University of Nottingham
UK | CHINA | MALAYSIA

research paper series

China and the World Economy Programme

Research Paper 2020/16

Trade-induced urbanization and the making of modern agriculture

Yuan Tian, Junjie Xia, and Rudai Yang

Trade-Induced Urbanization and the Making of Modern Agriculture

Yuan Tian, Junjie Xia, and Rudai Yang*

October 12, 2020

Abstract

Manufacturing growth can benefit the agricultural sector if the outflow of labor from agriculture improves land allocation efficiency and facilitates capital adoption. Using destination prefectures' trade shocks in the manufacturing sector driven by China's accession to the World Trade Organization (WTO) and the origin village's initial internal migration network, we construct the exposure to manufacturing trade shocks for a panel of 295 villages from 2001 to 2010. We find that villages with larger increases in trade exposure had larger increases in the share of non-agricultural laborers, more fluid local land markets, and faster modernization of production through the adoption of agricultural machinery. Village-level agricultural productivity improved through the allocation of land towards more productive farmers within a village. During the era we study, transaction costs declined in the agricultural land market. We use a quantitative model to show that the growth in non-agricultural productivity had a larger impact on urbanization and agricultural modernization than reductions in transaction costs.

JEL codes: F16, J24, O14, O4, Q12, Q15.

Keywords: Land Misallocation, Capital Adoption, Urbanization, Trade Liberalization, Agriculture Modernization

*Tian: University of Nottingham, School of Economics (email: yuan.tian1@nottingham.ac.uk); Xia: Peking University, Institute of New Structural Economics (email: junjiexia@nsd.pku.edu.cn); Yang: Peking University, School of Economics (email: rdyang@pku.edu.cn). We are grateful to Pablo Fajgelbaum, Brian Kovak, and Adriana Lleras-Muney for support and guidance. We received valuable feedback from Daniel Berkowitz, Lee Branstetter, Chaoran Chen, Yi Chen, Karen Clay, Matthew Delventhal, Jingting Fan, Martin Gaynor, Akshaya Jha, Rebecca Lessem, Bingjing Li, Daniel Nagin, Michael Poyker, Yogita Shamdasani, Allison Shertzer, Lowell Taylor, Xiaodong Zhu, and participants at various conferences and seminars. All errors are our own.

1 Introduction

No country has achieved a high level of income without a sharp reduction in agricultural employment and a concomitant modernization of agricultural production.[1] The economic literature on structural transformation hypothesizes two broad explanations on the drivers of this process: i) technological innovation in the agricultural sector, and ii) increased attractiveness of urban areas and the manufacturing sector.[2] Despite rich theoretical discussions of these two channels, empirical studies of structural transformation are scarce and focus mostly on the first channel.[3]

We fill this gap in the literature by providing empirical evidence on how structural transformation is initiated by manufacturing growth and reinforced by the development process in agriculture along a range of dimensions, including movements out of agricultural production, increased efficiency of land allocation, and machinery adoption. Using a unique dataset on agricultural households in China, we study how the manufacturing growth induced by international trade liberalization affected agricultural production in the context of China's WTO accession. The pull forces driving migration out of rural areas came from the reduction in tariffs faced by manufacturing exporters, which resulted in an increase in labor demand in urban areas. We measure a migrant destination prefecture's exposure to manufacturing trade as a weighted average of industry-level tariffs faced by Chinese exporters in their destination markets, with local industry-level employment as weights. Then, we connect an origin village with its migrant destination prefectures using the initial prefecture-to-prefecture migration network. A village was more exposed to manufacturing trade if its destination prefectures on average faced larger declines in tariffs in export markets.

We show that increased labor demand in the urban manufacturing sector led origin villages to shift employment from agriculture to non-agriculture. Rural land markets became thicker, and rental activity increased. Land allocation efficiency also improved; initially more productive households operated larger farms, especially in villages that experienced larger destination trade shocks. We also find that the villages facing larger trade shocks had faster modernization of production through the adoption of agricultural machinery and increased village-level agricultural productivity.

As in many other developing countries, rural land markets in China are incomplete, with high transaction costs of land leasing between farmers.[4] The initial land distribution was mostly proportional to household size, no matter how productive an individual or a house was in crop farming. Yet, before 2000 less than 4% of rural households rented out their land, suggesting that the land allocation efficiency was low. This situation improved with the substantial growth in manufacturing trade and fast urbanization after 2000. The share of rural labor working in the non-agricultural

[1] See Lewis [1954]; Schultz et al. [1964]; Taylor and Martin [2001]; Lucas [2004]; and Akram et al. [2017], among others.

[2] See Baumol [1967]; Murphy et al. [1989]; Kongsamut et al. [2001]; Gollin et al. [2002]; Ngai and Pissarides [2007]; and Yang and Zhu [2013], among others. Other explanations include the decline in the relative cost of obtaining non-agricultural skills, as shown in Caselli and Coleman [2001].

[3] Empirical studies on the channel of technological innovations in the agricultural sector include Foster and Rosenzweig [2004, 2007]; Nunn and Qian [2011]; Michaels et al. [2012]; Hornbeck and Keskin [2015]; and Bustos et al. [2016].

[4] See Adamopoulos and Restuccia [2014]; de Janvry et al. [2015]; Adamopoulos et al. [2017]; Chen [2017]; and Restuccia and Santaeulalia-Llopis [2017].

sector increased from 16% in 2000 to 25% in 2010, and the share of households with land rental income increased to 13%.

We use cross-sectional variation in the reduction of manufacturing tariffs to generate shocks to pull-factors for out-migration. Identifying the causal effect of out-migration on agricultural production is generally difficult. First, increases in agricultural productivity can lead to more out-migration if the technological change is labor-saving (Bustos et al. 2016). Second, expansion of rural transportation networks can incentivize out-migration *and* improve agricultural productivity through improved input quality. Third, economy-wide productivity shocks can generate correlated productivity growth in the manufacturing sector and the agricultural sector, and labor can flow out of agriculture if the manufacturing productivity growth dominates. We overcome these identification challenges by using the strong forces incentivizing internal migration generated by China's WTO accession in 2001 (Zi 2020; Facchini et al. 2019; and Tian 2020). Prefectures experienced differential shocks in the manufacturing sector due to different industrial compositions and different industry-level tariff reductions. These destination trade shocks were transmitted to origin villages through the initial migration network. Intuitively, a village was more affected by the trade shock in its major migration destinations, and villages differed in their initial distribution of migration destinations, due to different travel costs and other migration costs, such as cultural distances.

Our main data set is a nationally representative sample of rural households and villages from the National Fixed Point (NFP) Survey, with information on agricultural production and rural household living arrangements. The survey collects information on a panel of around 20,000 Chinese rural households in 300 villages. We use the 2001–2010 data for the main analysis and the 1995–2001 data to rule out confounding pre-trends.[5] We observe household-level occupation choices, land-in-operation, the amount of land rented from other households, and land transactions during the year. Using detailed information on output values, labor, capital, land, and intermediate inputs, we calculate household-level and village-level total factor productivity (TFP) in crop farming.

We employ a shift-share measure to construct a village's exposure to destination prefectures' trade shocks. Following the standard method in the literature on the local labor market effect of trade liberalization (Topalova 2010; Kovak 2013; McCaig and Pavcnik 2018; and Tian 2020), we measure a destination prefecture's exposure to manufacturing trade as a weighted average of industry-level output tariffs faced by exporters, with each industry's employment share as weights. A reduction in the average tariff faced by a prefecture acted as a positive demand shock for the goods produced in the prefecture, and it resulted in an increase in labor demand in the manufacturing sector. An origin village's exposure to manufacturing growth is measured as the interaction of its initial migration network and the destination prefectures' trade exposures. The prefecture-to-prefecture migration network is constructed using a sample of one million individuals from the 2000 population census, where an individual's residence prefecture in 1995 and current residence prefecture are observed.

[5] This is the best available dataset on agriculture production and rural households in China. Chari et al. [Forthcoming], Adamopoulos et al. [2017], and Kinnan et al. [2018], among others, have used different segments of the dataset. To our knowledge, we are the first to use the full sample of households for a time period that spans before and after China's international trade liberalization.

These output tariffs were *imposed by importing countries on Chinese exports*, and the industry-level post-WTO tariff reductions were uncorrelated with pre-WTO tariff changes and export growth. A prefecture's tariff reduction was uncorrelated with its pre-WTO wage and GDP growth (Tian 2020). We show that empirically, the change in the trade exposure of an origin village from 2001 to 2010 was uncorrelated with changes in the out-migration rate, the share of land leased, the value of agricultural machinery, and TFP in the 1995–2001 period.

We first show the impact of an origin village's trade exposure on rural workers' occupation choice. A one-standard-deviation larger decline in destination prefectures' output tariffs resulted in a three-percentage-point (or a 0.21-standard-deviation) larger increase in the share of non-agricultural labor in the origin village. The result is robust to (1) including agricultural trade shocks that could potentially be correlated with manufacturing trade shocks and affect the agricultural production directly, (2) using alternative measures of out-migration collected at the village level.

The results also imply larger effects for places that were less far along in the process of urbanization and agriculture modernization. We find that the effect was larger for villages with less land per agricultural worker in 2001. The initial land-to-labor ratio in agriculture was positively correlated with village characteristics that were pro-reallocation, such as the land market fluidity and the non-agricultural labor share, and was negatively correlated with factors that impeded land consolidation, such as the ruggedness. Villages with a smaller land-to-labor ratio also had a larger correlation of the output value and TFP across households, which is a direct measure for allocation efficiency. Hence, regions with a smaller land-to-labor ratio had larger factor misallocation at the beginning of the period, and they reacted more strongly to out-migration shocks. This finding is consistent with the observed irreversibility of urbanization; as pointed out in Lucas [2004], "this transition is an irreversible process that every industrializing society undergoes once and only once."

We then investigate the effect of trade exposure on the rural land market. The land rental market became more active in the face of the trade shock. A one-standard-deviation larger decline in the destination prefectures' output tariffs led to a 26% (or a 0.16-standard-deviation) larger increase in the stock of land leased, a 76% (or a 0.47-standard-deviation) larger increase in the flow of land leased within the year, and larger increases in the income from land rental. The effect was bigger for villages in prefectures where a larger share of migrants moved between prefectures, since the exposure to the destination prefectures' tariffs were larger for these villages. For villages with a between-prefecture migrant share that was sufficiently large (larger than 58%), we find an increased number of households with land larger than one-third hectare (the sample median land size).

We also provide household-level evidence on land reallocation. The shift of land from unproductive farmers to productive farms was stronger in villages that experienced larger shocks. Villages facing a one-standard-deviation larger tariff decline in migrant-destination prefectures had a 20% larger elasticity of land to TFP at the household level, indicating that the land allocation efficiency improved more in villages more exposed to the trade shocks.[6]

[6] We measure household-level productivity instead of plot-level productivity since we can only follow households over time, not plots. An alternative story could be that the plots are heterogeneous in land productivity and TFP is higher in the most efficient plots. When the outside option of non-agriculture increases, the workers abandon the

Declines in destination prefectures' output tariffs encouraged the adoption of agricultural machinery, especially for villages with more between-prefecture migration. Evaluated at the mean of the cross-prefecture migrant share, a one-standard-deviation larger decline in destination prefectures' tariff led to a 8% (or a 0.05-standard-deviation) larger increase in the value of agricultural machinery. This finding is consistent with Hornbeck and Naidu [2014] and Clemens et al. [2018] with out-migration shocks resulting from natural disasters and labor market policies, respectively. Capital adoption reinforces the urbanization process: once machines are in place, the production is not likely to revert back to the labor-intensive mode.[7]

There are several explanations for the increase in agricultural machinery. First, farmers may substitute labor with capital when labor costs increase (Manuelli and Seshadri 2014). Second, if there are scale-dependent returns to mechanization due to larger contiguous land areas, farmers may adopt machinery only when the land size is sufficiently large (Foster and Rosenzweig 2011, 2017). Third, the reduction in land misallocation can also lead to increased capital adoption, since productive agricultural households are able to lease in more land and use more capital. Fourth, migrant remittances can ease the household credit constraint and facilitate capital adoption. We find evidence supporting the first three explanations, but not the last one.

We additionally show that a one-standard-deviation larger decline in other prefectures' tariff led to a 30% (or a 0.33-standard-deviation) larger increase in the output weighted village-level TFP. The effect came from improved allocation efficiency, with increasingly more output produced by initially more productive households in villages with larger trade shocks. We do not find important roles for switching to high-value cash crops or husbandry. Overall, we find trade shocks inducing households to move out all forms of agricultural production.

To further understand the village-level TFP effect, we provide additional evidence on the selection into out-migration. Using individual-level information from 2003 to 2008, we show that non-agricultural income was positively correlated with factors such as an individual's education level and having non-agricultural occupational training, but that these factors were not good predictors of agricultural productivity. Initially unproductive farming households were more likely to leave agriculture, and they responded more strongly to the trade shocks.

We focus on the output tariff shocks since they generate shocks to pull factors of out-migration for origin villages; we do not intend to characterize the full impact of the WTO accession on the Chinese economy. The WTO accession affected manufacturing trade in China through multiple channels, including the reduction in output tariffs (Tian 2020), reduction in input tariffs (Zi 2020 and Brandt et al. 2017), and the reduction in trade uncertainty induced by the establishment of the U.S.–China permanent normal trade relationship (PNTR)(Pierce and Schott 2016; Handley and Limão 2017; Facchini et al. 2019; and Erten and Leight, Forthcoming). We focus on the reduction in output tariffs

village and go to manufacturing, wages in rural areas go up, so the less productive land plots shut down and are reallocated to more productive plots. This story involves no inefficiency and features Melitz-style reallocation among efficient producers. Our empirical evidence is not consistent with this story since when farmers move out of the villages, the land transactions and rents increase instead of decrease.

[7] Similar transitional paths can be found in the manufacturing sector. See, for example, Acemoglu and Restrepo [2017].

since the output tariff reduction generates intuitive demand shocks to the manufacturing sector. Yet, when we empirically evaluate the effects of the PNTR shock, we find effects of similar magnitudes on the land market, capital, and TFP, and insignificant effects on the occupation choice. The PNTR shock is approximately orthogonal to the output tariff shocks, in the sense that controlling for the PNTR shock does not affect the coefficient estimates of output tariff shocks (also see Handley and Limão 2017 and Tian 2020).

Finally, we present a simple two-sector open-economy model with land market transaction costs to investigate the role of migrant selection, land market frictions, and sectoral productivity shocks in determining the patterns of urbanization and agricultural modernization. We calibrate the model parameters using the NFP Survey data and conduct quantitative exercises. Overall, consistent with the empirical results, we find a mild positive correlation between an individual's agricultural and non-agricultural productivity (0.17). The land market transaction costs were substantial and declined slightly from 2001 to 2010. In this time period, the economy experienced much faster growth in the non-agricultural productivity than in the agricultural sector. Despite the large transaction costs in the agricultural land market, we find that reducing the transaction costs have smaller impacts on urbanization and agricultural modernization than the growth of non-agricultural productivity.[8]

The paper contributes to several strands of literature. First is the literature on structural transformation. Existing research mostly focuses on its initial causes: productivity growth in the agricultural sector (Ngai and Pissarides 2007 and Bustos et al. 2016), declines in the prices of agricultural machinery (Yang and Zhu 2013), and declines in relative cost of obtaining non-agricultural skills (Caselli and Coleman 2001). To our knowlege, this paper is the first to demonstrate that international trade in manufacturing accelerates the development process in agricultural sectors that are not directly exporters. Second, there is an extensive literature on how land misallocation affects agricultural productivity. The ambiguity of land rights in developing countries limits land reallocation and creates misallocation (Adamopoulos et al. 2017), and land reforms that clarify land rights can improve land allocation efficiency and increase agricultural productivity (de Janvry et al. 2015 and Chari et al. Forthcoming). We complement this literature by showing that in addition to the institutional barriers, the lack of land leasing activity is partially due to the fact that farmers usually do not have good outside options in non-agricultural sectors. The opportunity to work in the manufacturing sector can create a more fluid land market, potentially reducing misallocation. Third, the existing literature on out-migration of rural residents focuses on its effect on self-insurance (Kinnan et al. 2018), participation in local risk-sharing networks (Morten 2019), individual migrant outcomes (Johnson and Taylor 2019), and rural labor markets (Akram et al. 2017 and Dinkelman et al. 2017). We add to the literature by showing the first-order effect of out-migration on the organization of agricultural production. We demonstrate that when out-migration is not motivated

[8] Our model follows Adamopoulos et al. [2017] closely, while we explicitly model the source of the land market misallocation as the transaction cost of leasing. In our model, reductions in the land market transaction cost increase the land rental price and unproductive farmers leave agriculture, improving the allocation efficiency. Faster growth in the non-agriculture productivity increases the opportunity cost of remaining in agriculture, and also draws the unproductive farmers out of agriculture. These mechanisms complement Lagakos and Waugh [2013], where the subsistence food requirement keeps unproductive farmers in agriculture.

by income smoothing when facing temporary shocks, but as a part of the structural transformation, factor markets adjust, leading to productivity changes. Fourth, the paper contributes to the literature on the impact of international trade liberalization on the Chinese economy. Most literature discuss the impact of trade liberalization on manufacturing productivity and exports (Khandelwal et al. 2013; Brandt et al. 2017; Handley and Limão 2017, among others), migration (Zi 2020 and Facchini et al. 2019), sectoral employment shifts (Erten and Leight, Forthcoming), and reforms in labor institutions (Tian 2020). We show that manufacturing trade affects *the development process of the agricultural sector* through land reallocation, capital investment, and worker selection.

The rest of the paper is organized as follows. Section 2 provides background information on the land market and the WTO-induced trade shocks in China. Section 3 presents the key data sources and relevant measurement. Section 4 presents motivating facts linking out-migration to land reallocation and agricultural efficiency. Section 5 shows main empirical findings. Section 6 discusses the selection pattern. Section 7 presents the model and quantitative exercises. The last section concludes.

2 Background

2.1 The Rural Land Market in China

Land market conditions are an important aspect of the agricultural sector, since land is an essential input in agricultural production, and developing countries usually suffer from weak land right protection. Since the establishment of the Household Contract Responsible System (HCRS) in the early 1980s, Chinese agricultural land has been collectively owned by the village and contracted to households within the village commune. The initial land distribution was set mainly based on household sizes at the time of HCRS establishment.[9] The first round of contracts was set with a length of about 15 years and they were extended for another 30 years around 1998.[10] Households have the right to use the contracted land for agricultural production, and the right of leasing the land to other households within the same village commune was legalized in 1988.[11]

However, the rural land market remained very thin before the 2000s. The missing rural land market created misallocation of land across households, given the initial egalitarian distribution rule. Despite *de jure* tenure security, within-village land reallocation happened, and village leaders had discretion in the reallocation.[12] In addition, institutions that supported dispute resolution were

[9]Unfortunately, the NFP data does not have the contracted land size, but only land in agricultural operation. Rozelle et al. [2002] provides evidence on the distribution rule using survey data.

[10]The HCRS reform started in some regions in 1978 and was completed across the country in 1983. A 1993 regulation at the national level proposed the 30-year extension of the initial contract, and a 1995 regulation formalized the extension. In the 2003 *Law of the People's Republic of China on the Contracting of Rural Land*, the legal requirement for the length contract of farm land is 30 years.

[11]See the full description of timing of the reforms in Appendix A.2. For a summary of reforms of the rural land system, see this article by the Ministry of Agriculture and Rural Affairs: `www.moa.gov.cn/ztzl/xczx/rsgt/201812/t20181228_6165784.htm`.

[12]Rozelle et al. [2002] documents that in the 215 villages they surveyed, between 1983 and 1995, slightly less than 10% of all villages experienced land reallocation per year, and a reallocation entailed around half of the village's land

absent.[13] This insecurity of land rights decreased the incentive of land rental, since households feared that they could lose their land in the next round of land allocation if they didn't work on their own land (Benjamin and Brandt 2002; Rozelle et al. 2002; Adamopoulos et al. 2017). If the farmers did not have stable means of living other than agricultural production, they might not be willing to lease their land to other households even if other households were more productive, due to the perceived "use it or lose it" rule.

The land leasing market became more active alongside with urbanization, since the opportunity cost of remaining in crop farming became higher.[14] According to decennial population censuses, the share of population living in urban areas increased from 26% in 1990 to 36% in 2000, and then to 50% in 2010. Meanwhile, the share of households with land lease income increased from about 4% in the 1990s to 12% in 2010.[15]

2.2 Internal Migration and the WTO Shock

Urbanization came alongside with large flows of internal migration. The Chinese household registration system, i.e., the Hukou system, assigns all residents with a prefecture-sector-specific registration status, where a sector is either agriculture or non-agriculture. A person is an internal migrant when living and working in a prefecture-sector different from their registration. In 2000, 11% of the population were migrants, and the number increased to 20% in 2010. Migration is closely tied with sectoral employment shifts: 93% of migrants work in the non-agricultural sector.[16]

China's accession to the WTO was an important driver of the fast growth in the manufacturing sector and increased internal migration (Brandt et al. 2017, Zi 2020, Facchini et al. 2019, and Tian 2020). The WTO accession reduced international trade barriers, with the average tariff on Chinese exports declined from 3.7% in 2001 to 2.4% in 2010 (Figure 1, the line with solid dots). The average tariff is the weighted average across industries, using industry export values as weights. Industry level tariffs are calculated as the weighted average of tariffs on Chinese exports imposed by importing countries, using the 2001 export values as weights.[17] Manufacturing exports from China increased from less than 400 trillion dollars in 2001 to 1,750 trillions in 2010 (Figure 1, the line with hollow

and involved two-thirds of households.

[13]The *Law of the People's Republic of China on the Mediation and Arbitration of Rural Land Contract Disputes* was enacted in 2010. Before that, several regulations issued by the central government tried to address the issues of contract disputes starting from 1992.

[14]In addition, with more households leaving the agricultural sector, the local institution or norm on land allocation can change. For example, if only 10 rural households move to other regions or sectors in a village with 300 households, the norm would be that their land should be reallocated to other households. However, if 60 households make the move, the norm might change to accommodate the new situation. Thus, the effect of urbanization on the size of the rural land market can be higher if the transaction cost of land leasing decreases with the number of household due to coordination issues. However, without observing the transaction costs directly, we cannot directly test the hypothesis on changes in norms.

[15]Calculated using the NFP Survey.

[16]Calculated using the 2000 and 2010 population census.

[17]The export elasticity with respect to tariffs can be estimated using a panel regression of log exports on tariffs on the industry-year level, with industry fixed effects and year fixed effect. The elasticity estimate is 0.074 using 26 two-SIC code industries and 2001–2010 data. It indicates that a one-percentage point decline in the output tariff is correlated with a 7.4% increase in exports.

diamonds). This substantially increased demand for labor in the manufacturing sector, and resulted in increased internal migration (Tian 2020).[18]

[Figure 1 about here.]

Regions varied in the extent of the export demand shock based on their initial industrial composition, since the size of tariff reductions were different by industry. Figure 2 uses the 2001–2010 change as an example to show the variation in tariff reductions and export growth by industry. Similar to Tian [2020], we use the reduction in output tariffs on Chinese exports to measure the WTO shock. The reduction in tariffs was due to China's eligibility of the Most-Favored-Nation status, and a reduction in the output tariff effectively increased the output price. Empirically, the industry-level tariff changes in the post-WTO period (2001–2007) were uncorrelated with export growth and tariff changes in the pre-WTO period (1995–2001). We measure the prefecture-level tariff reductions as the weighted average of industry-level tariffs, using industry-level employment shares as weights. The prefecture-level tariff reductions in the post-WTO period were uncorrelated with pre-WTO wage and GDP growth.

[Figure 2 about here.]

Additionally, one might be concerned that the agricultural sector faced tariff changes on its output that were correlated with manufacturing sector tariff changes. Overall, the WTO's direct impact on the agricultural goods market is less clear. China's import tariff on soybeans declined from over 100 percentage points to zero, and soybean imports increased from less than 5 trillion dollars in 2001 to 25 trillion in 2010. For other crops, however, imports and exports fluctuated over time, and there were no clear patterns of tariff changes.[19]

3 Data and Measurement

3.1 Occupation Choice and Migration

In order to measure a rural resident's occupation choice and production activities, we use the NFP Survey, a longitudinal survey conducted by the Research Center for the Rural Economy under the Ministry of Agriculture in China. The survey started in 1986 and continues until now. Multi-stage sampling is used to get a nationally representative sample of around 300 villages and 20,000 households per year. Households and villages are followed with little attrition and are added over time for representativeness.[20] The core module of the household questionnaire remains stable, with household-level demographic summaries, agricultural production, assets, income, and expenditure.

[18] In addition, Tian [2020] shows that regions with more favorable export shocks started to provide more amenities for migrant workers, and these changes further increased the incentive to migrate.

[19] Oil crop, flax, vegetable and fruit experienced large increases in the value of import, although the scale was much smaller than soybean. See trade and tariff trends of major crops in Appendix B.5. In Appendix D.5, we show that our main results are robust to controlling for agricultural trade shocks.

[20] We provide evidence on the absence of selective attrition in Appendix B.2.1.

We use the 2001–2010 data for the main analysis of the post-WTO period and 1995–2001 data to check pre-trends.[21] Our main analysis is based on the household-level information and village-level measures aggregated from household data, and we supplement it with the village questionnaire and individual questionnaire.[22] An administrative village or subdistrict is the lowest level of government administration in China, with county, prefecture, and province as successively more aggregate levels of government.

One key element of our analysis is the definition of a rural resident's occupation. There are broadly three categories: laborer, entrepreneur, and public-sector employee. We focus on laborers who are wage earners: they are employed outside their own household and work for wages.[23] Thus, a wage earner can be (1) working in his own village and employed by other households, (2) employed by firms in his own village, (3) working outside the village, but within the same prefecture, or (4) working in a different prefecture.[24]

Empirically, Case (1) is not prevalent in rural China: hired labor days is on average only 2% of total labor days in any family operations, during the 2001–2010 period. The share of Type (2) workers is also likely to be small. According to the individual-level data between 2003 and 2010, a wage earner spends 212 days working *outside the village* on average. Thus, the majority of the wage earners are working outside the village, either in the urban areas of their own prefecture, or in other prefectures. In addition, the wage earners are predominantly working in non-agriculture. As shown in Appendix B.3 using the individual-level data between 2003 and 2010, 97% of the wage earners work in the non-agriculture sector.

Our main measure for the occupation choice is the non-agricultural labor share of a village, and it is the ratio of the total number of wage earners to the total number of labor in a village, where both numbers are aggregates from the information in the household questionnaire (Table 1). The village questionnaire has information on the number of laborers outside the village and decomposes the number into within-county, between-county, and between-provinces; we use the information in the robustness checks.[25]

[Table 1 about here.]

In order to measure pre-existing migration connections between prefectures, we use the 0.095% individual sample of the 2000 census to construct the prefecture-to-prefecture migration network. The census has information on the residence prefecture in 2000 and the residence prefecture in 1995

[21] The data is not available in 1992 and 1994. After 2003, demographic informations are collected at the individual level, and production information is more detailed, with inputs and outputs information by crop.

[22] The survey intends to get an accurate and consistent picture of agricultural production and rural household life for the support of policy making. For details of the survey disclosed by the survey center, see `http://jiuban.moa.gov.cn/sydw/ncjjzx/gcdgzdt/gzdtg/201302/t20130225_3225848.htm`. Also see Benjamin et al. [2005], Banerjee et al. [2012] for other descriptives of the data.

[23] Entrepreneurs are managers of firms and businesses. Public-sector employees include teachers, medical workers, and civil servants.

[24] A prefecture is composed of rural villages and village-equivalent urban units (town and districts). See the Chinese administrative units in Appendix B.1. There are 333 prefectures, and each prefecture has about 2200 villages.

[25] The village questionnaire is filled by village heads, and we think the village-level aggregates from household questionnaires are of higher quality, which are what we use in the main analysis.

(Table 1). Thus, we know the number of people who lived in prefecture i in 1995 moved to prefecture j before 2000 (m_{ij}).[26]

We also use the census to construct the share of cross-prefecture migrants out of all migrants. The share of cross-prefecture migrants is relevant for our analysis, since an origin village's exposure to other prefectures' trade shocks is larger if initially, a larger share of the village residents were working in those prefectures. The total number of cross-prefecture migrants from prefecture i is $m_i^{between} \equiv \sum_{j \neq i} m_{ij}$. The within-prefecture migrants are identified using the question on the registration place of Hukou. A person is defined as a within-prefecture migrant if his Hukou registration is in the same prefecture, but in different counties. Denoting the number of within-prefecture migrants as m_i^{within}, we calculate the share of cross-prefecture migrants as

$$s_i = \frac{m_i^{between}}{m_i^{between} + m_i^{within}}.$$

Given that prefecture is the finest geographical unit we can get for the migration information, we make the assumption that all villages (v) in a prefecture (i) shares the same migration network that connects other prefectures. If this assumption is violated, it will drive the empirical results toward zero. We also assume that the propensity of cross-prefecture migration is the same for all villages within a prefecture. Thus, $s_{v(i)} = s_i$ for all villages (v) in a prefecture (i).[27]

3.2 Regional Trade Shocks

We first construct the prefecture-year level trade exposure in the manufacturing sector. The output tariff (τ) on exports faced by prefecture i is calculated using applied tariffs from the World Bank TRAINS dataset on the 2-digit SIC level in the manufacturing sector. The output tariff on exports is the weighted average of tariffs on Chinese exports imposed by importing countries, with their 2001 import values as weights. Following Kovak [2013], the regional output tariff in prefecture i and in year t is

$$\tau_{it} = \sum_k \beta_{ik} \tau_{kt},$$

$$\text{where} \qquad \beta_{ik} = \frac{\lambda_{ik} \frac{1}{\theta_{ik}}}{\sum_{k'} \lambda_{ik'} \frac{1}{\theta_{ik'}}},$$

$\lambda_{ik} = \frac{L_{ik}}{\sum_{k'} L_{ik'}}$ is the fraction of regional labor allocated to industry k, and $1 - \theta_{ik}$ is the cost share of labor in industry k. λ_{ik} and θ_{ik} are calculated using the 2000 Industrial Enterprises Survey data,

[26] See Appendix C for descriptives of the migration network.

