Does Trade Liberalization Promote Opportunity? Evidence from China

Dan Liu*
(Shanghai University of Finance and Economics)

September 10, 2018

Very preliminary! Please don't circulate

Abstract

In this paper, we investigate the relationship between trade liberalization and intergenerational income mobility in China. We find that tariff cuts upon China's accession to WTO has a significant negative impact on overall mobility. Our results show that trade liberalization promotes upward mobility but hinders downward mobility, and the later effect dominates. The existence of this pattern does not depend on parents' income and education group. In addition, we find a potential spillover impact of trade liberalization from manufacturing to rural residents. The direct impact of trade liberalization on education decision is one potential explanation for our main results.

Key Words: Trade liberalization, Intergenerational mobility, China
JEL Classification: F16 F13

* Dan Liu, School of Economics, Shanghai University of Finance and Economics. Email:liu.dan@mail.shufe.edu.cn. Address: 777 Guoding Road, Shanghai 200433, PR China. Dan thanks financial support of the National Natural Science Foundation of China (No. 71703187).

1. Introduction

Social fairness has been an extremely important consideration in the process of economic development. Rising cross-sectional inequality and declining intergenerational mobility have drawn more and more attention in both developed countries (Corak, 2013; Clark, 2014; Chetty, et al, 2014b; Hilger, 2015; Chetty, et al, 2018) and developing countries (Gong, et al, 2012; Fan, et al, 2018). Does globalization contribute to this trend? A large number of studies have focused on the inequality aspect and found that globalization can lead to higher inequality (Goldberg & Pavcnik, 2007). In contrast, little attention has been paid to the impact of openness on intergenerational mobility. Recent influential works have documented spatial variation in intergenerational income mobility within a country using administrative data or census data (Chetty, et al., 2014a; Hilger, 2015) but leave the causality issues unanswered. In this paper, we investigate the impact of trade liberalization on intergenerational income mobility in China, which has experienced the fastest growth in trade over the last two decades. By exploring the plausibly exogenous tariff reduction upon China joining the WTO in 2001, we aim to identify the causal impact of tariff cuts on intergenerational income mobility.

To achieve this goal, we examine the spatial variation in intergenerational income mobility across cities in post-WTO China. There are three advantages to do so. First, it is widely documented that China's official accession to WTO in December, 2001 can be considered as an exogenous shock to the local labor markets (Yu, 2015; Brandt, et al., 2017; Dai, et al, 2018). After joining WTO, the tariffs were significantly lowered at industry level following the commitments made by the Chinese government at the stage of pre-WTO negotiations. Second, the geographic variation in a large country like China is of especial interest due to its unique Hukou system (Liu, 2005; Chan and Zhang, 1999), which leads to more segmented local markets by imposing relatively high migration costs both across regions and between rural and urban within a region. This special feature also generates interesting questions, such as the spillover effect of trade liberalization from urban to rural areas. Third, by focusing on the pattern of spatial variation, we plausibly alleviate the data requirements on examining intergenerational mobility and justify the use of census data in China for the first time to explore this relationship (Hilger, 2015).

We make three major contributions to the related literature. First, we show that more open to trade in fact has significant negative impact on mobility in China. In particular, locations with higher degree of trade liberalization have experienced higher upward mobility but at the same time much lower downward mobility, which plays a dominant role on overall mobility. Second, we show that trade liberalization has a positive spillover effect on rural upward mobility. Specifically, it is found that rural upward mobility is positively associated with the degree of manufacturing trade liberalization. Urban upward mobility is in fact not significantly affected and the negative impact on downward mobility is significantly higher for urban families. This indicates that the supply side force needs to be taken into consideration in China. Third, we find education, especially high school education, is one explanation of our story. One important policy implication from this result is that expanding high school education can promote the positive impact of trade liberalization on mobility.

This paper is the most closely related to Ahsana and Chatterjeeb (2017), who study the impact of trade liberalization on intergenerational occupational mobility in India. They find that tariff reduction has a positive impact on upward occupational mobility but no significant impact on downward mobility. Our paper is complimentary to their approach in two aspects: First, we focus on income mobility rather than occupational mobility. These two aspects of mobility are not necessarily equivalent, especially if the variation of income within the same occupation is significantly large (Autor and Handel, 2013). Second, our data is larger and more comprehensive with both rural and urban observations, which makes it feasible to explore the heterogeneous impacts of trade liberalization on the two groups and the potential spillover effect from urban to rural.

This research is also inspired by the large amount of recent literature examing the impacts of trade liberalization on local labor markets (Autor, et al., 2013; Sheng and Yang, 2016; Feenstra, et al., 2017, Dai, et al., 2018). This line of literature explore the geographic variation to identify the causal connection between tariff cuts and local labor market outcomes, such as employment, wage and inequality. Our paper contributes to this literature by exploring the effect of trade liberalization on local income mobility, which is an extremely important aspect of "inclusive growth" (Anand, et al. 2013). Our research provides another angel on the impact of globalization on social fairness.

The rest of this paper is organized as below. Section 2 introduces our data and the measures of mobility and trade liberalization. In Section 3, we explain the empirical strategy and establish the baseline results. Section 4 conducts some more robustness checks and Section 5 concludes.

2. Data and Measures

2.1 Data

To explore the spatial variation in intergenerational mobility, we use the mini census in 2005.[1] It is the largest national representative dataset in China available with matched child-parent pairs and income information for both generations, to our best knowledge. Our sample is a subset of the original survey, which contains about 2,585,000 individual observations randomly drawn from the original dataset at National Bureau of Statistics in China. We match children and parents based on the household ID number. It covers a wide range of information on job characteristics (industry, occupation, income, etc.), demographic characteristics (age, sex, minority, etc) and Hukou (rural or urban, location).

In our analysis, we only include sons, following Ahsan and Chatterjee (2017), for two reasons: First, the rate of co-living for son and parents is almost double of that for daughter and parents in the data, partially due to the tradition of "raising sons for the old"

[1] The notion of using census data to examine mobility issues was inspired by Hilger (2015) who shows that the census data presents extremely similar geographic pattern of intergenerational mobility to those from the administrative data in the US.

in China. It can potentially complicate our results if this preference varies across regions. Second, the related literature (Fan, et al., 2018; Gong, et al., 2012) have shown that the degree of daughter-parent and son-parent persistence are significantly different in China. This feature makes the potential variation in sex composition across regions further confound our results if both genders are included. In addition, difference in labor market participation between genders can also play a role. Excluding daughter-parent pairs, we efficiently minimize all those potential issues. Moreover, we also restrict our sample to sons between the age of 16 and 35.[2] One reason for setting an age limit is to minimize the potential measurement errors due to some special reasons for living with their parents after certain age, such as disabilities. It can be shown in our later analysis that our results are robust if we set the age cutoff to 32 or 40, or even if we remove the age cutoff. Furthermore, our analysis is restricted to family pairs with both son and father employed and earning positive income. A person is identified as "employed" as long as he worked for more than one hour a week including "Self-employed". There are 62,834 family pairs left with all the information needed for this analysis.

Our working sample includes both rural and urban residents. In China, under the household registration system (Hukou system), each person is registered as a resident of a city or its affiliated rural area[3]. It has two broad categories: urban Hukou and rural Hukou. The government adopts different policies for rural and urban residents in terms of housing, pension and education. For example, only people with a rural Hukou can be assigned farmlands. A child inherited his/her parents' Hukou identity when he/she was born. A person's Hukou status can be changed from "rural" into "urban" for three main reasons: going to college, marrying someone with urban Hukou, and due to illness or disability, having to live with children who have urban Hukou. In our sample, only less than 3.5% of family pairs have son's Hukou status different from father's status. In our later analysis, we investigate the heterogeneous impacts of trade liberalization on these two groups of families.

One concern using census data to examine intergenerational mobility is the potential bias due to the selection of co-residing sample. If the choice of co-living with his parents is affected by the degree of local trade liberalization exposure, then it can lead to potential bias in our estimate of the impact of trade liberalization. To deal with issue, we first compare our working sample, which only includes sons co-living with their parents, with the full sample that covers all males between the age of 16 and 35. The summary statistics of some main variables are reported in Table 1. It can be seen that males coliving with their parents are on average younger, more likely to be single, living in bigger households and with lower income. The biggest concern here is the age difference since it has been well documented that the degree of intergenerational persistence in income increases with sons' age (Haider and Solon, 2006; Chetty, et al., 2014b). This implies that the selection of our working sample can lead to an overestimate of intergenerational mobility. Whether this measurement error can result in further bias in our estimate of the

[2] It should be noted that the family pairs where a child's age is greater than 35 only accounts for about 0.9% of the total matched sample.

