UNIVERSITY OF CALIFORNIA
Los Angeles

Essays on Trade and Welfare

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Economics

by

Yuan Tian

2018

ABSTRACT OF THE DISSERTATION

Essays on Trade and Welfare

by

Yuan Tian

Doctor of Philosophy in Economics

University of California, Los Angeles, 2018

Professor Adriana Lleras-Muney, Chair

My dissertation contributes towards our understanding of effects of trade liberalization on institutional, economic and environmental outcomes. It consists of three chapters. The first, "International Trade Liberalization and Domestic Institutional Reform: Effects of WTO Accession on Chinese Internal Migration Policy" studies the effect of trade liberalization on migration regulations. I study how trade affects labor institutions in the context of China?s Hukou system that regulates internal migration. Chinese local governments were allowed to relax internal migration restrictions after China entered the WTO in 2001. I collect a new dataset on Chinese prefecture-level migration regulations that shows each region?s friendliness to migrant workers. Using these data, I document an increase in pro-migrant regulation around the time of WTO entry. I then consider the role of international trade in triggering this increase by estimating the impact of prefecture-level export and import shocks on migration regulations across 250 Chinese prefectures from 2001 to 2007. I find a positive and significant impact of export shocks on regulations that encourage in-migration. 17% of the impact of export shocks on migration and 9%-15% of their impact on growth operated through changes in regulation.

In my second chapter, "Was Entry into the WTO Worth it: Environmental Consequences of Trade Liberalization", I document that despite the enormous economic benefits from China's accession to the WTO in 2001, the overall welfare gains from trade liberalization may be compromised since pollution from production has also increased. Using plausibly exogenous tariff reductions on Chinese goods caused by the WTO accession, variation in industry composition across cities and variation in pollution intensity levels across industries, I study the effect of trade liberalization on income, pollution and health in China from 2000 to 2005. Using regional tariff shocks as instruments for changes in income and pollution levels, I show that cities which faced a 10% larger GDP per capita increase experienced a 6%-7% larger total mortality rate decline, and regions that faced a 10% larger increase in air pollution levels experienced a 4%-13% larger total mortality rate increase. Overall, if all exports were generated from non-polluting industries, the total mortality rate would have declined by 3.6% more. However, in terms of overall welfare, the gains from income growth outweigh losses from increases in pollution levels.

In the third and last chapter, "Hukou and Labor Misallocation in China", I propose to quantify the changes in misallocation cost due to the Hukou reform. The Hukou system impedes labor mobility across regions, and the marginal productivity of labor may not be equalized spatially. Following Hsieh and Moretti (2017), I plan to use a general equilibrium Rosen-Roback model to measure the effect of Hukou regulations on aggregate growth.

The dissertation of Yuan Tian is approved.

Paola Giuliano

Pablo David Fajgelbaum

Moshe Buchinsky

Adriana Lleras-Muney, Committee Chair

University of California, Los Angeles

2018

To the Sunny Days in Los Angeles

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am especially indebted to my Committee Chair Adriana Lleras-Muney, for her invaluable support and guidance throughout my PhD studies. I am extremely thankful for the support and advice from my other committee members Moshe Buchinsky, Paola Giuliano, and Pablo Fajgelbaum. I thank Brian Kovak, Walker Hanlon, Youssef Benzarti, Simon Board, Christian Dippel, Michela Giorcelli, Edward Kung, Rodrigo Pinto, Nico Voigtländer, and Till von Wachter for insightful discussion. I thank participants in UCLA applied micro proseminar and various conferences for helpful comments. I thank the UCLA Ziman Center for Real Estateâs Rosalinde and Arthur Gilbert Program in Real Estate, Finance and Urban Economics for generous funding. All remaining errors are mine.

Vita

Education

2014	MA, Economics, University of California, Los Angeles
2010	Bachelor, Economics, Tsinghua University

Awards

2017	UCLA Graduate School Dissertation Year Fellowship
2017	UCLA Graduate School Ph.D. Travel Grant
2016	UCLA Economics Travel Grant
2016	Royal Economic Society Conference Grant
2015-2016	Ziman Centerâs UCLA Rosalinde and Arthur Gilbert Program in Real Estate, Finance and Urban Economics, Research Grant
2015	UCLA Economics Travel Grant
2014	Honored Pass in Econometrics Field, UCLA
2011	Graduate Office Fellowship in Economics Department, MSU

2011	Award for Undergraduate Thesis, Tsinghua University
2010-2011	Tsinghua University-Tianyi Comprehensive Performance Scholarship
2010-2011	Tsinghua University-Morgan Stanley Scholarship
2007-2006	Tsinghua University Secondary Freshmen Scholarship

CHAPTER 1

International Trade Liberalization and Domestic Institutional Reform: Effects of WTO Accession on Chinese Internal Migration Policy

1.1 Introduction

Both trade policy and migration policy affect labor markets. Despite extensive research on the effects of each of the two policies on wages, employment, and economic growth, little evidence exists on the interaction of the two — how one policy affects the other, and how the two jointly determine economic growth. Trade shocks can lead to migration policy changes through the interaction of economic and political-economy forces. On the one hand, trade shocks affect labor demand directly as a result of import competition or export growth. On the other hand, labor supply is usually regulated by the government, and trade shocks can affect both the marginal benefit and the marginal cost of allowing for more migrants. Thus, policies on migration issues can be influenced by the direction and the size of trade shocks.

In this paper, I study how an exogenous positive trade shock has spurred relaxation of migration restrictions in China. Nailing down the causal effect of trade on migration policy is empirically difficult. Trade shocks and migration policies are usually determined at the country level, and cross-country comparisons can be biased since both trade policies and migration policies can be correlated with economic growth. In addition, while it is relatively easy to measure changes in trade policy, it is much harder to quantify changes in migration regulations. I solve both the simultaneity problem and the measurement problem by exploiting a specific institutional feature of labor mobility and a big trade shock in China. I find that liberalized trade policies, which increased demand for exports, led to relaxation of migration restrictions. I also find that both liberalizations contributed positively to economic growth.

The Chinese context has several unique features that facilitate the study. First, Chinese internal migration across regions is regulated like international migration is by other nations. The Hukou system ties individuals to their regions of origin, and migrants suffer legal discrimination in both job opportunities and access to public services; regional governments set their own migration policies on immigrant welfare. Second, China experienced a big trade shock after its accession to the World Trade Organization in 2001. Import tariffs on Chinese goods fell when China was granted most-favored-nation (MFN) status, and export volume growth followed. This aggregate shock affected regions within China differently, depending on their initial local industrial composition. Thus, I have a convenient setting, with many region-specific trade shocks and region-specific migration regulations, to study the effect of trade on migration polices.

First, I set up a simple model of multiple regions with government-regulation choice to elicit the connection between trade and migration policies. Regional governments seek to maximize tax revenue but need to provide local amenities to attract migrants. When there is a positive demand shock for goods produced locally, local governments have incentives to relax migration restrictions, but they also face the cost of providing public goods. The model predicts that positive trade shocks lead to relaxation of migration restrictions and that regions with larger output elasticity of migrants relax the restrictions more when trade shocks happen.

I then estimate the effect of trade shocks on migration regulations across 250 Chinese prefectures from 2001 to 2007. The first empirical challenge is the measurement of migration restriction changes. I build a novel dataset on prefecture-level government regulations related to migrant workers and construct a migration-friendliness index to summarize a local government's attitude toward migrant workers.[1] From 2001 to 2007, 168 cities relaxed their migration restrictions to some degree and implemented new regulations covering migrants related to workplace training, wage-arrears prevention, medical and social insurance, and school access.

The second empirical challenge is sorting out the trade effect. I identify trade shocks using a standard methodology; however, I use export shocks instead of import shocks which are more common in the literature. I follow Kovak (2013) and calculate a prefecture's exposure to trade shocks using the interaction of industry-level tariff reductions and prefecture-level industry employment shares. To address the concern that industry-level post-WTO trade shocks might be correlated with pre-WTO industry characteristics, I show that industry-level post-WTO tariff declines were not correlated with pre-WTO export growth or pre-WTO tariff declines. Tariff reductions have come from countries that import Chinese goods and should not be correlated with prefecture-level economic conditions. I show that prefecture-level post-WTO tariff declines could not be predicted by pre-WTO economic growth levels.

Overall, I find that regions that faced a 0.1 percentage point larger trade shock had a 7% higher increase in the migration regulation index. The regulation score of prefectures whose trade shocks were in the upper third of the distribution rose 22% higher than the regulation score of prefectures in the lower third. Further, prefectures with a higher demand for migrants responded more positively to the

[1] The measure is similar to Besley and Burgess (2004).

trade shock, which fits the model prediction.

Next, I evaluate the overall impact of liberalization on migrant flows, employment, wage and GDP growth. The trade shock impacts employment and wages through two channels: directly, through prices, and indirectly, through migration regulations. I use mediation analysis and an instrumental variable approach to identify the effect of regulation changes on economic outcomes. Compared to prefectures whose trade shocks were at the lower third of the distribution, prefectures at the upper third had a 76,000 greater increase in the number of migrants due to the trade shock overall, 17% of which was due to changes in regulation. I also find significant and sizable effects of trade shocks and regulation changes on wages, employment, and GDP growth, and about 9%-15% of the overall trade effect was through migration regulation changes.

The rest of the paper is organized as follows. In Section 1.2, I summarize this paper's contribution to the literature. In Section 1.3, I offer background on how the Hukou system restricts people's mobility and on local governments' power to decide the stringency of the immigration restrictions. I also discuss trade liberalization in China after the WTO accession and provide anecdotal evidence on policy makers seeing WTO accession as an opportunity to further economic reform. In Section 1.4, I present a simple framework to show government decisions in response to trade shocks. In Section 1.5, I describe my data collection and discuss how I measure regulation changes and regional trade shocks. In Section 1.6, I present the empirical results. In Section 1.7, I discuss potential extensions of this paper and conclude.

1.2 Contribution to Literature

This paper contributes to several strands of the economics literature. Recent papers in international trade study local labor market effects of import competition (for example, Autor et al. (2013), Kovak (2013), and Topalova (2010)). In this paper, I focus on export shocks. My key contribution is to study how export shocks affect local migration policies and, through these policies, economic outcomes such as wages, employment, and GDP.

My paper contributes to the literature on the relationship between institutions and economic growth, and I show the impact of trade on institutions. Khandelwal et al. (2013) shows how trade liberalization led to the elimination of quota-induced misallocation. Autor et al. (2016) and Dippel et al. (2017) study the impact of import competition on electoral outcomes in the United States and Germany, respectively. Acemoglu et al. (2005) studies how Atlantic trade intensity affected the protection of property rights and how this institution benefited later economic growth. Levchenko (2013) develops a theoretical model on how trade openness affects institutions in a two-country game. In this paper, I focus on labor mobility restrictions, and I investigate how institutions and economic outcomes affect each other.

This paper contributes to the literature on labor market frictions and trade. Several theoretical papers study the interaction of trade liberalization or migration liberalization: how welfare gains from trade liberalization are affected by labor market frictions or how the effect of migration liberalization is compromised by the existence of trade frictions.[2] Tombe and Zhu (2015), Fan (2015), and Ma and Tang (2016) focus on China, studying the aggregate and distributional effects of international and domestic trade on productivity where labor market frictions exist. While the literature takes labor market frictions as given, I endogenize them in a theoretical model. I also use a novel dataset to construct an empirical measure for the stringency of regulations. Finally, using a plausibly exogenous trade shock, I identify the effect of a trade shock on regulations that affect labor

[2]See Kambourov (2009), Alessandria and Delacroix (2008), Helpman and Itskhoki (2010), and Caliendo et al. (2017).

market frictions.

Several papers have studies the effect of China's WTO accession on both the Chinese economy and the world economy (Autor et al. (2013), Amiti et al. (2017), and Handley and Limão (2017)). Both the output tariff and input tariff cuts have been shown to contribute to the increased productivity of Chinese firms (Brandt et al. (2017)). The reduction in tariffs on intermediate inputs and the reduction in output tariff uncertainty led to both inter-regional migration and intra-region structural transformation (Zi (2017), Erten and Leight (2017), and Facchini et al. (2018)).[3] I focus on trade shocks coming from declines in output tariffs and show that China's WTO accession not only had big impacts on economic outcomes but also changed the Hukou system, a labor institution that affects all aspects of Chinese life.[4]

This study relates to the literature on fiscal competition (for example, Fajgelbaum et al. (2015) and summarized in Wilson (1999)). I show that regions compete to attract a common labor force by providing amenities or subsidies. However, I do not directly address efficiency issues related to such competition. Actually, when other distortions exist in the economy, this competition for labor could be welfare-improving for all. I discuss this possibility briefly in Section 1.7.

Lastly, this paper is related to the literature on the effects of migration on economic outcomes (for example, Card (1990, 2001), Borjas (2003), and Ottaviano and Peri (2012)). While most of the papers in this literature use exogenous increase in migrant flows, my paper goes back to the source of migrant flow changes and emphasizes the importance of regulatory forces.

[3] Tariff barriers are easier to measure than non-tariff barriers and are more commonly used in the literature to measure trade costs. The only paper on quota is Khandelwal et al. (2013), which uses the timing of quota elimination instead of the size of quotas. Goldberg and Pavcnik (2016) argues that tariff barriers and non-tariff barriers are usually highly correlated.

[4] I also investigate the impact of reduction in uncertainty and quotas on textile and clothing in Section 1.6.5.5, and I find that the effects are smaller than the output tariff reduction.

1.3 Background and Motivating Facts

In this section, I start by giving a brief background on the Hukou system. I then show the new data that I collect to measure the changes in the migration policies. Next, I present key aggregate trends that motivate the study. First, I show that the pace and pro-migrant nature of labor reforms in China increased around the time of WTO entry. Second, I show that across industries post-WTO output tariff shocks were associated with export growth. Third, I show that the pro-migrant nature of labor reform was on average stronger in prefectures that were specialized in sectors with higher export exposure.

These facts motivate the key hypothesis in this paper that trade reform increased the incentives for migration reform. I present a simple model to rationalize the decision making of local governments in Section 1.4 and inspect the hypothesis in detail in the econometric framework in Section 1.6.

1.3.1 The Hukou System in China

China's Hukou system is the internal registration system for Chinese citizens. Each individual has a Hukou status associated with a location and a sector (agricultural v. non-agricultural) based on parents' status. To switch sector or prefecture, an individual needs to obtain a temporary Hukou enabling legal migrant status. Even as legal migrants, individuals who switch Hukou are subject to diminished access to public services such as medical insurance or public schools.

Before 2000, the central government held a rigid stand on the Hukou system, and lower-level governments were not allowed to tailor the national policy on the local level. It was difficult for an urban resident to get a Hukou in other prefectures, unless he or she found an official job in an urban area that sponsored Hukou changes. The process was harder for those wishing to switch from agricultural to nonagricultural Hukou. There were very few annual quotas, most of which were assigned to people whose spouse held a nonagricultural Hukou.

The Hukou stystem has been linked to spatial disparities in income (Tombe and Zhu (2015) and Wang and Zuo (1999)). In 2000, 11% of the population was employed in a different prefecture-sector than their assigned Hukou. Migrant workers worked and lived under inferior conditions in terms of lack of protection of their legal rights at work and limited access to local schools and hospitals.[5]

Around 2000, the central government started to soften its stance on issues related to Hukou. The Tenth Five-Year Plan talked specifically about reducing political barriers to migration.[6] In addition, local governments were allowed some discretion to design their own reforms following the central government guidelines.[7] The timing of the reform coincided with the WTO accession; in research articles and interviews with government officials, the WTO accession was described as a chance to reform the Hukou system.[8]

The central government's evolving stance spurred big local responses. Cities started to carry out measures to improve the well-being of legal migrants and set up a pathway for some migrants to get local Hukou. They set up guidelines to protect migrant workers' legal rights in the workplace and also granted partial access to the social safety net and other local amenities. Some prefectures allowed migrant children to enroll in local primary and secondary schools. A few

[5] Source: http://www.gov.cn/zhuanti/2015-06/13/content_2878968.htm.

[6] From the Tenth Five-Year Plan: "We will adapt to the market-oriented employment mechanism ... to have an orderly and reasonable allocation of rural and urban labor." Source: www.people.com.cn/GB/shizheng/16/20010318/419582.html, or www.gov.cn/gongbao/content/2002/content_61966.htm.

[7] According to a 2001 document by the State Council of China, "Local governments should take into consideration local economic and social development levels and conduct reforms that balance population growth, infrastructure, employment and social security, and other welfare programs." Source: www.gov.cn/zhengce/content/2016-09/22/content_5110816.htm

[8] An interview with the Minister of Public Security, Division of Hukou Management, in 2001, writes: "The employment system, education system, and social security system are all evolving, and it is about the time to partially liberalize internal migration. Entering the WTO is an opportunity to change the Hukou system from management to service."

prefectures established a point-based system for application for a local Hukou.[9] Although migration was still regulated, the number of migrants increased. By 2010, the number of Chinese migrants was 260 million, almost double the 2000 figure, and a larger share of migrants moved between prefectures.

1.3.2 New Data on Labor Regulations in China

One of my key contributions is the assemblage of a new dataset that shows changes in labor regulations affecting migrants across Chinese prefectures. I collect all the prefecture-level migrant related regulations from 1995 to 2015 to document the change. For each regulation, I assign a score in $\{-2, -1, 0, 1, 2\}$ to measure how friendly the regulation is to migrants. A score of -2 means the regulation is strongly anti-migrant (for example, evicting all migrant workers who are living in rental houses without proper documents), while a score of 2 means that the regulation is strongly pro-migrant (for example, requiring firms to pay injury insurance for migrants). I explain the details of the index construction in Section 1.5.2.

1.3.3 Key Motivating Facts

1.3.3.1 Trend Break in Migration Regulations around the WTO Entry

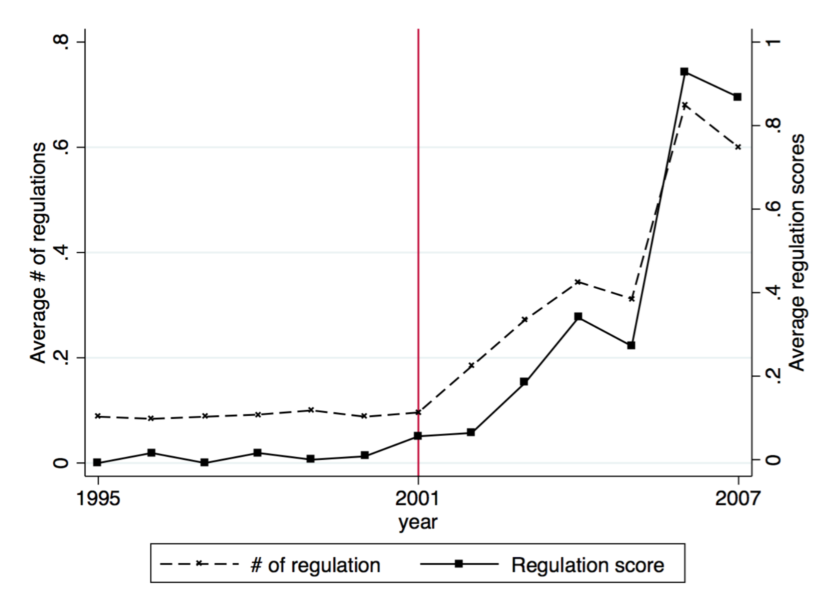
China entered the WTO in November 2001 as the 143rd member country. In the accession agreement, China and the partner countries made commitments regarding import tariff reduction, quota removal, and reduction of other non-tariff barriers. Specifically, China started to enjoy the most-favored-nation (MFN) status with other member countries. This means that among other things, whatever import tariff partner countries imposed on a certain product, the Chinese good

[9]This is similar to the point-based system in President Trump's immigration bill. www.cnn.com/2017/08/02/politics/cotton-perdue-trump-bill-point-system-merit-based/index.html

would be eligible for the same tariff level.[10]

In Figure 1.1, I plot the prefecture-level average number of new regulations on migrant issues and the regulation score. Each dot on the dashed line represents the total number of regulations in each year divided by the total number of cities. Each dot on the solid line represents regulation scores. The trend shows that before 2001, about 0.1 regulation per city per year addressed migrant issues. However, the migrant-friendliness score was essentially zero. After 2001, both the number of new regulations and the score of the regulation friendliness increased substantially. In 2006, for example, there was about one regulation per prefecture, and the average score was about 0.8, indicating that there were more regulations and those regulations were more favorable to migrants than before 2001.

Figure 1.1: Number of migrant/Hukou regulations and regulation score, prefecture-level average, 1995-2007



Note: Each dot is a prefecture-year average. The score is the sum of scores of all prefecture-level regulations related to migrants divided by the number of prefectures. Total number of prefectures is 250. The vertical line corresponds to China's accession to the WTO in 2001.

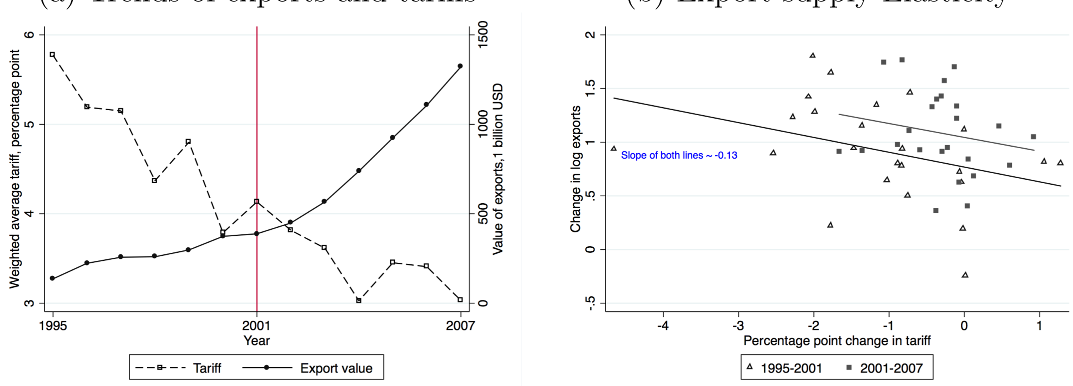
[10]See China's accession protocol: www.wto.org/english/thewto_e/acc_e/completeacc_e.htm.

1.3.3.2 WTO Accession, Tariff Reductions, and Export Growth

This paper focuses on the decline in output tariffs on Chinese goods imposed by countries that import from China, which I refer to as the "export tariff shock." First, I show that the output tariff decline was sizable and that industries that experienced bigger tariff declines also experienced bigger export volume growth. Figure 1.2 shows that the tariff on Chinese goods stood at about 5.8 percentage points in 1995, declined to 4.1 percentage points in 2001, and declined even further to 3 percentage points in 2007.[11] Figure 1.2 shows percentage point changes in tariff on the horizontal axis and change in log exports on the vertical axis. Each dot represents an industry. The hollow triangles are for 1995-2001 and the solid squares are for 2001-2007. In both periods, the fitted lines have a slope of -0.13, meaning that one percentage point reduction in the tariff faced by Chinese exporters induces a 13-14% increase in export values.

Figure 1.2: Declining output tariff and increasing export volume, 1995-2007

(a) Trends of exports and tariffs (b) Export supply Elasticity



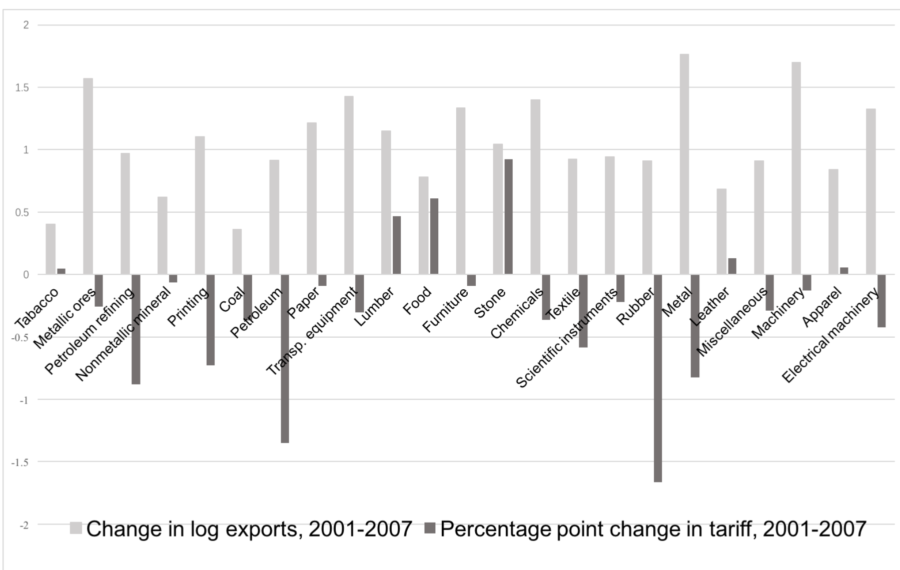
Note: In Panel (a), each dot on the red curve is the weighted average of industrial level tariffs, where the weights are shares of exports in this industry. The industry level tariff is constructed as the weighted average of destination-country tariffs on Chinese exports in the specific industry, where the weights are shares of exports in this destination country in the specific industry in 1995. In Panel (b), each dot is an industry-period. Crosses are for 1995-2001 and squares are for 2001-2007.

I study the post-WTO period of 2001-2007, comparing prefectures that had

[11] The 1995-2001 decline is bigger than the 2001-2007 decline when I use 1995 country-import shares as weights. If I use 2001 or 2007 country-import shares as weights, the decline in the later period is bigger. (See Appendix 1.7.)

bigger versus smaller export tariff shocks. Although it seems that there are no discontinuous changes on the overall tariff level from the pre-WTO period to the post-WTO period, there are substantial changes on the industry level. In Figure 1.3, I plot the change in log export (in light gray) and percentage-point change in tariff (in dark gray) in the 2001-2007 period. Each cluster is an industry represented by its 2-digit SIC code, and the clusters are sorted by the value of exports in the industry in 2000. We can see that changes in tariff levels varied greatly across industries.[12] My identification strategy will rely on differences in export specialization across Chinese prefectures. I will assign a larger export tariff decline to prefectures with larger employment shares in industries facing larger output tariff declines.

Figure 1.3: Distribution of tariff changes and export growth across industries, 2001-2007



Note: Each bar is an industry. Horizontally sorted by value of exports in the industry in 2000.

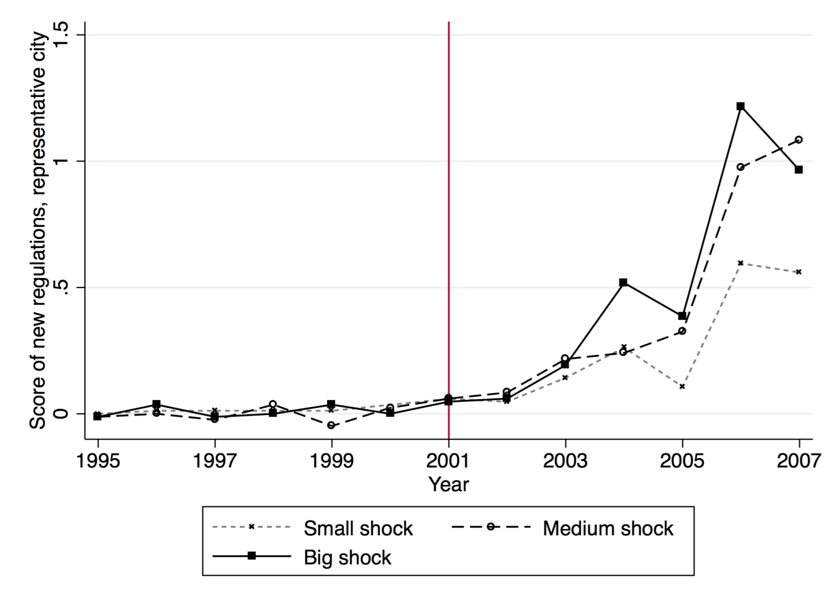
[12]Industry code and industry name can be found in Appendix 1.7.

1.3.3.3 More Exposed Prefectures Changed Migration Regulation More

Once prefectures were given the freedom to modify the Hukou system, their incentives to do so depended on the gains from a flexible labor market. The main conjecture of this paper is that these gains increase with export promotion. I now show some preliminary evidence that suggests this is the case. I will thoroughly inspect this hypothesis in the context of the econometric analysis of Section 1.6.

In Figure 1.4, I plot the trends of regulation scores, dividing prefectures into three groups: the solid line for prefectures with trade shocks in the upper third of the distribution (big shock), the dashed line for prefectures with trade shocks in the middle third (medium shock), and the dotted line for prefectures with trade shocks in the lower third (small shock). Here, the trade shock is calculated as weighted average tariff reductions from 2001 to 2007 (post-WTO period), with prefectures with bigger shocks experiencing larger reductions in output tariffs. It is clear from the figure that although the trends for the three groups of prefectures were very similar before 2001, places with bigger trade shocks after 2001 chose to relax migration restrictions more.

Figure 1.4: Regulation score, prefecture-level average, 1995-2007, three groups by the size of trade shock 2001-2007



Note: Each dot is a year-shock group. The score is the sum of all prefecture-level regulations related to migrants divided by the total number of prefectures. Trade shocks are constructed using the interaction of industry-level tariff declines with prefecture-level industry employment shares. Details of the data and measurement in Section 1.5.

I focus on the export tariff shock for several reasons. First, export tariff decline was direct and salient from the government officials' perspective. The decline in input tariffs also played an important role, but it was more indirect. Second, although policy discussions mentioned import competition, they mostly focused on the competition on high-end goods such as automobiles and on agricultural products. China's comparative advantage was thought to be on labor-intensive or low-skill-intensive goods, and the export expansion in these industries was likely to trigger migration regulation changes.

1.4 A Simple Model of Endogenous Migration Policy and Price Shocks

I build a simple framework of central and local governments' decision making in which the government chooses migration regulations by deciding the level of amenities for migrants. I use the model to elicit the channel through which trade shocks affect regulation changes and other economic outcomes. I show that an overall positive trade shock that increases the price of export goods among all cities will incentivize the central government to initiate Hukou reforms. Then I explain how local trade shocks drive the change of migration regulations at the local government level, conditional on the central government's choice. I also use the model to generate predictions about the heterogeneous effect of trade shocks based on the demand for migrants. Finally, I derive reduced-form equations connecting trade shocks with migrant flow, wage, employment, and GDP, which I estimate in the empirical section. The full model can be found in Appendix 1.7, and I only present the key driving forces and the results here.

I focus on the migration of rural residents to urban areas and assume that urban residents are immobile.[13] There are N cities in the economy and one rural area. The rural area has a supply of workers who can potentially move to one of the cities as migrant workers. Firms in each city produce a unique product that is sold on the international market. They have different output elasticities of migrants, and the product is different across cities. Trade shocks change international good prices and generate labor demand shocks for migrants in each city.

Migrant labor supply is determined by the wage and the amenity level in a city. There are two levels of governments: the local government and the central government. The local government's objective is to maximize its net fiscal profit,

[13] In the 2000 Census data, 76% of all between-city migrants were rural residents with agricultural Hukou.

which is the tax revenue on firms minus the expenditure on migrant amenities, by choosing the level of amenities for migrants.[14] The central government decides on the overall Hukou regime. In the strict Hukou regime, local governments cannot change the level of amenities; in the relaxed Hukou regime, local governments choose the level of amenities to optimize the number of migrants. The economy starts with the strict Hukou regime, and it is costly for the central government to switch to the relaxed Hukou regime. The switch happens only when the difference in national output is big enough between the relaxed regime and the strict regime.

When all cities experience a positive price shock, national output increases markedly when more migrants move to cities. Thus, the central government switches to the relaxed regime. Local governments then choose their own amenity level for migrants, and cities with bigger positive price shocks provide more amenities and attract more migrants.

When the price of goods produced in a city increases, even if the local government does not change amenities for migrants, the wage of migrants goes up, and more migrants move to the city; the inflow of migrants then pushes down wages. The key here is the choice of the local government; it equalizes the marginal benefit and the marginal cost of migrant flows. In the simplest case, the marginal benefit is the marginal tax contribution of a migrant worker, and it is proportional to the migrant wage. The marginal cost is the amenity expenditure. This means that as long as the net effect of price shocks on migrant wages is positive, the effect of price shocks on the migrant amenity is also positive.

Eventually, I solve the percentage increase in the amenity level ($\hat{A}_i$) as a function of the price shock ($\hat{p}_i$) :

$$\hat{A}_i = f(\alpha_i)\hat{p}_i \tag{1.1}$$

[14] I assume that the level of amenities for local workers is fixed.

where α_i is the output elasticity of migrant workers in the production function. When there is a positive price shock, the amenity level also increases ($f(\alpha_i)$ is positive as the solution of the model). In addition, when α_i is bigger, the amenity is more responsive ($f^{'}(\alpha_i) > 0$), meaning that in places that are more migrant intensive (or with higher output elasticity of migrants), a positive price shock leads to bigger changes in amenity level.

Positive price shocks increase migrant flows and total employment, and so does the induced higher amenity level. For migrant wages, the effect of price shocks is also positive, but the effect of the inflow of migrants due to higher amenity levels is negative. However, local workers benefit from the abundance of the migrant labor force.[15] Thus, the overall effect of the increased migrant amenities on wages can be positive or negative depending on the ratio of migrant vs. local labor.

1.5 Data and Measurement

1.5.1 The Definition of a Prefecture/City

China is divided into 31 provinces, which are further divided into 340 prefectures (including four municipalities: Beijing, Tianjin, Shanghai, and Chongqing). Each prefecture contains rural areas and urban areas. Thus, migrant flows could be within a single prefecture from rural area to urban area or between two prefectures.

Some of the 340 prefectures are purely rural areas.[16] The 2001 Prefecture Statistics Yearbook describes the GDP and other economic measures for 264 prefectures; I regard the 76 missing prefectures as rural areas and do not include

[15] I model migrant labor and local labor as different factors of productions. In 2000, the average years of education for urban residents age 15 and above was 10.3, while the number for migrant workers from rural areas was 8.2. The 2-year gap persisted until 2005. Thus, the migrant workers were relatively low-skilled compared to local urban residents. The division between local workers and migrant workers is similar to the division between high-skill workers and low-skill workers.

[16] For example, most prefectures in Yunnan, Gansu, Xinjiang and Tibet provinces.

them in the empirical analysis.[17]

1.5.2 Prefecture-level Measure of Regulation Changes

I collected government regulation documents from the website www.pkulaw.com. This fee-for-service website contains databases including law and regulations (22,148 items), legal news (16,696 items), legal cases (1,955 items), and other law and regulatory information in China.

I use the database of central and local government regulations. The website collected documents from official government websites, government gazettes, repositories of laws and regulations, as well as documents provided by relevant government units; all the sources are recognized by the legislation law.[18] The database contains at least one regulation document from 332 of China's 340 prefectures. Up to December 31, 2016, Shanghai, Beijing and Chongqing have more than 25,000 items; the median number of items per prefecture is 861.

I performed a keyword search of document titles for the following migration-related terms: non-Hukou population, migrant worker, temporary residence, and Hukou. I found 138 items from 1995-2001 and 673 items from 2001-2007, 44% related to labor issues (wage payment, labor union, training, etc.), 18% related to welfare programs (unemployment insurance, injury insurance, medical insurance, pension, etc.), 30% related to administrative issues (Hukou registration), and 9% related to birth control.

Some regulations are beneficial for migrant workers; others are not. Originally, regulations addressed how to manage the non-Hukou population, for example,

[17]The number of prefectures included in the yearbook increases over time. For example, there are 258 prefectures in the 1995 Yearbook. My final empirical analysis uses 250 prefectures; I will explain the sample size in Section 1.5.5.

[18]The local government database includes governmental regulations, regulatory documents, judicial documents and government rules by all provinces, autonomous regions, municipalities, capital cities of provinces, 19 large prefectures designated by the State Council, and other prefectures.

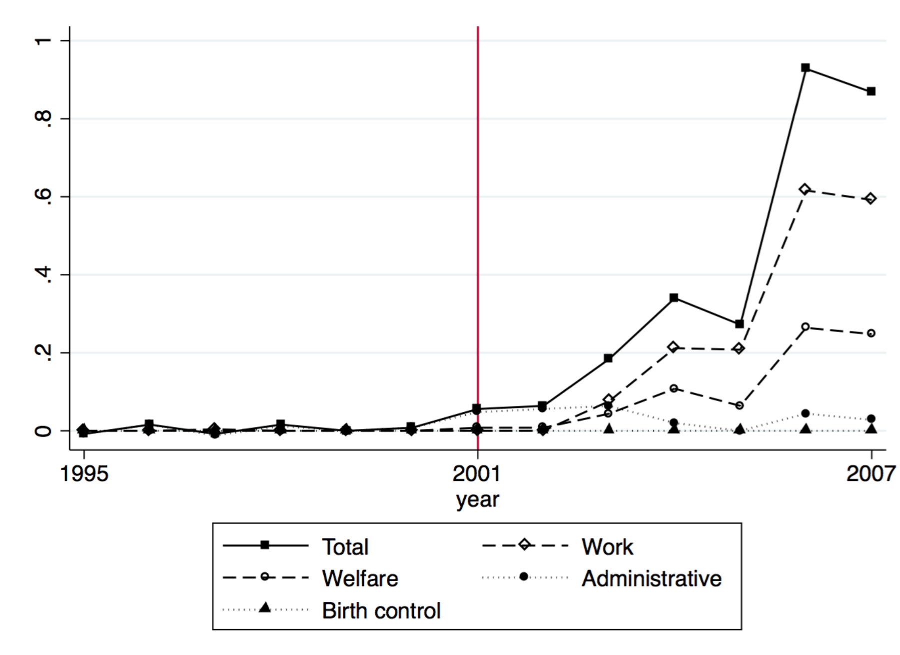
repatriation of migrant workers in rental houses. Starting in 2001, there were more regulations on fees for temporary residence and work permits, providing migrant children with compulsory schooling, urging firms to pay wages and sign contracts, and incorporating migrant workers into the social welfare system. To evaluate the migrant-friendliness of the regulations, I create an index in between of -2 to 2 for each item and sum the item index by prefecture to generate a prefecture-level index or score.[19] Thus, migrant amenities should be an increasing function of the regulation score.

Among the 250 prefectures I analyzed, the median score in the 1995-2001 period is 0, and the maximum is 7; for the 2001-2007 period, the median is 2 and the maximum is 38. Besides the four municipalities, prefectures with very high scores include Ningbo and Guangzhou, which had very strong export-oriented growth.

In Figure 1.5, I plot the regulation score from 1995 to 2007. Each dot is a year-prefecture average score of new regulations on migrant issues. The solid line with solid squares includes all new regulations, the dashed line with hollow diamonds is for work-related regulations, the dashed line with hollow circles is for welfare-related regulations, and the two dotted lines are for administrative- (solid circle) and birth-control-related (solid triangle) regulations. The figure shows that the increase in overall score of regulation is mainly driven by work- and welfare-related regulations. In 2007, for example, the average score for all regulations is about 1, where 0.62 is from work-related regulations, and 0.25 is from welfare-related ones.

[19] See Appendix 1.7 for details on how I constructed the index.

Figure 1.5: Regulation score, prefecture-level average and by topic, 1995-2007



Note: Each dot is a year-prefecture average. The score is the sum of all prefecture-level regulations related to migrants divided by the number of prefectures.

From Table 1.1, we can see that in the 2001-2007 period, 673 new regulations were enacted on migrant issues, with a mean score of 1.08, and 175 prefectures enacted at least one new regulation. In the 1995-2001 period, the numbers are much smaller: 138 new regulations in total and a mean score of 0.05.[20] Fifty prefectures have some regulation, but only 11 of them have positive regulations. Of the 11 positive-regulation prefectures, only one has regulations about work-related issues, but all 11 have administrative-related regulations. Among these 11 prefectures, nine are capital prefectures of provinces, with pro-migrant regulations about receiving local Hukou through purchase of commercial apartments and some specific issues on migrant workers.[21] There were only a few migrant-regulation

[20] One might be concerned that the greater number of regulations since 2001 is a result of more regulations overall. In Section 1.6.5.9, I show that regulations related to fiscal issues or natural resource issues do not have this trend break in 2001. Thus, the greater number of regulations on migrant issues since 2001 is due to more regulation being enacted, instead of being driven by the availability of data.

[21] The nine prefectures are Beijing, Shanghai, Chongqing, Guangzhou, Changsha, Wuhan,

changes before 2001, and they concentrated in few big prefectures on Hukou issues.

Table 1.1: Descriptive statistics on number of regulations and number of prefectures with positive regulations

	# of regulations	Mean score	# of prefectures with			
			Any	(+)	(0)	(-)
2001-2007						
Total	673	1.08	175	162	62	9
Administrative	199	0.35	76	56	39	9
Birth control	59	0.35	36	0	36	0
Work	296	1.56	128	128	0	0
Welfare	118	1.67	64	64	0	0
1995-2001						
Total	138	0.05	50	11	48	9
Administrative	110	0.04	46	11	43	9
Birth control	26	0	20	0	20	0
Work	1	1	1	1	0	0
Welfare	0	0	0	0	0	0

Note: Data source is www.pkulaw.com, and I assign a score to each regulation document on its migrant-friendliness. Administrative regulations are about administrative procedures, such as applying for temporary resident permit; a positive administrative regulation could be on simplifying the procedure and making it easier for migrants to obtain legal status. Regulations on Hukou-related issues are also classified as administrative; a positive Hukou regulation could be on lowering standards for obtaining local Hukou. Birth control related regulations are about enforcing birth control on the migrant population; a positive regulation could be about offering free medical check-ups for pregnant women in migrant families and vaccination for migration children. Work-related regulations address issues such as wage payment enforcement, contracts, and skill training; a positive regulation could be about setting up guidelines for wage payment enforcement. Welfare-related regulations discuss issues related to including migrant workers in the social safety net; a positive regulation could be about forcing firms to pay medical insurance for migrant workers.

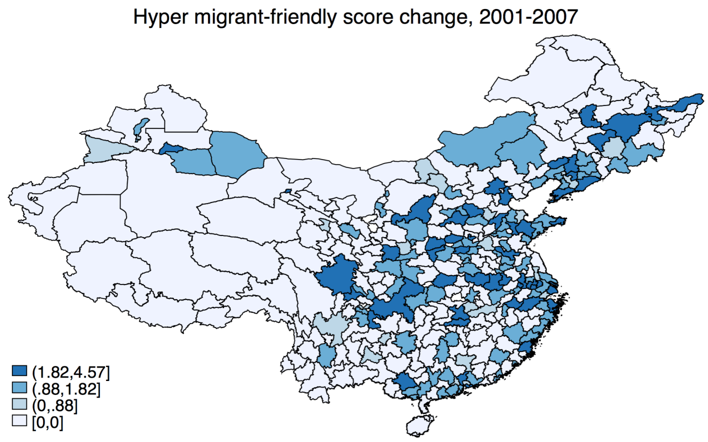
Figure 1.6 shows the geographical distribution of the new regulations. The total regulation score is the sum of all prefecture-level regulations related to migrants. Then I do an inverse-hyperbolic-sine transformation of the total regulation score.[22] Changes are taken from 2001 to 2007. Overall, coastal cities had more

Huhehaote, Wulumuqi, and Xi'an. The other two prefectures are Xiamen and Huizhou.

[22] I use the inverse-hyperbolic-sine transformation because there are 6 prefectures with negative total scores in 2001. However, the correlation between the changes in the inverse-hyperbolic-sine-transformed total regulation score and the changes in the log(regulation score+1) is 0.9925 for the 2001 to 2007 period, when I replace the negative regulation score of the 6 prefectures in 2001 to be 0. Thus, in the following text, I use the log transformation instead of the inverse-hyperbolic-sine transformation for the ease of interpretation. I show the replication of Table 1.2 using the inverse hyperbolic sine transformation in Appendix 1.7 to show the equivalence of the

changes, but many inland cities also made substantial changes.

Figure 1.6: Geographic distribution of regulation changes, 2001-2007



Note: Each bordered area is a prefecture. The regulation score is the sum of all prefecture-level regulations related to migrants. Then I do an inverse-hyperbolic-sine transformation of the total regulation score. Darker blue means that the prefecture became very migrant friendly from 2001 to 2007, and the lighter the color, the smaller the change.