[27] If this assumption is violated, we have measurement errors in the village-to-prefecture migration network. If the measurement error is classic, the exposure to trade shock will be measured with error, and the effect of trade exposure on the outcomes will be biased toward zero.

and only manufacturing industries are included.[28] τ_{kt} is the industry-year-specific tariff. A village v's (in prefecture i) exposure to its own prefecture's output tariff is

$$\tau_{v(i)t}^{own} = \tau_{it}.$$

Accordingly, its exposure to tariffs in other prefectures is

$$\tau_{v(i)t}^{other} \equiv \sum_{j \neq i} \frac{m_{ij}}{\sum_{j' \neq i} m_{ij'}} \tau_{jt},$$

where m_{ij} is the number of people who are in prefecture i in 1995 and reside in prefecture j in 2000.

Figure 3 shows the geographic distribution of own prefecture's output tariff reduction and other prefectures' output tariff reduction, using the 2001–2010 change as an example. Panel (a) shows the distribution of own prefecture's output tariff reduction, defined as $\tau_{2010}^{own} - \tau_{2001}^{own}$, and Panel (b) shows the distribution of other prefectures' output tariff reduction, $\tau_{2010}^{other} - \tau_{2001}^{other}$. Darker colors represent larger reductions. We can see that the distributions are different for the two types of tariff reductions. The declines in own prefecture's output tariff ranged from –3.96 to –0.14 and were rather scattered across the country. The declines in other prefectures' output tariff ranged from –1.69 to –0.54 and were larger in northern China.

[Figure 3 about here.]

3.3 Total Factor of Productivity (TFP) Estimation

Household TFP In order to track the allocation of land across households with different productivity levels, we construct household-level TFP. We estimate household-level revenue TFP instead of quantity TFP due to data constraints. Crop outputs are available in all years in quantities (kilos), but the input data varies by year. There are four types of inputs in crop farming: land (in hectares), labor (in labor days in agriculture), capital (in initial book value), and intermediate inputs (in value, including seed and seedlings, fertilizers, agricultural films, and pesticides). In our main analysis, we aggregate across crops to generate household-level outputs and inputs to estimate the household-level TFP.

The output value is constructed using a common vector of year-specific crop prices and household-level crop outputs. We include 11 types of crops that are consistently measured in the data: wheat, rice, corn, soy bean, cotton, oil crops, sugar crops, flax, tobacco, fruits, and vegetables.[29] For each crop, we calculate the sales price in yuan per kilo for all households with positive sales. The price of a crop in a particular year is calculated as the national average of all households. Then the household-level output is the sum of crop physical outputs evaluated at the common national prices.

[28] The 2-digit industry codes in the survey is different from the SIC code, and we provide concordance in Appendix B.4.

[29] Crop area of these 11 crops comprises 90% of total areas in our data, both as the sample mean for households and as the aggregate share. Other crops that do not show up in all years include potato, mulberry, tea, herbal medicine.

We deflate the household level output using the national output price indices to make output values comparable across years, using 1995 as the baseline year.[30,31]

We also make adjustments on the input side. The intermediate input value is the total value of inputs in all crops, deflated by province-level agricultural input price indices, using 1995 as the baseline year. The capital stock is recorded in initial book value. To take into account differential prices across years, depreciation of capital stock, and missing values of capital in some observations, we use the perpetual inventory method to reconstruct the capital stock at the household level.[32]

Assuming that agriculture production in crop farming follows a Cobb-Douglas form, we estimate the production function using the following equation:

$$\log(y_{h(v)t}) = \alpha \log(d_{h(v)t}) + \beta \log(k_{h(v)t}) + \gamma \log(l_{h(v)t}) + \delta \log(m_{h(v)t}) + \phi_{h(v)t}, \tag{1}$$

where $y_{h(v)t}$ is the output value in crop farming in household h, village v, and year t. Labor days in agriculture, capital, land, and intermediate inputs are $d_{h(v)t}$, $k_{h(v)t}$, $l_{h(v)t}$, and $m_{h(v)t}$, respectively. Cobb-Douglas parameters α, β, γ, and δ represent output elasticities with respect to each input, and are assumed to be constant over time and across households. We further decompose the log of TFP as follows:

$$\phi_{h(v)t} = \phi_{vt} + \phi_h + e_{h(v)t}.$$

Here, a household's productivity in a given year is comprised of factors common to its village in the year, ϕ_{vt}, such as weather and other aggregate shocks, its intrinsic ability to farm and other time-invariant household-level factors, ϕ_h, and idiosyncratic shocks, $e_{h(v)t}$.

We estimate Equation 1 controlling for village-year (ϕ_{vt}) fixed effects and household (ϕ_h) fixed effects.[33] The log of household TFP is measured as the following residual term,

$$\hat{\phi}_{h(v)t} \equiv \log(y_{h(v)t}) - \hat{\alpha} \log(d_{h(v)t}) - \hat{\beta} \log(k_{h(v)t}) - \hat{\gamma} \log(l_{h(v)t}) - \hat{\delta} \log(m_{h(v)t}).$$

Village TFP We are interested in measuring the village-level TFP since it is informative about the overall productivity of the village and reflects the efficiency of local land allocation, an important aspect of agricultural modernization. The village-level productivity (Φ_{vt}) is constructed as the

[30] The national output price deflator is the price index of crop farming from the National Statistics Yearbook of Agriculture. We use the national price deflator instead of province-level price deflators since the latter is only available after 2003.

[31] We use common prices to eliminate the price variation across households. In addition, we want to evaluate a household's output value even in the absence of crop sales, since the household may consume the output for food or for livestock feed.

[32] See the details of the method in Appendix B.2.3.

[33] See the details of the TFP estimation in Appendix D.2. The estimates for the output elasticity of inputs are similar as in Chow [1993], Cao and Birchenall [2013], and Chari et al. [Forthcoming]. An alternative method is to use the log value-added as the outcome variable (the output value minus the intermediate input value), and the estimated TFP is denoted as $\hat{\phi}^V_{hvt}$. To alleviate the concern that the input choices are correlated with unobserved idiosyncratic productivity shocks, we use the lagged inputs as instruments for the inputs in the current period. The TFP estimates with the IV method are very similar to the OLS estimates, so we use the OLS estimates directly.

weighted average of log household TFPs, using the output value ($y_{h(v)t}$) as weights,

$$\Phi_{vt} \equiv \sum_h w_{h(v)t} \hat{\phi}_{h(v)t} = \sum_h \frac{y_{h(v)t}}{\sum_{h'} y_{h'(v)t}} \hat{\phi}_{h(v)t}.$$

In addition, similar as in Chari et al. [Forthcoming], we decompose the village-level TFP in the following way,

$$\Phi_{vt} = \overline{\phi}_{vt} + \sum_h (w_{h(v)t} - \overline{w}_{vt})(\hat{\phi}_{h(v)t} - \overline{\phi}_{vt}) \equiv \overline{\phi}_{vt} + E_{vt},$$

where $\overline{\phi}_{vt} \equiv \frac{1}{N_h} \sum_h \hat{\phi}_{h(v)t}$ and $\bar{w}_{vt} \equiv \frac{1}{N_h} \sum_h w_{h(v)t} = \frac{1}{N_h}$ represent unweighted means, with N_h as the number of households in the village-year. The second term E_{vt} is the sample covariance between the household-level output weights and productivity multiplied by $N_h - 1$. A larger E_{vt} indicates that the productive household generate more output and have a larger weight in the calculation of the village-level TFP. Thus, we use it as one measure of allocation efficiency.

4 Key Motivating Facts

4.1 Less Labor, More Land Rental, More Capital, Higher Land and Labor Productivity in the Agricultural Sector after 2001

We first present the trends of labor, land, capital markets for the 1995–2010 period, with trend breaks around 2001. First, more and more households moved out of crop farming and started to work for wages (Figure 4 Panel a). The share of households whose main business were crop farming declined from 79% in 1995 to 75% in 2001, and further declined to 66% in 2010. This decline is mirrored by the increase in the share of households where the entire household worked in non-household business: while the 1995 to 2001 change was less than 2%, the share increased by 5% afterwards.[34]

The land rental market became active mostly after 2001 (Panel b). Less than 5% of households had income from land leasing between 1995 and 2001, and the number increased to 13% in 2010. The size of land-lease income also grew a lot in the post-2001 period, from 550 yuan per household to 2500 yuan in 2010.[35]

[Figure 4 about here.]

[34] Households can be either in the family-run business or in non-household business. Family-run business uses households as the unit of operation, relies entirely or mainly on household members' labor supply, utilizes family-owned or contracted factor inputs, directly organizes the production, does accounting independently, and bears its own gains or losses. There are eight categories for the family-run business: crop farming, forestry, husbandry, manufacturing, construction, transportation, service.

[35] Appendix D.1 shows that household occupation choices were correlated with how much land they decided to work on. In a household with three laborers, the probability of working on any land was six percentage points smaller when one more household member worked as a non-agricultural laborer; conditional on non-zero land in agricultural operation, the land size was 25% smaller.

Alongside the outflow of labor and land rental activities, the amount of capital increased. The dotted line in Panel (c) shows that the average value of total capital stock for households in agriculture increased from 4.8 thousand yuan in 2001 to eight thousand yuan in 2010, with a much smaller change before 2001. Similarly, the value of agricultural machinery had an increase of one thousand yuan after 2001, while the before 2001 changes was less than 0.5 thousand yuan.[36]

The patterns of labor and land productivity growth in Panel (d) were consistent with the increased capital input. Land productivity is defined as the log kilos per hectare, and labor productivity is the log kilo per labor day in agriculture. Take wheat as an example, the land productivity remained relatively stable before 2001 and experienced a 0.3 log-point increase afterwards, and the increase of labor productivity was 0.8 log-point after 2001.[37]

Overall, we find that the agricultural sector experienced outflow of labor, increased land rental activities, more capital adoption, and higher land and labor productivity through the 1995–2010 period, and the change was accelerated after 2001.

4.2 Tariff Reductions Led to Increases in Internal Migration

We argue that the outflow of labor from agriculture was closely related to the fast growth of manufacturing exports after 2001. Our main analysis focuses on how increased trade exposures in migration destinations pulled labor out of villages, i.e., from sending regions' perspective. Here, we first provide a mirror evidence from receiving regions' perspective, i.e., how the declines in the output tariff in a region pulled in labor. In Figure 5, each dot is a prefecture. The horizontal axis is the change in output tariffs from 2000 to 2010 ($\tau_{2010}^{own} - \tau_{2001}^{own}$), and the vertical axis is the change in the share of migrants, calculated using the 2000 and 2010 censuses. The slope is -0.016 and statistically significant at the 5% level, indicating that a one-percentage point larger decline in output tariffs in export markets resulted in a 1.6 percentage point larger increase in the share of migrants.

[Figure 5 about here.]

4.3 Larger Trade Shocks, Larger Correlations between Land and TFP within a Village

We show that the land allocation efficiency increased in villages with bigger trade shocks. In Figure 6, we split the villages into two group using the size of the trade shock they experienced from 2001 to 2010. The 2001 to 2010 trade shock is defined as the difference between a village's exposure to other prefectures' output tariff in 2001 and 2010, i.e., $(\tau_{2010}^{other} - \tau_{2001}^{other}) \times s$, where s is the share of cross-prefecture migrant share. The shocks larger than the median magnitude (in absolute values)

[36] Capital in all years are valued at 1995 yuan using the perpetual inventory method (see Appendix B.2.3 for details). There are eight types of capital in the survey: draft animals, hand farm tools valued at least at 50 yuan, agricultural machinery, industrial machinery, transportation machinery, facilities, fixed infrastructure, and other.

[37] On average, a household had 0.5 hectare land in 2001 and 0.48 hectare in 2010. In comparison, a household had 1.8 agricultural laborers in 2001 and 1.3 in 2010. Given the outflow of labor and relative stable total agricultural land, it is reasonable for the labor productivity to increase more than the land productivity. Appendix B.6 shows that the trends were similar for rice, corn, and soybean.

are defined as large shocks, and the shocks smaller than the median are defined as small shocks. Panel (a) shows the correlation between the log land and log TFP for households in the villages with small shocks. In 2001, the slope was -0.10, indicating that households that had larger productivity worked on smaller land (the squares and the solid lines). The slope became 0.10 in 2010, which suggests an improvement in land allocation efficiency (the crosses and the dashed lines). However, the increase was bigger for villages that experienced an above-median trade shock: the slope was -0.24 in 2001 and increased to 0.19 in 2010 (Panel b).

[Figure 6 about here.]

In sum, we find evidence that the Chinese agricultural sector experienced large changes after 2001, and the change was connected to the reductions in output tariffs in the manufacturing sector. We formally investigate this hypothesis in the next section.

5 Main Empirical Results

5.1 Village-Level Results with Trade Shocks: Empirical Specification

Our empirical analysis intends to show the impact of trade shocks on various outcomes in the agricultural sector. The baseline estimation equation is as follows:

$$y_{vt} = \beta_0 + \beta^{other}\tau^{other}_{v(i)t} + \beta^{own}\tau^{own}_{v(i)t} + X_{vt}\Gamma + I_{pt} + I_v + \epsilon_{vt}, \tag{2}$$

where y_{vt} is the outcome variables, including the share of non-agricultural labor, various measures of land, the log of agricultural machinery value, and the village-level TFP in village v and year t. $\tau^{own}_{v(i)t}$ and $\tau^{other}_{v(i)t}$ are village v's exposure to its own prefecture i's tariff and other prefectures' output tariff, respectively. We include a matrix of controls X_{vt}, including the log total number of laborers, the log total number of households, and the log government transfers plus one. Province-year fixed effects I_{pt} are controlled for to take into account unobserved province-year specific weather conditions and various government policies that potentially affect sectoral employment choices and land allocation rules.[38] Village fixed effects I_v control for all time-invariant village characteristics such as overall land quality, climate, other agro-geographical characteristics, and social norms regarding migration and land allocation. Standard errors are clustered at the province level and at the year level to take into account correlated shocks within provinces and within years.

β^{other} and β^{own} are the reduced-form parameters of the impact of tariffs on agricultural production. A reduction of output tariffs in the manufacturing sector effectively increases the price of goods received by exporters. Thus, wages in the manufacturing sector increase, and it acts as a pull factor for labor to move out of the agricultural sector. Both the trade shocks in one's own prefecture and other prefectures can impact the sectoral employment choices. β^{other} and β^{own} being negative means that lower output tariffs on manufacturing goods in export markets increase y_{vt}.

[38] For example, Chari et al. [Forthcoming] shows that provinces implemented the 2003 national land contract law at different times.

Our main parameter of interest is β^{other}. If both the own prefecture's output tariff and other prefectures' output tariff affect the agricultural production only through the labor market, we expect declines in either one leading to increased outflow of labor from agriculture. However, manufacturing trade can also affect agricultural production through other channels. For example, the positive shocks to manufacturing trade increase the income of urban residents, and the increase in income leads to higher demand for agricultural goods with larger income-elasticities. Then agricultural production is affected by the manufacturing growth through the agricultural goods market. If we assume that the agricultural goods market is relatively local, then we expect such demand effects to be captured in β^{own} rather than in β^{other}. In other words, we think that β^{other} is more likely to capture the labor demand effect.[39]

The key identification assumption is that the counterfactual changes in the outcome variables are the same across villages in the absence of trade shocks. Since the counterfactual is not observed, we use pre-trends to provide suggestive evidence on the exogeneity of trade shocks. The hypothesis is that there were no village-level trends in the share of non-agriculture labor, land rental, agricultural capital, and TFP before 2001 that were predictors of post-2001 changes in $\tau^{own}_{v(i)t}$ and $\tau^{other}_{v(i)t}$. We test this hypothesis empirically by running the following regression,

$$\tau^{o}_{v(i)t} = \gamma_o + \gamma_1 \tau^{o}_{v(i)2001} + \Pi Z_{v1995\text{–}2001} + I_p + \xi_v, \tag{3}$$

where $o = own, other$, $t = 2002, ..., 2010$, and $Z_{v1995-2001}$ is a vector of changes of village-level variables from 1995 to 2001, including changes in the share of non-agriculture labor, land rental, agricultural capital, and TFP, and I_p are province fixed effects. We find no evidence of differential trends of key outcome variables for villages with different sizes of trade exposures. The results are shown in Appendix D.4.

Our second main specification takes into account that regions differ in their share of cross-prefecture migrants $s_{v(i)}$ out of all migrants. The impact of pull factors of cross-prefecture migration is larger if in the beginning of the period (2000), a larger share of people moved across prefectures rather than within prefectures. In other words, $s_{v(i)}$ intensifies the impact of other prefectures' output tariff on the origin village's tendency to leave agriculture. Thus, our second main specification adds an interaction term of other prefectures' output tariff $\tau^{other}_{v(i)t}$ and the share of cross-prefecture migrants $s_{v(i)}$,

$$y_{vt} = \beta_0 + \beta^{other}\tau^{other}_{v(i)t} + \beta^{inter}\tau^{other}_{v(i)t} \times s_{v(i)} + \beta^{own}\tau^{own}_{v(i)t} + \Gamma X_{vt} + I_{pt} + I_v + \epsilon_{vt}, \tag{4}$$

and we expect the coefficient β^{inter} to be negative.

[39] In general, prefecture-to-prefecture goods trade and migration can still be correlated due to common costs, such as transportation.

5.2 Village-Level Results with Trade Shocks

Occupation Choice We first investigate the impact of output tariffs on the occupational choice for rural residents in Table 2. Column (1) regresses the share of non-agricultural labor on other prefectures' output tariff, controlling for province-year fixed effects and village fixed effects. The coefficient for other prefectures' output tariff is -0.08 and significant at the 1% level, indicating that a one-standard-deviation larger decline in other prefectures' output tariff resulted in a 3.5-percentage-point (or a 0.25-standard-deviation) larger increase in the share of non-agricultural labor. Column (2) adds own prefecture's output tariff, and the coefficient of other prefectures' output tariff becomes smaller and statistically significant at the 1% level. Column (3) follows the specification in Equation 2, adding the village-year specific controls (i.e., the log total number of laborers, the log total number of households, and the log government transfers plus one). The coefficient for other prefectures' output tariff remains stable. Column (4) follows the specification in Equation 4, adding an interaction of other prefectures' output tariff with the share of cross-prefecture migrants. The coefficient for the interaction is indeed negative as expected, indicating that villages with larger share of cross-prefecture migrants were impacted more strongly by other prefectures' trade exposures. Overall, we find that trade shocks in other regions pulled labor out of agriculture; more so for villages with higher shares of cross-prefecture migrants.

[Table 2 about here.]

We also find heterogeneous effects of other prefectures' output tariff with respect to the initial land-to-agriculture-labor ratio. Column (5) interacts others prefectures' tariff with the log agricultural land-to-labor ratio in 2001. The coefficient for other prefectures' output tariff is –0.16, and the interaction term is 0.05. The positive interaction means that for villages with a larger land-to-labor ratio in agriculture, the effect of a decline in other prefectures' output tariff was smaller. This message is clearer in Column (6), where we interact other prefectures' output tariff with quintile indicators of a village's land per agricultural worker in 2001. For villages with the smallest land per agricultural worker in 2001 (i.e., in the first quintile), a one-standard-deviation larger decline in other prefectures' output tariff led to a 8-percentage-point larger increase in the share of non-agricultural labor. For villages in the fifth quintile, the effect was 3 percentage points.[40]

We interpret the heterogeneous effect of other prefectures' output tariff with respect to the land-to-labor ratio as capturing the role of the extent of factor misallocation at the beginning of the period. In villages with more active land markets, land is likely to be allocated more efficiently across households according to their productivity, which allows workers with comparative advantage in non-agriculture to move out of agriculture. This is also related to the process of urbanization in general. In regions where the urbanization already took place and people moved out of agriculture, additional shocks to out-migration had smaller impacts.

[40]Note that the sample sizes are different in Columns (1)–(4) and Columns (5)(6) since not all villages in the 2001–2010 sample show up in the 2001 sample. When we use the 1971 village-year observations to run the baseline regression as in Column (1), the coefficient for other prefectures' output tariff is -0.09, which is in between of -0.16 for the first-quintile villages and -0.06 for the fifth-quintile villages.

This interpretation is supported by the descriptive evidence in Table 3. We regress different measures of baseline village characteristics in 2001 on the log land per agriculture worker, controlling for province fixed effects. Columns (1)–(3) use direct measures of land market fluidity as the outcome variables, and the size of land leased is positively correlated with the log land per agricultural worker. There are two measures for the land leased. The stock measure comes from the decomposition of the total land at the end of the year, and the flow is a separate measure on how much land a household leased during the year. Column (4) shows that a village's ruggedness is negatively correlated with the land-labor ratio. An explanation is that land consolidation is harder for villages with more rugged surface, thus land reallocation is limited.[41] Column (5) indicates that the share of non-agriculture labor is positively correlated with the land-labor ratio. Column (6) provides the most direct evidence: the allocation efficiency (E_{v2001}, which is the covariance between output and productivity) is positively correlated with the land-labor ratio.

[Table 3 about here.]

The coefficient estimate for own prefecture's output tariff is positive in all column in Table 2, indicating that a reduction in own prefecture's output tariff led to smaller labor outflows from agriculture. The positive estimates of the own prefecture's output tariff impact suggest that there could be alternative channels through which trade shocks affected the non-agricultural labor share. One potential channel is through the local demand for agricultural goods. A reduction in manufacturing output tariffs in prefecture i increased the wage, and the income effect could lead to higher demand for agricultural goods, especially food such as dairy products, vegetables, and fruits. Thus, it could be more profitable for farmers to stay in agriculture. We find evidence in Appendix D.10 on how an increase in own prefecture's trade exposure led to an increase in the revenue share of cash crops, although the effect is not statistically significant.

We provide robustness checks in Appendix D.5, by (1) including agricultural trade shocks, (2) controlling for the reduction in trade uncertainty induced by the establishment of the U.S.-China permanent normal trade relationship (PNTR), (3) using the migration-related information in the village questionnaire in Appendix D.6, (4) controlling for the initial crop patterns and contemporaneous crop patterns in Appendix D.12, and (5) controlling for the share of migrants who moved to the top 10 migrant destinations. We find that agricultural import and export tariffs and agricultural goods market access did not have significant effects on the occupation choice; this is consistent with the fact that China's most salient growth in trade was in the manufacturing sector instead of the agricultural sector. Including the PNTR shocks do not affect the estimates of our actual tariff effects (β^{other} and β^{own}). The village questionnaire includes village-level measures of the number of households exclusively in agriculture, the number of labor working outside the village, and excess labor. The occupation choice and migration results are consistent with the findings in Table 2. The effects on excess labor are insignificant. The crop patterns and the share of migrants who moved

[41] The ruggedness data is from Nunn and Puga [2012] with cells on a 30 arc-seconds grid. The cell-level data is aggregated to the county level, and our assumption is that villages within a county have the same ruggedness level.

to popular destinations did not affect the relationship between shocks through the migrant network and out-migration patterns.[42]

Overall, we find that villages with migrant connections to prefectures facing larger tariff declines in export markets had larger flows of labor from agriculture to non-agriculture. The effect was stronger for villages with larger shares of cross-prefecture migration, and also for villages that were in earlier stages of urbanization and had worse land allocation in the beginning of the period.

Land Market When more people moved out of agriculture, the rental market of agricultural land became more active. Table 4 shows the how output tariff reductions in other prefectures affected the origin village's land market fluidity. For each outcome variable that measures land market fluidity, we use the two main specifications as in Table 2 Columns (3) and (4). A one-standard-deviation larger decline in other prefectures' output tariff led to a 26% (or a 0.16-standard-deviation) larger increase in the stock of land leased, and the effect is significant at the 1% level (Column 1). The effect was larger for villages whose prefecture had higher between-prefecture migrant rates (Column 2). We find similar patterns when we use the log flow of land leased (Columns 3 and 4) and the log income from land leasing (Columns 5 and 6) as the measure for the activeness of the land market.

[Table 4 about here.]

How did the trade shocks affect the total land size and land distribution? Table 5 Column (1) shows that the total land size declined more in villages that experienced larger declines in other prefectures' output tariff. However, the estimates are not statistically significant. The effect on the log land per agricultural worker was also negative as expected, but not statistically significant either (Column 3 and 4). Columns (5) and (6) show the impact on the number of households with land larger than one-third hectare. The median land size of households in the 2001–2010 period was 0.32 hectare. Thus, we consider households with more than 1/3 hectare land the ones with relatively large land. The decline in other prefectures' output tariff led to the increase of the number of households with relatively large land in villages where the share of cross-prefecture migrants was larger than 58%.

[Table 5 about here.]

The active land rental market could also affect the land distribution across households, potentially shifting land from unproductive households to productive households. Thus, we further investigate the differential impacts of the exposure to other prefectures' trade shocks on households with different agricultural productivity in 2001, using the following baseline specification,

$$\log(land)_{h(v)t} = \alpha_0 + \alpha_1 \log(TFP)_{h(v)2001} + \alpha_2 \tau^{other}_{v(i)t} + \alpha_3 \log(TFP)_{h(v)2001} \times \tau^{other}_{v(i)t} + \alpha_4 \tau^{own}_{v(i)t} + I_{pt} + I_v + \epsilon_{h(v)t},$$

[42] The latter exercise intends to address the concern raised in Borusyak et al. [2019] on the how dispersed the migration network ("shares" in general terms in the shift-share design papers) was and whether the shock-level law of large numbers applies.

where the log land for household h in village v and year t is regressed on the households' initial productivity, $\log(TFP)_{h(v)2001}$, other prefectures' output tariff, $\tau^{other}_{v(i)t}$, and the interaction of the two. We control for a village's own tariff, province-year fixed effects, and village fixed effects. By controlling for the village fixed effects, we are essentially comparing households within a village and identifying the land allocation effect within villages rather than across villages. The parameter of interest is α_3, with a negative value indicating that a household with a relatively large TFP in 2001 gained more land than a household with a relatively small TFP in the same village, and the gap in the land increase was larger in villages with larger declines in other prefectures' output tariff.

Table 6 shows the household-level land allocation effect. We find that the shift of land from unproductive farmers to productive farms was stronger in villages that experienced larger shocks. The coefficient estimate of α_3 is –0.061 and is statistically significant at the 1% level. To interpret the coefficient, let us compare two households (A and B) in the same village. Household B has a one-standard-deviation larger log TFP in 2001 than household A. Holding the trade exposure constant, household B had 18% larger land during the 2001–2010 period than household A. In villages with a one-standard-deviation larger tariff decline in destination prefectures, household B had 21% larger land than household A. In other words, villages that experienced larger shocks of non-agricultural sector allocated land more toward the initially productive households. Results are similar in Column (2) where we control for household fixed effects instead of village effects, and in Column (3) and (4) where we use the measure for others prefectures' output tariff by taking into account the share of cross-prefecture migrants. Overall, villages with a one-standard-deviation larger tariff decline in other prefectures had a 20% larger elasticity of land to TFP at the household level.[43]

[Table 6 about here.]

In sum, we find several aspects of the agricultural land distribution being affected by tariffs facing exporters in migrant-connected prefectures. First, the land rental market became more fluid, measured by the size of rental transactions within a year and the stock of land that was rented. Second, the effects on the overall land size and the land per agricultural worker are insignificant. Third, there is some evidence on the land consolidation: the decline in other prefectures' output tariff led to the increase of the number of households with relatively large land, especially in villages with high cross-prefecture migration rates. Fourth, the within-village shift of land from unproductive farmers to productive farmers was more significant in villages with larger declines in other prefectures' output tariff.

Adoption of Agricultural Machinery We proceed to investigate the changes in the capital market. With labor leaving agriculture, capital adoption is likely to increase. The increase in capital can come from several forces. The first reason is the increased size of farms. Suppose that there is a fixed cost of purchasing agricultural machinery, then the farm size needs to be big enough

[43] The 2001 household TFP is highly correlated with TFP in later years. Appendix D.7 uses the current TFP instead of the initial TFP, and the results are very similar.

for farmers to adopt machinery. The second is due to the substitution between labor and capital. When the local labor costs increase due to the outflow of labor from agriculture, farms tend to substitute labor with capital. Third, capital adoption can increase when the land misallocation decreases. When land is reallocated from unproductive farmers to productive farmers, since the increase of the marginal product of capital for productive households is higher than the decrease of the marginal product of capital for unproductive households, the overall amount of capital can increase.

Consistent with the first hypothesis, in the land market section, we find larger increased number of households with relatively large lands in villages that experience larger declines in destination prefectures' tariff. However, we do not have plot-level information to know if the newly-rented plot was contiguous with the original plot or not.

Regarding the second hypothesis, the capital-labor substitution, we provide evidence on the increased labor cost in Table 8. An important fact about rural China is that hired labor was not prevalent, with a share of hired labor days of 2% on average during the 2001–2010 period. This includes hired labor *by rural households within the village* in both agricultural and non-agricultural production. Thus, the measurement errors can be big for wages and hired labor days: although the implicit labor cost increased with labor outflows, the wages and hired labor days could be noisily measured. Some villages even had no hired labor, so the number of observations in Table 8 is smaller than in Table 4. Overall, we find that tariff declines in migrant connected prefectures led to a decrease in the labor days in locally hired labor and an increase in wages of locally hired labor, for villages with large share of cross-prefecture migrants. However, the effects are statistically insignificant.

[Table 7 about here.]

Table 8 presents the impact of trade exposure on capital. Agricultural machinery increased more in villages that had larger declines in other prefectures' output tariff, and the effect was stronger where the share of cross-prefecture migrant was high (Columns 1 and 2). Evaluated at the mean of cross-prefecture migrant share (0.46), a one-standard-deviation larger decline in other prefectures' output tariff led to a 8% (or a 0.05-standard-deviation) larger increase in the value of agricultural machinery. We also find that the number of households with positive agricultural machinery and relatively large land increased in response to tariff declines (Columns 3 and 4), even when we control for the number of households with positive agricultural machinery and relatively small land (Columns 5 and 6). By controlling for the number of households with positive agricultural machinery and relatively small land, we address the potential concern that maybe in villages with larger trade shocks, *all* households increased capital adoption. We find that compared to the households with relatively small land, the households with relatively large land still had more intensive capital adoption. This result supports the third hypothesis on increased capital adoption.

[Table 8 about here.]