[3] Liu (2005), Chan and Zhang (1999) provide more detailed information on the history and the reform of the Hukou system.

impact of trade liberalization on mobility depends on how it correlated with trade liberalization. It can be shown in Section 4 that neither the probability of coliving with parents nor the average age difference between the coliving sample and the full sample is significantly associated with the degree of local trade liberalization. Despite this result, we further conduct rigorous robustness checks to deal with the coliving issue to reassure that our results are not driven by this. Moreover, we control for both sons' and fathers' ages to minimize the potential life cycle bias.

There are two concerns with income information from the census data. First, there is an upper bound of income in census, which is a common feature in many other self-report census data all over the world. Second, there can be potential bias caused by transitory income shocks since it is only available for one year. Ideally, it would require both sons' and parents' permanent income information to measure intergenerational mobility precisely. In the literature, average income over three or five years is used as a proxy of permanent income (Chetty, et al., 2014a; Gong, et al., 2012). There are two recent findings that justify the use of census data. Chetty, et al (2014a) examine the potential bias caused by transitory income shocks using the richest income information from the administrative data in the US. By examining the rank-rank intergenerational mobility measures using one year, two-year average and three-year average income, they find that the bias using one year income is relatively small. In addition, Hilger (2015) compares the mobility measures based on census income information with those from the administrative data. It is found that the two datasets exhibit extremely similar geographic variation.

2.2 Intergenerational mobility measures

We adopt the rank-rank style intergenerational income mobility in our main analysis. There are three major measures of intergenerational mobility: intergenerational income elasticity (IGE)[4], coefficient of correlation between son's and parents' income[5], and rank-rank mobility. There are three advantages to use the rank-rank measure. First, rank-rank measure is much less sensitive to the potential nonlinearity in the log-log relationship of income between the two generations. Second, rank measure is also more robust than the other two measures when there are potential measurement errors in income. Finally, rank-rank mobility can be easily adopted to examine the directions of mobility, upward or downward. However, it can be shown that our results do not depend on the particular measure, as shown in Section 4.

For each family pair, mobility is defined as an indicator, equal to one if son's income decile is different from parents' income decile and zero if otherwise. Income deciles are based on income distribution within each generation. Furthermore, if a son's income decile is higher than parents' income decile, it is defined as the case of upward mobility, and if a son's decile is lower than parents' income decile, it is identified as the case of downward mobility. Figure 1 illustrates the intergenerational persistence in income

4 IGE is the elasticity of a child's income on parents' income, estimated as the coefficient from a log-log regression. It is the most widely adopted measure of intergenerational transmission of income in the literature. The main concerns to this measure is the nonlinearity in log-log relationship, and the potential issue caused by change in income dispersion between the two generations. Please refer to () for more discussion on this.

5 This measure corrects the potential influence from income inequality but still suffers the non-linearity bias.

ranks shown by our data. Panel A considers the sons that were born to the families in the bottom income decile. Each bar shows the fraction of sons belonging to each decile. It can been seen that about a half of these sons keep falling into the bottom income decile in their own generation, with the rest a half moving up to other deciles. Similarly, Panel B examines those sons from the top family income decile. Around a third of them still maintain the top income. However, more than a half of those whose income decile falls move into the second top income decile, and only less than 10% fall below the medium. This implies that although the downward mobility for the top decile families seems to be higher, most of them continue maintaining high income ranks. This also leads to the potential concern regarding this rank measure with fixed bins.

To show that our results do not depend on the choice of percentile bins, we also consider two alternative rank measures of mobility. First, we replace decile bins with quintile bins and adjust the measures of mobility accordingly based on these new bins. Second, we construct moving bins around each observation. For example, mobility can be defined as an indicator equal to one if a son's income percentile is higher or lower than his parent's income percentile than 5%. Upward and downward mobility can be defined accordingly in a similar way. Moreover, we have tried replacing 5% by 10% to further show that our results are not sensitive to the width of the bins. It can be shown in Section 4, our results are robust to all these alternative measures.

In our analysis, we also consider rural and urban mobility separately. This is motivated by the special "rural-urban dual structure" in China. The central government has adopted extremely different policies in allocating resources between these two groups of people, including taxation, pension, education, medical care and so on. Therefore, it's meaningful to see if they exhibit various patterns. On the one hand, this will help us to identify if the composition matters to local mobility caused by the process of urbanization. On the other hand, it's interesting to examine the potential spillover effect of trade liberalization in manufacturing to the rural neighborhood. Figure 2 shows the overall degree of mobility measures among urban and rural people. The overall rural average mobility is about 4% higher than urban mobility. This is mainly attributed to the much higher downward mobility among rural residents. Urban residents in fact exhibit higher upward mobility. In Section 3, we will show how trade liberalization affects the two groups differently.

2.3 Local exposure to trade liberalization

Regarding local exposure to trade liberalization, we consider prefectural cities as our geographic units.[6] There are 345 prefectural cities in total covered by the census.[7] For each city, we construct the degree of tariff reduction following Autor, et al. (2013)[8]:

---

[6] As discussed in Zhang and Zhao (1998), a prefectural city includes not only the central city, but also the affiliated towns and counties with both rural and urban areas.

[7] It can be shown that our main results are robust to the exclusion of small cities.

[8] This is a widely adopted approach in the literature to construct local labor market shocks from natural trade liberalization at industry level.

$$\Delta\tau_c^{2001-2004} = \sum_i \left(\frac{l_{ic}^{1999}}{\sum_i l_{ic}^{1999}} \right) \times \Delta\tau_i^{2001-2004}$$

where *c* represent city and *i* represents industry. $\Delta\tau_i^{2001-2004}$ is the change of tariff in industry *i* from 2001 to 2004.[9] The tariff data are at 3-digit Chinese Industry Classification (CIC) level from custom dataset, covering all manufacturing industries.[10] l_{ic}^{1999} is the total employment of industry *i* in city *c* in 1999. The summation term as the denominator in the bracket is the total employment is thus the total employment of manufacturing in city *c*. The whole ratio represents the share of employment of industry *i* in total manufacturing employment in city *c*, which measures the importance of industry *i* in city *c*. We use the employment information in 1999 in order to avoid the potential endogenous impact of trade on local employment structure.[11] In a similar way, we also construct the change in other types of liberalization including input tariff reduction and FDI deregulation. The employment information at industry-city comes from the Annual Survey of Industrial Firms (ASIF).[12]

Our data shows significant tariff cuts in this period as indicated by Yu (2015) and Dai, et al. (2018). The average tariff reduction at industry level is about 45% of the initial average tariff in 2001. The city-level exposure to tariff cuts varies across cities, ranging from less than 10% to more than 90% of the initial tariff level. This variation in tariff reduction across cities allow us to explore the relationship between trade liberalization and local intergenerational mobility. The identification of the causal impact relies on the assumption that industry level tariff reduction is not associated with local labor market characteristics at city level, which is commonly considered plausible in the related literature. Finally, it should be noted that the degree of input tariff cuts is much lower than output tariff cuts.

3. Empirical Specification and Results

3.1 Model Specification

To investigate the impact of trade liberalization on local intergenerational mobility, we set up a simple Probit model as follows:[13]

$$\text{Prob}\ (\text{mobility}_{ic} = 1) = \Phi(\alpha_0 + \alpha_1 * \Delta\tau_c^{2001-2004} + A * X_i + B * Y_c^{2001} + \mu_p) \quad (1)$$

where i represents family pair and c is city. mobility_{ic} is an indicator equal to one if a son belongs to a different income decile from his father's. We also use two other measures of mobility, upward mobility and downward mobility. Upward mobility is an indicator equal to one if a son's income decile is higher than his father's, and downward mobility is also an indicator equal to one if a son's income decile is lower than his father's. The

9 We consider tariff reduction until 2004 because it could take some time for tariff change to affect local labor market. In addition, about 90% of tariff cuts in the post WTO period happened between 2000 and 2004, as shown by Figure 1 in Dai, et al. (2018).

10 Manufacturing industries account for more than 80% of total exports and imports in China.

11 We also tried using labor compositions in 2000 or 2001. The results are similar.

12 It should be noted that the survey only covers firms with sales equal or above 5 million yuan. However, as pointed out by Yu (2015), these firms account for more than 90% of total industrial output in China.

13 We also adopt a Logit or OLS model as our robustness checks later.

fundamental assumption here is that the probability of experiencing certain movement along the income ladders between the two generations depends on both individual and local labor market characteristics.

The impact of trade liberalization is captured by α_1 in front of $\Delta\tau_c^{2001-2004}$, which is the degree of tariff cut exposure of city c between 2001 and 2004, measured as the weighted average of industry-level tariff cuts. A city experiences higher exposure to trade liberalization if the industries with high tariff cuts accounted for large shares in total employment before the trade reform. X_i includes the common controls of son's characteristics, such as age, age squared, marital status, father's age and age squared, household size and minority indicator. The son's and father's age and age squared are used to control for the life cycle pattern in intergenerational mobility as documented in Chetty, et al. (2014a). Similar approach is adopted by Fan, et al. (2018). Household size controls for the potential differences in the allocation of resources among family members. In addition, we also control for fathers' education attainments in order to tease out the potential influence of parents' education on mobility.

