

FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

## **Persistence of Regional Inequality in China**

Christopher Candelaria,  
Stanford University

Mary Daly,  
Federal Reserve Bank of San Francisco

Galina Hale,  
Federal Reserve Bank of San Francisco

March 2013

Working Paper 2013-06

<http://www.frbsf.org/publications/economics/papers/2013/wp2013-06.pdf>

The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of New York, the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System. This paper was produced under the auspices of the Center for Pacific Basin Studies within the Economic Research Department of the Federal Reserve Bank of San Francisco.

# Persistence of Regional Inequality in China

Christopher Candelaria

Stanford University

Mary Daly

Federal Reserve Bank of San Francisco

Galina Hale\*

Federal Reserve Bank of San Francisco

3/25/13

## Abstract

Regional inequality in China appears to be persistent and even growing in the last two decades. We study potential explanations for this phenomenon. After making adjustments for the difference in the cost of living across provinces, we find that some of the inequality in real wages could be attributed to differences in quality of labor, industry composition, labor supply elasticities, and geographical location of provinces. These factors, taken together, explain about half of the cross-province real wage difference. Interestingly, we find that inter-province redistribution did not help offset regional inequality during our sample period. We also demonstrate that inter-province migration, while driven in part by levels and changes in wage differences across provinces, does not offset these differences. These results imply that cross-province labor market mobility in China is still limited, which contributes to the persistence of cross-province wage differences.

JEL classification: I3, J4, O5, R1

Key words: inequality, China

---

\*Contact: Federal Reserve Bank of San Francisco, 101 Market St., MS1130, San Francisco, CA 94105. galina.b.hale@sf.frb.org. We thank Shang-Jin Wei for helpful suggestions and Akshay Rao for outstanding research assistance. All errors are our own. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or any other person associated with the Federal Reserve System.

# 1 Introduction

Persistent inequalities across regions are a feature of many developed and developing nations. This fact conflicts with standard neoclassical theory which suggests that in a well-functioning economy regional inequalities should be eliminated through factor mobility, trade, or arbitrage. Although significant research literature has been devoted to resolving the conflict between theory and fact, no real consensus has been formed (Magrini, 2007). In this paper, we analyze the patterns and causes of persistent regional inequality in China.

China presents a unique and important opportunity to study regional inequality. In terms of income, China is a very unequal country.<sup>1</sup> According to World Bank estimates, China's income Gini coefficient far exceeds that of South Asian countries such as India, Bangladesh, and Pakistan. Income differences in China are large both inter- and intra-regionally, with inter-regional differences increasing over time (Fujita and Hu, 2001; Kanbur and Zhang, 1999). China is also a country in transition, altering its institutions and economy at a rapid pace but with considerable variation across provinces.<sup>2</sup>

Although most of the research on Chinese inequality has focused on the urban-rural income gap, recent studies have examined regional inequality. For example, Kanbur and Zhang (2005) construct long time series of regional inequality and show that it is explained, in different periods of time, by the share of heavy industry (which we also find), the degree of decentralization, and the degree of openness. In a recent paper, Wan, Lu, and Chen (2007) show that increasing globalization, uneven domestic capital accumulation, and privatization contribute to regional inequality in China, while the effects of location, urbanization and the dependency ratio have been declining. Yao and Zhang (2001) demonstrate that there is divergence between groups (or clubs) of Chinese provinces in terms of real per capita GDP. Meng, Gregory, and Want (2005) argue that discontinued provision of free education, housing, and medical care, as well as increased uncertainty increased the incidence of urban poverty when measured in terms of expenditure. Xia, Song, Shi, and Appleton (2013) find that reforms of the state sector have contributed to urban wage inequality across regions. In our

---

<sup>1</sup>Persistent and even growing cross-region inequality in China during the reform period has been documented in the past (Jian, Sachs, and Warner, 1996; Chen and Fleisher, 1996).

<sup>2</sup>For the survey of regional inequalities in transition economies, see Huber (2006).

paper we take it a step further to understand the relationship between the persistence of regional inequality and labor movement.

Some researchers argue that persistent inequalities in income are the result of unmeasured offsetting factors that work to equalize well-being across regional agents (Rice and Venables, 2003). In our study we find, not surprisingly, that cost-of-living differences are, indeed, an important offsetting factor. Thus, we adjust average nominal wage in each province by province-specific CPI to obtain average real wage. We find, however, that even in terms of real wage cross-province inequality in China is large and persistent.