An alternative hypothesis on the increase of capital stock is through migrant remittance. Suppose that there is no well-functioning credit market in rural areas, and farmers are not able to buy machinery due to their credit constraint. Out-migration of part of the agricultural household can increase the household income and ease the credit constraint, facilitating capital adoption. Consistent with De Brauw and Rozelle [2008], we don't find support for this hypothesis in our context.[44] As shown in Appendix D.8, the expenditure on productive fixed assets was negatively correlated with wage income, but was positively correlated with income from farming and government subsidies. This finding is consistent with the overall trend of urbanization, since people were not trying to fund their agricultural production through migration remittances, but they were rather leaving agriculture in the long-run. Only the workers who decided to remain in agriculture re-invest in their production.

Overall, we find positive impacts of the exposure to other prefectures' trade shocks on capital adoption. We provide suggestive evidence showing that the capital adoption was likely to be caused by increased local labor costs, land consolidation, and improved land allocation across households, rather than larger migrant remittance. The capital adoption in agriculture can further make the urbanization process irreversible. Once the agricultural sector modernizes, it is not likely to go back to labor intensive production. This process is similar to the modernization of manufacturing sector when robots and machines replace workers.

Productivity and Allocation The adjustment in the factor markets led to changes in the overall productivity of the villages. When land was more allocated toward the productive farmers and such farmers adopted more capital for production, they would be able to produce more output and have higher weights in the village-level TFP. We find that a one-standard-deviation larger decline in other prefectures' output tariff led to a 30% (or a 0.33-standard-deviation) larger increase in the output weighted village-level TFP (Columns 1 and 2 in Table 9), while there was no significant effect on the unweighted TFP (Columns 3 and 4). This allocation effect is directly demonstrated in Columns (5) and (6), where the allocation efficiency (i.e. the difference between the output weighted and unweighted TFP) increased more in villages that had larger declines in other prefectures' output tariff.[45]

[Table 9 about here.]

An alternative hypothesis on the TFP effect is that villages with more out-migration switches from cereal crops to cash crops, and this switch increases the TFP estimated using the output value since the cash crops have higher prices. We investigate this hypothesis in Appendix D.10. Overall, we don't find that big-shock regions had differential rates of switching from cereal crops to cash crops. The decline in other prefectures' output tariff had insignificant effects on the revenue

[44]Similarly, Dinkelman et al. 2017 find that migrant remittance increased capital formation in non-agricultural sectors.

[45]The results in Table 9 use the TFP calculated using the output method. Appendix D.9 shows that the results are similar with the TFP calculated using the value added method.

share of cash crops and led to a decline in the number of households in cash crop production. This finding is consistent with two facts. First, the share of households in cash crop production had a similar declining trend as the households with cereal crops from 2001 to 2010. Thus, households were moving out of all types of crop farming. Second, the cash crops can be more labor intensive than cereal crops. Take the most common cash crop, vegetables, as an example. As estimated in Chari et al. [Forthcoming], the output elasticity of labor is among the highest in all crops. Also, it is intuitive that with the harvest cycle of vegetables, more labor input is needed to attend to the production process. We indeed find households moving out of vegetable production in the face of positive out-migration shocks.

Although the focus of our paper is on crop farming, we also test the out-migration effect on husbandry. The effect is very similar to the one discussed above, on cash crops. We don't find any significant effect on husbandry.

In sum, the increase in trade exposures in destination prefectures attracted farmers to exit agriculture and enter manufacturing and service, and all sub-sectors in agriculture experienced a negative labor supply shock. We find that the land market became more fluid, with land allocated more toward productive farmers, that the capital adoption increased, and that village-level TFP improved. The trade-induced manufacturing growth led to an agricultural sector with better factor allocation and modern production technologies.

6 Migrant Selection

In this section, we investigate the role of migrant selection in driving productivity results documented in the previous section. When more workers choose to leave agriculture, there is an increase of land for lease on the market, and the rental price of land should decrease. Households with the higher agricultural productivity also have higher marginal product of land, thus the increase in the size of land leased-in should be larger than the one for the households with relatively low productivity. However, if the agricultural productivity and non-agricultural productivity are not strongly positively correlated, then the unproductive farmers are more likely to leave agriculture when there is a positive labor demand shock in non-agriculture, since the opportunity costs of leaving agriculture is smaller for unproductive farmers. In this case, the increase in the village-level TFP is even larger than the case where farmers with different agricultural productivity have equal probabilities of switching to non-agriculture.

Overall, we show some results on unproductive farmers leaving agriculture. We also provide evidence supporting the absence of a strong positive correlation between agricultural and non-agricultural productivity. In the next section, we present a simple two-sector-economy model with occupation choice and land market transaction costs and use the model to shed light on the correlation between an individual's agricultural versus non-agricultural abilities.

6.1 Skill Heterogeneity and Occupation Choice

Who were the productive farmers and productive non-agricultural workers? We investigate the characteristics of individuals that were correlated with productivity in Table 10, using the individual-level information from 2003 to 2008.[46] The baseline specification is as follows,

$$y_{dt} = \delta_0 + \delta_1 \text{edu}_{dt} + \delta_2 \text{train}_{dt}^{non-agr} + \delta_3 \text{train}_{dt}^{agr} + \delta_4 \text{age}_{dt} + \delta_5 \log(\text{labor})_{d(h)t} + I_t + \xi_{it},$$

where the outcome variable is either the log of TFP in agriculture for an individual d in household h and year t, or the log income from working outside the village for individual d in year t. The individual characteristics include years of education, a dummy variable indicating whether the person had non-agricultural occupational training, a dummy variable for agricultural training, age, and the size of the household. We also control for year fixed effects to take into account year-specific shocks to productivity and wages.

[Table 10 about here.]

Table 10 Columns (1) and (4) show the results on agricultural productivity and on wages, respectively. One additional year of education (or a 0.44-standard-deviation larger education) was correlated with a 3.5-percentage-point larger TFP (or a 0.05-standard-deviation increase in the log TFP). Individuals with agricultural training had 16-percentage-points larger TFP, and the individuals with non-agricultural training had 8-percentage-point lower TFP. For individuals with positive income from working outside the village, 75% report their industry as non-agriculture. We find a larger correlation of education with income from outside the village: a one-year larger education level was correlated with a 5.6-percentage-point larger income. Having occupation training was correlated with a 30-percentage-point larger income, and having agricultural training was correlated with a 4-percentage-point smaller income. We find small age effects on both outcomes.

Columns (1) and (4) suggest that educated individuals were likely to be in more productive agricultural households and had higher earning in non-agriculture. In addition, having sector-specific training was correlated with higher sector-specific productivity. Table 11 shows that only 2% of individuals had both agricultural trainings and non-agricultural trainings, while 6% only had non-agricultural training, and 5% only had agricultural training. Thus, there is evidence on sector-specific human capital investments.

[Table 11 about here.]

However, we find that the education and training effects were more significant and robust for the non-agricultural income than for the agricultural productivity. When including village fixed effects in Columns (2) and (5), the education and training effects on agricultural productivity become very

[46] The 2009 and 2010 questionnaire have different definition on occupational and agricultural training, so we only use the 2003–2008 data for consistency.

small, while the effects on non-agricultural income remain. In addition, the R^2 changes a lot from Column (1) to Column (2), and the change from Column (4) to Column (5) is smaller. The results suggest that observable characteristics of individuals explain only a small share (3%) of the variation in agricultural productivity, and village-specific time-invariant characteristics explain about 53% of the variation. In contrast, education levels and training explain a sizable share (11%) of the variation in the non-agricultural income, and the village fixed effects explain an additional 20%.

We also find that initial agricultural productivity is more informative about the agricultural productivity than about non-agricultural productivity in later years. Column (3) shows the persistence of the agricultural productivity. When regressing the log TFP on the log TFP in 2001, the coefficient indicates that a household that is 10% more productive in 2001 are 18% more productive in years 2003–2008. Thus, the initial agricultural productivity is a good indicator of the agricultural ability in later years. Column (6) regresses the non-agricultural income on the initial agricultural productivity, and the coefficient is not significant.

Overall, we find no evidence of a strong positive correlation between agricultural and non-agricultural ability. Individuals with higher non-agricultural productivity usually choose higher education, since education is often seen as a type of human capital investment that pays off in the non-agricultural sector. We do not find that education is highly correlated with the non-agricultural productivity, once the village fixed effects are controlled; neither do we find significant effects of training on agricultural productivity. In addition, the agricultural productivity in 2001 is a good predictor for later years' agricultural productivity, but not a good predictor for non-agricultural income once the individual works in the non-agricultural sector.

6.2 Out-Migration Patterns

We investigate the probability of leaving agriculture for households with different agricultural productivity, with the following baseline specification,

$$y_{h(v)t} = \delta_0 + \delta_1 \log(TFP)_{h(v)2001} + \delta_2 \text{labor}_{h(v)t} + I_t + I_v + \xi_{ht},$$

where $y_{h(v)t}$ represents the number of non-agricultural laborers of household h in village v and year t, $\log(TFP)_{h(v)t}$ is the household's productivity in 2001, and $\text{labor}_{h(v)t}$ is the total number of labor in the household. We control year fixed effects and village fixed effects, for the comparison to be within villages.

We find that the unproductive farmers were more likely to leave agriculture, but the effects are economically small. Table 12 Column (1) shows that a household with a one-standard-deviation (0.6) larger TFP was more likely to have 0.03 more non-agricultural labor. Column (2) includes the quadratic initial TFP. The result suggests an inverse-U shaped relationship, but for 80% of the households (i.e., households with the log TFP higher than 2.18), the negative correlation between the initial productivity and the number of non-agricultural laborers remains. Columns (3) and (4) repeat the exercise by using village-year fixed effects, and the results are similar as in Columns (1) and (2).

[Table 12 about here.]

We then investigate whether the responsiveness to trade shocks were different across households with different initial productivity. The baseline specification is as follows,

$$y_{h(v)t} = \delta_0 + \delta_1 \tau^{other}_{v(i)t} \times s_{v(i)} + \delta_2 \tau^{own}_{v(i)t} + \delta_3 \text{labor}_{h(v)t} + I_{pt} + I_h + \xi_{ht},$$

where $y_{h(v)t}$ represents the number of non-agricultural laborers of household h in village v and year t, and $\tau^{other}_{v(i)t} \times s_{v(i)}$ represents the exposure to other prefectures' output tariff. We control own prefecture's output tariff, the total number of labor in the household, province-year fixed effects, and household fixed effects.

Table 13 shows the unproductive farmers were more responsive to trade shocks. Column (1) uses the sample of households whose TFP in 2001 was below median, i.e., the relatively unproductive farmers, and Column (2) uses the above median sample. We find that with a one-standard-deviation larger decline in the exposure to other prefectures' output tariff (0.59), the unproductive households had a 0.24 larger increase in the number of non-agricultural laborers (Column 1). In comparison, the productive households had a 0.17 larger increase. Column (3) and (4) repeat the exercises by replacing the household fixed effects with the village fixed effects, and the results are similar.

[Table 13 about here.]

In sum, we find evidence that unproductive agricultural households were more likely to move out of agriculture, and were more responsive to trade shocks.

7 Model and Quantitative Exercise

In this section, we present a simple two-sector open-economy model with agricultural land market frictions. We calibrate the key model parameters using the 1995, 2001, and 2010 NFP Survey data and conduct several quantitative exercises. Importantly, we back out: (1) the correlation between agricultural and non-agricultural productivity; (2) the size of agricultural land market frictions; (3) the economy-wide average productivity in agriculture and non-agriculture. Using this information, we investigate the interaction between the push factors driving out-migration (i.e., the reduction in land market frictions) and the pull factors of out-migration (i.e., relative productivity growth in the two sectors).

Our two-sector economy model shares many features with Adamopoulos et al. [2017]. The key differences are as follows. First, instead of the closed-economy setup, we model a small open economy where the agricultural and non-agricultural prices are determined in the international market. This allows us to focus on the supply side of the economy and abstract from demand side factors such as non-homothetic preferences and subsistence constraints for the consumption of agricultural goods. Second, Adamopoulos et al. [2017] model the household-level frictions in the agricultural sector as a random variable that is correlated with the household's agricultural productivity, and the size of the friction is such that no land rental activity happens in equilibrium. In contrast, we provide a micro foundation for the source of misallocation in the form of land market transaction costs as in Chari et al. [Forthcoming]. Third, we calibrate the model parameters with data from different years to document changes in the key parameter values and to conduct counterfactual analysis using these changes.[47]

We present the model setup, the intuition of the calibration, and the results of the counterfactual analysis here. For details on the calibration and counterfactual analysis, see Appendix E.

7.1 Model Setup

Environment The economy is a small open economy with two sectors, one agricultural sector (a) and one non-agricultural sector (n). The relative price of the agricultural goods is p_a, and the non-agricultural good price is normalized to 1. Both prices are determined exogenously on the international market. There is a continuum of individuals with measure 1. Each individual i is endowed with a pair of productivity levels in the two sectors (s_{ai}, s_{ni}), land $\bar{l}_i$, and one unit of labor that is inelastically supplied. We assume that land is equally distribution across individuals, so $\bar{l}_i = \bar{l}$. The cost of capital r is determined exogenously on the world market. An individual chooses the sector with a higher income, so $I_i = \max\{I_{ai}, I_{ni}\}$, where I_i is the income of an individual, I_{ai} is the income in agriculture, and I_{ni} is the income in non-agricultural. Let $H_n = \{i : I_{ai} < I_{ni}\}$, and $H_a = \{i : I_{ai} \geq I_{ni}\}$.

Production The agricultural sector operates with individual farms that exhibit decreasing return to scale with respect to capital and land,

$$y_{ai} = (A_a s_{ai})^{1-\gamma}(l_i^{\alpha} k_i^{1-\alpha})^{\gamma},$$

where y_{ai} is the agricultural output in farm i, A_a is the agricultural productivity that is common across individuals, l_i and k_i are the land input and capital input, respectively. γ is the span-of-control parameter that governs the returns to scale, as in Lucas [1978].

The land market features a transaction cost in renting. If farmers use less than their endowed land $\bar{l}$, the rental rate is q. If they need to rent in land from other farmers, the rental rate is $q(1+\tau)$. Thus, there is a kink in the marginal cost curve. Think about three types of farmers with different agricultural productivity: $s_{a1} < s_{a2} < s_{a3}$. As shown in Figure 7, the marginal product of land for a Type I farmer at $\bar{l}$ is smaller than the rental rate q, so a Type I farmer will use the land up to the point where $MPL_1 = q$, and rent out the rest of land. The marginal product of land for a Type II farmer at $\bar{l}$ is between q and $q(1+\tau)$, so a Type II farmer will use exactly $\bar{l}$. For a Type III farmer with $MPL_{l_3=\bar{l}} > q(1+\tau)$, he/she will rent in land from other farmers.

[47]Note that the modeling of the non-agriculture sector is simplistic since it ignores the population that is always urban and only includes the population that moved from rural (agricultural sector) to urban (non-agricultural sector). In this sense, we also differ from Adamopoulos et al. [2017] since we do not match any aggregate moments but only focus on the part of the economy covered by the NFP Survey.

[Figure 7 about here.]

A farmer maximizes profit by choosing the optimal inputs,

$$\max_{k_i, l_i} \pi_i = p_a y_{ai} - r k_i - C(l_i),$$

where r is the rental rate of capital, and $C(l_i)$ is the cost of land, which takes the following form:

$$C(l_i) = \begin{cases} q l_i & \text{if } l_i \leq \bar{l}, \\ q\bar{l} + q(1+\tau)(l_i - \bar{l}) = q(1+\tau)l_i - q\tau\bar{l} & \text{if } l_i > \bar{l}. \end{cases}$$

Production in the non-agricultural sector employs a constant return to scale technology that uses effective labor only,

$$Y_n = A_n Z_n,$$

where Y_n is the non-agricultural output, A_n is non-agricultural productivity that is common to all individuals, and Z_n is the total amount of effective labor used. Thus,

$$Z_n = \int_{i \in H_n} s_{ni} di,$$

where H_n is the set of individuals who chooses to work in the non-agricultural sector. The total number of workers in the non-agricultural sector is

$$N_n = \int_{i \in H_n} di.$$

Occupation choice If a person chooses the agricultural sector, he receives the profit from running the farm. A person always receives the factor payment $q\bar{l}$ no matter which sector he works in. Thus, agricultural income $I_{ai} = \pi_i + q\bar{l}$. If a person chooses the non-agricultural sector, he receives wage of w_n. Thus, the non-agricultural income $I_{ni} = w_n s_{ni} + q\bar{l}$. The transaction cost τ represents the friction in the land rental market, so $\tau(l_i - \bar{l})$ is simply lost for $l_i > l$ and no agent gets it as income. The occupation choice is represented by

$$o(s_{ai}, s_{ni}) = \begin{cases} 1 & \text{if } i \in H_a, \\ 0 & \text{if } i \in H_n. \end{cases}$$

7.2 Model Analysis and Calibration

We calibrate the key model parameters using the 1995, 2001, and 2010 NFP Survey data. First, we assume that the ability (s_{ai}, s_{ni}) follows a bi-variate log-normal distribution with mean $(0, 0)$ and variance Σ and is fixed over time. The variance-covariance matrix Σ is important in the migrant selection analysis. Second, we want to recover the land market transaction cost τ. Third, the average productivity in the two sectors (A_a, A_n) is useful in analyzing the relative productivity changes in the two sectors.

The model implies that in any cross-section, there exist three types of farmers: Type I with land smaller than the endowment, Type II with land at the endowment, and Type III with land larger than the endowment. Importantly, given the bunching at the endowment point, we should observe an empirical mass point.

We define the empirical mass point to be the village mode. In Figure 8, we plot the distribution of the log land per agricultural worker (subtracting the village-year mode) in 1995, 2001, and 2010. The household-level land is divided by the number of agricultural workers to construct data at the individual level. We find that in all three years, there appears to be discontinuity in the distribution of the log land per agricultural worker at –0.16 and 0.16. Thus, we allow fuzziness in the definition of the village mode and define observations with the value between –0.16 and 0.16 as at the village mode. Then the observations to the left of the village mode are Type I farmers, the ones at the village mode are Type II farmers, and the ones to the right of the mode are Type III farmers. We find that the variance of the distribution of land per agricultural worker increased from 2001 to 2010, while the 1995–2001 change was very small.

[Figure 8 about here.]

Based on this definition, we use the information on (1) the probabilities of being Type I, Type II, and Type III farmers, (2) the probability of choosing agriculture over non-agriculture, (3) the variances of land for Type I and Type III farmers, (4) the mean income of workers who switched from agriculture to non-agriculture, (5) the mean value of agricultural output, and (6) the mean output-to-land and output-to-capital ratio to calibrate the variance of the joint distribution of agricultural and non-agricultural ability, Σ, the average productivity (A_a, A_n), and the land market transaction cost τ (see Appendix E.3 for details).

The calibration results are shown in Table 14. Here, we assume that the variance of agricultural ability and non-agricultural ability are the same. We find that the correlation of the sectoral ability is 0.17, a rather small positive correlation, which is consistent with the empirical results shown in Section 6. The land market transaction costs are 1.6 in 2001 and 1.2 in 2010. These costs are substantial and decline over the years, which is in line with the land reforms documented in Chari et al. [Forthcoming]. The growth of agricultural productivity (from 6.8 in 2001 to 7.2 in 2010) is smaller than the growth in the non-agricultural productivity (from 8.1 in 2001 to 9.6 in 2010), indicating strong forces for sectoral labor reallocation.

[Table 14 about here.]

7.3 Quantitative Exercise

Our quantitative exercise focuses on two aspects. First, we investigate the role of the land market transaction costs in determining the sectoral employment patterns and economic output.

Second, we experiment with different patterns of sectoral productivity growth and highlight the role of manufacturing productivity growth.

[Table 15 about here.]

The results of the counterfactual analysis are shown in Table 15. Column BE shows the aggregate statistics for the baseline economy in 2010. Our first experiment changes the value of the land market transaction cost. In Column C1, we set the 2010 transaction cost to be zero and eliminate all distortions in the land market. The demand for land rental increases, leading to increased land rental prices. Low productivity farmers exit and high productivity farmers enter, resulting in an increase in the average ability of farmers. On the other hand, the average productivity of non-agricultural workers declines due to the influx of the marginal workers who are less productive. Overall, this results in a slight increase in per capita GDP (2%), a slight decrease in the share of agricultural workers (one percentage point), and an increase in the capital per worker in the agricultural sector (9%).

One second experiment changes the non-agricultural productivity. In Column C2, we set the 2010 non-agricultural productivity to be the same as the 2001 non-agricultural productivity. Due to the decline in non-agricultural productivity, the demand for land increases and the land rental price also increases. The opportunity cost of farming declines, and more people stay in agriculture. This results in a decline in the average agricultural productivity and an increase in the average non-agricultural productivity, since only the most productive workers remain in non-agriculture. Overall, the effect is a substantial decrease in per capita GDP (17%), a large increase in the share of agricultural workers (42 percentage points), and a large decline in the capital per worker in the agricultural sector (31%).

In sum, we find a relatively small effect of reducing land-market transaction costs on the overall economy, measured as the per capita GDP and the share of employment in the agricultural sector. The intuition is that given the increased importance of the non-agricultural sector, the distortion in the land market in the agricultural sector was relatively unimportant. This mirrors Chari et al. [Forthcoming] where they find no impact of land reforms on out-migration in the 2003–2008 period. However, reducing the distortion still benefited the agricultural sector, in the form of increased capital adoption and agricultural TFP. In contrast, the effect of increasing non-agricultural productivity had very large impacts on the sectoral employment patterns and agricultural productivity. Only the most productive farmers remained in agriculture, and they substantially increased the amount of capital used. Overall, the pull factors of out-migration (i.e., relative productivity growth in the two sectors) had much larger impacts than the push factors of out-migration (i.e., the reduction of land market frictions) on both urbanization and agriculture modernization.

8 Conclusion

In this paper, we study the impact of manufacturing growth on agriculture production. Manufacturing growth induces the outflow of labor from agriculture, and this labor market linkage can have subsequent impacts on the rural land and capital market, resulting in productivity changes in agriculture. Started as a developing country with more than half of population in agriculture, China experienced fast growth in the manufacturing sector and substantial urbanization after 2001. China's accession to the WTO provides a unique context to study the structural change stemming from manufacturing growth. Using destination prefectures' trade shocks in the manufacturing sector resulted from China's accession to the WTO and origin village's initial migration network, we construct the exposure to other prefecture' manufacturing trade shocks for 210 villages from 2001 to 2010. We find that villages with larger exposures had larger increases in out-migration rates, higher land rental rates, and larger growth in agricultural productivity. The increase in productivity is through the allocation of land towards more productive farmers within a village. In addition, these villages modernized their production by adopting more agricultural machinery. The quantitative exercise further shows that the pull factors for out-migration had larger impacts on structural transformation compared to the impact of land market frictions.

As China moved along the global value chain and shifted many labor-intensive manufacturing jobs to other developing countries, such as Vietnam, the Philippines, and Indonesia, it is likely that the manufacturing growth in those countries can similarly accelerate the development process in the agriculture sector. In this light, international trade could generate welfare gains in the developing world through triggering the modernization of the economies. Studies of these alternative contexts with similar forces are left to future research.

The shift of employment from rural to urban areas can provide additional engines for economic growth, if human capital accumulation is faster in the urban areas, as hypothesized in Lucas [2004]. Through learning-by-doing in urban jobs and access to urban amenities, workers can acquire useful skills that can generate gains in life-time earnings. This can further increase the labor productivity in the economy. There is much room for future research on this important issue.

References

Daron Acemoglu and Pascual Restrepo. Robots and jobs: Evidence from us labor markets. Working Paper 23285, National Bureau of Economic Research, March 2017.

Tasso Adamopoulos and Diego Restuccia. The Size Distribution of Farms and International Productivity Differences. *American Economic Review*, 104(6):1667–97, June 2014. doi: 10.1257/aer.104.6.1667. URL `http://www.aeaweb.org/articles?id=10.1257/aer.104.6.1667`.

Tasso Adamopoulos, Loren Brandt, Jessica Leight, and Diego Restuccia. Misallocation, Selection and Productivity: A Quantitative Analysis with Panel Data from China. NBER Working Papers 23039, National Bureau of Economic Research, Inc, January 2017. URL `https://ideas.repec.org/p/nbr/nberwo/23039.html`.

Agha Ali Akram, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. Effects of Emigration on Rural Labor Markets. Technical report, National Bureau of Economic Research, 2017.

Manuel Arellano and Olympia Bover. Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics*, 68(1):29–51, 1995.

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, October 2013. doi: 10.1257/aer.103.6.2121. URL `http://www.aeaweb.org/articles?id=10.1257/aer.103.6.2121`.

Abhijit Banerjee, Esther Duflo, and Nancy Qian. On the road: Access to Transportation Infrastructure and Economic Growth in China. Technical report, National Bureau of Economic Research, 2012.

William J Baumol. Macroeconomics of Unbalanced Growth: the Anatomy of Urban Crisis. *The American Economic Review*, 57(3):415–426, 1967.

Dwayne Benjamin and Loren Brandt. Property Rights, Labour Markets, and Efficiency in a Transition Economy: The Case of Rural China. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 35(4):689–716, 2002. ISSN 00084085, 15405982. URL `http://www.jstor.org/stable/3131831`.

Dwayne Benjamin, Loren Brandt, and John Giles. The Evolution of Income Inequality in Rural China. *Economic Development and Cultural Change*, 53(4):769–824, 2005.

Kirill Borusyak, Peter Hull, and Xavier Jaravel. Quasi-Experimental Shift-Share Research Designs. Technical report, arXiv. org, 2019.

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`.

Paula Bustos, Bruno Caprettini, and Jacopo Ponticelli. Agricultural Productivity and Structural Transformation: Evidence from Brazil. *American Economic Review*, 106(6):1320–65, 2016.

Kang Hua Cao and Javier A Birchenall. Agricultural Productivity, Structural Change, and Economic Growth in Post-Reform China. *Journal of Development Economics*, 104:165–180, 2013.

Francesco Caselli and Wilbur John Coleman. The US Structural Transformation and Regional Convergence: A Reinterpretation. *Journal of Political Economy*, 109(3):584–616, 2001.

A.V. Chari, Elaine M Liu, Shing-Yi Wang, and Yongxiang Wang. Property Rights, Land Misallocation and Agricultural Efficiency in China. *Review of Economic Studies*, Forthcoming.

Chaoran Chen. Untitled Land, Occupational Choice, and Agricultural Productivity. *American Economic Journal: Macroeconomics*, 9(4):91–121, October 2017. doi: 10.1257/mac.20140171. URL `http://www.aeaweb.org/articles?id=10.1257/mac.20140171`.

Gregory C. Chow. Capital Formation and Economic Growth in China. *The Quarterly Journal of Economics*, 108(3):809–842, 1993. ISSN 00335533, 15314650. URL `http://www.jstor.org/stable/2118409`.

Michael A. Clemens, Ethan G. Lewis, and Hannah M. Postel. Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion. *American Economic Review*, 108(6):1468–87, June 2018. doi: 10.1257/aer.20170765. URL `http://www.aeaweb.org/articles?id=10.1257/aer.20170765`.

Alan De Brauw and Scott Rozelle. Migration and Household Investment in Rural China. *China Economic Review*, 19(2):320 – 335, 2008. ISSN 1043-951X. doi: https://doi.org/10.1016/j.chieco.2006.10.004. URL `http://www.sciencedirect.com/science/article/pii/S1043951X0600085X`.

Alain de Janvry, Kyle Emerick, Marco Gonzalez-Navarro, and Elisabeth Sadoulet. Delinking Land Rights from Land Use: Certification and Migration in Mexico. *American Economic Review*, 105(10):3125–49, October 2015. doi: 10.1257/aer.20130853. URL `http://www.aeaweb.org/articles?id=10.1257/aer.20130853`.

Taryn Dinkelman, Grace Kumchulesi, and Martine Mariotti. Labor Migration, Capital Accumulation, and the Structure of Rural Labor Markets. Technical report, 2017.

Bilge Erten and Jessica Leight. Exporting out of Agriculture: The Impact of WTO Accession on Structural Transformation in China. Technical report, Forthcoming.

Giovanni Facchini, Maggie Y Liu, Anna Maria Mayda, and Minghai Zhou. China's "Great Migration": The Impact of the Reduction in Trade Policy Uncertainty. *Journal of International Economics*, 2019.

Andrew D Foster and Mark R Rosenzweig. Agricultural Productivity Growth, Rural Economic Diversity, and Economic Reforms: India, 1970–2000. *Economic Development and Cultural Change*, 52(3):509–542, 2004.

Andrew D Foster and Mark R Rosenzweig. Economic Development and the Decline of Agricultural Employment. *Handbook of Development Economics*, 4:3051–3083, 2007.

Andrew D Foster and Mark R Rosenzweig. Are Indian Farms Too Small? Mechanization, Agency Costs, and Farm Efficiency. *Unpublished Manuscript, Brown University and Yale University*, 2011.

Andrew D Foster and Mark R Rosenzweig. Are There Too Many Farms in the World? Labor-Market Transaction Costs, Machine Capacities and Optimal Farm Size. Technical report, National Bureau of Economic Research, 2017.

Douglas Gollin, Stephen Parente, and Richard Rogerson. The Role of Agriculture in Development. *American Economic Review*, 92(2):160–164, 2002.

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`.

Richard Hornbeck and Pinar Keskin. Does Agriculture Generate Local Economic Spillovers? Short-Run and Long-Run Evidence from the Ogallala Aquifer. *American Economic Journal: Economic Policy*, 7(2):192–213, 2015.

Richard Hornbeck and Suresh Naidu. When the Levee Breaks: Black Migration and Economic Development in the American South. *American Economic Review*, 104(3):963–90, March 2014. doi: 10.1257/aer.104.3.963. URL `http://www.aeaweb.org/articles?id=10.1257/aer.104.3.963`.

Janna E Johnson and Evan J Taylor. The Long Run Health Consequences of Rural-Urban Migration. *Quantitative Economics*, 10(2):565–606, 2019.

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, October 2013. doi: 10.1257/aer.103.6.2169. URL `http://www.aeaweb.org/articles?id=10.1257/aer.103.6.2169`.

Cynthia Kinnan, Shing-Yi Wang, and Yongxiang Wang. Access to Migration for Rural Households. *American Economic Journal: Applied Economics*, 10(4):79–119, 2018.

