Trade Liberalization and Structural Changes: Prefecture-level Evidence from China

Xuan Fei
State Administration of Foreign Exchange Investment Center

February 1, 2022

Abstract

Notable heterogeneity regarding international trade openness and level of development across different regions with significant inter-provincial migration has been a prominent feature in contemporary China. This paper studies the particular role of international trade in the structural transformation across each region in China. Combining several data, we document that there exhibits a non-monotone, hump-shaped relationship between international trade freeness and prefecture-level manufacture labor share in China. We rationalize this stylized fact in a spatial equilibrium model with multi-region, multi-sector, costly trade and endogenous allocation of labor, capital, and land. The analysis highlights the role of internal geography in shaping patterns of trade-induced structural change.

Keywords: Structural changes, trade liberalization, China
JEL Classification: F11, F14, O13, O14

1 Introduction

Structural transformation, defined as reallocation of resources across the broad economic sectors, including agriculture, manufacture and service, is one of the leading stylized facts of development (Kuznets, 1966). It is also the target of policy interventions in various emerging economies and developed countries. However, there remains no consensus on the economic forces that drive the process of structural transformation. Recent literature has started focused on the role of global integration, especially international trade in shaping the pattern of structural transformation.[1] This linkage is natural since one of international trade's fundamental role is to facilitate specialization via an efficient reallocation of employment and other factors of production across sectors.

However, this linkage is often analyzed at the aggregate country level, implicitly viewing every country as a point in space. Realistically, countries differ substantially on their internal geography, including internal trade cost and internal migration cost. Recently, the role of internal trade cost and internal migration cost that prohibit countries from participating in the world market has also received renewed attention. A growing literature suggests that internal trade costs and internal migration costs can be substantial.[2] What is the role of international trade in the pattern of structural transformation at the disaggregated level, taking account of the geographic variations within a country?

We investigate this question in China's setting since China exhibits considerable economic differences within a country, which is a particular ideal setting for investigating this question. The economic distribution of China is shown in Figure 1. The manufacturing sector and export volume is concentrated in the coastal areas in Figure 1a and Figure 1b. Eastern coast tends to be more open to international trade, which attracts higher shares of migrant worker, suggesting the inter-provincial migration from the interior regions to coastal areas. After China joins WTO, we have witnessed increasing external integration driven by implementations of several phased tariff reductions.[3] The extraordinary growth of the exports in China is accompanied by massive inter-provincial migration (Chen et al., 2010) due to the geographic variations in income and trade openness across Chinese regions. The migrant workers mostly move to prefectures that are faced with relatively larger export shock.

[1] See, for example, Matsuyama (2009), Uy et al. (2013), and Swiecki et al. (2014).

[2] For instance, Atkin and Donaldson (2015) discuss the intra-national trade cost in African countries. Desmet et al. (2015) emphasize the importance of migration restrictions in welfare gains from trade and economic growth path, and Tombe and Zhu (2019) study the impact of large intra-national trade cost and labor migration cost in China on productivity.

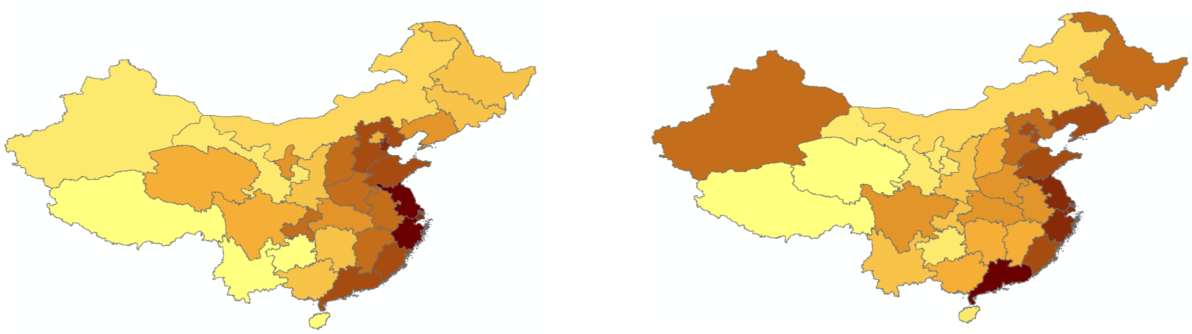
[3] China's simple average applied tariff was 17% in 2000 with the standard deviation across the six-digit Harmonized System (HS6) products being 12%. By the end of 2005, the average tariff level reduced to 6%, and the standard deviation almost halved (Zi, 2016).

External integration and its induced factor reallocation have a tremendous impact on the process of China structural transformation.

Figure 1: Economic Distribution in China

(a) Manufacture Labor Share

(b) Export Volume



Notes: The manufacture and export volume data are from China Data Online (2008). The darker the color is, the bigger the numbers are.

China has faced a spectacular export boom with an annual growth rate of 15% from the year 1999 with most of the exports concentrated in export-oriented manufacturing industries and exports were unevenly distributed across regions (most exports are concentrated in the east-coast areas). In the meantime, the aggregate manufacture output share in China has exhibited a hump-shape while the service output share in China has kept increasing from 39% in 1999 to 50% in 2015.[4] The increase in manufacturing exports doesn't naturally lead to the monotone increase in China's manufacturing output share.

To understand how trade liberalization affects structural change, in this paper, we emphasize the regional variations of trade exposures and the underlying factor reallocations instead of viewing the country as a point compared to previous macroeconomics literature. Specifically, we focus on the structural transformation from the manufacturing sector to the service sector within the urban regions, which deviates from previous literature that mainly targets the structural transformation from agriculture to the non-agriculture sector in China (Garriga et al., 2014; Tombe and Zhu, 2019).

To answer the question raised above, we first use the data from China City Year Statistics, China Customs data, TRAINS Database and China Industrial Production Data to provide cross-sectional evidence and document a stylized fact that there exists a non-monotone, hump-shaped relationship between trade openness and manufacture labor share after con-

[4]The manufacture output share was 45% in the year 1999 and then reached its highest value with 48% in the year 2006. But the share declined after that and came 41% in the year 2015.

trolling for income.[5] We also conduct a series of robustness checks, and find that our main empirical results hold under various circumstances.

To understand this stylized fact, we develop a tractable quantitative general equilibrium model in the setting of an open economy with multiple regions. The model allows us to capture economic activity distribution across regions and sectors, which are crucial to understanding the economic development process since this process is typically characterized by structural transformation and regional variations. Specifically, the model includes a home country with many locations and one foreign country representing the rest of the world. The model contains two sectors, including manufacture as tradable and service as non-tradable. Both sectors use land, capital, and labor as factors of production with various intensities, and production in each sector follows the structure of Eaton and Kortum (2002).The total land endowment and its allocation across manufacturing and service sector is exogenous given. Labor and capital in each location are endogenously allocated across manufacturing and service. Workers are mobile across regions and they endogenously choose their location to arbitrage away the real wage differences. The country's international trade cost affects the country's manufacturing labor share, and the manufacturing sector is a tradable sector.

The model shows the following. First, when the country has a high international cost and is far from the rest of the world, it can be viewed as a point that means that better market access will increase its manufacturing labor share. This rise in manufacture labor share is called *market access channel*. Second, when the country is close enough to the world market, internal factor reallocation matters. Since labor migrates to coastal areas due to higher wages, coastal areas tend to have higher population density and lower wage-rental ratio due to limited land supply. Coastal areas then choose to specialize in relatively more labor-intensive industries, which are the service sector, and decrease their manufacturing labor share. This is called *internal factor reallocation channel*.

To rationalize the empirical findings, we quantify the model and find that the model can generate the non-monotone, hump-shaped relationship between the trade openness and manufacture labor share, which echoes with the stylized fact in the empirical part. We illustrate the two channels of the non-linear relationship by changing international trade costs to show the two extreme cases. Furthermore, we carry the decomposition exercise to show the two channels' relative importance at the rising and falling part of the hump-shape pattern.

[5]The literature has documented that the higher income will lead to changes in labor shares across different sectors with non-homothetic preference (Uy et al., 2013; Swiecki et al., 2014). But we find that the hump-shaped pattern still exists after controlling for income, which means that the demand side does not just solely drive this hump-shaped pattern from the non-homothetic preference. In the paper, we propose that the regional variations of trade exposures and the underlying regional factor reallocation would also help understand the structural transformation.

We use the model's structure to undertake counterfactuals for changes in land policy and quantify its role in explaining differences in aggregate welfare. The process of urbanization and development has enlarged the urban land demand in the major cities. However, due to the considerations of an adequate food supply, the Chinese government imposes strict land supply limits to ensure a certain amount of arable land, which has received many critics concerning prohibiting the urbanization and giving rise to roaring housing prices in the major cities. We find that increasing the biggest city's land supply by 50%, while holding other variables constant, increases the real wages by 1% and reduces the rents in the biggest city by 19%, leading to a more flattening manufacture labor share across cities. Admittedly, the magnitude of the counterfactual results should be taken into more cautious consideration since the paper are abstract from accounting for the cost and benefits of the urban land provisions, which leaves us a very interesting topic for future research.

The paper is related to several literature. First, the work is closely related to literature about structural change and international trade. There is earlier work on structural change in an open economy setting. Echevarria (1995) employs a dynamic general equilibrium model to study the effect of trade on the composition of the national product and the economy's overall growth rate applying to OECD countries. Matsuyama (2009) presents a simple Ricardian model and shows that the growth in manufacturing productivity might not lead to a decline in manufacturing employment in an open economy. A more recent one is Uy et al. (2013) who specifically studies the importance of international trade in structural change theoretically in a Ricardian model following Eaton and Kortum (2002) and empirically. Other similar works include Caselli and Coleman (2001), Ngai and Pissarides (2007), Betts et al. (2013), Herrendorf et al. (2013), and Swiecki et al. (2014). However, none of them exploits the within-country factor reallocation effect. In the macroeconomics literature, countries are viewed as a point; in contrast, in this model, less remote areas have higher population density and lower wage-rental ratios, which leads to a structural transformation from manufacture to service. Literature has pointed out several determinants of structural changes, including sector-biased technological changes, non-homothetic preferences, international trade, and changes in wedges of factor costs across sectors (Swiecki et al., 2014). This paper is trying to provide another angle while considering internal geography.

This paper is also related to growing theoretical literature on new economic geography as synthesized by Fujita and Krugman (1999). Recently a small amount of paper start to develop quantitative models of trade with endogenous internal distributions of economic activity, including Redding (2012), Allen and Arkolakis (2013), Caliendo et al. (2014), Coşar and Fajgelbaum (2016). None of the papers examine the relationship between external integration, structural transformation, and economic development.

This paper also connects with the growing spatial models and exercises applying to China's structural transformation including Garriga et al. (2014), Zi (2016), Tombe and Zhu (2019) and Fan (2019). These papers focus on different aspects of research questions, but their models all contain one specific feature uniquely critical to recent China: *labor migration and reallocation*, which we also include in our paper. But deviating from these papers, which mainly focuses on the transformation from the agriculture sector to the non-agriculture sector, in this paper, we specifically would like to look at the structural transformation within the urban regions, that is, the shift from the manufacturing to the service sector. [6]

The remainder of the paper is structured as follows. Section 2 provides the main stylized fact from the data. Section 3 develops a theoretical model that rationalizes the stylized facts from Section 2. Section 4 shows the numerical result of the model and the counterfactual exercise. Section 5 concludes.

2 Empirical Motivation

We start the analysis by looking at some stylized facts from the data to see a broad impact of trade liberalization on the structural change within China. The cross-sectional evidence covers year 1998-2008.

2.1 Empirical Strategy

The unit of the analysis is a prefecture in China, which is an administrative division ranking between province and county. We construct a variable "*Export Shock*" (following the spirit of Topalova (2010), David et al. (2013) and Bombardini and Li (2020)) to proxy for the trade

[6] There is two paper that is highly related to my research question including Uy et al. (2013) and Fajgelbaum and Redding (2014). Still, their focal point is slightly different from the question in this paper. We are interested in international trade's role in impacting structural transformation, considering regional factor reallocation. Uy et al. (2013) theoretically and empirically emphasizes the role of international trade on structural change in a two-country multi-sector open economy model through comparative advantage effect of international trade. Still, they didn't look at the geographic variations within a country and the effect of regional factor reallocation. Fajgelbaum and Redding (2014) studies the role of external integration on structural change through the regional factor reallocation effect in a small open economy setting where they view price index in the rest of the world as exogenous. Although their small open economy assumption mainly simplifies their model analysis and is also realistic in the historic Argentina background, this assumption neglects the comparative advantage effect through international trade due to this *exogenous* price index level. If we model international trade following stylized Eaton and Kortum (2002) framework, the price index cannot be exogenous if allowing the comparative advantage of international trade.

openness for each prefecture:

$$ExportShock_{it} = \sum_k \frac{L_{ik,t-1}}{L_{k,t-1}} \frac{\Delta EX_{kt}}{L_{i,t-1}},$$

where i represents for prefecture, k represents for industry. $L_{ik,t-1}$ is the labor employment in industry k and region i at last time period[7], and ΔEX_{kt} is the change of exports in sector k at time t. We exploit the regional industry compositions to construct this variable.

The baseline regression is

$$\Delta y_{it} = \delta_1 + \delta_2 ExportShock_{it} + \delta_3 \Delta X_{it} + v_{it},$$

where y_{it} is our dependent variable such as manufacture output or labor share in prefecture i, X_{it} is a set of control variables including GDP, GDP per capita and sectoral productivity changes.

We're still concerned that the error term v_{it} might be affected by the demand and supply factors which are correlated with the export shock since local productivity or factor supply changes may have an impact on local output and exports. The Bartik approach measure *ExportShock* deals with this issue by not employing export expansion at the local level but instead using a weighted average of national export expansion. But this approach also assumes that other time-varying, region-specific determinants of the outcome variable are uncorrelated with (1) region's initial industry composition and (2) industry shocks at the national level. We tackle the first issue (1) by controlling for pre-existing trends in the dependent variable.

Another concern of using the Bartik approach is that an industry cluster in a particular location and the location also highly specializes in the industry. And then condition (2) is not valid since the national shock will co-move with the local shock. In this case, the Bartik approach is unable to provide an exogenous local shock apart from national shock. To deal with this possible endogeneity issue of *ExportShock*, we use the tariff faced by China when China exports to other countries and construct an IV denoted as *ExportTariff*:

$$ExportTariff_{kt} = \sum_j \frac{EX_{kj,t-1}}{EX_{k,t-1}} \tau_{kjt},$$

where τ_{kjt} is the tariff imposed by China's export destination country j in sector k at time t and $\frac{EX_{kj,t-1}}{EX_{k,t-1}}$ is the destination country j's export share in sector k at time $t-1$. We assume that the tariff imposed by other countries are politically determined by foreign countries and

[7]Here, We use 10-year-lag.

are uncorrelated with the economic circumstances in China. The variations purely come from the foreign demand. We sum up all the tariff across all the destination countries and they are weighted by their export shares. The following specification gives us the first-stage estimated exports in sector k at time t:

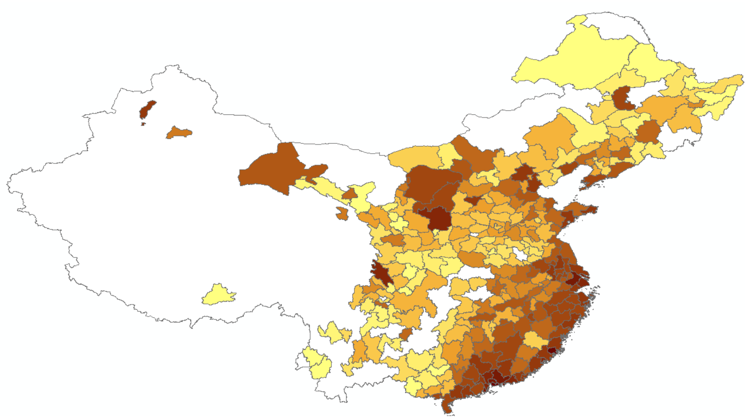
$$lnEX_{kt} = \eta_k + \phi_t + \gamma lnExportTariff_{kt} + \varepsilon_{kt},$$

where η_k is the sector fixed effect and ϕ_t is time fixed effect. We use the estimated exports $\hat{EX}_{kt}$ from the above specification to construct the estimated $ExportShock$ variable below:

$$Export\hat{Shock}_{i,t} = \sum_k \frac{L_{ik,t-1}}{L_{i,t-1}} \frac{\Delta \hat{EX}_{kt}}{L_{i,t-1}}, \tag{1}$$

where $\Delta \hat{EX}_{kt}$ is the change of expected export $\hat{EX}_{kt}$ in sector k at time t. The estimated export shock can be seen from Figure 2. It roughly matches with our intuition that coastal areas have higher export shock since they are closer to the world market while the hinterland areas are more remote to the rest of the world.[8] The first stage is reported in Figure A.1.