As pointed out by the related literature on trade liberalization and local labor market, such as Autor, et al. (2013), Ahsan and Chatterjee (2017) and Dai, et al. (2018), the estimate can be biased if there are missing variables associated with both local trade liberalization and mobility. Therefore, it's necessary to further control for such characteristics. Y_c^{2001} is a collection of city-level control variables for city size, share of the illiterate and share of agricultural employment in 2000.[14] For example, larger cities tend to be more unequal (Baum-Snow & Pavan, 2013) which is associated with lower mobility (Corak, 2013). This could lead to upward bias in the estimate of α_1 if city size is not controlled for. For similar reasons, education and industrial compositions are also controlled for. Finally, we also include province fixed effects as represented by μ_p in the specification to control for any unobserved state-level time invariant factors that can affect mobility.

3.2 Empirical Results

The baseline results from estimating equation (1) are reported in Table 2. In the first two columns, we examine the impact of trade liberalization on overall intergenerational mobility. The dependent variable is an indicator equal to one if a son's income decile is different from his parents' income decile. Column (1) controls for individual-level controls but not city-level controls. The estimate of the coefficient on tariff change is positive and significant at 1% level. This means that sons living in a city with a greater exposure to tariff cuts, i.e. tariff change is negative and bigger in absolute value, are less likely to move into an income decile different from their fathers'. In other words, trade liberalization is negatively associated with overall mobility in a city. This result continues to hold after we control for city characteristics in column (2). The estimate of the coefficient on tariff change is slightly smaller than in columns (1) as expected. However, this does not provide any information on the direction of mobility.

In columns (3) and (4), we investigate the relationship between tariff change and

[14] It can be shown that it doesn't matter if we use control variables in earlier years, such as 1999 or 1998.

upward mobility. The dependent variable is an indicator equal to one if sons' income decile is higher than parents' income decile, and zero if otherwise. It captures whether the sons climbed up the income decile ladders. Similarly, column (3) reports the results without the pre-WTO city-level controls. The estimate of the coefficient on tariff change is negative and significant at 5% level. This suggests that it's more likely for sons to move up into a higher income decile in cities with higher exposure to trade liberalization. In other words, trade liberalization promotes local upward mobility. Column (4) shows that this result holds after adding city-level controls. Combining this with those in columns (1) and (2), it implies that the negative impact of trade liberalization on total mobility is not caused by its effect on upward mobility.

We next examine the relationship between tariff change and downward intergenerational mobility. The results are reported in the last two columns. The dependent variable is an indicator equal to one if the son's income decile is lower than the parents' income decile, and zero if otherwise. In column (5), we estimate equation (1) without including city-level controls. It can be seen that the coefficient of interest is positive and highly significant at 1% level with a much bigger magnitude than those in the first four columns. The magnitude decreases slightly after adding the city-level controls but still highly significant. This result suggests that it's less likely for sons to move down into a lower income decile than their parents if they lived in cities with higher degree of exposure to tariff cuts. This means that trade liberalization in fact decreases local downward mobility, which results in a negative impact on local overall mobility as shown in the first two columns.

To better understand the magnitude of the impact of tariff cut on intergenerational mobility, we calculate the average marginal effect of tariff cut. If the tariff exposure at city level decreases by 1%, the probability for sons to move into a different decile from their parents' would decrease by about 0.20%, the probability of moving up into a higher decile would increase by about 0.16% and the probably of moving into a lower decile would decrease by 0.31%.[15] Given that the standard deviation of the probability of downward moving is 0.12, and the standard deviation of tariff change is about 17%, tariff cut can approximately explain about 42% of the difference in downward mobility between the 15th and the 85th percentile.

In addition, based on the estimates of the coefficients on age and age squared in columns (1) and (2), we find some interesting life cycle patterns. The coefficient on age is positive and that on age squared is negative. Both are significant at 1%. This suggests a non-monotonic relationship between mobility and age. Specifically, the probability of moving into a different income decile first decreases as the age of the group of sons is older, and then it increases after certain age. This is consistent with the finding from Figure 3 in Chetty, et al. (2014b). This also justifies the control of age and age squared to deal with the potential life-cycle bias.[16] The share of the illiterate is negatively associated with mobility as expected. Neither industry composition nor city size has significant connection with intergenerational mobility.

---

15 The marginal effects evaluated at the means of all variables is about 0.20%, 0.15% and 0.29%, respectively.

16 We also try using age dummies to better control for the potential life-cycle bias. The results are similar.

Next, we examine the potential heterogeneous impact of trade liberalization on intergenerational mobility depending on parents' income and education. The results are reported in Table 3. For each group, we consider both upward mobility (columns 1, 3 and 5) and downward mobility (columns 2, 4 and 6). All regressions include individual controls, city characteristics and province fixed effects. Panel A shows the estimates for various quartile groups based on parents' income. It can be seen that the coefficient of interest does vary across groups with the largest magnitude for the second family income quartile and smallest for the first quartile. This indicates that trade liberalization promote upward mobility and hinders downward mobility more for sons from middle and upper income families. The sign pattern is the same across all groups. This further confirms our earlier findings do not depend on family income.

Panel B reports the results for groups based on fathers' education groups: below middle school, middle school and above middle school. There are two main findings here: first, the impact of tariff change on upward mobility decreases with father's education with insignificant effect for those above middle school. This means that only lower and middle educated families benefit from higher upward mobility after tariff cuts. One potential explanation is that the parents above middle school mostly belong to the top decile which is unlikely for sons' to move upward. Second, the coefficient of tariff change is the largest for fathers' education above middle school. This suggests that sons with highly educated parents can benefit slightly more from trade liberalization in term of lower probability to move down along the income ladder.

In summary, from the above analysis we find that trade liberalization promotes upward mobility but hinders downward mobility in China, with a negative overall impact on mobility. On the one hand, the result on upward mobility is consistent with that in Chetty, et al. (2014a) which shows a positively connection between rank-rank slope and trade shock from China. It is also in line with the finding in Ahsan and Chatterjee (2017) that trade liberalization promotes upward occupational mobility in Indian. On the other hand, our finding on downward mobility also indicates that trade liberalization can contribute to the "class solidification" through its negative impact on downward mobility.

3.3 Mechanisms

To better explain the above findings, we investigate three potential mechanisms that can explain the above finding in this section. First, we examine the heterogeneous impact of trade liberalization on rural and urban residents/jobs, respectively, to illustrate the potential role of urbanization. Second, we study whether our result is actually driven by the industrial composition of jobs. Third, we explore if educational mobility can explain the finding on upward mobility.

Given the information available in the census, we separate parent-son pairs into urban and rural groups according to fathers' Hukou status or occupation. The unique Hukou system requires everyone to be identified as rural or urban. Only residents with rural Hukou identity were allocated farmlands. However, it's not necessarily the case that he would work on the farmland. An individual with rural Hukou can work with a non-agricultural occupation, such as drivers, managers and accountants, in an urban neighborhood. Therefore, these two methods of identifying people as rural or urban are complementary to each other.

Table 4 reports the results. The dependent variable is an indicator equal to one if the son's income decile is higher (columns 1 and 3) or lower (columns 2 and 4) than his father's income decile. The first two columns use Hukou status to identify rural and urban, and the last two columns use occupation. The coefficient of interest is the interaction term of tariff change with the urban indicator equal to one for fathers with urban Hukou and zero if otherwise. Column (1) shows a significantly positive coefficient on the interaction term. This suggests that sons born to fathers' with urban Hukou benefit less from the same tariff cut than those born to fathers' with rural Hukou in terms of moving into a higher decile than their parents. In other words, trade liberalization promotes upward mobility more among rural residents than urban residents. Column (3) shows similar result using occupation to identify rural and urban jobs. This rules out urbanization as a potential explanation for the positive association between trade liberalization and upward mobility.

Column (2) and (4) report the result on downward mobility. The coefficient of interest is positive and significant in column (2), which suggests that sons born to fathers' with urban Hukou benefit more from trade liberalization than sons born to fathers' with rural Hukou in term of lower downward mobility. This indicates that urbanization based on Hukou status may explain the negative association between downward mobility and trade liberalization. However, the result in column (4) shows the opposite sign if rural and urban are classified by occupation. This implies that trade liberalization could have led to tougher competition among non-agriculture jobs which further results in higher downward mobility.