Piyabha Kongsamut, Sergio Rebelo, and Danyang Xie. Beyond Balanced Growth. *The Review of Economic Studies*, 68(4):869–882, 2001.

Brian K. Kovak. Regional effects of trade reform: What is the correct measure of liberalization? *American Economic Review*, 103(5):1960–76, August 2013. doi: 10.1257/aer.103.5.1960. URL `http://www.aeaweb.org/articles?id=10.1257/aer.103.5.1960`.

David Lagakos and Michael E Waugh. Selection, Agriculture, and Cross-Country Productivity Differences. *American Economic Review*, 103(2):948–80, 2013.

W Arthur Lewis. Economic Development with Unlimited Supplies of Labour. *The Manchester School*, 22(2):139–191, 1954.

Robert E Lucas, Jr. On the Size Distribution of Business Firms. *The Bell Journal of Economics*, pages 508–523, 1978.

Robert E Lucas, Jr. Life Earnings and Rural-Urban Migration. *Journal of political economy*, 112(S1):S29–S59, 2004.

Rodolfo E Manuelli and Ananth Seshadri. Frictionless Technology Diffusion: The Case of Tractors. *American Economic Review*, 104(4):1368–91, 2014.

Brian McCaig and Nina Pavcnik. Export Markets and Labor Allocation in a Low-Income Country. *American Economic Review*, 108(7):1899–1941, 2018.

Guy Michaels, Ferdinand Rauch, and Stephen J Redding. Urbanization and Structural Transformation. *The Quarterly Journal of Economics*, 127(2):535–586, 2012.

Melanie Morten. Temporary Migration and Endogenous Risk Sharing in Village India. *Journal of Political Economy*, 127(1):1–46, 2019.

Kevin M Murphy, Andrei Shleifer, and Robert Vishny. Income Distribution, Market Size, and Industrialization. *The Quarterly Journal of Economics*, 104(3):537–564, 1989.

L Rachel Ngai and Christopher A Pissarides. Structural Change in a Multisector Model of Growth. *American Economic Review*, 97(1):429–443, 2007.

Nathan Nunn and Diego Puga. Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics*, 94(1):20–36, 2012.

Nathan Nunn and Nancy Qian. The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment. *The Quarterly Journal of Economics*, 126(2):593–650, 2011.

Justin R. Pierce and Peter K. Schott. The Surprisingly Swift Decline of US Manufacturing Employment. *American Economic Review*, 106(7):1632–62, July 2016. doi: 10.1257/aer.20131578. URL `http://www.aeaweb.org/articles?id=10.1257/aer.20131578`.

Diego Restuccia and Raul Santaeulalia-Llopis. Land Misallocation and Productivity. Working Paper 23128, National Bureau of Economic Research, February 2017. URL `http://www.nber.org/papers/w23128`.

Scott Rozelle, Loren Brandt, Li Guo, and Jikun Huang. Land Rights in China: Facts, Fictions, and Issues. *China Journal*, 47(1):67–97, 2002.

Theodore William Schultz et al. Transforming Traditional Agriculture. *Transforming Traditional Agriculture.*, 1964.

J Edward Taylor and Philip L Martin. Human Capital: Migration and Rural Population Change. *Handbook of Agricultural Economics*, 1:457–511, 2001.

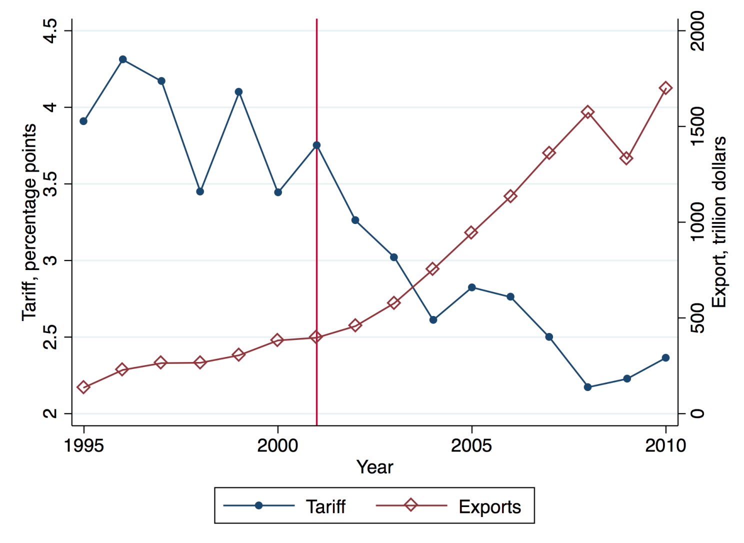
Yuan Tian. International Trade Liberalization and Domestic Institutional Reform: Effects of WTO Accession on Chinese Internal Migration Policy. 2020.

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.

Dennis Tao Yang and Xiaodong Zhu. Modernization of Agriculture and Long-Term Growth. *Journal of Monetary Economics*, 60(3):367–382, 2013.

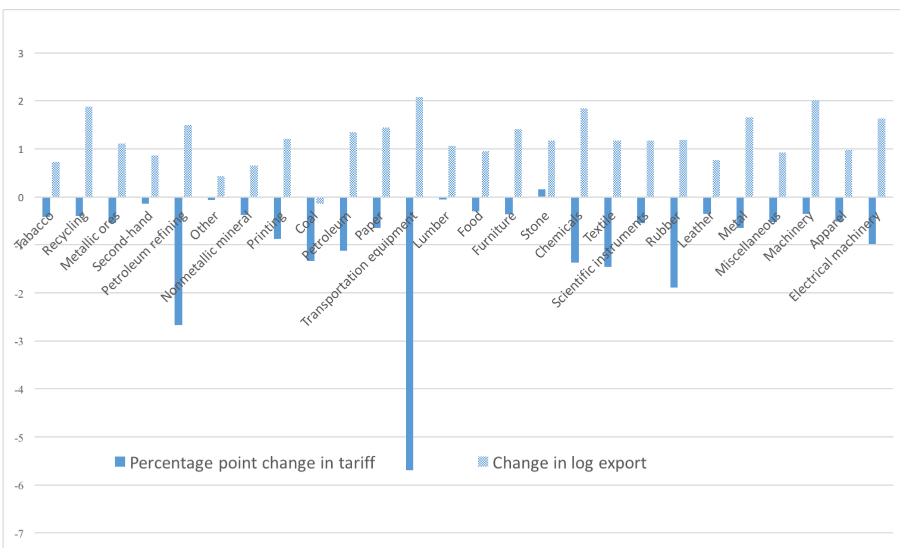
Yuan Zi. Trade Liberalization and the Great Labor Reallocation. Working paper, 2020.

Figure 1: Manufacturing tariffs declined and exports increased, especially after 2001



Note: This graph shows the trends of Chinese exports and tariffs on Chinese exports in the manufacturing sector, using the data from the World Bank TRAINS dataset. The solid dots are the weighted average of industry-level tariffs on Chinese exports in a year, using export values as weights. The diamonds are the total value of Chinese exports in a year. Industry level tariffs are calculated as the weighted average of tariffs on Chinese exports imposed by importing countries, using the 2001 import values as weights.

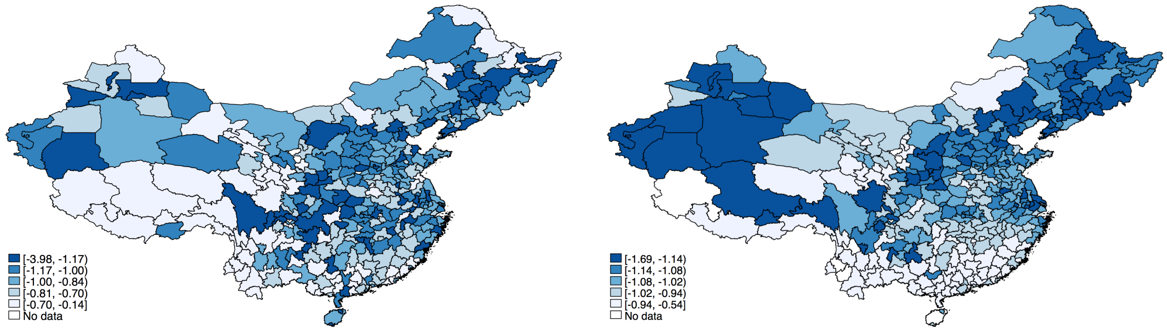
Figure 2: Variation in tariff reductions and export growth by industry, using 2001–2010 as an example



Note: This graph uses the data from the World Bank TRAINS dataset. The shaded light-color bars are the change in the log export from 2001 to 2010, and the dark color bars are the percentage point changes in tariff from 2001 to 2010. The industries are at the 2-digit SIC code level, sorted horizontally by the size of exports in 2001.

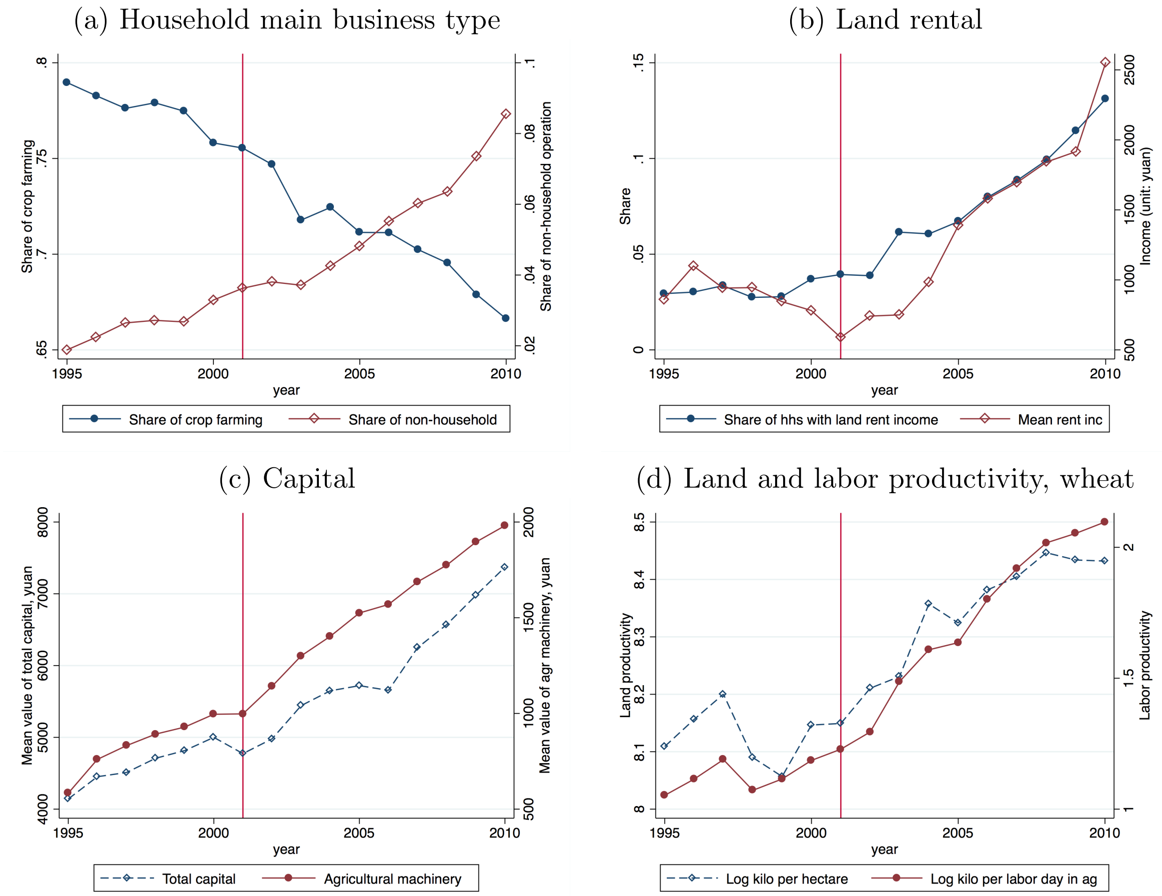
Figure 3: Geographic distribution of own prefecture's output tariff reduction and other prefectures' output tariff reduction, 2001–2010

(a) Own prefecture's output tariff reduction (b) Other prefectures' output tariff reduction



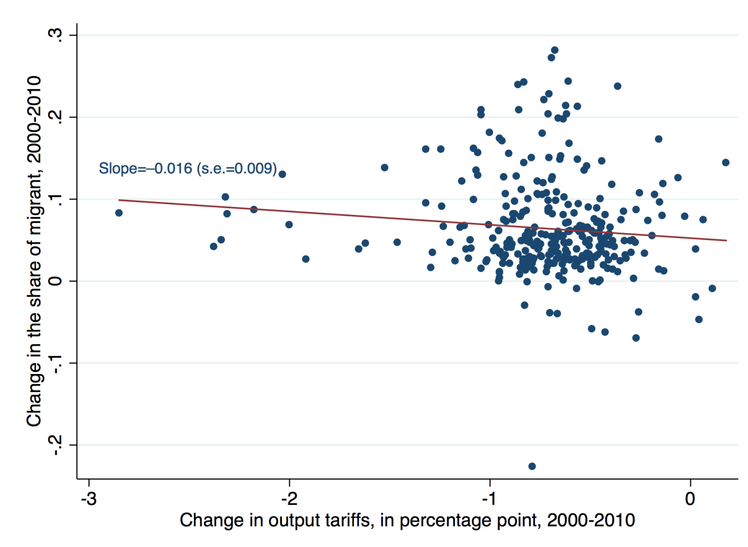
Note: This graph shows the geographical distribution of trade shocks from 2001 to 2010, using the industry-level tariff reductions, the prefecture-level industrial compositions, and the prefecture-to-prefecture migration networks. Each polygon is a prefecture. Panel (a) shows the change in a prefecture's own tariff on Chinese exports from 2001 to 2010 ($\tau^{own}_{2010} - \tau^{own}_{2001}$), and Panel (b) show the change in other prefectures' output tariffs on Chinese exports from 2001 to 2010 ($\tau^{other}_{2010} - \tau^{other}_{2001}$). Darker colors mean larger tariff reductions. In Panel (a), one prefecture (Jiyuan Prefecture in Henan Province) is missing because the prefecture is directly administered by the provincial government, and does not have corresponding information in the industrial survey. In Panel (b), there are 5 other prefectures missing (autonomous regions in Hainan Province, Baoshan, Lijiang, and Lincang in Yunnan Province, and Ngari in Tibet Province) because the 0.095% sample does not have their cross-prefecture migration information.

Figure 4: Trends in the agricultural sector



Note: This figure shows the trends of the agricultural sector, using information from the NFP household-level data. In Panel (a), the solid circles represent the share of households who are in crop-farming in a year, and the hollow diamonds represent the share of households who are in non-household businesses. In Panel (b), the solid circles represent the share of households with land-rental income, and the hollow diamonds represent the mean income from land rental, conditioning on having positive land-rental income. In Panel (c), the solid circles represent the mean value of agricultural machinery, and the hollow diamonds represent the mean value of total capital, both valued in 1995 yuan. In Panel (d), the solid circles represent the log kilo per labor day in agriculture for wheat, and the hollow diamonds represent the log kilo per hectare for wheat.

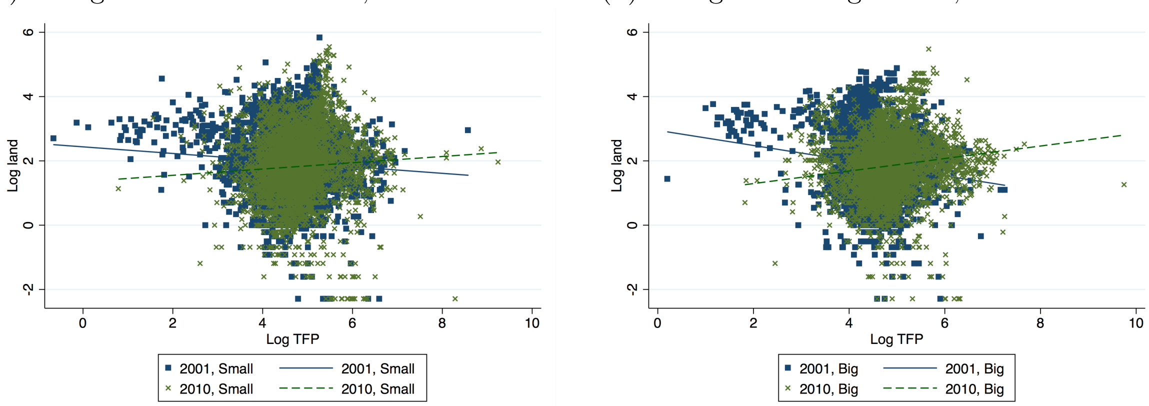
Figure 5: Prefectures with larger declines in output tariffs in the manufacturing sector experienced larger increases in the share of migrants, 2000–2010



Note: This figure shows the relationship between the declines in output tariffs and the changes in the share of migrants from 2000 to 2010, using the output tariffs and the number of migrants using the 2000 and 2010 censuses. Each dot is a prefecture. The horizontal axis shows the percentage point change in output tariffs in the manufacturing sector in a prefecture ($\tau_{2010}^{own} - \tau_{2001}^{own}$), and the vertical axis shows the change in the share of migrants. The pattern is robust to using binned scatter plot and dropping outliers on the left, see in Appendix D.3.

Figure 6: The correlation between land and TFP at the household-level increased more for villages with bigger shocks, 2001–2010

(a) Villages with small shocks, 2001 and 2010 (b) Villages with big shocks, 2001 and 2010



Note: This figure shows the correlation between land and TFP across households within villages, using tariff information and household-level data from the NFP Survey. Each dot is a household-year observation. The squares represent households in 2001 (with the solid line as the linear fitted line), and the crosses represent the households in 2010 (with the dashed line as the linear fitted line). Panel (a) shows the households in villages that experienced small trade shocks from 2001 to 2010. There are 5,411 households in 2001, and 4,033 households in 2010. Panel (b) shows the households in villages that experienced large trade shocks from 2001 to 2010. There are 4,994 households in 2001, and 4,216 households in 2010. The 2001 to 2010 trade shock is defined as the difference between a village's exposure to other prefectures' output tariff in 2001 and 2010, i.e., $(\tau_{2010}^{other} - \tau_{2001}^{other}) \times s$; the shocks above the median magnitude (in absolute values) are defined as large shocks, and the shocks below the median are defined as small shocks.

Figure 7: Marginal costs and products of land for farmers with different s_{ai}

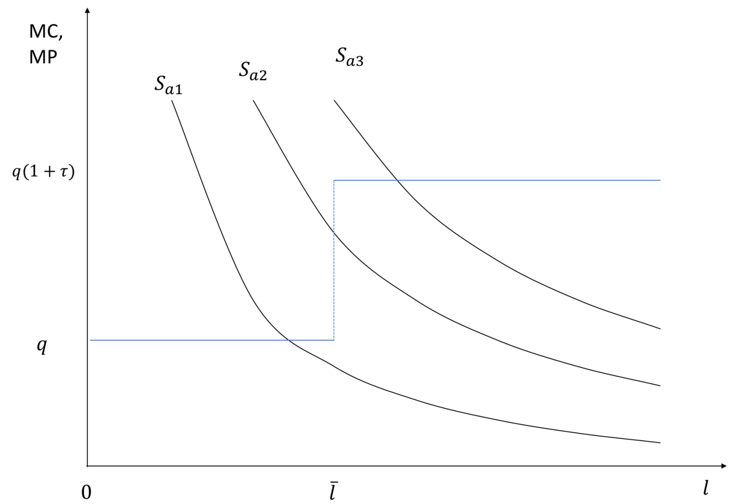
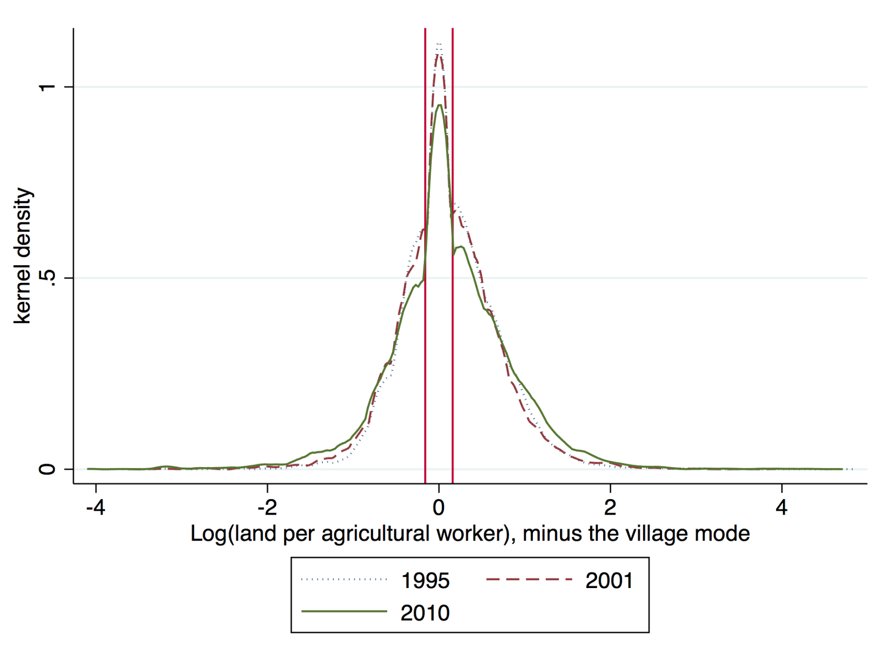


Figure 8: The distribution of land per agricultural worker outspread over time



Note: This figures shows the distribution of the log land per agricultural worker in 1995, 2001, and 2010, using the NFP Survey household level data. The distribution is for individuals in a year. To convert household-level information to individual-level information, we divide the total land in agricultural operation by the number of agricultural workers in the household, and duplicate the observation by the number of agricultural workers. Then we take the village-year mode as the village-year level land endowment, and deduct it from the log land per capita. The dotted line is for 1995, the dashed for 2001, and the solid for 2010. The vertical lines are at -0.16 and 0.16 on the x-axis.

Table 1: Data sources of the occupation and migration related variables

Source	Questionnaire	Information	Variable
NFP (1995–2010)	Household	# of laborers with occupation as wage earner	% non-agricultural laborers
	Village	# of laborers outside the village	
		within county	
		within province, between county	
		between province	
Pop Census 2000	Individual	Residence prefecture in 2000	Prefecture-to-prefecture migr network
		Residence prefecture in 1995	# of cross-pref migr
		Registration place of Hukou	
		Other county, same prefecture	# of within-pref migr

Note: This table summarizes the data sources for occupation and migration related variables. The first panel shows the information from the NFP Survey from 1995 to 2010. The household questionnaire records how many household members are working as wage earners. This information is aggregated at the village-level to construct the share of non-agricultural labor. The village-questionnaire has the number of laborers working outside the village. The second panel shows the information from the 2000 population census. The prefecture-to-prefecture migration network is measured using the current (2000) residence prefecture and the past (1995) residence prefecture. Accordingly, we calculate the number of people who moved across prefectures from 1995 to 2000. We compute the number of within-prefecture migrants as the total number of people whose registration place of Hukou in 2000 is the same as the current residence prefecture, but in different counties.

Table 2: Trade shocks in other regions pulled labor out of agriculture; more so for villages with higher shares of cross-prefecture migrants, and with smaller land-labor ratios

Y: % non-agricultural laborer	(1)	(2)	(3)	(4)	(5)	(6)
Other prefectures' output tariff	-0.08***	-0.06***	-0.06**	-0.04*	-0.16***	-0.16***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)
Other prefectures' output tariff × % cross-pref. migr				-0.04		
				(0.04)		
Log(land/labor) 2001 × Other pref. tariff					0.05***	
					(0.01)	
2nd land/labor quintile 2001× Other pref. tariff						0.03
						(0.02)
3rd land/labor quintile 2001 × Other pref. tariff						0.03
						(0.02)
4th land/labor quintile 2001 × Other pref. tariff						0.07**
						(0.03)
5th land/labor quintile 2001 × Other pref. tariff						0.10**
						(0.03)
Own prefecture tariff		0.02*	0.02*	0.02	0.03***	0.03***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Village-Year Specific Controls	No	No	Yes	Yes	Yes	Yes
Observations	2,333	2,333	2,333	2,333	1,971	1,971
R-squared	0.85	0.85	0.85	0.85	0.85	0.85

Note: This table shows the impact of other prefectures' output tariff on occupation choices of residents of a village. All columns control for province-year fixed effects and village fixed effects. Column (1) regresses the non-agricultural labor share on other prefectures's tariff. Column (2) adds the own prefecture's output tariff. Column (3) add controls, including the log labor, the log number of households, and the log government transfer +1. Column (4) adds the interaction between other prefectures' output tariff and the share of cross-prefecture migrants. Column (5) replaces the interaction term in Column (4) with the interaction between other prefectures' output tariff and the 2001 log land-to-labor ratio in agriculture. Column (6) replaces the interaction term in Column (4) with the interaction between other prefectures' output tariff and quintiles of the 2001 land-to-labor ratio. The mean (sd) of the share of non-agricultural labor is 0.18 (0.14), the mean (sd) of the log of land per agricultural worker in 2001 is 1.19 (0.67), and the mean (sd) of other prefectures' output tariff is 3.07 (0.44). Standard errors are clustered at the province and the year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: Places with higher land-labor ratios had more fluid land markets and higher allocation efficiency, 213 villages in 2001

	(1)	(2)	(3)
	% of land leased	Log(land leased+1)	
		Stock	Flow
Log(land/agr labor)	0.04**	1.06***	0.91***
	(0.01)	(0.25)	(0.24)
	(4)	(5)	(6)
	Ruggedness	% non-agr labor	Allocation efficiency
Log(land/agr labor)	-37.39**	0.04*	0.23***
	(16.66)	(0.02)	(0.07)

Note: This table presents the correlation between the land-to-land ratio and village characteristics in 2001. All columns control for province fixed effects. Each column (1–6) represents a separate regression of a characteristic of a village in 2001 on the log land-to-labor ratio in 2001. Standard errors are clustered at the province level. *** p<0.01, ** p<0.05, * p<0.