Figure 2: Constructed Trade Shock Distribution in China



Notes: The trade shock is estimated using Equation (1). The blank areas are regions with no trade statistics since they are remote. The darker the color is, the bigger the numbers are. The figure shows the inland China only.

[8] There are some dark areas in the west as well and this might be due to their abundant natural resources like oil or coal.

And then the second-stage of the IV regression is[9]

$$\Delta y_i = \delta_1 + \delta_2 Expor\hat{t}Shock_i + \delta_3 Expor\hat{t}Shock_i^2 + \delta_4 \Delta X_i + \delta_5 y_{ini,i} + \delta_6 Capital_i + v_i,$$

where y_i is our dependent variable such as manufacture output or labor share in prefecture i, X_i is a set of control variables including GDP, GDP per capita and sectoral productivity changes, $y_{ini,i}$ is the initial period's value of our dependent variable in prefecture i. We also include $Capital_i$ to control for prefectures which are the capital cities in a province.

2.2 Data

We use China's Customs Data to obtain the export data in prefecture level and sector level (at HS-8 digit) covering year 1997-2015. The employment data is from China's Manufacture Annual Firm Survey Data covering year 1998-2008. This firm-level survey data accounts for 60% of total manufacturing employment (Coşar and Fajgelbaum, 2016) though it doesn't cover all the firms in China. We use this dataset to measure the employment share for each region at the sector level. However, this dataset is using 4-digit CIC (Chinese industry code). So we created a crosswalk to map the 4-digit CIC code to the HS-8 digit code in the China's Customs Data by standardizing them into ISIC industry code.[10] The tariff data is TRAINS Database from World Bank. And the macroeconomic indicators (contains the agriculture, manufacture and service output and labor shares) at the prefecture level are from China's City Statistics Yearbook.

2.3 Empirical Result

The empirical result is presented in this section. We look at the impact of trade liberalization on the regional manufacture employment share among 1999-2008.[11]

Specifically, we would like to focus on the labor allocation from manufacturing industry to service industry within the urban regions in China, excluding the agriculture sector and the rural labor forces. We define the main dependent variable manufacture labor share as the following expression:

[9] Since we only have one 10-year time lag here(1998-2008), we neglect the time subscript here. And the change Δ is the value of year 2008 deducted by the value of year 1999.

[10] We end up with 142 manufacture industries in the ISIC code.

[11] We could extend the time period more recently, but due to the employment data limitation (which only covers until 2008), we choose this time period which has approximately 10 year gap and covering year 2001 when China joins WTO. We also notice that there are no significant changes in the intra-national trade cost during this period which rules out this confounding possibility since the high-speed railway started to be massively construct after 2008 and the major sizable and critical highways are almost constructed before 1998.

$$\text{Manufacture Labor Share}_i^* = \frac{emp_{Mi}}{emp_i - emp_{Ai}}, \tag{2}$$

where emp_{Mi} is the employment in the manufacturing sector, emp_i is the employment in prefecture i and emp_{Ai} is the employment in the agriculture sector.[12] Hence the denominator here indicates the total employment in the urban sector in each prefecture.[13]

Table 1: Trade Liberalization and Manufacture Labor Share

	Change in Labor Share		Level in Labor Share	
	OLS	2SLS	OLS	2SLS
ExportShock	0.309***	0.361***	0.262**	0.307**
	(0.034)	(0.040)	(0.052)	(0.063)
ExportShock^2	-0.043***	-0.059***	-0.038**	-0.052**
	(0.004)	(0.006)	(0.007)	(0.010)
R^2	0.30	0.30	0.64	0.64
N	188	188	188	188
Region FE	Y	Y	Y	Y
Initial Conditions	Y	Y	Y	Y
Contemporary Shocks	Y	Y	Y	Y
Capital City	Y	Y	Y	Y

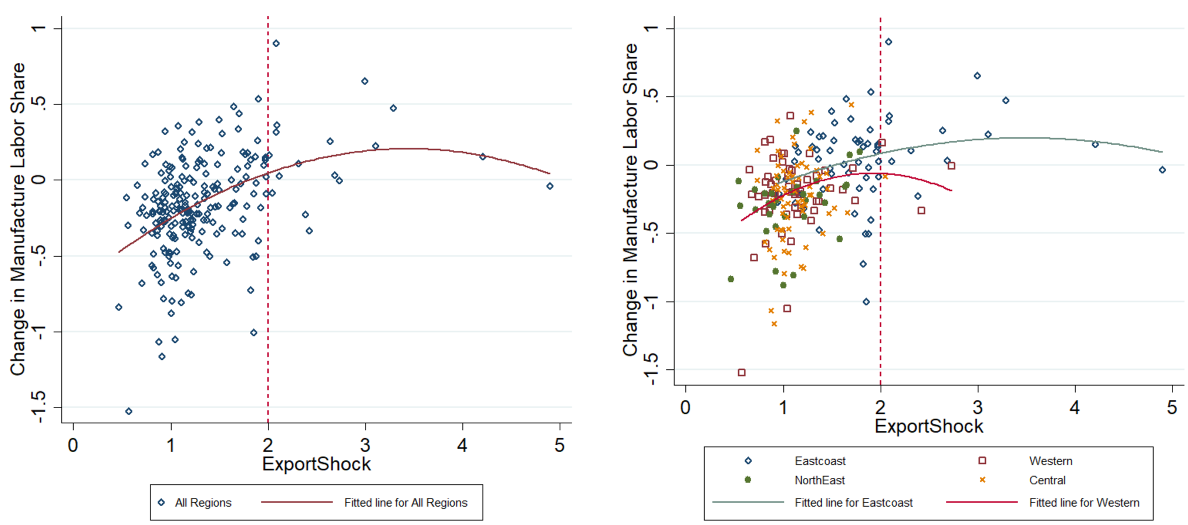
Notes: The dependent variables are the change or level of the manufacturing labor share. Here we define the manufacturing labor share as Equation (2). We also exclude mining, construction and electricity, gas and water industry in both the numerator and denominator of the manufacture labor share. Initial conditions include the start period's manufacture labor share. Contemporaneous shocks include change in log GDP per capita, change in log GDP, sectoral productivity changes and a dummy of provincial capital city. Region fixed effect is also included. Standard errors are clustered at the region level. ***p<0.01, **p<0.05, *p<0.1

[12]In this way, we focus on the structural transformation among the manufacturing and service sectors controlling for the rural-urban migration. Admittedly, the study-period 1998-2008 features the rapid urbanization in China, and might lead to the structural transformation from agriculture to industry sectors. Previous literature has focused on this issue (Tombe and Zhu, 2019), which is not the focus of our paper. Instead, we would like to control the rural-urban migration by focusing on the urban population in the empirical part. In the model part, we will extend the model and show that the rural-urban migration channel will not affect our main result of the hump-shape pattern.

[13]To be clear, all the dependent variable in the baseline regression and the robustness checks is the manufacturing labor share defined as here.

As shown in Table 1, there is a significant hump shape in the relationship between the trade shock and manufacturing labor share.[14] The result is robust to both the changes in labor share and the ending period of labor share as the dependent variable. We also show the scatter plot in Figure 3. In the left panel, we show the scatter and fitted lines for all regions. In the right panel, we divide prefectures into four different regions including east-coast, northcoast, western and central regions.[15] The scatter plot shows that east-coast and western China has exhibited a significant pattern of hump-shaped. In contrast, for central China and northeast China, the pattern is not very significant. This finding might suggest the structural transformation happens mainly among regions instead of within regions. It could suggest that central and northeast China has developed slowly compared to the east coast and has not entered the period of the hump-shape pattern from 1998 to 2008.[16]

Figure 3: Manufacture Labor Share and Export Shock



Note: The figure illustrates the relationship between prefecture-level export shock and manufacture labor share. In the left panel, we show the scatter and fitted lines for all regions. In the right panel, we divide prefectures into four different regions including eastcoast, northcoast, western and central regions and we also plot the fitted lines for the eastcoast and western regions.

[14]This hump-shaped relationship also robustly passes the Lind and Mehlum (2010) test for non-monotone relationships.

[15]The allocation of the four broad economics regions can be found in Figure A.2. There are four broad economic regions in China: East coast, Central China, Northeast China, and Western China.

[16]We also empirically test this and more details are in Appendix B.4.

2.4 Robustness Checks and Further Discussions

In this section, we discuss and provide a series of robustness checks of the main findings in the empirical part. Our main results remain robust in various robustness checks.

First, one might worry that the trade shock measure which we constructed in previous sections might not be a perfect measure since some areas have high international trade cost and high measure of "trade shock" due to their abundant natural resources. In addition, cities including Beijing and Shanghai have relatively low international trade cost but their trade shock measure is not high enough. These cities probably already start to specialize in service sector. Hence, we also use an alternative measure of the trade openness: the distance to the nearest port city as robustness check. Table A.1 finds that the hump-shape pattern between the manufacturing share and the trade shock remains robust if using this alternative proxy for the trade openness.

Second, to test whether the hump shape pattern is driven by a few cities on the right side in Figure 3, we formally conduct the empirical analysis by dropping the cities with large export shocks in the sample and rerun the regressions. Table A.2 shows that even after excluding cities with large export shock, the hump-shape pattern still exists.

Third, the major empirical result employs China's Manufacture Annual Firm Survey Data from year 1998-2008. The dataset accounts for 60% of total manufacturing employment but may neglect employment from small and medium sized enterprises. It could be that in small cities, the small and medium enterprises are small in individual size but large in total employment, especially in hinterland area in China. To test whether the selection problem in our sample might pollute our main result, we also double check this issue using data from prefecture-level city statistics yearbooks and compare it with the data from the manufacture firm survey. In Figure A.5, we present the scatter plot between manufacturing employment and city size by using both the city statistics yearbook and the firm survey data. We find little evidence on the systematic selection problems. To further check whether the small cities' selection problem might also mislead our main findings, we also exclude small cities and rerun the baseline regressions, and do not find significant changes of our major results in Table A.3.

Fourthly, we also conduct robustness checks using other key variables that proxy for the structural changes. In the main result, we use manufacture labor share as the dependent variable to resemble structural changes. In this part, we also use the manufacture GDP share as the dependent variable and find that the hump shape pattern still holds in Table A.5.

3 Model

In this section, we develop a two-sector multi-region spatial equilibrium model in an open economy with labor mobility and endogenous factor allocation building on Eaton and Kortum (2002) and Fajgelbaum and Redding (2014). We illustrate that this model can generate the non-monotone relationship between trade openness and the manufacture labor share within country which is found in Section 2.

3.1 Setting

The economy consists of n+1 multiple locations. These locations include n locations, and one extra location which represents the rest of the world (RoW). We denote the set of all Chinese locations **G**. There are two sectors k in the economy: manufacture (M) and service (S).[17]

Each location (i) is endowed with a fixed factor land N_i, capital K_i, and labor L_i which is frictionally mobile across regions.[18] We assume that labor is frictionally mobile across locations within China but there's no labor mobility between China and the rest of the world. Land endowment in manufacturing and service sector is exogenous given. We will use $i,\ j \in \mathbf{G} \cup \mathbf{RoW}$ to denote the origin and destination of trade and migration flows.[19]

3.2 Workers

3.2.1 Preferences

Workers' preferences are defined over consumptions of manufacture (M) and service (S) goods and are given by :

$$u_i = \left[\left(\rho^M\right)^{\frac{1}{\sigma}} \left(C_i^M\right)^{\frac{\sigma-1}{\sigma}} + \left(\rho^S\right)^{\frac{1}{\sigma}} \left(C_i^S\right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{3}$$

where $\rho^k (k \in \{M, S\})$ is the preference share parameter, $C_i^k (k \in \{M, S\})$ is the consumption of sector-k composite goods in location i. ρ^k are positive and sum to one across sectors. We

[17] As mentioned in the introduction, the main focus of the paper is to look at the structural transformation within the urban regions. Thus, in the model we are not including the agriculture sector. This is also consistent with the empirical part focusing on the manufacturing labor share within the urban employment.

[18] Here we only model one type of labor and not distinguish skill intensity in both the labor input and in various sectors. In section 4.6, we discuss this issue and illustrate that this will not change the paper's main finding of the hump-shape pattern.

[19] For migration flows, $i,\ j \in \mathbf{G}$ since labor is immobile across countries and only frictionally mobile within China by assumption.

assume inelastic elasticity of substitution ($0 < \sigma < 1$) following the literature in macroeconomics.[20] Demand for each sector's goods in location i is

$$C_i^k = \rho^k \frac{(P_i^k)^{-\sigma}}{P_i^{1-\sigma}} \nu_i \qquad \text{for} \quad k \in \{M, S\}, \tag{4}$$

where P_i^k is the price index of sector k goods in location i, P_i is the price index in location i, and ν_i is the nominal income in location i. The price index, P_i is then given by:

$$P_i = \left[\rho^M (P_i^M)^{1-\sigma} + \rho^S (P_i^S)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}.$$

3.2.2 Nominal and Real Income

Given each worker's wage is w_i, total labor income in each location i is $w_i L_i$. In addition to labor income, all payments to land in a given region are rebated to the workers of that region. Hence, the worker's total income comes from both labor income and land rents, which implies

$$\nu_i L_i = w_i L_i + r_i N_i, \tag{5}$$

where ν_i is the nominal income in location i, N_i is the fixed land endowment in location i and r_i is the land rent in location i.

3.2.3 Labor Migration

Labor is mobile within China, across regions and sectors but immobile between home country and the rest of the world.[21] Workers are registered to regions and migration is modeled as a once-for-life choice. Let π_{ij} denote the share of workers in region i who moved to region j to work.

We model the heterogeneity in the utility that workers obtain from living in different regions following Ahlfeldt et al. (2015). Upon birth, workers learn about their idiosyncratic

[20]See Ngai and Pissarides (2007), Herrendorf et al. (2013) and Uy et al. (2013). If $\sigma < 1$, the sectoral composite goods are complements. In the Cobb-Douglas case when σ =1, there is no structural change since sectoral expenditure and labor shares are equal and constant by the feature of Cobb-Douglas functional form. We assume $\sigma < 1$ here which is empirically more relevant.

[21]The focus of our paper is the structural transformation within the urban regions, so we mainly model the inter-provincial migration here. Admittedly, the rural-urban migration is critical in the period we study, so we further extend our model by adding the rural-urban migration in the Appendix E.2 and test whether adding rural-urban migration will change the major results of our paper. We model the rural-urban migration as the increase in urban labor supply. Although the slope of the hump-shape pattern changes a bit, we find that the hump-shape pattern itself remains robust in Figure A.9.

taste of living across different regions and decide where to work, taking into account their destination's specific component in direct utility of destination j, V_j, as well as the migration cost. We model these costs as discounted from income, where a worker from region i loses a fraction $1-1/\mu_{ij}$ of their income in region i.[22] We model migration is costly across regions. Once the worker is working in a particular region, he can freely choose which specific sector to work in.

Workers have idiosyncratic preference draws that vary by region and these draws create differences in worker migration incentives. Formally, worker ω's idiosyncratic preference draws $\{z_j(\omega)\}$ for each of the N regions - these are i.i.d. across workers, and regions.[23] Therefore, given the worker's preference draw, $\{z_j(\omega)\}$, worker chooses the destination j to maximize welfare

$$\max_{j \in \mathbf{G}} \left\{ \frac{z_j(\omega)\nu_i}{\mu_{ij}} \right\}, \tag{6}$$

where ν_i is the amenity-adjusted real wage rate in region j defined in Equation (5). Within this structure, we derive the migration flows. As $z_j(\omega)$ is a random variable across the continuum of individuals, the law of large numbers will ensure that the proportion of these workers who migrate to region j is

$$H_{ij} = Pr\left(\frac{z_j(\omega)\nu_j}{\mu_{ij}} \geq \frac{z_m(\omega)\nu_m}{\mu_{im}}, \forall \ m \in \mathbf{G} \right).$$

Specifically, assume that the idiosyncratic preference follows with Fréchet distribution F

$$F(z_j | j \in \mathbf{G}) = e^{-(z_j\tilde{\gamma})^{-\varepsilon}}, \tag{7}$$

where ε governs the degree of dispersion across individuals. A large ε means smaller dispersion. The parameter $\tilde{\gamma} = \Gamma(1-\varepsilon^{-1})$ is a normalizing constant so that the mean of z_j is one. Here Γ is the Gamma function.

Proposition 1. *Given the real income for each region ν_j, migration costs between all regions μ_{ij}, and*

[22] We model the migration cost as variable cost for simplicity, while in reality there are both fixed and variable cost when migrating across cities.