1.5.3 Regional Trade Shocks

I use tariff information from the World Bank TRAINS dataset to calculate regional trade shocks following Kovak (2013). I construct a shock that corresponds to the price change $\hat{p}_i$ in my theoretical model. The regional shock in prefecture i and from time t to t' is

$$RTC_{it} = \sum_j \beta_{ij} \hat{P}_{ijt}$$

$$\text{where} \qquad \beta_{ij} = \frac{\lambda_{ij} \frac{1}{\theta_{ij}}}{\sum_{j'} \lambda_{ij'} \frac{1}{\theta_{ij'}}}$$

and $\lambda_{ij} = \frac{Lij}{Li}$ is the fraction of regional labor allocated to industry j and $1 - \theta_{ij}$ as the cost share of labor in industry j. λ_{ij} and θ_{ij} can be calculated from the firm-level data. $\hat{P}_{ijt}$ is the price shock to industry j in region i from time t to t',

two measures.

and it is measured using export tariff (with superscript X), import tariff (with superscript M), and import tariff on intermediate goods (with superscript I):

$$\hat{P}^X_{ijt} = \hat{P}^X_{jt} = -\left[\ln(1 + \text{tariff}^X_{jt'}) - \ln(1 + \text{tariff}^X_{jt})\right]$$

$$\hat{P}^M_{ijt} = \hat{P}^M_{jt} = \ln(1 + \text{tariff}^M_{jt'}) - \ln(1 + \text{tariff}^M_{jt})$$

$$\hat{P}^I_{ijt} = \sum_{j'} \frac{\text{input}^{j'}_{ij}}{\sum_{j''} \text{input}^{j''}_{ij}} \left[-\ln(1 + \text{tariff}^M_{j't'}) + \ln(1 + \text{tariff}^M_{j't})\right]$$

I construct these measures for two periods: 1995-2001 and 2001-2007. I take industrial composition from the 2000 Industrial Enterprises Survey, which is conducted on Chinese manufacturing firms with annual sales of more than 500 million RMB and includes basic firm information such as name and address, financial information on sales, export values, fixed capital, wage payment and total sales cost, and total employment.[23] There are 145, 546 firms in 2000 with positive sales revenue and wage information, more than 10 employees, and a valid industry code. The industry code is the 4-digit Chinese Industry Code, which I aggregate to the 2-digit level. The 2-digit Chinese Industry Code is slightly finer than the 2-digit SIC code; I include the crosswalk between the codes in Table 1.17.

Tariff data is on the 2-digit SIC level from the World Bank.[24] The tariff on Chinese exports is calculated as the weighted average of import tariff imposed by each destination country, with the 1995 export share as weights:

$$\text{tariff}^X_{jt} = \sum_n \frac{X^{cn}_{j,1995}}{\sum_{n'} X^{cn'}_{j,1995}} \text{tariff}^{cn}_{jt}$$

[23]The 1995 Industrial Enterprise Survey data is not available.

[24]Source: wits.worldbank.org

where $X_{j,1995}^{cn}$ is the Chinese exports to country n in industry j in 1995 and tariff_{jt}^{cn} is the import tariff on Chinese exports to country n in industry j in year t. Chinese import tariffs are directly taken from the World Bank Database.

I use the input-output table from the 2002 Chinese Regional Input-Output Table to calculate each industry's contribution to a certain industry and to construct $\hat{P}_{ijt}^{I}$. The input-output table is available only on the province level; thus, my assumption here is that prefectures in the same province have the same input-output structure.

I also the use the 2001 Industrial Enterprises Survey data to calculate the prefecture-level share of total sales in state-owned enterprises.[25]

1.5.4 Data on Migrants

The information about the number of migrants is from the 2000 and 2010 censuses and a 1% population survey (2005). The 2000 individual data is a 0.1% random sample of the population, and the 2005 data is a 0.2% random sample of the population. I do not have the individual data for 2010, and I use instead aggregate prefecture-level data for the analysis.[26]

A person is defined as a migrant if he or she has been living in a place other than the Hukou registration place for more than six months or has left the Hukou registration location for more than six months. There were 144 million migrants in 2000, 0.39 million per prefecture. In 2010, the number increased to 261 million, 0.77 million per prefecture.

There is also information about how far the person migrated. If the Hukou prefecture and the residence prefecture are the same, I define the migrant as a

[25] I divide firms into SOEs and private firms.

[26] The census and 1% population survey are conducted via personal visits. To address potential issues related to under-reporting, the census bureau randomly samples some neighborhoods after the census concludes and check the non-response rate. The non-response rate in the 2000 census is 1.81%. www.stats.gov.cn/tjsj/ndsj/renkoupucha/2000pucha/html/append21.htm.

within-prefecture migrant. If the Hukou prefecture and the residence prefecture are different but in the same province, I define the migrant as a between-prefecture within-province migrant. If the Hukou province and the residence province are different, I define the migrant as a between-province migrant. I decompose the total number of migrants into these three categories to see if trade shocks and regulation changes affected them differently.

When I study the effect of the 2001-2007 trade shocks on migrant flows, I use the 2000 and 2010 data, because it could have taken time for the regulations to affect actual migrant flows. Later, I exploit the timing of the regulation change and migrant flows to show whether the regulation drives the migrant flows or the other way around, and I use all three years of data.

The 2005 Population Survey contains a wealth of information on respondents.[27] For example, the respondents were asked about their medical insurance, pension, unemployment insurance, terms of contract, and wages. I use the 2005 social insurance measures to check whether in places with more pro-migrant regulation the migrant welfare is higher, in the sense that migrants enjoy more social insurances and are paid higher wages. Also, industries are identified by 2-digit SIC code. I use the industry classification to construct industry-level migrant share of total employment, i.e., the industry-level migrant intensity.

1.5.5 Other Prefecture-Level Measures

Total population, total urban employment, wage, and GDP data at the prefecture level come from the Prefecture Statistics Yearbook. There are 258 prefectures in 1995, 264 in 2001, and 286 in 2007. The Yearbook contains primarily statistics for the urban part of the economy and intentionally excludes some rural prefectures. For example, Gansu province has 12 prefectures, but only six are included in the

[27] The 2000 Census also has the industry and occupation information, but the coding is not standard GB code. There is no information on social insurance or wages in the 2000 sample.

2001 Yearbook. The number of prefectures in the Yearbook increases over time as more prefectures become urbanized. My final sample includes 250 prefectures from the 2001-2007 period; I drop Yulin Prefecture in Guangxi Province due to its border change, one prefecture with missing industrial composition data, and 12 other prefectures where 20% of employment is in the petroleum industry. I drop these 12 prefectures because their cities differ from other cities in many dimensions, but the results are unchanged if I keep these 12 prefectures and include the petroleum industry employment share as a control.

The average wage data, from administrative reports, includes not only the wages of people working in firms but also of people working in the government and other administrative working units.[28] Total urban employment includes urban residents who working in the public sector and the private sector as well as individual laborers.

Since local government officials are promoted based on GDP growth rate, they might be incentivized to manipulate their prefecture-level GDP data. I use night-light satellite data to check the validity of the GDP data.[29] In 2001, the correlation between per capita GDP and night-light intensity is 0.7. I use the GDP from the Yearbook as my main measure of economic activities, but the results are similar when I use the principal component of per capita GDP and night-light intensity.

---

[28] Another way to calculate the average wage is to use the Industrial Enterprises Survey data. The correlation of the two wage measures is 0.8 across the 250 prefectures in 2001, and a linear regression with no constant term generates a coefficient of 1.08. I opt to use the wage data from the Yearbook because it covers all sectors of the economy.

[29] NASA night-light data can be downloaded from http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html.

1.6 Empirical Results

1.6.1 Effects of Trade Shocks on Regulation Changes

1.6.1.1 Econometric Framework and Identification

Did trade induce changes in labor regulations? I now use regression analysis to show effects of trade shocks on regulation changes in detail. I am interested in estimating the following equation:

$$\Delta \ln(\text{Regulation Score})_{it} = \alpha_0 + \alpha_1 \text{Trade Shock}_{it} + X_{it}\Gamma + \epsilon_{it},$$

where i represents a prefecture, and t represents a time period. X_{it} is the vector of potential confounding factors that could be correlated with the trade shocks. Δ represents the change during the time period, and I use change-in-log specifications to follow the model predictions closely. α_1 is the coefficient of export tariff shocks on changes in log migration regulation score.

Identification of α_1 requires that conditional on X_{it}, there are no unobservables that are correlated with the export tariff shocks and have a direct impact on migration regulation changes. I address two types of identification issues here. First, on the industry level, the WTO-induced decline in tariffs on Chinese exports should be uncorrelated with pre-WTO trends, such as pre-WTO export growth and pre-WTO tariff reduction. Otherwise, the regional trade shock will capture pre-existing industry characteristics instead of WTO shocks. Second, on the prefecture level, the WTO-induced regional trade shocks should not be correlated with pre-WTO trends, such as pre-WTO GDP growth and wage growth.

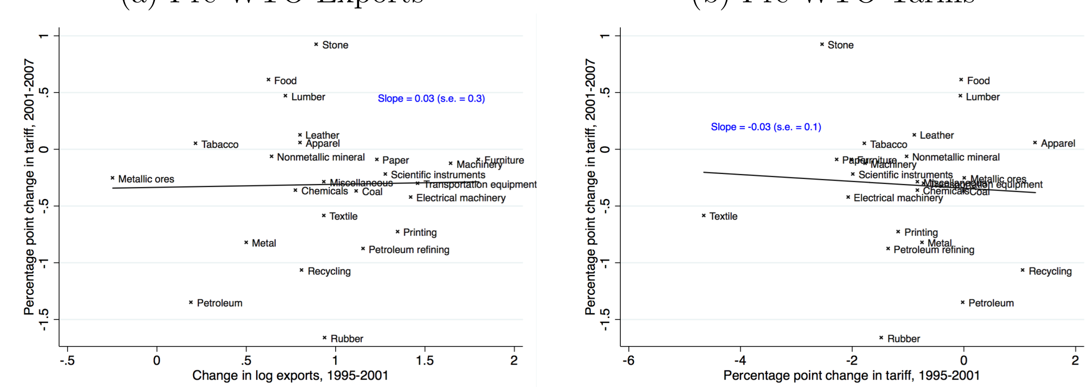
First, I argue that although China's WTO accession was a lengthy process involving lots of preparation and negotiation, the post-WTO tariff decline was still a shock to industries. China obtained MFN status after the WTO accession, and the resulting tariff reductions on Chinese exports were mainly determined by

WTO rules.

I then show that, empirically, the decline in tariffs was a shock in the sense that, across industries, the post-WTO tariff declines could not be predicted by either the pre-WTO export growth or the pre-WTO tariff decline. Thus, there were still relative industry "winners" and "losers" due to the WTO accession. In Figure 1.7, I plot the percentage-point change in tariff in the 2001-2007 period on the y-axis and the percentage change in exports in 1995-2001 on the x-axis. The linear fitted line has an insignificant coefficient of 0.03, meaning that the industries that had bigger pre-WTO export growth were not the ones that experienced bigger post-WTO tariff cuts. In Figure 1.7, I plot the percentage-point change in tariffs in 2001-2007 against the percentage-point change in tariff in 1995-2001. The linear fitted line has a coefficient of -0.03 and is statistically insignificant. This indicates that the industries that benefited more pre-WTO were not the ones that benefited more post-WTO.

Figure 1.7: 2001-2007 tariff declines against 1995-2001 export growth and tariff declines

(a) Pre-WTO Exports (b) Pre-WTO Tariffs



Note: Each dot is a 2-digit SIC industry.

Figure 1.7 and Figure 1.7 also help to address the concern that some unobserved domestic policies might target industries where tariffs happened to decline more or less than other industries in the post-WTO period.[30] If there were pre-WTO industry policies that intended to help or hurt certain industries, these policies were not correlated with post-WTO tariff changes; if there were post-WTO industrial policies that responded to pre-WTO export growth, they were also not correlated with the post-WTO tariff changes.

Second, I argue that decline in tariffs in various industries translated to prefecture-level shocks that should not be correlated with local economic conditions other than through the prefecture-level industrial composition. Except for the petroleum industry, which was concentrated in a few prefectures, other industries were located throughout China.[31] This indicates that shocks to a specific industry could affect all prefectures that produced in the industry and that the size of the impact would be determined by the importance of the industry in the prefecture (for example, the prefecture-level employment share of the industry). Figure 1.8 plots the trends in wages of prefectures with small, medium, and large trade shocks, and the three trends from 1995 to 2001 are not statistically different from each other. Figure 1.8 plots the trends in per capita GDP; here, there seems to be a slight divergence among the three groups from 1995-2001. The three trends are not statistically different from each other, but I will control for 1995-2001 wage and per capita GDP growth in the regressions to be conservative.[32]

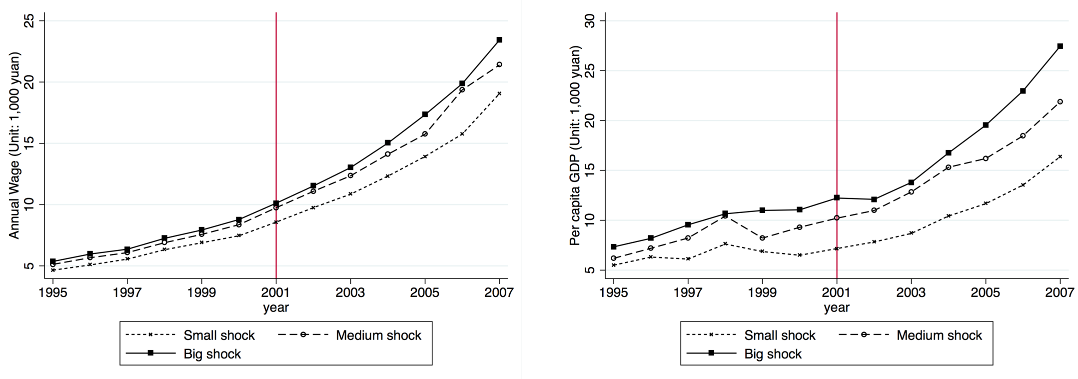
[30]For example, the Chinese government provided value-added tax rebates for exporting firms to encourage exports.

[31]See Herfindal Index Distribution in Appendix 1.7.

[32]See additional trend graphs in Appendix 1.7.

Figure 1.8: Trends of wage and per capita GDP, by size of trade shocks

(a) Wages (b) Per capita GDP



Note: The data is from prefecture yearbook. I divide prefectures into small-, mdedium-, and large-trade-shock groups as in Figure 1.4.

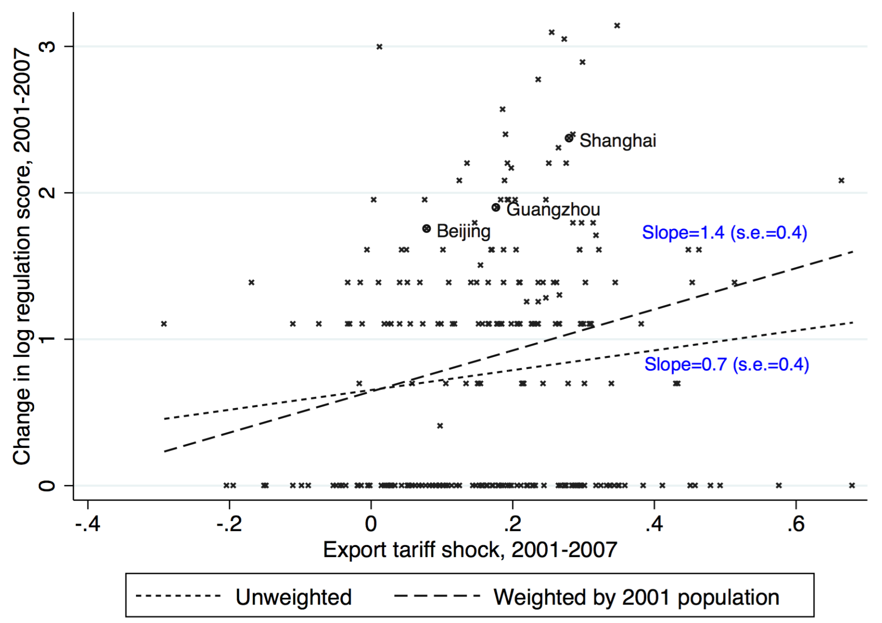
1.6.1.2 Main Results

I first investigate the relationship between trade shocks and migration regulation changes from 2001 to 2007 in Figure 1.9. The horizontal axis depicts the export tariff shock in 2001-2007; a bigger export tariff shock corresponds to lower export tariffs and effectively higher export prices. The vertical axis depicts the post-WTO change of log regulation score, and each dot is a prefecture. The dashed line is the linear fitted line with 2001 population size as weights, and the dotted line is the unweighted linear fitted line.

Figure 1.9 resembles the previous trend graphs for the post-WTO period: prefectures that experienced more positive trade shocks saw their regulation score rise, meaning they became more friendly to migrants. The slope ranges from 0.7 to 1.4, and are statistically significant at the 5% level. By comparison, the same regressions in the pre-WTO period give slopes of 0.02-0.03 and are statistically insignificant.[33]

[33] See the corresponding pre-WTO period graph in Section 1.6.5.1.

Figure 1.9: Bigger trade shocks, more pro-migrant regulation change, 2001-2007, 250 Chinese prefectures



Note: Each dot is a prefecture.

Table 1.2 shows the regression results of changes in log migration regulation score on trade shocks. I start with Columns (1)-(8) where the outcome variable is the change in log regulation score from 2001 to 2007. All columns have standard errors clustered at province level to account for potential spatial correlations of laws and regulations at the province level.

In Column (1), I add only export tariff shocks from 2001 to 2007. The coefficient 0.68 is statistically significant at the 5% level. It implies that a 1 percentage point higher export tariff shock increased the change in log regulation score by 0.71, which is equivalent to 71% bigger regulation score increase. As in Figure 1.4, I divide prefectures into three group: (1) prefectures with big trade shocks (0.33 percentage point tariff changes on average); (2) medium-shock ones (0.18 percentage point); (3) small-shock ones (0.02 percentage point). Thus, compared with small-shock prefectures, the big-shock prefectures experienced 21% higher increase in regulation score; the difference is equal to 0.26 standard deviation of the score increase in the 2001-2007 period.[34]

Table 1.2: Bigger trade shocks, more regulation relaxation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	Δ log regulation score, 2001-2007								Placebo, 1995-2001	
Export tariff shock	0.68**	1.19***	1.09***	1.09***	0.97**	0.94**	0.85*	0.76*	-0.02	
2001-2007	(0.32)	(0.39)	(0.39)	(0.39)	(0.42)	(0.41)	(0.43)	(0.42)	(0.03)	
Import tariff shock		-0.09**	-0.08**	-0.08**	-0.06	-0.07*	-0.09**	-0.06	-0.01*	
2001-2007		(0.03)	(0.04)	(0.04)	(0.06)	(0.04)	(0.04)	(0.06)	(0.00)	
Intermediate tariff shock		0.18	0.20*	0.20*	0.10	0.23*	0.13	0.11	0.00	
2001-2007		(0.11)	(0.11)	(0.11)	(0.17)	(0.12)	(0.13)	(0.17)	(0.03)	
Log regulation score			0.71***		0.69***	0.63***	0.60***	0.59***		
2001			(0.14)		(0.14)	(0.14)	(0.16)	(0.15)		
Δ log regulation score				0.71***						
1995-2001				(0.14)						
Export tariff shock					0.22			0.20		0.14
1995-2001					(0.17)			(0.18)		(0.10)
Import tariff shock					0.01			0.02		-0.02
1995-2001					(0.02)			(0.02)		(0.01)
Intermediate tariff shock					0.12			0.07		-0.03
1995-2001					(0.10)			(0.08)		(0.05)
Δ log wage						0.72**		0.49		
1995-2001						(0.30)		(0.35)		
Δ log GDP p.c.							0.49***	0.43***		
1995-2001							(0.13)	(0.15)		
Observations	250	250	250	250	237	237	237	237	238	241
R-squared	0.01	0.07	0.12	0.12	0.13	0.13	0.15	0.16	0.01	0.01

Note: Standard errors are clustered at the province level. Mean (sd) of Δ log regulation score, 2001-2007 is 0.77 (0.82), 1995-2001 is 0.06 (0.26).
Mean value of export tariff shock, 2001-2007 is 0.18 (0.15), 1995-2001 is 1.23 (0.40).
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

I add potential confounding factors in Columns (2)-(8) to check the robustness of the result. First, regarding trade shock, I focus on export tariff shocks, since export tariffs are imposed by countries that import Chinese goods and should not be correlated with local economic factors. However, if industries that experienced big export tariff reductions happen to be the ones that experienced big (small) import tariff reductions or big (small) reductions in import tariff on intermediate inputs, then the estimate of the export tariff shock effect will be biased. Thus, controlling for other types of trade shocks could be important. Second, since the export tariff shock is calculated as the interaction of industrial-level tariff reduction and prefecture-level industry shares, if industry shares were correlated with other variables that were important determinants of regulation change, then

[34] Mean of score increase is 77% and the standard deviation is 82%.

not including those variables would cause omitted variable bias. Such variables could be the baseline value of regulation scores, the lagged change in regulation score, the lagged trade shock, or lagged economic growth rates.

Columns (2)-(8) add import and intermediate tariff shocks. In addition to the three trade shocks, Column (3) and Column (4) control for the initial level (2001) and past changes (1995-2001) in regulation scores. Column (5)-(8) add past trade shocks and past economic growth. Column (5) controls for the lagged trade shocks, which are export tariff shocks, import tariff shocks, and intermediate tariff shocks from 1995 to 2001. Column (6) controls for change in log wages from 1995 to 2001, while Column (7) controls for changes in log per capita GDP from 1995 to 2001. Column (8) controls for all the variables mentioned. Columns (2)-(8) show estimates for the effect of export tariff shocks from 2001 to 2007, which range from 0.68 to 1.19 and are all statistically significant at the 10% level; all of them are within the 95% confidence interval of the estimate in Column (1), meaning that the effect of export shocks on regulation changes is quite robust.

It is useful to see whether the trade shocks from 2001 to 2007 could predict changes in log regulation scores in the pre-WTO period (1995-2001). If yes, then the industrial composition itself can predict the regulation change in any period, and tariff shocks that happened in the 2001-2007 period are relevant for regulation changes before. In this case, tariff changes in 2001-2007 are not good measures for the period-specific trade shocks, and we cannot identify the WTO effect. The result in Column (9) denies this possibility.

Finally, I check whether the trade-regulation relationship existed before the WTO accession. In Column (10), I regress the regulation change in 1995-2001 on trade shocks in 1995-2001, and the coefficients are insignificant. This is consistent with the observation in previous sections about the timing of the regulation change.[35] Although trade shocks happened in the previous period as well, they

[35] This is also consistent with the absence of correlation shown in Figure 1.13 in Section 1.6.5.1.

did not result in big regulation changes across prefectures, because the central government did not allow for the reform.

Overall, the findings are consistent with the hypothesis that in the post-WTO period, places with bigger trade shocks relaxed migration restrictions more.

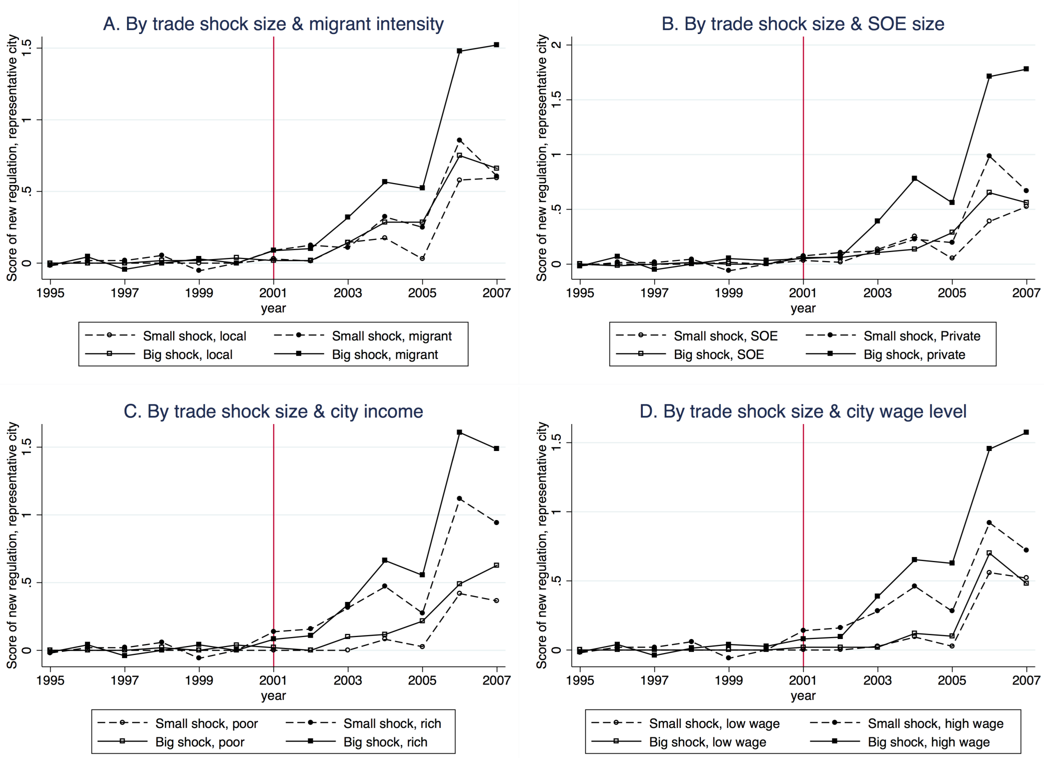
1.6.1.3 Heterogeneous Effects

In the model section, I find that places with bigger migrant-intensive industries should respond more to the trade shock. I investigate this heterogeneous effect using four empirical proxies for a prefecture's migrant intensity (which is α_i in the model). First, I calculate the migrant share of employment by industry using the 2005 mini-census data to measure industry level migrant intensity. Then I calculate each prefecture's migrant intensity by interacting each prefecture's employment share of manufacturing industries in 2000 with the industry-level migrant intensity. Second, state-owned enterprises (SOEs) are usually more restrictive in Hukou requirements and hire more locals than migrants. Thus, the employment share of private firms (or non-SOEs) will be positively correlated with migrant intensity. Third, I find that prefecture income level is empirically positively correlated with migrant intensity. This could be because richer prefectures tend to have more diversified industrial composition and rely less on SOEs. Thus, I use per capita GDP and wage as proxies for the migrant intensity.

In Figure 1.10, I divide prefectures into four groups, depending on the 2001-2007 trade shock size and one of the four proxies for migrant intensity in 2001, with the median value as the cutoff. In Panel A, the four groups are: (1) big trade shock and migrant-intensive prefectures (solid line with solid squares), (2) big trade shock and not-migrant-intensive prefectures (solid line with hollow squares), (3) small trade shock and migrant-intensive prefectures (dashed line with solid dots), and (4) small trade shock and not-migrant-intensive prefectures (dashed line with hollow dots). In Panel B, I use private firm employment share as the proxy for migrant intensity. In Panel C and Panel D, I use per capita GDP and wage as proxies, respectively.

The four graphs confirm the heterogeneous response to trade shocks predicted by the model: prefectures that experienced bigger trade shocks and were more migrant-intensive changed migrant regulations the most, and the ones that experienced smaller trade shocks and were less migrant intensive changed regulations the least.

Figure 1.10: Regulation score, prefecture-level average, 1995-2007, by the size of trade shock in 2001-2007, and by migrant intensity in 2001



Note: Each dot is a year-shock-type group. Average of all prefecture-level regulations related to migrants. Panel A divides prefectures into four groups. The small-shock and local group represents prefectures whose post-WTO trade shock was below the median and migrant intensity was below the median. Migrant intensity is defined as the interaction of prefecture-level industry employment share in 2000 interacted with industry-level migrant share of employment in 2005. Panel B uses the 2001 prefecture-level employment share of private firms as the measure for migrant intensity; Panel C uses the 2001 prefecture-level log per capita GDP; and Panel D uses the 2001 prefecture-level wage.

Table 1.3 shows similar findings with regression analysis. The regression equation is as follows:

$$\Delta \ln(\text{regulation score}_{it}) = \beta_0 + \beta_1 \text{TS}_{it} + \beta_2 \text{I}_{it} + \beta_3 \text{I}_{it} * \text{TS}_{it} + X_{it}\Gamma + \epsilon_{it}$$

where TS_{it} is the export tariff shock in prefecture i and time period starting at $t = 2001$, I_{it} is one of the four measures for migrant intensity in prefecture i and year $t = 2001$. In Table 1.3, Columns (1), (3), (5), and (7) show export tariff shock from 2001 to 2007, the variable I, and the interaction of export shock with I. Columns (2), (4), (6), and (8) show additional controls such as pre-WTO trade shocks and pre-WTO wage and GDP growth as in Table 1.2 Column (8). All columns control for import and intermediate trade shock, and log regulation score in 2001.

Table 1.3: More migrant-intensive prefectures responded more to trade shocks, 2001-2007

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Δ log regulation score, 2001-2007							
I for interaction, 2001	Migrant intensity		Private firm share		Log(wage)		Log(GDP p.c.)	
Export tariff shock, 2001-2007	-3.26	-4.40	-0.92	-1.37**	-25.28*	-30.19**	-8.47*	-10.22*
	(2.37)	(2.72)	(0.56)	(0.64)	(13.14)	(13.66)	(4.49)	(5.17)
I, 2001	-1.45	-2.79	-0.39	-0.68**	0.21	0.18	0.32**	0.29*
	(1.49)	(1.81)	(0.25)	(0.30)	(0.30)	(0.37)	(0.13)	(0.15)
Export tariff shock×I	12.33*	14.92*	3.93***	4.18***	2.84*	3.35**	1.00*	1.18*
	(7.20)	(8.17)	(0.87)	(0.91)	(1.45)	(1.50)	(0.51)	(0.58)
Controls		X		X		X		X
Observations	250	237	250	237	250	237	250	237
R-squared	0.13	0.17	0.16	0.20	0.17	0.21	0.23	0.26
Mean (s.d.) of I	0.34 (0.05)		0.55 (0.23)		9.11 (0.28)		8.94 (0.64)	

Note: Standard errors are clustered at the province level. Mean (sd) Δ log regulation score, 2001-2007 is 0.77 (0.82), mean (sd) export tariff shock is 0.18 (0.15). All columns control for import and intermediate tariff shocks, 2001-2007, log regulation score in 2001.
Columns (2)(4)(6)(8) also control for lagged trade shocks and lagged wage and gdp growth rate, 1995-2001, as in Table 2 Column (8) .
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Column (1) shows a positive interaction effect for migrant intensity and export tariff shock (12.33), and a negative coefficient for export tariff shock (-3.26). At the mean value of migrant intensity (0.34), the overall effect of export tariff shock becomes positive. This means that cities with bigger demand for migrants responded more positively to the export tariff shock. Column (2) shows similar results. Columns (3) and (4) use private-firm share of output, which is positively correlated with migrant intensity, and there is a positive interaction effect as well. It means that cities where private firms dominated responded less positively to the trade shock.

Column (5) shows a positive interaction effect for initial wage and export tariff shock (2.84), and a negative coefficient for export tariff shock (-25.28). Approximately at the mean value of log wages in 2001 (which is 9.11), the overall effect of export tariff shock becomes positive. This means that richer cities responded more positively to the export tariff shock. Column (6) has similar interpretations. Columns (7) and (8) use per capita GDP instead of wage, and the result is similar: richer prefectures responded more positively, and the overall effect of export tariff shock became positive at the mean value of log per capita GDP. Since income level and migrant intensity are positively correlated, results in Columns (5)-(8) confirm the earlier finding.

Overall, I find that migrant-intensive prefectures responded more positively to the trade shock, and the finding is consistent with the theoretical prediction.

1.6.2 Effects of Trade Shocks and Regulation Changes on Migrant Flows, Wages, and per capita GDP

1.6.2.1 Econometric Framework and Identification

Trade shocks affect economic outcomes (such as migrant flows, wages, employment, and per capita GDP) through two channels: directly, through prices, and indirectly, through migration policies. For migrant flows and employment, both effects are strictly positive: output price increase will attract more workers and so will more relaxed migration policy. Output price increase will also directly increase wages and per capita GDP, holding migrant flows and employment constant. However, the regulation change will decrease migrant wages and at the same time increase wages of local workers. Thus, the overall effect of regulation on wages and per capita GDP depends on the composition of the workforce.

First, I estimate reduced-form overall effects of trade shocks on economic outcomes by using the following regression equation:

$$\Delta Y_{it} = \gamma_0 + \gamma_1 \text{Trade shock}_{it} + X_{it}\Pi + \xi_{it}$$

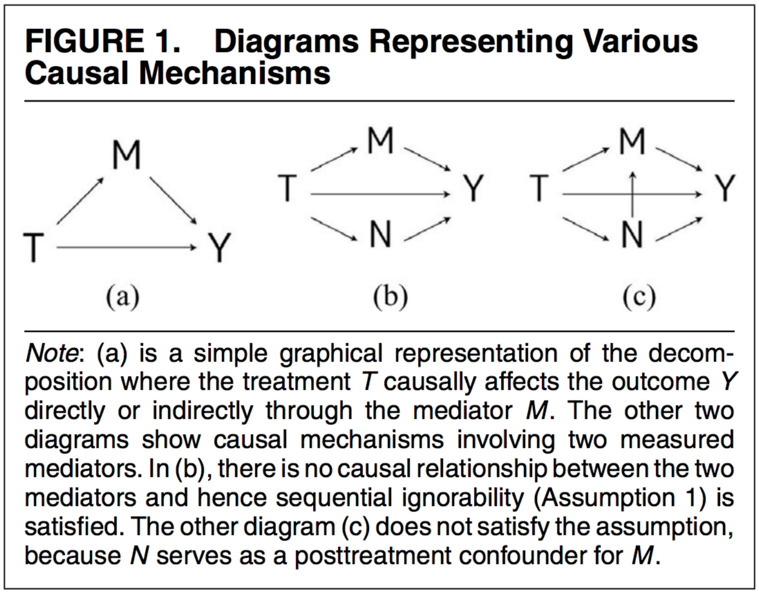
where ΔY_{it} can be the change of migrant share of population in prefecture i in the 2001-2007 period, the change in log migrant stock, the change in log wages, the change in total urban employment, or the change in per capita GDP. γ_1 will capture the reduced-form effect of trade shocks on outcome variables that include both the direct price channel and the indirect regulation channel. To identify γ_1, there should be no omitted variable that is correlated with the trade shock and affects the economic outcomes directly. The discussion in Section 1.6.1.1 shows that the prefecture-level post-WTO trade shocks are not correlated with pre-WTO wage and GDP growth. Thus, the identification assumption is likely to be satisfied.[36]

Second, I am also interested in identifying the regulation effect. The regulation changes (M) serve as a mediator in the relationship between trade shocks (T) and economic outcomes (Y). As discussed in Imai et al. (2011) and Dippel et al. (2017), the key identifying assumption for the effect of the mediator on outcome variables is sequential ignorability. In Figure 1.11 from Imai et al. (2011), Panel (a) shows a simple decomposition of the effect of T on Y: directly and indirectly through the mediator M. Panel (b) illustrates a more complicated case, where T affects Y indirectly both through the mediator M and the mediator N. However,

[36] I also control for pre-WTO wage and GDP growth in the regression to be conservative.

there is no causal relationship between M and N. In Panel (c), N has a direct impact on M. Both (a) and (b) satisfy the sequential ignorability, and the effect of the mediator M on the outcome Y can be estimated using the equation where both T and M are regressors. In (c), the identification assumption is violated.[37]

Figure 1.11: Mediation analysis



Note: The figure is from Imai et al. (2011).

Applying the analysis to my case, the identifying assumption for the regulation effect is that there are no other mediators such that the mediator affects both migration regulations and outcome variables. For example, trade shocks can affect other government policies, which in turn affect economic outcomes, but other government regulations should not directly affect migration regulations.

Another potential concern in my analysis is that in the model, both regulation and the economic outcomes change at the same time. Thus, it is not clear whether the causal relationship goes from M to Y or from Y to M. In Section 1.6.3.1, I use the timing of the regulation change and the migration flow changes, as well as at the leads and lags to show that it was the regulation change that drove the

[37]For the detailed discussion, refer to Imai et al. (2011) and Dippel et al. (2017).

migrant flows, not the other way around. Thus, the causal relationship should go from M to Y.

Under this assumption, I estimate the following equation:

$$\Delta Y_{it} = \pi_0 + \pi_1 \Delta \ln(\text{regulation score}_{it}) + \pi_2 \text{Trade shocks}_{it} + X_{it}\Phi + \zeta_{it}$$

where π_1 represents how regulation changes affect the outcome variables, and π_2 represents the direct effect of trade shocks on the outcome variables. Then, combining π_1 with α_1 in Section 1.6.1.2, the effect of trade shocks on outcome variables through the regulation channel will be $\pi_1 \cdot \alpha_1$. The regulation effect as a share of the total trade effect is $\frac{\pi_1 \cdot \alpha_1}{\gamma_1}$.

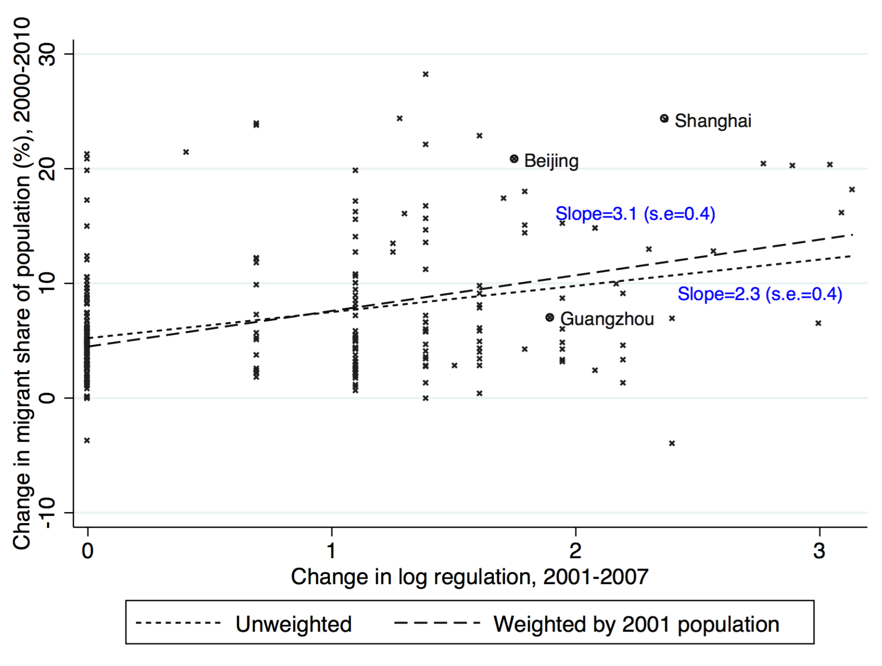
Another identification issue comes from the measurement of regulations. The first issue is that the regulations that I collect may not be the complete set of migration regulations that were enacted by the local government. The second issue is the coding of migrant-friendliness. The third issue is that enacting a regulation may not be equal to enforcing a regulation. I do not have a prior whether the prefectures with bigger changes in regulation scores enforced the regulations more strictly than prefectures with smaller changes. Overall, if the measurement error is random, the coefficient estimate for the effect of regulation changes on economic outcomes is downward biased.

To address the potential bias resulting from violation of sequential ignorability and the measurement errors, I also use an instrument for the regulation changes. I will explain the details of the IV approach in Section 1.6.2.3.

1.6.2.2 Trade Shocks, Regulation Changes, and Economic Outcomes: Mediation Analysis

Migrant Flows Did migration regulation relaxation lead to bigger migrant flows? I first investigate this relationship in Figure ??. The horizontal axis depicts the change in log regulation score from 2001 to 2007, the vertical axis depicts the change in migrant share of the population from 2000 to 2010, and each dot represents a prefecture. The graph shows that the more relaxed the regulation on migrants, the bigger the increase in migrant share. Megacities such as Beijing, Shanghai, and Guangzhou are not outliers. The graph suggests that the regulations are indeed useful, or the migration restrictions are indeed binding, in terms of affecting migration flows.

Figure 1.12: More regulation change from 2001 to 2007, greater change in migrant share of population from 2000 to 2010, 250 Chinese prefectures



Note: Each dot is a prefecture.

I present results of reduced-form trade effects and the OLS regression of migrant flows on regulation changes in Table 1.4. I first investigate how the migrant share of the population responds to trade shocks in Panel A. Columns (1) uses changes in migrant share of the population from 2000 to 2010 as the outcome variable, and the main regressor is the export tariff shock. I also control for the import tariff shock, the intermediate tariff shock, the migrant share of population in 2000, and the log of population in 2000. Column (1) shows that a 1 percentage point larger export tariff shock results in a 6.67 percentage point larger increase in the migrant share of the population. To alleviate the concern that trade shocks might be correlated with prefecture-level pre-WTO economic conditions, I add lagged trade shocks, wages and GDP growth in Column (2), and the coefficient becomes smaller and insignificant. Then I move on to Column (3) and (4) where I focus the effect of regulation changes. Column (3) shows that a 1% larger increase in regulation score from 2001 to 2007 results in a 0.018 percentage point larger increase in the migrant share of the population. Columns (4) gives similar results when I control for lagged trade shocks and lagged economic growth. I then add both trade shocks and regulation changes together in Column (5). Both the coefficient of the export tariff shock and the regulation change become smaller, but the significance level does not change. Evaluated at the coefficient estimate in Table 1.4 Column (2), big-shock prefectures had a 1.66 percentage point higher increase in the migrant share of the population than the small-shock ones. Using estimates from Table 1.4 Column (5) and Table 1.2 Column (1), big-shock ones had a 0.29 percentage point higher increase in the migrant share of the population through the regulation effect. The regulation effect is 17% of the overall trade effect. Given that the median size of prefecture population in 2001 (3.6 million), the big-shock prefectures had 76,000 greater increase in number of migrants than the small-shock prefectures, 13,000 of which was related to the change in regulation.

Table 1.4: Bigger regulation change (2001-2007), larger increase in number of migrants (2000-2010)

Panel A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	Δ migrant share of population					Δ log # of migrants, short-distance				
Export tariff shock	6.67*	5.37			4.58	1.31***	1.48***			1.36***
2001-2007	(3.37)	(3.24)			(2.95)	(0.31)	(0.33)			(0.27)
Δ log regulation score			1.76***	1.47***	1.38***			0.26***	0.29***	0.27***
2001-2007			(0.46)	(0.40)	(0.34)			(0.06)	(0.06)	(0.05)
Controls (lagged)		X		X	X		X		X	X
Observations	250	237	250	237	237	240	227	240	227	227
R-squared	0.14	0.26	0.15	0.28	0.29	0.19	0.27	0.16	0.28	0.33
Mean (s.d.) of depent.	6.99 (5.78)					-.83 (0.80)				
Panel B	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	Δ log # of migrants, medium-distance					Δ log # of migrants, long-distance				
Export tariff shock	0.42**	0.41**			0.38**	-0.11	-0.24			-0.30
2001-2007	(0.20)	(0.19)			(0.18)	(0.39)	(0.42)			(0.39)
Δ log regulation score			0.07*	0.06	0.04			0.14**	0.13*	0.13**
2001-2007			(0.04)	(0.03)	(0.03)			(0.06)	(0.06)	(0.06)
Controls (lagged)		X		X	X		X		X	X
Observations	250	237	250	237	237	249	236	249	236	236
R-squared	0.59	0.63	0.57	0.62	0.63	0.12	0.22	0.09	0.21	0.24
Mean (s.d.) of depent.	1.60 (0.66)					0.85 (0.65)				

Note: Standard errors are clustered at the province level. Dependent variables are changes from 2000 to 2010. Mean (sd) Δ log regulation score, 2001-2007 is 0.77 (0.82), mean (sd) export tariff shock is 0.18 (0.15). All columns control for import and intermediate tariff shock, log total population and level of dependent variable in 2000. Columns (2)(4)(6)(8) also control for lagged trade shocks and lagged wage and gdp growth rate, 1995-2001, as in Table 2 Column (8).

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Now I look at how migrants travel over various distances in response to trade shocks and regulation changes. Panel A Columns (6)-(10) use change in the log number of short-distance migrants as the outcome variable, and specifications are the same as in Columns (1)-(5). The short-distance migrants are the ones who migrate within a prefecture. Column (6) shows that a 1 percentage point higher export tariff shock results in a 131% larger increase in the number of short-distance migrants. Column (8) shows that a 1% higher increase in regulation score leads to a 0.26% larger increase in the number of short-distance migrants. Columns (7) and (9) generate similar results. I then add both trade shocks and regulation changes in Column (10), and coefficients for both terms become slightly smaller but remain statistically significant at the 1% level. The regulation effect is 12.4% of the overall trade effect. Given the median size of the short-distance migrant population in 2001 (167,000), the big-shock prefectures had a 69,000 larger increase in the number of migrants than the small-shock prefectures, 8,000 of which was due to the change in regulation.

Panel B Columns (1)-(5) use change in log number of medium-distance migrants as the outcome variable and follow the same specifications as in Panel A Columns (6)-(10). Medium-distance migrants are the ones who move between prefectures within a province. The effects of trade shocks and regulation changes are both smaller than the ones for short-distance migrants, and the effect of regulations become insignificant when both are added in Column (5). The regulation effect is 6.6% of the overall trade effect. Given the median size of the medium-distance migrant population in 2001 (64,500), the big-shock prefectures had a 8,000 larger increase in the number of migrants, 500 of which were due to the change in regulation.