Table 4: Land markets became more fluid in villages with larger declines in other prefectures' output tariff

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(land leased+1)				Log(land lease income+1)	
	Stock		Flow			
Other pref. tariff	-0.60***	-0.10	-1.73***	-1.10**	-3.04**	-2.37**
	(0.15)	(0.32)	(0.46)	(0.41)	(1.03)	(0.92)
Other pref. tariff		-1.40**		-1.78**		-1.87
× % cross-pref. migr		(0.54)		(0.62)		(1.97)
Own pref. tariff	-0.09*	-0.14***	0.10***	0.04	-0.27	-0.34
	(0.05)	(0.04)	(0.02)	(0.05)	(0.50)	(0.49)
Observations	2,333	2,333	2,333	2,333	2,333	2,333
R-squared	0.83	0.84	0.69	0.70	0.67	0.67

Note: This table shows the impact of other prefectures' output tariff on the land rental market of a village. All columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects, and village fixed effects. Columns (1)(3) and (5) have the same specification as Table 2 Column (3), and Columns (2)(4) and (6) have the same specification as Table 2 Column (4). The mean (sd) of the log stock of land leased is 3.03 (1.63), the mean (sd) of the log flow of land leased is 1.88 (1.63), and the mean (sd) of the log income from land leasing is 6.14 (4.17). Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: No significant impact on average land size; some effect on the number of households with relatively large land

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Log(land)		Log(land per worker)		Log(# of hhs>1/3 ha)	
Other pref. tariff	0.12	0.18	-0.02	-0.01	0.15	0.42*
	(0.13)	(0.17)	(0.16)	(0.18)	(0.16)	(0.22)
Other pref. tariff × % cross-pref. migr		-0.16		-0.02		-0.73**
		(0.18)		(0.15)		(0.25)
Own pref. tariff	-0.04	-0.05	-0.03	-0.03	0.01	-0.02
	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)	(0.04)
Observations	2,333	2,333	2,333	2,333	2,333	2,333
R-squared	0.96	0.96	0.93	0.93	0.95	0.95

Note: This table shows the impact of and other prefectures' output tariff on the total land size and the distribution of land of a village. All columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects, and village fixed effects. Columns (1)(3) and (5) have the same specification as Table 2 Column (3), and Columns (2)(4) and (6) have the same specification as Table 2 Column (4). The mean (sd) of the log land is 5.81 (0.89), the mean (sd) of the log land per agricultural worker is 1.35 (0.71), and the mean (sd) of the log number of households with relatively large land is 2.95 (1.18). Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Land allocated more towards initially productive households, 2001–2010

Y: log(land) in year t	(1)	(2)	(3)	(4)
Initial TFP (log TFP in 2001)	0.281***		0.201***	
	(0.039)		(0.041)	
Other pref. tariff	0.185	0.155		
	(0.157)	(0.169)		
Other pref. tariff ×Initial TFP	-0.061***	-0.054***		
	(0.013)	(0.015)		
Other pref. tariff*%cross-pref. migr			0.097	0.422**
			(0.166)	(0.171)
Other pref. tariff*%cross-pref. migr			-0.079**	-0.152***
× Initial TFP			(0.028)	(0.038)
Own pref. tariff	-0.033	-0.032	-0.046	-0.044
	(0.043)	(0.040)	(0.042)	(0.040)
Observations	103,027	102,262	103,027	102,262
R-squared	0.631	0.871	0.631	0.871
Province-Year FE	Yes	Yes	Yes	Yes
Village FE	Yes		Yes	
HH FE		Yes		Yes

Note: This table shows the land allocation across households within a village, in response to trade shocks. Column (1) regresses the log land size of households in year t on the household's TFP in 2001, other prefectures' output tariff, and the interaction of the two. We control for own prefecture's output tariff, province-year fixed effects, and village fixed effects. Column (2) has the same specification, except that we replace village fixed effects with households fixed effects. Columns (3) and (4) replicate Columns (1) and (2), replacing other prefectures' output tariff with the interaction of other prefectures' output tariff and the share of cross-prefecture migrants. The mean (sd) of the log land is 1.78 (0.99). The mean (sd) of the log TFP in 2001 is 4.62 (0.64), the mean (sd) of the log TFP is 4.72 (0.65), the mean (sd) of other prefectures' output tariff is 3.07 (0.45), and the mean (sd) of the product of other prefectures' output tariff and the share of cross-prefecture migrants is 1.41 (0.59). Standard errors are clustered at the province and the year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7: Fewer labor days supplied locally, higher local wages

	(1)	(2)	(3)	(4)
	Log(hired labor days)		Log(wage of hired labor)	
Other pref. tariff	-1.56	-1.99	0.16	0.37
	(1.49)	(1.55)	(0.35)	(0.38)
Other pref. tariff × % cross-pref. migr		1.27*		-0.64
		(0.59)		(0.39)
Own pref. tariff	-0.12	-0.09	0.00	-0.01
	(0.22)	(0.22)	(0.06)	(0.06)
Observations	1,879	1,879	1,742	1,742
R-squared	0.68	0.68	0.67	0.68

Note: This table shows the impact of other prefectures' output tariff on the hired-labor market of a village. All Columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects, and village fixed effects. Columns (1) and (3) have the same specification as Table 2 Column (3), and Columns (2) and (4) have the same specification as Table 2 Column (4). The mean (sd) of the log hired labor days is 5.81 (1.73), and the mean (sd) of the log wage for hired labor is 3.26 (0.75). Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: More machinery adopted, especially for households with large land

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Log(agr machine)		Log(# of hhs with positive ag machine & land larger than 1/3 hectare)			
Other pref. tariff	-0.07	0.34	-0.23	0.18	0.03	0.46
	(0.76)	(0.79)	(0.38)	(0.37)	(0.44)	(0.33)
Other pref. tariff × % cross-pref. migr		-1.15**		-1.23***		-1.39**
		(0.47)		(0.35)		(0.46)
Own pref. tariff	0.03	-0.01	-0.08	-0.12*	-0.05	-0.10*
	(0.06)	(0.08)	(0.05)	(0.07)	(0.04)	(0.05)
Log(# of hhs with ag machine>0					0.07	0.07
& land<1/3 ha)					(0.06)	(0.06)
Observations	2,333	2,333	2,181	2,181	1,413	1,413
R-squared	0.87	0.87	0.88	0.88	0.91	0.91

Note: This table shows the impact of other prefectures' output tariff on the capital market of a village. All columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects, and village fixed effects. Columns (1)(3) and (5) have the same specification as Table 2 Column (3), and Columns (2)(4) and (6) have the same specification as Table 2 Column (4). The mean (sd) of the log agricultural machinery is 10.18 (1.69), and the mean (sd) of the log number of household with positive value of agricultural machinery and have relatively large land is 2.00 (1.20). Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Weighted village TFP and allocation efficiency increased

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Log(village TFP)				Allocation efficiency	
	output weighted		unweighted			
Other pref. tariff	-0.69***	-0.45**	0.04	0.09	-0.74***	-0.54**
	(0.05)	(0.14)	(0.22)	(0.23)	(0.06)	(0.17)
Other pref. tariff × % cross-pref. migr		-0.69		-0.13		-0.55
		(0.59)		(0.08)		(0.55)
Own pref. tariff	0.01	-0.02	0.00	-0.00	0.00	-0.01
	(0.09)	(0.08)	(0.03)	(0.03)	(0.06)	(0.06)
Observations	2,333	2,333	2,333	2,333	2,333	2,333
R-squared	0.65	0.65	0.84	0.84	0.58	0.58

Note: This table shows the impact of other prefectures' output tariff on the village-level productivity and allocation efficiency of a village. All columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects, and village fixed effects. Columns (1)(3) and (5) have the same specification as Table 2 Column (3), and Columns (2)(4) and (6) have the same specification as Table 2 Column (4). The mean (sd) of the log weighted TFP is 4.52 (0.90), the mean (sd) of the log unweighted TFP is 4.75 (0.51), and the mean (sd) of allocation efficiency is -0.23 (0.82). Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Characteristics of productive farmers and productive non-agricultural workers (2003–2008), individual-level evidence

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(TFP) in agriculture			Log (income) from working outside village		
Education	0.035***	0.005***		0.056***	0.039***	
	(0.005)	(0.001)		(0.004)	(0.002)	
Non-agricultural training (=1)	-0.080***	-0.016		0.390***	0.303***	
	(0.026)	(0.013)		(0.020)	(0.016)	
Agricultural training (=1)	0.160***	0.011		-0.155***	-0.044*	
	(0.048)	(0.017)		(0.034)	(0.024)	
Age	0.002**	-0.000**		-0.001	-0.003***	
	(0.001)	(0.000)		(0.001)	(0.001)	
Log (TFP), 2001			0.177***			-0.008
			(0.013)			(0.018)
Village FE	No	Yes	Yes	No	Yes	Yes
Observations	127,029	127,029	165,515	85,104	85,104	63,000
R-squared	0.034	0.556	0.567	0.108	0.310	0.279

Note: This table shows the characteristics of individuals that are correlated with their households' agricultural TFP and their income when they work in non-agriculture. We include individuals who are not currently in school and aged 16 to 75. All columns control for year fixed effects and the log number of labor in the household. The outcome variable in Columns (1)–(3) is the log TFP, and the outcome variable in Columns (4)–(6) is the log income from working outside the village. Columns (1) and (4) do not include village fixed effects, and Columns (2)(3)(5) and (6) include village fixed effects. The mean (sd) of the log TFP is 4.77 (0.64) and the mean (sd) of the log income from working outside the village is 8.50 (0.95). The mean (sd) of age is 41 (15), the mean (sd) of education is 6.90 (2.29), the mean (sd) of the share of individuals with occupational training is 0.08 (0.27), and the mean (sd) of the share of individuals with agricultural training is 0.07 (0.25). Standard errors are clustered at the village-year level and at the household level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 11: Cross tabulation of individuals with non-agricultural training and agricultural training (2003–2008)

		Agricultural training	
		No	Yes
Non-agricultural training	No	87%	5%
	Yes	6%	2%

Note: This table tabulates the status of individuals in terms of agricultural training and non-agricultural training, using the NFP Survey data on individuals from 2003 to 2008. We include individuals who are not currently in school and aged 16 to 75. There are 256,184 observations in total.

Table 12: Households with lower initial productivity were more likely to switch to non-agricultural, inverse-U shaped relationship

Y: Number of non-agricultural laborers	(1)	(2)	(3)	(4)
Log(TFP), 2001	-0.043***	0.048**	-0.044***	0.051**
	(0.006)	(0.024)	(0.006)	(0.024)
Log(TFP), 2001, squared		-0.011***		-0.011***
		(0.003)		(0.003)
Number of laborers	0.251***	0.251***	0.255***	0.256***
	(0.005)	(0.005)	(0.005)	(0.005)
Constant	0.019	-0.174***	0.010	-0.191***
	(0.030)	(0.056)	(0.030)	(0.056)
Observations	110,244	110,244	110,243	110,243
R-squared	0.344	0.344	0.382	0.382
Year FE	Yes	Yes		
Village-Year FE			Yes	Yes
Village FE	Yes	Yes		

Note: This tables shows the correlation between initial agricultural TFP and the number of laborers choosing non-agriculture for households. Column (1) regresses the number of non-agricultural laborers of household h in year t on its TFP in year 2001, controlling for year fixed effects and village fixed effects. Column (2) adds the quadratic term of the initial TFP. Columns (3) and (4) replicates Columns (1) and (2), using village-year fixed effects instead. The mean (sd) of the number of non-agricultural laborers is 0.51 (0.87), and the mean (sd) of the number of laborers is 2.70 (1.29). The mean (sd) of the log TFP in 2001 is 4.62 (0.64). Standard errors are clustered at the village-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Households with lower initial productivity were more likely to switch to non-agricultural in response to trade shocks

	(1)	(2)	(3)	(4)
	TFP in 2001			
Y: Number of non-agricultural laborers	$<$Median	$\geq$Median	$<$Median	$\geq$Median
Other pref. tariff × % cross-pref migr	-0.40**	-0.28	-0.38**	-0.23
	(0.13)	(0.17)	(0.13)	(0.18)
Own pref. tariff	0.02	0.05	0.02	0.06
	(0.04)	(0.05)	(0.04)	(0.05)
Number of laborers	0.23***	0.23***	0.25***	0.25***
	(0.02)	(0.02)	(0.02)	(0.02)
Observations	55,715	53,904	56,050	54,192
R-squared	0.65	0.63	0.37	0.35
HH FE	Yes	Yes		
Village FE			Yes	Yes

Note: This table shows the responsiveness of households to trade shocks. All columns control for province-year fixed effects. Column (1) regresses the number of non-agricultural labor on the interaction of other prefectures' output tariff and the share of cross-prefecture migrants, own prefecture's output tariff, and total number of labor. Column (1) controls for household fixed effects, and use only the households with TFP in 2001 above median. Column (2) uses the households with TFP in 2001 below median instead. Columns (3) and (4) replicate Columns (1) and (2), replacing the household fixed effects with village fixed effects. The mean (sd) of other prefectures' output tariff × the share of cross-pref migrants is 1.41 (0.59), the mean (sd) of the number of non-agricultural laborers is 0.51 (0.87), and the mean (sd) of the number of laborers is 2.70 (1.29). Standard errors are clustered at the village and the year level. *** p$<$0.01, ** p$<$0.05, * p$<$0.1.

Table 14: Calibrated parameters

Year	2001	2010
Variance of sector-specific productivity, σ_a^2	0.7172	
Covariance between productivity in different sectors, σ_{an}	0.1243	
Land market transaction cost, τ	1.6	1.2
Agricultural productivity, $(1-\gamma)\log A_a$	6.8	7.2
Non-agricultural productivity, $\log A_n$	8.1	9.6

Note: This table shows the calibrated parameter values.

Table 15: Counterfactuals for 2010

Aggregate statistics	Benchmark Economy	C1 $\tau = 0$	C2 $A_n^{2010} = A_n^{2001}$
Real agricultural productivity (output per person)	1	1.09	0.69
Share of employment in agriculture	0.49	0.48	0.90
TFP in agriculture	1	1.06	0.91
Real non-agricultural productivity (output per person)	1	0.93	0.36
Average ability in agriculture (in log)	1	1.48	0.34
Average ability in non-agriculture (in log)	1	0.79	1.84
GDP per worker	1	1.02	0.83
Total agricultural capital per agricultural employment	1	1.09	0.69

Note: This table shows the results of counterfactual analysis for 2010. The column BE is the benchmark economy of 2010, and Columns C1 and C2 represent different counterfactual analysis. In Column C1, we set the distortion to be 0. In Column C2, we set the non-agricultural productivity in 2010 to be the same as the non-agricultural productivity in 2001. Real agricultural and non-agricultural productivity are the output per worker in the agricultural sector and the non-agricultural sector, respectively. TFP in agriculture is defined as $\frac{\sum y_{ai}}{\left(\sum_{i\in H_a} 1\right)^{1-\gamma} \text{land}^{\alpha\gamma} (\sum k_i)^{(1-\alpha)\gamma}}$. The average ability in agriculture is $\frac{\sum_{i\in H_a} s_{ai}}{\sum_{i\in H_a} 1}$, and the average ability in non-agriculture is $\frac{\sum_{i\in H_n} s_{ni}}{\sum_{i\in H_n} 1}$. GDP per worker is defined as the total value of output $p_a \sum y_{ai} + \sum y_{ni}$ divided by the total number of workers.

For Online Publication

A Appendix on the Institutional Context

A.1 Hukou and the Nature of Internal Migration in China

The Hukou status is tied with access to job opportunities and local public goods, and importantly for rural workers, the title of agricultural land. Switching Hukou to a new prefecture-sector is possible for college graduates who find jobs that sponsor the change, but it is much more difficult for low skilled workers. In addition, without a permanent Hukou, a stable job perspective and housing in the city, a rural resident who works in the city is not willing to switch to urban Hukou since the change comes with loss of land title. In this sense, the migration is not permanent. However, unlike in other developing country contexts (for example, Morten [2019]), consumption smoothing in the event of adverse productivity shocks in agricultural production is not the main motivation to migrate. Thus, the migration is not the typical temporary migration either. From the statistics in the NFP dataset, comparing households with non-agricultural wage earners in 2001 and households without, the probability of having at least one wage-earner outside agriculture in 2010 is 54% higher.

A.2 Rural Land Market Regulations

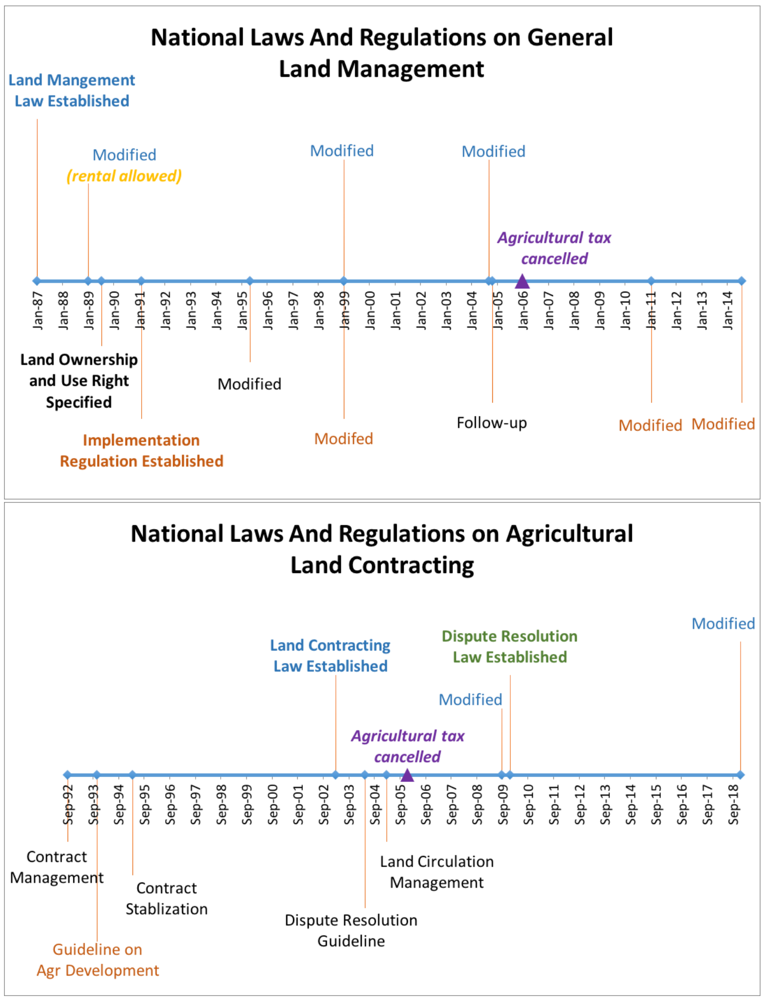
National Laws and Regulations Pertaining Rural Land The first set of laws on land management are at the national level, regarding land in both the rural and the urban areas.

The *Land Administration Law of the People's Republic of China* was first enacted on January 1, 1987, and was modified in 1988, 1998, and 2004. The 1988 modification legalized the transfer of land use rights. The 1998 modification emphasized forbidding the change of land use type, especially from agricultural land to construction land. Another important change was specifying that "within the term of land contractual operation, if the land contracted by individual contractors need to be adjusted, it must be approved by more than two-thirds of the members of the village meeting or more than two-thirds of the village representatives, and reported to the township (town) government and the county-level agricultural administrative department." The same procedure applies when units or individuals outside the village collective want to contract the farm land. The 2004 modification said that when the government expropriate land for the benefit of public interest, compensations need to be provided correspondingly.

The Land Administration Law has corresponding implementation regulations for details of implementation and interpretation. The first implementation regulation was enacted on February 1, 1991. More details were added in the 1998 modification, and minor changes were made in the 2011 and 2014 modifications. The implementation regulations were supplemented by several other regulations regarding practical issues in land ownership and use rights, in 1989, 1995, and 2004.

The second set of laws and regulations are specific to rural land, especially farmland. The *Law of the People's Republic of China on the Contracting of Rural Land* was enacted on March 1, 2003. It formalized legal rights of the contractors and the contract issuing party, the principle and procedure

Figure A.1: Land Laws and Regulations After the Establishment of HCRS



Note: These figures show the timing of enactment of laws and regulation at the national level on land-related issues, after the establishment of the HCRS. The top graph shows the laws and regulation on general land management, including both urban and rural land. The bottom graph shows the laws and regulations targeted at agricultural land contracting. The ones shown above the time line are laws, and the ones below the time line are regulations.

of contracting, terms of contracts, the protection of use rights of the contracted land, the protection of the transfer of use rights, and dispute resolution method and legal responsibilities. Importantly, on the issue of out-migration, during the contract period, if the contractor (i.e. the rural household) move to townships, the land use right remains, and the land can be rented to other households. If the contractor move to cities and obtain urban Hukou, then the contract terminates and land is given back to the contract issuing party.

Before that, several regulatory documents addressed some of these aspects, but with less details. The 1992 regulation on contract management mentioned that out of thirty-one provinces, seven provinces established local laws on contracting, while seventeen provinces issued related regulations. The 1993 regulation initiated the grain market reform, where the price of grain purchased by the government will be the same as the market price, and at the same time, the government provides a price floor. The 1995 regulation can be seen as the precursor of the 2003 law. There are several important points made: (1) Land contracts should be extended by another 30 years once the term ends; (2) The government encourages long-term contract relation between the village commune and farmers, with the principle that contract land sizes do not respond to changes in the household size; (3) In the case of substantial population changes or land occupation, adjustments at the village level should be agreed by the majority of villagers, and approved by the county government; (4) Land use right transfers are part of contract management, and transfers should have written contracts.

After the 2003 law was enacted, a 2005 notice on land circulation management specified additional execution issues. Minor modifications were made in 2009 and 2018.

Related, the *Law of the People's Republic of China on the Mediation and Arbitration of Rural Land Contract Disputes* was enacted on January 1, 2010. It formalized the procedure of mediation and arbitration, and noted that the costs of such dispute resolution are covered by and included in the government budget. This 2010 law was built on the guideline on dispute resolution issued in 2004.

Regional Laws and Regulations Vary by province.

B Data Appendix

B.1 Administrative Units in China

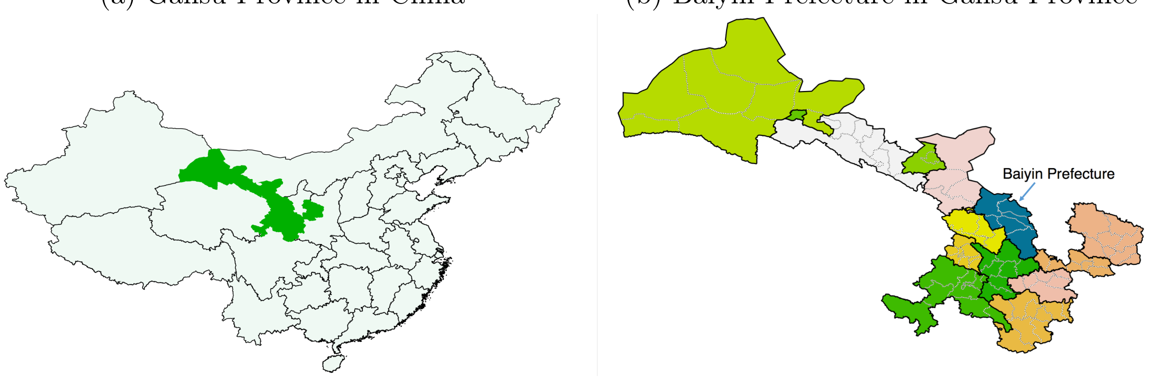
Table A.1: Levels of administrative units, 2000

Level	Number
Province	31
Prefecture	333
County	2,861
Village (1996)	748,340

Note: This table shows the hierarchy of Chinese administrative units. The units ranging from the largest to the smallest are: province, prefecture, county, and village. The number of provinces, prefectures, and counties are available in all sample years from National Bureau of Statistics of China, `data.stats.gov.cn`. The number of villages is available in 1996 only through the First Agricultural Census, `http://www.stats.gov.cn/tjsj/pcsj/nypc/dycnypc/200308/t20030826_39912.html`.

Figure A.2: Example of Baiyin Prefecture in Gansu Province in China

(a) Gansu Province in China (b) Baiyin Prefecture in Gansu Province



Note: This figure shows the hierarchy of Chinese administrative units using Gansu province as an example. Panel (a) shows the location of Gansu Province in China, along with other 30 provinces. Panel (b) show the location of Baiyin Prefecture in Gansu, along with other 13 prefectures. The light color borders within the prefectures are county borders.

B.2 Summary Statistics of the NFP Data

B.2.1 Sample Size and Coverage

We keep the villages with at least 20 households and at least 20 laborers in total. We also require that there is at least one household with land larger than 1/3 hectare and at least one household with agricultural machinery. Overall, in the 2001 to 2010 period, we have 295 villages, 2,333 village-year observations, and 148,327 household-year observations.

We find no evidence of selective attrition of households. We generate a dummy D_{hvt} that is equal to one if the household h is in a village(v)-year(t) sample, and zero otherwise, given that the village-year is in the sample, and the household is in at least one of the years between 2001 and 2010. Then we run the following regression:

$$D_{h(v)t} = \gamma_0 + \gamma_1 \tau^{other}_{v(i)t} + \gamma_2 \tau^{own}_{v(i)t} + I_{pt} + I_v + \epsilon_{hvt},$$

where $\tau^{own}_{v(i)t}$ and $\tau^{other}_{v(i)t}$ are village v's exposure to its own prefecture i's tariff and other prefectures' output tariff, respectively, I_{pt} are province-year fixed effects, and I_v are village fixed effects.

The regression result is shown in Table A.2. There is no significant effect of either own prefecture's output tariff and other prefectures' output tariff. A joint test of $\gamma_1 = \gamma_2 = 0$ generates a F-statistics of 0.06, and a p-value of 0.94. Thus, we fail to reject that there is no selective attrition.

Table A.2: No selective attrition

	(1)
	Dummy(=1 if the household is in the sample)
Other pref. tariff	-0.55
	(1.57)
Own pref. tariff	-0.36
	(1.16)
Constant	0.90***
	(0.08)
Observations	171,959
R-squared	0.18

Note: The table shows the regression results of sample attrition. The outcome variable is a dummy that is equal to one if a household shows up in a year, conditioning on the village shows up in the year, and the household ever shows up in the 2001–2010 sample. The regressors are the village's own prefecture's output tariff and other prefectures' output tariff. Province-year fixed effects and village fixed effects are controlled. Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

B.2.2 Summary of Statistics of Key Variables

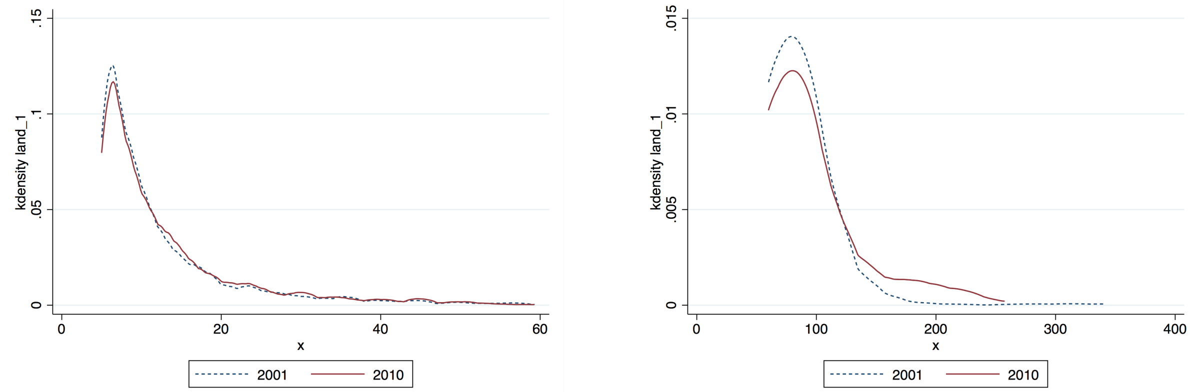
Table A.3: Summary of statistics of key variables

VARIABLES	N	mean	sd	min	max
Log(land)	2,333	5.82	0.89	2.30	9.12
Log(land leased+1), stock	2,333	3.03	1.63	0.00	8.54
Log(land leased+1), flow	2,333	1.88	1.63	0.00	8.53
Log(income from land leasing+1)	2,333	6.14	4.18	0.00	14.54
Log(land p.c.)	2,333	1.35	0.71	-1.91	5.18
Log(# of households with land >1/3 ha)	2,333	2.96	1.18	0.00	4.92
% non-agricultural laborer	2,333	0.18	0.14	0.00	0.87
Log(# laborer)	2,333	5.08	0.38	3.22	6.35
Log(# of households)	2,333	4.09	0.36	3.00	5.30
Log(gov transf+1)	2,333	9.18	2.04	0.00	13.60
Log(labor days), hired labor	1,879	5.81	1.73	0.00	12.62
Log(wage), hired labor	1,765	3.43	0.71	-1.39	7.79
Log(agricultural machinery)	2,333	10.18	1.69	0.00	16.65
Log(# HHs with positive ag machinery and land>1/3 ha.)	2,190	2.10	1.20	0.00	4.61
Log(# HHs with positive ag machinery and land<1/3 ha.)	2,016	2.46	1.16	0.00	4.54
Village TFP, weighted	2,333	4.52	0.90	0.01	12.92
Village TFP, unweighted	2,333	4.75	0.51	1.62	6.92
Allocation efficiency	2,333	-0.23	0.82	-6.66	8.02
Cash crop revenue/crop revenue	2,333	0.43	0.30	0.00	1.00
HHs with cash crops/HH with crops	2,333	0.76	0.33	0.00	1.00
Log(# of HHs with cash crops)	2,226	3.53	0.96	0.00	4.94
Vegetable revenue/crop revenue	2,333	0.24	0.25	0.00	1.00
HHs with vegetables/HH with crops	2,333	0.58	0.41	0.00	1.00
Log(# of HHs with vegetables)	2,002	3.18	1.24	0.00	4.85
Log(husbandry output value)	2,266	11.77	1.34	5.03	20.80
Log(# of HHs in husbandry)	2,266	3.17	1.03	0.00	5.05
Log(labor days in husbandry)	2,243	7.68	1.20	2.30	13.72
Own pref. tariff	2,333	3.05	0.92	0.72	8.18
Other pref. tariff	2,333	3.07	0.44	1.53	4.47
Other pref. tariff X % cross-pref. migr	2,333	1.43	0.58	0.19	3.27

Note: This figure shows the summary of statistics of key variables used in the empirical analysis. Overall, there are 2,333 village-year observations used in the main analysis. Some variables in logs have fewer numbers of observations due to zero values.

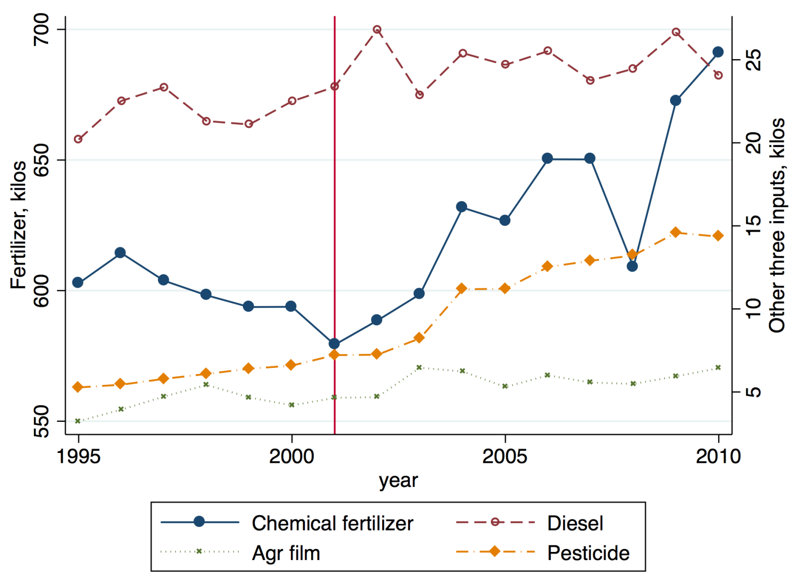
Figure A.3: Larger land sizes for households with 1/3 hectare and more, 2001 and 2010

(a) Land size between 1/3 hectare and 60 hectares (b) Land size larger than 60 hectares



Note: This figure plots the distribution of land for households with land larger than 1/3 hectare, using data from the NFP Survey. Panel (a) shows the distribution of land sizes conditioning on the land size being larger than one third hectare and smaller than 60 hectares. Panel (b) shows the distribution of land conditioning on the land size being larger than 60 hectares. Dotted lines are for 2001, and the solid lines are for 2010.

Figure A.4: Other inputs also increased



Note: This figure shows the trends of costs of intermediate inputs in agriculture, using the NFP Survey on households. Each dot is a household-year average across all households with positive land size. The solid line with solid dots represents chemical fertilizers, the dashed line with hollow circles represent diesel, the dotted line with small circles represents agricultural plastic films, and the dashed line with solid diamonds represents pesticides.

B.2.3 Perpetual Inventory Method for Capital

Initial-year real capital. Keep the initial year y_o of each household to enter the dataset. Assume that capital is zero for all households in 1986, and that for all years between 1986 and y_o, the capital growth rate of each household is the same as the national annual growth rate of capital. With the annual price index of capital (setting the 1995 index as 1) and the capital book value in y_o, impute the real investment for years between 1986 and y_o. Then reconstruct the capital stock in these years taking into account depreciation, with a depreciation rate of 0.04.[48] The result is the real value of capital stock in y_o.

Subsequent-year real capital. If capital is missing (or zero) in year t, and not missing (or zero) in year $t-1$, then use the capital growth rate from t to $t-1$ to impute for year t. If capital is also missing (or zero) in year $t-1$, use the $t-2$ for imputation. Then calculate the real investment. First, generate the nominal investment as the difference in nominal capital in year t and $t-1$. If the year $t-1$ capital is missing, use the closest non-missing year to calculate the annual difference. Second, deflate the nominal investment with the price index to generate real investment. Third, use the depreciation rate and nominal capital to generate nominal depreciation. Generate net nominal investment as the difference between nominal investment and nominal depreciation. Then use the price index to deflate the net nominal investment to generate net real investment. Fourth, use the y_o real capital and the net real investment to generate real capital series.