[23] The parametric assumption on distribution is also used by Hsieh et al. (2013), Ahlfeldt et al. (2015), and Bryan and Morten (2015).

heterogeneous preference distribution $F(z_j)$, the share of region i workers that migrate to region j is

$$H_{ij} = \frac{\left(\frac{\nu_j}{\mu_{ij}}\right)^{\varepsilon}}{\sum_{m\in G}\left(\frac{\nu_m}{\mu_{im}}\right)^{\varepsilon} dm}. \tag{8}$$

Proof. See Appendix.

Therefore the labor supply in each region i is

$$L_j = \sum_{i\in\mathbf{G}} H_{ij}\bar{L}_i dj. \tag{9}$$

Hence, we derive the expected utility of worker who is originally from region i

$$E(u_i) = \gamma\left(\sum_{m\in\mathbf{G}} (V_m/\mu_{im})^{\varepsilon}\right)^{\frac{1}{\varepsilon}}, \tag{10}$$

which measures the welfare of workers from location i. The more connected location i is to the labor markets of other region (smaller $\mu_{im}, \forall m \in \mathbf{G}$) and the more attractive the nearby locations are (greater $V_m, \forall m \in \mathbf{G}$), the higher utility the worker from location i will benefit. Note that, the expected utility does not depend on the destination location j for workers from the same region whose average welfare will be the same regardless of the location where they live. On the one hand, more attractive destination characteristics directly raise the welfare of a worker given his or her idiosyncratic taste draw, which increases the expected utility. On the other hand, more attractive destination characteristics attract workers with lower idiosyncratic taste draws, which reduces the average utility. With a Fréchet distribution of taste shocks, these two effects just cancel out on another for workers from the same place, which only depends on the characteristics of the original region. It is straightforward to illustrate the aggregate welfare is

$$W = \sum_{i\in\mathbf{G}} \gamma\left(\sum_{m\in\mathbf{G}} (V_m/\mu_{im})^{\varepsilon}\right)^{\frac{1}{\varepsilon}}. \tag{11}$$

Each worker is endowed with one unit of labor that is supplied inelastically with zero disutility. The total number of population in home country is $\bar{L}$ and satisfy:

$$\sum_{i=1}^{N} L_i = \bar{L}. \tag{12}$$

3.3 Production and Trade

The production side of the economy is a multi-sectoral version of Eaton and Kortum (2002). There are two production industries in the economy: manufacturing (M) and service(S).

3.3.1 Production

There's a continuum of variety z in each sector. Each country and each location consists of technologies for producing all the goods in all sectors. The composite good in each sector y_i^k is an aggregate of the individual goods $y_i^k(z)$,

$$y_i^k = \left[\int_0^1 y_i^k(z)^{\frac{\eta-1}{\eta}} dz \right]^{\frac{\eta}{\eta-1}}, \tag{13}$$

where the elasticity of substitution across goods within a sector is $\eta > 0$ and $y_i^k(z)$ is the quantity of variety z. For the service sector, each good z is produced locally; for the manufacture sector, each good is either produced locally or imported from other locations (within country or from abroad). The composite sectoral goods are directly used in the final consumption C_i^k.

Production is perfect competition in each sector and takes the Cobb-Douglas form. The production function for variety z at sector k is

$$y_i^k(z) = \varphi_i^k(z) \left(L_i^k(z) \right)^{\alpha^k} \left(K_i^k(z) \right)^{\beta^k} \left(N_i^k(z) \right)^{1-\alpha^k-\beta^k}.$$

Here, $y_i^k(z)$ is the output, $\varphi_i^k(z)$ is the productivity, $L_i^k(z)$ is the labor used for variety z in sector k in location i, $K_i^k(z)$ is the capital used and $N_i^k(z)$ is the land used. $0 < \alpha^k < 1$ is the labor intensity in sector k and $0 < \beta^k < 1$ is the capital intensity in sector k. Capital is freely tradable. Capital is provided competitively by absentee capitalists. The price of capital is fixed exogenously in international markets, and the stock of capital in the country adjusts to the demand of firms.[24]

The productivity $\varphi_i^k(z)$ is allowed to differ across locations i. The realizations of productivity $\varphi_i^k(z)$ is following Frechet distribution

$$P(\varphi_i^k(z) \leq \varphi) = e^{-T_i^k \varphi^{-\theta}} \qquad \text{for} \quad \forall \quad z, \tag{14}$$

where T_i^k controls the average productivity in sector k of location i for any variety z and θ controls the variance of the productivity.

[24]Featuring this input to the production function allows the model to capture the various input intensity which is one of the determinants of structural changes.

It is perfect competition in each sector. In each sector, firms choose labor, capital and land to maximize their profits given goods and factor prices and the location decisions of other firms and workers. Firms make zero profits in each location and each sector with positive production. The firm's profit maximization problem becomes the following:

$$\max_{L_i^k, K_i^k, N_i^k} \left\{ P_i^k y_i^k(z) - w_i L_i^k(z) - q K_i^k(z) - r_i^k N_i^k(z) \right\} \quad \text{for} \quad k \in \{M, S\}, \tag{15}$$

where w_i is the wage, q is the capital rent and r_i^k is the land rent. Especially, we allow land rents differ across sectors and thus the city's land rents in various sectors are determined by the exogenous sectors land allocations between manufacturing and service in different cities.[25]

Hence, with constant return to scale, the cost of producing one unit of good z in location i is given by

$$\omega_i^k(z) = \chi^k w_i^{\alpha^k} q^{\beta^k} (r_i^k)^{(1-\alpha^k-\beta^k)} / \varphi_i^k(z), \tag{16}$$

where χ^k is a constant: $\chi^k = (\alpha^k)^{-\alpha^k} (\beta^k)^{-\beta^k} (1 - \alpha^k - \beta^k)^{-(1-\alpha^k-\beta^k)}$.

3.3.2 Trade

Goods in manufacture industry (M) are tradable both domestically and internationally while varieties in the service sector (S) are non-tradable.[26] We model the iceberg trade cost as the literature: τ_{ij} is the iceberg trade cost from location i to location j. A firm producing good z in sector k of location i with productivity $\varphi_i^k(z)$ charges a buyer in location j, $p_{ij}^k(z) = \tau_{ij}\omega_i^k(z)$, where $\tau_{ij} \geq 1$ is an iceberg trade cost and $\omega_i^k(z)$ is the unit cost.

As in Eaton and Kortum (2002), the price of good z that is actually paid by the producer at location j is the lowest across all sources i,

$$p_j^k(z) = \min_i \left\{ p_{ij}^k(z) \right\}. \tag{17}$$

Therefore, we can directly show that among the expenditure spent on goods in location j,

[25] Here, we are abstract from modelling the land market distortions in China. In practice, land market distortions is an important issue to consider due to the regulation and corruptions in China's land market. We discuss how the land market distortions will not change our main findings of the hump-shape pattern in Appendix E.3.

[26] In the following, to simplify notations, we assume trade costs in the service sector are infinite and proceed as if services were tradable.

sector k, the share allocated to varieties produced in location i is

$$\pi_{ij}^k = \frac{T_i^k \left[\tau_{ij} w_i^{\alpha^k} (r_i^k)^{(1-\alpha^k-\beta^k)}\right]^{-\theta}}{\sum_m T_m^k \left[\tau_{mj} w_m^{\alpha^k} (r_m^k)^{(1-\alpha^k-\beta^k)}\right]^{-\theta}}. \tag{18}$$

The price index for each sector is given by

$$P_j^k = \gamma \left[\sum_m T_m^k (\tau_{mj} \chi^k w_m^{\alpha^k} q^{\beta^k} (r_m^k)^{(1-\alpha^k-\beta^k)})^{-\theta}\right]^{-1/\theta}, \tag{19}$$

where $\gamma = \left[\Gamma(1 + \frac{1-\eta}{\theta})\right]^{1/(1-\eta)}$ and $\eta < 1 + \theta$.[27]

3.4 Goods and Factor Markets Clearing Conditions

We now write down the goods and factor market clearing conditions to close the model. To start, we solve for the equilibrium of goods market, taken the labor costs as given, and then turn to the labor market equilibrium.

3.4.1 Goods Markets Clearing Conditions

Denote R_i^k the total revenue for goods produced in location i sector k. The total revenue for manufacture products produced in location i equals total sales to buyers in all locations:

$$R_i^M = \sum_j \pi_{ij}^M X_j^M, \tag{20}$$

where X_j^M is the total expenditure in manufacture sector in location j. In order to make varieties in sector k in region i, the producers in region i needs to hire workers and produce using Cobb-Douglas production function. Hence

$$R_i^k = \frac{w_i L_i^k}{\alpha^k}, \tag{21}$$

where L_i^k is the labor hired in location i working in sector k.

For the service sector, since it is nontradable, the revenue only comes from local buyers:

$$R_i^S = X_i^S. \tag{22}$$

[27] We need to assume $\eta < 1 + \theta$ to have a well-defined price index.

The total expenditure in location j sector k satisfy:

$$X_j^k = P_j^k C_j^k, \quad k \in \{M, S\}. \tag{23}$$

3.4.2 Factor Markets Clearing Conditions

The local labor market clearing conditions are the following:

$$L_i = L_i^M + L_i^S. \tag{24}$$

We now solve for the aggregate equilibrium given the structures described above. We define the competitive equilibrium of the economy in Appendix C.3.

4 Quantitative Analysis

4.1 Parameters

To provide intuition of the model to match the empirical finding, we illustrate the workings of the model with the quantitative analysis.

We assume that there are multiple locations in the home country and the home country and foreign country are relatively the same size in terms of labor, and land. Home country has comparative advantage in producing manufacture goods compared to the foreign country while in service sector, they have the same comparative advantage. The elasticity of substitution across different sectors is 0.5 following Ngai and Pissarides (2007). The elasticity of substitution across goods within a sector is set at 3 following Hsieh and Klenow (2009). The Fréchet distribution parameter for the migrants' idiosyncratic preference shocks ε is set as 1.5 following Tombe and Zhu (2019)[28]. Trade elasticity is set at 4 following Simonovska and Waugh (2014). We set the labor and capital share parameters following Valentinyi and Herrendorf (2008). To calibrate the land allocations between manufacturing and service sectors in different cities, we use the data from China Urban Construction Statistical Yearbook and the details are in Appendix D.1. The parameter table is in Table 2.[29]

[28] We set the parameter of Frechet distribution T_i by assumptions since the parameter is country-specific but we actually do not have one specific foreign country in our model – the foreign country in our model is the rest of the world. So we test whether our baseline results still hold under different assumptions of T_i in Figure A.7. We find that our main result is persistent and does not depend on the T_i's assumptions.

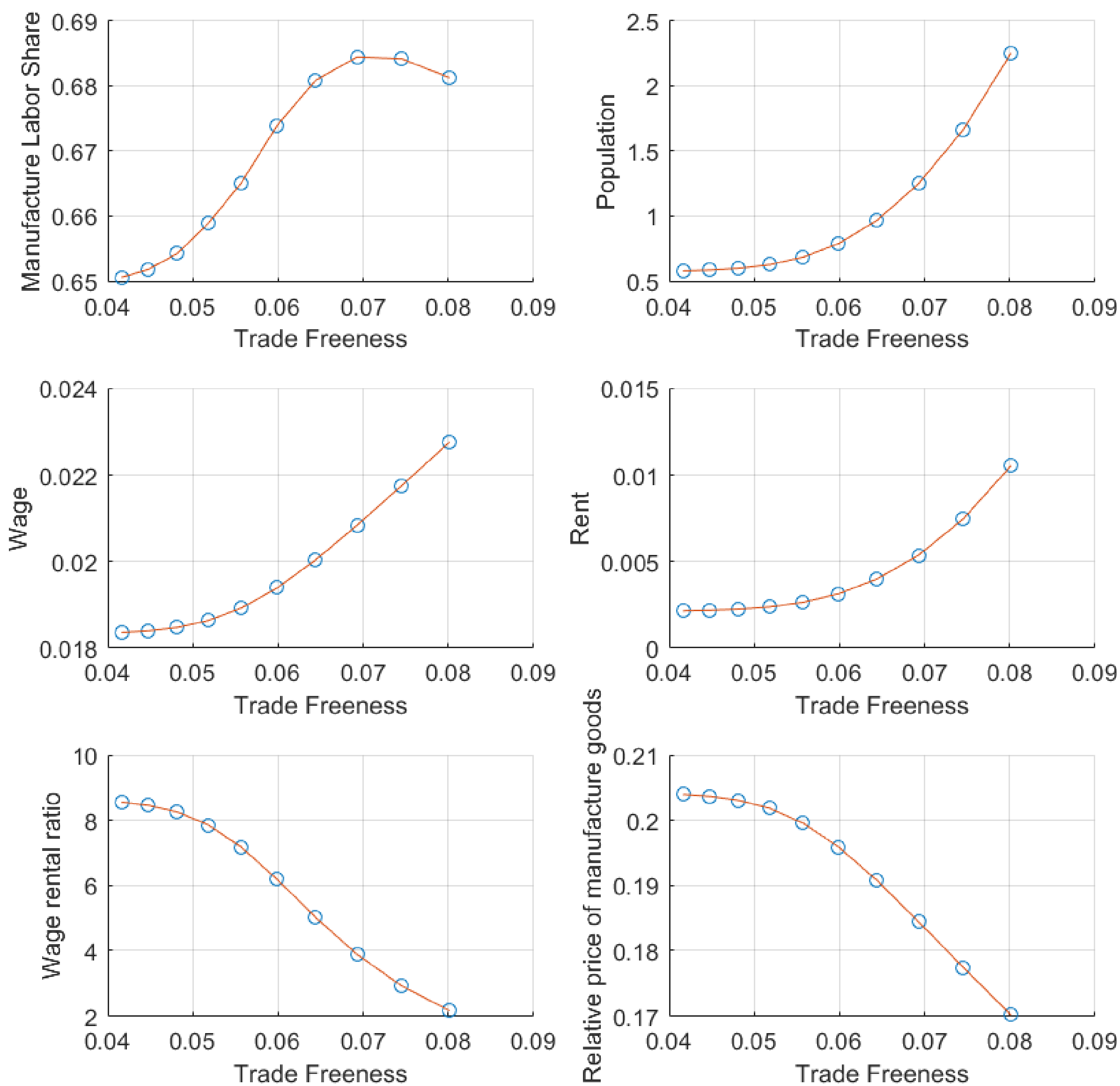
[29] We set homogeneous migration cost. Migration cost is set at 1.5 referring to Tombe and Zhu (2019), and Fan (2019) to capture the average migration cost in China. Intra-national trade cost is referring to the same literature to capture the average trade cost in China. We set homogeneous trade cost within each region. International trade cost differs across locations in the home country. For more details, refer to Appendix D.3.

Table 2: Parameters for the Numerical Analysis

Parameter	Definition	Value	Source/Target
n	Number of regions within the home country	10	Assumption
Consumer Parameters			
η	Elasticity of substitution among continuum of goods	3	Hsieh and Klenow (2009)
σ	Elasticity of substitution between manufacture and service sector	0.5	Ngai and Pissarides (2007)
Migration Parameters			
ε	Taste dispersions of migrant workers	1.5	Tombe and Zhu (2019)
μ_{ij}	Migration cost of moving from i to j	Data	Tombe and Zhu (2019)
Production Parameters			
θ	Trade elasticity	4	Simonovska and Waugh (2014)
T^M	Parameter of Frechet distribution in manufacture	10 (for home), 1 (for foreign)	Assumption
T^S	Parameter of Frechet distribution in service	1	Assumption
α^M, α^S	Labor intensity in manufacture, service sector	0.57, 0.64	Valentinyi and Herrendorf (2008)
β^M, β^S	Capital intensity in manufacture, service sector	0.4, 0.3	Valentinyi and Herrendorf (2008)
Endowment			
L	Labor endowment in home and foreign country endowment	Lh=10, Lf=10	Normalization
N	Land endowment in home and foreign country	Nf =10, Nh=1*10	Normalization
N_i^M, N_i^S	Regional land endowment in manufacturing, service sector in home country	See Appendix D.1	Data

Notes: This table summarizes the parameters used in the numerical analysis.

Figure 4: Numerical Baseline Result for the Home Country



Notes: The Figure shows the numerical baseline result for the home country of our model. The baseline case generates the hump-shaped pattern between the manufacture labor share and trade openness. Trade freeness is defined as $\tau^{-\theta} \in (0, 1)$, and τ is the geometric mean across all the destinations. The more detailed discussions and meaning of trade freeness is in Appendix D.3.

4.2 Numerical Result

Figure 4 illustrates the numerical result for the home country. The top left panel shows the change of manufacture labor share allocation with trade freeness for the home country and it matches the hump-shaped pattern which is consistent with empirical findings in Section 2. We refer this to the baseline model. Other panels also show that population is increasing with trade freeness, wages and rents are higher in locations that are more open to trade. However, wage-rental ratio is lower and relative price of manufacture goods are also lower. The intuition is that areas that are more open to trade will benefit from trade and receive higher wages, which attracts labor to come after they observe higher wages. However, land supply is fixed and this leads to higher rents and lower wage rental ratio. In order to further explain the hump-shape pattern of the manufacture labor share, we will explain this hump-shape pattern by looking at two extreme case: the rising part and the falling part. Each part is explained by a specific channel.