Next, we investigate the role of industrial compositions on explaining our baseline finding. Table 5 reports the results. The dependent variable is the indicator for upward mobility or downward mobility defined as in Table 4. We consider three broad industries: agriculture, manufacturing and service. Manufacturing is an indicator equal to one for sons working in manufacturing industries, and zero if otherwise. Similarly, service is an indicator equal to one for sons working in service industries. The coefficients of interest are the interaction term of tariff change with industry indicators. Column (1) and (2) shows that tariff change both promotes intergenerational upward mobility and decreases downward mobility more among manufacturing sons. Similarly results are reported in columns (3) and (4) for service industry. The last two columns consider both manufacturing and service in the regressions, and the results are of the same pattern. This indicates that the impact of tariff cuts on local mobility can be partially explained with its heterogeneous impacts across industries.

It has been shown that education is one determinant of intergenerational mobility in the literature (Fan, et al., 2018). In order to examine if education is one explanation to our baseline finding, we first separate the whole sample into two cohorts according to sons' age. The early cohort includes sons older than 25 in the survey year and the later cohort covers sons at the age of 25 or younger. The age threshold is chosen so that the early cohort would have finished college before China's accession to WTO in 2001, which means that it's unlikely for trade liberalization to have any impact on their education choice. If education is one potential mechanism, the direct impact of trade liberalization on education should be through the later cohort, whose education decision can be actually affected by trade liberalization. The results reported in Table 6 shows that our baseline finding on tariff change and mobility hold for both cohorts with larger impacts on the early cohort. This does not exclude the potential influence of trade liberalization on education for the later cohort. It also indicates that education may not be the only explanation for our results.

We then proceed to investigate if education is one potential explanation for the association between tariff cuts and mobility. The results are reported in Table 7. The first two columns examine whether tariff change has a direct impact on upward education mobility. The dependent variable is an indicator equal to one if sons had higher education attainment than fathers. Column (1) restricts the sample to sons at the age of going to college in 2001. It can be seen that tariff change is negatively associated with upward mobility but not significant. Column (2) restricts the sample to sons at the age of going to high school in 2001. The coefficient on tariff change is negative and significant at 1%. This indicates that tariff cut significantly promotes upward educational mobility among sons who made the decision on high school education after China joined WTO. Columns (3)-(5) continue focusing on sons younger with age below 24. Column (3) further restricts the sample to those family pairs with no educational mobility, i.e. sons had the same education with their fathers. The dependent variable is the indicator for upward income mobility. It can be seen that the coefficient of interest becomes insignificant. This implies that trade liberalization does not have any impact on intergenerational mobility for those families with no education mobility. Columns (4) and (5) restrict the sample to those family pairs with upward education mobility. It can be seen that the impact of trade liberalization on both upward mobility and downward mobility exit for this group of people. All these results combined together provide supporting evidence that trade liberalization promotes upward educational mobility and therefore upward income mobility. At the same time, this force also decreases the downward mobility.

Thus far, urbanization cannot explain the association between trade liberalization and upward mobility but Hukou based urbanization may explain the connection between tariff cut and downward mobility. In addition, we find that trade liberalization can directly promotes upward educational mobility and therefore promotes upward income mobility and hinders downward mobility. However, it should be noted that we do not exclude the potential possibility that education can also result in externality on those in the early cohort.

4. Robustness Checks

4.1 Selection Bias

As we have pointed out earlier, one concern with the census data to explore intergenerational mobility is that we can only observe the co-residing parent-son pairs. Although we have shown that tariff change is not significantly associated with the co-living pattern in our sample, it may still be a concern that the selection of our sample can lead to potential bias in our estimates since these families can be significant different from the representative sample. A major concern is that the sons in our working sample are significantly younger than the overall sample including both coliving and non-coliving sons. The younger sons' outcomes are less dependent on family background due to the life-cycle pattern of intergenerational mobility shown in Chetty, et al. (2014a). This can result in downward bias in our estimate of the impact of trade liberalization, which may only show up in the later stage of son's life. In this subsection, we deal with the potential selection bias rigorously using the propensity score weighting (PSW) method suggested by Francesconi and Nicoletti (2006) and is adopted by Ahsan and Chatterjee (2017) to deal with similar issues.

The first stage selection equation is specified as follows

$$\text{Prob}\ (coliving_i = 1) = f(\Gamma Z_i) \qquad (2)$$

where *i* represents sons. $coliving_i$ is an indicator equal to one if son *i* and parents were living together, and zero if otherwise. Z_i includes the observable characteristics that can determine whether son *i* lived with his parents or not. *f(.)* is a general function form to predict how the probability of son *i* living with his parents is associated with his observable characteristics in Z_i. The underlying assumption here is that the probability of coliving can be consistently predicted by the observables in Z_i. We include sons' cohort fixed effects, indicator for marital status, indicator for minorities and province fixed effects in the regressions to be consistent with Francesconi and Nicoletti (2006).

At the second stage, we used the inverse of the predicted probability of co-residing from the first stage as weights and re-estimate equation (1). The results are reported in Table 8. The first three columns use a Probit model to estimate equation (2) in the first stage, and the last three columns replace it with a Logit model. It can be seen that the coefficient on tariff change is in line with those in Table 2 without weighting in terms of both sign pattern and significance. In addition, the magnitude of the estimates are slightly bigger than those without using PSW correction. This is highly consistent with our conjecture and those in Ahsan and Chatterjee (2017).

Therefore, we can conclude from the above analysis that although the co-residing selection of sample may lead to a downward bias in mobility measures and the estimates of our main coefficient, the impact of trade liberalization does not change significantly after we correct these biases.

4.2 Alternative Mobility Measures

In this subsection, we conduct further robustness checks using alternative measures of intergenerational mobility. There are mainly three measures of intergenerational mobility: transition matrix, intergenerational income elasticity (IGE) and the slope from running a rank-rank regression. Our measure is based on the transition in deciles between the two generations, which is in line with the transition matrix measure. One advantage of this measure is that it's convenient to explore the directions of movement. The disadvantage is that it separates the sample into equal income bins and does not take the difference within the bins into consideration. For example, consider two father-son pairs. One pair is like this: a father's income is at the 19^{th} percentile, and his son's income is at the 21th percentile. According to our measure, this is counted as an upward movement from the second decile into the third decile. The other pair includes a father at the 11th income percentile and a son at the 19th percentile. Obviously, the improvement made by the latter family seems to be bigger but it is counted as no mobility according to our definition.

To deal with the above potential measurement issue, we define mobility using the difference between fathers' and sons' percentiles. We consider two threshold: 10 percentiles and 15 percentiles. For example, if a son's income percentile is higher than his father's income percentile by 10 or more, the indicator for upward mobility is equal to one and zero if otherwise. If a son's income percentile is lower than his father's income percentile by 10 or more, the indicator for downward mobility is equal to one and zero if otherwise. The indicator for mobility is equal to one if either upward mobility or downward mobility is equal to one and zero if otherwise. The results using these alternative measures are reported in the first three columns of Table 9. It can be seen that both the signs and significance pattern is in line with our earlier finding. Moreover, the last columns show the results using 15 percentiles as the threshold. The results are consistent.

In addition, we adopt the rank-rank measure to the robustness of our results. Solon (1999) first establish. Chetty, et al. (2014, a,b) and Fan, et al. (2018) also adopt it as the main measure of mobility. Following Fan, et al. (2018), we specify the rank-rank regression as below:

$$\mathrm{rank}_i^{son} = \gamma_0 + \gamma_1 * \mathrm{rank}_i^{parents} + \mathrm{E} * \mathrm{Z}_i + \xi_i$$

where rank_i^{son} is the son's income percentile rank for parent-son pair *i* in the national distribution of income for the sons' generation*;* $\mathrm{rank}_i^{parents}$ is the parents' income percentile in the national distribution of income for parents' generation. γ_1 captures how much the sons' rank is associated with parents' rank. Bigger γ_1 implies lower intergenerational mobility in income ranks. This is a measure of relative mobility. Z_i include son's age, age squared, marital status, minority indicator, father's age, age squared and education attainment. These variables are used to control for the potential life-cycle bias and other biases caused by missing variables.

Here, the coefficient of the interaction term captures the impact of tariff change on intergenerational mobility measured by the rank-rank slope. The results are reported in columns (1)-(3) in Table 10. In column (1), we run the standard rank-rank regression of son's income percentile on fathers' income percentile controlling for son's age, age squared, marital status, minority indicator, father's age and age squared, education attainment. The rank-rank slope is 0.532 which implies that for families with father's income percentile is higher by one percentile, son's income percentile is on average 0.532 percentile higher. The magnitude of this slope is highly consistent with those estimated by Fan, Yi and Zhang (2018) using an alternative dataset. In column (2), we include an interaction term of father's percentile with tariff change. The coefficient of this term captures the impact of tariff change on rank-rank slope. It can be seen that the estimate is negative and significant. This implies that a city with bigger exposure to tariff cut has higher rank-rank slope, i.e., lower relative mobility. In column (3), we further control for the interaction of father's income percentile with other city characteristics, such as population, share of illiterate and share of agriculture employment. The result does not change much.

