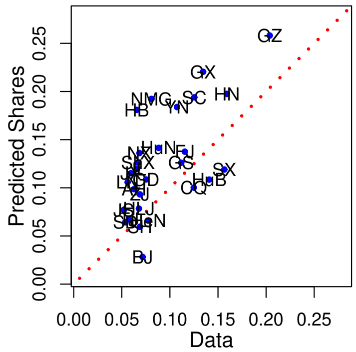


Figure 20: Predicted Food Shares in 2017 v.s. Data. This figure illustrates the predictive capacity of my model for consumption shares of foods. Predicted food consumption shares in 2017 are compared against the shares in data. First, I compute $d\ln(\vartheta_A - \omega_A)$ as $-\varepsilon \times d\ln e + \sum_s(\varepsilon \cdot \omega_s \times d\ln P_s)$, then combine it with food consumption share data in 2002 to predict the shares in 2017.

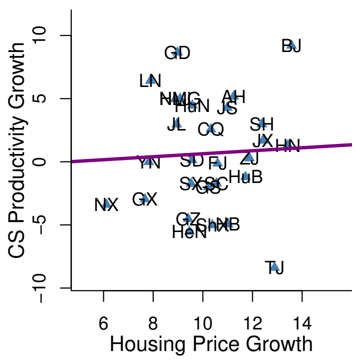


Figure 21: CS Productivity Growth v.s. Housing Price Growth. This figure plots the growth rates (in percentage points) of CS productivity growth against housing prices in the 28 Chinese provinces during 2002-2017 with a line fitted by OLS. Data on housing price growth are obtained from the NBS.

C Derivation of Equation 6

For this derivation, I suppress the individual, region, and time subscripts i, r, and t. From Equation 3 we know that λ_{ns} is the share of VA/GO of sector s used in final good n's production, which is also the income share of sector s when the firms maximize profits. Thus we have the following relationships between p_n's and P_s's:

$$p_n = P_F^{\lambda_{nF}} P_G^{\lambda_{nG}} P_{CS}^{\lambda_{nCS}} \tag{16}$$

Recall that ω_s's and β_s's are the asymptotic shares of F, G, and CS, thus we also have:

$$\omega_s = \int_n \lambda_{ns} \beta_n dn \tag{17}$$

Similarly, we have symmetric relationships between the current shares of sector s and the final goods consumption:

$$\vartheta_s = \int_n \lambda_{ns} \vartheta_n^{FE} dn \tag{18}$$

Finally, we can define ν_s's in the following way as combinations of κ_n's:

$$\nu_s = \int_n \lambda_{ns} \kappa_n dn \tag{19}$$

Now we are ready to map the consumption of final goods into that of the three sectors. Combining Equation 16 and Equation 17, we get:

$$\int_n p_{nrt}^{\beta_n} = P_{Frt}^{\omega_F} P_{Grt}^{\omega_G} P_{CSrt}^{\omega_{CS}} \tag{20}$$

Then we plug Equation 17, 19, and 20 into Equation 18 and substitute ϑ_n^{FE} by Equation 5 to get Equation 6.

D Estimation of ε

I estimate ε based on the regression of Equation 12 using household survey data from CHIP. Table 3 summarizes the results [19]. In column (1), I simply regress food consumption share on log consumption and province-year fixed effect. In column (2), additional controls of household size, number of workers, and household location (rural or urban) are included, which causes the estimated elasticity to increase from 0.341 to 0.412. In column (3) to column (8), I report the results of IV estimations, where the first-stage F statics are all significant. Besides column (4), the household principal's occupation serves as the instrument, and column (4) reports the results when an overidentification is done with the first three household members' (following the order in the survey data) occupations are instrumented. Columns (5) to column (8) report the results using CHIP 2002, CHIP 2007, CHIP 2013, and CHIP 2018, respectively. The estimated ε varies within a very small range with a length of 0.07 from 0.378 to 0.439. I thus choose 0.417 as the value for ε used in the calibration.

[19] In all the columns I trim the top and bottom 5% consumption levels to avoid bias caused by potential misreporting.

Table 3: Estimates of ε.