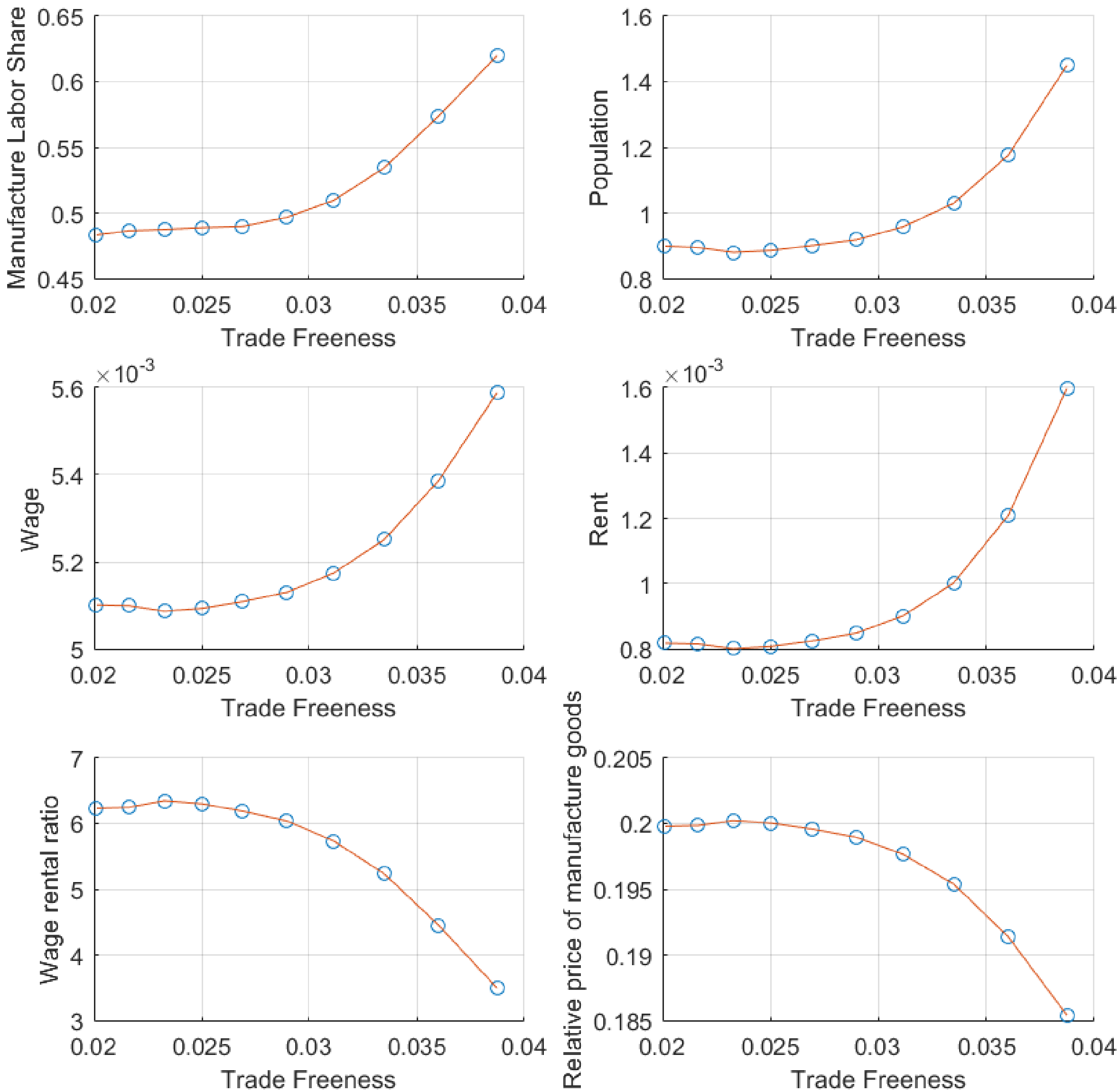
Market access channel. We first demonstrate the rising part by increasing the international trade cost and the numerical result is in Figure 5. In this extreme case, international trade costs are large enough while the trade freeness are small enough. International trade cost can also be viewed as the distance to the rest of the world. Hence if the international trade cost is large enough, which means that the home country is far away from the rest of the world, we could view the country as a point while neglecting the internal factor reallocation and geographic variations. Hence, we are back to the simple two-country case(Uy et al. (2013)). If the trade cost is lower, the manufacture (tradable goods) share increases due to better market access.[30] Other variables change in the similar way as the baseline case shown in Figure 4.

Internal factor reallocation channel. We then show the falling part by further decreasing the international trade cost and the numerical example is in Figure 6. When international trade cost is low enough, we cannot view the home country as a point anymore and the internal factor reallocation matters (Fajgelbaum and Redding (2014)). The intuition is that when labor is mobile, labor will migrant to the coastal areas where the international trade cost is lower and wage is higher. Similar to the Rybczynski theorem, a rise in the labor input will lead to the expansion in the relatively labor-intensive sector, which is the service sector in this case and result in the shrink in the other sector, manufacture. Other variables also follow the same pattern as Figure 4. Therefore, the internal factor reallocation channel explains the falling part.[31]

[30] We also formally prove this in a simplified model case in the appendix C.2.

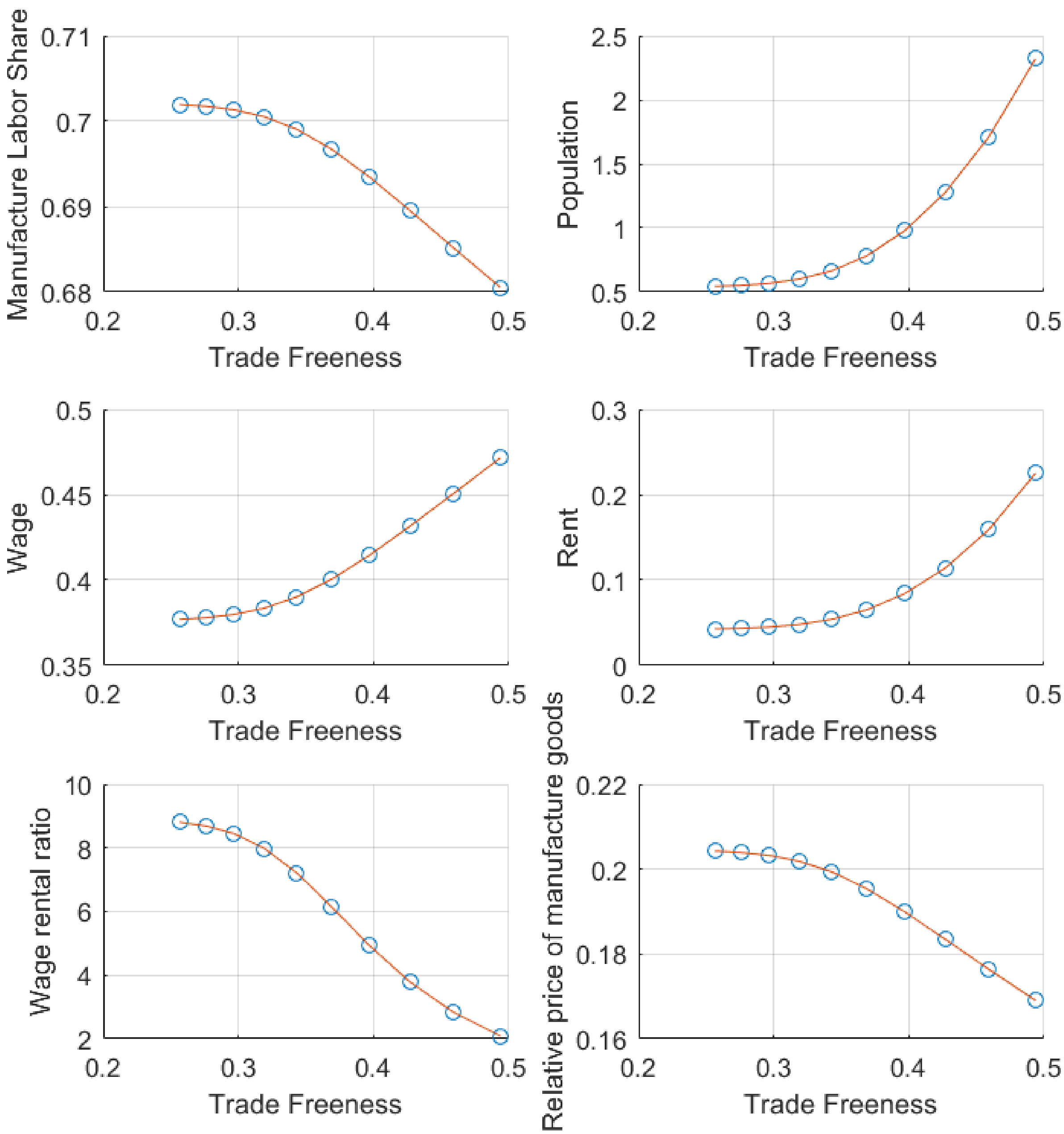
[31] This channel is also proposed by Fajgelbaum and Redding (2014) and they names this as "Spatial Balassa-Samuelsam" effect.

Figure 5: Numerical Extreme Case Result: High International Trade Cost Scenario



Notes: The Figure shows the numerical extreme case result for the home country of our model when the international trade cost is high. The scenario generates the rising part of the hump-shaped pattern in Figure 4 and shows the market access channel of our model. Trade freeness is defined as $\tau^{-\theta} \in (0, 1)$, and τ is the geometric mean across all the destinations. The more detailed discussions and meaning of trade freeness is in Appendix D.3.

Figure 6: Numerical Extreme Case Result: Low International Trade Cost Scenario



Notes: The Figure shows the numerical extreme case result for the home country of our model when the international trade cost is low. The scenario generates the falling part of the hump-shaped pattern in Figure 4 and shows the internal factor reallocation channel of our model. Trade freeness is defined as $\tau^{-\theta} \in (0, 1)$, and τ is the geometric mean across all destination. The more detailed discussions and meaning of trade freeness is in Appendix D.3.

There exists a cut-off point of trade freeness, after which the internal factor reallocation channel outweighs the market access channel generating the falling part of the manufacture labor share. This is different from the two-country model in Uy et al. (2013) which finds that the falling part of the manufacture share comes from the complete specialization of producing manufacture goods. In this model, we generate the falling part of the manufacture labor share without assuming the complete specialization of the manufacture goods. As long as the international trade cost decreases, and the internal factor like labor is mobile, the internal factors will reallocate to produce the hump-shape pattern of the manufacture labor share.

4.3 Channel Decomposition

In the previous section, we plot the two extreme cases by changing the level of trade freeness to show the two channels separately. In this section, we would like to split the two channels within the hump-shape pattern and show the relative importance of the two channels at different points of the hump-shape pattern.

To do so, we use two measures to proxy for the two channels in each city and decompose the hump-shape pattern into two channels. Each city is equipped with various levels of trade costs and migration flows from other cities in the model. So we could use the cross-sectional city variations to illustrate the relative importance of the two channels in the hump-shape relationship. For the market access channel, since the city's access to the international market would matter, we use each city's international trade cost to proxy the channel's magnitude,

$$MA_j = \tau_{jf},$$

where j is each city, and f is the rest of the world. For the internal factor reallocation channel, since land is a fixed factor, we use the summation of population migration flows to each city to proxy the channel's magnitude,

$$IFR_j = \sum_{i \in \mathbf{G}} H_{ij} \bar{L}_i = \sum_{i \in \mathbf{G}} \frac{\left(\frac{\nu_j}{\mu_{ij}}\right)^{\varepsilon}}{\sum_{m \in \mathbf{G}} \left(\frac{\nu_m}{\mu_{im}}\right)^{\varepsilon} dm} \bar{L}_i.$$

In Table 3, we separate the hump-shape pattern into the rising and falling part and then report each channel's average contribution to manufacturing labor share changes. We find that the market access channel dominates the internal factor reallocation channel initially at the rising part of the hump-shape. As the trade freeness increases, the internal factor reallocation channel outweighs the market access channel. Thus, the two channels' contribution

Table 3: Decomposing hump-shape pattern between manufacturing share and trade freeness

Contribution to manufacturing labor share changes		
	Rising Part	Falling Part
Market Access Channel	78%	10%
Internal Factor Reallocation Channel	22%	90%

Notes: This table decomposes the two channels' contribution to the manufacturing labor share at the rising and falling part of the hump-shape. We divide the hump-shape into rising and falling part according to the highest point of the hump-shape.

to the manufacturing labor share changes as the trade freeness varies.

4.4 Model Extensions

In this section, we would like to discuss a few model extensions related to the major results of our numerical analysis.

First, in the model, we assume the service sector is nontradable across countries and across regions within a country, which is a standard assumption in traditional trade models. But empirically, we can see a rising share of service trade across regions from China's regional input-output table (Mi et al., 2018). So we extend our model by allowing service trade within regions in the domestic country and rerun the analysis. Figure A.8 shows that the hump-shape pattern remains robust but the shape becomes flatter than before.[32]

Second, in the paper, we mainly focus on the structural transformation from manufacturing to the service sector within the urban areas. However, the study period 1998-2008 also features renowned and rapid rural-urban migration in China. Therefore, we extend our model by allowing rural-urban migration, and in Figure A.9, we find that the hump-shape pattern remains robust, but the slope is flatter. Adding the rural-urban migration will facilitate the internal factor reallocation channel, which makes the hump-shape slope becomes flatter.[33] The intuition is that adding the rural-urban migration will facilitate the internal factor reallocation channel, therefore dampens the market access channel in the rising part of the hump-shape, which makes the hump-shape's slope becomes flatter.

[32]When the service sector is non-tradable, international trade has a more critical impact on the tradable manufacturing sector. It acts as the primary driving force for the structural transformation in various cities. When the service sector is allowed to trade across regions, it will absorb part of the international trade effects since the factor reallocation through service trade is also easier. Thus, adding the service trade would make international trade's impact diminished and make the hump-shape pattern flatter compared to the baseline case.

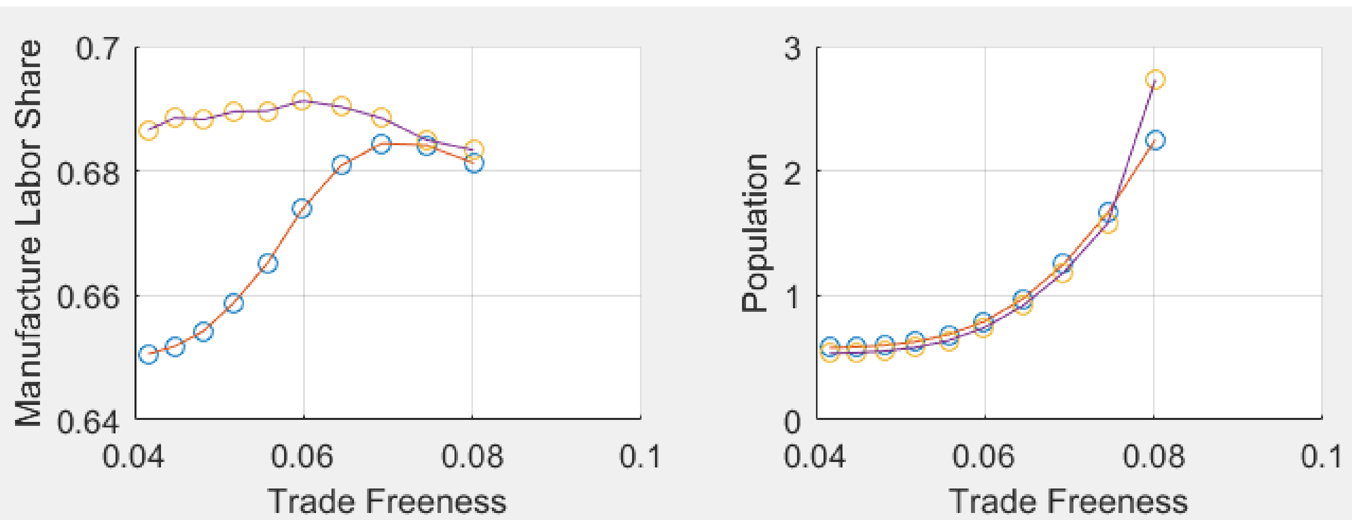
[33]More discussions and details are in Appendix E.2.

4.5 Counterfactual Analysis

In this section, we use the structural of the model to undertake an simple counterfactual exercise to show the impact of land supply on the distribution of economic activity.