B.3 Industry Distribution of Workers in Agriculture and Non-Agriculture

Table A.4: The industry distribution of people with agriculture/non-ag as occupation, using the 2003–2010 individual data

	(1)	(2)	(3)	(4)
	Agricultural labor		Non-ag laborer	
Industry	Freq	Percent	Freq	Percent
Agriculture	251,042	66%	2,732	3%
Industry	16,588	4%	38,397	39%
Construction	9,131	2%	15,119	15%
Transportation	9,923	3%	5,042	5%
Service	24,177	6%	17,833	18%
Other	66,831	18%	19,809	20%
Total	377,692		98,932	

Note: This table presents the distribution of industries for agricultural and non-agricultural laborers, using the information from individual questionnaires in the NFP Survey from 2003 to 2010.

B.4 Industry Crosswalk, from 2-digit GB Code to 2-digit SIC Code

The 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 costs, and total employment.[49] There are 145,546 firms in 2000 with

[48]This is the sample mean of the depreciation rate of all NFP households from 1995 to 2002. There is no such information from 2003 onward.

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

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 we aggregate to the 2-digit level. The 2-digit Chinese Industry Code is slightly finer than the 2-digit SIC code, with the crosswalk between the codes shown in Table A.5. The definition of the primary metal products and fabricated metal products are different in the Chinese industry code and the SIC code, so we combined the two industries into metal industry.

Table A.5: Crosswalk, 2-digit Chinese industry code (GB) to 2-digit U.S. industry code (SIC), manufacturing

GB code	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 nonferrous 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	Manu. of foods	20	Food and kindred products
15	Manu. of liquor beverages and refined tea	20	Food and kindred products
16	Manu. of tobacco	21	Tobacco products
17	Manu. of textile	22	Textile mill products
18	Manu. of textile fabrics wearing apparel and accessories	23	Apparel and related products
19	Manu. of leather, fur, feather and related products and footwear	31	Leather and leather products
20	Processing of timber, manufacture of wood, bamboo, rattan, palm and straw products	24	Lumber and wood products, except fuel
21	Manu. of furniture	25	Furniture and fixtures
22	Manu. of paper and paper products	26*	Paper and allied products
23	Printing, production of recording media	27	Printing, publishing, and allied products
24	Manu. of articles for culture, education, art, sport, and entertainment activities	26*	Paper and allied products
25	Processing of petroleum and coking	29	Petroleum refining and related products
26	Manu. of raw chemical material and chemical products	28	Chemicals and allied products
27	Manu. of medicines	28†	Chemicals and allied products
28	Manu. of chemical fibers	28	Chemicals and allied products
29	Manu. of rubber	30	Rubber and miscellaneous plastics products
30	Manu. of plastics products	28‡	Chemicals and allied products
31	Manu. of non-metallic mineral products	32	Stone, clay, glass, and concrete products
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	Manu. of metal products	300⋇	Metal processing and products
35	Manu. of general purpose machinery	35	Machinery, except electrical
36	Manu. of special purpose machinery	35	Machinery, except electrical
37	Manu. of transportation machinery	37	Transportation equipment
39	Manu. of electrical machinery and equipment	36	Electrical machinery, equipment, and products
40	Manu. of communication equipment, computers and other electric equipment	36	Electrical machinery, equipment, and products
41	Manu. of measuring instruments and machinery	38	Scientific and professional instruments
42	Manu. of artifacts and other manufacturing	39	Miscellaneous manufactured commodities
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).

B.5 Tariffs and Trade in the Agricultural Sector

The data on tariffs and national import and exports by crop are from TRAINS, with 3-digit SIC codes matching to the 11 crops in our data. The trends of tariffs, imports, and exports are shown in Figure A.5 (wheat, rice, corn, and soybean), Figure A.6 (cotton, oil crops, sugar crops, and flax), and Figure A.7 (tobacco, vegetable, and fruits).

Using this information, we calculate four types of agricultural shocks using the interaction of village-crop area shares in 2001 and crop-year tariffs/trade. The tariff on agricultural exports in village v and year t is calculated as follows:

$$\tau_{vt}^{EX} = \sum_{c} \frac{\text{area}_{vc2001}}{\sum_{c'} \text{area}_{vc'2001}} \tau_{ct}^{EX},$$

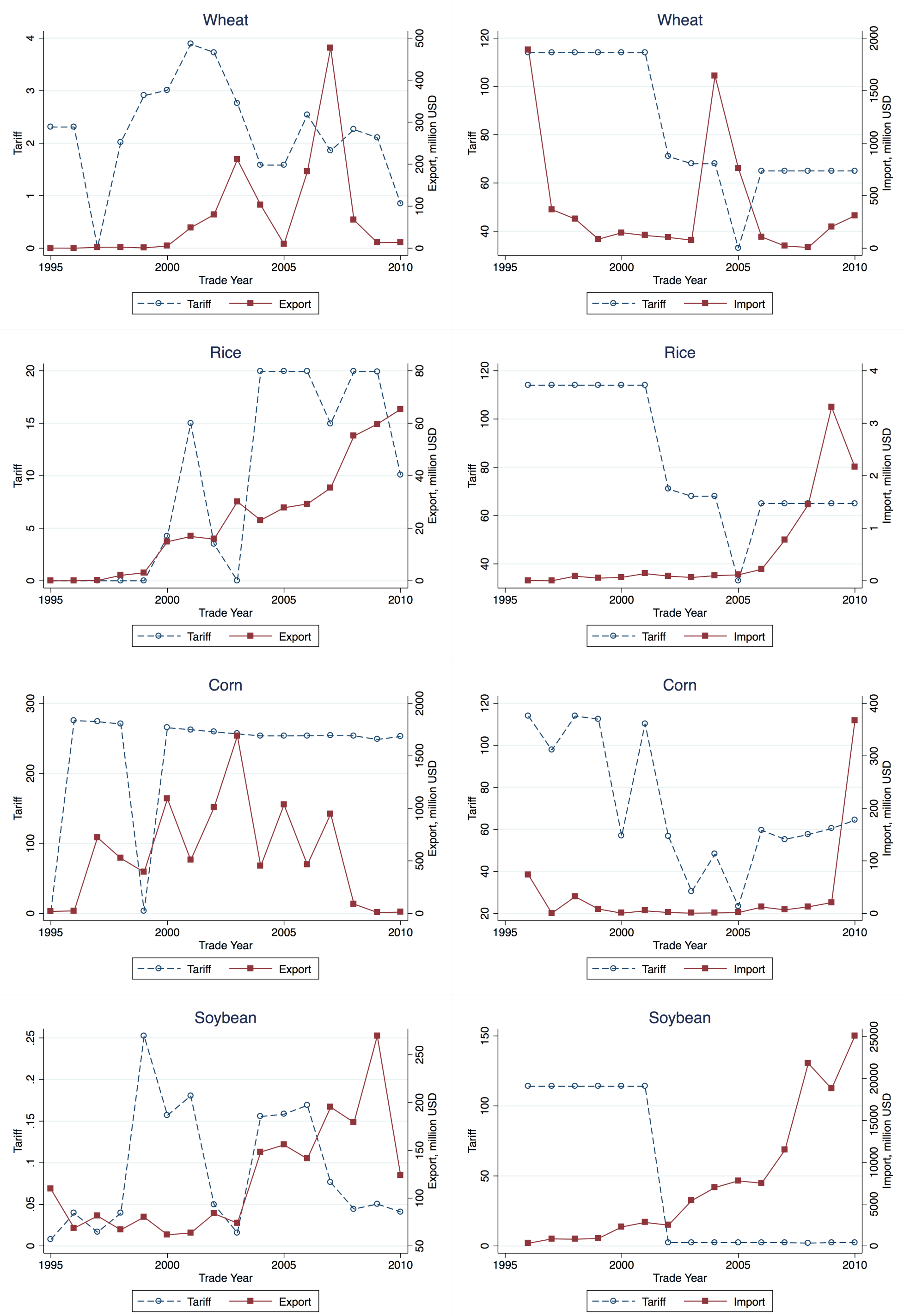
where τ_{ct}^{EX} is the tariff on Chinese exports for crop c in year t, and area_{vc2001} is the area of crop c in village v and year 2001. We include two types of measures, one for cereal crops only, and another for all 11 crops. The tariff on imports is calculated by replacing τ_{ct}^{EX} with τ_{ct}^{IM}.

Alternatively, we use the market-access type of measures of trade shocks as in Autor et al. [2013]. The exposure to agricultural exports in village v and year t is calculated as follows:

$$MA_{vt}^{EX} = \sum_{c} \frac{\text{area}_{vc2001}}{\sum_{c'} \text{area}_{vc'2001}} \frac{v_{ct}^{EX}}{\sum_{v'} \text{area}_{v'c2001}},$$

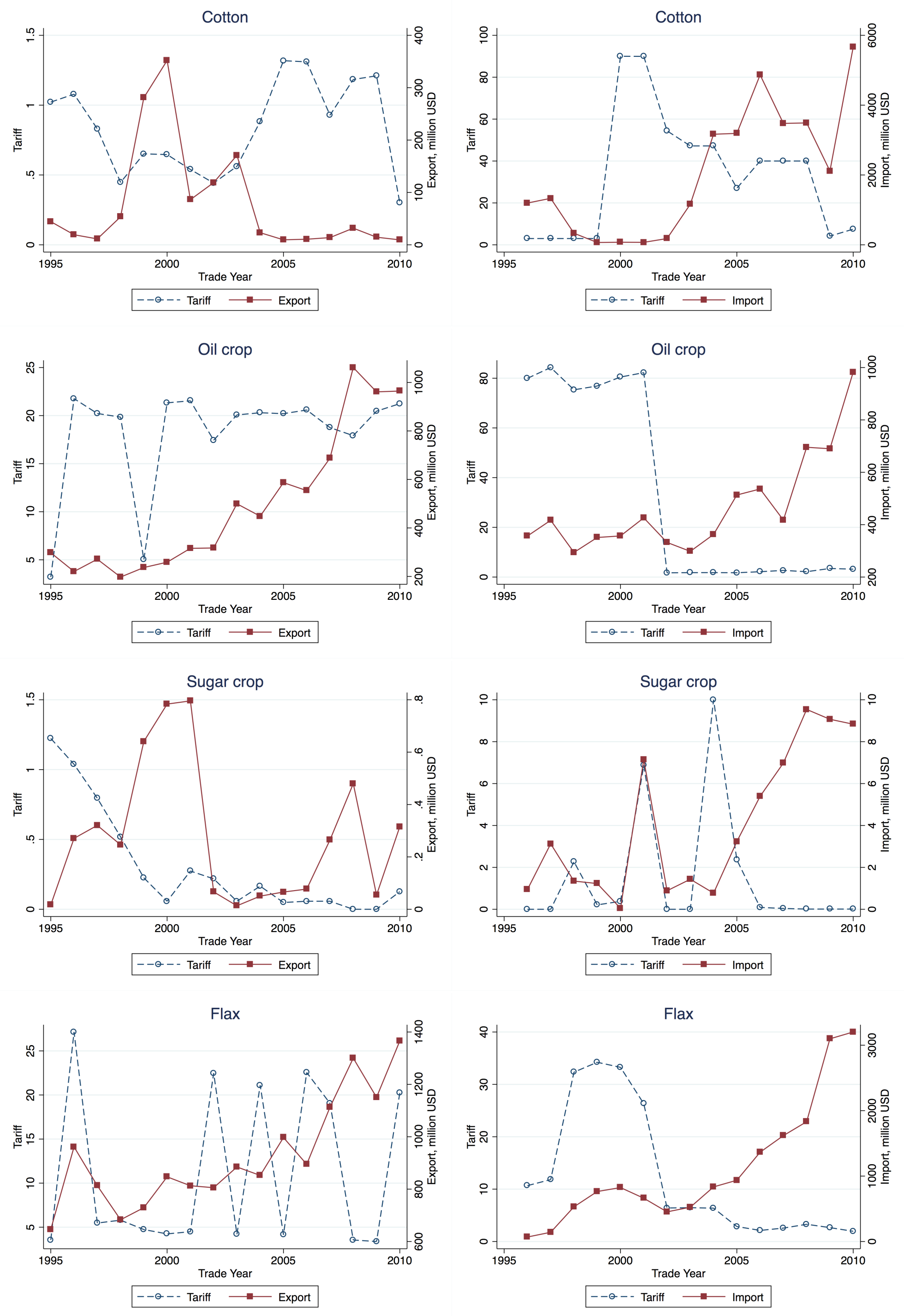
where v_{ct}^{EX} is the Chinese exports for crop c in year t, and area_{vc2001} is the area of crop c in village v and year 2001.

Figure A.5: Agricultural trade, 4 major cereal crops



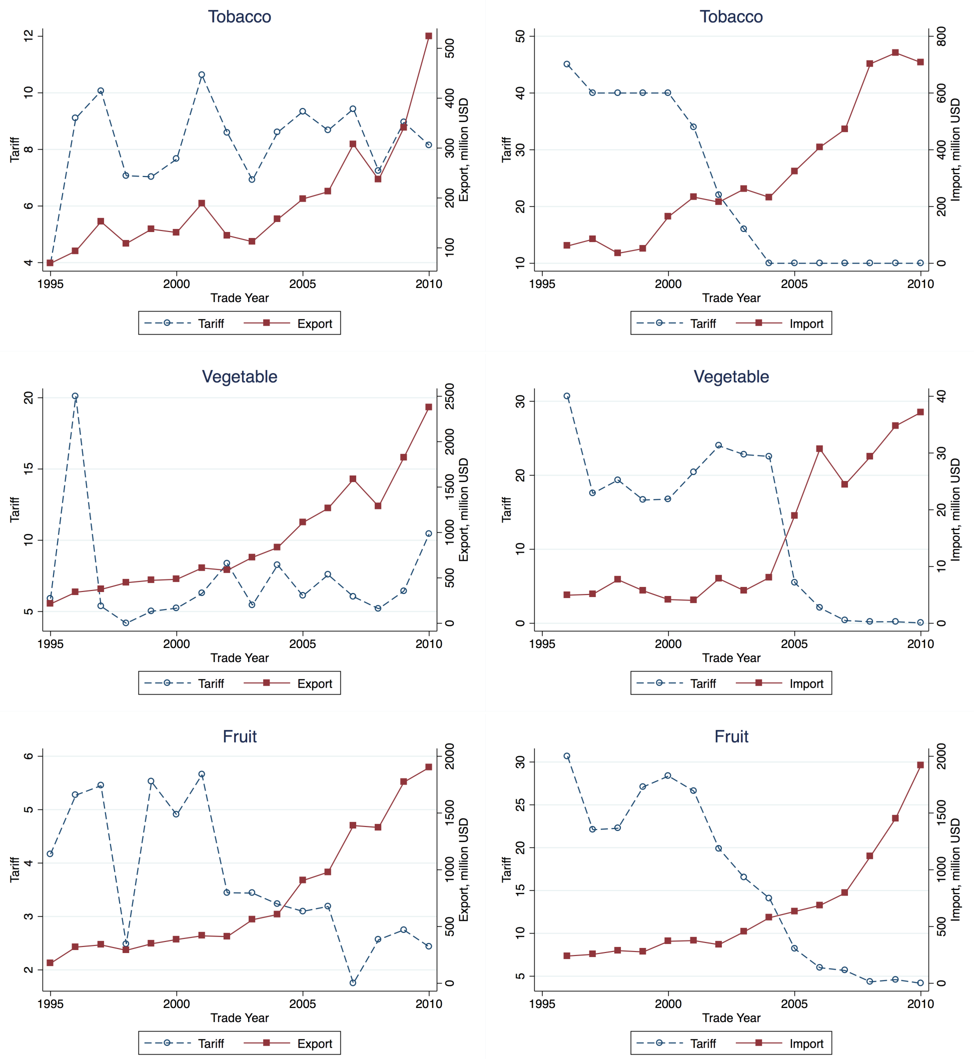
Note: This table presents the trends of Chinese imports, Chinese exports, Chinese tariffs on imports, and tariffs faced by Chinese exports for the four major cereal crops: wheat, rice, corn and soybean. The data source is the 3-digit SIC code trade data from TRAINS. The lines with hollow dots present tariffs, and the lines with solid squares present trade values.

Figure A.6: Cash crops (1): cotton, oil crops, sugar crops, flax



Note: This table presents the trends of Chinese imports, Chinese exports, Chinese tariffs on imports, and tariffs faced by Chinese exports for cotton, oil crops, sugar crops, and flax. The data source is the 3-digit SIC code trade data from TRAINS. The lines with hollow dots present tariffs, and the lines with solid squares present trade values.

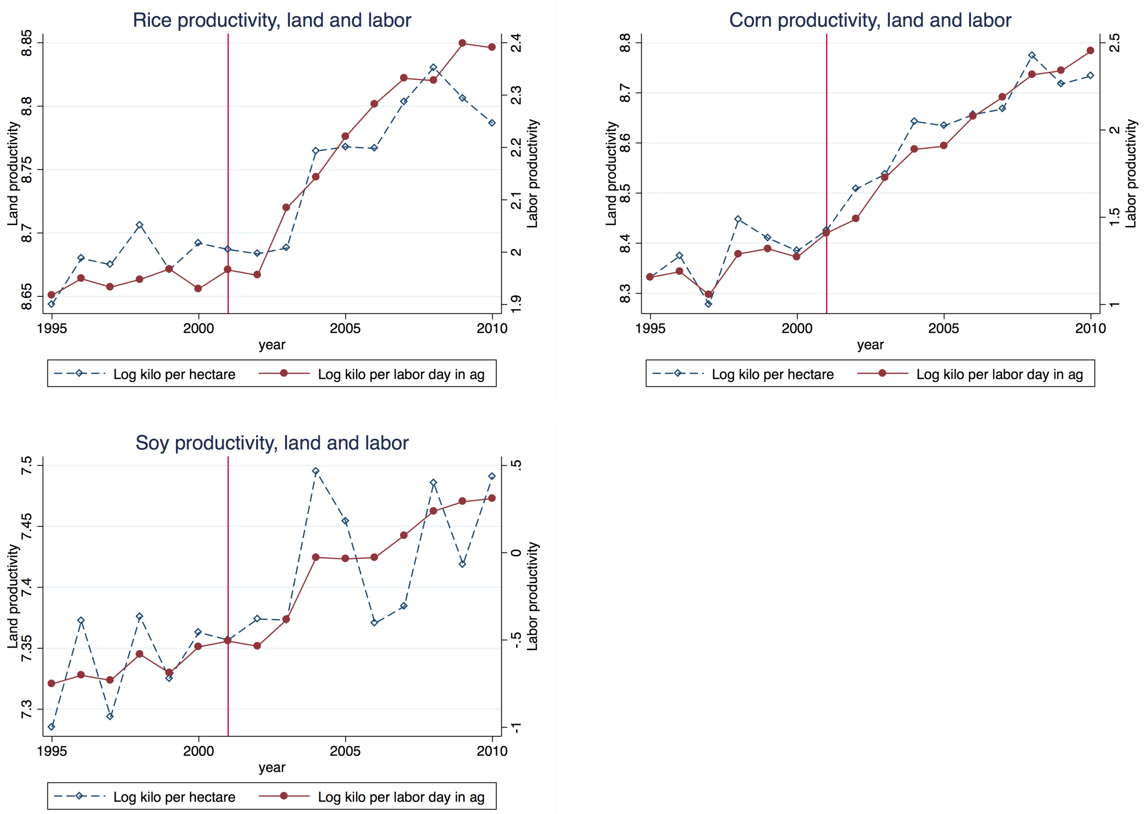
Figure A.7: Cash crops (2): tobacco, vegetable and fruits



Note: This table presents the trends of Chinese imports, Chinese exports, Chinese tariffs on imports, and tariffs faced by Chinese exports for tobacco, vegetables, and fruits. The data source is the 3-digit SIC code trade data from TRAINS. The lines with hollow dots present tariffs, and the lines with solid squares present trade values.

B.6 Productivity Growth for Cereal Crops

Figure A.8: Land and labor productivity for rice, corn, and soybean



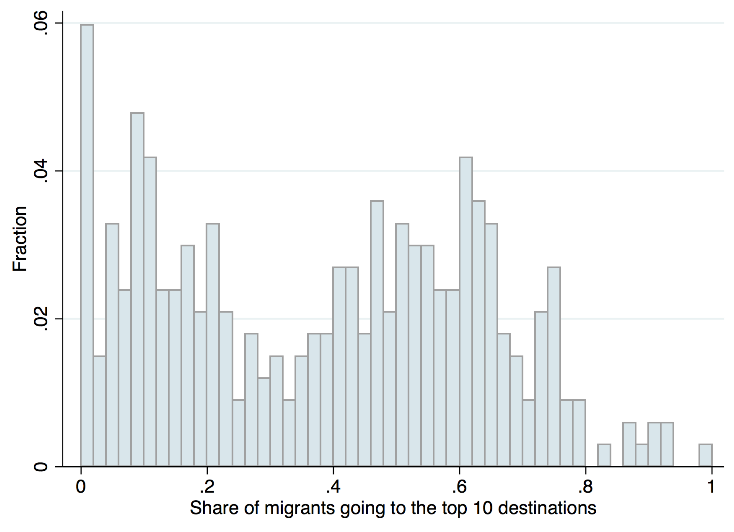
Note: This figure shows the trends of land productivity and labor productivity for rice, corn, and soybean, using the NFP Survey data. Each dot is a household-year average for households growing the crop. Land productivity is defined as the log output in kilo per hectare and is shown in dashed lines with hollow diamonds. Labor productivity is defined as the log output in kilo per labor day in agriculture and is shown in solid lines with solid circles.

C Migration Network Summary

Table A.6: Summary of Statistics of the 2000 Migration Network

Variable	Value
Total number of migrants	51,850
Total number of network links	10,491
Per Destination Prefecture	
Median number of migrants	54
Median number of source prefectures	21
Per Source Prefecture	
Median number of migrants	117
Median number of destination prefectures	28

Figure A.9: Distribution of the Share of Migrants Moving to the Top 10 Destinations



Note: The top 10 destination prefectures are Shenzhen, Dongguan, Guangzhou, Shanghai, Beijing, Foshan, Chongqing, Wenzhou, Wuhan, Quanzhou. These 10 prefectures absorbed 38% of total migrants in China in 2000.

D Additional Empirical Results

D.1 Occupation Choice and Land Rental, Household-Level Evidence, 2001–2010

Household occupation choice was correlated with how much land they decided to work on. Table A.7 Column (1) shows that in a household with 3 laborers, the probability of working on any land was 6 percentage points smaller when one more household member worked as a non-agricultural laborer. Columns (2) and (3) add village-year fixed effects and household fixed effects, and the coefficients are smaller than in Column (1), significant at the 1% level. Column (4) indicates that one more household member working as a non-agricultural laborer is correlated with a 25-percentage-point decline in the land size; while statistically significant, the numbers are again smaller with village-year fixed effects and household fixed effects.

Table A.7: Larger non-agricultural laborer share, less land, 2001–2010

	(1)	(2)	(3)	(4)	(5)	(6)
	Any land (=1)			Log(land size)		
% non-ag laborer	-0.18***	-0.15***	-0.13***	-0.84***	-0.21***	-0.10***
	(0.02)	(0.02)	(0.02)	(0.22)	(0.03)	(0.02)
Log(# labor)	0.08***	0.08***	0.03***	0.35***	0.40***	0.19***
	(0.01)	(0.01)	(0.01)	(0.04)	(0.03)	(0.02)
Constant	0.84***	0.84***	0.88***	1.54***	1.40***	1.59***
	(0.01)	(0.00)	(0.00)	(0.17)	(0.03)	(0.02)
Observations	144,675	144,675	142,513	128,884	128,883	126,528
R-squared	0.04	0.25	0.69	0.06	0.70	0.90
Year FE	Yes			Yes		
Village-Year FE		Yes	Yes		Yes	Yes
HH FE			Yes			Yes

Note: This table shows the results on the correlation between the household non-agricultural labor share and the land in agricultural operation, using the NFP Survey data on households. Column (1) regresses the a dummy for whether the household has any land in agricultural operation on the share of non-agricultural labor and the log number of labor, controlling for year fixed effects. Column (2) replaces the year fixed effects with village-year fixed effects. Column (3) adds household fixed effects. Columns (4)–(6) replicates the results in Columns (1)–(3), using the log land size as the outcome variable. Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

D.2 TFP Estimation

Our main TFP estimation specification is as follows,

$$\log(y_{hvt}) = \alpha \log(d_{hvt}) + \beta \log(k_{hvt}) + \gamma \log(l_{hvt}) + \delta \log(m_{hvt}) + I_h + I_{vt} + \epsilon_{hvt},$$

where y_{hvt} is the output value of crops. An alternative way is to the value-added as the outcome variable, defined as the difference between the output value and the intermediate input value ($y_{hvt} - m_{hvt}$), so

$$\log(V_{hvt}) = \alpha^V \log(d_{hvt}) + \beta^V \log(k_{hvt}) + \gamma^V \log(l_{hvt}) + I_h + I_{vt} + \epsilon^V_{hvt},$$

and again the TFP is measured as the residual

$$\hat{\phi}^V_{hvt} = \log(v_{hvt}) - \hat{\alpha}^V \log(d_{hvt}) - \hat{\beta}^V \log(k_{hvt}) - \hat{\gamma}^V \log(l_{hvt}).$$

The identification assumption the TFP estimation is that the inputs choices are uncorrelated with the idiosyncratic productivity shocks ϵ_{hvt}. However, if the household has information on the shock and make the input choices correspondingly, the estimation is biased. For example, if a household member has an adverse health shock, the household may choose to work on smaller land, supply less labor, and use less capital and intermediate goods. One solution is to use lagged input choices to instrument the current ones (Arellano and Bover 1995), and the identification assumption is that the lagged input choices are uncorrelated with the current period productivity shock.

Table D.2 shows results of various TFP estimation methods. Column (1) uses the output method, and the estimated output elasticity is 0.346 for land, 0.249 for labor, 0.018 for capital, and 0.271 for intermediate inputs. The labor and capital estimates are similar as in Chari et al. [Forthcoming] with by-crop quantity-based estimation, while the land estimate is smaller, closer to Chow [1993] and Cao and Birchenall [2013]. The sum of the coefficients is 0.88, and the F-test rejects constant return to scale. Column (2) uses the value-added method, and the coefficients for labor, land, and capital are larger than in Column (1). Column (3) and (4) use the output method, where Column (3) instruments all inputs with lagged values, and Column (4) instruments only labor and intermediate inputs, assuming that these two inputs are easily adjustable. The coefficients are similar to Column (1) for labor, smaller for capital, and larger for intermediate inputs.

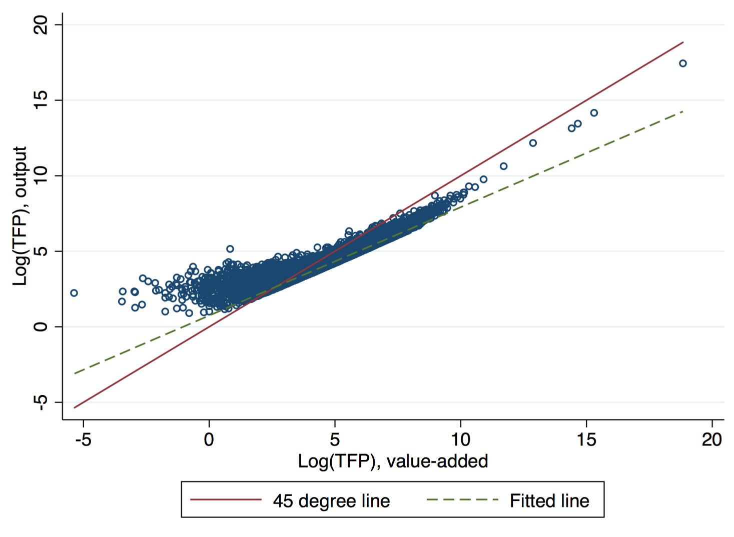
Table A.8: TFP estimation, output method and value-added method, 1995–2010

	(1)	(2)	(3)	(4)
	OLS		IV, log(output value)	
	Log(output value)	Log(value-added)	All inputs Lagged	Labor and intermediate lagged
Log(labor days in agriculture)	0.249***	0.343***	0.242***	0.261***
	(0.013)	(0.017)	(0.010)	(0.011)
Log(capital)	0.018***	0.024***	0.008***	0.010***
	(0.003)	(0.003)	(0.002)	(0.001)
Log(land)	0.346***	0.486***	0.248***	0.161***
	(0.014)	(0.017)	(0.010)	(0.007)
Log(intermediate input costs)	0.271***		0.466***	0.511***
	(0.014)		(0.012)	(0.012)
Observations	245,610	243,281	215,024	217,037
R-squared	0.892	0.846	0.385	0.371
Sum of the coefficients	.88	.85	.96	.94
CRS F-value	83.7	103.7	47.7	194.0
CRS p-value	0	0	0	0

Note: This table presents the results of TFP estimation using different methods. Column (1) uses the output method, where the log output value is regressed on the log land size, the log labor days, the log capital, and the log value of intermediate inputs. Column (2) uses the value-added method, where the log value-added (output value minus the input value) is regressed on the log land size, the log labor days, and the log capital. Column (3) uses the output method, and instruments all inputs with lagged values. Column (4) only instruments the log labor days in agriculture and the log intermediate input with lagged values. All columns control for village-year fixed effects and household fixed effects. Standard errors are clustered at the village and the year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure A.10 shows the relationship between the TFP using the output method and the value-added method, both with OLS estimation. The relationship is quite linear, and the fitted line has a smaller slope compared to the 45 degree line. We use the output method in the main result in Section 5.1 Table 9, and the results are similar in Appendix D.9.

Figure A.10: The relationship between the TFP from the output method and value-added method, OLS

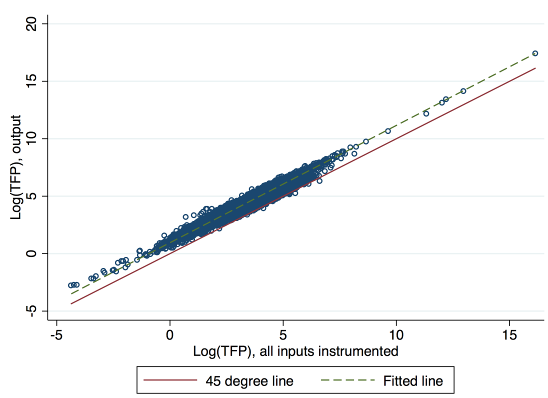


Note: This figure shows the correlation between the TFP calculated using the output method as in Table D.2 Column (1) and the TFP calculated using the value-added method as in Table D.2 Column (2). The solid line represents the 45 degree line, and the dash line presents the linear fitted line.