Panel B Columns (6)-(10) repeat the exercise for the long-distance migrants who migrate between provinces. The coefficient for the trade shocks is not significant. However, effects of regulation changes are big (0.13) and significant. This indicates that there are other mediators that are positively affected by trade shocks and affect long-distance migration negatively, or are negatively affected by trade shocks and affect long-distance migration positively.

The results in Table 1.4 show interesting heterogeneity of overall trade effects and regulation effects across different migration distances. For short-distance migration, both the overall trade shock and the regulation changes had a big impact on migrant flows. The overall effect of trade shocks is significant for medium-distance migrants, while for long-distance migrants, regulations matter more.

Wages, Employment, and per capita GDP Next, I discuss how trade shocks affected other economic outcomes such as wages, employment, and GDP growth.

I use the same specification, where the coefficient γ_1 in the first equation is the overall effect of trade shocks on outcome variables, and the effect through the regulation channel will be $\pi_1 \cdot \alpha_1$.

I first look at the effect of trade shocks and regulation changes on wages in Table 1.5 Columns (1)-(5). Column (1) shows that a 1 percentage point larger increase in trade shocks leads to a 16 % larger increase in wages. I then add pre-WTO trade shocks and economic growth controls in Column (2), and results remain the same. I investigate the effect of regulation changes on wages in Columns (3) and (4). Column (3) indicates that a 1% higher increase in regulation score results in a 0.03% higher increase in wages. Column (4) shows similar effects. In Column (5), I add both trade shocks and regulation changes, and both the size of the coefficients and the statistical significance remains the same. Big-shock prefectures had a 5% higher increase in wages than the small-shock prefectures. Given that the mean increase in wages is 82% (and the standard deviation is 14%), the overall trade effect is 6% of the mean (and 34% of one standard deviation) for changes in wages. The regulation effect is 15% of the total trade effect.

My finding of the positive effect of regulation changes on wages is similar to the finding in Lee et al. (2017), where the authors study the effect of the U.S. repatriation of Mexicans in the 1930s on local employment, and they find that the decrease in the number of Mexican worker was associated with small decreases in native employment and increases in native unemployment. Although my results point to the wage margin rather than the employment margin, the finding suggests that an inflow of migrant workers could be beneficial for local workers overall.

Columns (6)-(10) show that both the overall effect of trade shocks and the regulation effects are bigger for per capita GDP than for wages. Big-shock prefectures had a 20% higher increase in per capita GDP than the small-shock prefectures. Given that the mean increase in per capita GDP is 87% (and the standard deviation is 27%), the overall trade effect is 20% of the mean (and 65% standard deviation) for changes in per capita GDP. The regulation effect is 9% of the total trade effect.

Table 1.5: More regulation change, 2001-2007, and bigger increase in wages, employment, and per capita GDP, 2001-2007

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Dependent variable	Δ log wage, 2001-2007					Δ log GDP p.c., 2001-2007					Δ log total urban emp., 2001-2007				
Export tariff shock	0.16**	0.14**			0.13**	0.56***	0.53***			0.51***	0.37*	0.28			0.27
2001-2007	(0.07)	(0.06)			(0.06)	(0.13)	(0.14)			(0.13)	(0.20)	(0.18)			(0.17)
Δlog regulation score			0.03***	0.03***	0.03***			0.08***	0.07***	0.07***			0.09**	0.05	0.04
2001-2007			(0.01)	(0.01)	(0.01)			(0.03)	(0.02)	(0.02)			(0.04)	(0.03)	(0.03)
Controls (lagged)		X		X	X		X		X	X		X		X	X
Observations	250	237	250	237	237	250	237	250	237	237	249	236	249	236	236
R-squared	0.30	0.36	0.26	0.37	0.39	0.27	0.32	0.22	0.26	0.35	0.22	0.32	0.07	0.31	0.33
Mean (s.d.) of depent.	0.82 (0.14)					0.87 (0.27)					0.32 (0.39)				

Note: Standard errors are clustered at the province level. Mean (sd) Δ log regulation score, 2001-2007 is 0.77 (0.82), mean (sd) export tariff shock is 0.18 (0.15). All columns control for import and intermediate tariff shock, log total population and level of dependent variable in 2000. Columns (2)(4)(5)(7)(9)(10)(12)(14)(15) also control for lagged trade shocks and lagged wage and gdp growth rate, 1995-2001, as in Table 2 Column (8) .

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Columns (11)-(15) show change in log total urban employment. The coefficient estimates for both overall trade shock and the regulation changes are bigger than the ones in Columns (1)-(5), but some of them are significant. Big-shock prefectures had a 12% higher increase in total urban employment than the small-shock prefectures. Given that the mean increase in total urban employment is 32% (and the standard deviation is 39 percent), the overall trade effect is 36% of the mean (and 30% of one standard deviation) for changes in total urban employment. The regulation effect is 10% of the overall trade effect.

Overall, the trade effect on wages and income is statistically significant and economically large. The effect on per capita GDP is bigger than the effects on wages and employment, potentially capturing other channels through which trade shocks affected the economy (through payment to other factors, for example). The regulation channel is significant for wages, per capita GDP, and total urban employment, and the regulation effect is about 9%-15% of the total trade effect.

1.6.2.3 Trade Shocks, Regulation Changes, and Economic Outcomes: IV Approach

The natural growth rate of the population (birth rate minus death rate) predicts the future population size of a prefecture. Higher natural growth rate means that the prefecture will have a more abundant work force. At the same time, a prefecture needs infrastructures to accommodate a larger population. Also, given China's one child policy, a high natural population growth rate may indicate that the prefecture is not effective in enforcing the birth control policy. These factors are likely to make the prefecture government less willing to relax the migration policies.[38]

In Table 1.6 Column (1), I regress change in log regulation scores on trade shocks as in Table 1.2 Column (8) and control for the 2000 natural growth rate of population. The coefficient for the natural growth rate is negative and statistically significant, meaning that in prefectures with higher natural growth rates, the increase in migrant regulation score is smaller. I then repeat the OLS regression in the previous two tables regarding migrant flows, wages, per capita GDP, and employment, and I also use the 2000 natural growth rate and the 2000 regulation score as instruments for the change in regulation score from 2001 to 2007. I find that compared with the OLS estimates, the effect of changes in regulation scores on economic outcomes are bigger in the IV regressions. However, the IV standard errors are much bigger, and the difference between the OLS estimates and the IV estimates are not statistically significant according to the Hausman test.

Overall, I find that the OLS results from the mediation analysis is robust, and if anything, the OLS might underestimate the effect of regulations on economic outcomes.

[38] As shown in the data section, birth control is an important aspect of migration policies.

Table 1.6: Natural growth rate as IV for regulation change, first-stage and IV results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	Δ log reg. score	Δ migrant share of pop.		Δ log GDP p.c.		Δ log wage		Δ log total urban emp.	
2001-2007		OLS	IV	OLS	IV	OLS	IV	OLS	IV
Export tariff shock	0.70*	4.58*	3.43	0.51***	0.49***	0.13**	0.08	0.27	0.23
2001-2007	(0.38)	(2.58)	(2.97)	(0.12)	(0.13)	(0.06)	(0.07)	(0.17)	(0.19)
Δ log regulation score		1.38***	3.38	0.07***	0.13	0.03***	0.13**	0.04	0.22
2001-2007		(0.45)	(2.23)	(0.02)	(0.13)	(0.01)	(0.05)	(0.03)	(0.22)
Natural growth rate of	-0.06***								
pop., 2000	(0.02)								
Observations	237	237	237	237	237	237	237	236	236
R-squared	0.19	0.29	0.23	0.35	0.32	0.39	0.13	0.33	0.23
First-stage F stat.	-	5.51		6.95		5.64		6.26	
Hausman test p-value	-	1.00		1.00		0.97		1.00	
Mean (s.d.) of depent.	0.77 (0.82)	6.99 (5.78)		0.82 (0.14)		0.87 (0.27)		0.32 (0.39)	

Note: Mean (sd) export tariff shock is 0.18 (0.15). Column (1) has the same specification as in Table 2 Column (8), and adds the natural growth rate of population in 2000. Column (2) has the same specification as in Table 4, Panel A, Column (5). Column (3) instruments change in log regulation score with natural growth rate of population, and log regulation score in 2001. Column (4) is the same as in Table 5 Column (5), and Column (5) is the corresponding IV regression. Column (6) is the same as in Table 5 Column (10), and Column (7) is the corresponding IV regression. Column (8) is the same as in Table 5 Column (15), and Column (9) is the corresponding IV regression.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

1.6.3 Discussion on the Regulation Effects

1.6.3.1 Did Migrant Flow Drive Regulation Change, or Was It the Other Way Around?

On the one hand, migration regulation change can affect migrant amenities, making the city more or less attractive to migrants and leading to bigger or smaller migrant inflow. On the other hand, larger migrant inflows could put pressure on city infrastructure and local employment, and lead to regulation changes. Since trade shocks affect both migrant flows and migrant regulation, it could be useful to distinguish which happens first. To do this, I look at the timing of the regulation change and the migration flow changes, as well as at the leads and lags.

In previous sections, I use the migration flow from 2000 to 2010, since I cannot observe the number of migrants in 2007. In Table 1.7, I check the effect of regulation changes in different time periods on migrant flows from 2005 to 2010. Column (1) shows that a 1% increase in regulation score from 1995 to 2000 (two lagged periods) is related to a 0.44% increase in the number of migrants from

2005 to 2010. In Column (2), I use regulation changes from 2000 to 2005 (one lagged period), and the coefficient on the change in log regulation score declines to 0.32. Column (3) uses the contemporaneous regulation change from 2005 to 2010, and the coefficient declines to 0.09. This could be the mechanical effect from the fact that the mean change in regulation increases from 0.04 in Column (1) to 0.95 in Column (3). However, when we go to Column (4), although there is still a sizable change in log regulation score of 0.53 from 2010 to 2015, there is no longer a positive effect of regulation changes on migrant flow from 2005 to 2010.

In Columns (5)-(8), I use change in migrant share of population from 2005 to 2010 as the outcome variable, and the finding is similar to Columns (1)-(4).

Table 1.7: Regulation change and migrant flows, lagged, current, and lead, 254 prefectures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Δ log # of migrants, t=2005-2010				Δ migrant share of pop. (%), t=2005-2010			
Δ log regulation score, 1995-2000	0.44***				7.37***			
	(0.14)				(1.76)			
Δ log regulation score, 2000-2005		0.32***				2.43***		
		(0.07)				(0.78)		
Δ log regulation score, 2005-2010			0.09**				0.93**	
			(0.04)				(0.39)	
Δ log regulation score, 2010-2015				-0.15				-1.41**
				(0.09)				(0.57)
Mean regulation change	0.04	0.30	0.95	0.53	0.04	0.30	0.95	0.53
Observations	254	254	254	254	254	254	254	254
R-squared	0.24	0.31	0.23	0.23	0.10	0.09	0.05	0.05

Note: Standard errors are clustered at the province level. Mean (sd) of Δ log # of migrants from 2005 to 2010 is 0.55 (0.56). Mean value of Δ migrant share of population from 2005 to 2010 is 5.5 (5.4). All columns controls for Y and log total population in 2005.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Overall, I find a positive effect of lagged or current regulation change on migrant flow, but no effect of lead regulation changes. This finding reinforces the argument that regulations are indeed binding, and changes in regulation determine migration, rather than being the result of migrant flows.

1.6.3.2 Did Regulations Improve Migrant Welfare?

As explained in Section 1.5.2, migration regulations had specific targets: increasing migrant wages, forcing firms to sign contracts, providing social insurance to migrants, and giving migrant children access to local primary and secondary schools. Thus, it would be helpful to see whether the regulations indeed help migrants and improve migrant welfare.

Unfortunately, the only available source that include these measures is the 2005 mini-census data. Thus, I cannot see how regulation changes affect change in migrant welfare, and I can only investigate in the cross-section whether in prefectures with more pro-migrant regulations, migrants report greater access to local amenities. To alleviate the concern that prefectures with more migrant-friendly regulations could be essentially different from other prefectures, I will control for local-worker welfare measures, log per capita GDP, and log number of migrant adults in 2005. Results are shown in Table 1.8 Panel A.

Table 1.8 Panel A indicates that the prefectures with higher regulation scores are also the ones with higher migrant welfare, concerning social insurance, income levels, and contract issues. Column (1) shows that a 1 unit increase in regulation score is related to a 0.3 percentage point increase in the unemployment insurance rate for migrants. Given that the mean of insurance rate for migrant is 21% and for locals is 37%, a 1 unit increase in regulation score will close 5.3% of the migrant-local gap. However, the coefficient is not statistically significant. Columns (2) and (3) show similar patterns, but the effects on pension and medical insurance rates are statistically significant at the 5% level and at the 10% level, respectively. Column (4) show a significant effect of regulation scores on the length of contracts: a 1 unit increase in regulation score is related to a 0.06 month increase in the length of contracts, which is 10% of the gap between locals and migrants. Column (5) indicates that a 1 unit increase in regulation score is related to an income increase of 13 yuan per month, which is 22% of the wage gap between locals and migrants. Column (6) shows that the regulation score has no significant impact of school enrollment rates among migrant children at the school age. Column (7) is about whether a regulation score increase is correlated with more migrant children brought to prefectures where their parents are working; the result is insignificant.

Table 1.8: Regulation score and migrant welfare in 2005, 247 prefectures

Panel A	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable Y for migrants, 2005	Unemployment Insurance (rate)	Pension (rate)	Medical insurance (rate)	Terms of Contract (in months)	Monthly Income (in yuan)	School Enrolment (rate)	Log # of children
Regulation score, 2005	0.003	0.005**	0.005*	0.06**	12.7***	0.001	-0.005
	(0.002)	(0.002)	(0.002)	(0.03)	(4.6)	(0.001)	(0.006)
Y for local, 2005	0.495***	0.391***	0.391***	0.48***	0.30***	0.487***	0.147***
	(0.051)	(0.048)	(0.040)	(0.03)	(0.09)	(0.117)	(0.035)
Mean (s.d.) Y for migrants	0.21(0.11)	0.35 (0.12)	0.36 (0.12)	4.50 (2.64)	924 (190)	0.94 (0.04)	4.47 (1.00)
Mean (s.d.) Y for local	0.38 (0.16)	0.60 (0.14)	0.60 (0.15)	5.05 (2.88)	982 (336)	0.95 (0.02)	6.81 (0.75)
Observations	247	247	247	247	247	247	247
R-squared	0.382	0.213	0.265	0.343	0.599	0.123	0.940
Panel B							
Regulation score, 2005	0.008***	0.000	0.001	0.14***	7.54	0.001**	-0.018**
(Y for local as dependent var.)	(0.002)	(0.002)	(0.002)	(0.04)	(8.21)	(0.000)	(0.007)

Note: Standard errors are clustered at the province level. Mean (sd) regulation score in 2005 is 0.93 (2.67). In Panel A, I regress migrant welfare measures on regulation score and local welfare measures, controlling for log number of adult migrants and log GDP p.c. in 2005. In Panel B, I regress local welfare measures on regulation score, controlling for log number of adult locals and log GDP p.c. in 2005.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

In Panel B, I regress welfare measures for local workers on the regulation score, controlling for local population size and GDP. Columns (1)(4) and (6) suggest that a higher regulation score is correlated with a higher unemployment insurance rate, longer terms of contracts, and a higher school enrollment rate for locals. These results might capture the fact that higher income prefectures usually provide more amenities for both locals and migrants. Column (7) shows that a larger number of local children is correlated with less generous migration regulations, suggesting potential congestion forces regarding education resources. It is reassuring that Columns (2)(3) and (5) do not show significant effects of regulations on local welfare measures, indicating that regulation effects are not only the proxy for local socioeconomic levels that could affect migrant welfare directly, but also the actual improvement through implementation and enforcement of the regulation contents.

Overall, the results in Table 1.8 show that prefectures with higher regulation scores also have higher migrant well-being, although the estimates are relatively small. The significant effects concentrate on pensions, medical insurances, terms of contract, and wages, and all these aspects are the focus of many migration regulations. These results suggest that more pro-migrant regulations were associated with improvements on well-being of migrants in the workplace. The outcomes related to migrant children were not significantly affected by the regulations, and there are several potential explanations. First, school capacities were limited, and it was very costly for prefecture governments to expand the capacity in the short-run. The second explanation is that prefecture governments only wanted the migrant workforce and were reluctant to make substantial changes to incentivize migrant workers to settle down with their family. Alternatively, migrant workers might view their migration as temporary and did not want to bring their family, especially considering the fact that migrant children are still not allowed to take the college entrance examination outside their Hukou location.

1.6.4 Migrant Network, Transportation Network and Migrant Flow Responses

Another issue is that the effect of change of regulations and trade shocks might differ depending on how connected the prefecture is through the transportation network. Migrants can travel more easily to prefectures where transportation cost is low; they are able to travel back to their hometown when needed, and this can further incentivize temporary migration. Yang (2017) shows the Chinese highway system expanded substantially from 1995 to 2015 as a result of a national-level infrastructure construction plan.[39] I construct the change of overall connectedness of a prefecture using the interaction of the prefecture-level initial migrant network

[39] For details of the highway expansion, please see Yang (2017).

in 2000 and the reduction in transportation cost between prefectures from 2000 to 2005, and study how regulation changes interact with the improved connectedness in determining migrant flows. The change of connectedness of prefecture i is given by:

$$\Delta \text{Connection}^{I}_{i,2000-2005} = \sum_{j} (T_{ij,2000} - T_{ij,2005})$$

where $T_{ij,2000}$ is the number of hours to travel from prefecture j to prefecture i through the least cost path in 2000, and $T_{ij,2005}$ is for 2005. $T_{ij,2000}$ and $T_{ij,2005}$ are from Yang (2017), using the highway and non-highway network in China, with the assumption that the speed of travel is 90 kilometers per hour on highways, 25 kilometers per hour on national and provincial level non-highways, and 15 kilometers on local roads.[40]

Alternatively, I can take the migrant network into account and construct the change in connectedness as:

$$\Delta \text{Connection}^{II}_{i,2000-2005} = \sum_{j} \frac{m_{ij}}{\sum_{j'} m_{ij'}} (T_{ij,2000} - T_{ij,2005})$$

where m_{ij} is the number of migrants who are from prefecture j and reside in prefecture i, and I calculate the bilateral migrant flows using the 2000 census data.

I estimate the effect of change in prefecture connectedness and regulation changes on migrant flows using the following equation:

$$\Delta \ln M_i = \pi_0 + \pi_1 ExportShock_i + \pi_2 \Delta \ln(\text{reg. score}_i) + \pi_3 \Delta C_i + \pi_4 \Delta \ln(\text{reg. score}_i) \cdot \Delta C_i + X_i \Phi + \zeta_i$$

[40]National and provincial level nonhighways are of better conditions.

$$\Delta \ln M_i = \pi_0 + \pi_1 ExportShock_i + \pi_2 \Delta \ln(\text{reg. score}_i) + \pi_3 \Delta C_i + \pi_4 ExportShock_i \cdot \Delta C_i + X_i \Phi + \zeta_i$$

where $\Delta \ln M_i$ is the change in log number of migrants in prefecture i from 2000 to 2010, $ExportShock_i$ is the shock due to export tariff changes from 2001 to 2007, $\Delta \ln(\text{reg score}_i)$ is the change in log regulation score from 2001 to 2007, and ΔC_i is the change in connectedness from 2000 to 2005. I also control for log number of migrants in 2000, import and intermediate shocks from 2001 to 2007, trade shocks from 1995 to 2001, and wage and GDP growth from 1995 to 2001 as in Table 1.4 Columns (5) and (10).

Results using only the change in transportation network ($\Delta\text{Connection}^I$) are shown in Table 1.9. In Column (1) I control for trade shocks and regulation changes and add the change in connectedness, and there is no significant effect of the change in connectedness on short-run migration flows. Column (2) introduces the interaction of regulation changes and connectedness changes, where the connectedness measure incorporates the network measure. The regulation effect becomes small (-0.12) and insignificant; the interaction effect is significant and positive. This indicates that the effect of regulation changes was amplified in prefectures which became more connected. For short-distance migrants, they migrate within a prefecture, and the positive interaction effect can be from the positive correlation between prefecture connectedness and economic growth expectation. The change in connectedness has negative effects on migrant flows (-0.04), and a possible explanation of the result is that when a prefecture is more connected with other prefectures, local rural workers can migrate out more easily. Combining the level effect with the interaction effect, the overall effect of connectedness is positive only when there are relatively big regulation changes (with change in log regulation score bigger than 1). Column (3) adds the interaction between the export shock and connectedness changes, and there is again a positive significant interaction effect.

Columns (4)-(6) show effects on medium-distance migrants. Coefficients for the interaction with connectedness is a significant 0.02. This indicates that more connected prefectures attract more migrants from other prefectures when they relax migration regulations. Again, there is a negative effect of the change in connectedness on migrant flows (-0.01), potentially due to between-prefecture competition. The overall effect of connectedness is positive only when there are relatively big regulation changes. The interaction with the export shock is insignificant, and much smaller than short-run coefficients.

Columns (7)-(9) show effects on long-distance migrants. The interaction effect with regulation changes is insignificant, but the size is comparable with medium-distance results. The interaction with the export shock is again not significant.

Overall, this finding shows that when a prefecture becomes more connected, trade shocks and changes migration regulations have bigger effects on the attracting migrant inflows, especially for short- and medium-distance migrants.

Table 1.9: Interaction effects of regulation changes (2001-2007) and prefecture connection (2000-2005)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	Δ log # of migrants, short-distance			Δ log # of migrants, medium-distance			Δ log # of migrants, long-distance		
Export tariff shock	1.51***	1.45***	0.70	0.44**	0.42**	0.37	-0.17	-0.20	0.03
2001-2007	(0.27)	(0.29)	(0.59)	(0.19)	(0.19)	(0.40)	(0.43)	(0.41)	(0.87)
Δ log regulation score	0.27***	-0.12	0.26***	0.04	-0.12*	0.04	0.13**	-0.07	0.13**
2001-2007	(0.05)	(0.10)	(0.05)	(0.03)	(0.07)	(0.03)	(0.06)	(0.14)	(0.06)
Δ Connection	0.00	-0.04**	-0.02	0.00	-0.01	0.00	0.02	-0.00	0.02
2000-2005	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Δ log regulation score		0.04***			0.02**			0.02	
× Δ Connection I		(0.01)			(0.01)			(0.01)	
Export tariff shock			0.08*			0.01			-0.02
× Δ Connection I			(0.05)			(0.03)			(0.05)
Observations	225	225	225	235	235	235	234	234	234
R-squared	0.35	0.38	0.35	0.64	0.64	0.64	0.25	0.25	0.25
Mean (s.d.) of depent.	-0.81 (0.80)			1.60 (0.67)			0.84 (0.66)		

Note: Standard errors are clustered at the province level. Dependent variables are changes from 2000 to 2010. Mean (sd) Δ log regulation score, 2001-2007 is 0.77 (0.82), mean (sd) export tariff shock is 0.18 (0.15), and mean (sd) Δ Connection, 2000-2005 is 7.23 (4.01). All columns control for import and intermediate tariff shock, log total population, level of dependent variable in 2000, and lagged trade shocks and lagged wage and gdp growth rate, 1995-2001, as in Table 2 Column (8).
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Then I repeat the exercise in Table 1.9 in Table 1.10, replacing the connectedness measure with the one that with the migrant network (ΔConnection^{II}). The overall finding is similar to Table 1.9, but less significant.

Table 1.10: Interaction effects of regulation changes (2001-2007) and prefecture connection (2000-2005)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	Δ log # of migrants, short-distance			Δ log # of migrants, medium-distance			Δ log # of migrants, long-distance		
Export tariff shock	1.59***	1.50***	0.82	0.45**	0.43**	0.42	-0.15	-0.19	-0.05
2001-2007	(0.29)	(0.30)	(0.50)	(0.20)	(0.19)	(0.35)	(0.43)	(0.42)	(0.80)
Δ log regulation score	0.27***	-0.08	0.25***	0.04	-0.07	0.04	0.13**	-0.03	0.13**
2001-2007	(0.06)	(0.13)	(0.06)	(0.03)	(0.08)	(0.03)	(0.06)	(0.12)	(0.06)
Δ Connection	-0.02	-0.07**	-0.05**	-0.00	-0.02	-0.00	-0.00	-0.02	0.00
2000-2005	(0.02)	(0.03)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Δ log regulation score		0.05**			0.02			0.02	
× Δ Connection II		(0.02)			(0.01)			(0.01)	
Export tariff shock			0.11**			0.00			-0.01
× Δ Connection II			(0.05)			(0.03)			(0.06)
Observations	225	225	225	235	235	235	234	234	234
R-squared	0.35	0.38	0.34	0.64	0.64	0.64	0.24	0.25	0.24
Mean (s.d.) of depent.	-0.81 (0.80)			1.60 (0.67)			0.84 (0.66)		

Note: Standard errors are clustered at the province level. Dependent variables are changes from 2000 to 2010. Mean (sd) Δ log regulation score, 2001-2007 is 0.77 (0.82), mean (sd) export tariff shock is 0.18 (0.15), and mean (sd) Δ Connection, 2000-2005 is 7.23 (4.01). All columns control for import and intermediate tariff shock, log total population, level of dependent variable in 2000, and lagged trade shocks and lagged wage and gdp growth rate, 1995-2001, as in Table 2 Column (8).

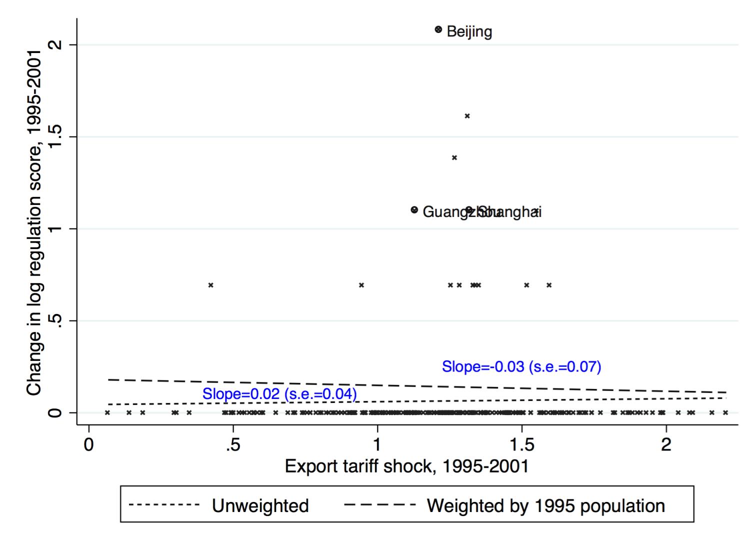
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

1.6.5 Robustness Checks

1.6.5.1 Trade Shock-Regulation Relationship in 1995-2001

In Figure 1.13, I plot the relationship between change in log regulation score and export tariff shock in the 1995-2001 period to compare with Figure 1.9. It is clear that before the WTO accession, there were few changes in migrant-related regulations (with insignificant coefficients of -0.03 to 0.02, while the coefficients are 0.7 to 1.4 and statistically significant in the post-WTO period), and the few cities that changed migrant regulation were big capital cities. This reinforces the argument about the significance of the WTO effect.

Figure 1.13: Effect of trade shocks on regulation change, 1995-2001, 250 prefectures



Note: Each dot is a prefecture.

1.6.5.2 Change the Sample of Prefectures

In the main analysis, I focus on 250 prefectures with complete data on economic conditions such as GDP and wages from the Prefecture Statistic Yearbook. In addition, I can include all 340 prefectures in China to check the robustness of the result with respect to the sample selection. In Columns (1), Table 1.11, I include 333 prefectures and regress the change in log regulation score on export tariff shocks.[41] The point estimate for export tariff shocks remains similar in Columns (2)-(3) when I add import tariff shocks, intermediate tariff shocks and log regulation score in 2001. As I mentioned in the main analysis, prefectures with high employment concentration in the petroleum industries are outliers in the analysis. They experienced big and positive export tariff shocks, but the petroleum industry is mostly state-owned. Thus, the response in regulation changes was

[41] 7 Tibetan prefectures are not included because there is no input-output table for Tibet, and I cannot construct the intermediate tariff shock. The result in Column (1) holds if I include the 7 prefectures, but I drop them in Column (1) to be comparable with Columns (2)-(4).

small in those industries despite the big trade shocks. In Column (4), I include those prefectures in the analysis and control the employment share of petroleum industry. In Column (5), I drop prefectures with share of employment in the petroleum industry higher than 20% as in the main analysis. The coefficients for export tariff shocks are comparable in these two columns, but bigger than in Column (1)-(3). This is consistent with the outlier story.

In Figure 1.9, 114 prefectures experienced no regulation changes from 2001 to 2007. Thus, it is useful to distinguish whether the result of trade shocks on regulations is driven by the comparison between prefectures with no changes and prefectures with changes, or between the prefectures with big positive changes and small positive changes. In Table 1.11 Columns (6)-(9), I only include prefectures with nonzero changes. The coefficient estimates are 15-40% smaller than in Table 1.2 Columns (1)-(3) and remain statistically significant at the 5% level. This result suggests the both the extensive margin and the intensive margin of regulation changes are important in estimating the trade effects.

Table 1.11: Effects of trade shocks on regulation change, 2001-2007, different sample size

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ log regulation score 2001-2007	All prefectures					Prefectures with nonzero changes		
Export tariff shock	0.72***	0.81***	0.75***	1.02***	1.03***	0.57**	0.67**	0.64**
2001-2007	(0.19)	(0.21)	(0.22)	(0.35)	(0.36)	(0.24)	(0.28)	(0.28)
Import tariff shock		-0.03	-0.03	-0.04	-0.04		-0.03	-0.03
2001-2007		(0.02)	(0.03)	(0.03)	(0.03)		(0.03)	(0.03)
Intermediate tariff shock		0.40**	0.40**	0.37**	0.39**		-0.04	-0.03
2001-2007		(0.15)	(0.15)	(0.15)	(0.15)		(0.10)	(0.11)
Log regulation score			0.81***	0.79**	0.78**			0.17
2001			(0.30)	(0.30)	(0.31)			(0.11)
Employment share in				-1.05*				
petroleum ind, 2000				(0.58)				
Observations	333	333	333	333	323	148	148	148
R-squared	0.03	0.10	0.15	0.16	0.16	0.02	0.03	0.04
Mean (s.d.) of depent.	0.64 (0.78)				0.64 (0.79)	1.37 (0.54)		
Mean (s.d.) of expor shock	0.16 (0.20)				0.14 (0.18)	0.20 (0.19)		

Note: Standard errors are clustered at the province level. Columns (1)-(4) include all prefectures in China expect for prefectures in Tibet, since there is no input-output table for Tibet and intermediate tariff shocks are missing. Column (4) controls for the employment share in petroleum industry in 2000, and Column (5) drops the prefectures with employment share in petroleum industry higher than 20%. Columns (6)-(8) include all prefectures with nonzero changes from 2001 to 2007, excluding prefectures with employment share in petroleum industry higher than 20%.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

1.6.5.3 Alternative Measure of Regulation Changes

One important aspect of the data is the coding of regulations' migrant-friendliness. In the main specification, I create the regulation score on a -2 to 2 scale, with -2 as the least migrant-friendly and 2 as the most migrant-friendly. I follow a specific rule and cross-check the coding result with two independent research assistants with law degrees, but there might still be concerns on the objectivity of the coding.[42] Thus, instead, I use a "negative (-1), neutral (0) and positive (+1)" scale and also a simple count of number of regulations to check the robustness of the result.

Also, I decompose the regulations by topics into work-related, welfare-related, and administrative to investigate the effect of trade shocks on each categories.

In Table 1.12, I use the same specification as in Table 1.2 Column (3). Column (1) replicates Table 1.2 Column (3), with the outcome variable using the five-level coding. Column (2) uses the three-level coding and Column (3) uses log of number of regulations. Columns (4)-(6) use the five-level coding by topic.

The results show that the effect of trade shocks on regulation changes is robust to variation in the regulation measure. The five-level coding is the most informative about the migrant-friendliness, and the effect of export tariff shocks is also the biggest and most significant among the first three columns. In the latter three columns, trade shocks that affected work-related regulations were most significant and administrative ones the least. Overall, all columns are consistent with the main result.

[42] See details in Appendix 1.7.

Table 1.12: Alternative measure of regulation change

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Δ log regulation score		Δ log num.	Δ log regulation score		
2001-2007	5 levels	3 levels	regulations	Work	Welfare	Admn.
Export tariff shock	1.09***	0.83**	0.46*	1.03***	0.61**	0.32*
2001-2007	(0.39)	(0.32)	(0.27)	(0.36)	(0.26)	(0.16)
Import tariff shock	-0.08**	-0.05*	-0.02	-0.07**	-0.07**	-0.02
2001-2007	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.01)
Intermediate tariff shock	0.20*	0.23**	0.20**	0.06	-0.01	0.14**
2001-2007	(0.11)	(0.09)	(0.08)	(0.11)	(0.08)	(0.06)
Y, 2001	0.71***	0.75***	0.33***	3.16***	1.25***	0.14
	(0.14)	(0.14)	(0.06)	(0.10)	(0.05)	(0.16)
Observations	250	250	250	250	250	250
R-squared	0.12	0.16	0.16	0.07	0.06	0.08
Mean (s.d.) of depent.	0.77 (0.82)	0.57 (0.65)	0.57 (0.58)	0.61 (0.77)	0.29 (0.58)	0.13 (0.31)

Note: Standard errors are clustered at the province level. Mean value (sd) of export tariff shock, 2001-2007 is 0.18 (0.15).
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

1.6.5.4 Alternative Measure of Bartik-Style Trade Shocks.

To check the robustness of the main results with respect to the measure of trade shock, I use industry labor share as weights directly: $\beta'_{ij} = \lambda_{ij}$ in Table 1.13 Column (2). Compared to Column (1), which replicates Table 1.2 Column (3), the coefficient on the export tariff shock is very similar.

Alternatively, I follow the Autor et al. (2013) measure of local labor market trade shock and construct local market access shocks. The market access shock is also a Bartik-style measure, with industry-level export growth distributed across regions, weighted by local industry labor shares. The difference with the Autor et al. (2013) measure is that I use export growth instead of import growth, since the export one is more relevant in the Chinese context. Also, since Autor et al. (2013) analyze the effect of exposure to Chinese exports on the U.S. economy, the authors use Chinese exports to other developed countries as an instrument to capture the Chinese productivity growth effect. In my case, I want to capture the demand-side forces that led to the expansion of Chinese exports, so I use the

GDP growth of the importing countries as an instrument. An alternative measure would be the change in country dummies from a bilateral trade gravity regression.

The details of the construction of the measures can be found in Appendix 1.7. Table 1.13 Columns (3)-(7) show the results when I use the market-access-based shocks. Column (3) contains only the export shocks, Column (4) adds the import and intermediate shocks, and Column (5) adds urban share of the prefecture as a control. Column (6) instruments the export shock with the GDP-based instrument. Column (7) uses the gravity-dummy-based instrument. The size of the coefficient on the export shock is robust across these specifications, but the IV coefficients are less significant. The results show that a $1,000 per worker increase in exports led to a 2% increase in regulation score changes. Again, I divide prefectures into big-, medium- and small-shock ones, and the difference in export shocks between the big- and small-shock ones is $14,000 per worker. This translates into a 26% higher increase in regulation scores and it is comparable to the 21% difference found in the main regression with tariff shocks.[43]

[43] The per capita export was about $300 in 2001 and $1,000 in 2007. The number of employed workers in the Industrial Enterprises Survey in 2000 is 50 million. Thus, the $14,000 per worker difference is equivalent to $580 per person and is comparable to the $700 mean increase from 2001 to 2007.

Table 1.13: Alternative measure of trade shocks, Bartik-Style

Depend. Var. (2001-2007) Δlog regulation score	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Tariff based trade		Market access based trade shock				
Trade shock measures:	Main	Labor share	OLS	OLS	OLS	GDP IV	Gravity IV
Export shock	1.09***	1.02***	0.03***	0.02***	0.02**	0.02*	0.02
2001-2007	(0.39)	(0.37)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Import shock	-0.08**	-0.07*		0.01	0.00	-0.00	0.00
2001-2007	(0.04)	(0.04)		(0.01)	(0.01)	(0.01)	(0.01)
Intermediate shock	0.20*	0.20*		0.08	0.08	0.06	0.07
2001-2007	(0.11)	(0.11)		(0.09)	(0.10)	(0.10)	(0.10)
Log regulation score, 2001	0.71***	0.71***			0.46**	0.44**	0.45***
	(0.14)	(0.14)			(0.18)	(0.17)	(0.17)
Urban share, 2001					0.01**	0.01***	0.01***
					(0.00)	(0.00)	(0.00)
Observations	250	250	250	250	250	250	250
R-squared	0.12	0.11	0.06	0.07	0.16	0.16	0.16
Mean (s.d.) of export shock	0.18 (0.15)	0.17 (0.15)			16.40 (6.63)		
First-stage F-stat	-	-	-	-	-	493	488

Note: Standard errors are clustered at the province level. Mean value (sd) of Δ regulation score, 2001-2007 is 0.77 (0.82).

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

1.6.5.5 Additional Trade-Shock Measures: Uncertainty and Quota

In addition to the decline in tariffs, WTO also led to two other kinds of reduction in trade barriers. First, Handley and Limão (2017) shows that the United States applied MFN tariffs on Chinese exports even before the WTO accession. However, before 2001, there was large uncertainty regarding the U.S. trade policy: the MFN status had to be voted every year to be approved by the Congress and the House, otherwise, the Column 2 tariff would be applied to Chinese exports. Handley and Limão (2017) argues that the reduction in policy uncertainty was the main impact of the WTO accession on the U.S.- China trade relationship. Second, Khandelwal et al. (2013) shows that the Chinese textile and clothing exports to the United States, the European Union, and Canada were subject to Multifiber Arrangement (MFA) quota restrictions until January 2005. The removal of the quota restrictions increased Chinese exports a lot.

I investigate these two factors in Table 1.14. I use the 2000 customs data by firm, eight-digit Harmonized System (HS) category, and destination country, and combine it with the information on the 2000 Column 2 tariffs and MFN tariffs by eight-digit HS category by the United States from Handley and Limão (2017).[44] With this data, I construct the following variable

$$Column2_{i,2000} = \sum_p \frac{export^{US}_{p,i,2000}}{\sum_{p'} export^{US}_{p,i,2000}} (Column2^{US}_{p,2000} - MFN^{US}_{p,2000})$$

where i is a prefecture, p is a six-digit HS category, $export^{US}_{p,i,2000}$ the exports from Chinese prefecture i to the United States in category p in 2000, $Column2^{US}_{p,2000}$ is the U.S. Column 2 tariff on category p in 2000, and $MFN^{US}_{p,2000}$ is the U.S. MFN tariff. I also construct the U.S. export share as

$$\text{US export share}_{i,2000} = \frac{export^{US}_{i,2000}}{export^{W}_{i,2000}}$$

where $export^{US}_i$ is the total exports from Chinese prefecture i to the United States in 2000, and $export^{W}_{i,2000}$ is the total exports from China to the rest of the world in 2000. To take into account the fact that different prefectures' output share of output is different, and it might affect the exposure to trade shocks, I also construct the export share as

$$\text{Export share}_{i,2000} = \frac{export_{i,2000}}{output_{i,2000}}$$

where $export_{i,2000}$ is the value of export from prefecture i in 2000, and $output_{i,2000}$ is prefecture i's total sales revenue in 2000 – both are taken from the 2000 Industrial Enterprises Survey.

[44] I convert the eight-digit HS codes to six-digit ones in both datasets to increase the matching probability.

I also combine the customs data with the MFA quota restrictions to measure the quota removal effect. I construct a prefecture's exposure to MFA restrictions as

$$\text{Value of textile w/quota}_{i,2000} = \sum_p \frac{export_{p,i,2000}}{\sum_{p'} export_{p',i,2000}} \cdot D(MFA_p^{2001-2005} = 1)$$

where p is a 8-digit HS category, $export_{p,i,2000}$ is the export of product p from Chinese prefecture i to the world, and $D(MFA_p^{2001-2005} = 1)$ is an indicator variable that takes the value of 1 if the export is to the United States, Canada, and the European Union, and product p is subject to the MFA quota in any period between 2001 and 2005.

I then add the uncertainty controls and the MFA controls into the baseline regression as in Table 1.2 Column 3. Table 1.14 Column (1) replicates Table 1.2 Column 3, and in Column (2) I add the interaction between $Column2$ and US export share as the measure of uncertainty. Column (3) further interacts the uncertainty measure with the export share of output. Both columns show a small positive effect, indicating that the reduction in trade uncertainty indeed contributes to the change in regulations. However, the magnitude is relatively small, given the mean of 0.04 and 0.006.

Column (4) and (5) show the MFA effect. In Column (4), I control for Value of textile w/quota , and I interact it with export share of output in Column (5). Both columns show a positive and significant effect, indicating that the removal of textile and clothing quota increased the migrant-friendliness of a prefecture. The effects are also relatively small, evaluated at the mean. In addition, the distribution of Value of textile w/quota is very skewed to the right: the median value is 0.004 while the mean is 0.02. Thus, a few prefectures with big share of exports in the textile and clothing industries are the major source of variation.

In all columns, the coefficients on the export tariff shock remain largely unchanged. Overall, I find a robust effect of the tariff shocks on regulation changes, and given the positive estimates on the uncertainty and MFA effects, the tariff effect can act as a lower bound for the overall WTO effect.

Table 1.14: Alternative measures of trade shocks, uncertainty and textile quotas

Dependent variable: Δ log regulation score, 2001-2007	(1)	(2)	(3)	(4)	(5)
Export tariff shock, 2001-2007	1.09***	1.12***	1.07**	1.08***	1.02**
	(0.39)	(0.39)	(0.39)	(0.39)	(0.38)
Column2×US export share		2.06*			
		(1.13)			
Column2×US export share×Export share			4.35		
			(3.62)		
Value of textile w/ quota				3.95**	
				(1.50)	
Value of textile w/ quota×Export share					38.74**
					(15.52)
Observations	250	250	250	250	250
R-squared	0.12	0.13	0.12	0.13	0.13
Mean (s.d.) of control		0.04 (0.04)	0.006 (0.01)	0.01 (0.03)	0.002 (0.003)

Note: Standard errors are clustered at the province level. All columns control for import tariff shocks, intermediate tariff shocks, and log regulation score in 2001. "Column2" of a prefecture is the weighted average of differences between each product's U.S. Column 2 ad valorem tariff and its U.S. MFN tariff, with the the product's share of U.S. imports from the prefecture. Each product is on the HS6 level. "Export share" is a prefecture's output share of exports. "Value of textile w/ quota" is a prefecture's share of exports that is subject to textile quota restrictions during the 2001 to 2005 period.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

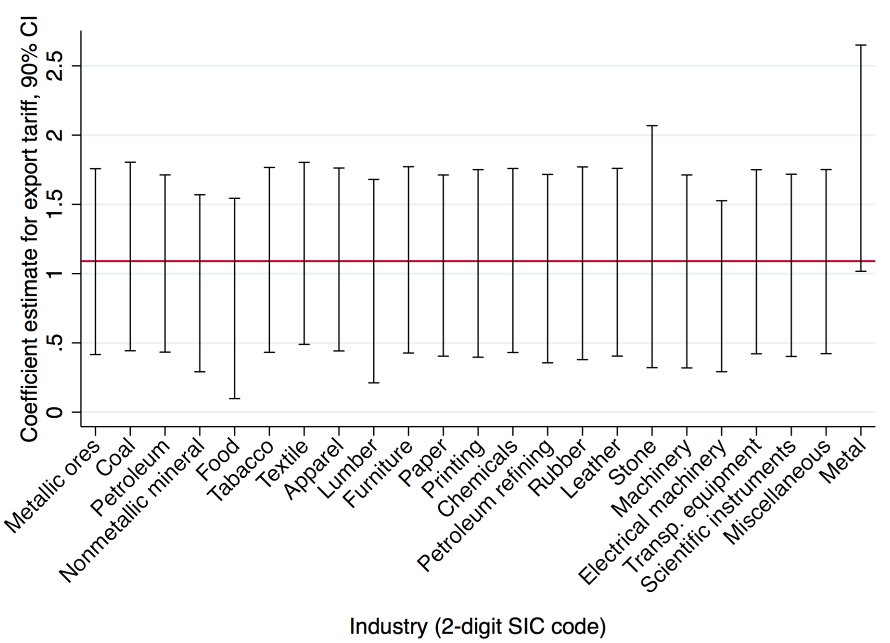
1.6.5.6 Adding Industrial Composition Controls

The regional tariff shocks are generated using the interaction of prefecture-level industrial composition and industry-level tariff reductions. If certain industries drive variation and are correlated with other local factors that affect regulation changes directly, then the estimates for regional tariff shock effects would be biased. To check whether such an industry exists, I add industry employment share one at a time and run the regression in Table 1.2 Column (3).