	Food Consumption share							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln c$	-0.341	-0.412	-0.417	-0.392	-0.423	-0.441	-0.439	-0.378
	(0.032)	(0.033)	(0.020)	(0.020)	(0.042)	(0.065)	(0.059)	(0.029)
Year	ALL	ALL	ALL	ALL	2002	2007	2013	2018
IV	N	N	Y	Y (3)	Y	Y	Y	Y
Addtl. Controls	N	Y	Y	Y	Y	Y	Y	Y
N	52970	52967	52967	52967	8280	11709	14569	18406
R^2	0.411	0.445	0.254	0.254	0.412	0.244	0.242	0.205
First-Stage F	-	-	420.458	371.386	102.174	45.754	55.871	174.400

Notes. This table summarizes the estimated coefficient ε of Equation 12.

E Measurement of CS Share

E.1 Deduction

In this paper, CS VA are defined as VA of services that never serve food and goods producers. In China's provincial I-O tables, the final consumption (FC) column reports the GO of each industry consumed by residents (containing inflowed and imported items). The consumption is measured by producer's price. Thus, the value of transportation, wholesale, and retail in final goods classified as food and goods are separately reported as final consumption in these service industries. In this situation, CS VA can only be recovered from the FC of service industries. To illustrate how to measure CS VA, first consider an economy with only three industries: non-service (NS), service 1 ($S1$), and service 2 ($S2$). The I-O tables tell us, to produce one unit GO of $S1$, how much GO of $S1$, $S2$, NS, and VA are input. Put differently, one unit FC of $S1$ could be decomposed to GO of NS, of $S2$, of itself, and VA generated in the final step. Then, the intermediate GO inputs can be further decomposed in similar ways (See Figure 22).

In the figure, I denote the number of units of CS VA contained in 1 unit of $S1$ FC as x_1. To produce 1 unit GO (FC) of $S1$, a_1 units GO of $S1$, b_1 units GO of $S2$, and v_1 units VA are input ($a_1 + b_1 + v_1 \leq 1$ as I abstract NS from the figure for simplicity). Similarly, x_2, a_2, b_2, and v_2 are symmetric notations for $S2$. The VA in red parts in Figure 22 never enter the production of NS. Thus, I should add all of them to obtain the CS VA recovered from FC of $S1$, and the rest are combinations of VA from PS and NS. Figure 22 can be simplified to Figure 23.

Or, I can express this figure in a linear system:

$$\begin{cases} x_1 = v_1 + a_1 x_1 + b_1 x_2 \\ x_2 = v_2 + a_2 x_1 + b_2 x_2 \end{cases} \tag{21}$$

Solving the system to obtain:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1-a_1 & -b_1 \\ -a_2 & 1-b_2 \end{bmatrix}^{-1} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{22}$$

All parameters on the right-hand side of the above equation are observable from the I-O

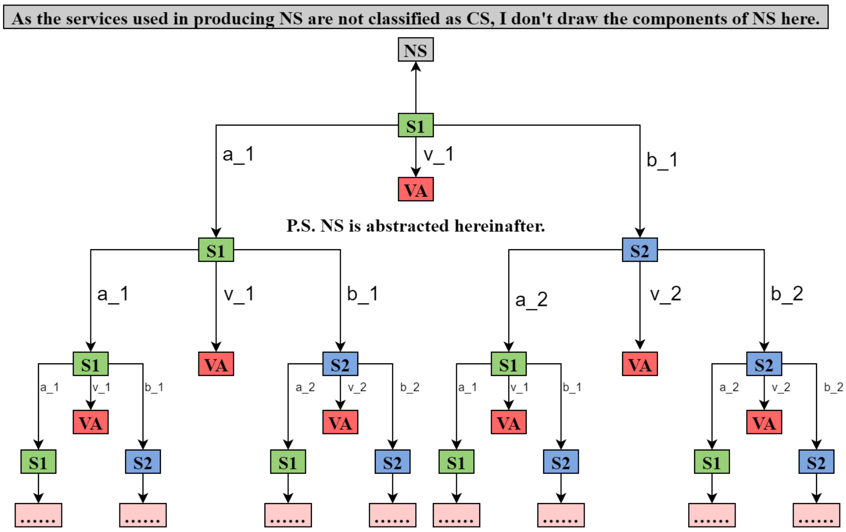


Figure 22: Decomposition of final consumption on S1.