China has unique and strict urban land supply policy[34]. To ensure the adequate food production and supply, the government has identified a minimum threshold of "redline" of 120 million hectares of cultivated land. This policy has recently received critiques in terms of the government intervention of free market and land supply restrictions in the urban areas. With the process of urbanization and development, population is increasing in big cities with rising land demand while the land supply is still restricted by the "redline" policy. This imbalance of land demand and supply has given rise to many urban issues like roaring urban housing prices and the inefficiency in the agriculture sector as well.

Figure 7: The Numerical Result of Counterfactual Exercise



Notes: The countefactual exercise here is to increase the largest city's land supply by 50%. The Figure shows the counterfactual exercise result for the home country of our model when the international trade cost is low. Trade freeness is defined as $\tau^{-\theta} \in (0, 1)$, and τ is the geometric mean across all destination. The yellow dots are the counterfactual result and the blue dots are the baseline results. The more detailed discussions and meaning of trade freeness is in Appendix D.3.

Thus, we are interested in the changing the land supply to observe the counterfactual impact. Based on the baseline model[35], we take the counterfactual by increasing the biggest

[34] In China, the land is owned by government.

[35] Since for simplicity, we model the urban part of an economy by assuming manufacture and service take place in urban areas while agriculture is in rural areas. Hence, the land in the model can be referred as the urban land.

cities' urban land supply by 50% assuming that government could release some rural area land and turn them into the urban land supply to satisfy the growing urban land demand.[36] Comparing with the baseline model result, this increases the home country welfare by 1% and the rent in the largest city falls by 19%. And the structural change pattern is shown in Figure 7. The hump-shape of the manufacture labor share still exists but is less sharp compared to the baseline model. The population is higher in the largest city since it is able to accommodate more people.

The counterfactual results could indicate a direction for the benefits and improved welfare of the land reforms. The actual magnitude could only provide a rough estimate of the possible revenue from such a land market provision project but should be taken into more careful considerations. To fully and properly measure the cost and benefit of the urban land provision, it requires additional detailed data on the government's land transaction and fiscal balance sheets. On the modeling side, it also requires a government sector that collects the land revenue from the urban land provision and bears the land concession cost at the same time. The government would redistribute the additional profits to the residents after collecting the revenue. All this information seems pretty scarce on the table and may also be out of this paper's scope. And it would be a fascinating topic for future research.[37]

4.6 Discussion of Limitations

While our model manages to capture the stylized empirical fact of the hump-shape pattern between the trade openness and manufacturing labor share, because of the model simplification, there are some issues in which one should interpret our results with caution. The first concern is the land market distortions. In China, the land market is highly regulated and distorted, which could pose a threat to affect our main results. The relatively cheap industrial land in China (Xu et al., 2017) might lead to the overestimation of the land price ratio of manufacturing land to service land in the baseline model. If the actual manufacturing land price is relatively cheaper compared to the service sector, this could facilitate the development of the manufacturing sector, which lead to a steeper hump-shaped pattern between the manufacturing share and the trade freeness. But we argue that this will not change the hump-shape pattern itself. [38]

Another concern is the heterogeneity of skill intensity in various sectors. We do not incorporate human capital input in the baseline model, but big cities could specialize in

[36] Though the counterfactual exercise is not practical in reality but it shows a direction of the impact of land supply changes on the housing market and overall welfare in contemporary China.

[37] We discuss this issue in more details in Appendix E.4.

[38] For more details, please refer to Appendix E.3.

human capital intensive service sectors. We do not add the skill intensity heterogeneity in various sectors because this will not change the hump shape pattern itself between trade freeness and manufacturing share. But it will change the slope of the hump-shape pattern, which depends on the type of trade shock. For instance, if the trade shock is skill-intensive (Li, 2018), the trade shock will lead to higher demand for the skill-intensive manufacturing sector, i.e., human capital intensive sectors. More skilled labor will migrate to the coastal areas to work in the skill-intensive manufacturing sector. Skill premiums in the coastal areas will increase due to a higher demand for skilled labor, which also is beneficial for the service sector since the service sector is more skill-intensive than the manufacturing sector (Buera and Kaboski, 2012; Blum, 2008). The rising inflow of skilled labor accelerates the generation of the hump-shaped patterns, i.e., the hump-shape will be steeper. [39]

5 Conclusion

This paper examines the role of external integration on the structural transformation process. Focusing on China, we provide a direct empirical evidence on the role of trade liberalization on the manufacture labor share by applying the cross-sectional data. We find that there exists a nonlinear, hump-shaped pattern between trade openness and manufacture labor share. The rising part of this linkage is explained by improving market access while the falling part comes from the internal factor reallocation. Less remote areas have higher population density, lower wage-rental ratio, and therefore, specialize in a relatively more labor-intensive sector. Counterfactual exercise shows that a possible direction of improving aggregate welfare is to exploit and increase urban land supply in contrast to the limited and restrictive land policy in contemporary China.

While the determinants of structural changes have a long and extensive literature in macroeconomics, this paper shows that the internal geography is also critical while considering trade liberalization's impact on the structural changes. Within the role of trade liberalization in influencing the structure change, we emphasize that the internal reallocations of resources across regions and sectors within one country are also central to the economic development. Thus, the policy should be place-based and the migration, trade and land policy need to make corresponding changes according to each region's stage of structural transformation.

[39] On the contrary, if the trade shock is less skill-intensive, i.e., more labor-intensive, the trade shock will result in higher demand for unskilled labor in the coastal areas. More unskilled labor will migrate to the coastal regions to specialize in the labor-intensive manufacturing sector. Skill premium in the coastal regions will decrease due to the higher demand for unskilled labor. Thus the rising inflows of the unskilled labor will dampen the service sector, and the hump shape will be flatter.

Acknowledgment

I am greatly indebted to the editor, two anonymous referees who help me improve the paper. I also thank Colin Carter, Ben Faber, Thibault Fally, Li Fang, Xiaoyi Han, Anna Ignatenko, Ding Lu, Ming Lu, Ziho Park, Sarah Quincy, Andrés Rodríguez-Clare, Luis Felipe Saenz, Yongseok Shin, Yong Wang, Wing Woo and Yingyan Zhao for their helpful comments and discussions, as well as seminar and conference participants at UC Davis, 6th International Workshop on Regional, Urban and Spatial Economics in China, 3rd Biennial Conference of China Development Studies, 12th Economics Graduate Students Conference at Washington University in St. Louis., 5th International Annual Conference of NSE at PKU. I'm also extremely grateful to my advisers, Robert C. Feenstra, Christopher M. Meissner, Giovanni Peri, Ina Simonovska for their guidance and support. Financial support from UC Davis summer research travel grant is gratefully acknowledged. Please direct correspondence to: Xuan Fei, State Administration of Foreign Exchange Investment Center (SAFEIC), Beijing, China; email: xuafei@ucdavis.edu. The views expressed herein are those of the author and do not necessarily reflect the views of SAFEIC of the People's Republic of China. All errors are mine.

References

AHLFELDT, G. M., S. J. REDDING, D. M. STURM, AND N. WOLF (2015): "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 83, 2127–2189.

ALLEN, T. AND C. ARKOLAKIS (2013): "Trade and the Topography of the Spatial Economy," Tech. rep., National Bureau of Economic Research.

ATKIN, D. AND D. DONALDSON (2015): "Who's Getting Globalized? The Size and Implications of Intranational Trade Costs," Tech. rep., National Bureau of Economic Research.

BETTS, C., R. GIRI, AND R. VERMA (2013): "Trade, reform, and structural transformation in South Korea," *CAFE Research Paper*.

BLUM, B. S. (2008): "Trade, technology, and the rise of the service sector: The effects on US wage inequality," *Journal of International Economics*, 74, 441–458.

BOMBARDINI, M. AND B. LI (2020): "Trade, pollution and mortality in china," *Journal of International Economics*, 103321.

BRYAN, G. AND M. MORTEN (2015): "Economic development and the spatial allocation of labor: Evidence from Indonesia," *Manuscript, London School of Economics and Stanford University*, 1671–1748.

BUERA, F. J. AND J. P. KABOSKI (2012): "The rise of the service economy," *American Economic Review*, 102, 2540–69.

CALIENDO, L. AND F. PARRO (2015): "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 82, 1–44.

CALIENDO, L., F. PARRO, E. ROSSI-HANSBERG, AND P.-D. SARTE (2014): "The impact of regional and sectoral productivity changes on the US economy," Tech. rep., National Bureau of Economic Research.

CASELLI, F. AND W. J. COLEMAN (2001): "The US structural transformation and regional convergence: A reinterpretation," *Journal of political Economy*, 109, 584–616.

CHEN, Q. AND Z. SONG (2014): "Accounting for China's urbanization," *China Economic Review*, 30, 485–494.

CHEN, Y., G. Z. JIN, AND Y. YUE (2010): "Peer migration in China," Tech. rep., National Bureau of Economic Research.

COŞAR, A. K. AND P. D. FAJGELBAUM (2016): "Internal geography, international trade, and regional specialization," *American Economic Journal: Microeconomics*, 8, 24–56.

DAVID, H., D. DORN, AND G. H. HANSON (2013): "The China syndrome: Local labor market effects of import competition in the United States," *The American Economic Review*, 103, 2121–2168.

DENG, Y., Y. TANG, P. WANG, AND J. WU (2020): "Spatial Misallocation in Chinese Housing and Land Markets," Tech. rep., National Bureau of Economic Research.

DESMET, K., D. K. NAGY, AND E. ROSSI-HANSBERG (2015): "The geography of development: Evaluating migration restrictions and coastal flooding," Tech. rep., National Bureau of Economic Research.

DESWAL, P. (2014): "International Trade Costs in Services," *Suffolk University Working Paper*.

EATON, J. AND S. KORTUM (2002): "Technology, geography, and trade," *Econometrica*, 70, 1741–1779.

ECHEVARRIA, C. (1995): "Agricultural development vs. industrialization: Effects of trade," *Canadian Journal of Economics*, 631–647.

FAJGELBAUM, P. AND S. J. REDDING (2014): "External integration, structural transformation and economic development: Evidence from argentina 1870-1914," Tech. rep., National Bureau of Economic Research.

FAN, J. (2019): "Internal geography, labor mobility, and the distributional impacts of trade," *American Economic Journal: Macroeconomics*, 11, 252–88.

FEI, X. (2020): "The Misallocation in the Chinese Land Market," *BOFIT Discussion Paper*.

FUJITA, M. AND P. KRUGMAN (1999): "A. Venables (1999), The Spatial Economy: Cities, Regions, and International Trade," .

GARRIGA, C., Y. TANG, P. WANG, ET AL. (2014): "Rural-Urban Migration, Structural Transformation, and Housing Markets in China," *Federal Reserve Bank of St. Louis Working Paper Series*.

HERRENDORF, B., R. ROGERSON, AND Á. VALENTINYI (2013): "Growth and structural transformation," Tech. rep., National Bureau of Economic Research.

HSIEH, C.-T., E. HURST, C. I. JONES, AND P. J. KLENOW (2013): "The allocation of talent and us economic growth," Tech. rep., National Bureau of Economic Research.

HSIEH, C.-T. AND P. J. KLENOW (2009): "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 124, 1403–1448.

KUZNETS, S. (1966): *Economic growth and structure*, Heinemann Educational.

LI, B. (2018): "Export expansion, skill acquisition and industry specialization: Evidence from China," *Journal of International Economics*, 114, 346–361.

LIND, J. T. AND H. MEHLUM (2010): "With or without U? the Appropriate Test for a U-Shaped Relationship," *Oxford bulletin of economics and statistics*, 72, 109–118.

MATSUYAMA, K. (2009): "Structural change in an interdependent world: A global view of manufacturing decline," *Journal of the European Economic Association*, 7, 478–486.

MI, Z., J. MENG, H. ZHENG, Y. SHAN, Y.-M. WEI, AND D. GUAN (2018): "A multi-regional input-output table mapping China's economic outputs and interdependencies in 2012," *Scientific data*, 5, 180155.

MIROUDOT, S., J. SAUVAGE, AND B. SHEPHERD (2013): "Measuring the cost of international trade in services," *World Trade Rev.*, 12, 719.

NGAI, L. R. AND C. A. PISSARIDES (2007): "Structural change in a multisector model of growth," *The American Economic Review*, 97, 429–443.

REDDING, S. J. (2012): "Goods trade, factor mobility and welfare," Tech. rep., National Bureau of Economic Research.

SIMONOVSKA, I. AND M. E. WAUGH (2014): "The elasticity of trade: Estimates and evidence," *Journal of international Economics*, 92, 34–50.

SWIECKI, T. ET AL. (2014): "Determinants of Structural Change," in *2014 Meeting Papers*, Society for Economic Dynamics, 870.

THE WORLD BANK (2014): *Urban China: Toward efficient, inclusive, and sustainable urbanization*, World Bank Group.

TOMBE, T. AND X. ZHU (2019): "Trade, migration, and productivity: A quantitative analysis of china," *American Economic Review*, 109, 1843–72.

TOPALOVA, P. (2010): "Factor immobility and regional impacts of trade liberalization: Evidence on poverty from India," *American Economic Journal: Applied Economics*, 2, 1–41.

UY, T., K.-M. YI, AND J. ZHANG (2013): "Structural change in an open economy," *Journal of Monetary Economics*, 60, 667–682.

VALENTINYI, A. AND B. HERRENDORF (2008): "Measuring factor income shares at the sectoral level," *Review of Economic Dynamics*, 11, 820–835.

WANG, Y. AND E. C.-M. HUI (2017): "Are local governments maximizing land revenue? Evidence from China," *China Economic Review*, 43, 196–215.

WU, Y., X. ZHANG, M. SKITMORE, Y. SONG, AND E. C. HUI (2014): "Industrial land price and its impact on urban growth: A Chinese case study," *Land Use Policy*, 36, 199–209.

XU, M. (2020): "Globalization, the skill premium, and income distribution: the role of selection into entrepreneurship," *Review of World Economics*, 1–36.

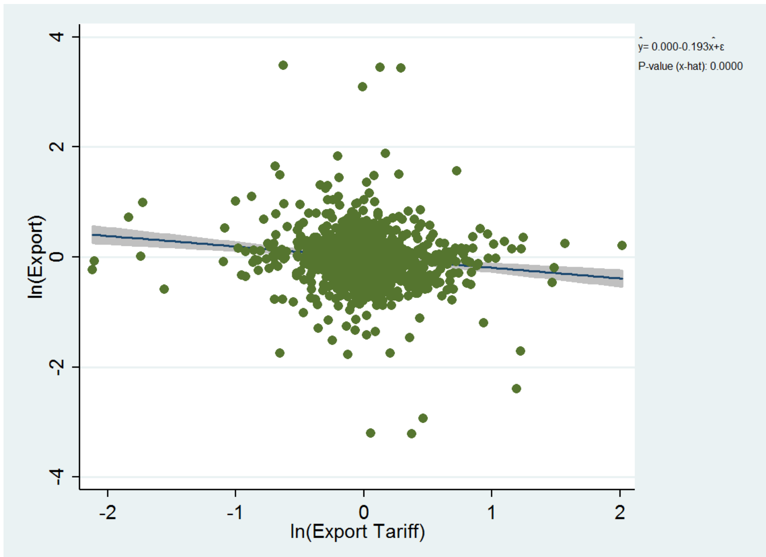
XU, Z., J. HUANG, AND F. JIANG (2017): "Subsidy competition, industrial land price distortions and overinvestment: empirical evidence from China's manufacturing enterprises," *Applied Economics*, 49, 4851–4870.

ZI, Y. (2016): "Trade Liberalization and the Great Labor Reallocation," *mimeograph*.

Appendix

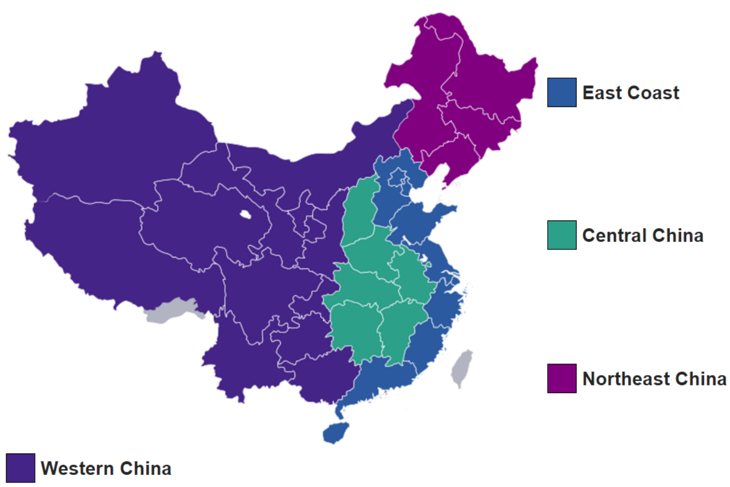
A More tables and graphs

Figure A.1: Relationship between the Export Tariff and Export Volume



Note: This figure shows the first-stage result of our baseline regression. It shows that the export tariff is significantly correlated with the export volume.