We analyze factors that might explain this persistent inequality, such as the quality of labor, which we proxy by measures of education level in each province; labor productivity, which we proxy by the industrial composition in each province; labor supply elasticity, which we measure as a share of agricultural population in the province; access to export markets, which we measure as the size and presence of large sea ports in the province;<sup>3</sup> cross-province government transfers. We find that, with the exception of government transfers, all these factors help explain a substantial portion of wage differentials. Nevertheless, about half of the variance in real wages across provinces remains after controlling for these factors.

During our sample period from 1993 to 2011, the restrictions on migration resulting from the system of permanent registration (*hukou*) are being gradually relaxed, if not *de jure*, then *de facto* (Bao, Bodvarsson, Hou, and Zhao, 2011).<sup>4</sup> It is, therefore, natural to expect that such persistent unexplained difference in real wages across provinces would generate cross-province migration. Using the summary data from 2000 and 2010 population censuses we find that cross-province migration in the second half of the 1990s and the second half of 2000s is indeed partly driven by real wage differences and by differences in growth rates of real wages.

Even though cross-province migration is partly driven by wage differences, it still appears quite limited. In particular, migration flows do not seem to reduce regional inequality, once we control for factors that we find to affect wage inequality. This suggests that the migration is still quite

---

<sup>3</sup>Much of the literature emphasizes inland-coastal differences. However, as demonstrated in Candelaria, Daly, and Hale (2009), it is the presence of a large port that makes all the difference, not simply the presence of the coastline.

<sup>4</sup>We briefly describe China's migration policies in Section 2.

constrained by the remaining barriers. This finding is consistent with the literature. Chan and Buckingham (2008) argue that recent reforms of the hukou system did not make migration any easier, but may in fact further reduced mobility, especially between rural and urban population. Whalley and Zhang (2007) calibrate the effect of removal of the hukou system and show that the resulting reduction in regional inequality is sizable, indicating that such removal has not occurred yet. More recently, Bao, Bodvarsson, Hou, and Zhao (2011) show empirically that migration is still very sensitive to hukou constraints, even in 2000-05, although the sensitivity has declined relative to 1995-2000.

Taken together, our results indicate that regional inequality is likely to persist in China for quite some time, since it is due to structural and long-term factors such as education, industry composition, regional labor supply, and geographical location. While further removing barriers to labor mobility may alleviate some of the regional wage differences, it is likely to be a slow process, because it needs to be accompanied by urban development.<sup>5</sup> In the meantime, cross-province income redistribution may be useful to address social and economic tensions that regional inequality can bring. Our study finds that in the period of consideration inter-province transfers did not help offset regional wage differences.

The paper is organized as follows: We begin by describing the policies that restrict labor movement in China in Section 2. We describe our data and recent trends in Section 3. In Section 4 we present the results of our empirical analysis. We conclude in Section 5.

## 2 Migration policy in modern China

In 1958, Mao Zedong set up an hereditary residency permit (Hukou) system defining where people could work. This system classifies individuals as either “rural” or “urban” workers and assigned them to a specific geographic area. Under the hukou system, one cannot acquire a legal permanent residence and the numerous community-based rights, opportunities, benefits and privileges in places other than the location assigned in his hukou. Only through proper authorization of the government can one permanently change his hukou location. For longer than one-month stay and especially

---

<sup>5</sup>Chan (2009) argues that the hukou system is a major obstacle to China’s development.

when seeking local employment, one must apply and be approved for a temporary residential permit. Violators are subject to fines, detention, and forced repatriation (partially relaxed in 2003).<sup>6</sup>

The main impact of the hukou system is on rural-to-urban mobility, which explains an extraordinary low degree of urbanization in China (Chan and Zhang, 1999). In addition to limiting the rural-to-urban migration of the workers, this system limited migration across regions: migrants would need to obtain registration in the new area before they could be allowed to work legally. Because of bureaucracy and red tape, obtaining registration could take months or years. Hence, the migration process in China was in many ways similar to the migration process across national borders.

Until 1978, the system was enforced rather strictly: police would periodically round up those without valid residence permit, place them in detention centers, and expel them from cities. After Chinese market reforms began in 1978, a private sector appeared. Unlike state-owned enterprises (SOEs), private firms frequently do not require registration from potential employees. Economic reforms created pressures to encourage migration from the interior to the coast and provided incentives for officials not to enforce regulations on migration. However, even if migrants are hired without the registration, they would not have access to social services such as child care, schools, or health care.