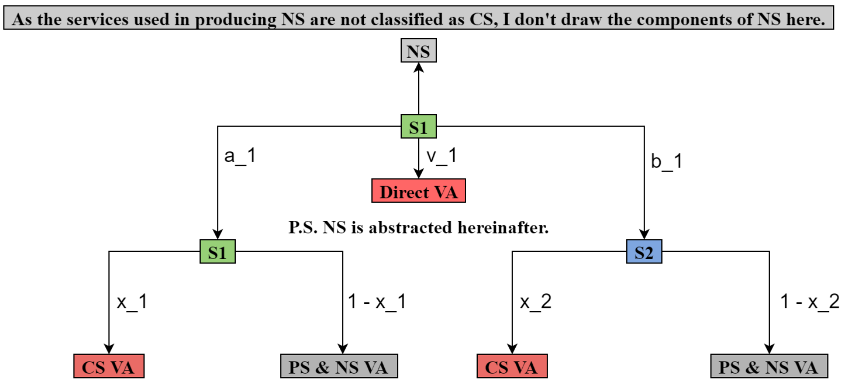


Figure 23: Decomposition of final expenditure on S1 (Simplified)

tables. I can easily generalize this two-dimension formula to cases with more service industries and calculate CS VA as:

$$CS \text{ VA} = \mathbf{C}_S{}'\mathbf{Q}\mathbf{V}_S \tag{23}$$

E.2 Discussions

In this section, I discuss the implicit assumptions of the method used to calculate the CS VA shares. When the assumptions are satisfied, the CS shares calculated align exactly with the definition of CS in the model. I exploit two sets of I-O tables (single-region and multi-region) to discuss the extent to which the assumption can be satisfied when possible and the potential results of them being violated. Three assumptions are needed to unbiasedly calculate the CS VA shares:

Assumption 1 A tertiary industry's input structure (consisting of shares of intermediate inputs and own VA) is the same no matter whether it's produced for goods producers or consumers. If this assumption is violated, the calculated CS share is biased either upward or downward depending on whether CS contains more VA or uses more intermediate inputs than PS. In my opinion, this assumption is reasonable in many tertiary industries. For example, a chief's VA share in each dish is almost the same no matter whether he works for a firm or a restaurant. A more obvious example is that a software firm sells the same products (thus the same share of VA contained) to both consumers and producers. However, it's hard to statistically examine this assumption given the data I have at hand. Reassuringly, even though the production of CS and PS have different input structures, if this difference is uniform in all provinces, the results of productivity growth in terms of provincial rankings are not affected too much.

Assumption 2 Direct service inputs for the production of CS (thus also CS) are produced locally. If this assumption is violated, local consumers may indirectly consume CS VA from other provinces and my method will classify some inflowed CS VA as locally produced. If CS VA is produced more efficiently in other provinces, I may overestimate the local CS productivity. To see the severity of this problem, I look at the multi-regional I-O table published in 2012. On average, to produce one unit of final service product (either PS or CS), 0.2 units of service (final product) are input, in which 87% is from local and only 13% is inflowed. However, if the 13% unit is only used in the production of PS (and PS and CS still share the same input structure), this would not cause any problems. Intuitively, the production of CS is usually small in scale and is not able to pay for inflowed services due to high trade costs.

Assumption 3 The final consumption vector should not contain inflowed and imported services. Similar to the last one, if this assumption is violated, local consumers may directly consume CS VA from other provinces and my method will classify some inflowed CS VA as locally produced. If CS VA is produced more efficiently in other provinces, we may finally overestimate the local CS productivity. However, this assumption can be easily violated considering industries like online shopping, transportation, etc. To see the severity of this issue, I again look at the multi-regional IO table in 2012. If I exclude Beijing and Tianjin and exclude retail & wholesale and transportation, most of the services produced by one province are consumed by this province's residents. The local consumed shares averaged higher than 90%, and almost all of them are above 80%. However, the issue is still severe if a province has a high share of retail & wholesale or transportation. Reassuringly, for some CS noted as inflowed, it's actually that local residents consume services in other localities when they're traveling, in which case the "trade" of services would not affect the non-tradability of CS and its implications [20]. The only assumption we need for these kind of "inflowed" services is that they are produced by the same technology in and out of the consumers' residential location.