In Figure 1.14, I plot the coefficient estimates with 90% confidence intervals, and each bar is from a regression, including a specific-industry employment share.

The coefficient estimates are relatively stable around 1.09, which is the estimate in Table 1.2 Column (3). Thus, the results are not sensitive to specific-industry effects.

Figure 1.14: Coefficients from main regression by adding industrial composition controls one by one



Note: Each bar is the 90% confidence interval of the coefficient estimate of export tariff shocks from a regression as in Table 1.2 Column (3), controlling for a specific industry share of total employment. The horizontal bar is the point estimate of 1.09 from Table 1.2 Column (3).

1.6.5.7 Decomposition of Migrant Flow

Table 1.4 classifies migrant flows into short-, medium-, and long-distance categories. As a robustness check, I use alternative classifications: (1) by the purpose of migration (Table 1.15 Columns (1)-(4)); (2) by the time since migrating (Table 1.15 Columns (5)-(6)); and (3) by years of education (Table 1.15 Columns (7)-(8)). The specifications here are the same as in Table 1.4 Panel A Column (5).

I find that the relaxation of migration restrictions affected people who migrated for work the most and people who migrated for marriage the least. This is a reasonable result, since the regulations were mostly work-related. I also find that regulation changes had bigger effects in the later period (migrated in the nearest three years) than the current period (migrated in more than three years ago). This finding is consistent with Table 1.7 that regulations take time to impact migrant flows. Finally, I find that the regulation changes affected the migrants with more than 12 years of education the most. In the 2000-2010 period, the medium- and long-distance migrant flows increased a lot, and it seems that more-educated migrants were the driving force.

Table 1.15: Regulation change (2001-2007) and migrant flow in subcategories (2000-2010)

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δlog # of migrants	Purpose of migration				Time since migrated		Years of education	
Subcategory	Work	Family	Marriage	Other	<= 3 years	> 3 years	<=12 years	>12 years
Δlog regulation score	0.13***	0.11**	0.06	0.13**	0.13***	0.08*	0.07*	0.16***
2001-2007	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)
Y, 2001	-0.32***	-0.24***	-0.42***	-0.10	-0.29***	0.11**	-0.06	-0.23***
	(0.04)	(0.04)	(0.05)	(0.06)	(0.03)	(0.05)	(0.04)	(0.05)
Log population, 2001	0.18**	0.09	0.27***	0.06	0.18***	-0.07	0.04	0.12
	(0.08)	(0.06)	(0.07)	(0.04)	(0.06)	(0.05)	(0.05)	(0.08)
Observations	250	250	248	250	250	250	250	249
R-squared	0.35	0.13	0.30	0.04	0.29	0.07	0.02	0.15
Mean (s.d.) of depent.	1.47 (0.56)	1.15 (0.48)	0.92 (0.53)	-0.44 (0.48)	0.98 (0.45)	0.48 (0.44)	0.50 (0.35)	1.50 (0.53)

Note: Standard errors are clustered at the province level. Mean value (sd) of Δ regulation score, 2001-2007 is 0.77 (0.82).

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

1.6.5.8 Emigration Instead of Immigration

The 2000 and 2010 censuses also collected information on emigration, since each household were asked to report the number of family members who left their Hukou location for more than six months. In Table 1.16, I replicate the results in Table 1.4 Panel A by replacing the immigration share of population with emigration share of population and replacing the change in log number of short-distance migrants by the change in log number of out-migrants. Overall, there is no consistent significant effect of either trade shocks or regulation changes on emigration. Columns (1), (2) and (5), (6) show that bigger local export shocks decreased the outflow of people, but the results are not precisely measured. The effect of regulation changes on emigration is mixed and only significant in Column (8).

The results for emigration are consistent with the immigration results. Positive local shocks will make people less likely to migrate to other regions to work. Regulation changes centered mostly on improving the well-being of people who migrated to the region. This could still increase the incentive of within-prefecture migration, which might be captured by the positive effect in Column (8).

Table 1.16: Did trade shocks and regulation changes affect emigration?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Δ out-migrant share of population				Δ log # of out-migrants			
Export tariff shock	-4.57	-2.54			-0.01	-0.02		
2001-2007	(3.24)	(2.90)			(0.15)	(0.14)		
Δ log regulation score			-0.52	0.26			0.04	0.06*
2001-2007			(0.61)	(0.50)			(0.03)	(0.03)
Controls (lagged)		X		X		X		X
Observations	250	237	250	237	250	237	250	237
R-squared	0.04	0.16	0.01	0.16	0.68	0.70	0.67	0.70
Mean (s.d.) of depent.	11.4 (6.4)				1.1 (0.55)			

Note: Standard errors are clustered at the province level. Dependent variables are changes from 2000 to 2010. Mean (sd) Δ log regulation score, 2001-2007 is 0.77 (0.82), mean (sd) export tariff shock is 0.18 (0.15). All columns control for import and intermediate tariff shock, log total population and level of dependent variable in 2000. Columns (2)(4)(6)(8) also control for lagged trade shocks and lagged wage and gdp growth rate, 1995-2001, as in Table 2 Column (8) .

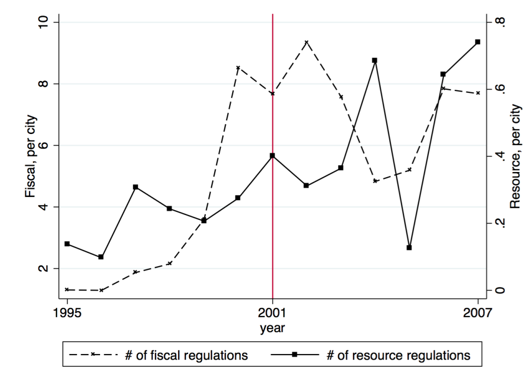
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

1.6.5.9 Placebo-Regulation Time Trend

One potential concern about the regulation data is that the number of regulations collected on the website is increasing over time, so the increase in migrant-related regulation might be just because of improved data availability. To alleviate the concern, I also count the total number of regulations on fiscal topics and resource topics.[45] In Figure 1.15, I plot the fiscal regulations with the dashed line and the resource regulation with the solid line. Both of them act as a placebo regulation type; neither line demonstrates a clear pattern, and there is no trend break around 2001.

[45]The website allows users to search by topic.

Figure 1.15: Trend of number of regulations, fiscal-related vs. resource related



1.7 Concluding Remarks

This paper uses the trade shock that happened after China entered the WTO to study the effects of trade liberalization on migration regulations. Instead of taking labor mobility as given, I use a simple political economy model to highlight the potential channel through which trade shocks can affect mobility restrictions and how these changes in regulation would affect the labor market outcomes and economic growth in general.

This empirical study shows that these indirect trade effects are statistically significant and economically sizable. Economic institutions have important interaction effects with specific economic conditions and economic shocks, so careful investigation and analysis on the interaction effect would deepen our understanding of the underlying economic forces.

One potential extension of this paper is to study the interaction of different types of regulations and how different local governments choose the set of regulation changes to achieve their objectives. One important aspect is industrial policies. For example, local governments can also provide tax credits or discounts for land usage fee to firms to attract capital flows. The complementarity between the industrial policies and labor policies might be important in evaluating their effects on economic outcomes. Importantly, the identification of a migration regulation effect in this paper relies on the assumption that other policies that were affected by trade liberalization should not affect migration policies. If migration policies and industrial policies are complements, and local governments decides them simultaneously, then the identification assumption will be violated. I hope to address these issues in subsequent papers.

This paper focuses on migration regulations and the Hukou system overall. However, trade liberalization can affect other types of economic institutions as well.[46] The external force of WTO rules and the pressure of competing with a bigger international market forced Chinese governments to take measures to improve efficiency and increase transparency. Establishing the rule of law not only affects contemporaneous outcomes but also has long-run impacts on the economy. How to measure the effect of trade liberalization on these broader institutional features is also left to future study.

Another related topic is the competition between prefectures. In this paper, prefectures provide amenities to attract migrant workers. Thus, the increase in the overall migrant welfare puts pressure on each individual prefecture to increase their amenity level. This competition between prefectures can actually decrease the fiscal profit of prefecture governments. If we think about the fiscal profit of local governments as economic rents, then the competition is welfare-improving for the economy since rents become smaller.

[46] According to the Deputy Director of Foreign Affairs Department, Legal Affairs Office, State Council of China: "After joining the WTO, a new set of rules must be applied through China's domestic law... According to the State Council Legislative Affairs Office's incomplete statistics, as of December 2002, the central government developed, modified, or abolished more than 1,000 laws, administrative regulations, departmental rules, and policy measures. All localities began to clean up in September 2001 in accordance with the unified arrangements. By the end of June 2002, 31 provinces, autonomous regions and municipalities had cleared more than 2 million pieces. Since then, the central and local conditions have continued to be modified and adjusted on a timely and planned basis." Source: Zhang, Zhoulai and Lei, Min, "The Largest-scale Regulation Change within the 5 Years after the WTO Accession." Xinhua News, 2006-12-10. The article is republished on the website of the Ministry of Commerce: www.mofcom.gov.cn/article/zt_rswzn/subjectm/200612/20061204045235.shtml.

However, in the current relaxed Hukou regime, prefectures with different economic fundamentals set amenities at different levels. As a result, wages are not equalized across prefectures. This creates spatial misallocation and decreases efficiency. The ideal system would be a completely free mobility regime where the central government provides a uniform level of migrant amenities, and rural workers can choose to work in any prefecture. The complete abolishment of the Hukou regime could benefit the overall economic growth, but the establishment of the nation-wide welfare system is politically difficult.

Data Appendix

Coding of Regulation Score

I need the regulations that potentially affect the utility of migrant workers, by either changing the income they get, welfare or amenity, or give them access to local Hukou (which will indirectly provide income, welfare and amenity benefits). I follow the following steps to extract these information:

1. I have a migrant-friendliness index for each regulation, and I will refer to this index as score in the paper. The score has the scale of five levels: -2 as very against migrants, -1 as against migrants, 0 as neutral, 1 as favorable to migrants, and 2 as very favorable to migrants.

2. Rate the documents with very short length (fewer than 200 Chinese characters) as 0, since they are usually purely administrative regulations for notices.

3. Rate the documents with pure administrative contents (for example, informing the logistics of getting some documents, certificates or proofs.) as 0.

4. Rate the documents related to birth control as 0, since people are subject to birth control both in their home regions and in the regions where they live temporarily, and it is not clear which rules are more strict. In some cases, these documents mention providing healthcare services to pregnant women and free vaccination to kids, and I code them as 1.

5. For the documents related to temporary residence, most of them are coded as 0. In most places there are still temporary residence registration requirements, and although there have been revisions of the terms, the revisions tend to be minor. In some cases, these documents mention reducing the

fee for registration, and simplifying the procedures significantly, and I code them as 1.

6. For documents of all other topics, the coding rule is: (1) If the document is about setting up a complete and executable guideline for one specific issue (for example, how to guarantee payment of wages to migrant workers, what is the rules for firms to purchase injury insurance and medical insurance for migrant workers, etc.), then I code it as 2 (-2 if it is against migrants); (2) If the document addresses one issue, but is more about enforcement of the specified rules (for example, guarantee the payment of wages before the Chinese new year), then I code it as 1 (-1 if it is against migrants); in some cases, the enforcement is very detailed and contains some guideline components, and I code it as 2 (-2 if it is against migrants); (3) If the document addresses two or more issues, and are either about guidelines or enforcement, then I code it as 2.

Figure 1.16 shows the wordclouds of the very pro-migrant and very anti-migrant regulations. A wordcloud shows the words used in highest frequencies, and the size of a word is positively correlated with the frequency. Panel (a) shows the wordcloud of the regulations of a score of -2, and the words with highest frequencies include "administrative penalties", "fines", "remedy", "warn", "deport", and "illegal". Panel (b) shows the wordcloud of the regulations of a score of 2, and the words with highest frequencies include "training", "loans", "wages", "service", "injury insurance", and "wage arrears".

Figure 1.16: Wordclouds for very pro-migrant regulations and very anti-migrant regulations

(a) Very anti-migrant (a score of -2) (b) Very anti-migrant (a score of 2)

Industry Crosswalk, from 2-digit GB Code to 2-digit SIC Code

Table 1.17: Crosswalk, 2-digit Chinese industry code (GB) to 2-digit U.S. industry code (SIC), secondary sector

GB	GB description	SIC	SIC description
6	Mining and washing of coal	12	Coal and lignite
7	Extraction of petroleum and natural gas	13	Crude petroleum and natural gas
8	Mining and processing of ferrous metal ores	10	Metallic ores and concentrates
9	Mining and processing of non-ferrous metal ores	10	Metallic ores and concentrates
10	Mining and processing of nonmetal ores	14	Nonmetallic minerals, except fuels
11	Mining of other ores	14	Nonmetallic minerals, except fuels
13	Processing of food from agricultural products	20	Food and kindred products
14	Manufacture of foods	20	Food and kindred products
15	Manufacture of liquor beverages and refined tea	20	Food and kindred products
16	Manufacture of tobacco	21	Tobacco manufactures
17	Manufacture of textile	22	Textile mill products
18	Manufacture of textile fabrics wearing apparel etc.	23	Apparel and related products
19	Manufacture of leather, fur, feather etc.	31	Leather and leather products
20	Processing of timber, manufacture of wood, etc.	24	Lumber and wood products, ex. fuel
21	Manufacture of furniture	25	Furniture and fixtures
22	Manufacture of paper and paper products	26	Paper and allied products
23	Printing, production of recording media	27	Printing, publishing, and etc.
24	Manufacture of articles for culture, education, etc.	26*	Paper and allied products
25	Processing of petroleum and coking	29	Petroleum refining and related prod
26	Manufacture of raw chemical material etc.	28	Chemicals and allied products
27	Manufacture of medicines	28†	Chemicals and allied products
28	Manufacture of chemical fibers	28	Chemicals and allied products
29	Manufacture of rubber	30	Rubber and miscellaneous plastics

*https://www.osha.gov/pls/imis/sicsearch.html?p_sic=&p_search=stationery.
†https://www.osha.gov/pls/imis/sicsearch.html?p_sic=&p_search=drug.
‡https://www.osha.gov/pls/imis/sicsearch.html?p_sic=&p_search=plastic.
⋇https://www.osha.gov/pls/imis/sicsearch.html?p_sic=&p_search=metal. Here the SIC 300 will be the weighted average of SIC 33 (Primary metal products) and 34 (Fabricated metal products).

Table 1.18: Crosswalk, 2-digit Chinese industry code (GB) to 2-digit U.S. industry code (SIC), secondary sector, continued.

GB	GB description	SIC	SIC description
30	Manufacture of plastics products	28‡	Chemicals and allied products
31	Manufacture of non-metallic mineral products	32	Stone, clay, glass, and concrete
32	Smelting and pressing of ferrous metals	300⋇	Metal processing and products
33	Smelting and pressing of non-ferrous metals	300⋇	Metal processing and products
34	Manufacture of metal products	300⋇	Metal processing and products
35	Manufacture of general purpose machinery	35	Machinery, except electrical
36	Manufacture of special purpose machinery	35	Machinery, except electrical
37	Manufacture of transportation machinery	37	Transportation equipment
39	Manufacture of electrical machinery and equipment	36	Electrical machinery, equipment, etc.
40	Manufacture of communication equipment, etc.	36	Electrical machinery, equipment, etc.
41	Manufacture of measuring instruments etc.	38	Scientific and professional instruments
42	Manufacture of artifacts and etc.	39	Miscellaneous manufacturing
43	Recycling	91	Scrap and waste material

*https://www.osha.gov/pls/imis/sicsearch.html?p_sic=&p_search=stationery.
†https://www.osha.gov/pls/imis/sicsearch.html?p_sic=&p_search=drug.
‡https://www.osha.gov/pls/imis/sicsearch.html?p_sic=&p_search=plastic.
⋇https://www.osha.gov/pls/imis/sicsearch.html?p_sic=&p_search=metal. Here the SIC 300 will be the weighted average of SIC 33 (Primary metal products) and 34 (Fabricated metal products).

Firm Information from the Enterprises Survey

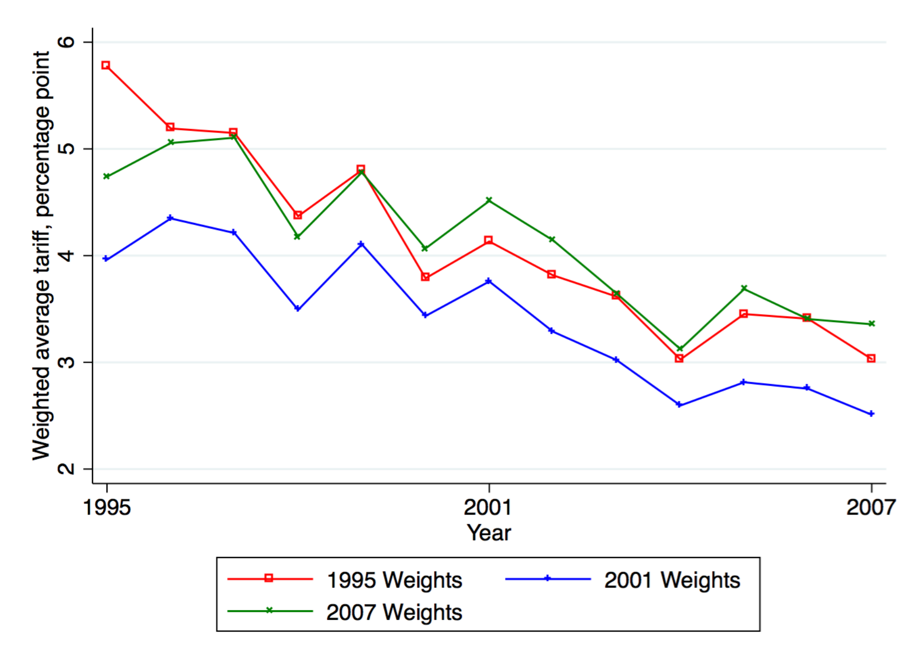
Table 1.19: Summary of statistics, firms

Year	Number of Firms (100,000)		Employment (10 million)		Sales Revenue (Billion RMB)		Export Value (Billion RMB)	
	Total	SOE %	Total	SOE %	Total	SOE %	Total	SOE %
2000	1.4	46%	5	57%	7.5	42%	1.4	21%
2001	1.5	35%	4.8	46%	8.5	32%	1.6	14%
2002	1.6	28%	5.1	40%	10	29%	2	12%
2003	1.8	21%	5.3	32%	13.1	23%	2.7	9%
2005	2.5	10%	6.5	19%	22.5	16%	4.7	5%
2006	2.8	8%	6.9	16%	28.3	14%	6	5%
2007	3.2	6%	7.4	14%	36.3	12%	7.3	5%

Trade Data Description

Alternative way to construct tariff values by industry.

Figure 1.17: Time trend of average tariff faced by Chinese exporters, 1995-2007, using different weights



*Note: each dot here is weighted average of industrial level tariffs, where the weights are shares of exports in this industry. For the industry level tariff, it is constructed as the weighted average of destination country tariffs on Chinese exports in the specific industry, where the weights are shares of exports in this destination country in the specific industry.

Number of partner countries.

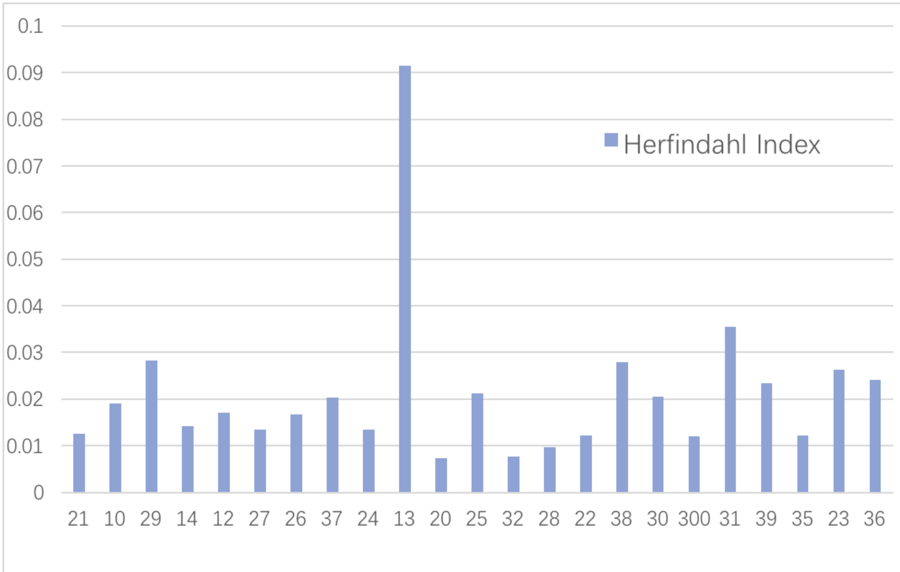
Table 1.20: Number of partner countries

Year	1995	2001	2007
Number of partner countries	46	115	111

Industry concentration across cities, Herfindahl Index, 2000 data.

To show that most industries are quite spread out.

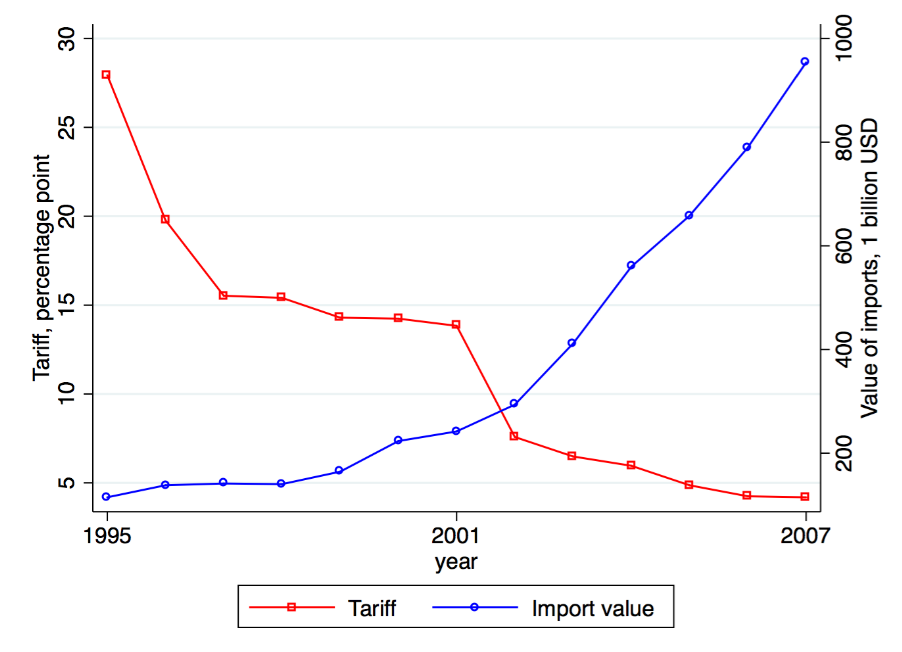
Figure 1.18: Industry concentration across cities, Herfindahl Index, 2000



*Note: each bar is a industry. Horizontally sorted by value of exports in the industry in 2000.

Trend in import tariff and import values.

Figure 1.19: Time trend of average import tariff and overall import volume, 1995-2007



*Note: each dot here is weighted average of industrial level tariffs on Chinese imports.

Number of partner countries in the bilateral trade flow data and distance data

Table 1.21: Number of partner countries in the bilateral trade flow data and distance data

Year	# of importers	# of exporters	# of bilateral trade flows	# of countries in the distance data
1995	113	226	14,398	224
2001	165	238	22,416	224
2007	173	236	26,268	224

Census Data Description

Questions to Identify Migrants in 2000, 2005 data.

Table 1.22: Questions to identify migrant workers

Year	Question‡	Options	Migrant?
2000	(1) Hukou registration status	a) this village/street/town	No
		b) other village/street/town, lived here for more than half a year	Yes
		c) other village/street/town, left Hukou place for more than half a year	Yes
		d) not determined.	-
		e) other countries	-
	For the ones answering b) c) to (1):		
		a) this county, but different village	Yes
		b) this county, but different town	Yes
		c) this county, but different street	Yes
		d) this city , but different (county and) village	Yes
		e) this city, but different (county and) town	Yes
		f) this city, but different (county and) street	Yes
		g) this province, but different city	Yes
		h) other province (if yes, need to specify the exact province)	Yes

Year	Question†	Options	Migrant?
2005	(1) Hukou registration place	a) this village/street/town	No
		b) this county, but different village/street/town	Yes
		c) other county (if yes, need to specify the exact county)	Yes
		d) not determined.	-
			Keep?
	(2) Residence at the time of survey	a) this neighborhood	Yes
		b) this village/street/town, but different neighborhood	No
		c) this county, but different village/street/town	No
		d) other county (if yes, need to specify the exact county)	No
		e) other countries	No

†http://www.stats.gov.cn/tjsj/ndsj/renkou/2005/html/03.htm
‡http://www.stats.gov.cn/tjsj/ndsj/renkoupucha/2000pucha/html/appen4.htm

Description of the 2005 mini-Census and Summary of Statistics for the Migrant Data

The 2005 mini-census took place on November 1, 2005, and intended to get 1% population sample in the whole country. The mini-census used stratified, multi-stage, clustered probability sampling method, where the final sampling unit was neighborhood (cluster of buildings that shared common management services). All people surveyed were Chinese citizens and residents. The overall sample size was 17.05 million, and the estimated total population of mainland China was 1,307.56 million, so the actual sampling rate was 1.31%.[47]

The survey questionnaire is filled in by each household. Within the household, two kinds of people should be registered: (1) the ones who are living in the household at the night of census; (2) the ones whose Hukou is in the household, but is not living in the household. This means that if someone is living in the place outside Hukou registration, he/she could be registered twice if only part of the household migrated but part of the household stays at the Hukou place. This is different from the 2000 Population Census where each person is only registered once in the "permanent residence".

Overall, the survey covers basic demographic information of household members, living conditions, Hukou registration places, reasons of migration, and the time since leaving Hukou place. For people aged 15 and above, there are questions about working status, industry and occupation, and social insurances (unemployment insurance, pension, and medical insurance). There are also marital status and basic birth and death information. The data I use has 2, 585, 480 individuals, and is a 0.2% sample of the whole population (or 15% sample of the whole sample). [48]

[47]Data source: http://www.stats.gov.cn/tjsj/tjgb/rkpcgb/qgrkpcgb/200603/t20060316_30326.html

[48]Summary of statistics, questionnaires, and explanation of the questionnaires can be found: http://www.stats.gov.cn/tjsj/ndsj/renkou/2005/renkou.htm

Table 1.23: Summary of statistics for migration flow data

Year	2000			2005			2010
	Sample	Population¶	Sampling rate	Sample	Population†	Sampling rate	Population‡
# of persons	1,180,111	1,295,330,000	0.1%	2,585,480	1,306,280,000	0.2%	1,339,724,852
# of households	345,167	348,370,000	0.1%	996,607	395,190,000	0.25%	401,517,330
# of cities	340	340		340	334?		
# of migrants	124,519	144,390,748	0.08%	263,534	147,350,000	0.18%	261,386,075

¶www.stats.gov.cn/tjsj/tjgb/rkpcgb/qgrkpcgb/200203/t20020331_30314.html
†www.stats.gov.cn/tjsj/tjgb/rkpcgb/qgrkpcgb/200603/t20060316_30326.html
‡www.stats.gov.cn/tjsj/tjgb/rkpcgb/qgrkpcgb/201104/t20110428_30327.html

Table 1.24: Summary of statistics

Mean (sd) in million persons	2000	2005	2010
Total population	3.5 (2.7)	3.8 (4.0)	3.9 (3.2)
# of locals	3.1 (2.3)	3.2 (3.3)	3.1 (2.3)
# of migrants	.39 (.63)	.56 (1.20)	.77 (1.31)
By migration distance			
within cities (short-distance)	.19 (.22)	.25 (.45)	.11 (.28)
across cities (medium-distance)	.09 (.15)	.08 (.22)	.39 (.41)
across prov (long-distance)	.11 (.37)	.23 (.88)	.25 (.85)
By reason of migration			
Work	.12 (.36)	.26 (.83)	.41 (.80)
Family	.04 (.06)	.09 (.18)	.13 (.15)
Marriage	.01 (.02)	.04 (.08)	.04 (.04)
Other	.20 (.26)	.16 (.38)	.15 (.24)
By years of education			
<=12 years of education	.34 (.55)	.45 (1.11)	.59 (0.99)
> 12 years of education	.03 (.07)	.06 (.23)	.13 (.29)
By years since moved here			
<=3 years	.19 (.40)	.25 (.69)	.43 (.70)
> 3years	.16 (.22)	.26 (.68)	.32 (.62)

Prefecture-Level Market Access Shock

This section shows the construction of market-access-based trade shocks. The idea is that suppose the overall export (import) volume increases in a certain industry over time at the national level, and I can calculate per capita export growth by dividing the increase in export (import) volume by the total number of people employed in the industry. Then I can distribute the per capita export growth across regions according to the share of employment in the industry in a certain region, and then generate the overall regional trade shock by summing over industries. Specifically, following Autor et al. (2013), the formula to calculate regional export exposure is as follows:

$$\Delta IPW_{it}^{M} = \sum_{j} \frac{L_{ijt}}{L_{jt}} \frac{\Delta M_{jt}}{L_{it}} = \sum_{j} \frac{L_{ijt}}{L_{it}} \frac{\Delta M_{jt}}{L_{jt}}$$

$$\Delta IPW_{it}^{X} = \sum_{j} \frac{L_{ijt}}{L_{jt}} \frac{\Delta X_{jt}}{L_{it}} = \sum_{j} \frac{L_{ijt}}{L_{it}} \frac{\Delta X_{jt}}{L_{jt}}$$

where L_{it} is the start of period employment (year t) in region i and L_{jt} is the start period employment in industry j, L_{ijt} is the start of period employment in region i and industry j. ΔM_{jt} is the observed change in China's import from the rest of the world in industry j between the start and the end of the period. Labor market exposure to import competition is the change in import exposure per worker in a region (in Autor et al. (2013) is the change in Chinese import exposure), where imports are apportioned to the region according to its share of national industry employment. And the export exposure is calculated by replacing observed change in China's import from the world ΔM_{jt} with China's export to the world ΔX_{jt}.

The primary measure of interest in my paper is ΔIPW_{it}^{X}. The Bartik instrument uses the overall national growth to generate regional growth, by interacting with initial conditions. The benefit here is that it will be free of other local shocks that are correlated with local export growth. However, using observed trade volume increase might still be problematic, since the overall trade increase might still be correlated with overall economic growth, and then the result will capture the "economic growth effect" instead of "trade growth effect". Thus, I will use two ways to instrument the trade volume change further. The first is to use the importing country's income growth to instrument for China's export growth. The second is to use gravity dummies instead of GDP.

The GDP based instrument is constructed in the following way. Suppose a country's fraction of income allocated to different industries' consumption (import) do not change over time, and the fraction of import in an industry that comes from China also do not change over time, then the growth of demand for Chinese goods will be from the growth of importing country's income level. Specifically, the import value of Chinese goods in industry j (which I use as the super script instead of sub script) and year t is constructed as the following:

$$
\begin{aligned}
\mathrm{X}_{jt} &= \sum_n \frac{X^{cn}_{jt*}}{\sum_{n'} X^{n'n}_{jt*}} \frac{\sum_{n'} X^{n'n}_{jt*}}{\sum_{j'} \sum_{n'} X^{n'n}_{j't*}} GDP^n_t \\
&= \sum_n \frac{X^{cn}_{jt*}}{X^n_{jt*}} \frac{X^n_{jt*}}{X^n_{t*}} GDP^n_t \\
&= \sum_n \frac{X^{cn}_{jt*}}{X^n_{t*}} GDP^j_t
\end{aligned}
$$

where $\frac{X^{cn}_{jt*}}{X^n_{jt*}}$ is the fraction of import in industry j and country n that comes from China in baseline year $t*$, and $\frac{X^n_{jt*}}{X^n_{t*}}$ is the fraction of import in industry j and country n out of the total import value. Then the export market access shock with the GDP measure in industry j between year t and t' is defined as

$$
\Delta X^{GDP}_{jt} = \sum_n \frac{X^{cn}_{jt*}}{X^n_{t*}} \left(log(GDP^n_{t'}) - log(GDP^n_t)\right)
$$

In Figure 1.20, I sort the industries by the value of export in 2000, and show the industry level shock constructed from the importing country GDP growth (ΔX^{GDP}_{jt}).

Figure 1.20: Industry-level trade shock, using importing country GDP growth, 2001-2007

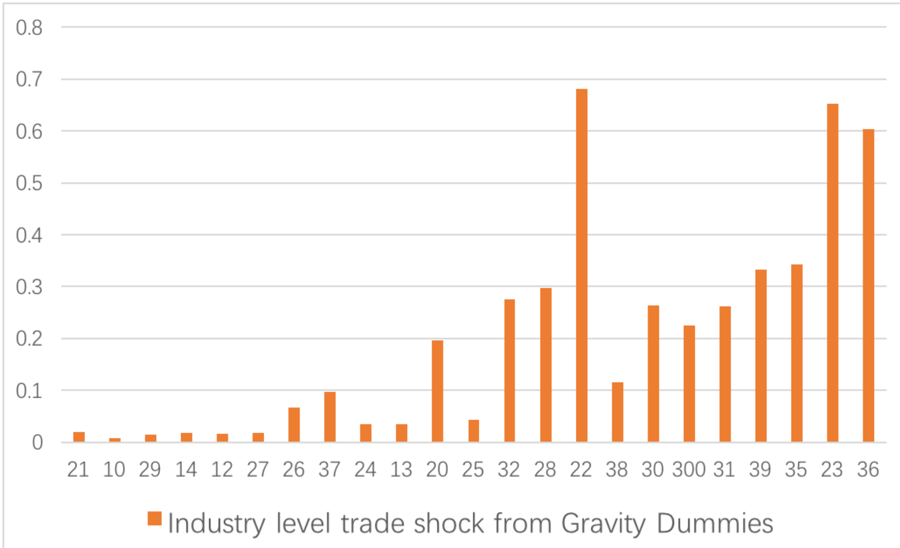
Note: Each bar is an industry. Horizontally sorted by value of exports in the industry in 2000.

Alternatively, I can use gravity dummies instead of GDP growth. First I run a regression of log pairwise country imports on origin and destination country dummies, controlling for geographic distances. Then I construct the export market access shock with gravity measure as

$$\Delta X_{jt}^{Gravity} = \sum_{n} \frac{X_{jt*}^{cn}}{X_{t*}^{n}} \left(D_{n,t'} - D_{n,t}\right)$$

In Figure 1.21, I sort the industries by the value of export in 2000, and show the industry level shock constructed from the importing country gravity dummy changes ($\Delta X_{jt}^{Gravity}$). We can see that the results are quite similar to the ones in Figure 1.20.

Figure 1.21: Industry-level trade shock, using gravity dummies, 2001-2007



Note: Each bar is a industry. Horizontally sorted by value of exports in the industry in 2000.

After getting ΔX_{jt}^{GDP} and $\Delta X_{jt}^{Gravity}$, I use them instead of ΔX_{jt} to calculate ΔIPW_{it}^{X}.

I also add a measure which is the intermediate goods market access shock.

$$\Delta IPW_{it}^{I} = \sum_{j} \frac{L_{ijt}}{L_{it}} \hat{I}_{ijt}$$

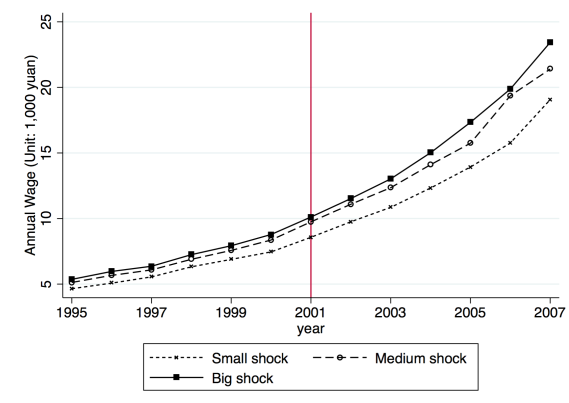
and

$$\hat{I}_{ijt} = \sum_{j'} \frac{input_{j'}^{ij}}{\sum_{j''} input_{j''}^{ij}} \left[\frac{M_{j't'}}{L_{j't}} - \frac{M_{j't}}{L_{j't}} \right]$$

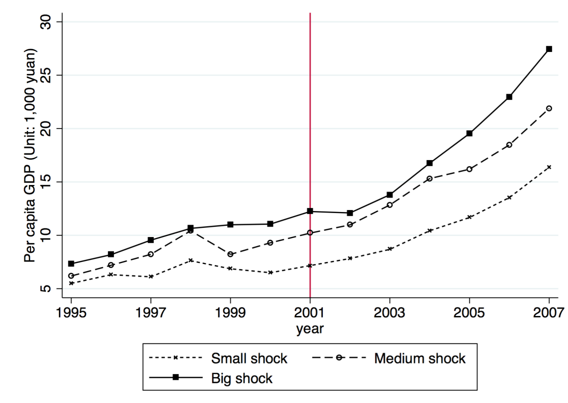
Trends of Wages and Employment, 1995-2007

Figure 1.22: Trends of wage, per capita GDP, and total urban employment, prefecture yearbook data

(a) Wage



(b) Per capita GDP



(c) Urban employment

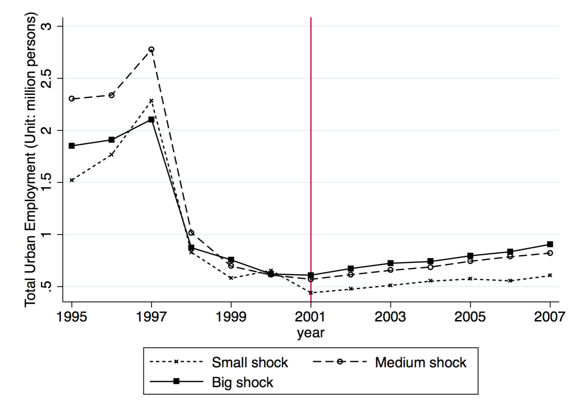
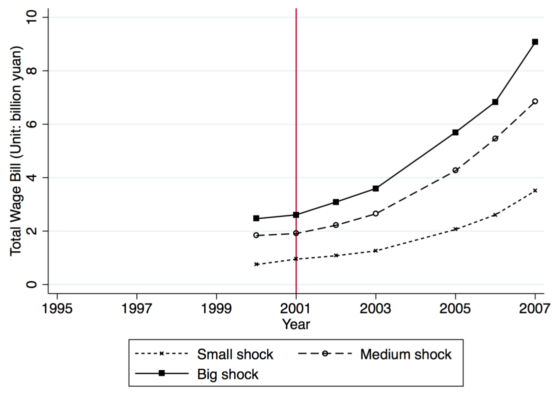
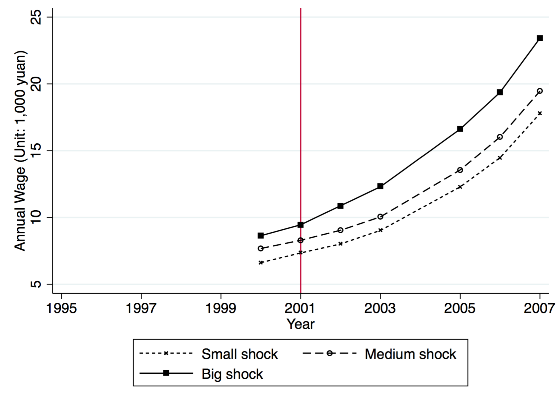


Figure 1.23: Trends of wagebill, wage, total employment, sales, and exports, industrial survey data

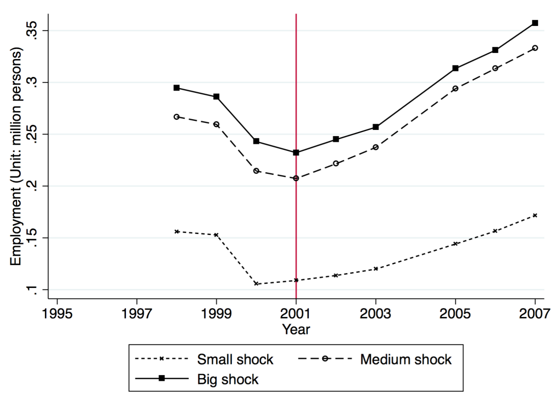
(a) Wage bill



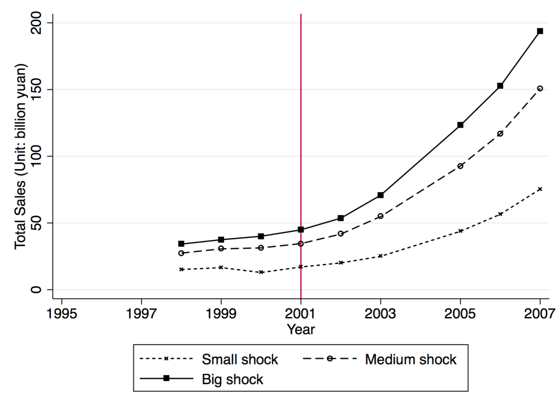
(b) Wage



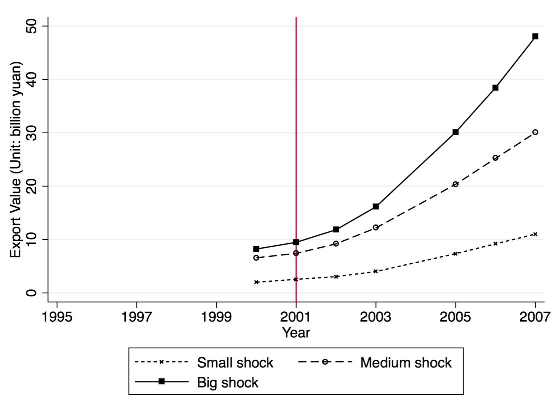
(c) Employment



(d) Sales



(e) Exports



Trends of Wages and Employment, 1995-2007

Migrant Intensity by Industry

Table 1.25: Industry composition (agriculture vs. others), migrant workers vs. local workers

	Local	Migrant
Agricultural	66%	7%
Non-agricultural	34%	93%

Table 1.26: Industry composition (only non-agricultural), migrant workers vs. local workers

Industry category	Local	Migrant
Mining	3.15	1.47
Manufacturing	28.01	39.15
Production and supply of power, gas and	2.17	0.99
Construction	7.75	7.32
Transportation, storage and post	9.28	5.59
Information communication, computer and	0.97	1.26
Wholesale and retail	15.85	18.12
Catering and accommodation	3.86	6.45
Finance	1.69	1.24
Real estate	0.96	1.23
Rental and business services	1.24	1.67
Scientific research, technology service	0.82	0.68
Water, environment and public infrastruc	0.96	0.63
Residential and other services	4.21	5.2
Education	7.1	2.92
Health, social security and social welf	3.28	1.59
Cultural, sports and entertainment	1.04	1.16
Public administration and social organi	7.64	3.32
International organization	0	0.01
Total	100	100

Table 1.27: Share of migrant workers, among manufacturing industries

ind	Share of migrants	# workers
Mining and processing of ferrous metal ores	15%	1129
Mining and washing of coal	18%	8736
Extraction of petroleum and natuaral gas	19%	1253
Mining and processing of non-metal ores	20%	2234
Processing of food from agricultural products	22%	6005
Other mining and processing	22%	479
Manufacture of tobacco	22%	585
Smelting and pressing of ferrous metals	23%	4770
Processing of petroleum, coking, processing of nuclear fuel	23%	1432
Manufacture of chemical fibers	23%	606
Manufacture of beverages	25%	2207
Mining and processing of non-ferrous metal ores	25%	1221
Manufacture of raw chemical materials and chemicals	26%	7474
Manufacture of non-metallic mineral products	26%	13181
Processing of timber, manufacturing of wood, bamboo, etc.	27%	4966
Smelting and pressing of non-ferrous metals	30%	2638
Manufacture of medicines	30%	2425
Manufacture of general purpose machinery	31%	7896
Manufacture of foods	33%	5870
Manufacture of paper and paper products	33%	4176
Manufacture of textile	33%	14912
Manufacture of transport equipment	35%	7622
Manufacture of special purpose machinery	35%	6340
Manufacture of artwork and other manufacturing	38%	7966
Manufacture of rubber	46%	1800
Manufacture of metal products	46%	11747
Manufacture of textile wearing apparel, footwear, and caps	47%	19474
Printing, reproduction of recording media	48%	3181
Recycling and disposal of waste	50%	1623
Manufacture of plastics	54%	6241
Manufacture of furniture	55%	4167
Manufacture of measuring instruments and machinery etc.	56%	2969
Manufacture of electrical machinery and equipment	56%	9161
Manufacture of leather, fur, feather and related products	63%	6666
Manufacture of communication equipment, computers and etc.	68%	12135
Manufacture of articles for culture, education and sport activity	69%	3794

Theory Appendix

Model Setup

Model Overview

I model an open economy with N+1 regions with $n \in \{1, 2, ..., N, r\}$. All $n \leq N$ index cities, and $n = r$ is a rural area . [49] Each city is endowed with immobile local labor L_i and an immobile fixed factor R_i (for example, land). A mass of $\bar{M}$ rural residents can either live in the rural area or move to one of the cities and work as migrant workers. The production technology in each city uses the migrant labor, local labor, and the fixed factor as inputs, and produces a unique product that is sold on the international market. The unique product can be viewed as a composite good. The technology of production is different across cities.