Forced repatriation rules were still in place after the beginning of market reforms. In fact, the 1982 “Measures of Detaining and Repatriating Floating and Begging People in the Cities” streamlined the relevant legislation. It was not until 2003, after the tragic incident in Guangzhou, where a Hubei resident was beaten to death in jail in the process of repatriation, and the outcry that followed, that the repatriation law was relaxed. “Measures on Managing and Assisting Urban Homeless Beggars without Income” adopted on June 20, 2003, established new rules governing the handling and assisting of destitute migrants. Many cities, including the most controlled Beijing municipality, decided soon after that hukou-less migrants must be dealt with more care; they are no longer automatically subject to detention, fines, or forced repatriation, unless they have become homeless, paupers, or criminals.

---

<sup>6</sup>This information and some of what follows in this section is partly quoted from Wang (2005). For more detail, see Wang (2004).

While recent reforms made it easier for migrants to survive without hukou, they did not make it easier to obtain one. Thus, many people, especially those with families, are hesitant to permanently change their location. Moreover, fees charged for temporary residency permits, while fairly affordable, add to the cost of moving. Nevertheless, according to the census data, in 2000 almost 30 percent of the population was residing in a province other than the one in which they had their hukou.

There has been more progress in reforming the hukou system at the provincial level than at the national level. Between 2001 and 2007, a number of provinces adopted various measures, mostly addressing the rural–urban migration limitations. These measures ranged from changing the administration of the hukou system to completely abolishing the differences between rural and urban hukou.<sup>7</sup>

For the purpose of our project, we have to recognize that the barriers to cross–province migration in China still exist but were weakening throughout our sample period. Moreover, we have to take into account the fact that the progress of reform was uneven across provinces.

### 3 Data

In this section, we describe our data sources for inequality, explanatory variables, and migration. We also describe recent trends and patterns in regional inequality and migration.

#### 3.1 Inequality and related data

We use national and provincial level data that are reported in the *China Statistical Yearbook*. This annual publication is compiled by the National Bureau of Statistics of China and is published by the China Statistics Press. In conjunction with these data, we also use CEIC Data’s “China Premium Database” because it provides time series of the data in the *China Statistical Yearbook*. Using these sources, we construct an annual data sample from 1993 to 2011, covering a period of fast economic growth in China. Due to data limitations, not all series have coverage for the entire

---

<sup>7</sup>See “Recent Chinese Hukou Reforms” published by the Congressional Executive Commission on China, accessed at <http://www.cecc.gov/pages/virtualAcad/Residency/hreform.php> in October 2007.

period.

Table A.1 provides definitions and Table A.2 gives summary statistics for the variables we use in this paper. Because of the high level of economic growth in China, the variables are trending for all the provinces; thus, for measures of inequality, we express the majority of our variables in terms of percentage deviation from the country average. In doing so, we obtain unit-free, common trend-free, comparable variables for all the provinces. The variables that are not expressed in percentage deviation are berth capacity, inter-provincial migration, and rural-to-urban intra-provincial migration.

In 1997 the city of Chongqing in Sichuan province was raised to a status of provincial city, resulting in the creation of a new province and also affecting all the population-based statistics for the Sichuan province. Moreover, since Chongqing was the largest and wealthiest part of Sichuan, average income in the Sichuan province was also affected through a composition change. We take two approaches to dealing with Chongqing: we either drop Sichuan and Chongqing from our sample, or we construct weighted averages of variables for Chongqing and Sichuan after 1997. Measures of inequality are weighted by population. In the rest of the paper, we present the results which treat Chongqing and Sichuan as a combined province, but our results do not change if we exclude both of them.

We omit the provinces of Hubei and Tibet from our data sample. Hubei has been excluded because of the vast relocation of residents due to construction of the Three Gorges dam beginning in December of 1994. We do not include Tibet due to a lack of a CPI index that extends back to 1993. Because of this, we cannot construct real measures of key variables that are used in our analysis. With the omission of Hubei and Tibet, and the combination of Chongqing and Sichuan into a single province, we have a data set with a total of 28 provinces.

### **3.2 Migration data**

We use migration data that is collected by the National Bureau of Statistics of China and compiled by China Data Online. The data we use come from the 2000 and 2010 population censuses, which measure migration between 1995 and 2000 (2000 census), and between 2005 and



2010 (2010 census). To our advantage, the 2000 data provide information not only on the province of origin and destination, but also the area within each province that the migrant came from and went to. This is especially useful to track patterns of migration such as the movement of people from rural to urban areas—both intra- and inter-provincially.<sup>8</sup>

In both censuses, individuals 5 years of age or older were asked if in the census year they resided in a different subcounty-level unit than that in which they lived five years prior. If an individual moved his hukou to the new location or he resided in the new location for more than six months, he was counted as a migrant. While the migration data includes in-migration to Chinese provinces from Hong Kong, Macau, and abroad, we exclude these numbers from our analysis.