[20] For example, assume that you live in Fujian and take a trip to Beijing. During the trip, you find that the catering services in Beijing charge much higher prices than in Fujian. But you probably will accept the high prices as you cannot quit eating in Beijing and compensate yourself after you are back. As a result, the CS price in a locality still can not affect the CS price in another locality.

E.3 Dealing with Potential Measurement Error

In the 2017 I-O tables of Hebei, Zhejiang, and Jilin, the reported real estate final consumptions (service industry) are extremely low and only account for 1%-2% of total consumption (the national average is 10%). Likely, these provinces did not impute rents for housing owners. There is strong evidence for this conjecture. For example, in 2017, the urban per capita equivalent rent for owned housing was 3138.95 yuan (*Hebei Economic Yearbook*) for Hebei. However, the urban per capita real estate final consumption reported in the I-O table was just 335.73 yuan, way below. To deal with this issue, I manually compute the equivalent rent for owned housing (per capita equivalent rent for owned housing × population) for the three provinces and add the numbers to the reported total real estate final consumption.

F Measurement of Other Variables

This section introduces the data and methods I use to calculate several key variables in the estimation process, which are provincial household consumption levels, provincial wage levels, provincial user cost of capital, and price indexes for food and goods. A statistical summary of these variables is provided in Table 4.

F.1 Household Consumption Level

I use provincial I-O tables to first calculate provincial total consumption. Total consumption is then divided by average population obtained from the *China Population & Employment Statistics Yearbook* to derive consumption per capita. Finally, consumption per capita is multiplied by the average number of household members to obtain consumption per household. Provincial data on the average number of household members come from the NBS, calculated as the ratio of the number of residents to the number of households in the annual sample survey of the household population.

F.2 Wage

Assuming uniform wage levels across sectors within a province, wages are computed as total labor compensation divided by total human capital. Total human capital is calculated as:

$$H_{rt} = \text{population} \times e^{0.1 \times \text{average education year}} \tag{24}$$

The return to education (0.1) is estimated by CPSZ using CHIP. The shares of people at each education level in the population are obtained from the *China Statistical Yearbook*. I multiplied the shares by corresponding education years to calculate the average education year.

F.3 User Cost of Capital

Provincial user costs of capital are calculated based on the following formula:

$$q_{rt} = P_{rt}R_t + P_{r,t+1} - (1-\delta)P_{r,t+1} = P_{rt}R_t + \delta P_{r,t+1} \tag{25}$$

where P_{rt} is the replacement cost of capital in region r at time t. R_t is the universal interest rate of China at time t. δ is the depreciation rate of capital set to be 0.05 following the literature.

F.4 Price Indexes for Food and Goods

Provincial price indexes for food and goods are derived from the difference between sectoral real and nominal VA published by the NBS, namely sectoral GDP deflators. Although GDP is measured in VA, their deflators are GO prices calculated from farm-gate and factory-gate prices.

Table 4: Annual Growth Rates (%) of Variables

Region	$\bar{c}_{rt}$	P_{rFt}	P_{rGt}	w_{rt}	q_{rt}	Region	$\bar{c}_{rt}$	P_{rFt}	P_{rGt}	w_{rt}	q_{rt}
BJ	10.66	4.64	1.71	10.19	0.88	TJ	12.49	1.10	-0.38	9.64	1.29
HB	11.90	3.54	3.13	12.52	1.47	SX	12.82	5.22	3.44	11.53	1.94
NMG	11.58	4.91	-0.75	13.72	1.68	LN	10.56	3.92	-0.61	8.89	1.39
JL	7.87	2.14	1.84	9.10	1.01	HLJ	9.51	5.82	-3.53	12.25	1.64
SH	9.74	5.51	0.88	9.47	1.17	JS	14.75	5.45	1.54	11.81	1.92
ZJ	11.66	4.23	1.84	11.64	1.48	AH	13.64	3.84	1.98	13.50	1.83
FJ	11.02	4.10	1.14	12.87	1.14	JX	13.40	3.02	1.56	11.60	2.12
SD	13.25	4.38	0.90	11.50	1.57	HeN	12.98	3.38	1.47	13.47	1.92
HuB	11.80	5.17	2.40	13.45	2.12	HuN	12.93	3.74	1.89	12.64	2.58
GD	9.33	4.10	1.21	9.70	1.42	GX	13.14	5.19	2.57	12.79	1.15
HN	12.12	2.62	1.33	11.78	2.03	CQ	12.64	3.83	-0.37	13.75	1.64
SC	12.62	4.54	0.37	13.68	2.07	GZ	15.39	7.05	3.35	15.84	1.54
YN	13.12	4.67	0.98	13.51	1.81	ShX	13.04	6.15	4.12	13.35	1.99
GS	11.44	2.98	0.39	11.27	1.56	NX	16.28	5.08	3.95	12.71	1.75