Figure A.2: The economic region in China



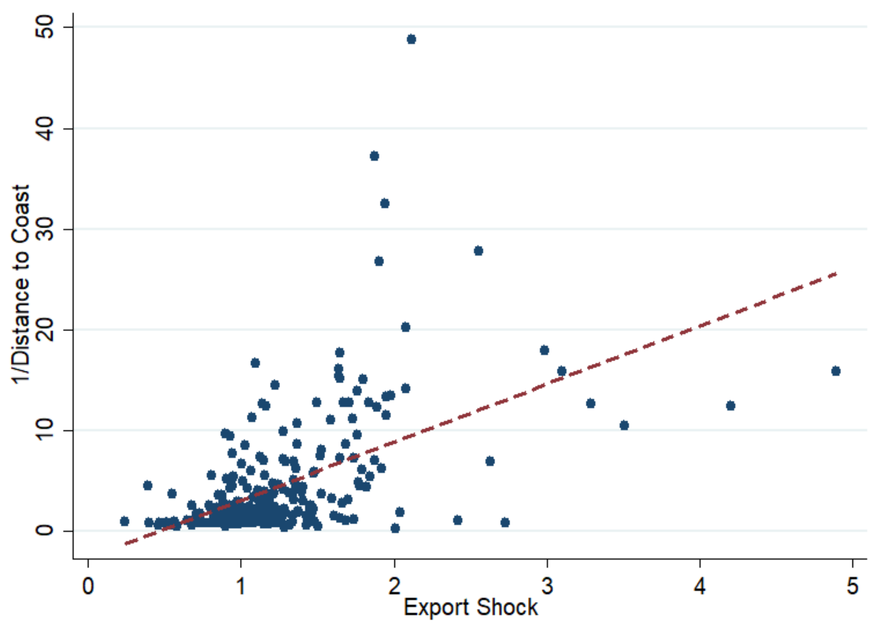
Source: Wikipedia

B Robustness Checks of Empirical Part

B.1 Using Distance as Proxy for Trade Shock

The trade shock measures the prefecture's openness to the international trade. However, the measure is not perfect since some areas have high international trade cost and high measure of "trade shock" due to their abundant natural resources. What's more, cities including Beijing and Shanghai have relatively low international trade cost but their trade shock measure is not high enough. These cities probably already start to specialize in service sector. Hence, in this section we use an alternative measure of the trade openness: the distance to the nearest port city as robustness check. We use a group of port cities and calculate each city's distance to their nearest port.[40]

Figure A.3: Export Shock and the Inverse of Distance to the Coast



Notes: The vertical axis is the reciprocal of the city's distance to the coast, which proxies the city's export shock. The horizontal axis is the export shock we constructed in the paper. Here, we define the distance to the coast is the shortest distance to the nearest port city.

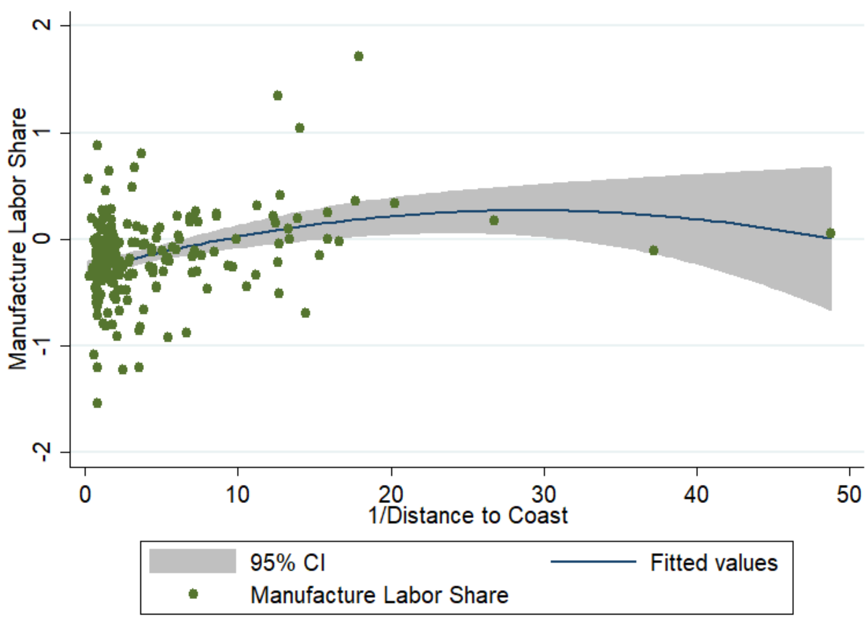
The empirical result is in Table A.1[41] and the scatter plot is in Figure A.4. We use the reciprocal of "Distance to Coast" to proxy for the constructed export shock and find that the

[40]The group of the port cities are: Dalian, Dandong, Jinzhou, Qinhuangdao, Tianjin, Yantai, Qingdao, Shanghai, Hangzhou, Ningbo, Chaozhou, Zhangjiang, Beihai, Haikou, Wenchang, Qionghai, Wanning, Lianyungang, Nantong, Wenzhou, Fuzhou, Guangzhou, Zhanjiang, Xiamen, Nanjing.

[41]In Table A.1, the observations are 167 and 118 which are different from the 188 observations in Table 1. The

result is still consistent with the empirical findings in the section 2. In Table A.1, the coefficient signs are the same as our baseline result. To check whether the hump shape pattern is driven by a few cities on the right-hand side of Figure A.4, we also exclude the cities with the Distance to Coast's reciprocal is higher than 75th percentile of its own distribution, which are the cities on the right hand side of Figure A.4. And we find the hump shape pattern still holds in the sub-sample regression if we use the level of manufacturing labor share in the ending period as the dependent variables [42].

Figure A.4: Manufacture Labor Share and the Inverse of Distance to the Coast



Notes: The horizontal axis is the reciprocal of the city's distance to the coast, which proxies the city's export shock. Here, we define the distance to the coast is the shortest distance to the nearest port city. The figure illustrates the relationship between prefecture-level distance to the coast's reciprocal and manufacture labor share. We also include the 95% confidence interval as the shaded gray areas.

attrition comes from the construction of trade shock data since some of the prefectures don't have the trade information from the customs data.

[42] If we use the log change in labor share as the dependent variable, the sub-sample coefficients are not significant. It could be that the sample size becomes small if we use the sub-sample and the variations of labor share's log change are small. But still the coefficients' signs are consistent with our other regressions.

Table A.1: Trade Liberalization and Manufacture Labor Share using the Inverse of the Distance to the Coast as Proxy for Export Shock

Dep. Var.:	Change in Labor Share		Level in Labor Share	
Sample:	Full	Sub-Sample	Full	Sub-Sample
1/Distance to Coast	0.037***	0.037	0.036**	0.125*
	(0.014)	(0.042)	(0.015)	(0.033)
1/Distance to Coast^2	-0.001**	-0.002	-0.001**	-0.019**
	(0.000)	(0.007)	(0.000)	(0.003)
R^2	0.31	0.25	0.63	0.67
N	167	118	167	118
Region FE	Y	Y	Y	Y
Initial Conditions	Y	Y	Y	Y
Contemporary Shocks	Y	Y	Y	Y
Capital City	Y	Y	Y	Y

Notes: The sub-sample excludes cities with 1/Distance to Coast is higher than 75th percentile of its own distribution, which are the cities on the right hand side of Figure A.4. Initial conditions include the start period's manufacture labor share. Contemporaneous shocks include change in log GDP per capita, change in log GDP, sectoral productivity changes and a dummy of provincial capital city. Region fixed effect is also included. Standard errors are clustered at the region level. ***p<0.01, **p<0.05, *p<0.1

B.2 Outlier Problems in Export Shock Variables

To test whether the hump-shape pattern between trade liberalization and manufacturing labor share is driven by a few cities with very large export shock, we conduct the robustness checks by excluding cities with very large export shock. We exclude outlier cities by imposing the conditions if the city's export shock is larger than its 50th, 75th, 90th, 95th percentile of the "ExportShock" distribution. In Table A.2, we find that even after excluding cities with large export shock, the hump-shape pattern still exists. Hence, the hump shape does not depend on those few cities with very large export shocks or outliers, but these cities indeed determine the shape parameters of the hump shape pattern. Excluding these cities with large export shocks does not change the hump shape pattern but change the hump shape pattern's slope.

For the outlier issues in Figure A.4, we also conduct similar robustness checks in Table A.1 and find the pattern remains robust even after excluding a few outliers.

Table A.2: Trade Liberalization and Manufacture Labor Share Excluding Cities with Large Export Shock

Dep. Var.:	Change in Labor Share				Level in Labor Share			
Sample: if ExportShock<	p(50)	p(75)	p(90)	p(95)	p(50)	p(75)	p(90)	p(95)
Method:	2SLS							
ExportShock	5.777**	2.295*	1.726**	1.044**	6.775**	2.775*	2.062**	1.241**
	(1.604)	(0.910)	(0.366)	(0.217)	(1.987)	(1.003)	(0.433)	(0.250)
ExportShock^2	-2.553**	-0.803	-0.515**	-0.253**	-3.509**	-1.141*	-0.723**	-0.354**
	(0.690)	(0.370)	(0.127)	(0.066)	(1.017)	(0.473)	(0.178)	(0.090)
R^2	0.76	0.72	0.70	0.70	0.76	0.72	0.70	0.70
N	86	133	164	175	86	133	164	175
Region FE	Y	Y	Y	Y	Y	Y	Y	Y
Initial Conditions	Y	Y	Y	Y	Y	Y	Y	Y
Contemporary Shocks	Y	Y	Y	Y	Y	Y	Y	Y
Capital City	Y	Y	Y	Y	Y	Y	Y	Y

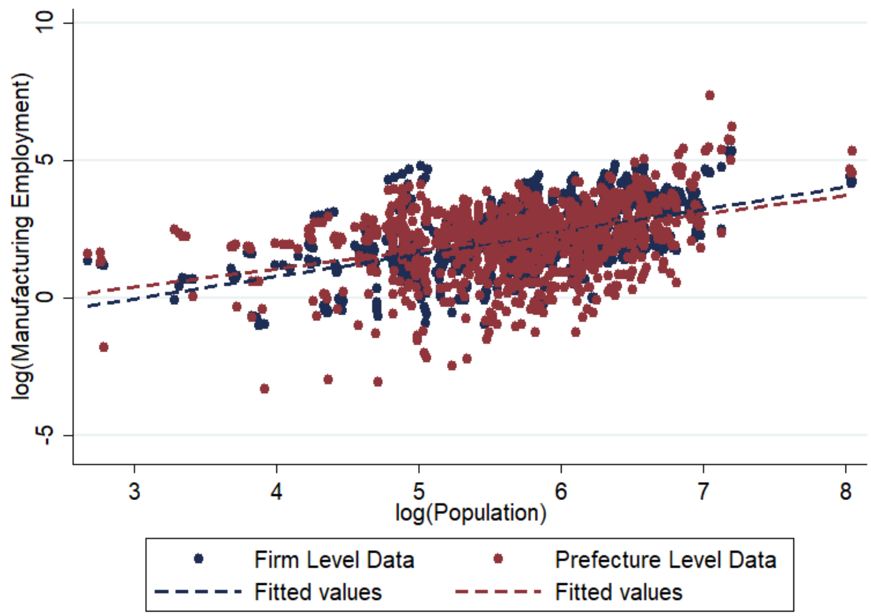
Notes: The dependent variables are the change or level of manufacturing labor share. We exclude cities with large export shocks by imposing the conditions if the city's export shock is larger than its 50th, 75th, 90th, 95th percentile of the "ExportShock" distribution. Initial conditions include the start period's manufacture labor share. Contemporaneous shocks include change in log GDP per capita, change in log GDP, sectoral productivity changes, and a dummy of the provincial capital city. Region fixed effect is also included. Standard errors are clustered at the regional level. ***p<0.01, **p<0.05, *p<0.1

B.3 Selection Problem in the Firm-level Data

We use the Chinese manufacturing survey data to calculate the manufacturing labor share from the year 1998-2008. The dataset is comprehensive but might neglect employment from small and medium-sized firms. The selection problem could be more prominent in small cities since small cities might have more small firms than large cities. If the firm-survey data exists a selection problem, the trade shock in small cities might be underestimated.

To check whether the selection problem exists in the firm-level survey data, we compare the manufacturing employment calculated from the firm survey data with the prefecture-level data in the City Statistics Yearbooks in Figure A.5. We plot the scatter plots and fitted lines between the city size and manufacturing employment using both the firm survey data and city statistics yearbooks. We plot the scatter plots and fitted lines between the city size and manufacturing employment using both the firm survey data and city statistics yearbooks. We find little evidence on the differences of the selection problem in the two datasets. The slope of the two fitted lines using firm-level survey data and prefecture-level yearbook data are very close to each other.

Figure A.5: Scatter Plot between Manufacturing Employment and City Size



Notes: The firm-level manufacturing employment data is aggregated from China's Manufacture Annual Firm Survey Data. The city-level manufacturing employment data is directly from the City Statistics Yearbooks.

To address the concern of the selection problem in small cities, we further provide robustness checks by excluding small cities in the baseline regressions. The result is in Table A.3. We exclude small cities by excluding the sample if the log of the city's population size is smaller than its 1st, 5th, 10th percentile of the overall city size distribution. We find that the hump-shaped pattern between the manufacturing labor share and trade openness still exists after excluding the small cities. And the coefficients also are at a similar magnitude to our baseline result.

Table A.3: Trade Liberalization and Manufacture Labor Share Excluding Small Cities

Dep. Var.:	Change in Labor Share			Level in Labor Share		
Sample: if log(pop)>	p(1)	p(5)	p(10)	p(1)	p(5)	p(10)
Method:	2SLS					
ExportShock	0.363***	0.380**	0.429**	0.311***	0.325**	0.367**
	(0.042)	(0.085)	(0.092)	(0.036)	(0.075)	(0.080)
ExportShock^2	-0.059***	-0.063**	-0.073**	-0.043***	-0.046**	-0.054**
	(0.006)	(0.013)	(0.015)	(0.005)	(0.010)	(0.011)
R^2	0.29	0.33	0.33	0.29	0.33	0.33
N	186	177	169	186	177	169
Region FE	Y	Y	Y	Y	Y	Y
Initial Conditions	Y	Y	Y	Y	Y	Y
Contemporary Shocks	Y	Y	Y	Y	Y	Y
Capital City	Y	Y	Y	Y	Y	Y

Notes: The dependent variables are changes in the manufacturing labor share and the ending period of manufacturing labor share. We exclude small cities by imposing the conditions if the log of population size is smaller than its 1st, 5th, 10th percentile of city size distribution. Initial conditions include the start period's manufacture labor share. Contemporaneous shocks include change in log GDP per capita, change in log GDP, sectoral productivity changes and a dummy of the provincial capital city. Region fixed effect is also included. Standard errors are clustered at the regional level. ***p<0.01, **p<0.05, *p<0.1

To conduct the Bartik IV approach in the baseline empirical analysis, we need the detailed industry-level manufacturing employment data, which the city or provincial level yearbooks do not provide. In fact, the manufacture annual firm survey dataset is also widely used in the similar exercise setting in other related literature such as Li (2018), Bombardini and Li (2020) and Xu (2020).

B.4 Other Robustness Checks

The structural transformation might not happen within regions, and thus, we test this in this session. First, if we add the square term of the export shock in the regression, the coefficients are not significant for the central and northeast regions. After excluding the square term of the export shock, the export shock's coefficient becomes positive, indicating that the manufacturing labor share is still positively correlated with the export shock in the central and northeast regions. The results suggest that these two regions are still in their early development stage and have not reached the hump shape structural transformation yet. The empirical test shows that the structural change happens among regions instead of within the region.

To further validate our baseline result, we also conduct robustness checks using other dependent variables that resemble the structural transformation, such as the GDP or the output manufacture share. The regression result is in Table A.5. We find that the hump shape pattern still holds if we use the manufacturing output or GDP share as the dependent variables.