Each city has an endowed natural amenity C_i.[50] Rural residents who live in City i enjoy the wage, the natural amenity C_i, and a local public good A_i provided by the city government. If they choose to stay in the rural area, they will enjoy a fixed amenity level of C_r and a fixed wage of w_r.[51]

Given the technology and factor endowment, the city government chooses the amount of amenities it provides for the migrant labor, conditioning on the Hukou system set by the central government.[52] The central government chooses one of the two systems $S \in \{0, 1\}$. $S = 0$ is the strict Hukou system, where the local government is not allowed to manipulate the level of amenities offered to migrants.

[49] This is a simplifying assumption, since I group the rural areas of all prefectures into one.

[50] The natural amenity can include the air quality, transportation infrastructure, landscape, and other nonexclusive features.

[51] C_r could include the value of the attachment to homeland, eligibility to be part of the rural social network, and the right to use farmland. For example, Munshi and Rosenzweig (2016) shows that local risk-sharing networks provide informal insurance and restrict migration in India.

[52] The amenity is equivalent to a migrant subsidy.

$S = 1$ is the relaxed Hukou system, where local governments are free to choose the amenity level for migrants.

Preferences and the Worker's Decision Problem

The indirect utility of rural worker l living in region n is $v_n^l = v_n \epsilon_n^l$, where ϵ_n^l represents worker l's idiosyncratic taste for living in region n, and v_i is common for all rural workers who live in region n. v_n is determined by the amenity level and wage level in the following way:

$$v_n = (C_n + A_n)^{\beta_1} w_n^{\beta_2}$$

where $\beta_1 > 0$ and $\beta_2 > 0$.

In rural areas, amenities are normalized to zero, $A_r = 0$.

Worker l will choose to live in the region n that maximizes their utility, so $n = \text{argmax}_{n'} v_{n'} \epsilon_{n'}^l$. There is an idiosyncratic taste draw ϵ_n^l for each worker and area, and the draw is i.i.d. across workers and areas from a Fréchet distribution, $Pr(\epsilon_n^l \leq x) = e^{-x^{-\epsilon}}$, with $\epsilon > 1$. Thus, the number of migrant workers who live in area n is

$$M_n = \left(\frac{v_n}{v}\right)^{\epsilon} \bar{M},$$

where $v \equiv \left(\sum_n v_n^{\epsilon}\right)^{1/\epsilon}$. Under the Fréchet distribution, the average utility of migrant workers in all cities will be the same and proportional to v, thus v will be used as the measure of worker utility.

Plugging in the expression for v_n, the labor supply equation for migrant labor is

$$M_n = \left(\frac{(C_n + A_n)^{\beta_1} w_n^{\beta_2}}{v}\right)^{\epsilon} \bar{M} \tag{1.2}$$

Production Technology and the Firm's Problem

Firms in each city i use the local labor, the migrant labor, and a fixed factor to produce a unique product, but the product is different across cities.[53] Both local labor and the local fixed factor are supplied inelastically: $L_i = \bar{L}_i$, $R_i = \bar{R}_i$. The price of the product i, p_i, is determined on the international market. The production function is Cobb-Douglas with constant return to scale, and output Y_i can be written as:

$$Y_i = p_i \mu_i' L_i^{\alpha_{1,i}} M_i^{\alpha_{2,i}} R_i^{1-\alpha_{1,i}-\alpha_{2,i}}$$

where $\alpha_{1,i}$ and $\alpha_{2,i}$ are both positive, and $\alpha_{1,i} + \alpha_{2,i} < 1$.[54]

This is equivalent to writing the production function as

$$Y_i \equiv p_i \mu_i M_i^{\alpha_i}$$

where $\mu_i \equiv \mu_i' L_i^{\alpha_{1,i}} R_i^{1-\alpha_{1,i}-\alpha_{2,i}}$, and $\alpha_i \equiv \alpha_{2,i}$.

The firm's output is subject to a sales tax t. The firm maximizes profits by choosing

$$\max_{M_i} \quad (1-t)Y_i - w_i M_i$$

where w_i is the wage of migrant workers in City i. The market is perfectly competitive, and each firm earns zero profit.

[53] In 2000, the average years of education for urban residents age 15 and above was 10.3, while the number for migrant workers from rural areas was 8.2. The 2-year gap persisted until 2005. Thus, the migrant workers were relatively low-skilled compared to local urban residents. I calculate these numbers by using the 2005 mini-census.

[54] I use the Cobb-Douglas production to keep the model predictions simple. With CES production function, the intuition of the model remains, while results are more complicated. The local fixed factor can be fixed capital or land. Adding mobile capital will not change the main results of the model.

First-order condition of the firm gives:

$$w_i = \alpha_i(1-t)p_i\mu_i M_i^{\alpha_i - 1} \tag{1.3}$$

The City Government's Problem

The local government's objective is to maximize net fiscal profit, which is equal to tax revenue minus expenditure on public services. In the literature on political economy in China, Li and Zhou (2005), Jia et al. (2015), and Jia (2017a) argue that local governments are driven by promotion incentives, and as a result, they seek to maximize GDP growth. Instead, I adopt the fiscal incentive approach as in Gordon and Li (2012). The rationale for the objective function is that only few local government officials are promoted, and the rest of them stay in the system because they can get the rent or net fiscal profit. The advantage of this approach here is that I can model the direct benefit of migrant inflow and the direct cost explicitly in the objective function: benefits in terms of increasing the tax income, and costs in terms of amenity provision.[55] These tradeoffs are present not only in China but also in cases of low-skilled immigration in the United States and Europe.[56]

In addition, I assume that each city government has a negligible impact on the overall migrant welfare v and takes v as given. There are 340 prefectures in China. In 2000, the city of Shenzhen had the biggest number of migrants,

[55] This specification has an alternative interpretation, where there is no tax on firm revenues, but the city government prefers a bigger economy and smaller expenditure on migrants, and t is the weight for utility from a bigger economy when the weight for disutility of migrant expenditure is normalized to 1.

[56] For example, in the United States, there is a debate on whether to provide immigrant children with Medicaid. (www.latimes.com/local/politics/la-me-immigrants-medi-cal-20160427-story.html). In Europe, there is a discussion on how welfare program generosity affects migrant skill mix and in turn affects the strength of the welfare-state institution (voxeu.org/article/immigration-and-welfare-state-new-evidence-eu).

5,622,000; however, this was only 4% of the total national migrant population.[57] Thus, the assumption is very reasonable in this context due to the overall size of the Chinese population and the number of cities.

(1) Relaxed Hukou System In the relaxed Hukou system, where $S = 1$, the city government maximizes the net fiscal profit by choosing the amenity level A_i for migrant workers and the number of migrants M_i:

$$\max_{M_i, A_i} \quad t \cdot Y_i - A_i M_i$$

subject to labor supply (Equation 1.2) and labor demand (Equation 1.3) constraints.

The city government has full information about production and labor supply. Given the labor supply equation, A_i can be solved as a function of M_i. Thus the first-order condition of the local government is given by

$$t \cdot \frac{\partial Y_i}{\partial M_i} = A_i + \frac{\partial A_i}{\partial M_i} M_i. \tag{1.4}$$

(2) Strict Hukou System In the strict Hukou System, where $S = 0$, the city government faces the same maximization problem subject to the constraint that $A_i = 0$. So there is no maximization: M_i is determined by the labor supply equation and labor demand equation, and the city government will take it as given.

The Central Government's Problem

The central government is interested in increasing the total output (or GDP), which is $Y = \sum_{i=1}^{N} Y_i + M_r w_r$. At the same time, a sudden, large inflow of

[57] I caculate this by using the 2000 Census data.

population into the city might cause regime instability, and cause the city to incur burdensome administrative and bureaucratic costs; these will impose a cost to the central government.[58] The central government has full information about the economy and the decision process of the city governments and chooses $S \in \{0, 1\}$.

Assume that the central government starts with the strict Hukou system and the decision rule is as follows:

$$S = \mathbf{1}(Y_{S=1} - Y_{S=0} \geq \underline{Y}) \tag{1.5}$$

where $\mathbf{1}(\cdot)$ is an indicator function. This means that if the cost of immobility $(Y_{S=1} - Y_{S=0})$ is big enough, the central government will switch to the relaxed Hukou system. However, once in the relaxed system, the cost of switching back to the strict system is very high (suppose it is impossible for now).

General Equilibrium

A general equilibrium of this economy consists of the distribution of workers $\{M_n\}_{n\in\{1,2,...\}}$, city output values $\{Y_i\}_{i\in\{1,2,...\}}$, wages $\{w_i\}_{i\in\{1,...,N\}}$, amenities $\{A_i\}_{i\in\{1,...,N\}}$, the migrant welfare measure v, the type of Hukou system S, and economy-wide GDP Y such that (1) firms make optimal decision about production; (2) rural workers make optimal location decision; (3) city governments make the optimal decision about amenity provision; (4) the central government makes the optimal decision about the state of mobility; (5) city-level labor markets clear; and (6) the national labor market clears, i.e., $\sum_n M_n = \bar{M}$.

[58]For example, a report on the website of National Bureau of Statistics of China points out that the crime rate among temporary residents is 12.8%, which is 4 times the average crime rate. Source: www.stats.gov.cn/ztjc/ztfx/fxbg/200306/t20030606_14197.html.

Comparative Statics[59]

Relaxed Hukou System $S = 1$:

Proposition 1: When there is a positive price shock in City i ($p_i\uparrow$), then the local government will provide more amenities for migrants ($A_i\uparrow$), and migrants will flow into the city ($M_i\uparrow$). Overall output in City i will increase ($Y_i\uparrow$).

Strict Hukou System $S = 0$:

Proposition 2: When there is a positive price shock in City i ($p_i\uparrow$), migrants will flow into the city ($M_i\uparrow$) and overall output in City i will increase ($Y_i\uparrow$). However, both the increase in number of migrants and the increase in output are smaller than in the Relaxed Hukou system.

Implication for the Overall Economy

Proposition 3: In the symmetric case, when all cities are the same and wage in the rural area is small enough, the overall output Y is an increasing function of the number of people who migrated. When there is an economy-wide positive price shock, both $Y_{S=1}$ and $Y_{S=0}$ will increase, and $Y_{S=1}$ will increase more. Thus, the central government will switch to the relaxed Hukou system.

Estimation Equations

I use $\hat{x} \equiv d\ln x$ to present percentage changes. I log-linearize the equilibrium equations and solve for the percentage changes of endogenous variables (amenity, migrant inflow, wage, total employment, and per capita GDP) as functions of the exogenous trade shock $\hat{p}_i$.[60]

[59] Please see proofs of propositions in Appendix 1.7.

[60] Assume that price changes are small and higher-order terms are negligible.

The key variable of interest is the percentage increase in the amenity level ($\hat{A}_i$):

$$\hat{A}_i = \frac{1}{S_{2,i}} \frac{S_{1,i} \cdot \frac{1}{\epsilon}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} \hat{p}_i \equiv f(\alpha_i)\hat{p}_i. \tag{1.6}$$

Equation 1.6 shows that when there is a positive price shock, the amenity level also increases ($f(\alpha_i) > 0$).[61] In addition, when α_i is bigger, the amenity is more responsive ($f'(\alpha_i) > 0$), meaning that in places that are more migrant intensive (or with higher output elasticity of output), a positive price shock leads to bigger changes in amenity level.

I then solve for the percentage change in the migrant inflow ($\hat{M}_i$), total urban employment ($\hat{E}_i = \widehat{L_i + M_i}$), migrant wages ($\hat{w}_i$), local wages ($\hat{w}_i^L$), wages of the total employment ($\hat{w}_i^T$) , and per capita GDP as functions of the exogenous trade shock $\hat{p}_i$. All of them are increasing functions of the trade shock. [62]

Additional Discussion on the Model Features

In the model, I group the rural areas of all prefectures into one and focus on migration from a single rural area into multiple urban areas. Also, when I talk about the prefecture-level government, I take the stand that it is only interested in the urban areas, and that its regulations are geared toward providing amenities that attract migrants to the urban areas.

This simplification is based on several features of the data and the institutional background. First, in the regulation documents, the treatment of migrants does not depend on their Hukou origin. This means that a migrant worker from another prefecture is usually treated the same as a migrant from the rural area of his or her

[61] $S_{2,i} = \frac{A_i}{A_i + C_i}$, $S_{1,i} = \frac{\alpha_i t(1-t) p_i \mu_i M_i^{\alpha_i - 1}}{\alpha_i t(1-t) p_i \mu_i M_i^{\alpha_i - 1} + C_i} = \frac{B_i}{B_i + C_i}$. B_n is the baseline per capita migrant contribution in taxes, A_i is the baseline government-supplied amenity level, and C_i is the baseline natural amenity level.

[62] Please see the details of the expressions in Appendix 1.7.

own prefecture. Second, the average agriculture share of prefecture tax revenue in 2000 was just 13%, meaning that the urban area is the major contributor to prefecture tax revenue.[63] Third, local governments have few mechanisms and little incentive to restrict rural residents from emigrating. Migrants usually earn higher wages in the urban areas and remit part of their income to their family back in their hometown. This remittance helps to alleviate rural poverty.

Admittedly, when a rural resident decides to migrate, the geographic distance between the origin and the destination is correlated with both the transportation cost and the cultural and language differences. Thus the supply of migrants could vary across prefectures. To keep the representation simple, the model does not include this feature, but the main implications of the model do not depend on this assumption.

Proof of Propositions.

Proposition 1: In the relaxed Hukou system ($S = 1$), when there is a positive price shock in City i ($p_i\uparrow$), then the local government will provide more amenities for migrants ($A_i\uparrow$), and migrants will flow into the city ($M_i\uparrow$). Overall output in City i will increase ($Y_i\uparrow$).

Proof:

Plug in the wage expression from the labor demand equation (Equation 1.3) into the labor supply equation (1.2)

$$M_i = \left(\frac{(C_i + A_i)^{\beta_1}\left(\alpha_i(1-t)p_i\mu_i M_i^{\alpha_i - 1}\right)^{\beta_2}}{v}\right)^{\epsilon}\bar{M}$$

Solving M_i as a function of A_i,

[63] I calculated this by using prefecture level fiscal revenue and expenditure data.

$$M_i = \left(\frac{\bar{M}}{v^\epsilon}(\alpha_i(1-t)p_i\mu_i)^{\beta_2\epsilon}\right)^{\frac{1}{1+(1-\alpha_i)\beta_2\epsilon}} (C_i + A_i)^{\frac{\beta_1\epsilon}{1+(1-\alpha_i)\beta_2\epsilon}}$$

Since $S = 1$, I can solve A_i as a function of M_i.

$$A_i = \frac{M_i^{\frac{1}{\beta_1\epsilon}+(1-\alpha_i)\frac{\beta_2}{\beta_1}}}{\left(\frac{\bar{M}}{v^\epsilon}\right)^{\frac{1}{\beta_1\epsilon}} (\alpha_i(1-t)p_i\mu_i)^{\frac{\beta_2}{\beta_1}}} - C_i$$

And $\frac{\partial A_i}{\partial M_i}$ can also be solved as a function of M_i.

$$\frac{\partial A_i}{\partial M_i} = \left(\frac{1}{\beta_1\epsilon} + (1-\alpha_i)\frac{\beta_2}{\beta_1}\right) \frac{M_i^{\frac{1}{\beta_1\epsilon}+(1-\alpha_i)\frac{\beta_2}{\beta_1}-1}}{\left(\frac{\bar{M}}{v^\epsilon}\right)^{\frac{1}{\beta_1\epsilon}} (\alpha_i(1-t)p_i\mu_i)^{\frac{\beta_2}{\beta_1}}}$$

Plug A_i , $\frac{\partial A_i}{\partial M_i}$ and w_i into Equation 1.4

$$t \cdot \alpha_i(1-t)p_i\mu_i M_i^{\alpha_i-1} = (1 + \frac{1}{\beta_1\epsilon} + (1-\alpha_i)\frac{\beta_2}{\beta_1}) \frac{M_i^{\frac{1}{\beta_1\epsilon}+(1-\alpha_i)\frac{\beta_2}{\beta_1}}}{\left(\frac{\bar{M}}{v^\epsilon}\right)^{\frac{1}{\beta_1\epsilon}} (\alpha_i(1-t)p_i\mu_i)^{\frac{\beta_2}{\beta_1}}} - C_i$$

Rearranging the terms,

$$t \cdot \alpha_i(1-t)p_i\mu_i M_i^{\alpha_i-1} + C_i = (1 + \frac{1}{\beta_1\epsilon} + (1-\alpha_i)\frac{\beta_2}{\beta_1}) \frac{M_i^{\frac{1}{\beta_1\epsilon}+(1-\alpha_i)\frac{\beta_2}{\beta_1}}}{\left(\frac{\bar{M}}{v^\epsilon}\right)^{\frac{1}{\beta_1\epsilon}} (\alpha_i(1-t)p_i\mu_i)^{\frac{\beta_2}{\beta_1}}}$$

Suppose that when p_i increases, M_i decreases. Thus, the left-hand side of the above equation increases. At the same time, the right-hand side decreases; the equation will not hold. Thus, M_i has to increase.

Log-linearizing the equation:

$$\hat{M}_i = \frac{\beta_2 + \beta_1 S_{1,i}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} \hat{p}_i - \frac{1}{\frac{1}{\epsilon} + (1-\alpha_i)\left(\beta_2 + S_{1,i}\beta_1\right)} \hat{v} \tag{1.7}$$

where $S_{1,i} = \frac{\alpha_i t(1-t) p_i A_i M_i^{\alpha_i - 1}}{\alpha_i t(1-t) p_i A_i M_i^{\alpha_i - 1} + C_i} = \frac{B_i}{B_i + C_i}$, and B_i is the baseline per capita migrant contribution in taxes.

Similarly, I solve the percentage change in amenity as

$$\hat{A}_i = \frac{1}{S_{2,i}} \frac{S_{1,i} \cdot \frac{1}{\epsilon}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} \hat{p}_i + \frac{1}{S_{2,\mathrm{i}}} \frac{S_{1,\mathrm{i}} \cdot (1-\alpha_i)}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} \hat{v}$$

where $S_{2,i} = \frac{A_i}{A_i + C_i}$. Thus, when p_i increases, A_i increases.

Total regional GDP is $Y_i = p_i \mu_i M_i^{\alpha_i}$. Then the percentage change in GDP is

$$\hat{Y}_i = \hat{p}_i + \alpha_i \hat{M}_i = \frac{\beta_2 + \beta_1 S_{1,i} + \frac{1}{\epsilon}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} \hat{p}_i - \frac{\alpha_i}{\frac{1}{\epsilon} + (1-\alpha_i)\left(\beta_2 + S_{1,i}\beta_1\right)} \hat{v}.$$

where $S_{3,i} = \frac{M_i}{L_i + M_i}$. Thus, when p_i increases, G_i increases.

For each city, the impact of p_i on v is negligible, and the term with $\hat{v}$ in the above equations can be dropped. ■

Proposition 2:

In the strict Hukou System ($S = 0$), when there is a positive price shock in City i ($p_i \uparrow$), migrants will flow into the city ($M_i \uparrow$) and overall output in City i will increase ($Y_i \uparrow$). However, both the increase in number of migrants and the increase in output are smaller than in the Relaxed Hukou system.

Proof:

Since $S = 0$, $A_i = 0$, and

$$M_i = \left(\frac{\bar{M}}{v^\epsilon}(\alpha_i(1-t)\mu_i)^{\beta_2\epsilon}\right)^{\frac{1}{1+(1-\alpha_i)\beta_2\epsilon}} C_i^{\frac{\beta_1\epsilon}{1+(1-\alpha_i)\beta_2\epsilon}} p_i^{\frac{\beta_2\epsilon}{1+(1-\alpha_i)\beta_2\epsilon}}$$

And M_i is an increasing function of p_i. Log-linearize the equation:

$$\hat{M}_i = \frac{\beta_2}{(1-\alpha_i)\beta_2 + \frac{1}{\epsilon}}\hat{p}_i - \frac{1}{\frac{1}{\epsilon} + (1-\alpha_i)\beta_2}\hat{v}$$

Compared with Equation 1.7 in the proof for Proposition 1, the coefficient of $\hat{p}_i$ is smaller, meaning that the impact of price shocks on migrant flows is smaller in the strict Hukou system than in the relaxed Hukou system:

$$\frac{\beta_2 + \beta_1 S_{1,i}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} - \frac{\beta_2}{(1-\alpha_i)\beta_2 + \frac{1}{\epsilon}} = \frac{\beta_1 S_{1,i} \cdot \frac{1}{\epsilon}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} > 0$$

Since Y_i is an increasing function of M_i, given p_i, the overall output increase will also be smaller. ■

Proposition 3: In the symmetric case, when all cities are the same and the rural area has a very small wage, the overall output Y is an increasing function of the number of people who migrated. When there is an economy-wide positive price shock, both $Y_{S=1}$ and $Y_{S=0}$ will increase, and $Y_{S=1}$ will increase more. Thus, the central government is more likely to switch to the relaxed Hukou system.

Proof:

Suppose that all cities are the same in terms of economic fundamentals and prices shocks, the total output in cities is

$$\sum_i Y_i = N \cdot Y_i = N \cdot p_i \mu_i M_i^{\alpha_i},$$

which is a strictly increasing function in M_i. Using the national-level labor market clearing condition, $M_r = \bar{M} - N \cdot M_i$. The national total output is

$$Y = \sum_i Y_i + M_r w_r = N \cdot p_i \mu_i M_i^{\alpha_i} + (\bar{M} - N \cdot M_i) w_r.$$

The national total output will be a strictly increasing function in M_i when w_r is small enough. As shown in Proposition 2, when there is a positive price shock, the increase in number of migrants is bigger in the relaxed Hukou system than in the strict Hukou system; thus, overall output increase will also be bigger and the central government is more likely to switch to the relaxed Hukou system.

Additional Estimation Equations

The percentage increase in migrant inflow ($\hat{M}_i$) is

$$\hat{M}_i = \frac{\beta_2 + \beta_1 S_{1,i}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} \hat{p}_i \tag{1.8}$$

The total urban employment is the sum of local labor and migrant labor, and the percentage increase in total employment ($\hat{E}_i$) is

$$\hat{E}_i = \widehat{L_i + M_i} = S_{3,i} \hat{M}_i = \frac{(\beta_2 + \beta_1 S_{1,i}) S_{3,i}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} \hat{p}_i \tag{1.9}$$

where $S_{3,i} = \frac{M_i}{L_i + M_i}$.

The percentage increase in migrant wages ($\hat{w}_i$) is

$$\hat{w}_i = \hat{p}_i + (\alpha_i - 1)\hat{M}_i = \frac{\frac{1}{\epsilon}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} \hat{p}_i$$

And the percentage increase in local wages ($\hat{w}_i^L$) is

$$\hat{w}_i^L = \hat{p}_i + \alpha_i \hat{M}_i = \frac{\beta_2 + \beta_1 S_{1,i} + \frac{1}{\epsilon}}{(1-\alpha_i)(\beta_2 + \beta_1 S_i) + \frac{1}{\epsilon}} \hat{p}_i$$

When pooling the migrants with the local labor, the percentage change in mean wage for the total employment ($\hat{w}_i^T$) is

$$\hat{w}_i^T = (1 - S_{3,i})\hat{w}_i^L + S_{3,i}\hat{w}_i = \frac{(\beta_2 + \beta_1 S_{1,i})(1\text{-}S_{3,i}) + \frac{1}{\epsilon}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} \hat{p}_i \tag{1.10}$$

where $S_{3,i} = \frac{M_i}{L_i + M_i}$.

Define per capita GDP as $G_i = \frac{Y_i}{L_i + M_i}$. Then the percentage change in per capita GDP is

$$\hat{G}_i = \hat{Y}_i - S_{3,i}\hat{M}_i = \hat{p}_i + \alpha_2 \hat{M}_i - S_{3,i}\hat{M}_i = \frac{(\beta_2 + \beta_1 S_{1,i})(1 - S_{3,i}) + \frac{1}{\epsilon}}{(1-\alpha_i)(\beta_2 + \beta_1 S_{1,i}) + \frac{1}{\epsilon}} \hat{p}_i.$$

Additional Empirical Results

Competition between Prefectures in Regulation Changes

Prefecture i's regulation change and trade shock can not only affect its own regulation, but also affect other prefectures. The most direct measure of the intensity of competition is to focus on nearby prefectures. Table 1.28 Column (1) replicates the result in Table 1.2 Column (3). Column (2)-(4) consider the competition with other prefectures in the same province. Column (2) adds trade shocks, Column (3) adds regulation changes, and Column (4) controls for both. Column (5)-(7) repeat the exercise by considering the competition with five nearby prefectures.[64] Overall, I find no significant competition effect.

[64]The five nearby prefectures are the five closest prefectures by euclidian distance calculated from the longitude and the latitude.

Table 1.28: Competition between prefectures: by industrial composition, population size, and income similarity

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ log regulation score, 2001-2007		All other prefectures in the same prov.			5 closest prefectures		
Export tariff shock, own	1.09***	0.89**	0.90**	0.86**	1.06**	0.89*	0.88*
2001-2007	(0.39)	(0.38)	(0.39)	(0.39)	(0.41)	(0.44)	(0.45)
Export tariff shock, other pref.		0.59		0.39			
2001-2007		(0.77)		(0.57)			
Δ log regulation score, other pref.			0.29	0.28			
2001-2007			(0.19)	(0.18)			
Export tariff shock, nearby pref.					0.08		0.05
2001-2007					(0.14)		(0.13)
Δ log regulation score, nearby pref.						0.89	0.85
2001-2007						(0.65)	(0.63)
Observations	250	244	244	244	250	250	250
R-squared	0.12	0.09	0.10	0.10	0.12	0.13	0.13

Note: Standard errors are clustered at the province level. Mean (sd) of Δ reg regulation score, 2001-2007 is 0.77 (0.82), 1995-2001 is 0.06 (0.26). Mean value of own export tariff shock, 2001-2007 is 0.18 (0.15), 1995-2001 is 1.23 (0.40). Column (2) controls for the trade shock in all other prefectures in the same province. Column (3) controls for the regulation change in all other prefectures in the same province. Column (4) controls for both. Columns (5)-(7) repeat the exercise by controling for the variables in the 5 closest prefectures.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

In addition to focusing on nearby prefectures, I use three ways to measure a prefecture's exposure to competition with all other prefectures in terms of trade shocks and regulation changes. First, I measure the distance between prefectures using similarities in the industrial composition. The distance between prefecture o and prefecture d is the sum of squared differences in employment shares in each industry:

$$D^{ind}_{o,d} = \sum_j (EmpShare^{2001}_{o,j} - EmpShare^{2001}_{d,j})^2$$

where $EmpShare^{2001}_{i,j}$ is the employment share in industry j prefecture i in 2001, $i \in \{o, j\}$.

Second, I measure the distance between prefectures using similarities in the population size. The distance between prefecture o and prefecture d is the squared differences in log population in 2001:

$$D^{pop}_{o,d} = (\log(population^{2001}_o) - \log(population^{2001}_d))^2.$$

Third, I measure the distance between prefectures using similarities in per capita GDP:

$$D^{gdp}_{o,d} = (\log(\text{GDP p.c.}^{2001}_{o}) - \log(\text{GDP p.c.}^{2001}_{o}))^2.$$

I then construct the weight assigned to each destination prefecture d with respect to an origin prefecture o by taking the inverse of the distance measure as above, combined with the inverse of geographic distance:

$$w^S_{o,d} = \frac{1}{D^S_{o,d}} \cdot \frac{1}{D^{geodist}_{o,d}}$$

where $S \in \{ind, pop, gdp\}$, and $D^{geodist}_{o,d}$ is the travel time between prefecture o and prefecture d in 2001.[65]

The trade shock in competing prefectures of prefecture o is measured as

$$TS^S_o = \sum_d \frac{w^S_{o,d}}{\sum_{d'} w^S_{o,d'}} TS_d$$

and regulation change in in competing prefectures is measured as

$$R^S_o = \sum_d \frac{w^S_{o,d}}{\sum_{d'} w^S_{o,d'}} R_d$$

where $S \in \{ind, pop, gdp\}$.

I test whether the trade shocks and regulation changes in competing prefectures increase a prefecture's incentive to change its own regulation. In Table 1.29, I include a prefecture's own trade shocks and initial regulation score, and then add change in regulation score in competing prefecture in terms of industrial composition. The coefficient on other prefectures' regulation change is positive but not significant. In Column (2)-(4), I focus on competition by population size.

[65] The data on travel time is described in Section 1.6.4.

Column (2) includes trade shocks of competing prefectures, Column (3) includes regulation changes, and Column (4) includes both. None of the coefficients are significant. I do the same exercise in Column (5)-(7), focusing on competition by per capital GDP. I find positive and significant effects of both trade shocks and regulation changes: one unit change in the export tariff shock in competing prefectures has almost the same effect as one unit change in a prefecture's own export tariff shock (0.75-0.94 compared to 0.82-1.05); the elasticity between own regulation change and competing regulation change is 0.28-0.30.

Overall, I find that including competing prefectures' trade shocks and regulation changes does not affect the coefficient on own trade shocks a lot. However, there is some evidence that prefectures are competing in regulations with other prefectures that are similar in terms of income. This indicates that prefectures with similar income compete for the same pool of migrants, and there is a significant spillover effect in both trade shocks and regulation changes.

Table 1.29: Competition between prefectures: by industrial composition, population size, and income similarity

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ log regulation score, 2001-2007	By industries	By the size of population			By the size of GDP p.c.		
Export tariff shock, own	1.04***	1.09***	1.11***	1.12***	1.05***	0.84**	0.82**
2001-2007	(0.36)	(0.39)	(0.39)	(0.39)	(0.38)	(0.33)	(0.33)
Δ log regulation score, other, ind.	0.13						
2001-2007	(0.33)						
Export tariff shock, other, pop.		0.04		-0.05			
2001-2007		(0.35)		(0.37)			
Δ log regulation score, other, pop.			0.10	0.11			
2001-2007			(0.09)	(0.09)			
Export tariff shock, other, gdp					0.94***		0.75**
2001-2007					(0.31)		(0.33)
Δ log regulation score, other, gdp						0.30***	0.28***
2001-2007						(0.10)	(0.09)
Observations	250	250	250	250	250	250	250
R-squared	0.12	0.12	0.12	0.12	0.12	0.13	0.13

Note: Standard errors are clustered at the province level. Mean (sd) of Δ reg regulation score, 2001-2007 is 0.77 (0.82), 1995-2001 is 0.06 (0.26). Mean value of own export tariff shock, 2001-2007 is 0.18 (0.15), 1995-2001 is 1.23 (0.40). Column (1) controls for regulation changes in other prefectures, using distance in hours of travel and closeness of the industrial composition as weights. Column (2) controls for trade shocks in other prefectures, using distance in hours of travel and closeness of the population size as weights. Column (3) controls for regulation changes in other prefectures, using the same weights as in Column (2); Column (4) controls for both trade shocks and regulation changes in other prefectures. Columns (5)-(7) repeats the exercise in Columns (2)-(4) using distance in hours of travel and closeness of GDP p.c. as weights.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Migrant Supply

The potential supply of migrants can affect the responsiveness of migrant flow to trade shocks and regulation changes. For prefecture o, the distance weighted agricultural population is

$$\log(agrPOP)_o^{2001} = \sum_d \frac{w_{o,d}}{\sum_{d'} w_{o,d}} \log(agrPOP)_d^{2001}$$

where $w_{o,d} = \frac{1}{D_{geodist_{o,d}}}$, which is inverse of travel time between prefecture o and prefecture d in 2001, and $\log(agrPOP)_d^{2001}$ is the log agricultural population in prefecture d in 2001.

I investigate the impact of migrant supply on the equilibrium migrant flow in Table 1.30. Columns (1)-(3) uses the change in the migrant share of population as the outcome, and control for agricultural population measured as above. In addition, Column (2) adds the interaction between trade shocks and agricultural population, and Column (3) adds the interaction between the regulation change and agricultural population. I find no significant effect either on the agricultural population or on the interaction. Columns (4)-(6) investigate the effect on short-distance migrant flows, where migrants move within a prefecture. Thus, I use the agricultural population in the same prefecture. There is no significant interaction effect, but there is some evidence that places with a larger agricultural population to begin with do not move much either. One possible interpretation is that these prefectures have some fixed characteristics that lead to low mobility. Column (7)-(9) shows the effect on medium-distance migrant flows, where migrants move within a province across different prefectures. I use the agricultural population in the whole province as the measure for the potential pool of migrant supply. I find a positive interaction effect between the regulation change and migrant supply: a prefecture that is part of a province with a lot of agricultural population has a bigger inflow of migrant workers once the regulation is relaxed.

Table 1.30: Interaction effects of migrant supply and migrant demand

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	Δ migrant share of population			Δ log # of migrants, short-distance			Δ log # of migrants, medium distance		
Export tariff shock	6.08*	-15.33	6.45**	1.01***	4.36	1.03***	0.37**	-1.60	0.45**
2001-2007	(3.19)	(19.15)	(3.15)	(0.26)	(4.30)	(0.25)	(0.18)	(2.09)	(0.16)
Δ log regulation score	1.65***	1.68***	-4.86	0.24***	0.23***	-0.32	0.06	0.06	-1.53**
2001-2007	(0.41)	(0.41)	(4.40)	(0.05)	(0.05)	(0.76)	(0.04)	(0.04)	(0.66)
Log(agr. pop.)	-0.72	-1.04*	-1.08*	-0.42**	-0.35	-0.48**	-0.04	-0.06	-0.13***
2001	(0.50)	(0.55)	(0.56)	(0.19)	(0.25)	(0.22)	(0.05)	(0.05)	(0.05)
Export tariff shock		1.93			-0.25			0.12	
×Log(agr. pop.)		(1.67)			(0.32)			(0.13)	
Δ log regulation score			0.56			0.04			0.10**
×Log(agr. pop.)			(0.38)			(0.05)			(0.04)
Observations	249	249	249	240	240	240	250	250	250
R-squared	0.20	0.21	0.20	0.28	0.29	0.29	0.60	0.60	0.61
Mean (s.d.) of depent.	6.99 (5.78)			-0.81 (0.80)			1.60 (0.67)		

Note: Standard errors are clustered at the province level. Mean (sd) Δ log regulation score, 2001-2007 is 0.77 (0.82), mean (sd) export tariff shock is 0.18 (0.15).
All columns control for import and intermediate tariff shock, log total population and level of dependent variable in 2000. Columns (1)-(3) use the weighted average agricultu
population. Columns (4)-(6) use the agricultural population in the same prefecture. Columns (7)-(9) use the agricultural population in the same province.
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Hyperbolic-sine Transformation

Instead of log transformation, I use hyperbolic-sine transformation to allow for both positive and negative changes. The results are essentially the same as in Table 1.2.

Table 1.31: Bigger trade shocks, more migrant-friendly

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	Δ hyper regulation score, 2001-2007								Placebo, 1995-2001	
Export tariff shock	0.89**	1.56***	1.57***	1.57***	1.39**	1.32**	1.19**	1.04*	-0.02	
2001-2007	(0.41)	(0.51)	(0.51)	(0.51)	(0.56)	(0.54)	(0.57)	(0.56)	(0.03)	
Import tariff shock		-0.11**	-0.11**	-0.11**	-0.10	-0.10**	-0.13***	-0.10	-0.01*	
2001-2007		(0.04)	(0.04)	(0.04)	(0.07)	(0.05)	(0.04)	(0.08)	(0.00)	
Intermediate tariff shock		0.24*	0.24	0.24	0.06	0.29	0.13	0.09	0.00	
2001-2007		(0.14)	(0.14)	(0.14)	(0.23)	(0.18)	(0.16)	(0.23)	(0.03)	
Hyper regulation score			-0.08		-0.09	-0.10	-0.13	-0.13		
2001			(0.26)		(0.25)	(0.23)	(0.24)	(0.23)		
Δ hyper regulation score				-0.08						
1995-2001				(0.26)						
Export tariff shock					0.23			0.20		0.14
1995-2001					(0.24)			(0.25)		(0.10)
Import tariff shock					0.01			0.02		-0.02
1995-2001					(0.03)			(0.02)		(0.01)
Intermediate tariff shock					0.21			0.12		-0.03
1995-2001					(0.13)			(0.13)		(0.05)
Δ log wage						1.19**		0.84		
1995-2001						(0.44)		(0.52)		
Δ log GDP p.c.							0.75***	0.64***		
1995-2001							(0.17)	(0.20)		
Observations	250	250	250	250	237	237	237	237	238	237
R-squared	0.01	0.07	0.07	0.07	0.08	0.10	0.12	0.13	0.00	0.00

Note: Standard errors are clustered at the province level. Mean (sd) of Δ hyper regulation score, 2001-2007 is 1.01(1.08), 1995-2001 is 0.05 (0.40).
Mean value of export tariff shock, 2001-2007 is 0.18 (0.15), 1995-2001 is 1.23 (0.40).
*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Prices

I show that nominal wage and GDP increased where there is an inflow of migrant workers. However, if prices rise too much, then real income may not rise as much. There is no price index on the prefecture level, thus I measure price change on the province level as the product of consumer price indexes (CPI) from 2002 to 2007, since the CPI takes the previous year as the base year. The CPI on the province level is from the website of the Bureau of Statistics of China.

I replicate the results of Table 1.5 in Table 1.32. The results are largely unchanged, and the price effect seems not to affect the relationship between trade shocks, regulation changes, and welfare.

Table 1.32: Trade shocks, regulation changes, and welfare measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Dependent variable	Δ log wage, 2001-2007					Δ log GDP p.c., 2001-2007					Δ log total urban emp., 2001-2007				
Export tariff shock	0.15**	0.13**			0.12*	0.54***	0.51***			0.49***	0.37*	0.28			0.27
2001-2007	(0.07)	(0.06)			(0.06)	(0.13)	(0.14)			(0.13)	(0.20)	(0.18)			(0.17)
Δlog regulation score			0.03***	0.03***	0.03***			0.09***	0.08***	0.07***			0.09**	0.05	0.04
2001-2007			(0.01)	(0.01)	(0.01)			(0.02)	(0.02)	(0.02)			(0.03)	(0.03)	(0.03)
Δprice	-0.03	-0.03	-0.02	-0.03	-0.04*	-0.07	-0.10**	-0.08**	-0.12***	-0.11**	-0.00	0.05	-0.14**	0.01	0.04
2001-2007	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)	(0.06)	(0.05)	(0.05)
Controls (lagged)		X		X	X		X		X	X		X		X	X
Observations	249	236	249	236	236	250	237	250	237	237	250	237	250	237	237
R-squared	0.22	0.33	0.14	0.31	0.33	0.32	0.38	0.27	0.38	0.41	0.29	0.36	0.26	0.32	0.39
Mean (s.d.) of depent.	0.82 (0.14)					0.87 (0.27)					0.32 (0.39)				

Note: Standard errors are clustered at the province level. Mean (sd) Δ log regulation score, 2001-2007 is 0.77 (0.82), mean (sd) export tariff shock is 0.18 (0.15). All columns control for import and intermediate tariff shock, log total population and level of dependent variable in 2000. Columns (2)(4)(5)(7)(9)(10)(12)(14)(15) also control for lagged trade shocks and lagged wage and gdp growth rate, 1995-2001, as in Table 2 Column (8) .

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Additional Theoretical Results

An alternative way of setting up the model is to assume that the local government objective function is to maximize the weighted average of utilities of locals and migrants, while keeping the budget balanced. The local government has to choose two parameters: B which is government expenditure and it is equal to the value of public goods provided by the government, and r which is the share of public goods accessible by migrant workers, and $r \in [0, 1]$. There is congestion in the consumption of public goods: the public goods have to be divided among the locals and the migrants, with $\frac{B}{L+rM}$ allocated to a local, and $\frac{rB}{L+rM}$ allocated to migrants.[66]

The utility of a local worker (U_L) and the utility of a migrant worker (U_M) are

$$U_L = \left(\frac{B}{L + rM}\right)^{\beta_1} w_L^{\beta_2},$$

$$U_M = \left(\frac{rB}{L + rM}\right)^{\beta_1} w_M^{\beta_2},$$

[66]The total value of public goods is $\frac{B}{L+rM} \cdot L + \frac{rB}{L+rM} \cdot M = B$.

where L is the number of local workers, M is the number of migrant workers, w_L and w_M are wages for local workers and migrant workers, respectively.

The government assign a weight of δ to the utility of local workers and a weight of $1-\delta$ to the utility of migrant workers, and $\delta \in (0,1]$. Thus, the problem of the government is as follows:

$$\max_{B,r} \quad \left[\left(\frac{B}{L+rM} \right)^{\beta_1} w_L^{\beta_2} \right]^{\delta} \left[\left(\frac{rB}{L+rM} \right)^{\beta_1} w_M^{\beta_2} \right]^{1-\delta}$$

where

$$\text{s.t. } tR = B$$

Output is produced with a Cobb-Douglas production function where p is the price of output, A is the productivity, and local labor and migrant labor are inputs ($\alpha_1 > 0, \alpha_2 > 0$, and $\alpha_1 + \alpha_2 < 1$). To capture the potential effect that the amenity level of migrant workers might affect productivity, I add a term with r^{γ}, where $\gamma \in (-\infty, +\infty)$.

$$R = pAr^{\gamma} L^{\alpha_1} M^{\alpha_2}$$

Thus, wages of local workers and migrant workers can be solved:

$$w_L = \alpha_1 (1-t) pAr^{\gamma} L^{\alpha_1 - 1} M^{\alpha_2},$$

$$w_M = \alpha_2 (1-t) pAr^{\gamma} L^{\alpha_1} M^{\alpha_2 - 1}.$$

Again, I assume that local workers are immobile, and migrant workers need to make a migration choice. Worker l will choose to live in the region n that maximizes their utility, so $n = \text{argmax}_{n'} v_{n'} \epsilon_{n'}^{l}$. There is an idiosyncratic taste draw ϵ_n^l for each worker and area, and the draw is i.i.d. across workers and areas from a Fréchet distribution, $Pr(\epsilon_n^l \leq x) = e^{-x^{-\epsilon}}$, with $\epsilon > 1$. Thus, the number of migrant workers who live in area n is

$$M = \left[\frac{\left(\frac{rB}{L+rM}\right)^{\beta_1} w_M^{\beta_2}}{v} \right]^{\epsilon},$$

normalizing the total supply of migrants to be 1.