Furthermore, we use the IGE measure as widely adopted in the mobility literature. The results are reported in the last two columns of Table 10. Column (4) shows the standard log-log regression of son's income on father income under similar controls. The coefficient on father's income measures approximately how much son's income would increase if his parents' income were 1% higher. Our estimate is 0.591 which is also in line with those in the literature. The last column include the interaction term of tariff change. The coefficient is negative and highly significant. This implies that sons' income is more associated with parents' income in cities with higher tariff cuts.

In addition to the measures reported here, we also tried using alternative thresholds for column (1)-(3). The result continue to hold. Moreover, we tried running city-level regressions with province fixed effects. The results show similar pictures. In a short summary, we show in this subsection that our main finding is highly robust to the alternative measures of mobility.

4.3 Alternative Estimates

In the above analysis, we assume that the probability of intergenerational movement across percentile bins follows is determined by a probit model. Here, we relax this assumption and consider logit model and OLS estimate. The results are reported in Table 11. It can be seen that both models lead to similar sign pattern of the estimates. Importantly, the average marginal effects of tariff change on mobility are highly consistent too across the three models, ranging from 0.22 to 0.31 for overall mobility, 0.17 to 0.21 for upward mobility, 0.34 to 0.42 for downward mobility. This result further confirm that our main finding is robust to the choice of model.

Finally, we take other forms of liberalization into consideration. The commitments made by China in the WTO negotiations include other reforms besides tariff reduction. Allowing foreign investment to enter certain industries is an important aspect of these reforms. Neglecting these changes can lead to overestimate of the impact from tariff cuts on final output. In Table 12, we consider two major changes upon China's accession to WTO: FDI deregulation and input tariff reduction. It shows that neither FDI deregulation nor input tariff reduction is significantly associated with local intergenerational mobility. Importantly, the estimates of the coefficient on output tariff change does decrease slightly in magnitude, which is consistent with our conjecture on the direction of the potential bias.

In a nut shell, we provide robust and consistent evidence to support our main finding that trade liberalization has resulted in lower overall mobility. It promotes upward mobility but has a negative impact on downward mobility.

5. Conclusions

In this paper, we investigate the relationship between trade liberalization and intergenerational income mobility in China. We find that tariff cuts upon China's accession to WTO has a significant negative impact on overall mobility. In addition, we explore upward mobility and downward mobility separately. Our evidence shows that in a city with greater exposure to tariff cut, it's more likely for sons to move up along the income decile ladder. At the same time, the probability of moving into a lower decile than their parents' is also significantly lower in the cities with greater tariff cuts. Furthermore, we show that these results are robust to alternative measures of mobility and various model specifications. We also show that it is not caused by other forms of liberalization in the same period.

Furthermore, we examine the potential mechanisms and find that the direct impact of trade liberalization on education decision seems to be one channel through which tariff cut can increase upward mobility. On the one hand, we find that tariff cut can increase upward educational mobility. On the other hand, our evidence shows that the positive impact of tariff cut on upward mobility only exists for those parent-son pairs with upward educational mobility. These two aspects combined together imply that educational mobility is one explanation for our finding on upward mobility. In addition, upward educational mobility is one potential reason for the negative impact of tariff cut on downward mobility at the same time.

Our result also indicates that there is potential spillover effects of manufacturing tariff reduction on rural regions. Although trade liberalization can increase the competition in manufacturing and thus make it more difficult, especially for those from the low-income families, to move up along the income ladders, it seems that tariff cuts have created more equal opportunities for those living in the rural regions. There are two potential explanations for this: first, after China's accession to WTO, manufacturing has kept expanding extremely fast. This means that a large number of rural residents moved into manufacturing in this process. Tougher competition can lead to more innovation which in fact can actually generate positive spillover effects on productivity in rural areas, from which everyone can benefit regardless of family background. Second, the development of manufacturing is accompanied by higher demand on service provided by rural area, such as sightseeing and family restaurants. This creates more job opportunities for rural residents which can be more equally spread.

There are many other interesting topics along this line of research, such as how trade liberalization affects the segregation pattern and thus mobility, how trade liberalization affects educational mobility through income inequality. We leave these questions for further research when more data is available.

References:

Ahsana, R. N., Chatterjeeb A. 2017. Trade liberalization and intergenerational occupational mobility in urban India. Journal of International Economics 109, 138–152.

Anand R., Mishra S., Peiris S. J. (2013). Inclusive Growth: Measurement and Determinants. IMF Working Paper WP/13/135.

Autor, D. H., Dorn, D., Hanson, G. H. 2013. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. American Economic Review 103, 2121-68.

Autor D. H., Handel M. J. 2013. Putting Tasks to the Test: Human Capital, Job Tasks, and Wages, Journal of Labor Economics, University of Chicago Press, vol. 31(S1), pages S59 - S96.

Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang, "WTO Accession and Performance of Chinese Manufacturing Firms," American Economic Review, September 2017, 107 (9), 2784–2820

Chan, K. W., & Zhang, L. (1999). The Hukou System and Rural-Urban Migration: Processes and Changes, The China Quarterly, 160(1), pp. 818-855.

Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. The American Economic Review, 106(4), 855 – 902.

Chetty, R., Hendren, N., Kline, P., & Saez, E. 2014a. Where is the land of opportunity? The Geography of intergenerational mobility in the United States Quarterly Journal of Economics, 129(4), 1553 – 1623.

Chetty, R., Hendren, N., Kline, P., Saez, E., & Turner, N. (2014b). Is the United States still a land of opportunity? Recent trends in intergenerational mobility, NBER Working Paper No. 19844.

Clark, Gregory (2014). The son also rises: Surnames and the history of social mobility. Princeton, NJ: Princeton University Press.

Corak, Miles (2013). Income inequality, equality of opportunity, and intergenerational mobility. The Journal of Economic Perspectives, 27(3), 79 – 102.

Dai, M., Huang, W., Zhang, Y. 2018. How Do Households Adjust to Trade Liberalization: Evidence from China's WTO Accession. Memo.

Dix-Carneiro, Rafael and Brian K Kovak, "Trade Liberalization and the Skill Premium: A Local Labor Markets Approach," American Economic Review, 2015, 105 (5), 551–557

Fan, Y., Yi, J., & Zhang, J. 2015. The great Gatsby curve in China: Cross-sectional inequality and intergenerational Mobility, memo.

Feenstra, Robert C., Hong Ma, Yuan Xu, 2017. US Exports and Employment. NBER Working Paper No. 24056.

Francesconi, M., Nicoletti, C., 2006. Intergenerational mobility and sample selection in short panels. J. Appl. Econ. 21 (8), 1265 – 1293.

Goldberg, P. K., Pavcnik, N. (2007). Distributional Effects of Globalization in Developing Countries. Journal of Economic Literature 45 (1), 39 – 82.

Gong, H., Leigh, A., & Meng, X. (2012). Intergenerational income mobility in urban China. Review of Income and Wealth, 58(3), 481 – 503.

Haider, S., Solon, G., 2006. Life-cycle variation in the association between current and lifetime earnings. Am. Econ. Rev. 96 (4), 1308 – 1320.

Hilger, Nathaniel G. 2015. The Great Escape: Intergenerational Mobility since 1940. NBER Working Paper No. 21217.

Kovak, Brian K, "Regional Effects of Trade Reform: What is the Correct Measure of Liberalization?," American Economic Review, 2013, 103 (5), 1960–1976

Liu, Z. Q. (2005). Institution and inequality: the Hukou System in China. Journal of Comparative Economics 33(1): 133–157.

Sheng, L., Yang, D. T., 2016. The Ownership Structure of Offshoring and Wage Inequality: Theory and Evidence from China, memo.

Solon, Gary. 1999. Intergenerational Mobility in the Labor Market." in O. Ashenfelter and D. Card, eds., Handbook of Labor Economics, Vol. 3, Elsevier, pp. 1761-1800.

Yu, Miaojie, 2015. Processing Trade, Tariff Reductions and Firm Productivity: Evidence from Chinese Firms, Economic Journal, 125 (585), 943–988.

Zhang, W., Li, X., & Hongyan, C. (2005). China's inter-census survey in 2005. Beijing: NBS.

Zhang, L., & Zhao, X. B. (1998). Re-examining China's "urban" concept and level of urbanization. The China Quarterly, 154, 330 – 381.

Table 1: Summary statistics of main variables

	(1)	(2)
Sample	Coliving sample	Full sample
Age	24.47	28.40
	[5.16]	[5.28]
Married	0.33	0.66
	[0.47]	[0.52]
Race	0.85	0.86
	[0.36]	[0.34]
Income	588.75	756.38
	[616.38]	[945.28]
Education	3.10	3.14
	[0.85]	[0.96]
Household size	4.56	3.68
	[1.48]	[1.48]
Observations	65483	176631

Notes: this table reports the mean and standard deviation (in parenthesis) of the sons' characteristics considered in our analysis. The second column only includes the sons co-residing with their fathers, and the third column include all working-age male either living with their fathers or as their own household leaders. There are seven categories of education: (1) illiterate, (2) primary school, (3) secondary school, (4) high school, (5) professional college, (6) college, (7) graduate or above.