Table A.5: Trade Liberalization and Manufacture Output Share

	Change in Output Share		Level in Output Share	
	OLS	2SLS	OLS	2SLS
ExportShock	0.115**	0.133**	0.104*	0.121*
	(0.050)	(0.058)	(0.056)	(0.065)
ExportShock^2	-0.024***	-0.032***	-0.021**	-0.028**
	(0.009)	(0.012)	(0.010)	(0.013)
R^2	0.48	0.48	0.62	0.62
N	203	203	203	203
Region FE	Y	Y	Y	Y
Initial Conditions	Y	Y	Y	Y
Contemporary Shocks	Y	Y	Y	Y
Capital City	Y	Y	Y	Y

Notes: The dependent variables are the change or level of the manufacturing output share. Here we define the manufacturing output share as manufacturing value-added divided by the total value-added in each prefecture. Initial conditions include the start period's manufacture labor share. Contemporaneous shocks include change in log GDP per capita, change in log GDP, sectoral productivity changes and a dummy of provincial capital city. Region fixed effect is also included. Standard errors are clustered at the region level. ***p<0.01, **p<0.05, *p<0.1

Table A.4: Trade liberalization and Manufacture Labor Share by Regions

Panel A: Change in Labor Share

	East and West		Central and Northeast			
	OLS	2SLS	OLS	2SLS	OLS	2SLS
ExportShock	0.743***	0.864***	-0.331	-0.395	0.317**	0.375**
	(0.202)	(0.233)	(0.804)	(0.947)	(0.136)	(0.159)
ExportShock^2	-0.096***	-0.130***	0.237	0.330		
	(0.033)	(0.044)	(0.290)	(0.400)		
R^2	0.54	0.54	0.32	0.33	0.31	0.31
N	60	60	31	31	31	31
Region FE	Y	Y	Y	Y	Y	Y
Initial Conditions	Y	Y	Y	Y	Y	Y
Contemporary Shocks	Y	Y	Y	Y	Y	Y
Capital City	Y	Y	Y	Y	Y	Y

Panel B: Level in Labor Share

	East and West		Central and Northeast			
	OLS	2SLS	OLS	2SLS	OLS	2SLS
ExportShock	0.514**	0.600**	0.496	0.570	0.443***	0.520***
	(0.239)	(0.275)	(0.791)	(0.933)	(0.132)	(0.155)
ExportShock^2	-0.066*	-0.090*	-0.019	-0.021		
	(0.039)	(0.052)	(0.285)	(0.394)		
R^2	0.52	0.52	0.89	0.89	0.89	0.89
N	60	60	31	31	31	31
Region FE	Y	Y	Y	Y	Y	Y
Initial Conditions	Y	Y	Y	Y	Y	Y
Contemporary Shocks	Y	Y	Y	Y	Y	Y
Capital City	Y	Y	Y	Y	Y	Y

Notes: The dependent variables are the log changes or level of manufacturing labor share within the 10 year time period. Initial conditions include the start period's manufacture labor share. Contemporaneous shocks include change in log GDP per capita, change in log GDP, sectoral productivity changes, and a provincial capital city dummy. Region fixed effect is also included. Standard errors are clustered at the regional level. ***p<0.01, **p<0.05, *p<0.1

C Theoretical Proofs

C.1 Proof of Proposition

Proposition 2. *Given the real income for each region ν_j, migration costs between all regions μ_{ij}, and heterogeneous preference distribution $F(z_j)$, the share of region i workers that migrate to region j is*

$$H_{ij} = \frac{\left(\frac{\nu_j}{\mu_{ij}}\right)^{\varepsilon}}{\sum_{m \in \boldsymbol{G}} \left(\frac{\nu_m}{\mu_{im}}\right)^{\varepsilon} dm}. \tag{25}$$

Proof. The probability that a worker ω is moving from origin i to destination j is[43]

$$\begin{aligned}
H_{ij} &= \mathbf{Pr}\left(\frac{z_{ij}(\omega)\nu_j}{\mu_{ij}} \geq \frac{z_{ig}(\omega)\nu_g}{\mu_{ig}}, \forall \ g \in \mathbf{G}\right) \\
&= \mathbf{Pr}\left(z_{ig} \leq \frac{\nu_j/\mu_{ij}}{\nu_g/\mu_{ig}} z_{ij}, \forall \ g \in \mathbf{G}\right) \\
&= \mathbf{Pr}\left(z_{ig} \leq \min\left(\frac{\nu_j/\mu_{ij}}{\nu_g/\mu_{ig}} z_{ij}, \forall \ g \in \mathbf{G}\right)\right) \\
&= \int_0^{+\infty} \prod_{g \in \mathbf{G}} \left[1 - F_{ij}\left(\frac{\nu_j/\mu_{ij}}{\nu_g/\mu_{ig}} z_{ij}\right)\right] dF_{ig}(z)
\end{aligned}$$

.

Given the formula of F, it follows that

$$\begin{aligned}
H_{ij} &= \int_0^{+\infty} \varepsilon z^{-\varepsilon-1} exp\left\{ -z^{-\varepsilon} \times \frac{\int_g \left(\frac{\nu_g}{\mu_{ig}}\right)^{\varepsilon} dg}{\left(\frac{\nu_j}{\mu_{ij}}\right)^{\varepsilon}} \right\} dz \\
&= \frac{\left(\frac{\nu_j}{\mu_{ij}}\right)^{\varepsilon}}{\int_g \left(\frac{\nu_g}{\mu_{ig}}\right)^{\varepsilon} dg}.
\end{aligned}$$

[43] For simplicity, we drop ω in the last few steps of derivation.

$\square$

C.2 Market Access Channel

We illustrate the market access channel in an simplified two symmetric country case following Uy et al. (2013). We assume two symmetric country and three sectors: agriculture(A), manufacture(M) and service(S) and agriculture and manufacture are tradable while service is non-tradable.[44] Production follows the same characteristics as before: perfect competition, Frechet technology distribution and symmetric iceberg trade cost τ. However, here we even simplify to one factor of production, labor instead of labor and land. The Frechet technology parameter follows $T_1^M = T_2^A > T_2^M = T_1^A$ which means home country 1 has comparative advantage in producing manufacture goods while foreign country 2 has comparative advantage in producing agriculture goods.

From the goods market condition in manufacture sector, we can write down the equation below

$$w_1 L_1^M = \pi_{11}^M P_1^M C_1^M + \pi_{12}^M P_2^M C_2^M. \tag{26}$$

Country 1's manufacture revenue comes from the sales in country 1 $P_1^M C_1^M$ and country 2 $P_2^M C_2^M$. P_1^M is the price index of manufacture industry in country 1 and C_1^M is the consumption of manufacture in country 1. π_{12}^M is the country 2's trade share of importing goods from country 1 and π_{11}^M is country 1's domestic trade share.

Dividing Equation (26) both sides by $w_1 L_1$, we can derive the manufacture labor share l_1^M:

$$\begin{aligned} l_1^M = \frac{L_1^M}{L_1} &= \pi_{11}^M \frac{P_1^M C_1^M}{w_1 L_1} + \pi_{12}^M \frac{P_2^M C_2^M}{w_1 L_1} \\ &= \pi_{11}^M X_1^M + \pi_{12}^M \frac{P_2^M C_2^M}{w_1 L_1} \\ &= (1 - \pi_{21}^M) X_1^M + \pi_{12}^M \frac{P_2^M C_2^M}{w_1 L_1}, \end{aligned}$$

where X_1^M is the manufacture expenditure share in country 1. After further algebra arrange-

[44] We add one more tradable sector agriculture to simply the model and equalize the wages between the two countries. If there's only one tradable sector manufacture and the comparative advantages of manufacture differ across two countries, the wages won't be the same across the countries.

ments, we can derive

$$
\begin{aligned}
l_1^M &= X_1^M + \frac{\pi_{12}^M X_2^M w_2 L_2 - \pi_{21}^M X_1^M w_1 L_1}{w_1 L_1} \\
&= X_1^M + NX_1^M,
\end{aligned}
$$

where NX_1^M is the net export share in country 1. Hence, the change in manufacture labor share comes from two sources: expenditure share effect and net export effect.

We further assume that the elasticity of substitution σ between the three sectors A,M,S is 1 and this simplify back to the Cobb-Douglas preference. Based on the property of the Cobb-Douglas preference, the expenditure share in each sector is constant so this eliminates the expenditure share effect. Hence, X_1^M is not changing with trade cost: $\frac{\partial X_1^M}{\partial \tau} = 0$. Given the symmetry across the two countries, wages are equalized $w_1 = w_2$. We write the trade share equation as following

$$
\begin{aligned}
\pi_{21}^M &= \frac{T_1^M (w_1 \tau)^{-\theta}}{T_2^M w_2^{-\theta} + T_1^M (w_1 \tau)^{-\theta}} = \left[\frac{T_2^M}{T_1^M} \tau^\theta + 1 \right]^{-1} \\
\pi_{12}^M &= \frac{T_2^M (w_2 \tau)^{-\theta}}{T_1^M w_1^{-\theta} + T_2^M (w_2 \tau)^{-\theta}} = \left[\frac{T_1^M}{T_2^M} \tau^\theta + 1 \right]^{-1}.
\end{aligned}
$$

Hence,

$$
\frac{\partial NX_1^M}{\partial \tau} = X_1^M \theta \tau^{\theta - 1} \left[-(\pi_{12}^M)^2 \frac{T_1^M}{T_2^M} + (\pi_{21}^M)^2 \frac{T_2^M}{T_1^M} \right] < 0,
$$

since $\pi_{21}^M > \pi_{12}^M$. The manufacture labor share increases as trade cost declines.

C.3 Definition of Equilibrium

Definition C.1. A competitive equilibrium of the economy is defined as a set of prices and allocations such that the following conditions are satisfied:

1. Workers maximize utility and choose their location optimally considering migration frictions, that is, Equation (8) and (9) are satisfied.

2. Firms maximize profits given input prices and goods prices.

3. Firms make zero profits.

4. Factors and goods markets clear and specifically, labor market clears in each city.

D Calibration Parameters

D.1 Land Share Parameters

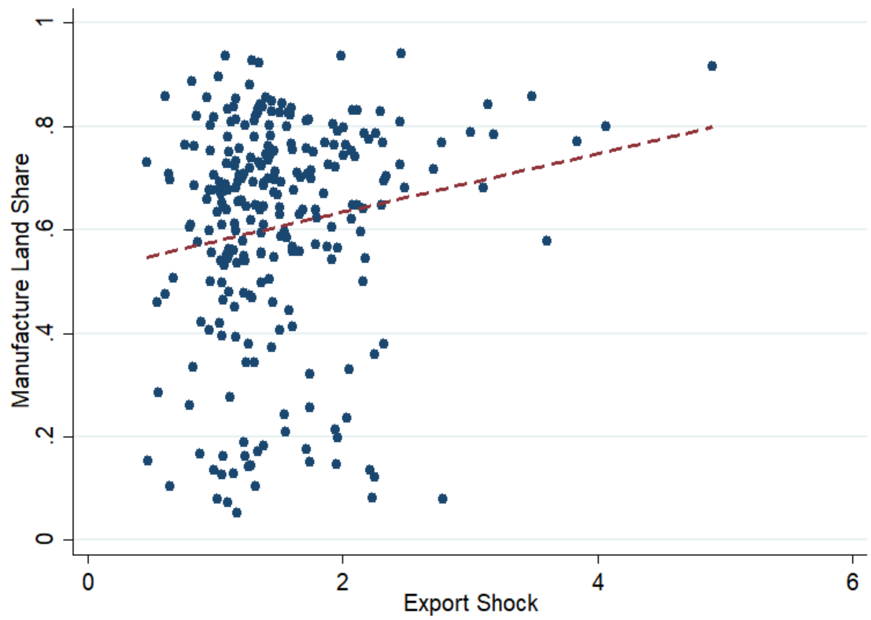
To calibrate the land allocations between manufacturing and service sectors in different c-ities, we use the data from China Urban Construction Statistical Yearbook[45]. The dataset provides the area of urban construction land, including industrial and manufacturing land and commercial and business facilities land. We infer the land share of manufacturing and service sectors directly from the data, and the manufacturing land share is defined as

$$\text{Manufacture Labor Share}_i = \frac{emp_{Mi}}{emp_{Mi} + emp_{Si}}, \tag{27}$$

where emp_{Mi} is the employment in the manufacture sector in prefecture i, and emp_{Si} is the employment in the service sector in prefecture i. To validate the data, we plot the relationship between manufacturing land share and export shock in Figure A.6. We find a roughly positive relationship between export shock and manufacturing land share. The summary statistics of manufacturing land share is in Table A.6 .

[45]Since the service sector land data, listed as "commercial and business facilities land", is only available starting from the year 2012, we use the year 2012 data as a proxy for the year 2008 to calibrate the land share allocations between manufacturing and service sector. Previous years does not include the commercial and business facilities land in the yearbook, so we could not infer the land shares of manufacturing and service in different cities before 2012.

Figure A.6: Scatter Plot between Manufacturing Land Share and Export Shock



Notes: The city-level manufacturing land data is from China Urban Construction Statistical Yearbook. The export shock data is constructed from Section 2.

Table A.6: Summary Statistics of Manufacturing and Service Land Share

	mean	S.D.	p10	p25	p50	p75	p90
Manufacturing Land Share	0.607	0.214	0.213	0.507	0.670	0.762	0.829
Service Land Share	0.393	0.214	0.171	0.238	0.330	0.493	0.787

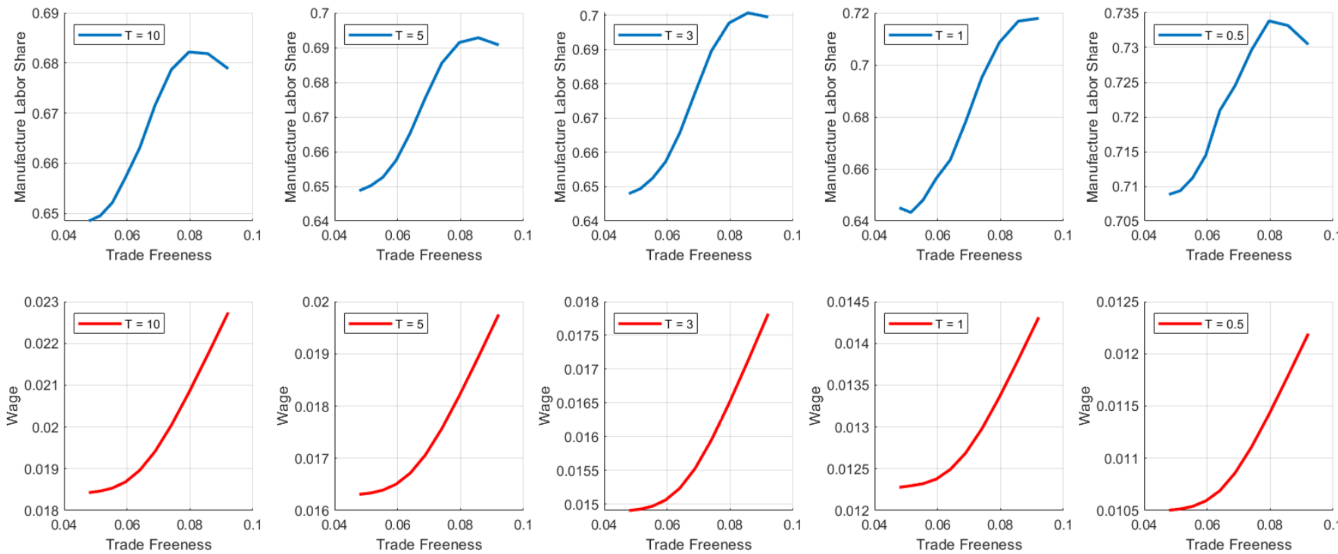
Notes: The table presents the summary statistics of manufacturing and service land share distribution. We will use the data to calibrate the land endowment of different cities in the quantitative analysis part.

D.2 Choice of Frechet Distribution Parameters

In the numeric analysis part, we choose the Frechet distribution parameters by assumptions, and in the section, we would test whether the baseline result still holds by choosing different Frechet Distribution parameters.

The country-specific frechet distribution parameter, i.e., state of technology T_i governs the location of the Frechet distribution and implies that a high-efficiency draw for any good is

Figure A.7: The Hump-shape Relationship between Manufacture Labor Share and Trade Freeness at Various Choices of Frechet Distribution



Note: The figure illustrates the relationship between manufacture labor share and trade freeness at various choices of Frechet distribution technology parameter T_i in the manufacturing sector in the domestic country. We fix the foreign country's Frechet distribution parameter in the manufacturing sector as 1.

more likely. The technology parameters T_i are hard to estimate in the literature as mentioned in Caliendo and Parro (2015), especially when we do not have one specific foreign country in our model (i.e., the rest of the world represents all other foreign countries except for China). As in the model, we assume China has comparative advantages in manufacturing sector compared to the rest of the world, we set China's T_i^M in the manufacturing sector larger than the ones in the foreign country, and set the relative magnitude of T_i^M between domestic and foreign following Eaton and Kortum (2002) [46].

We carry out the robustness checks by adopting various Frechet distribution parameters T_i, and the results are in Figure A.7. The hump-shape pattern still holds under different choices of Frechet parameters T_i and T_i only change the hump-shape pattern's turning points. On average, the wages are higher when T_i is higher since higher T_i indicates higher productivity efficiency.