With this alternative model setup, the optimal policy r^* can be solved as a function of p, and the coefficient on $log(p)$ is negative, indicating that a positive price shock to a region leads to a smaller share of amenity for migrants. This prediction is opposite to the prediction in the model with local governments maximizing net fiscal profits, and does not fit the empirical findings either.

$$\log(r^*) = \left[\frac{-1}{1 - \alpha_2 + \frac{1}{\epsilon}} \right] \log(p) + const$$

Mediation Analysis

Figure 1.24: Identification of the mediation effect

(a) One mediator (b) Two mediators (c) Two mediators, interaction

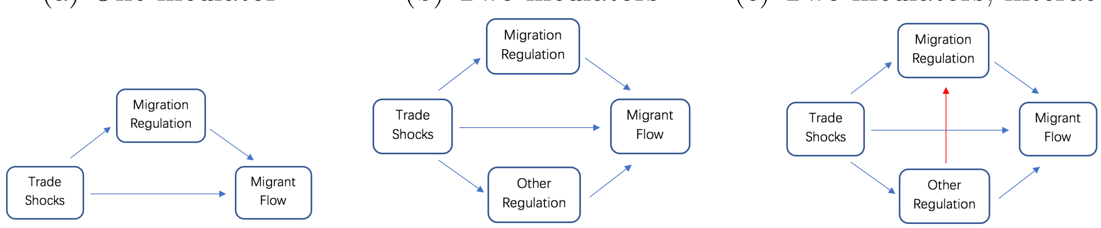
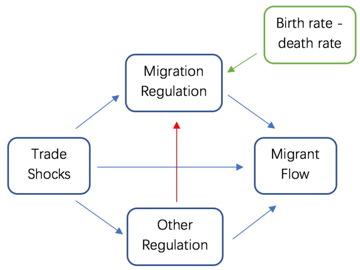
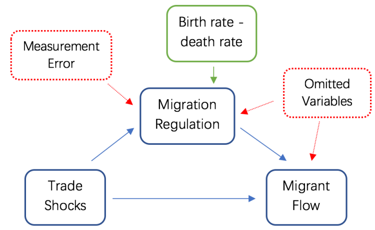


Figure 1.25: Using the natural population growth rate as the IV for regulation changes

(a) Two mediator, interaction



(b) Other identification challenges



CHAPTER 2

Was Entry into the WTO Worth it: Environmental Consequences of Trade Liberalization

2.1 Introduction

China's accession into the World Trade Organization (WTO) in 2001 represents one of the biggest trade liberalization events during the last two decades. Looking back in 2011, China achieved great economic growth since it joined the WTO, with an annual growth rate of around 10%. In terms of GDP, China grew from the 6th largest economy in the world in 2001 to the 2nd largest in 2011; moreover, it became the 1st largest merchandise exporter, 2nd largest merchandise importer, 1st largest investor and first destination for foreign direct investment (FDI) among the developing countries. The growth in exports, imports and overall GDP has benefited the quality of life for Chinese citizens. Both urban disposable income and rural income per capita tripled from 2001 to 2011: urban income rose from 7,000 RMB to 22,000 RMB, while rural income rose from 1,800 RMB to 7,000 RMB.[1] In a speech given at a forum in Beijing commemorating the 10th anniversary, Director-General Pascal Lamy commented: "WTO membership has served as a stabilizer and accelerator in China's economic take-off."[2]

However, this economic prosperity came along with associated environmental

[1] https://www.wto.org/english/thewto_e/acc_e/s7lu_e.pdf

[2] https://www.wto.org/english/news_e/sppl_e/sppl211_e.htm

costs. During the period of 2001-2011, CO2 emissions increased from 2.7 to 6.7 metric tons per capita;[3] total volume of industrial waste gas emissions (in 100 million cubic meters) increased from 138,145 to 674,509.[4] Waste gas emissions led to a severe ambient air pollution problem. According to the World Bank, the mean annual exposure of PM 2.5 in China was 44.2 $\mu g/m^3$ in 2000, and rose to 54.1 $\mu g/m^3$ in 2011,[5] while the WHO recommended threshold for a healthy environment is 10 $\mu g/m^3$.[6] Beginning in 2012, hazy weather in Beijing stimulated a heated discussion about the environmental situation in China. News reports featuring the gloomy weather and environmental regulation issues of China appeared in both Chinese and international media.[7] In early 2015, a documentary movie called *Under the Dome* also promoted discussions all over China about the cause, the results and the possible solutions to the country's urgent environmental problem.

The WTO's primary goal is to promote trade negotiations between economies, reduce trade barriers and create a competitive and transparent international trade environment. Thus, the focus of tariff negotiations is mainly economic concerns rather than environmental issues. However, in the case of China, which was already a highly polluted country before the WTO accession by WHO standards, the effect of trade liberalization on the environment can not be overlooked. If the tariff regime incentivizes the production of goods with large environmental and health costs in China, what are the true gains from trade in terms of consumer welfare once the costs are accounted for?

On the other hand, when evaluating the effect of environment on health, most

---

[3] http://data.worldbank.org/country/china?view=chart

[4] *Chinese Environmental Statistics Yearbook*, 2012, Bureau of Statistics of China.

[5] http://databank.worldbank.org/data/reports.aspx?source=2&series=EN.ATM.PM25.MC.M3&country=

[6] http://apps.who.int/iris/bitstream/10665/69477/1/WHO_SDE_PHE_OEH_06.02_eng.pdf

[7] *The Economist*: "The East is grey", August 10, 2013. *The Wall Street Journal*: "Beijing Choking on Air Pollution", Feb 24, 2014.

discussions do not pay as much attention to the role of economic or income growth. There have been observations that people in developing countries have a relatively low marginal willingness to pay for environmental protection, and one possible reason is that the marginal utility of income or consumption is high in such cases. (Greenstone and Jack (2015)) Also, in terms of health outcomes, it is not that people in developing countries do not suffer from pollution, but that the health effect related to income can be bigger. Thus, it is very important to understand the cost of environmental protection in the context of low income countries, and estimate the joint effect of income and pollution on health.

This paper aims to evaluate China's gains from the WTO accession by taking into account both the income growth benefits and the associated pollution and health costs. By comparing cities in China with differential income and pollution changes, I investigate how the costs and benefits are ultimately distributed across areas, thus assessing the extent to which welfare inequality has changed as a result of trade liberalization. Although China experienced the trade liberalization as a whole, the income and environmental effects differed across cities depending on a city's initial industrial composition. Cities grew differentially in terms of exports based on which industries they specialized in and which industries grew most due to the overall trade shock. For cities with a similar overall export value growth, the one specializing in more pollution-intensive industries might suffer more from increased pollution.

There are several challenges when studying the environmental consequences of trade liberalization. The first challenge is that in developing countries such as China, reliable environmental data is not readily available. In 2000, only 41 major Chinese cities (among a total of 340 cities) detailed air quality index reports on the website of the Chinese Environment Protection Bureau; 85 cities in 2010, and only in 2014 the number increased to 289. In addition, it is purported that local Chinese governments manipulate the reading of monitors out of career concerns (see Ghanem and Zhang (2014)). Thus, for the period that I am studying, 2000 as pre-WTO and 2005 as post-WTO, there is no good city-level air quality data available. As an alternative, I use the satellite data by NASA on aerosol optical depth to infer ground air pollution concentration levels by city, which is a commonly accepted method in the literature. (Foster et al. (2009); Gutierrez (2010); Jia (2017b))

The second challenge is that export growth of cities can be correlated with many other city-level characteristics, especially local government policies. For example, more affluent cities might both discourage the development of high-polluting industries, and invest more in the public health system. Thus, the effect of increased exports on health may be confounded by other non-trade-related factors. In order to solve this problem, I use regional tariff shocks as an instrument of export growth. After China entered the WTO, other WTO countries had to offer the most-favorable-nation (MFN) tariff for Chinese export goods, and this overall tariff cut is not likely to be correlated with local economic conditions and government policies.

First, I use plausibly exogenous tariff reductions caused by WTO accession, variations in industrial compositions across cities and variations in pollution intensities across industries to measure the income shock and the pollution shock by city. In order to generate overall and pollution-related regional tariff shocks, I use a simple specific-factor trade model to guide the empirical measure. To address quality concerns about data reported by the Chinese government, I use satellite data to measure economic activities and pollution intensities by city. Lastly, I perform a back-of-the-envelope calculation to derive an overall welfare measure of trade liberalization by accounting for both the overall health effects and consumption effects.

This paper contributes to several strands of literature related to trade and health. First, there has been a significant literature regarding the effect of international trade on the environment. Antweiler et al. (2001) decomposes the trade effect on pollution into scale, technique, and composition effects and then tests the theory using cross-country data on sulfur dioxide concentrations. Copeland and Taylor (2004) builds a unified framework of economic growth, international trade and environmental consequences. However, both theoretical and empirical work in this literature focuses on cross-country analysis without an overall welfare evaluation. Second, the paper is closely related to the literature on the distributional effect of trade liberalization. Topalova (2010) studies the regional effect of trade liberalization in terms of poverty alleviation in India. Kovak (2013) is the first paper to construct a theoretically consistent regional trade shock measure, and studies the effect of trade liberalization on local wages. Autor et al. (2013) studies how the shock of Chinese exports affects local labor markets in the United States. However, most studies in this area of literature focus solely on economic outcomes. In addition, my paper is closely related to Becker et al. (2005) about measuring welfare by taking into account quantity and quality of life. Instead of using a cross-country comparison, I conduct the analysis on the city level, eliminating more uncontrolled regional heterogeneity. Finally, my research contributes to the literature studying the effect of pollution on health. Chay and Greenstone (2003) uses the economic downturn in the United States at the beginning of 1980s to acquire exogenous changes in pollution levels, and studies the effect of total particulate matters on the infant mortality rate. Arceo et al. (2015) uses daily pollution and infant mortality data in the Mexico City to provide evidence of health costs due to pollution in a developing country context. Chen et al. (2013) uses the regression discontinuity generated by the collective winter heating system in China to study the effect of sustained air pollution on life expectancy. In the literature, by omitting the income effect of industrial activities, the effect of pollution on health may not be correctly measured.

The paper is organized as follows. Section 2 presents a description of data and measurement as well as a summary of statistics. Section 3 introduces background information on China's WTO accession and provides preliminary evidence about the trade induced income, pollution and health effects. Section 4 presents a theoretical model to show how tariff cuts will translate into production increases, and how different industries and cities benefit. Section 5 details the main empirical estimations, while Section 6 includes additional discussions and various robustness checks. The last section concludes.

2.2 Data and Measurement

2.2.1 Data

2.2.1.1 Health data

The major health measure used in the paper is the mortality rate. I use city-level populations and the number of deaths by age groups from the 2000 China Population Census and 2005 1% Population Survey. The 2000 Census covers the entire population of mainland China in 340 cities of 31 provinces, surveyed during the period from 1999 to 2000. The 2005 Population Survey was conducted using stratified multistage clustered probability sampling methods during the period from 2004 to 2005. This survey also covered all cities in mainland China, but the city-level data is not available for all provinces, as I was only able to obtain city-level measures for 136 cities in 16 provinces. The infant mortality rate (mortality rate in age 0) is used in the robustness check instead of the main regression, since on the one hand side, I would like to conduct the welfare analysis on the entire population, and on the other hand, there are more reporting errors (under-reporting) for infant deaths than overall deaths in the census.

I also use the aggregate mortality rate by causes of death from the Disease Surveillance Point (DSP) on the city level from 1991 to 2000, to check for mortality trends prior to China's WTO accession. The DSP system started in 1989 as a basic health monitoring system by the Chinese Center of Disease Control, with 145 surveillance points in 31 provinces chosen by multi-stage cluster population probability sampling to form a representative national population sample. Each surveillance point monitors a population of about 30 to 100 thousand, and records a total of 50 thousand death cases, 100 thousand birth cases and many contagious disease epidemic cases each year. For death cases, basic demographic information and cause of death is recorded as an ICD-9 code.[8]

2.2.1.2 Trade and production data

The tariff information, delineated by a 2-digit Harmonized System (HS) Code is obtained from the World Bank's World Integrated Trade Solution (WITS). Weighted effectively applied import tariffs on 96 Chinese products are reported by the importing countries, from 2000 to 2010. In the next subsection, I discuss in greater details about how I construct the tariff by product, industry and city.

Firm-level data in 2000 and 2005 is obtained from the Database of Chinese Industrial Enterprises Survey (IES) collected by the National Bureau of Statistics of China. The database contains the information of all state-owned industrial enterprises and privately-owned enterprises whose sales revenues exceed 5 million RMB annually. The information includes the basic information (e.g. address, legal entity, capital ownership and industry code), financial information (e.g. total sales, exports, assets and taxes) and product information(e.g. primary product and secondary product). In this paper, the information used includes the city code, industry code, total sales revenue, total export value, total cost of sales, total capital, fixed capital and wage bills. There were 162,883 firms in 2000 and 268,330 in 2005.

[8] www.phsciencedata.cn/Share/en/data.jsp?id=7253a104-63ac-40f7-a0ac-b04c1096ae52a&show=1

2.2.1.3 Pollution data by industry and by city

One important element in this paper is the pollution intensity for each industry. First, I will use a simple rule to classify industries into two categories: polluting vs. non-polluting industries. Second, the 2003 Chinese Environmental Yearbook details the national industry pollution intensity by pollutant type (soot, water, waste gas, SO_2). This is the earliest available report on this measure.

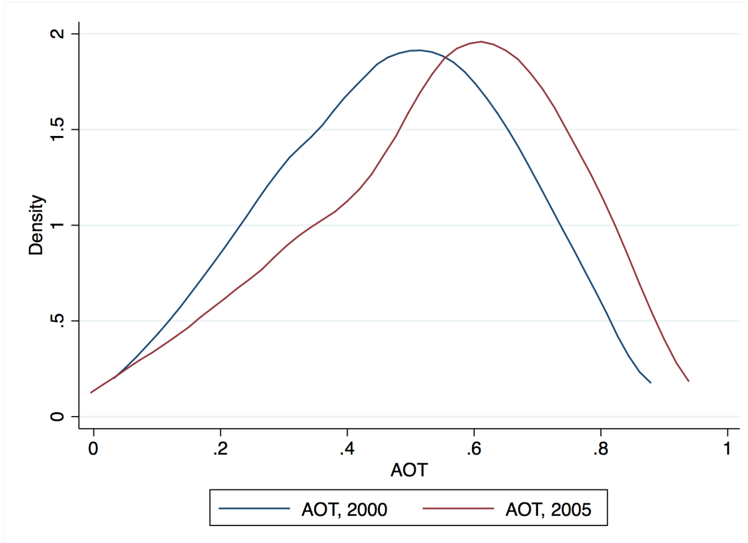
The ambient air pollution level for each city is calculated from NASA's MODIS data set. "The MODIS Aerosol Product monitors the ambient aerosol optical thickness (AOT) over the ocean globally and over a portion of continents."[9] Aerosols are tiny solid and liquid particles suspended in the atmosphere, and can either come from natural sources, such as windblown dust, sea salts and volcanic ash, or from human activities such as smoke from fire and pollution from factories. In terms of ambient air pollutants, aerosol levels are most closely related to particulates. Wang and Christopher (2003) shows that the correlation between AOT and $PM_{2.5}$ is 0.7 in Jefferson County, Alabama. The data used in this paper is NASA MODIS Daily Level 2 data, with a spatial resolution of 10 by 10 kilometers pixel.[10] Due to lack of credible measures of ambient air pollution levels from ground monitoring stations in China, this satellite data has been used as a proxy for air pollution in several papers (e.g., Jia (2017b); Long et al. (2014)).

Figure 2.1 plots the distribution of AOT levels across 106 Chinese cities used later in the regression. The blue line is the distribution in 2000, while the red line is the distribution in 2005. We can see that the distribution shifted to the right over the time period, indicating that air pollution measures as per AOT levels worsened, consistent with the World Bank report mentioned previously.

[9] http://modis-atmos.gsfc.nasa.gov/MOD04_L2/index.html

[10] http://neo.sci.gsfc.nasa.gov/view.php?datasetId=MODAL2_M_AER_OD&year=2000

Figure 2.1: Pollution level distribution across cities, 2000 and 2005



Note: Data for aerosol optical thickness (AOT) is taken from NASA satellite information. Distributions are across 106 cities used in the final regressions. Kernel density is estimated with bandwidth 0.11.

2.2.1.4 Other city-level data

For city-level characteristics, variables including GDP, population, total employment, and population density are taken from the City Statistics Yearbook in 2000 and 2005.

I also use night-time light intensity data from NASA as a proxy for economic activities following Henderson et al. (2012), to address quality concerns about government-reported GDP.[11]

2.2.2 Measurement

2.2.2.1 Mortality rate

The major outcome variable used is the standardized total mortality. The formula for calculating the age-standardized total mortality rate in city i is as follows:

[11] http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html

$$MR_i^{std} = \sum_{g=1}^{G} MR_{gi} PS_g$$

where MR_{gi} is the mortality rate in age group g in county i, and PS_g is the share of population in age group g in the entire country. This measure has been used in Hanlon and Tian (2015) and Beach and Hanlon (2017). Chen et al. (2013) uses life expectancy, which is a similar measure. In order to have comparable mortality rates over time and eliminate the age structure variation, I use the age structure in 2000 to standardize the mortality rates in 2005.

The infant mortality is the most common health measure in the environmental and health literature (Chay and Greenstone (2003), Almond and Currie (2011)), as infant's exposure to detrimental environmental conditions are relatively short and not confounded with factors that contribute to health conditions. However, the infant mortality in China's population census is relatively poorly measured, especially in rural areas. Thus, the effect of trade on the infant mortality rate will be postponed until the robustness check is conducted.

2.2.2.2 Tariff

The second important problem is how to construct the weighted tariff by product, by industry and by city.

The tariff is reported by WITS, with 96 product categories, and by WTO countries. For example, in 2000, China exported to 68 countries, and in 2005, the number of partner countries increased to 106. For each product, I generate the average tariff faced by Chinese exporters by weighting the tariff of each importing country by the share of exports out of the total export value of the product.

For example, for aluminum, the total export value of China in 2005 is 100 units, and Country A imports 40 units and Country B imports 60 units. The import tariff of Country A for aluminum from China is 2.3 (%) and that of Country B is 1.3

(%). Then, the tariff faced by Chinese aluminum exporters is 2.3*40%+1.3*60%.

However, the import value and import tariff can be correlated. For example, it is very likely that countries which import a significant amount of goods will impose lower import tariffs. Because later I will argue that reduced tariffs will induce higher export volumes on the supply side without affecting the demand side, this correlation is troublesome. Thus, I use the trade volumes in the baseline year (2000) as weights. Using the same example, if out of the 100 units, if Country A imports 40 units and Country B 60 units in 2005, and 20 and 80 units in 2000, I will calculate the tariff in 2005 as 2.3*20%+1.3*80% instead of 2.3*40%+1.3*60%.

After constructing the weighted average of tariffs faced by Chinese exporters by product category, I construct the weighted average of tariffs by industry. The problem here is that trade product categories are classified using a 2-digit HS code, which focuses mainly on product characteristics, while industry classifications are done using 4-digit U.S. Standardized Industrial Classification (SIC) codes, which take into account the production procedure. I use the data and code provided in Pierce and Schott (2009) to match HS codes with SIC codes. The second problem is that the Chinese Industrial Enterprise Survey uses a Chinese version of SIC code, which is slightly different from the U.S. SIC code. Thus, I perform a manual matching between the two sets of industry codes. Details about the matching is shown in Appendix A. After doing the matching between product and industry classifications, I construct the tariffs faced by an industry in the following way: supposing that Industry 1 has three HS categories, the tariff faced by Industry 1 is the mean of the tariffs in the three HS categories.

After constructing the weighted average of tariffs faced by Chinese exporters by product categories, I construct the weighted average of tariffs by city. Tariff cuts can have two effects on the production: the price effect and the cost effect. On the one hand side, since the tariffs faced by exports are reduced, the effective price of the good increases. In other words, in order to sell products to other countries at the same price as before, now Chinese exporters can charge higher prices and pay lower tariffs. On the other hand, without perfect factor mobility across regions, and in the short-run, if we assume that the total amount of production factors is constant in a region, then the price increase of goods will lead to higher prices of input factors. In order to highlight these channels, in Section 4 I will use a specific-factor model to show how to construct the regional tariff shocks properly.

2.2.2.3 Division of polluting vs. non-polluting industries

Following Hanlon and Tian (2015), I classify the industries into polluting and non-polluting ones. This classification is made according to an official document of the Chinese government. In the document issued by the Chinese Environment Protection Bureau ([2003] No.10) named *About Inspection of Environmental Qualification of Companies that are Applying for Listing and Refinancing*, the heavy-polluting industries are: metallurgical, chemical, petrochemical, coal, thermal power, building materials, paper, brewing, pharmaceutical, fermentation, textile, leather and mining.

2.2.2.4 National-level industry pollution density

Using industry-level pollutant emission and industrial output (values from IES), I then calculate the pollution intensity of each industry, in terms of waste gas, SO_2, soot and water emission per dollar of output. As can be seen in Table 2.1, most of industries ranking high on the pollution intensity list are also defined as polluting industry in my classification. Electric, gas and sanitary services rank at the top in the pollution intensity list, but they are not exporting industries. I use both the division of polluting vs. non-polluting industries and the industry pollution intensity measures in the empirical part.

Table 2.1: Industry rank of pollutant production per dollar of output, 2003.

Industry name	SO2	Soot	Waste gas	Water	Polluting industry?
Electric, Gas and Sanitary Services	1	1	4	5	No
Stone, Clay, Glass and Concrete Products	2	2	21	10	Yes
Paper and Allied Products	3	5	16	1	Yes
Petroleum Refining and Related Industries	4	3	23	9	Yes
Coal Mining	5	7	7	3	Yes
Primary Metal Industries	6	4	9	11	Yes
Mining and Quarrying of Nonmetallic Min.	7	16	3	8	Yes
Metal Mining	8	13	2	4	Yes
Lumber and Wood Products, Except Furn.	9	11	6	12	No
Textile Mill Products	10	9	14	6	Yes
Chemicals and Allied Products	11	6	15	7	Yes
Food and Kindred Products	12	8	20	15	No
Industrial and Commercial Mach. and Comp.	13	12	19	19	No
Oil and Gas Extraction	14	17	12	18	Yes
Tobacco Products	15	19	13	20	No
Communications	16	25	10	2	No
Transportation Equipment	17	10	22	17	No
Fabricated Metal Products	18	14	5	16	No
Leather and Leather Products	19	20	17	13	Yes
Instruments	20	22	8	14	No
Rubber and Miscellaneous Plastic Products	21	18	11	23	Yes
Furniture and Fixtures	22	23	25	24	No
Printing, Publishing and Allied Industries	23	24	1	22	No
Apparel, Finished Products from Fabrics	24	21	18	21	No
Electronic, Electrical Equip. & Comp.	25	15	24	25	No

Note: Pollutant emissions by industry are from the Chinese Environmental Statistics Yearbook. Sales revenues are from Chinese Industrial Enterprise Survey (IES). The list here is by 2-digit Standardized Industrial Classification (SIC) codes. The rank is calculated using 2003 data, since this is the earliest year available with detailed pollutant information and industry classifications.

2.3 Background

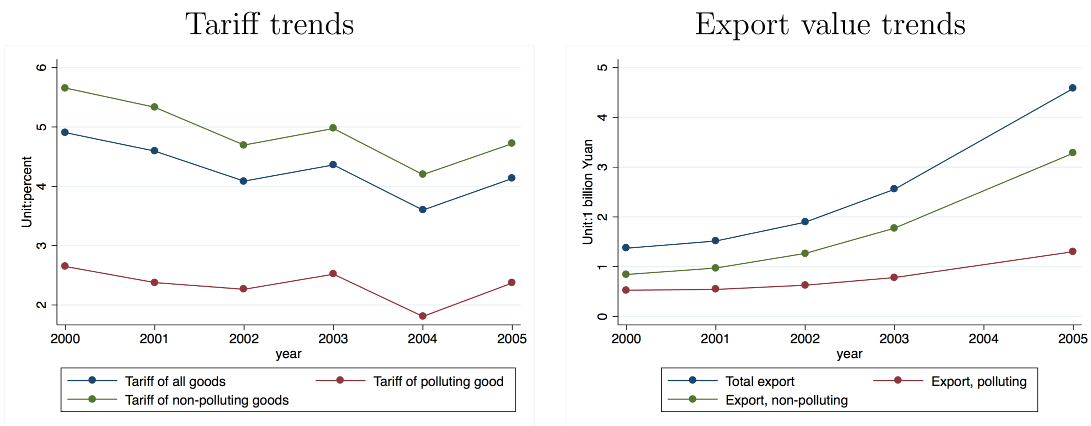
The World Trade Organization (WTO) is an international organization that promotes and runs trade negotiations between countries, intending to create a competitive, transparent and predictable international trade environment. The WTO was created in January 1, 1995 as part of the Uruguay Round negotiations, following its predecessor, the General Agreement on Tariffs and Trade (GATT). The GATT was originally signed by 23 nations in Geneva on October 30, 1947, and the WTO has 161 member countries as of April 26, 2015.

China was among the 23 original signatories of the GATT in 1948, but pulled out of the agreement after 1949 due to domestic political and economical situations. In 1986, China notified the GATT of its wish to resume its status as a GATT contracting party, and started to work on a series of economic reforms to transform the economy into a more market-oriented one. On November 10, 2001, China's accession was approved in the 4th Ministerial Conference in Doha, and China became the 143rd member of the WTO.

Upon accession, China committed to undertaking a series of important commitments to open and liberalize its regime to better integrate into the word economy, providing a more predictable environment for trade and foreign investments according to the WTO rules. Specifically, China eliminated dual pricing practices, differences in the treatment of goods produced for domestic sales and the goods produced for exports; reduced price controls intended to protect domestic producers; eliminated import quotas and bound import tariffs; and did not maintain or introduce any export subsidies on agricultural products.

At the same time, China would have Most-Favorable-Nation (MFN) status with all other WTO member countries, which meant that there would be upper bounds on the import tariff that other countries could impose on Chinese goods, and those tariffs should be equalized among all MFN countries. Also, other WTO member countries needed to gradually phase out import quotas on Chinese goods. During China's phase-in period into the WTO from 2001 to 2005, the average tariff faced by Chinese exporters decreased from 4.9% in 2000 to 4.1% in 2005, with an annual decrease rate of 0.15 percentage point. In response, China experienced substantial export growth, rising from 0.25 billion dollars in 2000 to 0.75 billion dollars in 2005 (see Figure 2.2).

Figure 2.2: Total Chinese export value and average tariff, 2000-2005

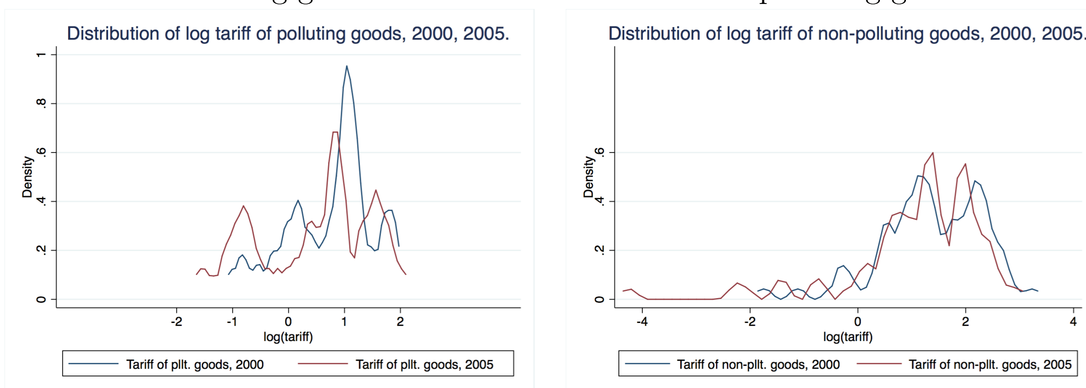


Note: Tariff data and export values are taken from World Integrated Trade Solution (WITS) for WTO member countries. Each data point on the left graph represents the average tariff faced by Chinese exporters in a year. For a year, the tariff of all goods is the simple mean across 96 product categories. The tariff for polluting (non-polluting) goods is the simple mean across 24 (72) non-polluting goods. Polluting and non-polluting goods are defined according to their corresponding industry characteristics. Each data point on the right graph represents the total export value in a year. For division of pollution vs. non-polluting industries, see the data section.

Export growth and tariff cuts varied across industries and products. As shown in Figure 2.2, on average, import tariffs were higher in non-polluting industries than in polluting industries, and export value levels were also higher for non-polluting industry than for polluting ones. In terms of percentage changes, if we take a look at the distribution of the log of tariffs, we can see that actually the distribution of the log of tariffs on polluting goods shifts further to the left than the log of tariffs for non-polluting goods. (Figure 2.3). When interacting the tariff reductions with initial city-level industrial compositions, cities would experience differential sizes of regional tariff shocks, both in the polluting and the non-polluting industries.

Figure 2.3: Density of log of import tariffs on Chinese goods, 2000 and 2005, polluting and non-polluting goods

Polluting goods | Non-polluting goods

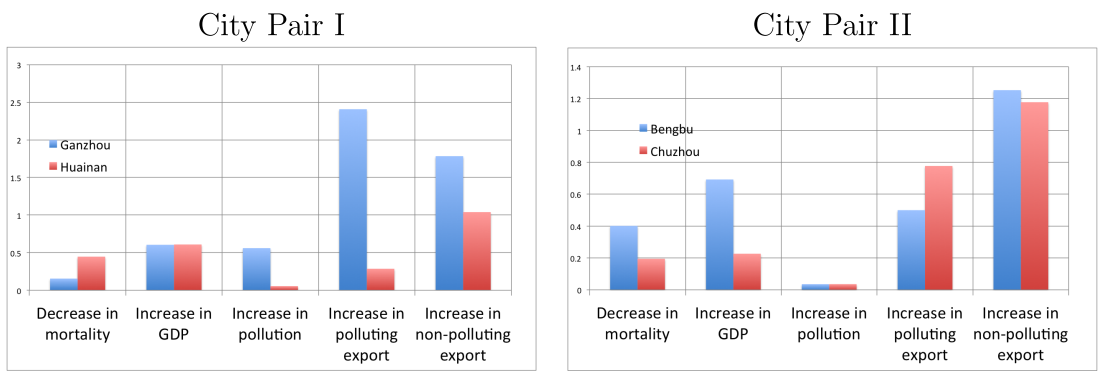


Note: Tariff data is from World Integrated Trade Solution (WITS) for WTO countries. The figure on the left is the density of the log of tariffs for polluting goods, and the figure on the right is the density of the log of tariffs for non-polluting goods. There are 96 product categories in total, and 24 are defined as polluting goods. The kernel density is estimated with a bandwidth 0.11.

Making use of differential regional tariff shocks and industry pollution intensity levels, I aim to disentangle the income effect and pollution effect of trade liberalization, and see how the welfare gains of trade is distributed across cities. In order to highlight the intuition of the empirical work, we can take a look at two pairs of cities in China. In Figure 2.4, City Pair I, *Ganzhou* and *Huainan* experienced similar GDP growth, while pollution in *Ganzhou* increased much more. As a result, *Huainan* was able to decrease its mortality rate more. If we take a closer look at the industrial composition, we can see that *Ganzhou* experienced a much higher growth in polluting industries than in non-polluting industries, while *Huainan* had the opposite trends. In City Pair II, *Bengbu* and *Chuzhou*, the cities experienced similar industrial growth in polluting and non-polluting industries, and a similar decrease in pollution levels, while *Bengbu* grew more in terms of GDP. As a result, *Bengbu* experienced a higher decrease in its mortality rate.

These two examples highlight the importance of working with both the income and the pollution effect when it comes to the discussion of trade liberalization effect on health, and how industrial composition maps to differential income and pollution growth. Combined with exogenous shocks of tariff reductions after the WTO accession, I aim to identify the environmental and health consequences of trade liberalization.

Figure 2.4: Exports, GDP, pollution and mortality rates: two pairs of cities, 2000 to 2005 changes



Note: Export information is from Chinese Industrial Enterprise Survey (IES). Pollution data is from NASA aerosol optical thickness (AOT) satellite images. GDP is from city statistics yearbooks, and mortality rates are from population census. All changes are expressed in log terms. Ganzhou is a city in Jiangxi province, and Huainan, Bengbu and Chuzhou are cities in Anhui province.

2.4 Theoretical model

2.4.1 Motivating theoretical trade model

When import tariffs (which itself is a sales tax) on Chinese goods decrease, *ceteris paribus*, Chinese exporters will receive higher prices. An increase in the price of goods will affect the allocation of production factors across industries within a city, thus affecting industry output levels. To formalize this idea, I use a specific-factor model as in Jones (1975), Kovak (2013) and Hanlon and Miscio (2017).

The economy in this model is a Ricardo-Viner economy with r regions. All regions have the same technology in production; the only difference between regions is the endowment of factors. In each region, there are N industries, each producing a homogeneous good. The production of each industry i requires a common input, capital, and an industry-specific factor. K is the endowment of capital and T_i is the endowment of an industry-i specific factor. Due to the specific factor here, the model allows for a spatial distribution of production in different industries, even with perfect competition.

First, from the industry-specific cost minimization problem, we can obtain a_{K_i} and a_{T_i}, the capital and industry-specific factor respectively needed for production of one unit of output. Both of these are functions of factor prices. Suppose that Y_i is the output in industry i. By factor market clearing conditions, we have

$$a_{T_i} Y_i = T_i \quad \forall i \tag{2.1}$$

$$\sum_i a_{K_i} Y_i = K. \tag{2.2}$$

Under perfect competition, the profit is zero. If P_i is the price of good i, r is the interest rate for capital, and R_i is the price of the industry-i specific factor, we then have

$$a_{T_i} R_i + a_{K_i} r = P_i \quad \forall i \tag{2.3}$$

Log-linearizing (2.3), and let hat variables represent proportional changes, we have

$$(1-\theta_i)\hat{r} + \theta_i \hat{R}_i + (1-\theta_i)\hat{a}_{K_i} + \theta_i \hat{a}_{T_i} = \hat{P}_i, \tag{2.4}$$

where $\theta_i = \frac{a_{T_i} R_i}{a_{t_i} R_i + a_{K_i} r}$ presents the cost share of the specific factor in industry i.

Cost minimization implies that a_{T_i} and a_{K_i} will adjust such that small changes of factor prices will not affect costs. Using this envelop condition, we have

$$(1-\theta_i)\hat{a}_{K_i} + \theta_i \hat{a}_{T_i} = 0, \tag{2.5}$$

and thus

$$(1-\theta_i)\hat{r} + \theta_i \hat{R}_i = \hat{P}_i \quad \forall i \tag{2.6}$$

Suppose that T_i is fixed in all industries, and K might change. Log-linearizing (2.1) gives us

$$\hat{Y}_i = -\hat{a}_{T_i} \quad \forall i \tag{2.7}$$

Log-linearizing (2.2) and substituting (2.7) into the equation results in

$$\sum_i \lambda_i (\hat{a}_{K_i} - \hat{a}_{T_i}) = \hat{K}, \tag{2.8}$$

where $\lambda_i = \frac{K_i}{K}$ is the fraction of region capital used in industry i. Supposing σ_i is the elasticity of substitution between T_i and K_i in production,

$$\sigma_i = \frac{dln(a_{T_i}/a_{K_i})}{dln(r/R_i)} \tag{2.9}$$

we have

$$\hat{a}_{T_i} - \hat{a}_{K_i} = \sigma_i(\hat{r} - \hat{R}_i) \quad \forall i \tag{2.10}$$

Substituting (2.10) into (2.8), we have

$$\sum_i \lambda_i \sigma_i (\hat{R}_i - \hat{r}) = \hat{K} \tag{2.11}$$

Solving the system of equations of (2.6) and (2.11) with $N+1$ equations of $N+1$ unknowns ($\hat{R}_i$ and $\hat{r}$), and taking $\hat{P}_i$ and $\hat{K}$ as given, we have

$$\hat{r} = \frac{-\hat{K}}{\sum_{i'} \lambda_{i'} \sigma_{i'} / \theta_{i'}} + \sum_i \beta_i \hat{P}_i, \tag{2.12}$$

where

$$\beta_i = \frac{\lambda_i \sigma_i / \theta_i}{\sum_{i'} \lambda_{i'} \sigma_{i'} / \theta_{i'}}, \tag{2.13}$$

and

$$\hat{R}_i = \frac{\hat{P}_i - (1 - \theta_i)\hat{r}}{\theta_i}. \tag{2.14}$$

Moreover, we can solve for the output level in industry i using (2.7), (2.5) and (2.10).

$$\hat{Y}_i = \frac{1 - \theta_i}{\theta_i} \sigma_i (\hat{P}_i - \hat{r}). \tag{2.15}$$

$$\begin{aligned}
\widehat{GDP} = \widehat{ExportValue} &= (\widehat{\sum_i P_i Y_i}) \\
&= \sum_i \frac{P_i Y_i}{\sum_{i'} P_{i'} Y_{i'}} (\hat{P}_i + \hat{Y}_i) \\
&= \sum_i \frac{P_i Y_i}{\sum_{i'} P_{i'} Y_{i'}} (\frac{1 - \theta_i}{\theta_i} \sigma_i + 1) \hat{P}_i - \hat{r} \sum_i \frac{P_i Y_i}{\sum_{i'} P_{i'} Y_{i'}} \frac{1 - \theta_i}{\theta_i} \sigma_i \\
&= Price^t - Cost^t
\end{aligned} \tag{2.16}$$

where

$$Price^t = \sum_i \frac{P_i Y_i}{\sum_{i'} P_{i'} Y_{i'}} (\frac{1 - \theta_i}{\theta_i} \sigma_i + 1) \hat{P}_i,$$

$$Cost^t = \hat{r} \sum_i \frac{P_i Y_i}{\sum_{i'} P_{i'} Y_{i'}} \frac{1 - \theta_i}{\theta_i} \sigma_i.$$

$Price^t$ represents the direct effect of price changes: industries that experience larger price increases will have bigger increases in production, and the regional price shock is a weighted average of industry shocks. $Cost^t$ represents the indirect effect of price changes: with all industries wanting to produce more, the price of capital will increase.

In the empirical regression, I construct both price and cost using industry information and tariff changes. For simplicity, I assume the Cobb-Douglas production function, where σ_i is equal to 1. θ_i is the cost share of the specific factor, and is measured using the share of wage bill out of the total cost from the IES firm data. K_i is measured using total capital in the industry. P_iY_i is the total sales revenue in industry i. Most importantly, $\hat{P}_i$ is measured as $-\Delta \ln(1 + tariff_i/100)$.[12]

Also, Equation 17 generates a testable hypothesis: If I regress total export value on price and cost, the coefficients should be 1 on price and -1 on cost. Later in the empirical part, I will test this hypothesis.

2.4.1.1 Constant pollution intensity across industries, and constant pollution intensity over time.

Suppose that among industries $1, ..., N$, the first k industries are more capital intensive, and generates a unit of pollution for each dollar of output. Also, the pollution intensity of industries does not change over time.

$$Plln_i = aP_iY_i$$

Then, the total volume of pollution in region r is

[12] Let the price paid by consumer be P_i^c, and the price charge by exporters be P_i. Then $P_i = P_i^c/(1 + tariff_i/100)$, and $\hat{P}_i = \hat{P}_i^c - \Delta \ln(1 + tariff_i/100)$. Supposing that the price paid by the consumer does not change, we get $\hat{P}_i = -\Delta \ln(1 + tariff_i/100)$.

$$\begin{aligned}
\widehat{Plln} &= a\sum_{i=1}^{k}\frac{P_iY_i}{\sum_{i'=1}^{k}P_{i'}Y_{i'}}(\hat{P}_i+\hat{Y}_i) \\
&= a\sum_{i=1}^{k}\frac{P_iY_i}{\sum_{i'=1}^{k}P_{i'}Y_{i'}}(\frac{1-\theta_i}{\theta_i}\sigma_i+1)\hat{P}_i - a\hat{r}\sum_{i=1}^{k}\frac{P_iY_i}{\sum_{i'=1}^{k}P_{i'}Y_{i'}}\frac{1-\theta_i}{\theta_i}\sigma_i \\
&= aPrice^{p,I} - aCost^{p,I} \qquad (2.17)
\end{aligned}$$

Similar to the export value part, I use a price factor ($Price^{p,I}$, first block) and a cost factor ($Cost^{p,I}$, second block), with equal coefficients (but opposite signs). I use superscript I to indicate that I use an indicator variable to define pollution vs. non-polluting industries.

2.4.1.2 Different pollution intensities across industries, and constant pollution intensity over time.

Suppose that Industry i generates a_i unit of pollution for each dollar of output.

$$Plln_i = a_iP_iY_i$$

Then, the total volume of pollution in region r is

$$\begin{aligned}
\widehat{Plln} &= \sum_{i=1}^{n}\frac{a_iP_iY_i}{\sum_{i'=1}^{n}a_{i'}P_{i'}Y_{i'}}(\hat{P}_i+\hat{Y}_i) \\
&= \sum_{i=1}^{n}\frac{a_iP_iY_i}{\sum_{i'=1}^{n}a_{i'}P_{i'}Y_{i'}}(\frac{1-\theta_i}{\theta_i}\sigma_i+1)\hat{P}_i - \hat{r}\sum_{i=1}^{n}\frac{a_iP_iY_i}{\sum_{i'=1}^{n}a_{i'}P_{i'}Y_{i'}}\frac{1-\theta_i}{\theta_i}\sigma_i \\
&= Price^{p,A} - Cost^{p,A} \qquad (2.18)
\end{aligned}$$

In this specification, a_i is the pollution intensity of industry i. In the empirical part, it is measured using pollutant emission/output from national-level data. I use superscript A to indicate that this is an alternative measure, and a_i represents the pollution intensity in terms of SO_2 or soot. I use both the polluting/non-polluting division method and the pollution intensity method in the main regression.

2.5 Empirical model

2.5.1 Model setup

First, to study the effect of trade liberalization on city air pollution levels, I run the following regression:

$$\triangle \ln(AOT)_c = \alpha_0 + \alpha_1 \triangle \ln(export^p)_c + \Gamma X_c + \epsilon_c$$

where $\triangle \ln(AOT)_c$ is the proportional change of aerosol optical thickness in city c from 2000 to 2005, and $\triangle \ln(export^p)_c$ is the proportional change of exports in polluting industries. X_c is a vector of other city-level characteristics, including the proportional change of population density and the baseline air pollution level. However, if cities that grow substantially in polluting industries are also the ones that experience a substantial increase in vehicle use, then the effect of exports on air pollution will likely be overestimated. To solve the potential omitted variable bias problem, I will use the price factor and cost factor generated by exogenous tariff reductions as instruments for $\Delta \ln export^p$.

Similarly, to study the effect of trade liberalization on city income levels, I run the following regression:

$$\triangle \ln(GDP)_c = \beta_0 + \beta_1 \triangle \ln(export^t)_c + \Gamma X_c + \epsilon_c$$

where $\triangle \ln(GDP)_c$ is the change in the log of GDP in city c, and $\triangle \ln(export^t)_c$

is the change in the log of exports in all industries. However, if cities that grow substantially in exports are also the ones that experience a substantial growth in non-export goods, then the effect of exports on city income levels will be overestimated. Thus, I use regional tariff changes as instruments for total export growth.

Finally, I would like to see how the income and pollution jointly influence people's health. The main regression is

$$\triangle \ln MR_c = \gamma_0 + \gamma_1 \triangle \ln(GDP)_c + \gamma_2 \triangle \ln(AOT)_c + \Gamma X_c + \epsilon_c$$

where $\triangle \ln MR_c$ is the change in the total mortality rate in city c. There can be extra confounding factors, including government investment in the medical system. If cities with higher economic growth also invest more in their health system and discourage growth of polluting industries, estimates of γ_1 and γ_2 will be biased. Regional tariff changes will help to disentangle the income and pollution effect after trade liberalization.

2.5.2 The pre-trend of mortality rates

One threat to the identification of the income and pollution effects in the main regression on the mortality rate is whether the mortality rate pre-trend by city is correlated with tariff changes in later periods. In other words, if the cities that experienced higher regional tariff shocks (either price factor or cost factor) had had a declining (or increasing) trend in their mortality rate even before the WTO accession, then we cannot take tariff shocks as exogeneous.

In order to check the mortality rate pre-trend, I will use the mortality information from the DSP system. The regression is as follows:

$$MR_{2000,c} = \alpha_0 + \alpha_1 Price_c^p + \alpha_2 Cost_c^p + \alpha_3 Price_c^t + \alpha_4 Cost_c^t + \alpha_5 MR_{t,c} + \epsilon_c$$

where $MR_{2000,c}$ is the total mortality rate in city c and year 2000, and $MR_{t,c}$ is the total mortality rate in city c and baseline year t, where t can be any year between 1991 and 1999. $Price_c^t$, $Cost_c^t$, $Price_c^p$, and $Cost_c^p$, are regional tariff shocks that affect the price and cost of production in city c that happened between 2000 and 2005. If there is a pre-trend of the mortality rate (before 2000) that correlates with the tariff shocks a later period (after 2000), we would expect some of α_1 to α_4 to be significant. Thus, after running the regression, I then test the joint hypothesis:

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$$

If I fail to reject the joint hypothesis, I make the conclusion that there is no evidence in the existence of a pre-trend.

Table 2.2: Is the past trend in mortality rates correlated with tariff shocks?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Baseline year	1991	1992	1993	1994	1995	1996	1997	1998	1999
$\text{Price}^{p,I}$	27.6	12.8	3.5	15.0	-14.1	-39.5	1.3	7.4	9.1
	(32.3)	(32.2)	(39.1)	(33.0)	(30.2)	(27.1)	(20.3)	(17.3)	(16.9)
$\text{Cost}^{p,I}$	11.6	32.6	34.5	11.9	2.8	42.9	-11.0	10.7	-9.2
	(29.9)	(34.8)	(32.7)	(29.7)	(35.7)	(31.0)	(27.0)	(16.9)	(15.3)
Price^{t}	1.5	11.9	44.7	1.6	24.7	47.2	17.4	6.8	11.9
	(41.2)	(31.9)	(44.3)	(40.0)	(36.9)	(31.8)	(25.1)	(19.5)	(14.7)
Cost^{t}	1.6	-17.5	-48.3	7.0	-11.7	-46.9	26.6	-2.3	0.8
	(42.1)	(36.8)	(49.0)	(42.7)	(44.4)	(37.7)	(35.2)	(23.3)	(15.8)
MR baseline	0.4***	0.4***	0.4***	0.4***	0.6***	0.6***	0.6***	0.8***	0.8***
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
Constant	29.0***	29.4***	27.1***	29.7***	20.6***	19.3***	17.3***	8.6*	8.0
	(5.7)	(6.9)	(5.6)	(6.0)	(6.0)	(4.6)	(5.3)	(5.0)	(5.7)
Observations	113	109	107	111	109	114	114	112	111
R-squared	0.2	0.2	0.3	0.2	0.5	0.4	0.4	0.6	0.7
F-stat for H_0	0.6	0.6	0.4	0.6	0.9	0.5	0.3	0.8	0.3

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Errors clustered at the province level. Different columns use different baseline years (1991-1999).