Notes. This table shows the annual growth rates of consumptions per household, food prices, goods prices, wages, and user costs of capital during 2002-2017 in the 28 provinces.

G Alternative Functional Form: IA

In Alder, Boppart, and Müller (2022), the authors propose the inter-temporal aggregatable (IA) preference. They argue that this preference is more flexible than PIGL in that it can explain the hump-shaped trend in manufacturing shares, allowing a sector to be a luxury and a necessity at different income levels. One form of IA preference given as indirect utility is [21]:

$$\mathcal{V}\left(c_i, \mathbf{P}\right) = \frac{1-\epsilon}{\epsilon}\left(\frac{c_i}{\mathbf{B}(\mathbf{P})} - \mathbf{A}(\mathbf{P})\right)^{\epsilon} - \mathbf{D}(\mathbf{P}) \tag{26}$$

where $\mathbf{A}(\mathbf{P})$ and $\mathbf{D}(\mathbf{P})$ are functions homogeneous of degree zero in prices, and $\mathbf{B}(\mathbf{P})$ is a linearly homogeneous function of prices. I impose the following parametrization:

$$\mathbf{B}\left(\mathbf{P}\right) = \prod_s \left(P_s\right)^{\omega_s}; \quad \mathbf{A}\left(\mathbf{P}\right) = \mathbf{B}\left(\mathbf{P}\right)^{-1}\sum_{s\in S} P_s \bar{c}_s; \quad \mathbf{D}\left(\mathbf{P}\right) = \sum_s \nu_s \ln P_s \tag{27}$$

[21] I suppress the time and region subscripts in this part.

where $\omega_s \geq 0$, $\sum_{s\in S}\omega_s = 1$, and $\sum_{s\in S}\nu_s = 0$. After applying Roy's Identity, it gives the following consumption share function:

$$\vartheta_s\left(c_i, \mathbf{P}\right) = \omega_s + \frac{P_s\bar{c}_s}{c_i} - \omega_s \cdot \frac{\sum_{j\in S} P_j\bar{c}_j}{c_i} + \nu_s\left(\frac{c_i - \sum_{j\in S} P_j\bar{c}_j}{\prod_j \left(P_j\right)^{\omega_j}}\right)^{1-\varepsilon} \times \frac{\prod_j \left(P_j\right)^{\omega_j}}{c_i} \tag{28}$$

In this case, if $\bar{c}_s = 0$, $\forall s$, then the IA preference here is the same as the PIGL form specified in this paper. I estimate parameters in Equation 28 using the iterated feasible generalized nonlinear least square (IFGNLS) method with pooled data of consumption shares on food and CS VA of the 28 Chinese provinces. Since the expenditure shares sum to one, I can drop the equation for goods consumption share. More details about the IFGNLS estimation can be found in Herrendorf, Rogerson, and Valentinyi (2013), Alder, Boppart, and Müller (2022), and the Stata document for the *nlsur* command.

Table 5: IA Estimates

	ω_A	ω_M	ω_{CS}	ε	ν_A	ν_{CS}	$\bar{c}_A$	$\bar{c}_M$	$\bar{c}_{CS}$
Coefficient	0.068	0.437	0.496	0.816	0.061	-0.708	0.017	-0.698	0.401
Std. err.	0.014	0.044	0.048	0.083	0.027	0.209	0.017	0.281	0.171
P-value	0.000	0.000	0.000	0.000	0.023	0.001	0.317	0.013	0.019

Table 5 reports the estimation results. The estimates of $\bar{c}_s$'s are very close to 0. This supports the use of PIGL preference in this paper.