Table 2: Trade liberalization and intergenerational income mobility

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Var.	Mobility		Upward mobility		Downward mobility	
Tariff change	0.841***	0.730***	-0.463**	-0.487**	1.154***	1.089***
	[0.304]	[0.271]	[0.203]	[0.224]	[0.332]	[0.282]
Age	-0.113***	-0.114***	0.150***	0.156***	-0.175***	-0.180***
	[0.013]	[0.014]	[0.014]	[0.015]	[0.014]	[0.015]
Age squared	0.002***	0.002***	-0.002***	-0.002***	0.003***	0.003***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Household size	-0.002	0.001	-0.010	-0.007	0.008	0.008
	[0.006]	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]
Rate of illiterate		-0.746***		-0.512		-0.300
		[0.283]		[0.351]		[0.265]
Ln(population)		-0.039		-0.032		-0.003
		[0.027]		[0.028]		[0.026]
Ag share		0.185		-0.703***		0.804***
		[0.141]		[0.147]		[0.142]
Constant	2.551***	3.614***	-4.602***	-3.281***	2.791***	2.645***
	[0.359]	[0.606]	[0.349]	[0.581]	[0.352]	[0.585]
Prov FE	Yes	Yes	Yes	Yes	Yes	Yes
N	68,734	62,872	68,734	62,872	68,734	62,872
Pseudo R-squared	0.029	0.031	0.097	0.099	0.085	0.088

Notes: the dependent variable in columns (1) and (2) is an indicator equal to one if a son and his father belong to the same income decile and zero otherwise. The dependent variable in columns (3) and (4) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. The dependent variable in columns (5) and (6) is an indicator equal to one if a son's income decile is lower than his father's income decile and zero otherwise. Tariff change is the output tariff change between 2003 and 2000 at city level as constructed in the text. Rate of illiterate, ln(population) and the share of agricultural employment (Ag share) are constructed from the 2000 census. All regressions include controls for household size, the son's marital status, race, the father's age, age squared and indicators of father's education attainment. Province fixed effects are also included in all regressions. The standard errors are robust and clustered at the city level.

Table 3: Heterogeneity by family background

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Var.	Upward mobility	Downward mobility	Upward mobility	Downward mobility	Upward mobility	Downward mobility
Panel A: By family income						
Sample	1st quartile family		2nd quartile family		3th quartile family	
Tariff change	-0.892*	1.157**	-1.990***	2.031***	-1.485***	1.398***
	[0.463]	[0.502]	[0.433]	[0.481]	[0.323]	[0.320]
Constant	-2.302**	1.595**	-2.884***	2.677***	-5.639***	5.121***
	[0.896]	[0.757]	[0.937]	[0.867]	[1.045]	[1.043]
N	13,835	13,835	15,977	15,977	16,746	16,746
Pseudo R-squared	0.196	0.107	0.140	0.132	0.136	0.122
Panel B: By father's education						
Sample	below middle school		middle school		above middle school	
Tariff change	-0.500*	0.744***	-0.365*	0.731***	0.011	0.980**
	[0.259]	[0.292]	[0.212]	[0.204]	[0.243]	[0.487]
Constant	-2.711***	3.107***	-4.670***	4.389***	-1.856	0.520
	[0.602]	[0.600]	[1.064]	[1.042]	[1.746]	[1.695]
N	31,033	31,033	24,862	24,862	6,946	6,974
Pseudo R-squared	0.104	0.079	0.089	0.083	0.067	0.079

Notes: the dependent variable in columns (1), (3) and (5) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. The dependent variable in columns (2), (4) and (6) is an indicator equal to one if a son's income decile is lower than his father's income decile and zero otherwise. Tariff change is the output tariff change between 2004 and 2001 at city level as constructed in the text. In Panel A, we restrict the sample to sons from families with income in the first quartile for columns (1) and (2), the second quartile for columns (3) and (4), and the third quartile for the last two columns. In Panel B, we restrict the sample to sons with father's education below middle school for columns (1) and (2), equal to middle school for columns (3) and (4), and above middle school for the last two columns. City level controls include rate of illiterate, ln(population) and the share of agricultural employment which are constructed from the 2000 census. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. Province fixed effects are also included in all regressions. The standard errors are robust and clustered at the city level.

Table 4: Heterogeneity by father's Hukou status and occupation

	(1)	(2)	(3)	(4)
Dependent Var.	Upward mobility	Downward mobility	Upward mobility	Downward mobility
	Hukou status		Occupation	
Tariff change	-0.461**	0.728***	-0.840***	0.921***
	[0.210]	[0.231]	[0.214]	[0.269]
Tariff change * urban hukou	0.699**	0.503***		
	[0.301]	[0.162]		
Tariff change * non-agricultural occup.			1.120***	-0.319*
			[0.232]	[0.187]
constant	-3.467***	3.343***	-3.498***	3.433***
	[0.547]	[0.549]	[0.547]	[0.551]
N	62,834	62,834	62,834	62,834
Pseudo R-squared	0.099	0.088	0.111	0.099

Notes: the dependent variable in columns (1) and (3) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. The dependent variable in columns (2) and (4) is an indicator equal to one if a son's income decile is lower than his father's income decile and zero otherwise. Tariff change is the output tariff change between 2003 and 2000 at city level as constructed in the text. Urban hukou is an indicator equal to one if the father's hukou status is urban and zero otherwise. Non-agricultural occup. is an indicator equal to one if the father's occupation is not related to agriculture, forestry, animal husbandry or fishing production activities, and zero otherwise. City level controls include rate of illiterate, ln(population) and the share of agricultural employment which are constructed from the 2000 census. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. Province fixed effects are also included in all regressions. The standard errors are robust and clustered at the city level.

Table 5: Heterogeneity by industry

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Var.	Upward mobility	Downward mobility	Upward mobility	Downward mobility	Upward mobility	Downward mobility
Tariff change	-0.468**	0.865***	-0.532***	0.861***	-0.574***	0.815***
	[0.206]	[0.213]	[0.190]	[0.230]	[0.185]	[0.204]
Tariff change * Manufacturing	-1.367***	1.343***			-1.440***	1.221***
	[0.243]	[0.185]			[0.211]	[0.188]
Manufacturing	0.739***	-0.685***			0.894***	-0.798***
	[0.041]	[0.035]			[0.039]	[0.034]
Tariff change * Service			-1.237***	0.779***	-1.192***	0.673***
			[0.251]	[0.164]	[0.217]	[0.216]
Service			0.618***	-0.547***	0.779***	-0.671***
			[0.036]	[0.028]	[0.035]	[0.028]
Constant	-3.896***	3.692***	-3.493***	3.398***	-3.911***	3.696***
	[0.533]	[0.547]	[0.543]	[0.554]	[0.536]	[0.553]
N	62,834	62,834	62,834	62,834	62,834	62,834
Pseudo R-squared	0.100	0.088	0.099	0.088	0.103	0.091

Notes: the dependent variable in columns (1), (3) and (5) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. The dependent variable in columns (2), (4) and (6) is an indicator equal to one if a son's income decile is lower than his father's income decile and zero otherwise. Tariff change is the output tariff change between 2003 and 2000 at city level as constructed in the text. Manufacturing is an indicator equal to one if the son works in manufacturing industry and zero otherwise. Service is an indicator equal to one if the son works in manufacturing industry and zero otherwise. City level controls include rate of illiterate, ln(population) and the share of agricultural employment which are constructed from the 2000 census. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. Province fixed effects are also included in all regressions. The standard errors are robust and clustered at the city level.

Table 6: Heterogeneity by cohort

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent var.	Mobility	Upward mobility	Downward mobility	Mobility	Upward mobility	Downward mobility
Sample	Early cohort: son's age>25			Late cohort: son's age<26		
Tariff change	0.788**	-0.568**	1.409***	0.611**	-0.420*	0.841***
	[0.322]	[0.283]	[0.281]	[0.256]	[0.236]	[0.322]
Constant	-0.981	-3.616**	0.283	3.557***	-3.549***	2.391***
	[1.749]	[1.758]	[1.745]	[1.003]	[0.931]	[0.923]
N	19,388	19,388	19,386	39,583	39,579	39,583
Pseudo R-squared	0.030	0.059	0.050	0.032	0.069	0.062

Notes: the dependent variable in columns (1) and (4) is an indicator equal to one if a son's income decile is different from his father's income decile and zero otherwise. The dependent variable in columns (2) and (5) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. The dependent variable in columns (3) and (6) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. Tariff change is the output tariff change between 2003 and 2000 at city level as constructed in the text. We restrict the sample to sons with age above 25 for columns (1) to (3), and below or equal to 25 for columns (4) to (6). City level controls include rate of illiterate, ln(population) and the share of agricultural employment which are constructed from the 2000 census. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. Province fixed effects are also included in all regressions. The standard errors are robust and clustered at the city level.