As shown in Table 2.2, no matter which baseline year I use, there is no evidence that regional tariff shocks after 2000 are correlated with an associated change in the mortality rate before 2000 (F statistics for joint test range from 0.3 to 0.9). In terms of the baseline year mortality rate, it is evident that indeed mortality rates in years closer to 2000 predicts mortality rate in 2000 better ($\hat{\alpha}_5$ is 0.4 in early years and 0.8 in later years). To conclude, I find that tariff shocks are exogenous to local health condition trends.

2.5.3 Main specification

2.5.3.1 Exports and GDP, first stage, reduced form.

First, I check the relationship between regional tariff shocks, production and income. The regression is

$$\Delta Y_c = \beta_0 + \beta_1 Price_c + \beta_2 Cost_c + \Gamma X_c + \epsilon_c$$

where Y_c is the log of per capita GDP or the log of total export values in city c, and $Price_c$ is price factor $Price^t_c$ and $Cost_c$ is cost factor $Cost^t_c$. When Y_c is the log of export values in polluting industries, then $Price_c$ =$Price^{p,I}_c$ and $Cost_c = Cost^{p,I}_c$. I also control for other factors, such as the proportional change in the population density and the proportional change in the export share of the total sales revenue.

The model also generates a testable of hypothesis of $\beta_1 = -\beta_2 = 1$. The regression results are as detailed in Table 2.3. We indeed see a positive price effect and a negative cost effect, and the joint test failed to reject the null hypothesis. In the GDP regression, both price and cost factors have significant effects, while in the export regressions, we can only observe a significant cost effect.

Table 2.3: How tariff shocks affect export and income

	$\triangle \ln(GDP_{pc})$		$\triangle \ln(Export^t)$		$\triangle \ln(Export^p)$	
Mean(Y)	0.63		1.10		0.99	
	(1)	(2)	(3)	(4)	(5)	(6)
Price	0.84***	0.93***	1.11	0.34	1.54	0.47
	(0.26)	(0.26)	(1.78)	(0.65)	(2.08)	(0.80)
Cost	-0.53**	-0.67**	-1.56	-2.27*	-4.77*	-3.22**
	(0.22)	(0.24)	(3.22)	(1.25)	(2.31)	(1.24)
Observations	106	106	106	106	105	105
R-squared	0.06	0.08	0.01	0.78	0.04	0.84
Other controls	No	Yes	No	Yes	No	Yes
Test: $\beta_1 = -\beta_2 = 1$	0.13	0.33	0.98	0.21	0.29	0.17

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Errors clustered at the province level. Compared with Columns (1)(3)(5), Columns (2)(4)(6) are controlled with changes in the log of the population density and changes in export share of output. Regressions weighted by the 2000 city population size.

2.5.3.2 Exports and pollution, first stage, reduced form.

From the theoretical model, we have the relationship between city air pollution levels and regional tariff shocks as follows:

$$\Delta \ln(AOT)_c = \beta_0 + \beta_1 Price_c^p + \beta_2 Cost_c^p + \Gamma X_c + \epsilon_c$$

Also, depending on my assumption of industrial pollution intensity, I have four specifications in Table 2.4: Column (1) and Column (2) with $Price^{p,I}$ and $Cost^{p,I}$, and Column (3) and Column (4) with $Price^{p,A}$ and $Cost^{p,A}$. Column (3) uses the SO_2 emission intensity as the measure of a_i and Column (4) uses the soot intensity. Column (1) has no additional controls, while the latter three columns have a proportional change in the population density and initial air pollution levels.

In Table 2.4, we can see that in all specifications where we have extra controls (Column (2)-(4)), we have a significantly positive price effect, but not a significant cost effect. In the joint test of $H_o : \beta_1 = -\beta_2$, two out of four specifications fail to reject the null hypothesis.

Table 2.4: How tariff shocks affect pollution levels

$\triangle \ln(AOT)$	I(Polluting)		SO2	Soot
	(1)	(2)	(3)	(4)
$Price^p$	0.66	0.70*	1.61***	0.04***
	(0.50)	(0.35)	(0.39)	(0.01)
$Cost^p$	-0.73	-0.26	0.15	0.00
	(0.52)	(0.43)	(0.29)	(0.01)
Observations	106	106	106	106
R-squared	0.06	0.15	0.22	0.23
Controls	No	Yes	Yes	Yes
Test: $\beta_1 = -\beta_2$	0.90	0.24	0.00	0.00

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Errors clustered at the province level. Compared with Column (1), Columns (2)(3)(4) are controlled with changes in the log of population density and initial air pollution level ($\log(AOT)_{2000}$). Regressions weighted by 2000 city population.

2.5.3.3 Mortality regression: OLS and 2SLS

Finally, I will combine the income regression and pollution regression to see how trade liberalization affect people's health.

$$\triangle \ln MR_c = \gamma_0 + \gamma_1 \triangle \ln(GDP)_c + \gamma_2 \triangle \ln(AOT)_c + \Gamma X_c + \epsilon_c$$

Due to possible endogeneity concerns discussed previously, I use regional tariff shocks to instrument for changes in GDP and changes in pollution levels. In Table 2.5, Column (1) is a OLS regression with changes in GDP and changes in AOT only, with no other controls, and Column (2) adds controls such as proportional changes in the population density, proportional changes in the export share of output, proportional changes in employment, and baseline GDP and AOT levels. In Column (3) to (5), I instrument changes in GDP and changes in AOT as shown in the previous first-stage analysis, with different specifications. From the OLS regression with controls, regions that faced a 10 percent larger GDP per capita increase experienced a 2.5 percent larger mortality decline, and a 10 percent large air pollution increase led to a 2.9 percentage larger mortality rate increase. In the 2SLS regressions, the effect of GDP ranges from 0.55 to 0.65, and the effect of air pollution ranges from 0.43 to 1.34. Overall, in the 2SLS, both the beneficial effect of GDP and the harmful effect of air pollution become larger. At the sample mean of $\triangle \ln(GDP_{pc})$ (0.6) and the mean of $\Delta(AOT)$ (0.12), in the absence of GDP growth, $ln(MR)$ increases by 0.4, and in the absence of an increase in air pollution, $ln(MR)$ decreases by 0.14. Both factors are important determinants of the total mortality rate and should be accounted for when analyzing the health consequences of trade liberalization.

Table 2.5: Total mortality and trade shock, main result

$\triangle\ln(MR)$	OLS		2SLS		
Mean: -0.19			I_p	SO2	Soot
	(1)	(2)	(3)	(4)	(5)
$\triangle\ln(GDP_{pc})$	-0.16*	-0.25**	-0.65	-0.55**	-0.56**
	(0.09)	(0.11)	(0.38)	(0.24)	(0.21)
$\triangle\ln(AOT)$	0.15	0.29**	1.34*	0.43*	0.58**
	(0.12)	(0.11)	(0.70)	(0.22)	(0.23)
$\triangle\ln(PopDen)$		-0.49*	-0.67*	-0.55*	-0.57**
		(0.23)	(0.36)	(0.28)	(0.26)
$\triangle\ln(E\%)$		-0.07**	-0.09***	-0.07**	-0.07**
		(0.03)	(0.02)	(0.03)	(0.03)
$\triangle\ln(Emp)$		-0.06	-0.33	-0.13	-0.16
		(0.14)	(0.23)	(0.16)	(0.16)
$\ln(GDP_{pc})_{2000}$		-0.04	-0.05	-0.05	-0.05
		(0.04)	(0.05)	(0.04)	(0.04)
$\ln(AOT)_{2000}$		0.14**	0.31**	0.18**	0.20**
		(0.06)	(0.14)	(0.08)	(0.07)
Observations	106	106	106	106	106
R-square	0.06	0.29		0.18	0.12
F-stat for GDP	-	-	2.53	2.40	2.51
F-stat for AOT	-	-	2.68	3.63	3.96

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Errors clustered at the province level. Columns (1) (2) are OLS regressions. Columns (3)-(5) areIV regressions where both GDP and AOT are instrumented using price factors and cost factors. Column (3) uses $Price^{p,I}$ and $Cost^{p,I}$, and Columns (4)-(5) use different pollutant emission intensities. F-stat for GDP and F-stat for AOT are F-statistics forfirst stage regressions. Regressions weighted by the 2000 city population.

2.6 Discussion and robustness checks

2.6.1 Discussion of the government's role

One concern about identification using tariff shocks is the potential role of the government. If local governments use different policy instruments to either offset or reinforce the regional tariff shocks, then the effect that I identify is trade shocks confounded with policy reactions. Possible policies that local governments use to promote exports include providing value-added tax (VAT) rebates, low interest rate loans, and low cost land to exporting firms. Although it would be good to test all possible policy effects, due to data constraints, I only check if the VAT rebate is correlated with tariff shocks.

The VAT rebate is allowed by the WTO rule to avoid tax multiplicity. In China, the VAT rebate has been part of the export-related fiscal policies ever since the 1985. Since 1994, the VAT rate has been set to 17% for most manufacturing goods, and the official rebate rate was set to be either a full rebate (17%) or a partial rebate (5%, 13% 15%) for export goods. However, the real rebate rates vary significantly by region (and a firm's ability to claim the rebate), especially after the 2004 reform when the central government was no longer the only payer for the rebate, and started to have 25% payment by local governments and 75% by the central government conditional on local payment.[13]

Using real VAT rebate rates calculated from firm-level data, Chandra and Long (2013) shows that each percentage point increase in the VAT rebate rate leads to an increase of the export value by 13%. The formula to calculate the firm-level rebate rate is as follows:

$$RebateRate_{VAT} = (0.17 * revenue - VAT_{throughput} - VAT_{payable})/export$$

The downside of this approach is that the firm-level VAT information can have considerable noise due to reporting errors or a lag in payment for rebates. Thus, I calculate the real rebate by firm, and take mean values by industry-city to construct averages in 2000 and 2005. Also, I drop top and bottom 1% firm level rebate rates, and either top and bottom 1% or 5% city-industry level rebate rates. From Equation (2.15), the relationship between industry output and prices is

[13] For details about the VAT rebate system and the 2004 reform in China, see Chandra and Long (2013).

$$\widehat{P_iY_i} = \hat{P}_i + \frac{1-\theta_i}{\theta_i}\sigma_i(\hat{P}_i - \hat{r}) = (\frac{1-\theta_i}{\theta_i}\sigma_i + 1)\hat{P}_i - \frac{1-\theta_i}{\theta_i}\sigma_i\hat{r}$$

where $\hat{P}_i$ can be decomposed into $-\Delta\ln(1+tariff_i/100)$ and $-\Delta\ln(1+17\%-rebate_i)$, where the first refers to the tariff reduction and the second refers to the tax rebate. Thus, I have

$$\widehat{P_iY_i} = -(\frac{1-\theta_i}{\theta_i}\sigma_i+1)\Delta\ln(1+\frac{tariff_i}{100}) - (\frac{1-\theta_i}{\theta_i}\sigma_i+1)\Delta\ln(117\%-rebate_i) - \frac{1-\theta_i}{\theta_i}\sigma_i\hat{r}$$

I call the first block $Price_i$, the second block $Rebate_i$ and the third block $Cost_i$. Stacking all cities and industries together, I run the following regression:

$$\Delta\ln(export)_{ic} = \theta_0 + \theta_1 Price_{ic} + \theta_2 Cost_{ic} + \theta_3 Rebate_{ic} + \Gamma X_{ic} + \epsilon_{ic}$$

where i represents the industry and c represents the city. I first assume that $\theta_3 = 0$, and then add $Rebate_{ic}$ into the regression. If the estimates for θ_1 and θ_2 do not change considerably after adding $Rebate_{ic}$, and I fail to reject $H_o : \theta_3 = 0$, then I conclude that there is no evidence that the local government VAT rebate is correlated with regional tariff shocks.

The regression results are shown in Table 2.6. In the first three columns, I drop the city-industries pairs that fall in the top or bottom 1% in terms of rebate rates. Column (1) controls for *Price*, *Cost* and an indicator variable for whether the industry is a polluting industry or not. Column (2) controls for *Rebate*, and Column (3) adds an interaction of *Rebate* and polluting industry dummy. We can see that the coefficients for *Price* and *Cost* do not change much across specifications, and the coefficient for *Rebate* is not significant. Column (4)-(6) are the same as Column (1)-(3) except that I drop the top and bottom 5%, and the results are similar.

Overall, I find no evidence that local governments are trying to reinforce or offset the regional tariff shocks using export VAT rebates. It might be true that they are trying to promote export growth using different types of policies, but as long as these policies are not correlated with tariff shocks, the identification strategy still works.

Table 2.6: Are local VAT rebate rates correlated with tariff shocks?

	Dropped top and bottom 1%			Dropped top and bottom 5%		
$\Delta \ln(export)$	(1)	(2)	(3)	(4)	(5)	(6)
Price	1.00***	1.00***	1.01***	0.87***	0.87***	0.86***
	(0.23)	(0.23)	(0.23)	(0.28)	(0.28)	(0.28)
Cost	-0.23	-0.23	-0.21	-0.47	-0.48	-0.49
	(0.40)	(0.40)	(0.40)	(0.57)	(0.57)	(0.57)
Polluting industry (=1)	-0.24***	-0.24***	-0.25***	-0.26***	-0.26***	-0.26***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Rebate		-0.00	-0.01		0.01	0.01
		(0.01)	(0.01)		(0.01)	(0.02)
Interaction			0.01			-0.01
			(0.01)			(0.02)
Log(export_{2000})	-0.20***	-0.20***	-0.20***	-0.16***	-0.16***	-0.16***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	3.30***	3.30***	3.30***	2.86***	2.85***	2.84***
	(0.26)	(0.26)	(0.26)	(0.26)	(0.26)	(0.26)
Observations	2,118	2,118	2,118	1,865	1,865	1,865
R-squared	0.10	0.10	0.11	0.08	0.08	0.08

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Errors clustered at the city level. Column (1)-(3) drop city-industry pairs that fall in the top or bottom 1

2.6.2 Robustness check

2.6.2.1 Infant mortality

One concern for using the total mortality rate is that adult mortality is based on a life-time of accumulated effects, while the infant mortality rate would be a cleaner measure since infants have limited time exposure to outside factors. However, due to possible reporting errors in census data for infants, I only use infant mortality as one robustness check.

As shown in Table 2.7, when using changes in the log of infant mortality rates as the outcome, the beneficial effect of income is not significant and has mixed signs across specifications. There is still some evidence pointing to the harmful effect of pollution, and the magnitude is much larger than in the total mortality regression. This confirms that infants are more sensitive to air pollution. On the income side, one hypothesis is that the infant mortality that is affected on the margin by income growth might be for the low-income group only (who cannot afford to deliver babies in hospitals for example), while for the higher-income group, the infant mortality rate is already quite low and is not sensitive to income growth. However, without detailed individual-level data, I am not able to test this hypothesis.

Table 2.7: Infant Mortality and trade shock

$\triangle \ln(IMR)$	OLS		2SLS		
Mean: -0.78			I_p	SO2	Soot
$\triangle \ln(GDP_{pc})$	-0.43	-0.31	0.63	0.41	0.00
	(0.42)	(0.51)	(2.24)	(1.80)	(1.49)
$\triangle \ln(AOT)$	-0.05	0.23	3.69*	2.58	1.83*
	(0.32)	(0.43)	(2.03)	(1.45)	(0.92)
$\triangle \ln(PopDen)$		-3.07	-5.18	-4.54	-4.00
		(2.37)	(3.20)	(2.60)	(2.48)
$\triangle \ln(E\%)$		-0.13	-0.19	-0.17*	-0.16*
		(0.08)	(0.11)	(0.09)	(0.09)
$\triangle \ln(Emp)$		0.07	-0.42	-0.25	-0.17
		(0.68)	(0.72)	(0.62)	(0.59)
$\ln(GDP_{pc})_{2000}$		0.25	0.27	0.26	0.26
		(0.20)	(0.23)	(0.22)	(0.21)
$\ln(AOT)_{2000}$		0.17	0.55	0.42	0.36
		(0.24)	(0.51)	(0.49)	(0.40)
Observations	93	93	93	93	93
R-squared	0.01	0.12			0.00

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Errors clustered at the province level. Column (1) (2) are OLS regressions. Column (3)-(5) areIV regressions where both GDP and AOT are instrumented using price factors and cost factors. Column (3) uses the polluting vs. non-polluting division, and Column (4)-(5) use different pollutant emission intensities. Regressions weighted by the 2000 city population.

2.6.3 Adjust calculation of return to capital

From Equation (2.12), to calculate the return to capital $\hat{r}$ (which is an input in the calculation of $Cost^t$ and $Cost^p$), I need information regarding $\hat{K}$ and $\lambda_i = K_i/K$ since

$$\hat{r} = \frac{-\hat{K}}{\sum_{i'} \lambda_{i'}\sigma_{i'}/\theta_{i'}} + \sum_i \beta_i \hat{P}_i$$

where

$$\beta_i = \frac{\lambda_i \sigma_i/\theta_i}{\sum_{i'} \lambda_{i'}\sigma_{i'}/\theta_{i'}}$$

In the main regression (Table 2.5), I choose $\hat{K} = 0$ and capital (K_i) as the total capital in industry i reported by firms. In order to check the robustness of previous results, I use different sets of assumptions about capital mobility and definitions of capital.

In Table 2.8, Column (1) is the OLS regression as in Column (2), Table 2.5. Column (2) uses $\hat{K} = K_{2005} - K_{2000}$, and K is total capital in a certain city. Column (3)-(5) maintains the assumption of $\hat{K} = 0$, but changes the definition of capital. Column (3) uses $K_i^{new} = export_i/sales_i * K_i$, adjusting for the export intensity of the industry. Column (4) uses fixed capital instead of total capital. Column (5) uses $K_i^{new} = export_i/sales_i * K_i^{fixed}$. In all 2SLS results, for simplicity I use the same instruments as in Column (5), Table 2.5, where the pollutant intensity is measured by soot emission rates.

The results are quite robust across different specifications, thus there is no evidence that the choice of parameters in calculation of return on capital is driving the results.

Table 2.8: Total mortality and trade shock, different $\hat{r}$

$\triangle \ln(MR)$	OLS	2SLS			
Mean: -0.19	(1)	(2)	(3)	(4)	(5)
$\triangle \ln(GDP_{pc})$	-0.25**	-0.39	-0.67*	-0.61**	-0.67
	(0.11)	(0.27)	(0.33)	(0.26)	(0.39)
$\triangle \ln(AOT)$	0.29**	0.58**	0.68*	0.56**	0.72**
	(0.11)	(0.24)	(0.31)	(0.22)	(0.31)
$\triangle \ln(PopDen)$	-0.49*	-0.55**	-0.59*	-0.57*	-0.60*
	(0.23)	(0.22)	(0.29)	(0.27)	(0.28)
$\triangle \ln(E\%)$	-0.07**	-0.07**	-0.07**	-0.07**	-0.07**
	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
$\triangle \ln(Emp)$	-0.06	-0.14	-0.20	-0.17	-0.21
	(0.14)	(0.16)	(0.16)	(0.16)	(0.18)
$\ln(GDP_{pc})_{2000}$	-0.04	-0.05	-0.05	-0.05	-0.05
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
$\ln(AOT)_{2000}$	0.14**	0.19**	0.23**	0.21**	0.23***
	(0.06)	(0.08)	(0.07)	(0.08)	(0.08)
Observations	106	106	106	106	106
R-squared	0.29	0.19		0.10	

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Errors clustered at the province level. Column (1) is the OLS regression. Column (2)-(5) areIV regressions where both GDP and AOT are instrumented using price factors and cost factors. Column (2) allows for across-region capital mobility. Column (2)-(5)use different ways to calculate return to capita. Column (3) uses capital adjusted by export intensity. Column (4) uses fixed capital. Column (5) uses fixed capitaladjusted by export intensity. All 2SLS regressions assume soot as the main pollutant. Regressions weighted by the 2000 city population.

2.6.4 Night-time light intensity instead of GDP

Another concern is that GDP data might be subject to errors in reporting and manipulation by the local government. The night-time light intensity has been proven to provide an alternative measure of the economic activity intensity, and the data quality is not subject to political conditions. Thus, in Table 2.9, I replace GDP per capita with the night-time light intensity, and the effect of air pollution still remains.

Table 2.9: Total mortality and trade shock

$\triangle \ln(MR)$	OLS		2SLS		
Mean: -0.19			I_p	SO2	Soot
$\triangle \ln(light)$	-0.08	-0.11	-0.48	-0.47*	-0.44**
	(0.07)	(0.07)	(0.36)	(0.22)	(0.20)
$\triangle \ln(AOT)$	0.16	0.31**	1.52*	0.63	0.75**
	(0.11)	(0.13)	(0.75)	(0.39)	(0.34)
$\triangle \ln(PopDen)$		-0.39*	-0.25	-0.17	-0.20
		(0.21)	(0.42)	(0.23)	(0.24)
$\triangle \ln(E\%)$		-0.07**	-0.10***	-0.08***	-0.08***
		(0.03)	(0.02)	(0.02)	(0.02)
$\triangle \ln(Emp)$		-0.04	-0.24	-0.08	-0.10
		(0.12)	(0.23)	(0.16)	(0.16)
$\ln(Light)_{2000}$		-0.02	-0.07	-0.06	-0.06
		(0.03)	(0.04)	(0.03)	(0.03)
$\ln(AOT)_{2000}$		0.15*	0.41**	0.28**	0.29**
		(0.07)	(0.17)	(0.11)	(0.11)
Observations	106	106	106	106	106
R-squared	0.04	0.24			

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Errors clustered at the province level. Column (1) (2) are OLS regressions. Column (3)-(5) areIV regressions where both light and AOT are instrumented using price factors and cost factors. Column (3) uses polluting vs. non-polluting division, and Column (4)-(5) use different pollutant emission intensities. Regressions weighted by the 2000 city population.

2.6.5 Alternative check of the model

I also conduct an alternative test of the theory by replacing $Price^{p,I}$ and $Cost^{p,I}$ with $Price^{np}$ and $Cost^{np}$[14], and use them in both the first-stage for trade shocks and air pollution, as well as the final mortality regression.

In Table 2.10, the first two columns are the same as in Table 2.4, while in Column (3) and (4), I use $Price^{np}$ and $Cost^{np}$ instead of $Price^{p}$ and $Cost^{p}$. We can see that regional tariff shocks in polluting and non-polluting industries have exactly the opposite effects on change in pollution level. This result reiterates

[14] $Price^{np}$ and $Cost^{np}$ can be calculated using the two blocks in Equation (2.17), by replace the index of $i = 1, ..., k$ with $i = k + 1, ..., n$.

the intuition of the model: production factors are allocation among polluting and non-polluting industries, and good shock in one sector is an indirect bad shock to the other sector.

In Table 2.11, Column (1) and (2) are OLS regression as in Table 2.5, Column (3) is the 2SLS regression when pollution change is instrumented with polluting industry shocks, and Column (4) is the 2SLS regression when instrumented with non-polluting industry shock. Coefficient estimates are very close to each other in Column (3) and (4), and both first-stage in Table 2.10 and 2SLS in Table 2.11 confirm that the model's intuition carries through in the empirical regressions.

Table 2.10: How tariff shocks affect pollution, alternative test

$\triangle \ln(AOT)$	I(Polluting)		I(Non-polluting)	
Mean: 0.12	(1)	(2)	(3)	(4)
Price	0.66	0.70*	-0.36**	-0.42**
	(0.50)	(0.35)	(0.16)	(0.18)
Cost	-0.73	-0.26	0.30	0.62**
	(0.52)	(0.43)	(0.27)	(0.24)
Observations	106	106	106	106
R-squared	0.06	0.15	0.03	0.15
Test: $-\beta_1 = \beta_2$	0.90	0.24	0.80	0.32
Controls	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Errors clustered at the province level. Column (1) and (2) use$Price^{p,I}$ and $Cost^{p,I}$ as regressors, and Column (3) and (4) use $Price^{np}$ and $Cost^{np}$ as regressors. Column (2) and (4) controlfor changes in the log of the population density and the initial air pollutionlevel ($\log(AOT)_{2000}$). Regression weighted by the 2000 city population.

Table 2.11: Total mortality and trade shock, alternative test

$\triangle \ln(MR)$	OLS		2SLS	
Mean: -0.19			I_p	I_{np}
$\triangle \ln(GDP_{pc})$	-0.16*	-0.25**	-0.65	-0.61*
	(0.09)	(0.11)	(0.38)	(0.30)
$\triangle \ln(AOT)$	0.15	0.29**	1.34*	1.06**
	(0.12)	(0.11)	(0.70)	(0.39)
$\triangle \ln(PopDen)$		-0.49*	-0.67*	-0.63*
		(0.23)	(0.36)	(0.32)
$\triangle \ln(E\%)$		-0.07**	-0.09***	-0.08***
		(0.03)	(0.02)	(0.02)
$\triangle \ln(Emp)$		-0.06	-0.33	-0.27
		(0.14)	(0.23)	(0.16)
$\ln(GDP_{pc})_{2000}$		-0.04	-0.05	-0.05
		(0.04)	(0.05)	(0.05)
$\ln(AOT)_{2000}$		0.14**	0.31**	0.27**
		(0.06)	(0.14)	(0.10)
Observations	106	106	106	106
R-squared	0.06	0.29		

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Errors clustered at province level. Column (1) (2) are OLS regressions. Column (3)-(4) are IV regressions where both GDP and AOT are instrumented using price factors and cost factors. Column (3) uses $Price^{p,I}$ and $Cost^{p,I}$ to instrument for pollution, and Column (4) uses $Price^{np}$ and $Cost^{np}$. Regressions weighted by 2000 city population.

2.6.6 Back-of-the-envelope calculation

In the previous analysis, pollution is a natural output of increased production in the polluting sector. Only consumption enters into the consumer's utility function, and the cost of pollution in terms of worsening health conditions is not taken into account. However, supposing that we put the health cost of the pollution into the welfare calculation, what would the net welfare gain from trade amount to? Following Becker et al. (2005), I perform a simple welfare analysis. Suppose that the overall welfare

$$V(y, MR) = (1 - MR)u(y)/r$$

where $1 - MR$ is the probability of survival, $u(y)$ is the utility from consumption and r is the interest rate. The utility function takes the form of

$$u(y) = \frac{y^{1-1/\gamma}}{1 - 1/\gamma}$$

Log-linearizing the function $V(y, MR)$

$$\hat{V}(y, MR) = (\widehat{1 - MR}) + (1 - 1/\gamma)\hat{y} \approx -(MR_{2005} - MR_{2000}) + (1 - 1/\gamma)\hat{y}$$

$$\begin{aligned} \hat{V}(y, MR) &= (\widehat{1 - MR}) + (1 - 1/\gamma)\hat{y} \\ &\approx -(MR_{2005} - MR_{2000}) + (1 - 1/\gamma)\hat{y} \\ &\approx -(e^{\hat{MR}} - 1) * MR_{2000} + (1 - 1/\gamma)\hat{y} \end{aligned}$$

I will use regression results from Section 5 to generate predicted values of GDP per capita, air pollution, and mortality rates, and then calculate welfare gains, under three sets of assumptions. Income is predicted using Table 2.3, Column (2); air pollution is predicted using Table 2.4, Column (2); the mortality rate is predicted using Table 2.5, Column (3). In the first case, I predict income with $Price^t = Cost^t = 0$, and air pollution using $Price^{p,I} = Cost^{p,I} = 0$. In the second case, I predict income with current $Price^t$ and $Cost^t$ levels, and predict air pollution using $Price^{p,I} = Cost^{p,I} = 0$. In the last case, I use income and air pollution predicted at current price and cost factor levels. Using the same parameter values as in Becker et al. (2005), I take $\gamma = 1.25$.

Results are summarized in Table 2.12. Case 1 assumes that there is no tariff shock; thus, income and pollution are not affected by trade. Case 2 allows tariff shocks to affect income, but assumes that tariff shocks will not affect pollution levels (i.e. as if all production were clean). Case 3 represents the real-life case where tariff shocks affect both income and pollution levels. Comparing the three cases, we can see that the growth rate in GDP per capita is about 1% larger with trade, and the growth rate in air pollution levels is 3.6% larger with trade. In terms of a decline in the total mortality rate, Case 2 is the highest since people enjoy the income benefits and do not suffer from pollution increases. In the real-life case, the rate of increase in income and pollution levels is even worse than in Case 1, meaning that people are better off in terms of health without trade. However, if we look at overall welfare measure, it is higher in Case 3 than in Case 1, meaning that overall, people are still better off with trade, although ideally we would like to have Case 2.

In Case 3, the total mortality rate declines by 17.2%[15], while in Case 2, it declines by 20.8%[16]. Thus, in the absence of pollution, the total mortality rate would have decline by an additional 3.6%. The difference in welfare in Case 2 and Case 3 is small, and one reason is that the cost of air pollution only enters the utility function through the survival rate, while in real life, an increase in medical costs, loss in property value and other avoidance costs can also be significant.

[15] Calculated using exp(-0.187)-1

[16] Calculated using exp(-0.234)-1

Table 2.12: Welfare analysis

		Case 1 Trade=0, Pollution=0	Case 2 Trade=1, Pollution=0	Case 3 Trade=1, Pollution=1
$\widehat{GDP_{p.c.}}$	Mean	0.608	0.617	0.617
	Std.	0.046	0.054	0.054
$\widehat{AOT}$	Mean	0.109	0.109	0.145
	Std.	0.064	0.064	0.069
$\widehat{MR}$	Mean	-0.225	-0.234	-0.187
	Std.	0.109	0.116	0.113
$\widehat{welfare}$	Mean	0.123	0.126	0.125
	Std.	0.009	0.011	0.011

2.7 Conclusion

This paper examines the evidence of how China's accession into the WTO in 2001 affected income, air pollution level and mortality rates across 106 Chinese cities. Using regional tariff shocks as instruments for changes in income and pollution levels, I show that cities that faced a 10% larger GDP per capita increase experienced a 6%-7% larger total mortality decline, and regions that faced a 10% larger air pollution increase experienced a 4%-13% larger total mortality increase. Overall, if all exports were generated from non-polluting industries, the total mortality rate would have declined by an additional 3.6%. The results are robust across different specifications, and there is no evidence that local governments are trying to reinforce or cancel out the tariff shocks using export policies. However, in terms of overall welfare, the gains from income growth outweight losses from increased pollution levels.

Appendix

Matching U.S. Industry with Chinese Industry

Table 2.13: List of relevant U.S. industry, by 2-digit SIC code

Industry Code	Industry Name
1	Agricultural Production - Crops
2	Agricultural Production - Livestock and Animal Specialties
7	Agricultural Services
8	Forestry
9	Fishing, Hunting and Trapping
10	Metal Mining
12	Coal Mining
13	Oil and Gas Extraction
14	Mining and Quarrying of Nonmetallic Minerals, Except Fuels
15	Building Cnstrctn - General Contractors & Operative Builders
16	Heavy Cnstrctn, Except Building Construction - Contractors
17	Construction - Special Trade Contractors
20	Food and Kindred Products
21	Tobacco Products
22	Textile Mill Products
23	Apparel, Finished Prdcts from Fabrics & Similar Materials
24	Lumber and Wood Products, Except Furniture
25	Furniture and Fixtures
26	Paper and Allied Products
27	Printing, Publishing and Allied Industries
28	Chemicals and Allied Products
29	Petroleum Refining and Related Industries
30	Rubber and Miscellaneous Plastic Products
31	Leather and Leather Products
32	Stone, Clay, Glass, and Concrete Products
33	Primary Metal Industries
34	Fabricated Metal Prdcts, Except Machinery & Transport Eqpmnt
35	Industrial and Commercial Machinery and Computer Equipment
36	Electronic, Elctrcl Eqpmnt & Cmpnts, Excpt Computer Eqpmnt
37	Transportation Equipment
38	Mesr/Anlyz/Cntrl Instrmnts; Photo/Med/Opt Gds; Watchs/Clocks
39	Miscellaneous Manufacturing Industries

Note: Data from https://www.osha.gov/pls/imis/sic_manual.html

Table 2.14: List of Relevant Chinese Industries by 2-digit GB code

Industry Code	Industry name
6	Agriculture
7	Mining and Washing of Coal
8	Extraction of Petroleum and Natural Gas
9	Mining and Processing of Ferrous Metal Ores
10	Mining and Processing of Non-ferrous Metal Ores
11	Mining and Processing of Nonmetal Ores
13	Mining of Other Ores
14	Processing of Food from Agricultural Products
15	Manufacture of Foods
16	Manufacture of Beverages
17	Manufacture of Tobacco
18	Manufacture of Textile
19	Manufacture of Textile Wearing Apparel, Footware, and Caps
20	Manufacture of Leather, Fur,Feather and Related Products Feather etc.
21	Processing of Timber, Manufacture of Wood, Bamboo,Rattan, Palm, etc.
22	Manufacture of Furniture
23	Manufacture of Paper and Paper Products
24	Printing,Reproduction of Recording Media
25	Manufacture of Articles for Culture, Education and Sport Activity
26	Processing of Petroleum, Coking, Processing of Nuclear Fuel
27	Manufacture of Raw Chemical Materials and Chemical Products
28	Manufacture of Medicines
29	Manufacture of Chemical Fibers
30	Manufacture of Rubber
31	Manufacture of Plastics
32	Manufacture of Non-metallic Mineral Products
33	Smelting and Pressing of Ferrous Metals
34	Smelting and Pressing of Non-ferrous Metals
35	Manufacture of Metal Products
36	Manufacture of General Purpose Machinery
37	Manufacture of Special Purpose Machinery
39	Manufacture of Transport Equipment
40	Manufacture of Electrical Machinery and Equipment
41	Manufacture of Communication Equipment, Computers and etc.
42	Manufacture of Measuring Instruments and Machinery etc.
43	Manufacture of Artwork and Other Manufacturing

Note: Data from http://www.stats.gov.cn/tjsj/tjbz/201301/t20130114_8675.html

Table 2.15: Matching U.S. and Chinese industries

Chinese Ind. Code	U.S. Ind. Code	Chinese Ind. Code	U.S. Ind. Code
6	1, 2, 7, 8, 9	25	39
7	12	26	29
8	13	27	28
9	10	28	28
10	10	29	28
11	14	30	30
13	14	31	30
14	20	32	32
15	20	33	33
16	20	34	33
17	21	35	34
18	22	36	35
19	23	37	36
20	31	39	37
21	24	40	36
22	25	41	36
23	26	42	38
24	27	43	39

Note: Matched by the author.

Summary of statistics

In this section I will present summary of statistics for the variable that will be used in the main regressions.

Table 2.16: Summary of statistics

VARIABLES	(1) N	(2) Mean	(3) S.D.	(4) Min	(5) Max
$\Delta\ln(MR)$	106	-0.19	0.15	-0.73	0.14
$\Delta\ln(GDP_{p.c.})$	106	0.63	0.15	-0.29	1.38
$\Delta\ln(AOT)$	106	0.13	0.15	-0.76	0.66
$\Delta\ln(Employment)$	106	-0.08	0.14	-0.74	0.43
$\Delta\ln(PopulationDensity)$	106	0.03	0.05	-0.44	0.94
$\ln(MR)_{2000}$	106	1.76	0.14	1.38	2.18
$\ln(GDP_{p.c.})_{2000}$	106	8.72	0.59	7.73	10.55
$\ln(AOT)_{2000}$	106	-0.73	0.33	-1.95	-0.26
$\Delta\ln(Export^t)$	106	1.24	0.85	-1.69	4.38
$\Delta\ln(Export^p)$	105	1.06	0.97	-1.69	5.65
$\Delta\ln(Export^t/Sales^t) = \Delta\ln(E\%)^t$	106	0.15	0.67	-2.19	2.52
$\Delta\ln(Export^p/Sales^p) = \Delta\ln(E\%)^p$	105	-0.08	0.91	-2.28	5.32
$Price^t$	106	0.10	0.06	-0.02	0.31
$Price^{p,I}$	106	0.09	0.05	-0.02	0.24
$Cost^t$	106	0.09	0.04	-0.01	0.30
$Cost^{p,I}$	106	0.10	0.05	-0.00	0.34
$Price^{p,so2}$	106.00	1.80	1.17	0.07	9.40
$Price^{p,soot}$	106.00	0.04	0.03	0.00	0.22
$Cost^{p,so2}$	106.00	0.07	0.06	-0.00	0.35
$Cost^{p,soot}$	106.00	3.23	2.41	-0.06	12.32

CHAPTER 3

Hukou and Labor Misallocation in China

3.1 Introduction

Economic theories argue that efficient allocation of production factors across different production units maximizes the overall output. However, in real-life, different kinds of technological and institutional frictions create obstacles that impede factor mobility, and result in misallocation. Misallocation can happen between firms (e.g., Hsieh and Klenow (2009)) or between regions within a country (e.g., Fajgelbaum et al. (2015)). Among various sources of resources misallocation, one important source is restrictions on labor mobility.[1] For example, Hsieh and Moretti (2017) shows that the land regulations in productive U.S. cities make housing prices too high such that not enough people move into these cities. They find that stringent housing supply restrictions lowered aggregate U.S. growth by more than 50% from 1964 to 2009.

In China, the Hukou stystem has been linked to spatial disparities in income (Tombe and Zhu (2015) and Wang and Zuo (1999)). China's Hukou system is the internal registration system for Chinese citizens. Each individual has a Hukou status associated with a location and a sector (agricultural vs. non-agricultural) based on parents' status. To switch sector or prefecture, an individual needs to obtain a temporary Hukou enabling legal migrant status. Even legal migrants are subject to diminished access to public services such as medical insurance or public

[1] See Restuccia and Rogerson (2008, 2017) and Hopenhayn (2014) for surveys of the literature.

schools. By creating large costs of migration through formal administrative and bureaucratic measures, the Hukou system impedes labor mobility across regions and creates spatial misallocation in China. In a developing country such as China where the average income is lower than in the United States and the variance of the income across regions is bigger, the cost of misallocation for the economic growth can be even larger.[2]

First, I propose to quantify the Hukou-induced cost of misallocation in terms of total output. Following Hsieh and Moretti (2017), I define misallocation of labor by dispersion of wages. I use a general equilibrium Rosen-Roback model to measure the effect of Hukou regulations on aggregate growth. There are three major differences. First, instead of focusing on housing supply restrictions, I focus on amenity supply restrictions, which is the relevant constraint in the Hukou system. Second, while in Hsieh and Moretti (2017), the housing supply elasticity is measured in 2007 with survey data, I collect a new dataset on Chinese prefecture-level migrant regulations from 1995 to 2015, where I document and measure the changes in Hukou regulations. With this data, I can quantify the changes in the model-implied amenity levels attributable to observed changes in Hukou regulations. Third, the U.S. case focuses on the misallocation of labor across different U.S. regions; the Chinese case involves both the movement of urban workers between prefectures and the movement of workers from rural areas to the urban areas. Thus, my work will also feature a urbanization dimension which is more relevant in a developing country context.

Second, I plan to study misallocation resulting from the decentralized political decision making. While the Chinese central government indeed choose to make some regulatory decisions on the national level, such as the cancellation of

[2] In 2000, the Chinese mean wage was one thousand dollars, and the U.S. mean wage was 29 thousand dollars. The standard deviation of average wages across Chinese prefectures was 0.32 times the mean, while the U.S. standard deviation was 0.22 times the mean. See details in Table 3.1.

the agricultural tax in 2006, the Hukou reform is designed to be decentralized. Currently, different Chinese prefectures make their own decisions on Hukou regulations, which depend on the objective function of prefecture-level government officials. Thus, the observed Hukou regulation changes can be different from the choices by a social-planner on the central government level. This decentralization is not unique to the Chinese context: Hsieh and Moretti (2017) discusses the housing supply regulations by U.S. city governments, and Fajgelbaum et al. (2015) focuses on the distortion in factor allocations due to state taxes. I would like to quantify the loss in output and in welfare due to the decentralized Hukou reform.

3.2 Motivating Facts: Wage Dispersion between Chinese Prefectures

3.2.1 Benchmarking the U.S. Case

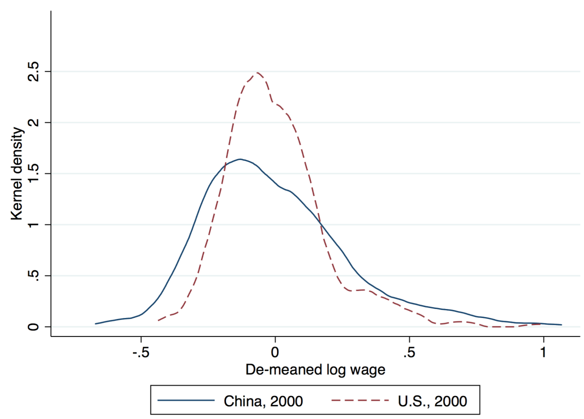
I first show that the pattern of spatial dispersion of wages in the United States and in China in 2000. In Figure 3.1, blue line is the kernel density of de-meaned log wages across 262 Chinese prefectures in 2000, and the red line is across 315 U.S. metropolitan statistical areas (MSAs).[3] We can see that the Chinese wage distribution is more spread-out than the U.S. one: the U.S. density concentrates more around zero, while the Chinese density has more mass on both the left tail and the right tail. Similar results are shown in Table 3.1. The standard-deviation-to-mean ratio is 0.32 for Chinese prefecture-level wages, and it is 0.22 for U.S. MSA-level wages.

While the efficient allocation of labor dictates the equalization of wages across regions, the more spread-out the distribution is, the larger misallocation there

[3] I use "city" and "prefecture" interchangeably in later texts.

is. Thus, we have reasons to believe that the size of spatial misallocation of labor is higher in China than in the United States. This is consistent with the institutional background that labor is freely mobile in the United States, and in China, the Hukou system creates big obstacles to inter-regional migration.

Figure 3.1: Distributions of de-meaned log wages across U.S. MSAs and Chinese prefectures, 2000



Note: The Chinese wage data is from the 2000 Prefecture Statistics Yearbook, with 262 prefectures in total. The U.S. wage data is from the 2000 County Business Pattern with 315 MSAs. The curves are kernel densities of log wages (subtracted by the mean of log wages).

Table 3.1: Summary of Statistics of Wages in 2000, United States vs China, in 2000 USD

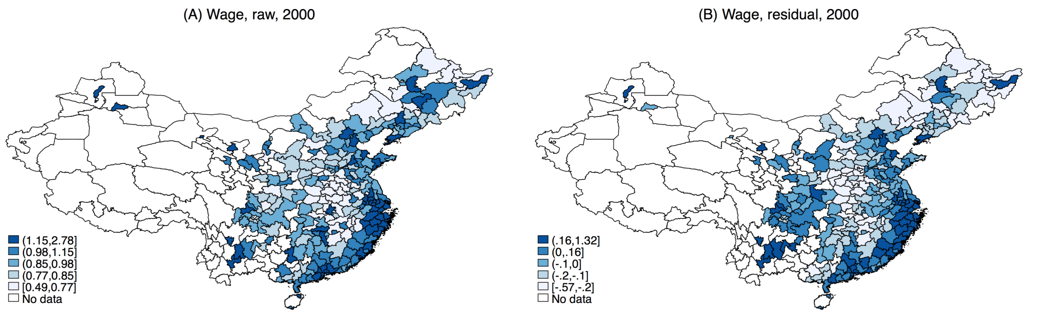
Country	Mean (Std)	Median	Min	Max	Std/Mean	p90/p10
China	1.00 (0.32)	0.92	0.49	2.78	0.32	1.92
United States	29.07 (6.36)	27.80	18.42	76.82	0.22	1.54

Note: the Chinese wage data is from the 2000 Prefecture Statistics Yearbooks. The U.S. wage data is from the 2000 County Business Pattern with 315 MSAs. RMB to USD exchange rate was 8.2784 in 2000.