D.3 The Choice of the Trade Freeness

In this section, we provide a further illustration of how the trade freeness is calculated and the meaning of the level of trade freeness.

[46]However, Eaton and Kortum (2002) estimates T_i among a set of OECD countries, which do not include China. So the estimates might not directly to apply in our model's case.

Table A.7: The Choice of Trade Freeness and its Corresponding Meaning

Assumption: theta = 4				
Corresponding Region	Western	Northeast	Central	Eastcoast
Average Intra-regional trade cost	2.1	1.9	1.9	1.8
International trade cost	4.2	3.5	2.8	2.6
Corresponding Trade Freeness	0.05	0.06	0.07	0.09

Notes: The table illustrates the mapping of the intra-regional trade cost and international trade cost to the corresponding level of trade freeness. We choose the level of intra-regional trade cost and international trade following Tombe and Zhu (2019), which are the average trade costs in different regions in China from the year 2002-2007. We calculate the trade freeness following the equation (28).

To be more specific, we define the trade freeness as the following expression:

$$\text{trade freeness in region } i = \Big(\sum_{j \neq i, j=1}^{N+1} \tau_{ij}\Big)^{-\theta}, \tag{28}$$

where $\tau_i j$ is the trade cost between region i and region j. The range of the trade freeness in the region i depends on the intra-regional and international trade cost, trade elasticity θ, and the number of cities in the numerical calibration. The higher the trade freeness, the lower its trade cost to the other regions and the rest of the world as showed in Table A.7.

E Further Discussions of the Major Numerical Result

E.1 Service Sector Trade

In this section, we would like to test whether the major results still hold if adding the cross-regional service sector trade within a country. Empirically, we can see a rising share of service trade across regions from China's regional input-output table (Mi et al., 2018). So we add a scenario that allows cross-regional service trade within a country but does not allow trade in the service sector across countries. We calibrate the service sector's trade cost following the literature (Miroudot et al., 2013; Deswal, 2014). The result is in Figure A.8. We find that the hump-shape pattern still holds, but the slope becomes flatter compare to the baseline.

When the service sector is non-tradable, international trade has a more critical impact on the tradable manufacturing sector. It acts as the primary driving force for the structural transformation in various cities. When the service sector is allowed to trade across regions, it will absorb part of the international trade effects since the factor reallocation through service trade is also easier. Thus, adding the service trade would make international trade's impact diminished and make the hump-shape pattern flatter compared to the baseline case.

E.2 Add Rural-urban Migration

Although the primary focus of the paper studies the transformation among service and manufacturing sectors, the study period 1998-2008 is renowned for the rapid urbanization in China, which features the labor allocation from the agriculture sector the industry sectors.[47] In the model part, we would like to test whether the hump-shape pattern still holds if adding the rural-urban migration.

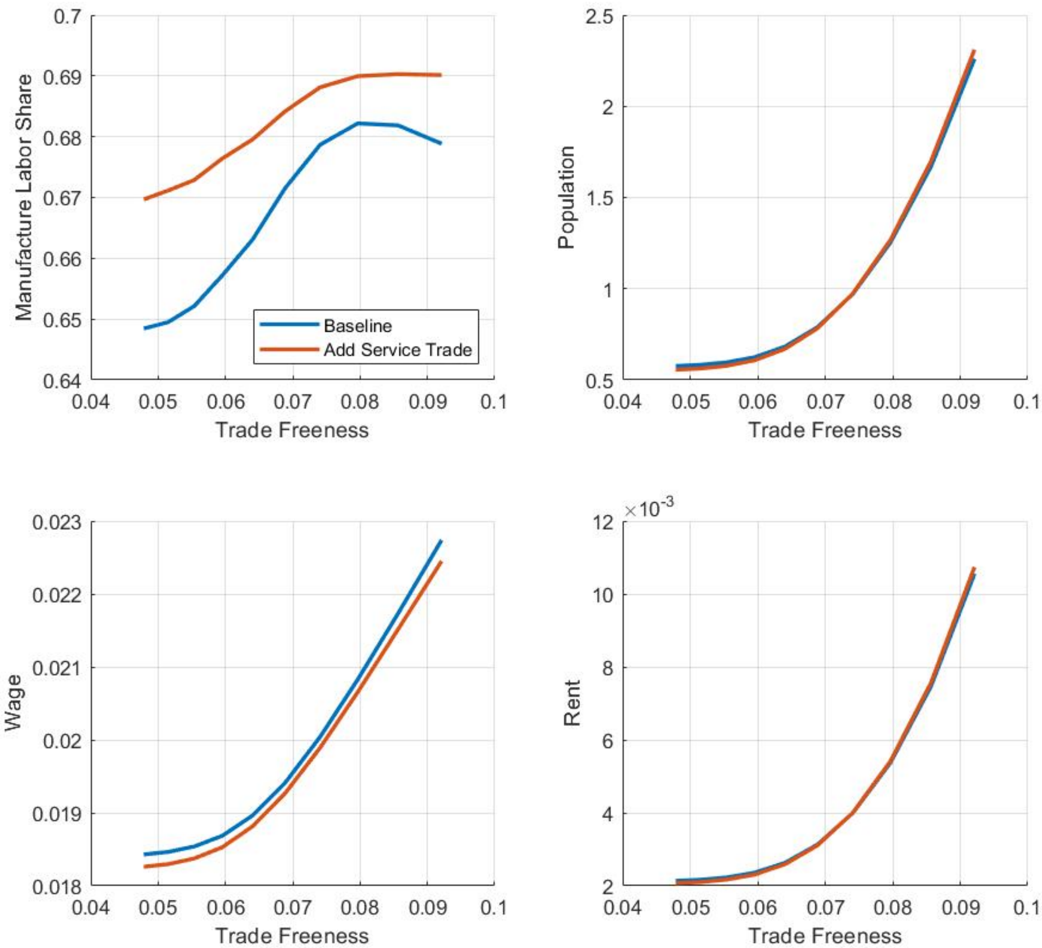
We model the rural-urban migration as the exogenous supply shock to the urban population supply[48]. We calibrate the urban population's level using the urban population statistics from 1998 to 2008 from the National Bureau of Statistics. Figure A.9 shows that the hump-shape pattern still holds if we add the rural-urban migration. However, the hump-shape pattern's slope is flatter compared to the baseline case without rural-urban population migration. The intuition is that adding the rural-urban migration will facilitate the internal factor reallocation channel, therefore dampens the market access channel in the rising part of the hump-shape, which makes the hump-shape's slope becomes flatter.

The paper's major focus is to understand the within-urban labor allocation between manufacturing and service sectors, and the agriculture to the industry or the rural to urban labor allocation is not our main concentration of the paper.

[47] The Chinese government starts to formally institute a Hukou registration system in 1958 to control the population mobility. Every Chinese citizen is assigned with a *Hukou* (registration status), classified as "rural" or "urban" in an administrative unit. Suppose an individual wants to change their status (from rural to urban) or change the location of *Hukou* registration. In that case, they need approval from the government, which is extremely hard to obtain. Hence, before the economic reform started in 1978, working out of one's *Hukou* registration location is severely prohibited. And then government relaxes the policy in the 1980s, but before 2003, workers without local *Hukou* still need to apply for a temporary residence permit. After 2003, many provinces, especially the coastal provinces, remove the temporary residence permit requirement for migrant workers due to the rise in demand for workers in manufacturing, construction, and labor intensive service industries. This leads a massive rural-to-urban and inter-provincial migration: the fraction of the urban population in the total population increased from the low level of 20 percent in 1980 to 60 percent in 2008, and the annual level of migration from rural areas to urban areas has ranged from 0.2% to 3.2%, with a 1.5% average.

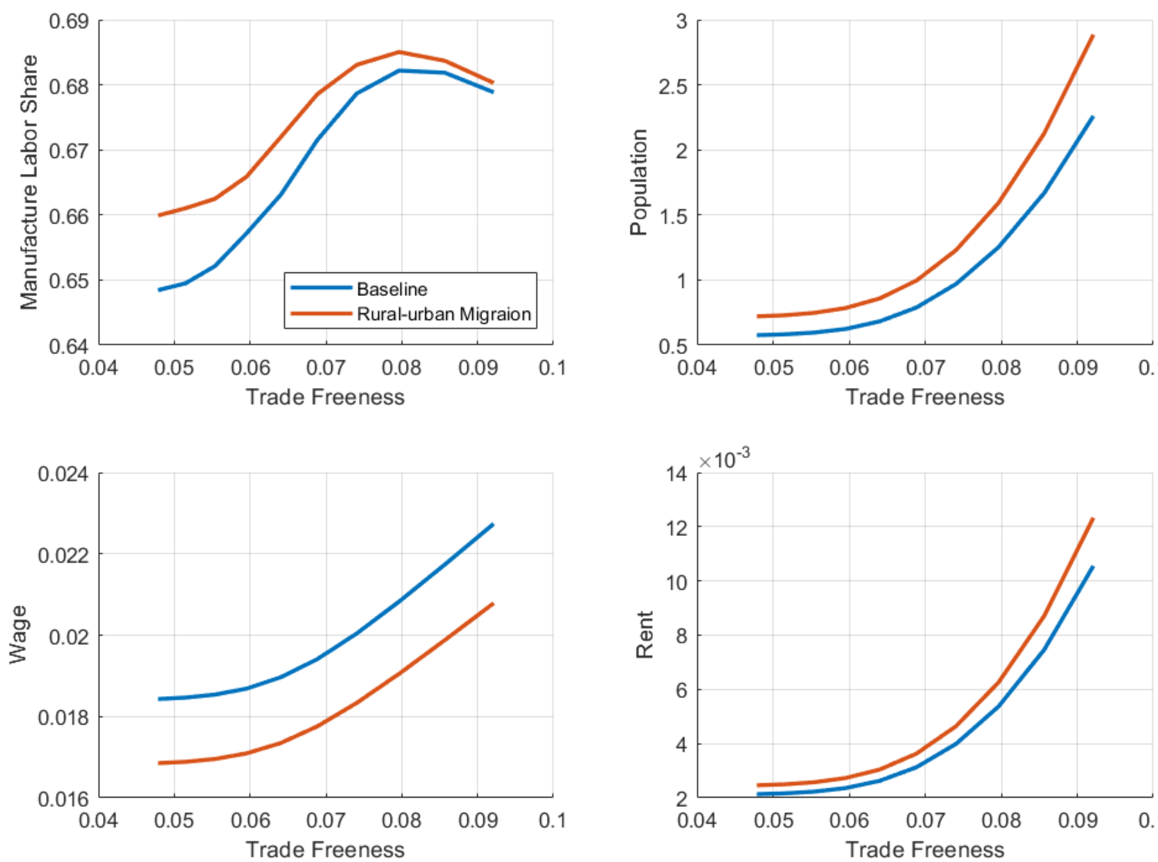
[48] From 1998 to 2008, the urban population has increased by 50%, which turns into 2.08 billion in the number of the total urban population. According to the Chen and Song (2014), the rural-urban migration accounts for 53.5% of the entire urban population increase in the period of 1990-2000. Therefore, the urban population' rise by around 26.7% due to the rural-urban migration.

Figure A.8: Counterfactual Analysis: Adding Service Trade



Note: The figure illustrates our baseline scenario's numerical results and the scenario adding the service trade across regions within a country. But we do not allow international trade in the service sector in both scenarios. Trade freeness is defined as $\tau^{-\theta} \in (0, 1)$, and τ is the geometric mean across all the destinations.

Figure A.9: Counterfactual Analysis: Adding Rural-urban migration



Note: The figure illustrates the numerical results of our baseline scenario and the scenario adding the rural-urban migration. Trade freeness is defined as $\tau^{-\theta} \in (0, 1)$, and τ is the geometric mean across all the destinations.

E.3 Land Distortion Problem

In China, the land market is faced with strict state regulations, and the regional government usually distort the manufacturing land price allocations to attract investment. One the one hand, urban land planning is under rigorous planning and restrictions; on the other hand, the provision of manufacturing land is under local government's distortions, and the place-based policies are usually in place to attract firms and boost the local economy (Fei, 2020). Hence, the land market distortions might have an impact on our main result in the paper. We discuss this issue in this section.

First, we could view the land market distortion as the distorted land price in the model. Following Hsieh and Klenow (2009), the distortions could be modeled as a wedge in front of the land rent price.

Hence, the profit maximization problem of firms is given by

$$\max_{L_i^k, K_i^k, N_i^k} \left\{ P_i^k y_i^k(z) - w_i L_i^k(z) - qK_i^k(z) - (1+\tau_{Nzi}^k) r_i^k N_i^k(z) \right\} \quad \text{for} \quad k \in \{M, S\}, \tag{29}$$

where τ_{Nzi}^k is the land market distortion for firm z in sector k and region i. For example, τ_N would be lower for firms in inland cities where the local government imposes the place-based policies attracting firms to boost the local capital investment and employment. τ_N could also be higher in large cities if the city is under strict urban planning regulations, such as Beijing and Shanghai.

One closely related work estimating land wedges in China is Deng et al. (2020), whose estimations in 2008 span from -0.11 to 3.41, the median and standard deviation are 1.36 and 1.91 respectively, which reflects the land wedges vary a lot across different cities. However, to be noted, their estimations focus only on the residential land and housing market, and this is different from our model's setting focusing on the manufacturing and service sectors. And we note that the wedges of the residential land and manufacturing land are also be different. Therefore, their estimations should not be directly applied in our model's setting.

Although the accurate estimates of the land wedges are not directly measured in the literature, we could infer the signs of the land wedges from other related literature. Previous literature finds that the industrial land price distortions significantly affect the overinvestment of manufacturing firms (Xu et al., 2017). Hence, in general, China's industrial land price is distorted towards the low-end while the service land price might be over-priced (Wu et al., 2014). To put it in another way, τ_N^M is usually negative while τ_N^S is positive. We are overestimating the land price ratio of manufacturing land to service land in the baseline model if we do not include the land market distortions.

As the paper points out, the hump shape pattern between the manufacturing labor share and trade liberation is shaped through two channels: the market access channel and the internal factor reallocation channel. Since the market access channel is related to the international trade cost, which is generally exogenous, the land market distortions will not directly affect the structural transformation through this channel. However, the land market distortions will affect the internal factor reallocation channel through the underestimated land price ratio of manufacturing land to service land. As the paper illustrates, the internal factor reallocation channel's intuition is that labor migrates to the coastal areas where the wage is higher. The rise in the labor input will lead to the expansion in the relatively labor-intensive sector, the service sector. If the actual manufacturing land price is lower, this will still facilitate the development of the manufacturing sector and slow down the structural transformation towards the service sector.

In this way, the hump shape pattern will be flatter if considering the land market distortions, but the hump shape pattern in the main findings remain robust.

E.4 Urban Land Provision Problem

One key issue of the counterfactual exercise is that the cost of new urban land provision is not endogenized into the model. Adding the agriculture sector in the model would partly help offset the missing land provision cost since transforming the agriculture sector's land into urban land would reduce the agriculture's output, which functions as one source of land provision cost.

However, the literature suggests that even adding the agriculture sector alone cannot fully address the urban land provision cost issue. In fact, the major urban land provision cost is not from the economic side, but from the policy burdens, for instance, the resettlement and subsidy cost for the affected farmers. the World Bank (2014) points out that in China, the government needs to bear the cost for the affected farmers' resettlement and subsidies and the land development cost while transforming the agricultural land into urban land. For example, the World Bank (2014) provides some rough statistics for the composition of the total government land concession cost (including both the rural and urban land transformations concession cost). Within total government concession cost, the resettlement cost accounts for around 66%, the subsidies to farmers amount to only 3%, the land development cost is 17%, and the rest are subsidies to workers (14%) and land administration cost (1%). Besides, since the property rights of sub-urban and rural land are not crystal clear from 1998 to 2008, there exists some illegal rural land occupations and informal urbanization in many cities [49].

[49] Please refer more details to the World Bank (2014), which provides comprehensive descriptions about China's urban land provision process, reforms and challenges .

In fact, increasing urban land provision does not always bring the cost. It could sometimes increase additional revenue for the local government as a critical source of fiscal income in recent years (Wang and Hui, 2017). the World Bank (2014) summarizes the rough revenue and cost of the urban land provision in recent years and report that government collects land concession income which local governments receive from leasing the use rights to state-owned land to investors.

To sum up, to fully and properly measure the cost and benefit of the urban land provision, adding the agriculture sector in the model alone would not be enough but undoubtedly add the model computational burden. For instance, it requires additional detailed data on the government's land transaction and fiscal balance sheets. On the modeling side, it also requires a government sector that collects the land revenue from the urban land provision and bears the land concession cost at the same time. The government would also redistribute the additional profits to the residents after collecting the revenue. All this information seems pretty scarce on the table and it may also be out of the scope of this paper. And it would be a fascinating topic for future research.