Table 7: Mechanism I: educational mobility

	(1)	(2)	(3)	(4)	(5)
Dependent var.	Upward education mobility		Upward income mobility	Upward income mobility	Downward income mobility
Sample	Age<24	Age<21	No edu. mobility	Upward edu. mobility	Upward edu. mobility
Tariff change	-0.320	-0.351*	-0.099	-0.450*	0.479*
	[0.218]	[0.186]	[0.264]	[0.246]	[0.270]
Constant	0.123	-1.706	-7.616***	-1.901	2.316
	[0.953]	[1.961]	[1.675]	[1.484]	[1.446]
N	34,958	19,792	16,594	14,017	14,017

Notes: the dependent variable in columns (1) and (2) is an indicator equal to one if the son's education attainment is higher than his father's and zero otherwise. The dependent variable in columns (3) and (4) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. The dependent variable in column (5) is an indicator equal to one if a son's income decile is lower than his father's income decile and zero otherwise. Column (1) restricts to the sample with son's age lower than 24, and column (2) with son's age lower than 21. Column (3) considers sample with son and father of the same education attainment, column (4) with son's education higher than father's education, and column (5) with son's education lower than father's education. There are seven categories of education: (1) illiterate, (2) primary, (3) secondary, (4) high school, (5) professional college, (6) college, (7) graduate or above. Here, we only considers two occupation categories: agricultural job and non-agriculture job. Tariff change, input tariff change and FDI regulation change are at city level as constructed in the text. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. City level controls include rate of illiterate, ln(population) and the share of agricultural employment (Ag share) which are constructed from the 2000 census. Province fixed effects are included in all regressions. The standard errors are robust and clustered at the city level.

Table 8: Mechanism II: occupational mobility

	(1)	(2)	(3)	(4)	(5)
Dependent Var.	Upward	Income mobility			
	Occup.	Upward	Downward	Upward	Downward
Sample	All	Upward occup.=1		All	All
Tariff change	-0.058***	0.030	-0.018	0.008	0.039**
	[0.017]	[0.027]	[0.027]	[0.022]	[0.017]
Occup. mobility	No	No	No	Yes	Yes
Constant	-2.526***	-2.399*	5.775***	-3.675***	3.267***
	[0.702]	[1.290]	[1.399]	[0.547]	[0.585]
N	62,862	7,976	7,976	62,872	62,872

Notes: the dependent variable in columns (1) is an indicator equal to one if the son's occupation is better than his father's and zero otherwise. The dependent variable in columns (2) and (4) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. The dependent variable in columns (3) and (5) is an indicator equal to one if a son's income decile is lower than his father's income decile and zero otherwise. Columns (2) and (3) restrict to the sample with upward occupational mobility. Columns (4) and (5) further control for son-father occupational mobility indicator. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. City level controls include rate of illiterate, ln(population) and the share of agricultural employment (Ag share) which are constructed from the 2000 census. Province fixed effects are included in all regressions. The standard errors are robust and clustered at the city level.

Table 9: Selection Bias

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent var.	Mobility	Upward mobility	Downward mobility	Mobility	Upward mobility	Downward mobility
First stage model	Probit	Probit	Probit	Logit	Logit	Logit
Tariff change	0.411	-0.811***	1.354***	0.578**	-0.630***	1.240***
	[0.302]	[0.272]	[0.361]	[0.294]	[0.227]	[0.301]
Constant	4.854***	-1.911**	2.286***	4.734***	-1.937**	2.298***
	[1.156]	[0.863]	[0.789]	[1.024]	[0.766]	[0.779]
N	62,831	62,831	62,831	62,831	62,831	62,831

Notes: the dependent variable in columns (1) and (4) is an indicator equal to one if a son and his father belong to the same income decile and zero otherwise. The dependent variable in columns (2) and (5) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. The dependent variable in columns (3) and (6) is an indicator equal to one if a son's income decile is lower than his father's income decile and zero otherwise. The first three columns use weights estimated from probit model at the first stage, and the last three columns use weights estimated from logit model. Tariff change is the output tariff change between 2003 and 2000 at city level as constructed in the text. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. City level controls include rate of illiterate, ln(population) and the share of agricultural employment (Ag share) which are constructed from the 2000 census. Province fixed effects are also included in all regressions. The standard errors are robust and clustered at the city level.

Table 10: Alternative measures of mobility

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent var.	Mobility	Upward mobility	Downward mobility	Mobility	Upward mobility	Downward mobility
Mobility measure	By 10 percentiles			By 15 percentiles		
Tariff Change	0.506**	-0.451**	1.080***	0.371	-0.369*	0.954***
	[0.223]	[0.213]	[0.278]	[0.233]	[0.206]	[0.321]
Constant	4.346***	-2.933***	2.823***	4.241***	-2.623***	2.199***
	[0.580]	[0.631]	[0.609]	[0.570]	[0.642]	[0.547]
N	62,834	62,824	62,834	62,834	62,824	62,834

Notes: the dependent variable in columns (1) is an indicator equal to one if a son's income percentile is higher or lower than his father's income percentile by 10, and zero otherwise. The dependent variable in columns (2) is an indicator equal to one if a son's income percentile is higher than his father's income percentile by 10 and zero otherwise. The dependent variable in columns (3) is an indicator equal to one if a son's income percentile is lower than his father's income percentile by 10 and zero otherwise. The dependent variable in columns (4), (5) and (6) is a similar indicator with the difference between son's income and father's income percentile equal or bigger than 15. Tariff change is the output tariff change between 2003 and 2000 at city level as constructed in the text. City level controls include rate of illiterate, ln(population) and the share of agricultural employment (Ag share) which are constructed from the 2000 census. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. Province fixed effects are also included in all regressions. The standard errors are robust and clustered at the city level.

Table 11: Rank-rank mobility and IGE

	(1)	(2)	(3)	(4)	(5)
Dependent var.	Son's income percentile			Ln(son's income)	
Family income percentile	0.532***	0.522***	0.902***		
	[0.010]	[0.010]	[0.134]		
Family income percentile * tariff change		-1.700*	-1.902**		
		[0.890]	[0.811]		
Ln(family income)				0.591***	0.735***
				[0.121]	[0.413]
Ln(family income) * tariff change					-0.71***
					[0.22]
Constant	-0.492***	-0.510***	-0.542***	0.582***	0.585***
	[0.054]	[0.056]	[0.059]	[0.151]	[0.155]
Other interaction terms	No	No	Yes	No	Yes
R^2	0.51	0.51	0.50	0.55	0.54
N	68,693	68,693	62,834	68,693	62,834

Notes: the dependent variable is the percentile of son's income in columns (1)-(3), and the dependent variable is the log of son's income in columns (4)-(5). Tariff change is the output tariff change between 2003 and 2000 at city level as constructed in the text. Columns (3) and (5) also includes interaction terms of family income percentile with city level controls, such as rate of illiterate, ln(population) and the share of agricultural employment which are constructed from the 2000 census. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. Province fixed effects are also included in all regressions. The standard errors are robust and clustered at the city level.

Table 12: OLS and Logit model

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent var.	Mobility	Upward mobility	Downward mobility	Mobility	Upward mobility	Downward mobility
Model	logit	logit	logit	OLS	OLS	OLS
Tariff change	1.278***	-0.841**	1.792***	0.221***	-0.171**	0.391***
	[0.461]	[0.353]	[0.461]	[0.079]	[0.074]	[0.103]
Constant	6.088***	-6.091***	4.264***	1.566***	-0.076	1.643***
	[1.066]	[0.988]	[0.970]	[0.179]	[0.179]	[0.210]
N	62,834	62,834	62,834	62,834	62,834	62,834
R^2	-	-	-	0.03	0.12	0.12

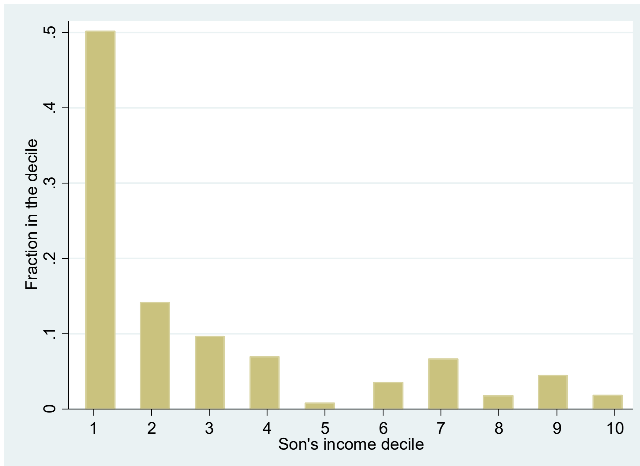
Notes: the dependent variable in columns (1) and (4) is an indicator equal to one if a son and his father belong to the same income decile and zero otherwise. The dependent variable in columns (2) and (5) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. The dependent variable in columns (3) and (6) is an indicator equal to one if a son's income decile is lower than his father's income decile and zero otherwise. The first three columns report the estimates from logit model, and the last three columns are OLS estimates. Tariff change is the output tariff change between 2003 and 2000 at city level as constructed in the text. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. City level controls include rate of illiterate, ln(population) and the share of agricultural employment (Ag share) which are constructed from the 2000 census. Province fixed effects are also included in all regressions. The standard errors are robust and clustered at the city level.