3.2.2 Geographical Distribution of Wages, 2000

Figure 3.2 Panel A shows the geographical dispersion of average wages across Chinese prefectures in 2000. Darker colors represents higher wages. Coastal prefectures feature high wages of 1.15 to 2.78 thousand dollars, while some inland prefectures have wages of 0.49 to 0.77 thousand dollars. To address the concern that wage differences represents underlying skill differences, I regress wages on years of schooling and plot the residuals in Figure 3.2 Panel B, and the patterns remain largely unchanged.

Figure 3.2: Distributions of average wages across Chinese prefectures, 2000

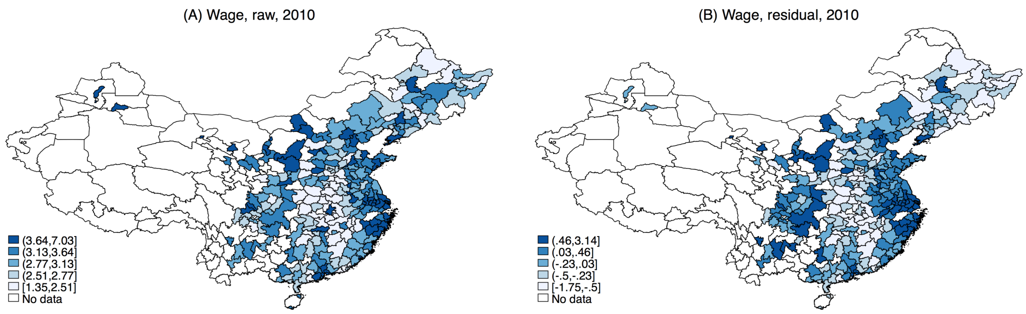


Note: Wage data is from Prefecture Statistics Yearbooks, and average years of education is from the 2000 Population Census. Panel (A) plots the distribution of prefecture-level wages. Panel (B) plots the residual of wages regressed on average years of education. There are 262 prefectures with wage data in 2000. Wages are in 1,000 USD. RMB to USD exchange rate was 8.2784 in 2000.

3.2.3 Changes from 2000 to 2010

Similar geographical distribution of wages can be found in 2010, as shown in Figure 3.3 Panel A and B. However, the dispersion is reduced. The standard-deviation-to-mean ratio is 0.24, and the wage ratio of the 90th percentile over the 10th percentile prefecture declines to 1.73: both are smaller than the 2000 Chinese numbers, and are closer to the 2000 U.S. benchmark. (Table 3.2) In addition, the wage advantage of coastal prefecture in southeastern China seems to become small while northern prefectures catch up.

Figure 3.3: Distributions of average wages across Chinese prefectures, 2010



Note: Wage data is from Prefecture Statistics Yearbooks, and average years of education is from the 2010 Population Census. Panel (A) plots the distribution of prefecture-level wages. Panel (B) plots the residual of wages regressed on average years of education. Wages are in 1,000 yuan and in 2000 values. The price deflator between 2000 and 2010 is from the website of the Chinese Bureau of Statistics website. There are 287 prefectures with wage data in 2010.

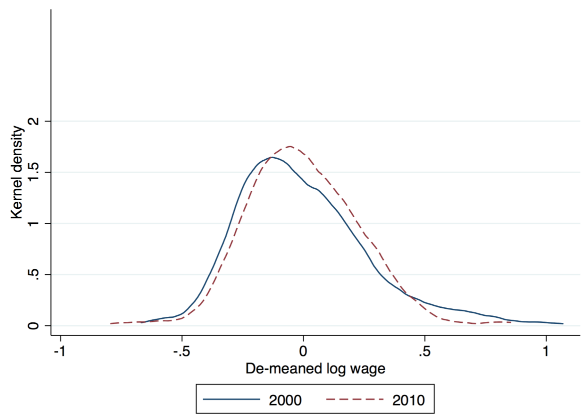
Table 3.2: Summary of Statistics of Wages, 2000 and 2010

Year	Mean (Std)	Median	Min	Max	Std/Mean	p90/p10
2000	1.00 (0.32)	0.92	0.49	2.78	0.32	1.92
2010	3.06 (0.73)	2.93	1.35	7.03	0.24	1.73

Note: Wage data is from Prefecture Statistics Yearbooks, 2000 and 2010. Wages are in 1,000 USD and in 2000 values. The price deflator between 2000 and 2010 is from the website of the Chinese Bureau of Statistics website.

Figure 3.4 shows the shift of distribution of de-meaned log wages across Chinese prefectures from 2000 to 2010. We can see that the distribution shifts to be more centered around 0 in 2010, while the 2000 distribution has a bigger mass on the left side on the right tail, meaning that in 2000 there are more prefectures with very big sizes and more prefectures with sizes right below the mean.

Figure 3.4: Distributions of de-meaned log wages across Chinese prefectures, 2000 and 2010



Note: Wage data is from the 2000 and 2010 Prefecture Statistics Yearbooks. Wages are in 1,000 dollars and in 2000 values. There are 262 prefectures with wage data in 2000. The curves are kernel densities of log wages (subtracted by the mean of log wages).

3.3 Hukou Reform and Changes in Misallocation

3.3.1 The Hukou Reform: Institutional Background and a New Data Source

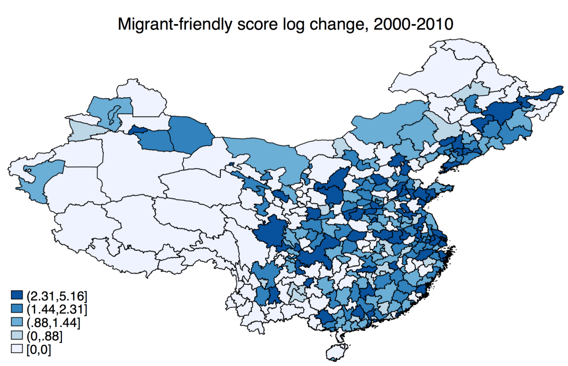
To quantify the effect of Hukou restrictions on misallocation, I use an institutional feature and collect a new dataset on Hukou regulations. While the Hukou restrictions were uniformly strict across prefectures before 2000, the central government in China started to reform the Hukou system in 2000. Local governments were allowed some discretion to design their own reforms, and prefectures started to carry out measures to improve the well-being of legal migrants and set up a pathway for some migrants to get local Hukou.[4] These regulation changes increased migrant workers' access to local amenities and made the cities more migrant-friendly. Thus, the more pro-migrant regulations lead to a more relaxed Hukou system.

[4]According to a 2001 document by the State Council of China, "Local governments should take into consideration local economic and social development levels and conduct reforms that balance population growth, infrastructure, employment and social security, and other welfare programs." Source: www.gov.cn/zhengce/content/2016-09/22/content_5110816.htm.

To measure the changes in the Hukou regulation, I collect prefecture-level government regulation documents from the fee-for-service website www.pkulaw.com. Topics of the regulations span wage payment, vocational training, social insurances, and administrative issues. I evaluate the migrant-friendliness of the regulations by creating an index in between of -2 to 2 for each item. Then I sum all the item indices by prefecture to generate a prefecture-level index or score.[5] The regulation score measures a prefecture's migrant friendliness, and it should be an increasing function of migrant amenities.

Figure 3.5 depicts the distribution of changes in migrant-friendliness of prefectures from 2000 to 2010. Darker blue ones indicate bigger increase in migrant friendliness of a prefecture. Overall, coastal prefectures made bigger changes than inland prefectures. Southeastern prefectures seemed to have more uniform changes, while northern ones had a larger variation.

Figure 3.5: Prefecture-level changes in Hukou regulations/migrant friendliness, 2000-2010



Note: Each prefecture's migrant friendliness in a year is the sum of scores of all migrant-related regulations enacted up to that year. "log changes" are actually changes in inverse hyperbolic-sine transformed regulations scores.

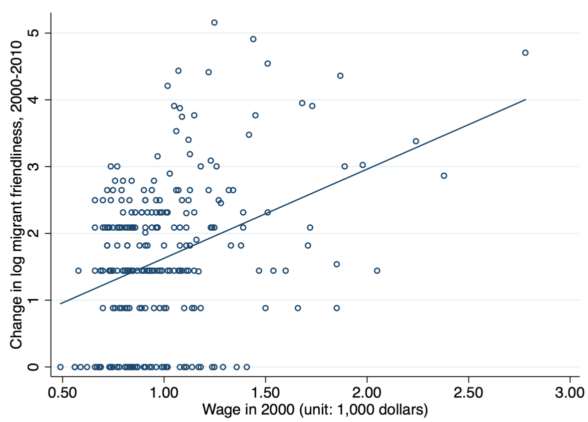
3.3.2 Connecting the Hukou Reform with the Initial Wages

Suppose that the wage dispersion is the result of different regulations on migrant amenities. If in this situation, prefectures with initially higher wages relax the

[5]See Tian (2018) for details on how I constructed the index.

Hukou system more and become more migrant-friendly, the overall wage dispersion will be smaller. This will result in smaller misallocation. Figure 3.6 shows that it is indeed the case.

Figure 3.6: Initial wage levels in 2000 and changes in migrant friendliness from 2000 to 2010



Note: Each dot is a prefecture. Wage data is from the 2000 Prefecture Statistics Yearbook. Wages are in 1,000 dollars and in 2000 values. There are 262 prefectures with wage data in 2000. Changes in log migrant friendliness is the same as in Figure 3.5.

3.4 Quantifying the Misallocation Cost: Replication of Hsieh and Moretti (2017)

3.4.1 Model Setup

I will first follow the theoretical model in Hsieh and Moretti (2017) closely. City i produces a homogeneous good using a Cobb-Douglas production function:

$$Y_i = A_i L_i^{\alpha} K_i^{\eta} T_i^{1-\alpha-\eta},$$

where A_i is the total factor of productivity, L_i is labor, K_i is capital, and T_i is the amount of land available for production. Capital is freely mobile on the international market and gives a return of R. Labor is freely mobile across cities, and the indirect utility of a worker in City i is:

$$V = \frac{W_i Z_i}{P_i^{\beta}},$$

where W_i is the wage, Z_i is the amount of amenity, P_i is the housing price, and β is the share of expenditures on housing. However, since the focus of my exercise is not the housing market, I model the local housing price in a simpler way:

$$P_i = L_i^{\gamma},$$

where γ is the inverse elasticity of housing supply with respect to the number of workers in all cities. The amenity level of a city is:

$$Z_i = \bar{Z}_i L_i^{\delta_i},$$

where $\bar{Z}_i$ is the fixed amenity level of a city that does not change with the number of workers, and δ_i is the inverse elasticity of amenity supply with respect to the number of workers. Cities with a limited amount of public infrastructure or stringent migration policies (i.e. low migrant-friendliness) have a lower elasticity of amenity supply (a larger δ_i). Increases in the number of workers in a city have larger effects on amenity levels when the elasticity of amenity supply is smaller (δ_i is large).

Solving the model, equilibrium employment in a city is given by:

$$L_i = \left[\frac{\alpha^{1-\eta}\eta^{\eta}}{R^{\eta}V^{1-\eta}} A_i T_i^{1-\alpha-\eta} \left(\bar{Z}_i\right)^{1-\eta} \right]^{\frac{1}{1-\alpha-\eta+\beta\gamma-\delta_i(1-\eta)}}.$$

After imposing the national labor market clearing condition $L = \sum_i L_i$ and normalizing L to be one, aggregate output $Y \equiv \sum_i Y_i$ is given by:

$$Y = \left(\frac{\eta}{R}\right)^{\frac{\eta}{1-\eta}} \left[\sum_i A_i^{\frac{1}{1-\alpha-\eta}} T_i \left(\frac{\bar{Q}}{Q_i}\right)^{\frac{1-\eta}{1-\alpha-\eta}} \right]^{\frac{1-\alpha-\eta}{1-\eta}},$$

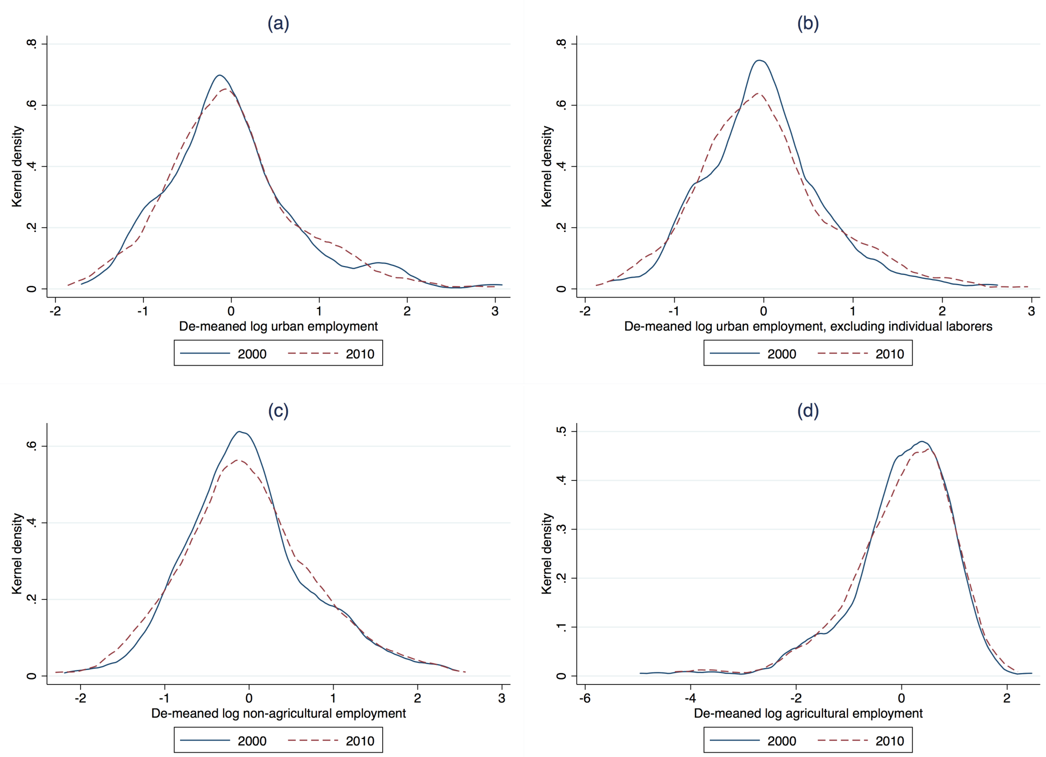
where $Q_i \equiv \frac{P_i^\beta}{Z_i}$ is the ratio of the housing price P_i to local amenities Z_i, and $\bar{Q} \equiv \sum_i L_i Q_i$ is the employment weighted average Q_i across all cities. Q_i will be referred to as the "local price". Since $\frac{1-\eta}{1-\alpha-\eta} > 1$, a mean preserving spread of local prices lowers aggregate output. Conversely, when the variance of local prices increases and the mean keeps the same, aggregate output will increase.

3.4.2 Stylized Facts

Employment I plot the kernel density of de-meaned log employment in 2000 and 2010 in Figure 3.7. The employment data by city comes from two sources: the Prefecture Statistics Yearbooks have total urban employment (Panel (a)) and urban employment excluding individual laborers (Panel (b)); the 2000 and 2010 censuses have total employment by agricultural (Panel (d)) and non-agricultural sector (Panel (c)). All four panels include 261 prefectures observed in both 2000 and 2010 and in both datasets.

Panels (a)(b) and (c) convey the same information: from 2000 to 2010, the distribution of employment across cities became more spread-out. The density around 0 decreased and overall there are more cities on the left and the right tail.

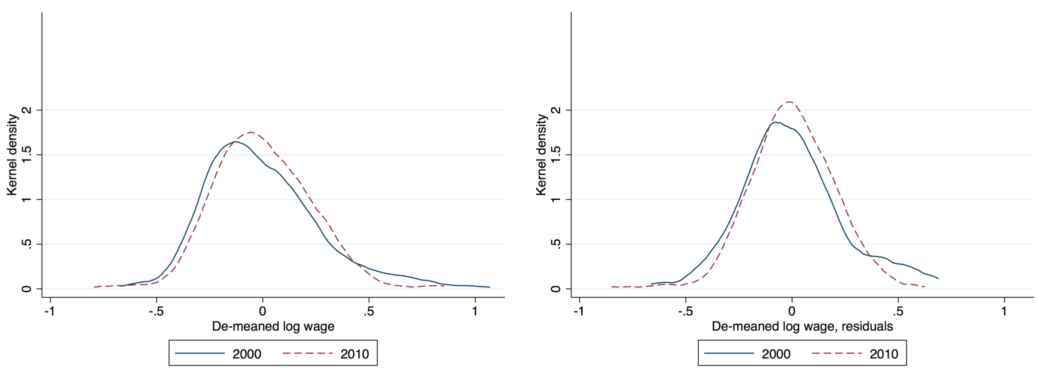
Figure 3.7: Distributions of de-meaned log employment across Chinese prefectures, 2000 and 2010



Note: Urban employment data is from the 2000 and 2010 Prefecture Statistics Yearbooks. Agricultural and non-agricultural employments are from the 2000 and 2010 Population Census. The curves are kernel densities of log employment (subtracted by the mean of log employments). The sample includes 261 prefectures observed in both 2000 and 2010.

Wages The bigger the spatial wage dispersion, the higher the misallocation in output. I plot Kernel densities of de-meaned log wages in 2000 and 2010 in Figure 3.8. The left graph uses the raw wage data from the 2000 and 2010 Prefecture Statistics Yearbooks, and right graph plots residuals of de-meaned log wages regressed on average years of education and male-female population ratios. The message is the same in the two graphs: in 2010, the wage distribution is less spread-out than in 2000, and there is more density around 0 and less density especially on the right tail of the distribution.

Figure 3.8: Distributions of de-meaned log wages across Chinese prefectures, 2000 and 2010



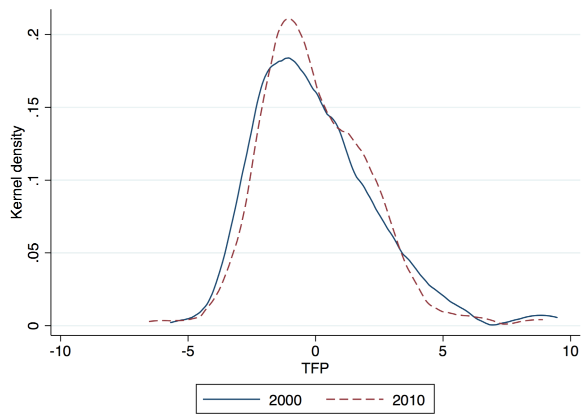
Note: Wage data is from the 2000 and 2010 Prefecture Statistics Yearbooks. Wages are in 1,000 dollars and in 2000 values. The left curves are kernel densities of log wages (subtracted by the mean of log wages). The right curves are kernel densities of the residuals of regressions of de-meaned log wages regressed on male-female ratio and average years of education. The sample includes 261 prefectures observed in both 2000 and 2010.

TFP Local TFP can be solved as a function of local employment and wages:

$$A_i^{\frac{1}{1-\alpha-\eta}} T_i \propto L_i W_i^{\frac{1-\eta}{1-\alpha-\eta}}.$$

I follow Hsieh and Moretti (2017) and use $\alpha = 0.65$ and $\eta = 0.25$. The distributions of log TFPs across Chinese prefectures in 2000 and 2010 are plotted in Figure 3.9. Overall, the TFP distribution became more centered around 0 in 2010 than in 2000. Specifically, the right tail of the TFP distribution in 2000 was much fatter, while in 2010 the density increased more around 0.

Figure 3.9: Distributions of log TFPs across Chinese prefectures, 2000 and 2010



Note: : Local TFP is defined as $A_i^{\frac{1}{1-\alpha-\eta}} T_i$. The graph shows the distribution of de-meaned log local TFP in the relevant year. The sample includes 261 prefectures observed in both 2000 and 2010.

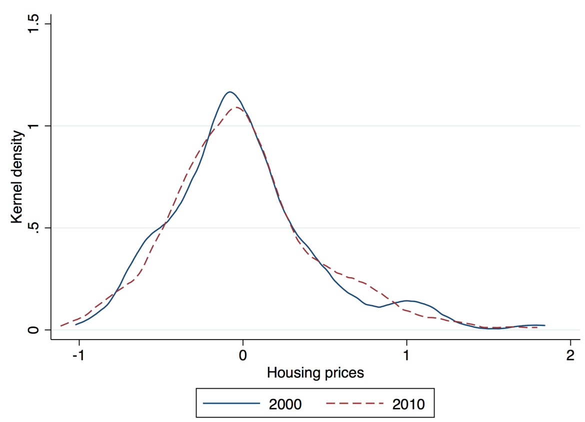
Housing Prices Due to data limitation, I back out housing price data using the model and observed employment data.[6] The housing price of a city is

$$P_i = L_i^{\gamma},$$

where I take $\gamma = \frac{1}{1.67}$, and 1.67 is the housing supply elasticity of a median city in the United States in Saiz (2010). The distributions of log housing prices across Chinese prefectures in 2000 and in 2010 are plotted in Figure 3.10. The overall shape of the housing price distribution is similar to the employment distribution, while the change from 2000 is 2010 is slightly smaller since $\gamma < 1$.

[6] Only 67 big cities report housing prices.

Figure 3.10: Distributions of log housing prices across Chinese prefectures, 2000 and 2010



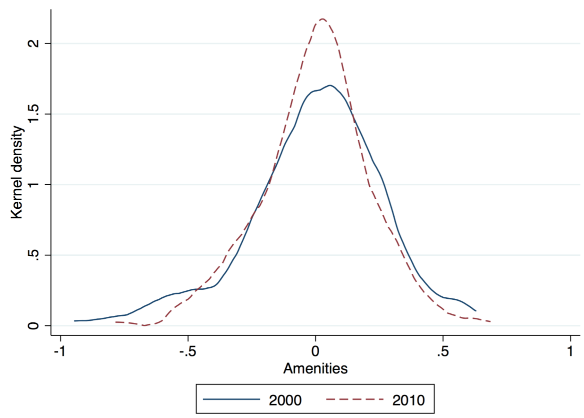
Note: : Local housing price P_i is defined as $L_i^{1/1.67}$, where 1.67 is the housing supply elasticity of a median U.S. city as in Saiz (2010), and L_i is the total employment. The graph shows the distribution of de-meaned log local housing prices in the relevant year. The sample includes 261 prefectures observed in both 2000 and 2010.

Amenities With the housing price data and wage data, I solve local amenities as the following:

$$Z_i = \frac{P_i^\beta}{W_i}.$$

Figure 3.11 plots the distributions of log amenities across Chinese prefectures in 2000 and 2010. Overall, the distribution of amenities became less spread-out in 2010 than in 2000. This indicates that Chinese prefectures are more equalized in amenities in 2010, and given similar distribution of housing prices, wages are also more equalized.

Figure 3.11: Distributions of log amenities across Chinese prefectures, 2000 and 2010



Note: : Local amenity is defined as $P_i^{0.32}/W_i$, where 0.32 is the share of household expenditure on housing, P_i is the housing price, and W_i is the wage. The graph shows the distribution of de-meaned log local amenities in the relevant year. The sample includes 261 prefectures observed in both 2000 and 2010.

3.4.3 Decomposing Aggregate Growth

I will proceed in three steps to decompose the aggregate growth into local effects.

First, I estimate how local growth contributes to aggregate growth through the lenses of the model and compare it to the accounting estimate. In an accounting sense, aggregate growth can be decomposed into local growth: $g_t = \frac{Y_t}{Y_{t-1}} = \frac{\sum_i Y_{i,t}}{Y_{t-1}} = \sum_i \frac{Y_{i,t}}{Y_{i,t-1}} \frac{Y_{i,t-1}}{Y_{t-1}} = \sum_i \frac{Y_{i,t-1}}{Y_{t-1}} g_{i,t}$, and city i's contribution to aggregate growth is the share of output in city i times the growth rate of city i's GDP from period t to t'. With the model structure, I allow city i's local TFP, local amenities and housing prices change as in the data while holding these factors in all other cities constant at (1) period $t-1$ level; (2) period t level. Then I solve for wages and employment endogenously for (1) and (2) in both periods and calculate the aggregate growth rate. City i's contribution to aggregate growth is the geometric mean of growth rates in (1) and (2).

Second, I isolate the effect of housing price growth and amenity changes on aggregate growth. I hold the distribution of local prices $Q_i = \frac{P_i^\beta}{Z_i}$ in all Chinese cities constant at the 2000 level and allow the local TFPs to change as in the data. I then solve for the allocation of employment and corresponding wages in 2010 and calculate the aggregate output growth rate. I also compute the counterfactual growth rate (1) holding the housing prices constant and allow TFPs and amenities to change as in the data; (2) holding the amenity level constant and allow TFPs and housing prices to change as in the data.

To estimate each individual city's contribution to aggregate growth due to amenity changes, I hold city i's amenity at its 2000 level, and allow amenity levels in all other cities and TFP and housing prices in all cities to change as in the data. Then I solve for the 2010 wages and employment and calculate the counterfactual aggregate growth.

Third, I estimate how the changes in amenities are driven by amenity supply elasticities.

3.4.4 Preliminary Evidence on the Connection between Hukou Regulation and Amenities

With the estimated changes in log amenity, I test whether it is correlated with Hukou regulation changes observed in the data. In Table 3.3, I regress changes in log model-implied amenity levels on changes in migrant friendliness from 2000 to 2010, weighted by the total employment in 2000. Changes in migrant friendliness is measure as the change in inverse hyperbolic-sine transformed regulation score. Column (1) shows a positive coefficient of 0.032, indicating that one unit change in migrant friendliness is correlated with 3.2% increase in the amenity level. I control for log population size in 2000 and log employment in 2000 in Column (2) and (3), respectively, and the estimates are similar.

Overall, I find a positive correlation between the changes in model-implied amenity levels and changes in migrant friendliness. Of course, the changes in migrant friendliness is not exogenous. Tian (2018) shows that local governments relax migration regulations more in places with bigger trade shocks, indicating that the TFP shocks and changes in amenity supply elasticities are correlated.[7] I will investigate other possible ways to identify the regulation change effect.

Table 3.3: Correlation between model-implied change in log amenity and change in migrant friendliness

Changes in log amenity, model-implied	(1)	(2)	(3)
Changes in migrant friendliness, 2000-2010	0.032**	0.038**	0.054***
	(0.014)	(0.015)	(0.014)
Log population size, 2000		-0.038	
		(0.029)	
Log employment, 2000			-0.073***
			(0.017)
Constant	-1.25***	-1.031***	-0.979***
	(0.034)	(0.169)	(0.058)
Observations	259	259	259
R-squared	0.03	0.05	0.1

Note: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

3.5 Additional Thoughts

3.5.1 Urbanization

In Hsieh and Moretti (2017), the mobility of labor is across different MSAs in the United States. However, in the Chinese case, the movement of labor can be both across urban areas different prefectures and from rural areas to urban areas, or from the agricultural sector to the non-agricultural sector. If I use a city-sector as the unit of measurement, then the misallocation is across more units than in previous analysis.

I plot the de-meaned log employment across Chinese prefectures in 2000 and

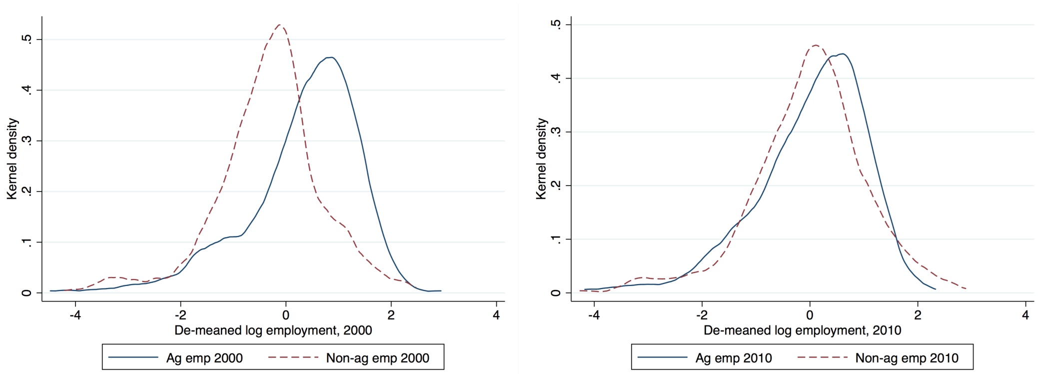
[7] Trade shocks affect the price of good produced in a city and is equivalent to a TFP shock.

2010 in Figure 3.12. The data is from the 2000 and 2010 population censuses, and there are overall 365 prefectures, divided into an agricultural sector and a non-agricultural sector. On the left graph, the distribution of agricultural sector employment is to the right of the distribution of non-agricultural sector employment, indicating that there are overall more people employed in the agricultural sector than the non-agricultural sector. The two distributions are much closer in 2010. This is consistent with the increased urbanization rate in China: it is 35.39% in 2000 and 49.68% in 2010 according to the censuses.

One important aspect of the Hukou reform is the promotion of the rural-to-urban migration. Previously, rural workers were not able to work and live in urban areas and the main source of income is agricultural production. With the relaxed Hukou system, rural workers can migrant seasonally to urban areas and complement their income with wages. Thus, the misallocation due to the rural-urban divide could be larger than the prefecture-to-prefecture divide.

The difficulty in using the city-sector measurement is that it is usually hard to measure wages earned from agricultural production. I will try to supplement the urban wage data with Rural Household Survey data and Chinese Household Income Project data to quantify the gains from urbanization.

Figure 3.12: Distributions of de-meaned log employment across Chinese prefectures, 2000 and 2010



Note: Agricultural and non-agricultural employments are from the 2000 and 2010 Population Census. The curves are kernel densities of log employment (subtracted by the mean of log employments). The sample includes 365 prefectures observed in both 2000 and 2010.

3.5.2 Calibrating Model Parameters Using Chinese Data

The model parameters for Chinese prefectures can be different from the U.S. ones. I plan to calibrate the parameters with corresponding Chinese data. One important caveat of my exercise is that I don't observe housing prices for all cities in 2000 and 2010. I follow Zhang et al. (2017) to construct housing price data using the perpetual inventory method and combine information from the 2005 1% Population Survey, the housing price index and information on new home sales.

I also plan to estimate the production function parameters α and η using the Annual Survey of Manufacturing Firms and Prefecture Statistics Yearbooks. Expenditure share of housing β can be estimated with Urban Household Survey data. Amenity supply elasticity can be estimated using the Prefecture Statistics Yearbook and the Hukou regulation data.

Table 3.4: Parameter Calibration

Parameter	Hsieh and Moretti (2017) Source	Chinese Data Source
α and η	BEA (2013); Karabarbounis and Neiman (2013)	Annual Survey of Manufacturing Firms
	Piketty (2014)	Prefecture Statistics Yearbook
β	Albouy (2008)	Urban Household Survey
γ	Saiz (2010)	Zhang et al. (2017)
δ	-	Prefecture Statistics Yearbook
		Hukou Regulation from Tian (2018)

3.6 Conclusion

The Hukou system in China creates obstacles in labor mobility and has been linked to spatial wage dispersion. The dispersion in wages across Chinese prefecture decreased from 2000 to 2010. Thus, potential spatial misallocation of labor decreased. During the same time period, Hukou regulations became more migrant-friendly across prefectures when the central government granted local discretion in Hukou reform. In this paper, I first propose to quantify the reduction in the misallocation cost due to the Hukou reform. Second, I propose to investigate the potential misallocation due to the decentralized reform, since local governments' choice of regulations may not coincide with the decision made by a planner on the central government level.

Bibliography

Daron Acemoglu, Simon Johnson, and James Robinson. The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth. American Economic Review, 95(3):546–579, 2005. doi: 10.1257/0002828054201305. URL http://www.aeaweb.org/articles.php?doi=10.1257/0002828054201305.

David Albouy. Are Big Cities Bad Places to Live? Estimating Quality of Life Across Metropolitan Areas. Technical report, National Bureau of Economic Research, 2008.

George Alessandria and Alain Delacroix. Trade and the (Dis)Incentive to Reform Labor Markets: The Case of Reform in the European Union. Journal of International Economics, 75(1):151 – 166, 2008. ISSN 0022-1996. doi: https://doi.org/10.1016/j.jinteco.2008.02.001. URL http://www.sciencedirect.com/science/article/pii/S0022199608000111.

Douglas Almond and Janet Currie. Killing Me Softly: The Fetal Origins Hypothesis. The Journal of Economic Perspectives, 25(3):153–172, 2011.

Mary Amiti, Mi Dai, Robert C. Feenstra, and John Romalis. How Did China's WTO Entry Benefit U.S. Consumers? Working Paper 23487, National Bureau of Economic Research, June 2017. URL http://www.nber.org/papers/w23487.

Werner Antweiler, Brian R. Copeland, and M. Scott Taylor. Is Free Trade Good for the Environment? American Economic Review, 91(4):877–908, 2001. doi: 10.1257/aer.91.4.877. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.91.4.877.

Eva Arceo, Rema Hanna, and Paulina Oliva. Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from

Mexico City. The Economic Journal, pages 257–280, 2015. ISSN 1468-0297. doi: 10.1111/ecoj.12273. URL http://dx.doi.org/10.1111/ecoj.12273.

David Autor, David Dorn, Gordon Hanson, and Kaveh Majlesi. Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure. Working Paper 22637, National Bureau of Economic Research, September 2016. URL http://www.nber.org/papers/w22637.

David H. Autor, David Dorn, and Gordon H. Hanson. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. American Economic Review, 103(6):2121–68, 2013. doi: 10.1257/aer.103.6.2121. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.103.6.2121.

BEA. National Economic Accounts. 2013.

Brian Beach and W Walker Hanlon. Coal Smoke and Mortality in an Early Industrial Economy. The Economic Journal, 2017.

Gary S. Becker, Tomas J. Philipson, and Rodrigo R. Soares. The Quantity and Quality of Life and the Evolution of World Inequality. American Economic Review, 95(1):277–291, 2005. doi: 10.1257/0002828053828563. URL http://www.aeaweb.org/articles.php?doi=10.1257/0002828053828563.

Timothy Besley and Robin Burgess. Can Labor Regulation Hinder Economic Performance? Evidence from India. The Quarterly Journal of Economics, 119 (1):91–134, 2004.

George J. Borjas. The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. The Quarterly Journal of Economics, 118(4):1335–1374, 2003. doi: 10.1162/003355303322552810. URL +http://dx.doi.org/10.1162/003355303322552810.

Loren Brandt, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang. WTO Accession and Performance of Chinese Manufacturing Firms. American Economic Review, 107(9):2784–2820, September 2017. doi: 10.1257/aer.20121266. URL http://www.aeaweb.org/articles?id=10.1257/aer.20121266.

Lorenzo Caliendo, Luca David Opromolla, Fernando Parro, and Alessandro Sforza. Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement. Working Paper 23695, National Bureau of Economic Research, August 2017. URL http://www.nber.org/papers/w23695.

David Card. The Impact of the Mariel Boatlift on the Miami Labor Market. Industrial and Labor Relations Review, 43(2):245–257, 1990. ISSN 00197939, 2162271X. URL http://www.jstor.org/stable/2523702.

David Card. Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. Journal of Labor Economics, 19(1):22–64, 2001. ISSN 0734306X, 15375307. URL http://www.jstor.org/stable/10.1086/209979.

Piyush Chandra and Cheryl Long. VAT Rebates and Export Performance in China: Firm-level Evidence. Journal of Public Economics, 102:13–22, 2013.

Kenneth Y. Chay and Michael Greenstone. The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession. The Quarterly Journal of Economics, 118(3):1121–1167, 2003. doi: 10.1162/00335530360698513. URL http://qje.oxfordjournals.org/content/118/3/1121.abstract.

Yuyu Chen, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy. Proceedings of the National Academy of Sciences, 110(32):12936–12941, 2013. doi: 10.1073/pnas.1300018110. URL http://www.pnas.org/content/110/32/12936.abstract.

Brian R. Copeland and M. Scott Taylor. Trade, Growth, and the Environment. Journal of Economic Literature, 42(1):7–71, 2004. doi: 10.1257/002205104773558047. URL http://www.aeaweb.org/articles.php?doi=10.1257/002205104773558047.

Christian Dippel, Robert Gold, Stephan Heblich, and Rodrigo Pinto. Instrumental Variables and Causal Mechanisms: Unpacking The Effect of Trade on Workers and Voters. Working Paper 23209, National Bureau of Economic Research, March 2017. URL http://www.nber.org/papers/w23209.

Bilge Erten and Jessica Leight. Exporting out of Agriculture: The Impact of WTO Accession on Structural Transformation in China. Technical report, mimeo, 2017.

Giovanni Facchini, Maggie Y Liu, Anna Maria Mayda, Minghai Zhou, et al. China's "Great Migration": The impact of the reduction in trade policy uncertainty. Technical report, CEPR Discussion Papers, 2018.

Pablo D. Fajgelbaum, Eduardo Morales, Juan Carlos SuÃ¡rez Serrato, and Owen M. Zidar. State Taxes and Spatial Misallocation. Working Paper 21760, National Bureau of Economic Research, November 2015. URL http://www.nber.org/papers/w21760.

Jingting Fan. Internal Geography, Labor Mobility, and the Distributional Impacts of Trade. Working paper, January 2015. URL http://dx.doi.org/10.2139/ssrn.2573873.

Andrew Foster, Emilio Gutierrez, and Naresh Kumar. Voluntary Compliance, Pollution Levels, and Infant Mortality in Mexico. The American Economic Review, pages 191–197, 2009.

Dalia Ghanem and Junjie Zhang. Effortless Perfection: Do Chinese Cities Manipulate Air Pollution Data? Journal of Environmental Economics and Management, 68(2):203 – 225, 2014. ISSN 0095-0696. doi: http://dx.doi.org/10.1016/j.jeem.2014.05.003. URL http://www.sciencedirect.com/science/article/pii/S0095069614000400.

Pinelopi K. Goldberg and Nina Pavcnik. The Effects of Trade Policy. Working Paper 21957, National Bureau of Economic Research, February 2016.

Roger H Gordon and Wei Li. Provincial and Local Governments in China: Fiscal Institutions and Government Behavior. In Capitalizing China, pages 337–369. University of Chicago Press, 2012.

Michael Greenstone and B Kelsey Jack. Envirodevonomics: A Research Agenda for an Emerging Field. Journal of Economic Literature, 53(1):5–42, 2015.

Emilio Gutierrez. Using Satellite Imagery to Measure the Relationship between Air Quality and Infant Mortality: an Empirical Study for Mexico. Population and Environment, 31(4):203–222, 2010.

Kyle Handley and Nuno Limão. Policy Uncertainty, Trade, and Welfare: Theory and Evidence for China and the United States. American Economic Review, 107(9):2731–83, September 2017. doi: 10.1257/aer.20141419. URL http://www.aeaweb.org/articles?id=10.1257/aer.20141419.

W Walker Hanlon and Antonio Miscio. Agglomeration: A long-run Panel Data Approach. Journal of Urban Economics, 99:1–14, 2017.

W Walker Hanlon and Yuan Tian. Killer Cities: Past and Present. American Economic Review, 105(5):570–75, 2015.

Elhanan Helpman and Oleg Itskhoki. Labour Market Rigidities, Trade and Unemployment. The Review of Economic Studies, 77(3):1100–1137, 2010.

J Vernon Henderson, Adam Storeygard, and David N Weil. Measuring Economic Growth from Outer Space. The American Economic Review, 102(2):994–1028, 2012.

Hugo A Hopenhayn. Firms, Misallocation, and Aggregate Productivity: A Review. Annual Review of Economics, 6(1):735–770, 2014.

Chang-Tai Hsieh and Peter J Klenow. Misallocation and Manufacturing TFP in China and India. The Quarterly Journal of Economics, 124(4):1403–1448, 2009.

Chang-Tai Hsieh and Enrico Moretti. Housing Constraints and Spatial Misallocation. 2017.

Kosuke Imai, Luke Keele, Dustin Tingley, and Teppei Yamamoto. Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies. American Political Science Review, 105(4):765–789, 2011.

Ruixue Jia. Pollution for Promotion. Working paper, March 2017a.

Ruixue Jia. Pollution for Promotion. Working paper, 2017b.

Ruixue Jia, Masayuki Kudamatsu, and David Seim. Political Selection in China: the Complementary Roles of Connections and Performance. Journal of the European Economic Association, 13(4):631–668, 2015. ISSN 1542-4774. doi: 10.1111/jeea.12124. URL http://dx.doi.org/10.1111/jeea.12124.

Ronald W Jones. Income Distribution and Effective Protection in a Multicommodity Trade Model. Journal of Economic Theory, 11(1):1–15, 1975.

Gueorgui Kambourov. Labour Market Regulations and the Sectoral Reallocation of Workers: The Case of Trade Reforms. The Review of Economic Studies, 76 (4):1321–1358, 2009.

Loukas Karabarbounis and Brent Neiman. The Global Decline of the Labor Share. The Quarterly Journal of Economics, 129(1):61–103, 2013.

Amit K. Khandelwal, Peter K. Schott, and Shang-Jin Wei. Trade Liberalization and Embedded Institutional Reform: Evidence from Chinese Exporters. American Economic Review, 103(6):2169–95, 2013. doi: 10.1257/aer.103.6.2169. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.103.6.2169.

Brian K. Kovak. Regional Effects of Trade Reform: What Is the Correct Measure of Liberalization? American Economic Review, 103(5):1960–76, 2013. doi: 10.1257/aer.103.5.1960. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.103.5.1960.

Jongkwan Lee, Giovanni Peri, and Vasil Yasenov. The Employment Effects of Mexican Repatriations: Evidence from the 1930's. Working Paper 23885, National Bureau of Economic Research, September 2017. URL http://www.nber.org/papers/w23885.

Andrei A. Levchenko. International Trade and Institutional Change. The Journal of Law, Economics, and Organization, 29(5):1145–1181, 2013. doi: 10.1093/jleo/ews008. URL +http://dx.doi.org/10.1093/jleo/ews008.

Hongbin Li and Li-An Zhou. Political Turnover and Economic Performance: the Incentive Role of Personnel Control in China. Journal of Public Economics, 89(9):1743 – 1762, 2005. ISSN 0047-2727. doi: http://dx.doi.org/10.1016/j.jpubeco.2004.06.009. URL http://www.sciencedirect.com/science/article/pii/S0047272704001355.

Ying Long, Jianghao Wang, Kang Wu, and Junjie Zhang. Population Exposure to Ambient PM 2.5 at the Subdistrict Level in China. Available at SSRN 2486602, 2014.

Lin Ma and Yang Tang. Geography, Trade, and Internal Migration in China, 2016.

Kaivan Munshi and Mark Rosenzweig. Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap. The American Economic Review, 106(1):46–98, 2016.

Gianmarco I. P. Ottaviano and Giovanni Peri. Rethinking the Effect of Immigration on Wage. Journal of the European Economic Association, 10(1): 152–197, 2012. ISSN 1542-4774. doi: 10.1111/j.1542-4774.2011.01052.x. URL http://dx.doi.org/10.1111/j.1542-4774.2011.01052.x.

JR Pierce and PK Schott. Concording US Harmonized System Codes Over Time. Journal of Official Statistics, 2009.

Thomas Piketty. Capital in the 21st Century. 2014.

Diego Restuccia and Richard Rogerson. Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. Review of Economic Dynamics, 11(4):707–720, 2008.

Diego Restuccia and Richard Rogerson. The Causes and Costs of Misallocation. Journal of Economic Perspectives, 31(3):151–74, 2017.

Albert Saiz. The Geographic Determinants of Housing Supply. The Quarterly Journal of Economics, 125(3):1253–1296, 2010.

Yuan Tian. International Trade Liberalization and Domestic Institutional Reform: Effects of WTO Accession on Chinese Internal Migration Policy. 2018.

Trevor Tombe and Xiaodong Zhu. Trade, Migration and Productivity: A Quantitative Analysis of China. Manuscript, University of Toronto, 2015.

Petia Topalova. Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India. American Economic Journal: Applied

Economics, 2(4):1–41, 2010. doi: 10.1257/app.2.4.1. URL http://www.aeaweb.org/articles.php?doi=10.1257/app.2.4.1.

Feng Wang and Xuejin Zuo. Inside China's Cities: Institutional Barriers and Opportunities for Urban Migrants. The American Economic Review, 89(2): 276–280, 1999.

Jun Wang and Sundar A Christopher. Intercomparison between Satellite-derived Aerosol Optical Thickness and PM2. 5 Mass: Implications for Air Quality Studies. Geophysical research letters, 30(21), 2003.

John Douglas Wilson. Theories of Tax Competition. National Tax Journal, pages 269–304, 1999.

Yang Yang. Transport Infrastructure, City Productivity Growth and Sectoral Reallocation: Evidence from China. Working paper, January 2017.

Yu Zhang et al. Liquidity Constraints, Transition Dynamics, and the Chinese Housing Return Premium. In 2017 Meeting Papers, number 1217. Society for Economic Dynamics, 2017.

Yuan Zi. Trade Liberalization and the Great Labor Reallocation. Working paper, December 2017. URL ssrn.com/abstract=2877106.