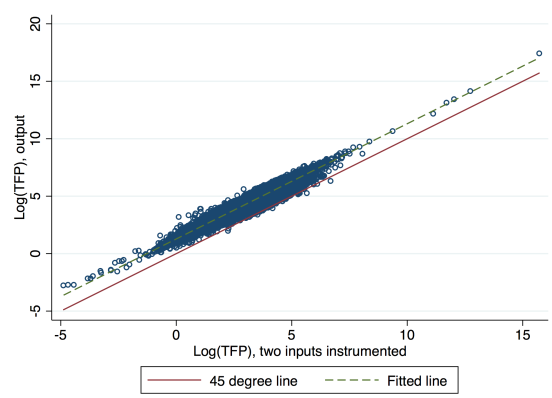
Figure A.11 compares the OLS estimates with the IV estimates using the output method. In both cases (four instruments vs two instruments), the IV estimate is linearly correlated with the OLS one, and slope is very close to one. Thus, we only use the OLS estimate in the analysis.

Figure A.11: The correlation between OLS and IV generated TFP, output method

(a) All inputs instrumented with lags



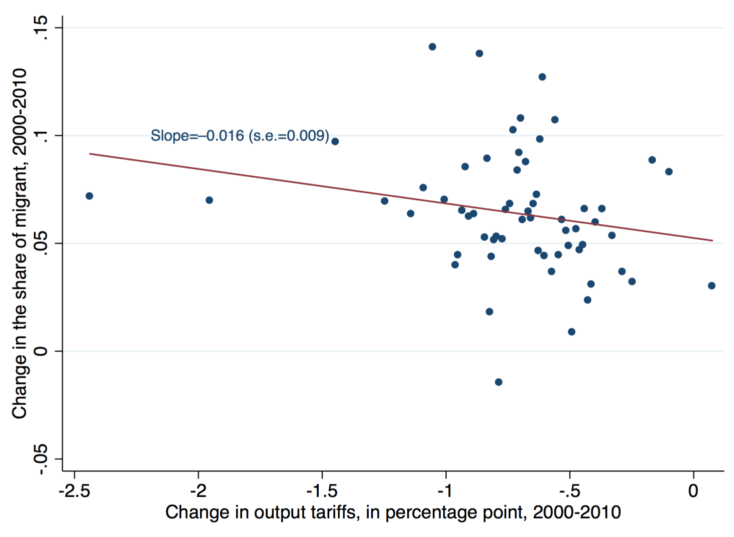
(b) Two inputs instrumented with lags



Note: This figure shows the correlation between the TFP calculated using the OLS output method as in Table D.2 Column (1) and the TFP calculated using IV as in Table D.2 Columns (3) and (4). The solid line represents the 45 degree line, and the dash line presents the linear fitted line.

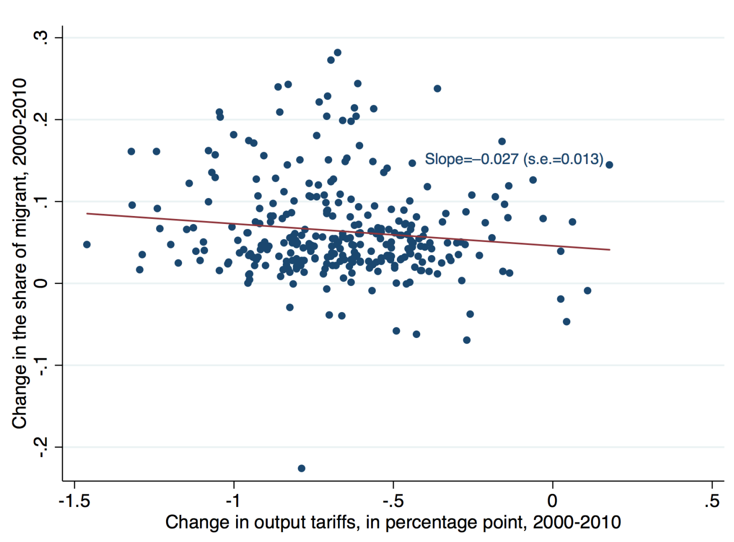
D.3 Graphs of Trade Shocks on Migrant Flows, Robustness

Figure A.12: Prefectures with larger declines in output tariffs in the manufacturing sector experienced larger increases in the share of migrants, 2000–2010, binned scatter plot



Note: This figure shows the relationship between the declines in output tariffs and the changes in the share of migrants from 2000 to 2010, using the output tariffs and the number of migrants using the 2000 and 2010 censuses. Each dot is a prefecture-group, defined by the size of the change in output tariffs, with 60 groups in total. The horizontal axis shows the percentage point change in output tariffs in the manufacturing sector in a prefecture ($\tau_{2010}^{own} - \tau_{2001}^{own}$), and the vertical axis shows the change in the share of migrants.

Figure A.13: Prefectures with larger declines in output tariffs in the manufacturing sector experienced larger increases in the share of migrants, 2000–2010, dropping outliers on the left



Note: This figure shows the relationship between the declines in output tariffs and the changes in the share of migrants from 2000 to 2010, using the output tariffs and the number of migrants using the 2000 and 2010 censuses. Each dot is a prefecture. The horizontal axis shows the percentage point change in output tariffs in the manufacturing sector in a prefecture ($\tau_{2010}^{own} - \tau_{2001}^{own}$), and the vertical axis shows the change in the share of migrants. Here 12 outliers out of 318 on the left-hand side are dropped.

D.4 Tests of Absence of Pre-Trends

As shown in Table A.9, for $o = other$ and $o = own$, we fail to reject the hypothesis of $\Pi = 0$ for 17 out of the 18 tests for $t = 2002, ..., 2010$, when 2 outlier villages are excluded. Results in Section 5 are not affected by dropping the two outlier villages.

Table A.9: P-values for F tests on pre-trends

Year	2010	2009	2008	2007	2006	2005	2004	2003	2002
p-value for $o = own$	0.61	0.33	0.73	0.36	0.47	0.30	0.07	0.64	0.30
p-value for $o = other$	0.27	0.47	0.21	0.24	0.16	0.34	0.47	0.13	0.51

Note: This table presents the p-value for the joint test of coefficients for $\Pi = 0$ in Equation 3. Each cell represents the p-value for the joint test from one regression.

D.5 Other Trade Shocks

First, we provide evidence on the absence of agricultural trade effect in this section. Table A.10 adds measure of agricultural trade shocks as controls, and the specification is the same as in Section 5.1 Table 2 Column (1). Column (1) is a direct replication with all villages with no missing agricultural trade shock measures. Columns (2) and (3) add the agricultural tariff shocks, for four cereal crops and for all eleven crops, respectively. Column (4) and (5) use market-access-based trade exposure measure. In all last four columns, the coefficients on the agricultural trade shocks are insignificant, and the coefficient on manufacturing tariffs (other prefectures' and own prefecture's) are the same as in Column (1).

Another potential concern is that whether the tariff reduction was the only trade shock induced by China's WTO accession. For example, Pierce and Schott [2016], Handley and Limão [2017], and Erten and Leight [Forthcoming] emphasize the important of the reduction in tariff uncertainty between the United States and China. They argue that United States applied MFN tariffs on Chinese exports even before the WTO accession. However, before 2001, there was great uncertainty regarding the U.S. trade policy: the MFN status had to be approved each year by the Senate and the House; otherwise, the Column 2 tariff would be applied to Chinese exports.

To address this concern, we construct the U.S. uncertainty related tariff. We use the 2000 customs data by firm, eight-digit Harmonized System (HS) category, and destination country, then combining 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].[50] With these data, the potential U.S. tariff on prefecture i in 2001 is:

$$\tau^{US}{}_{i2001} = \sum_p \frac{export^{US}_{p,i,2000}}{\sum_{p'} export^{US}_{p,i,2000}} Column2^{US}_{p,2000},$$

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

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

p in 2000. The U.S. tariff on prefecture i in year $t = 2002, ..., 2010$, is

$$\tau^{US}{}_{it} = \sum_{p} \frac{export^{US}_{p,i,2000}}{\sum_{p'} export^{US}_{p,i,2000}} MFN^{US}_{p,t},$$

where $MFN^{US}_{p,2000}$ is the U.S. MFN tariff.

Then we have a village's exposure to its own prefectures' output tariff as

$$\tau^{own,US}_{vit} = \tau^{US}_{it},$$

and its exposure to tariffs in other prefectures as

$$\tau^{other,US}_{vit} = \sum_{j \neq i} \frac{m_{ij}}{\sum_{j' \neq i} m_{ij'}} \tau^{US}_{jt},$$

where m_{ij} is the number of migrants who are from prefecture i and reside in prefecture j in 2000.

Table A.10: Occupation choice results, controlling for the agricultural trade shocks and uncertainty shock

	(1)	(2)	(3)	(4)	(5)	(6)
Y: % non-ag laborer		Agricultural tariff		Agricultural Market Access		Uncertainty
		Cereal	All	Cereal	All	
Other pref. tariff	-0.09***	-0.09**	-0.09***	-0.09***	-0.09***	-0.06**
	(0.02)	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)
Own pref. tariff	0.02**	0.02**	0.02*	0.02**	0.02**	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Agricultural export shock		-0.07	-0.11	0.06	-0.02	
		(0.16)	(0.19)	(0.10)	(0.03)	
Agricultural import shock		-0.09	-0.02	0.00	-0.00	
		(0.05)	(0.03)	(0.01)	(0.00)	
Uncertainty, own pref.						-0.11***
						(0.03)
Uncertainty, other pref.						0.15
						(0.13)
Observations	1,964	1,964	1,964	1,964	1,964	2,333
R-squared	0.85	0.85	0.85	0.85	0.85	0.85

Note: This table shows the robustness of result in Table 2 Column (3). Column (1) replicates Table 2 Column (3), restricting the sample to the villages that had non-missing crop area information in 2001. Columns (2) and (3) add agricultural export and import shocks calculated from agricultural tariff reductions, including only cereal crops and all crops, respectively. Columns (4) and (5) use agricultural export and import shocks using the market access approach, including only cereal crops and all crops, respectively. Column (6) replicates Table 2 Column (3), adding the uncertainty shocks. All columns control for the log labor, the log number of households, the log government transfer +1, province-year fixed effects, and village fixed effects. The mean (sd) of the share of non-agricultural labor is 0.19 (0.15). Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

In Table A.10 Column (6), we control for the U.S. uncertainty-related tariffs, both in a village's own prefecture, and in other prefectures. We find that the coefficients for the actual tariffs in the first two rows are exactly the same as in Table 2 Column (1), meaning that our actual tariff results are robust. We don't find a significant effect of other prefectures' U.S. uncertainty-related tariff. One-standard deviation larger decline in own prefecture's U.S. uncertainty-related tariff resulted in a one-percentage point larger increase in the share of non-agricultural labor. This result is consistent with the finding in Erten and Leight [Forthcoming], but contrary to our actual own prefecture effect.

D.6 Village-Level Results for Occupation Choice and Out-Migration

Section 5.1 Table 2 uses the share of non-agricultural labor in a village as the outcome variable to investigate the impact of trade exposure to rural residents' occupation choice. Table A.11 uses data from the village questionnaire, with the specification is the same as in Table 2 Column (1). Column (1) shows that a one-standard-deviation larger decline in other prefectures' output tariff led to a 17% larger decline in the share of households whose sole business were agriculture. Column (2)–(4) focus on the share of labor working outside the village. Column (2) shows that a one-standard-deviation larger decline in other prefectures' output tariff led to a 10% larger increase in the share of labor working outside the village. The effect is mainly driven by within-province migration (Column 3) instead of between-province migration (Column 4). The decline in tariff also led to a decline in the share of excess labor, but the effect is not statistically significant (Column 5).[51]

Table A.11: Occupation choice and out-migration, village questionnaire result, 2001–2010

	(1)	(2)	(3)	(4)	(5)
	% in agriculture		% out of village		% excess labor
		Any	% within province	% between province	
Other pref. tariff	0.17***	-0.10*	-0.12**	0.02	0.03
	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)
Own pref. tariff	0.01	-0.01**	-0.02*	0.00	0.01
	(0.02)	(0.00)	(0.01)	(0.01)	(0.01)
Observations	2,255	2,257	2,256	2,257	2,253
R-squared	0.86	0.80	0.78	0.88	0.70
Mean (sd) Y	0.50 (0.31)	0.27 (0.18)	0.17 (0.14)	0.10 (0.13)	0.08 (0.11)

Note: This table shows the labor market outcomes in response to the trade shocks, using NFP Survey data from village questionnaires. All columns control for the log labor, the log number of households, the log government transfer +1, province-year fixed effects, and village fixed effects. All columns have the same specification as Table 2 Column (3). Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

D.7 Land Reallocation Results with Current TFP

Table A.12 shows the results of household-year-level regressions of the log TFP in year t, $t = 2002, ..., 2010$, regressed on the log TFP in 2001. Column (1) controls for year fixed effects, and we

[51] Excess labor is defined in labor units. It is calculated as $\frac{\text{total labor}\times 300-\text{total labor days}}{300}$.

find that a 10% increase in the TFP in 2001 was correlated with a 6% increase in the TFP in later years. Columns (2)–(6) have similar results when we control for the quadratic form of the initial TFP and different fixed effects.

Table A.12: TFP persistence, 2001–2010

Y: Log(TFP), year t	(1)	(2)	(3)	(4)	(5)	(6)
Log(TFP), 2001	0.586***	0.149	0.308***	0.093	0.308***	0.088
	(0.029)	(0.168)	(0.016)	(0.086)	(0.016)	(0.086)
Log(TFP), 2001, squared		0.052***		0.025***		0.025***
		(0.017)		(0.009)		(0.009)
Constant	2.010***	2.904***	3.298***	3.751***	3.300***	3.764***
	(0.136)	(0.411)	(0.075)	(0.211)	(0.073)	(0.210)
Observations	95,937	95,937	95,937	95,937	95,934	95,934
R-squared	0.339	0.346	0.543	0.544	0.653	0.654
Year FE	Yes	Yes	Yes	Yes		
Village-Year FE					Yes	Yes
Village FE			Yes	Yes		

Note: This table shows the correlation between current TFP (in year t) and initial TFP (in year 2001). Column (1) regresses the TFP of household h in year t on the TFP of household i in year 2001, controlling for year fixed effects. Column (2) adds the quadratic term of the initial TFP. Columns (3) and (4) replicates Columns (1) and (2), adding village fixed effects. Columns (5) and (6) use village-year fixed effects instead. Standard errors are clustered at the village-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.13 replicates the results of Table 6, by replacing the TFP in 2001 with the concurrent TFP. Given the high serial correlation of TFP over time, the results are very similar to the ones in Table 6.

Table A.13: Land allocated more towards initially productive households, 2001–2010

Y: log(land)	(1)	(2)	(3)	(4)
Log(TFP)	0.267***	0.135***	0.215***	0.091***
	(0.055)	(0.034)	(0.045)	(0.025)
Other pref. tariff	0.184	0.144		
	(0.142)	(0.138)		
Other pref. tariff × log(TFP)	-0.062***	-0.049***		
	(0.016)	(0.010)		
Other pref. tariff*%cross-pref. migr			0.135	0.057
			(0.186)	(0.158)
Other pref. tariff*%cross-pref. migr			-0.100**	-0.076***
× Log(TFP)			(0.032)	(0.019)
own prefecture's output tariff	-0.016	-0.017	-0.027	-0.026
	(0.035)	(0.036)	(0.035)	(0.036)
Observations	119,760	117,671	119,760	117,671
R-squared	0.648	0.885	0.648	0.885
Village-Year FE	Yes	Yes	Yes	Yes
Village FE	Yes		Yes	
HH FE		Yes		Yes

Note: This table replicates Table 6, replacing TFP in 2001 by TFP in the contemporaneous year t. Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

D.8 Elasticity of Investment to Income from Different Sources

In this section, we show how investment in fixed capital is correlated with income from different sources, both on the extensive margin and on the intensive margin. In Table A.14, the outcome variable in Panel A is the log of household expenditure on purchase on productive fixed assets plus one, and the regressors are the share of income from different sources. In Panel B, the outcome variable is the probability of having positive investment. All columns control for the log total income, household fixed effects and year fixed effects.

When the share of income from wages increases by 10 percentage points, the investment decreases by 4%, which is 7% of the mean of growth rate of the investment. The probability of investment decreases by 0.5 percentage point, which is 6% of the mean of investment probability. On the other hand, income from farm operation and from government agricultural subsidies increase both the size of the investment and the probability of investment.

Table A.14: Expenditure on capital investment negatively correlated with wage income, 1995–2010

	(1)	(2)	(3)	(4)	(5)
	Wage	Farm	Land rent	Interest	Government
Panel A: Y = Log(investment+1)					
Share of income from...	-0.41***	0.29***	0.17	0.20	0.75***
	(0.04)	(0.04)	(0.10)	(0.29)	(0.22)
Panel B: Y= I(investment>0)*100					
Share of income from...	-4.80***	3.44***	1.85	1.72	9.12***
	(0.64)	(0.63)	(1.50)	(3.12)	(2.96)
Mean (sd) share of income from...	0.22 (0.28)	0.62 (0.33)	0.01 (0.05)	0.005(0.003)	0.02 (0.05)

Note: This table shows the correlation between investment in fixed capital and the share of income from different sources. The outcome variable in Panel A is the log of investment +1, and the outcome variable in Panel B is the probability of having positive investment. All columns control for the log total income, household fixed effects and year fixed effects. The mean (sd) of the log expenditure on capital +1 is 0.57 (1.99), and the mean (sd) of the probability of investing in expenditure is 0.08 (0.27). Standard errors are clustered at the village and the year level. *** p<0.01, ** p<0.05, * p<0.1.

D.9 TFP Results with Value-Added Method

The TFP results are similar using the value-added method, compared to the results in Table 9 with the output method.

Table A.15: Villages with larger trade exposures increased allocation efficiency

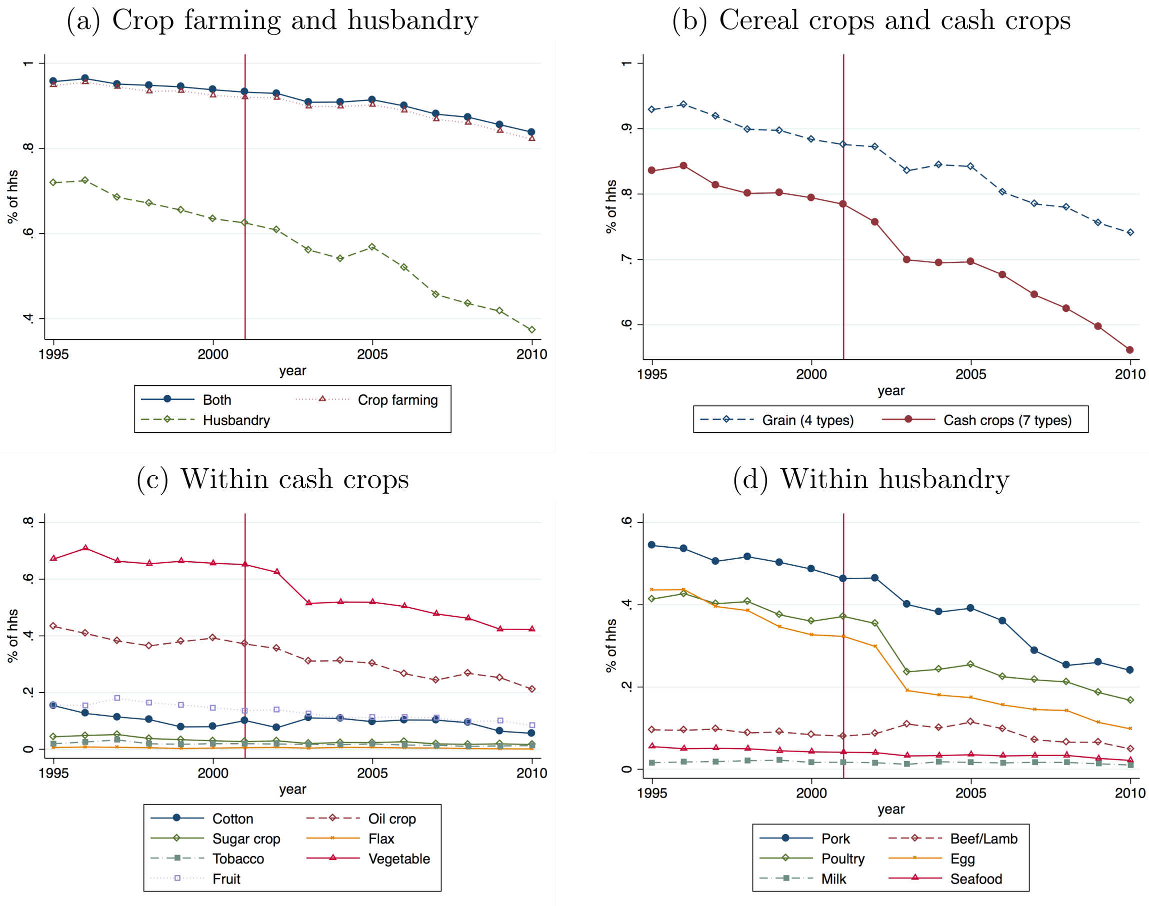
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Log(village TFP)				Allocation efficiency	
	output weighted		unweighted			
Other pref. tariff	-0.68***	-0.39	0.03	0.11	-0.70***	-0.49
	(0.16)	(0.21)	(0.34)	(0.35)	(0.14)	(0.27)
Other pref. tariff × % cross-pref. migr		-0.81		-0.23		-0.59
		(0.71)		(0.13)		(0.66)
Own pref. tariff	0.04	0.02	0.02	0.01	0.02	0.00
	(0.12)	(0.11)	(0.04)	(0.04)	(0.09)	(0.08)
Observations	2,332	2,332	2,332	2,332	2,332	2,332
R-squared	0.67	0.67	0.85	0.85	0.59	0.59

Note: This table shows the equivalence of results with Table 9 when we use the TFP from the valued-added method instead of the TFP from the output method. All columns control for the log labor, the log number of households, the log government transfer +1, province-year fixed effects and village fixed effects. Columns (1)(3) and (5) have the same specification as Table 2 Column (3), and Columns (2)(4) and (6) have the same specification as Table 2 Column (4). The mean (sd) of the log weighted TFP is 5.32 (1.13), the mean (sd) of the log unweighted TFP is 5.61 (0.68), and the mean (sd) of allocation efficiency is -0.29 (0.97). Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

D.10 Crop Mix and Husbandry

In this section, we provide some facts on crop mix and husbandry, as well as the impact of trade on these outcomes. Figure A.14 shows the trend of the share of households in each type of production. Overall, the share declined for all types, especially after 2001.

Figure A.14: Crop mix and husbandry



Note: This figure shows the trends of husbandry and crop farming, using the NFP Survey data on households. Panel (a) depicts the share of households in either crop farming or husbandry (the solid line with solid circles), the share of households in crop farming only (the dotted line with hollow triangles), and the share of households in husbandry only (the dashed line with hollow diamonds). Panel (b) depicts the share of households growing grains (the dashed line with hollow diamonds), and the share of households with cash crops (the solid line with solid circles). Panel (c) shows the share of households with different types of cash crops. Panel (d) shows the share of households with different types of husbandry. Oil crops include peanut, sesame, rapeseed, sunflower, benne, and castor bean. Sugar crops include sugar cane and beets.

Table A.16 tabulates the share of households with each crop. Wheat and rice are the most common cereal crops, while oil crops and vegetables are the most common cash crops.

Table A.16: Cross tabulation, % of HHs with certain crops, 1995–2010

1995–2010	Wheat	Rice	Corn	Bean
Wheat	.39			
Rice	.14	.45		
Corn	.28	.18	.51	
Soy bean	.14	.14	.19	.28

1995–2010	Cotton	Oil	Sugar	Flax	Tobacco	Vegetable	Fruit
Cotton	.10						
Oil crop	.06	.33					
Sugar crop	-	.01	.03				
Flax	-	-	-	-			
Tobacco	-	-	-	-	.02		
Vegetable	.07	.24	.02	-	.01	.57	
Fruit	.02	.04	-	-	-	.08	.13

Note: This table tabulates the share of households growing one or two types of cereal crops or one or two types of cash crops, using the NFP Survey data on households. These shares pool 1995 to 2010 data together.

Table A.17 presents the effect of trade on cash crop production. A one-percentage point larger decline in other prefectures' output tariff led to a 11% larger decline in the share of households with cash crops among all crop farming households (Columns 3 and 4); similar effects can be found the in the number of households in cash crop farming (Column 5 and 6). We find no significant effect on the revenue share of cash crops.

Table A.17: Cash crop effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Cash crop revenue /crop revenue		# HHs with cash crops /HH with crops		Log(# of HHs) with cash crops	
Other pref. tariff	-0.04	-0.05	0.11***	0.11**	0.20*	0.13
	(0.08)	(0.07)	(0.03)	(0.03)	(0.11)	(0.17)
Other pref. tariff ×		0.02		0.02		0.20
% cross-pref. migr		(0.08)		(0.09)		(0.42)
Own pref. tariff	-0.02	-0.01	0.01	0.01	-0.00	0.01
	(0.01)	(0.01)	(0.02)	(0.02)	(0.08)	(0.09)
Observations	2,333	2,333	2,333	2,333	2,221	2,221
R-squared	0.92	0.92	0.89	0.89	0.87	0.87
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table shows the effect of other prefectures' output tariff on cash crop farming. All columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects, and village fixed effects. Columns (1)(3) and (5) have the same specification as Table 2 Column (3), and Columns (2)(4) and (6) have the same specification as Table 2 Column (4). The mean (sd) of the revenue share of cash crops is 0.42 (0.30), the mean (sd) of the household share of cash crops is 0.76 (0.33), and the mean (sd) of the log number of households with cash crops is 3.53 (0.95). Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

Similar effect are found in Table A.18 with vegetable production. Households left vegetable production more in villages with larger increase in trade exposures.

Table A.18: Vegetable effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Vegetable revenue /crop revenue		# HHs with vegetable /HH with crops		Log(# of HHs) with vege production	
Other pref. tariff	0.01	-0.02	0.16*	0.14	1.12***	0.87***
	(0.06)	(0.07)	(0.07)	(0.09)	(0.13)	(0.20)
Other pref. tariff × % cross-pref. migr		0.06		0.05		0.72
		(0.06)		(0.13)		(0.57)
Own pref. tariff	-0.00	-0.00	0.01	0.01	-0.02	0.00
	(0.01)	(0.01)	(0.02)	(0.02)	(0.08)	(0.09)
Observations	2,333	2,333	2,333	2,333	1,993	1,993
R-squared	0.92	0.92	0.90	0.90	0.87	0.87

Note: This table shows the effect of other prefectures' output tariff on vegetable farming. All columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects, and village fixed effects. Columns (1)(3) and (5) have the same specification as Table 2 Column (3), and Columns (2)(4) and (6) have the same specification as Table 2 Column (4). The mean(sd) of the revenue share of vegetables is 0.42 (0.30), the mean (sd) of the household share of vegetables is 0.76 (0.33), and the mean (sd) of the log number of households with vegetables is 3.53 (0.95). Standard errors are clustered at the province and the year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

We find no significant effect on husbandry, as shown in Table A.19.

Table A.19: Husbandry effects

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Log (husb. output)		Log (# of HHs in husb.)		Log(labordays in husb.)	
Other pref. tariff	0.21	0.20	0.25	0.42	-0.00	0.13
	(0.26)	(0.23)	(0.32)	(0.35)	(0.48)	(0.58)
Other pref. tariff ×		0.05		-0.46		-0.37
% cross-pref. migr		(0.47)		(0.31)		(0.51)
Own pref. tariff	0.12	0.12	0.07	0.06	0.10	0.09
	(0.07)	(0.08)	(0.04)	(0.04)	(0.09)	(0.08)
Observations	2,263	2,263	2,263	2,263	2,241	2,241
R-squared	0.79	0.79	0.87	0.87	0.82	0.82

Note: This table shows the effect of other prefectures' output tariff on husbandry. All columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects and village fixed effects. Columns (1)(3) and (5) have the same specification as Table 2 Column (3), and Columns (2)(4) and (6) have the same specification as Table 2 Column (4). The mean (sd) of the log value of husbandry is 11.77 (1.33), the mean (sd) of the log number of households with husbandry is 3.16 (1.04), and the mean (sd) of the log labor days in husbandry is 7.67 (1.21). Standard errors are clustered at the province and the year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

D.11 Controlling for the Share of Migrants Going to the Top 10 Destinations

Table A.20: Differential initial share of migrants going to top 10 destinations didn't affect the impact of trade shocks

VARIABLES	(1) % non-ag laborer	(2) Log(land leased+1)	(3) Log(agr machine)	(4) 	(5) Village TFP
Other pref. tariff	-0.06**	-0.65***	-0.12	0.27	-0.79***
	(0.02)	(0.15)	(0.79)	(0.82)	(0.14)
Other pref. tariff × % cross-pref. migr				-1.04*	
				(0.52)	
Own pref. tariff	0.02*	-0.09*	0.02	-0.01	0.00
	(0.01)	(0.05)	(0.07)	(0.08)	(0.08)
Share to top ten × year trend	0.01	0.12	0.12	0.08	0.22*
	(0.01)	(0.11)	(0.07)	(0.08)	(0.11)
Constant	0.49***	-0.52	2.05	2.56	4.88***
	(0.09)	(1.95)	(2.69)	(2.69)	(0.65)
Observations	2,333	2,333	2,333	2,333	2,333
R-squared	0.85	0.83	0.87	0.87	0.65
Province-Year FE	Yes	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes	Yes

Note: This table shows the effect of other prefectures' output tariff on the major outcomes of interest, when we control for linear trends interacted with the share of migrants going to the top 10 destination prefectures in 2000. All columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects and village fixed effects. The average share of people moving to the top 10 destination prefectures is 29%. Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

D.12 Controlling for the Crop Patterns

Table A.21: Results on occupation choices robust to controlling for concurrent crop patterns

	(1)	(2)	(3)	(4)
Y=% non-ag laborer	Wheat	Rice	Corn	Soybean
Other pref. tariff	-0.06**	-0.06**	-0.06**	-0.06**
	(0.02)	(0.02)	(0.02)	(0.02)
Own pref. tariff	0.02*	0.02	0.02	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)
Share of output value coming from crop X	0.13*	0.03	0.01	0.02
	(0.06)	(0.05)	(0.04)	(0.06)
Observations	2,333	2,333	2,333	2,333
R-squared	0.85	0.85	0.85	0.85
Province-Year FE	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: This table shows the effect of other prefectures' output tariff on the occupation choice pattern, when we control for the share of output value coming from one of the four major crops. All columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects and village fixed effects. The average value share of wheat is 11%, rice 23%, corn 19%, and soy bean 4%. Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.22: Results on occupation choices robust to controlling for initial crop patterns

	(1)	(2)	(3)	(4)
Y=% non-ag laborer	Wheat	Rice	Corn	Soybean
Other pref. tariff	-0.09***	-0.10***	-0.09***	-0.09***
	(0.02)	(0.02)	(0.02)	(0.02)
Own pref. tariff	0.02**	0.02**	0.02***	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)
2001 crop share × Year FE	Yes	Yes	Yes	Yes
Observations	1,971	1,971	1,971	1,971
R-squared	0.85	0.85	0.85	0.85
Province-Year FE	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: This table shows the effect of other prefectures' output tariff on the occupation choice pattern, when we control for the share of output value coming from one of the four major crops in the initial year (interacted with year fixed effects). All columns control for own prefecture's output tariff, the log labor, the log number of households, the log government transfer +1, province-year fixed effects and village fixed effects. The average initial value share of wheat is 13%, rice 24%, corn 16%, and soybean 5%. Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

D.13 Village-Level OLS Evidence

We first provide evidence on the correlation between non-agricultural laborer share and various measure on the land market. Our estimation equation is as follow:

$$y_{vt} = \alpha_0 + \alpha_1 \left(\frac{\text{non-ag laborer}}{\text{labor}} \right)_{vt} + \Gamma X_{vt} + I_{pt} + I_v + \epsilon_{vt}, \tag{5}$$

where y_{vt} can be the log land area that is leased, the log total land in agricultural operation, the log land area per agricultural laborer, and the log number of households with more than 1/3 hectare land in village v and year t. The main explanatory variable is the share of non-agricultural laborer. Our parameter of interest is α_1, which measures how non-agricultural labor shares affect land allocation.