Table 13: Other forms of liberalization

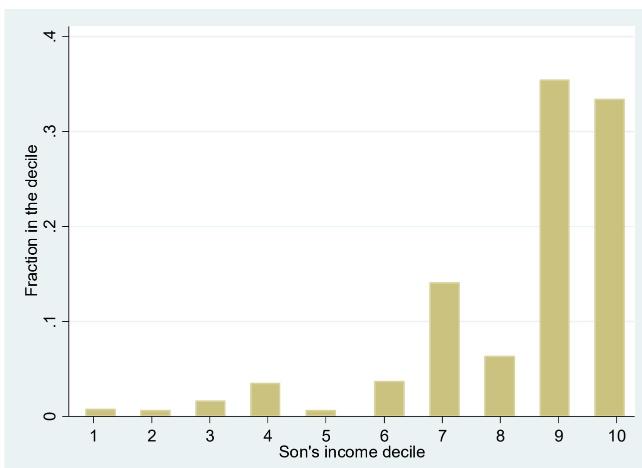
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent var.	Mobility	Upward mobility	Downward mobility	Mobility	Upward mobility	Downward mobility
Tariff change	0.658**	-0.537**	1.068***	0.657**	-0.507**	1.040***
	[0.281]	[0.241]	[0.323]	[0.271]	[0.238]	[0.315]
Input tariff change	1.641	1.726	-0.212	2.078	3.680	-1.636
	[3.056]	[3.001]	[3.540]	[2.218]	[3.233]	[3.730]
FDI regulation change				-0.045	-0.064	0.032
				[0.060]	[0.054]	[0.077]
Constant	3.453***	-2.914***	2.275***	3.318***	-3.154***	2.405***
	[0.636]	[0.600]	[0.620]	[0.678]	[0.611]	[0.661]
N	62,834	62,834	62,834	62,834	62,834	62,834

Notes: the dependent variable in columns (1) and (4) is an indicator equal to one if a son and his father belong to the same income decile and zero otherwise. The dependent variable in columns (2) and (5) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. The dependent variable in columns (3) and (6) is an indicator equal to one if a son's income decile is lower than his father's income decile and zero otherwise. Tariff change, input tariff change and FDI regulation change are at city level as constructed in the text. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. City level controls include rate of illiterate, ln(population) and the share of agricultural employment (Ag share) which are constructed from the 2000 census. Province fixed effects are included in all regressions. The standard errors are robust and clustered at the city level.

Figure 1: Son’s income deciles by family background

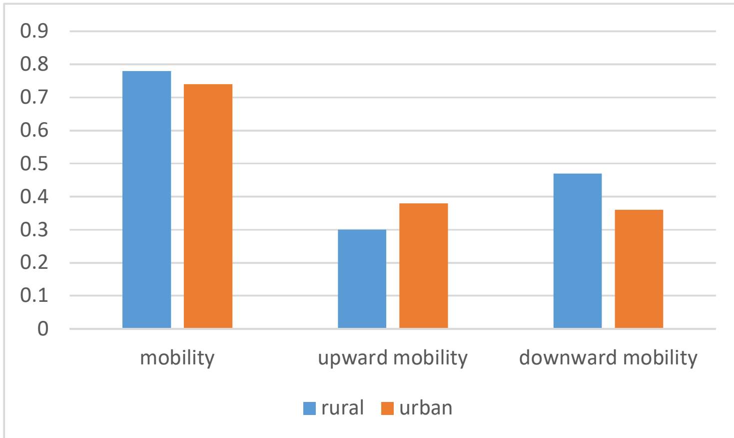


a. Son’s income deciles born to bottom decile family



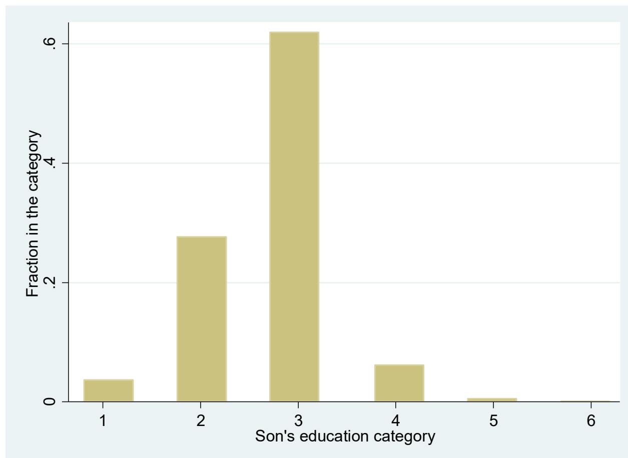
b. Son’s income deciles born to top decile family

Figure 2: Average Mobility by Hukou Status

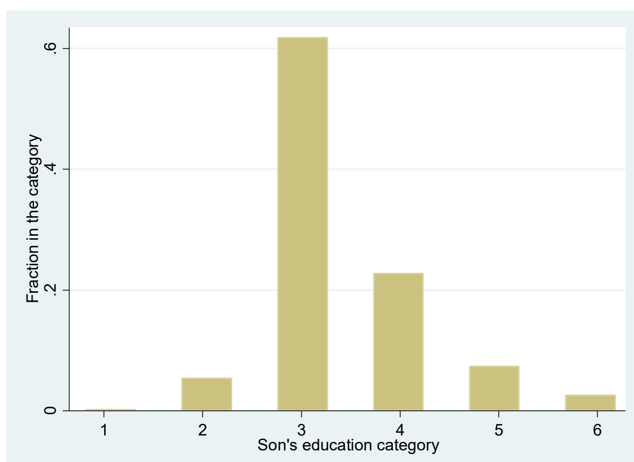


Notes: mobility is an indicator equal to one if a son and his father belong to the same income decile and zero otherwise. Upward mobility is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. Downward mobility is an indicator equal to one if a son's income decile is lower than his father's income decile and zero otherwise. The sample is separated into rural and urban based on son's hukou status.

Figure 3: Son's education attainments by father's education



a. Son's education attainments with fathers' education as primary school or below



b. Son's education attainments with fathers' education as high school or above

Appendix

Table A1: Mechanism - occupational mobility

	(1)	(2)	(3)	(4)	(5)
Dependent var.	Upward occup. mobility	Upward mobility	Upward mobility	Downward mobility	Downward mobility
Sample	All	No upward occup. mobility	Upward occup. mobility	No upward occup. mobility	Upward occup. mobility
Tariff change	-0.356	-0.470*	0.573**	0.862***	-0.420*
	[0.263]	[0.251]	[0.241]	[0.218]	[0.218]
Constant	-2.334***	-3.745***	-2.564*	3.306***	3.558**
	[0.714]	[0.578]	[1.321]	[0.590]	[1.404]
N	62,577	52,477	10,357	52,477	10,357

Notes: the dependent variable in columns (1) and (2) is an indicator equal to one if the son's education attainment is higher than his father's and zero otherwise. The dependent variable in columns (3) to (6) is an indicator equal to one if a son's income decile is higher than his father's income decile and zero otherwise. Column (1) restricts to the sample with son's age lower than 24, and column (2) with son's age lower than 21. Column (3) considers sample with son and father of the same education attainment, and column (4) with son's education higher than father's education. There are seven categories of education: (1) illiterate, (2) primary, (3) secondary, (4) high school, (5) professional college, (6) college, (7) graduate or above. Column (5) restricts the sample with sons and fathers of the same occupation. Column (6) considers the sample with fathers as agricultural occupation and sons with non-agriculture occupation. Here, we only considers two occupation categories: agricultural job and non-agriculture job. Tariff change, input tariff change and FDI regulation change are at city level as constructed in the text. All regressions include controls for household size, the son's age, age squared, marital status, race, the father's age, age squared and indicators of father's education attainment. City pre-WTO characteristics, such as rate of illiterate, ln(population) and the share of agricultural employment, constructed from the 2000 census, are also controlled for. Province fixed effects are included in all regressions. The standard errors are robust and clustered at the city level.