Table A.23 shows that regions with larger non-agriculture labor shares had more land leasing activities. A 10% increase in the non-agricultural labor share is correlated with a 7.6% increase in the area of land leased, a 4.9% decrease in total land areas in agricultural operation, and a 7% increase in the land area per agricultural worker. This indicates that about 50% of all households lease their land out when moving out of agricultural production. The correlation between the share of non-agricultural labor and the log number of households with more than 1/3 hectare land is insignificant. Overall, the evidence is consistent with the individual evidence in Appendix D.1, where households with more labor in non-agriculture choose to work on smaller land.

Table A.23: Larger non-agriculture labor share, more land leasing, 2001–2010

	(1)	(2)	(3)	(4)
	Log(land leased+1)	Log(land)	Log(land p.c.)	Log(# of hhs>1/3 ha)
% non-ag laborer	0.76**	-0.49*	0.68***	-0.03
	(0.31)	(0.25)	(0.15)	(0.15)
Log(# laborer)	0.01	0.20***	-0.86***	0.00
	(0.25)	(0.06)	(0.07)	(0.04)
Log(land)	1.35***			0.97***
	(0.20)			(0.10)
Observations	2,333	2,333	2,333	2,333
R-squared	0.86	0.96	0.93	0.97

Note: This table shows how non-agricultural labor shares are correlated with land distributions. All columns control for the log number of households, the log government transfer +1, province-year fixed effects and village fixed effects. The mean of % non-agr laborer is 0.19. Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

We proceed to investigate how capital and TFP are correlated with non-agricultural labor shares in Table A.24. A 10-percentage-point increase in the share of non-agricultural labor is correlated with a 7.5% increase in the value of agricultural machinery and a 3.9% increase in the village TFP.

Table A.24: More non-agricultural workers, more machinery, higher TFP

	(1)	(2)
	Log(agricultural machinery)	Village TFP
% non-ag laborer	0.75*	0.39
	(0.40)	(0.37)
Log(# laborer)	0.55	0.08
	(0.35)	(0.17)
Observations	2,333	2,333
R-squared	0.87	0.65

Note: This table shows the correlation between the non-agricultural labor share, agricultural machinery, and village level TFP. All columns control for the log land, the log number of households, the log government transfer +1, province-year fixed effects and village fixed effects. The mean of % non-agricultural laborer is 0.19. Standard errors are clustered at the province and the year level. *** p<0.01, ** p<0.05, * p<0.1.

These OLS estimates are not sufficient to pin down the causal impact of increased out-migration on agricultural production. The first issue is reverse causality. Think about a rural worker's occupation choice. Villages with more land leased can be the ones with larger correlation between the household TFP and land, and the access to the land market allow farmers with a comparative advantage in non-agriculture to move out of the agriculture. Given TFP, villages with more capital have higher wages in agriculture, which disincentives out-migration. If there is a labor-saving change in agricultural technology, out-migration will increase; if the technological change is land-augmenting, out-migration will decrease.[52]

The second concern is on the omitted variables that affect both the migration decision and agricultural production. For example, when a village becomes connected through the expansion of transportation networks, both trade costs and migration costs are reduced. If the quality of fertilizer improves due to increased trade, TFP will increase. At the same time, the cost of migrating out is smaller and more villagers move out of agriculture. In this case, the co-movement of TFP and out-migration are not the cause of each other but the consequence of transportation network expansion. Another concern is common shocks to productivity across sectors. Suppose that there is a positive shock in productivity in both the manufacturing sector and the agriculture sector. The shock in the manufacturing sector is larger, inducing labor reallocation; however, the labor outflow is not the cause of the productivity growth in the agricultural sector.

To overcome these identification challenges, we use the manufacturing trade shock as an exogenous change to the pull factors of out-migration, and show the causal impact of the increased non-agricultural labor share on agricultural production in the following section.

[52] As shown in Bustos et al. [2016], an example of labor-saving technology is the introduction of genetically modified seeds. An example of land-augmenting technology is the adoption of two season farming.

E Model Appendix

E.1 Definition of equilibrium

Given international prices for goods and capital, $\{p_a, r\}$, a competitive equilibrium is a land price q, an allocation for each farm $\{l_i, k_i, y_{ai}\}$, and allocation for the non-agricultural firm $\{Y_n, N_n\}$, an occupation choice $\{o(s_{ai}, s_{ni})\}$, a consumption allocation $\{(c_{ai}, c_{ni})\}$ for each individual such that: (1) individuals maximizes their utility; (2) firms maximize their profit; (3) farmers maximize their profits; (4) individuals maximizes their income by choosing their occupation; and (5) the land market and labor market clear.

E.2 Model Analysis

Assume that the ability (s_{ai}, s_{ni}) follows a bi-variate log-normal distribution with mean $(0,0)$ and variance

$$\Sigma = \begin{pmatrix} \sigma_a^2 & \sigma_{an} \\ \sigma_{an} & \sigma_n^2 \end{pmatrix}.$$

The correlation between the agricultural and non-agricultural ability is $\rho_{an} = \frac{\sigma_{an}}{\sigma_a \sigma_n}$.

The first order condition of the non-agricultural sector gives

$$w_n = A_n. \tag{6}$$

Denote the marginal product of land for individual i as q_i. The first order conditions for land and capital in the agricultural sector are

$$l_i = A_a(\gamma p_a)^{\frac{1}{1-\gamma}} \left(\frac{1-\alpha}{r}\right)^{\frac{\gamma(1-\alpha)}{1-\gamma}} \left(\frac{\alpha}{q_i}\right)^{\frac{1-\gamma(1-\alpha)}{1-\gamma}} s_{ai}, \tag{7}$$

$$k_i = A_a(\gamma p_a)^{\frac{1}{1-\gamma}} \left(\frac{1-\alpha}{r}\right)^{\frac{1-\alpha\gamma}{1-\gamma}} \left(\frac{\alpha}{q_i}\right)^{\frac{\alpha\gamma}{1-\gamma}} s_{ai}. \tag{8}$$

The profit of farmer i and farm output are

$$\pi_i = A_a \frac{1-\gamma}{\gamma}(\gamma p_a)^{\frac{1}{1-\gamma}} \left(\frac{1-\alpha}{r}\right)^{\frac{\gamma(1-\alpha)}{1-\gamma}} \left(\frac{\alpha}{q_i}\right)^{\frac{\alpha\gamma}{1-\gamma}} s_{ai}, \tag{9}$$

$$y_i = A_a(\gamma p_a)^{\frac{\gamma}{1-\gamma}} \left(\frac{1-\alpha}{r}\right)^{\frac{\gamma(1-\alpha)}{1-\gamma}} \left(\frac{\alpha}{q_i}\right)^{\frac{\alpha\gamma}{1-\gamma}} s_{ai}, \tag{10}$$

For Type II farmers, they are at the kink, $l_i = \bar{l}$, and $q_i \in (q, q(1+\tau)]$, and the value of q_i is determined by plugging $l_i = \bar{l}$ into Equation 7. Type I farmers have $l_i < \bar{l}$ and $q_i = q$. Type III farmers have $l_i > \bar{l}$ and $q_i = q(1+\tau)$. Thus, for all individuals, $q_i = q\phi_i$, with $1 \leq \phi_i \leq 1+\tau$.

With the new representation, we can re-write the expressions for l_i, k_i and π_i as the follows,

$$l_i = l\phi_i^{-\frac{1-\gamma(1-\alpha)}{1-\gamma}} s_{ai}, \tag{11}$$

$$k_i = k\phi_i^{-\frac{\alpha\gamma}{1-\gamma}} s_{ai},$$

$$\pi_i = \pi\phi_i^{-\frac{\alpha\gamma}{1-\gamma}} s_{ai},$$

$$y_i = y\phi_i^{-\frac{\alpha\gamma}{1-\gamma}} s_{ai},$$

where $l = w_a\alpha\gamma/q$, $k = w_a(1-\alpha)\gamma/r$, $\pi = (1-\gamma)w_a$, $y = \frac{w_a}{p_a}$, and

$$w_a \equiv A_a\gamma^{\frac{\gamma}{1-\gamma}} p_a^{\frac{1}{1-\gamma}} \left(\frac{1-\alpha}{r}\right)^{\frac{\gamma(1-\alpha)}{1-\gamma}} \left(\frac{\alpha}{q}\right)^{\frac{\alpha\gamma}{1-\gamma}}$$

are common for all farmers. The terms with ϕ_i represent the extent of distortion for the farmers, and they are decreasing functions with respective to ϕ_i, meaning that the distortions are larger for households who want to utilize larger lands. Holding the non-agricultural ability constant, these farmers are the ones with higher agricultural productivity.

Given the log normal distribution of (s_{ai}, s_{ni}), define the following terms

$$u_{ai} \equiv \log(s_{ai}),$$

$$u_{ni} \equiv \log(s_{ni}).$$

Thus, $E(u_{ai}) = E(u_{ni}) = 0$, $Var(u_{ai}) = \sigma_a^2$, $Var(u_{ni}) = \sigma_n^2$, $cov(u_{ai}, u_{ni}) = \sigma_{an}$. Also, $u_{ni} - u_{ai}$ follows a normal distribution with mean 0 and variance $\sigma^2 \equiv \sigma_a^2 + \sigma_n^2 - 2\sigma_{an}$.

The probability of individual i choosing the agricultural sector is

$$\begin{aligned}
n_a^i &= \Pr(I_{ai} \geq I_{ni}) \\
&= \Pr(\pi_i \geq w_n s_{ni}) \\
&= \Pr(\log(\pi_i) \geq \log(w_n s_{ni})) \\
&= \Pr(\log((1-\gamma)w_a\phi_i^{-\frac{\alpha\gamma}{1-\gamma}} s_{ai}) \geq \log(w_n s_{ni})) \\
&= \Pr(\log(s_{ni}) - \log(s_{ai}) \leq \log((1-\gamma)w_a\phi_i^{-\frac{\alpha\gamma}{1-\gamma}}) - \log(w_n)) \\
&= \Pr(\frac{u_{ni} - u_{ai}}{\sigma} \leq \frac{b_a^i - b_n}{\sigma}),
\end{aligned}$$

where $b_a^i \equiv \log((1-\gamma)w_a\phi_i^{-\frac{\alpha\gamma}{1-\gamma}})$, and $b_n \equiv \log(w_n)$.

Depending on the equilibrium value of l_i, n_a^i can take three forms. For Type I farmers, $\phi_i = 1$,

$b_a = b_1$. For Type III farmers, $\phi_i = 1+\tau$, $b_a = b_3$. For Type II farmers, we first solve the ϕ_i from Equation (11),

$$\phi_i = \left(\frac{w_a \alpha\gamma s_{ai}}{q\bar{l}}\right)^{\frac{1-\gamma}{1-\gamma(1-\alpha)}}$$

As a result,

$$\begin{aligned}
n_a^i &= \Pr(\log((1-\gamma)w_a\phi_i^{-\frac{\alpha\gamma}{1-\gamma}} s_{ai}) \geq \log(w_n s_{ni})) \\
&= \Pr(\log((1-\gamma)w_a\left(\frac{w_a\alpha\gamma s_{ai}}{q\bar{l}}\right)^{-\frac{\alpha\gamma}{1-\gamma(1-\alpha)}} s_{ai}) \geq \log(w_n s_{ni})) \\
&= \Pr(\log((1-\gamma)w_a^{\frac{1-\gamma}{1-\gamma(1-\alpha)}}\left(\frac{\alpha\gamma}{q\bar{l}}\right)^{-\frac{\alpha\gamma}{1-\gamma(1-\alpha)}} s_{ai}^{\frac{1-\gamma}{1-\gamma(1-\alpha)}}) \geq \log(w_n s_{ni})) \\
&= \Pr(\log(s_{ni}) - \frac{1-\gamma}{1-\gamma(1-\alpha)}\log(s_{ai}) \leq \log((1-\gamma)w_a^{\frac{1-\gamma}{1-\gamma(1-\alpha)}} c) - \log(w_n)) \\
&= \Pr(\frac{u_{ni}-\widetilde{u}_{ai}}{\widetilde{\sigma}} \leq \frac{b_2 - b_n}{\widetilde{\sigma}}),
\end{aligned}$$

where $\widetilde{u}_{ai} \equiv \frac{1-\gamma}{1-\gamma(1-\alpha)}u_{ai}$, $\widetilde{\sigma}^2 \equiv \left(\frac{1-\gamma}{1-\gamma(1-\alpha)}\right)^2\sigma_a^2 + \sigma_n^2 - 2\frac{1-\gamma}{1-\gamma(1-\alpha)}\sigma_{an}$, $b_2 \equiv \log((1-\gamma)w_a^{\frac{1-\gamma}{1-\gamma(1-\alpha)}} c)$, and $c \equiv \left(\frac{\alpha\gamma}{q\bar{l}}\right)^{-\frac{\alpha\gamma}{1-\gamma(1-\alpha)}}$.

Now we want to solve for the range of s_{ai} for the three situations. In Equation 11, set $l_i = \bar{l}$, and $\phi_i = 1$, we solve

$$s_{ai} = \frac{\bar{l}q}{w_a\alpha\gamma} \equiv \underline{s},$$

and set $l_i = \bar{l}$, and $\phi_i = 1+\tau$, we solve

$$s_{ai} = \frac{\bar{l}q}{w_a\alpha\gamma}(1+\tau)^{\frac{1-\gamma(1-\alpha)}{1-\gamma}} \equiv \bar{s}.$$

Thus, for any individual to choose agriculture over non-agriculture, the probability is

$$\begin{aligned}
N_a = &\Pr(\frac{u_{ni}-u_{ai}}{\sigma} \leq \frac{b_1-b_n}{\sigma}, \frac{u_{ai}}{\sigma_a} \leq \frac{\log(\underline{s})}{\sigma_a}) \\
&+ \Pr(\frac{u_{ni}-\widetilde{u}_{ai}}{\widetilde{\sigma}} \leq \frac{b_2-b_n}{\widetilde{\sigma}}, \frac{\log(\underline{s})}{\sigma_a} < \frac{u_{ai}}{\sigma_a} \leq \frac{\log(\bar{s})}{\sigma_a}) \\
&+ \Pr(\frac{u_{ni}-u_{ai}}{\sigma} \leq \frac{b_3-b_n}{\sigma}, \frac{u_{ai}}{\sigma_a} > \frac{\log(\bar{s})}{\sigma_a}) \\
&= P(A) + P(B) + P(C).
\end{aligned}$$

E.3 Calibration

Step 1. Simplification First assume that the variance of agricultural ability and non-agricultural ability are the same, i.e., $\sigma_n^2 = \sigma_a^2$. The following terms can be simplified,

$$\sigma^2 \equiv \sigma_a^2 + \sigma_n^2 - 2\sigma_{an} = 2\sigma_a^2 - 2\sigma_{an}.$$

For Type I farmer, denote $X \equiv \frac{u_{ai}}{\sigma_a}$, $Y \equiv \frac{u_{ni}-u_{ai}}{\sigma}$, $a^I \equiv \frac{\log(\bar{s})}{\sigma_a}$, and $b^I \equiv \frac{b_1 - b_n}{\sigma}$. Then we know that $(X, Y) \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$, where $\rho \equiv \frac{1}{\sigma\sigma_a}(\sigma_{an} - \sigma_a^2) = \frac{1}{\sqrt{2\sigma_a^2 - 2\sigma_{an}}}\frac{1}{\sigma_a}(\sigma_{an} - \sigma_a^2) = -\frac{\sqrt{\sigma_a^2 - \sigma_{an}}}{\sqrt{2}\sigma_a}$, and

$$P(A) = P(X \leq a^I, Y \leq b^I).$$

For Type III farmer, denote $a^{III} \equiv \frac{\log(\underline{s})}{\sigma_a}$, $b^{III} \equiv \frac{b_3 - b_n}{\sigma}$, and

$$P(C) = P(X > a^{III}, Y \leq b^{III}).$$

The relationship between the cutoffs are as follows. For (a^I, a^{III}),

$$a^I = \frac{\log(\underline{s})}{\sigma_a} = \frac{1}{\sigma_a} \log \frac{\bar{l}q}{w_a \alpha\gamma},$$

$$a^{III} = \frac{\log(\bar{s})}{\sigma_a} = \frac{1}{\sigma_a} \log \frac{\bar{l}q}{w_a \alpha\gamma}(1+\tau)^{\frac{1-\gamma(1-\alpha)}{1-\gamma}} = a^I + \frac{1-\gamma(1-\alpha)}{1-\gamma}\frac{\log(1+\tau)}{\sigma_a}.$$

For (b^I, b^{III}),

$$b^I = \frac{\log((1-\gamma)w_a) - b_n}{\sigma},$$

$$b^{III} = \frac{\log((1-\gamma)w_a(1+\tau)^{-\frac{\alpha\gamma}{1-\gamma}}) - b_n}{\sigma} = b^I - \frac{\alpha\gamma}{1-\gamma}\frac{\log(1+\tau)}{\sigma}.$$

Also

$$\sigma = \sqrt{2}\sqrt{\sigma_a^2 - \sigma_{an}} = -2\rho\sigma_a.$$

For Type II farmer, denote $Z \equiv \frac{u_{ni} - \widetilde{u}_{ai}}{\widetilde{\sigma}}$, and $b^{II} \equiv \frac{b_2 - b_n}{\widetilde{\sigma}}$. Then we know that $(X, Z) \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{bmatrix}\right)$, where $\rho_2 \equiv \frac{1}{\widetilde{\sigma}\sigma_a}(\sigma_{an} - \frac{1-\gamma}{1-\gamma(1-\alpha)}\sigma_a^2)$, and

$$P(B) = P(a^I < X \leq a^{III}, Z \leq b^{II}).$$

The cutoff b^{II} is related to (a^I, b^I),

$$
\begin{aligned}
b^I &\equiv \frac{\log((1-\gamma)w_a^{\frac{1-\gamma}{1-\gamma(1-\alpha)}}c) - \log(w_n)}{\widetilde{\sigma}} \\
&= \frac{\sigma \cdot b^I + \frac{\alpha\gamma}{1-\gamma(1-\alpha)}\sigma_a \cdot a^I}{\widetilde{\sigma}},
\end{aligned}
$$

where $\widetilde{\sigma} = \left(\frac{1-\gamma}{1-\gamma(1-\alpha)}\right)^2 \sigma_a^2 + \sigma_a^2 - 2\frac{1-\gamma}{1-\gamma(1-\alpha)}\sigma_{an}$.

Step 2. Calibration of Parameters using Values from Literature We take the production function parameters α and γ directly from the literature, as shown in Table A.25.

Table A.25: Parameter values from the literature

Parameter	Value	Source
α	0.66	Adamopoulos et al. [2017]
γ	0.54	

Step 3. Calibration of the Variance-Covariance Matrix and The Distortion The value to be recovered: $(\rho, \sigma_a^2, a^I, b^I, \tau)$. We use the following five equations to recover them.

1. The probability of being a Type I farmer.

$$
P(A) = P(X \leq a^I, Y \leq b^I). \tag{12}
$$

2. The probability of being a Type II farmer.

$$
P(C) = P(a^I < X \leq a^I + \frac{1-\gamma(1-\alpha)}{1-\gamma}\frac{\log(1+\tau)}{\sigma_a}, Z \leq \frac{\sigma \cdot b^I + \frac{\alpha\gamma}{1-\gamma(1-\alpha)}\sigma_a \cdot a^I}{\widetilde{\sigma}}). \tag{13}
$$

3. The probability of being a Type III farmer.

$$
P(C) = P(X > a^I + \frac{1-\gamma(1-\alpha)}{1-\gamma}\frac{\log(1+\tau)}{\sigma_a}, Y \leq b^I - \frac{\alpha\gamma}{1-\gamma}\frac{\log(1+\tau)}{\sigma}). \tag{14}
$$

4. The variance of land for Type I farmers.

$$
\begin{aligned}
var(\log(l_i) \mid A) =& var(\log(l_i) \mid X \leq a^I, Y \leq b^I) \\
=& var(\log(s_{ai}) \mid X \leq a^I, Y \leq b^I) \\
=& \sigma_a^2 var(\frac{u_{ai}}{\sigma_a} \mid X \leq a^I, Y \leq b^I) \\
=& \sigma_a^2 var(X \mid X \leq a^I, Y \leq b^I).
\end{aligned} \tag{15}
$$

5. The variance of land for Type III farmers.

$$
\begin{aligned}
var(\log(l_i) \mid C) =& var(\log(l_i) \mid X > a^{III}, Y \leq b^{III}) \\
=& \sigma_a^2 var(X \mid X > a^I + \frac{1-\gamma(1-\alpha)}{1-\gamma}\log(1+\tau), Y \leq b^I - \frac{\alpha\gamma}{1-\gamma}\frac{\log(1+\tau)}{\sigma}). \quad (16)
\end{aligned}
$$

Take the 1995, 2001, and 2010 data. First, we plot the distribution of land per agricultural worker in Figure 8. The data is at the individual level, and deducted by the village-year mode. Observations with the value smaller than –0.16 are defined as Type I farmers, and observations with the value larger than 0.16 are defined as Type III farmers. The farmers with the value between –0.16 and 0.16 are Type II farmers.

With the definition of different types of farmers, the following information can be calculated from the data.

Table A.26: Data moments

Year	1995	2001	2010
Probability of choosing agriculture	.75	.71	.49
Total number of farmers	21,521	19,564	15,158
Number of Type I farmers	5,963	5,769	4,560
Number of Type II farmers	6,242	5,673	3,966
Number of Type III farmers	9,316	8,122	6,632
Prob of Type I farmers, conditioning on being a farmer	.28	.29	.30
Prob of Type II farmers, conditioning on being a farmer	.29	.29	.26
Prob of Type III farmers, conditioning on being a farmer	.43	.42	.44
Variance of log land, Type I	.097	.136	.237
Variance of log land, Type III	.158	.161	.232

Note: This table shows the data moments in 2001 and 2010, using the NFP Survey household level data. Type I farmers are the ones with land (minus the village-year mode) smaller than –0.16, Type II farmers are the ones with land (minus the village-year mode) bigger than –0.16 and smaller than 0.16, and Type III farmers are the ones with land (minus the village-year mode) larger than 0.16.

We proceed in the following order using the 2001 information. (1) Guess (ρ, τ, σ_a). (2) Solve the cutoff points (a^I, b^I) from Equations 12 and 14. (3) Solve σ_a from Equation 15. (4) Update σ_a such that the guess and the solution are close. (4) Choose τ such that the difference between LHS and RHS of Equation 13 is the smallest. (5) Choose ρ such that the difference between LHS and RHS of Equation 16 is the smallest.

Table A.27: Calibrated parameters

Year	1995	2001	2010
ρ		-0.6429	
σ_a		0.8469	
τ	1.6	1.6	1.2
a^I	-0.3939	-0.3333	-0.1515
a^{III}	0.2840	0.3446	0.4079
b^I	0.8981	0.7541	0.2036
b^{II}	0.8332	0.6987	0.1659
b^{III}	0.1724	0.1191	-0.4016

Note: This table shows the calibrated parameter values following the procedures in Step 3. The 2001 column calibrates all parameter values using the 2001 data. The 1995 and 2010 columns take the value of ρ and σ_a from the 2001 column and calibrates the rest of parameters with the 1995, and 2010 data, respectively.

Suppose that we now use the variance-covariance matrix calibrated using the 2001 data to back out other parameter values in 1995 and 2010. We proceed in the following order. (1) Guess τ. (2) Solve the cutoff points (a^I, b^I) from Equations 12 and 14. (3) Calculate the difference between the LHS and the RHS of Equation 13, and choose τ such that the difference is the smallest.

Step 4. Calibration of the Remaining Parameters

- Solve (w_a, w_n). Think about the individuals who are (1) Type I farmers in 1995, and (2) Type I workers in 2001. Then we know that the following conditions must be satisfied: $X \leq a^I_{1995}, Y \leq b^I_{1995}, X \leq a^I_{2001}, Y > b^I_{2001}$. Denote the set of these individual as H. Consider these switchers from 1995 to 2001, then we know that their non-agricultural income is,

$$\begin{aligned} E\log(I^{2001}_{ni} \mid i \in H) &= \log(w^{2001}_n) + E(\log(s_{ni}) \mid i \in H) \\ &= \log(w^{2001}_n) + E(u_{ni} - u_{ai} + u_{ai} \mid i \in H) \\ &= \log(w^{2001}_n) + \sigma E(Y \mid i \in H) + \sigma_a E(X \mid i \in H). \end{aligned}$$

Empirically, given the values of $(a^I_{1995}, b^I_{1995}, a^I_{2001}, b^I_{2001})$, the individuals with $i \in H$ are just the ones who worked in agriculture in 1995 as Type I farmer, and moved to non-agriculture in 2001.

Then w_n solved. Using the cutoff, $b^I = \frac{\log((1-\gamma)w_a) - \log(w_n)}{\sigma}$, then w_a solved.

Similarly, we use the 2001 to 2010 switchers to solve the (w_a, w_n) in 2010.

- Solve (r, p_a). Consider the Type I farmer in 2001. The value of output from farming is

$$\begin{aligned} E\log(p_a y_i \mid A) &= \log(w_a) + p_a \cdot E(\log(s_{ai}) \mid A) \\ &= \log(w_a) + p_a \sigma_a \cdot E(X \mid A). \end{aligned}$$

Then p_a solved.

Ratio of the value of output w.r.t. to inputs are as follows,

$$\frac{p_a y_i}{k_i} = \frac{1}{(1-\alpha)\gamma} r,$$

$$\frac{p_a y_i}{l_i} = \frac{1}{\alpha\gamma} q.$$

then $\{r, q\}$ can be solved.

- (A_a, A_n) can be solved using the following two equations:

$$w_n = A_n,$$

$$w_a = A_a \gamma^{\frac{\gamma}{1-\gamma}} p_a^{\frac{1}{1-\gamma}} \left(\frac{1-\alpha}{r}\right)^{\frac{\gamma(1-\alpha)}{1-\gamma}} \left(\frac{\alpha}{q}\right)^{\frac{\alpha\gamma}{1-\gamma}}.$$

- $\bar{l}$ can be solved using the following equation:

$$a^I = \frac{\log(\underline{s}) - \mu_a}{\sigma_a} = \frac{1}{\sigma_a} \log(\frac{\bar{l} q}{w_a \alpha\gamma}).$$

Table A.28: Calibration of remaining parameters

Panel A: Data moments	2001	2010
Mean of log non-agricultural income for switchers	8.10	9.31
Mean of log agricultural output value for type I farmers	7.33	7.67
Mean agricultural output to land ratio for type I farmers	1271	1777
Mean agricultural output to capital ratio for type I farmers	3.72	4.83
Panel B: Calibrated parameters		
$\bar{l}$	3.65	5.29
p_a	2.64	3.66
r	1.80	3.25
$\log(A_n)$	8.1	9.6
$(1-\gamma)\log(A_a)$	6.8	7.2

Note: This table shows the data moments used and the calibrated parameter values in Step 4. The non-agricultural income includes wage income and income from land leasing, and it is calculated by dividing the household-level income by the number of non-agricultural workers. Switchers are the ones who were working in the agricultural sector in the previous year as a Type I farmer and work in the current period as a wage earner. The log agricultural output value is calculated as the household-level crop output value (minus the cost of intermediate inputs) divided by the number of agricultural workers.

E.4 Quantitative Experiments

Step 1. Generate a sample of 1,000,000 individuals with agricultural and non-agricultural ability with the calibrated variance-covariance matrix.

Step 2. Benchmark economy: use the cut-off points calculated from the 2010 data to determine occupation choice and type of farmer. Then calculate the capital and land allocation accordingly. Generate the total land size.

Step 3. Counterfactual. First guess the price q. Given the vector, determine the cutoff points $(a^I, a^{III}, b^I, b^{II}, b^{III})$. Then do the same exercise as in the BE, such that the land market clears.