### 3.3 Trends in inter-province inequality

We now turn to the description of the trends and patterns in our data. Figure 1 shows median average levels of disposable income for urban and rural households and for urban households in different subgroups of provinces. One can immediately see the growing gap between urban and rural income that was much discussed in the literature. We can also see large differences in urban income across provinces, especially between coastal and inland provinces, as has also been discussed in the literature. As Caballero, Candelaria, and Hale (2009) show, for coastal provinces, the presence of a large commercial port is important.

In this paper we leave aside the important rural-urban inequality, since it is well studied in the existing literature, and focus on inter-provincial inequality in urban income. We have some concerns about the accuracy of disposable income data at the provincial level and therefore for the rest of the paper we focus on the average wage data. Average wage data is not reported separately for rural and urban households, but the detailed description of these series (see Table A.1) suggests that most of the input into average wage comes from the urban wages. We, therefore, frame the rest of the paper as an analysis of inter-provincial inequality in urban income.

Unlike most countries, China provides province by province consumer price index (CPI) data. We take advantage of this fact to calculate real wages using individual province CPI, which allows

---

<sup>8</sup>See Fan (2005) and Lavelly (2001) for a detailed discussion of 2000 migration data.

us to account for differences in cost of living in different provinces. Figure 2 depicts the coefficient of variation of nominal and real wages across provinces as a measures of cross-province inequality. We can see that rapid increase in urban wage dispersion from 1990 to 2001 followed by a gradual decline. Figure 2 also shows that part of the measured inequality in wages disappears once we account for the cost of living. For example, in 2002 the coefficient of variation of nominal wage across provinces was 0.34, while the coefficient of variation of real wage was 0.26. Because real wage inequality and not nominal wage inequality is economically important, we continue our analysis in terms of average real wage. In terms of real wage, we also see a rapid increase in cross-province inequality from 1990 to 1999, followed by a very small decline from 2006 to 2011.

Despite some decline in cross-province inequality in recent years, it remains very large, even in real terms. At the end of our sample period, the coefficient of variation of real wage is over 0.22, which means the standard deviation of the real wage across provinces is almost a quarter of the mean real wage.

### 3.4 Patterns in inter-province migration

We now describe trends in the migration patterns in China. We focus on the migration that took place between 1995 and 2000, and between 2005 and 2010, using information from the 2000 and 2010 population censuses.

For the 2010 census data we observe total migration within and across provinces. The 2000 census data is more detailed. We have emigration (out-migration) from each province broken down into rural and urban migrants. The urban migrants are those coming from the sub-county level units of neighborhood committee of the town and street. The rural migrants are the migrating population from townships and village committees of the town. With respect to immigration (in-migration), we are able to classify migrants as entering a town, city, or county. We broadly define city and town as an urban region and county as a rural region.<sup>9</sup>

A large part of migration took place within provinces, with population moving from rural to

---

<sup>9</sup>Chan and Hu (2003) include an appendix of major points in defining urban population for the 2000 census. In this appendix they state that urban population of China consists of City and Town population.

urban areas. Thirty five percent of the intra-province migrating population moved from rural to urban areas in 1995-2000.<sup>10</sup> Inter-province migration was also important, with 76 percent of all inter-provincial migrants going into urban areas and almost none going to rural areas.

Tables A.3a and A.3b report migration for all provinces between 1995 and 2000 and between 2005 and 2010, respectively, both cross-province and rural-to-urban migration within the province, in levels and as a percent of host province population in 1997 and 2007, respectively. Not surprisingly, the provincial cities of Beijing and Shanghai, as well as the Guandong province that is adjacent to Hong Kong, attracted the most cross-province migrants as a share of their population in both time periods. We also note that coastal provinces tended to attract more cross-province migrants than inland provinces as a share of their population. Trends in intra-provincial migration are more difficult to identify. For intra-province migration, the variance in rural-to-urban intra-province migration across provinces is much smaller than that of the cross-province migration. We also note that within province migration increased substantially from 2.7 percent to 6.1 percent between the 2000 and 2010 census periods, indicating that, indeed, some relaxation of hukou controls was happening.

## 4 What explains persistent inter-province inequality in China

In this section we study explanations for persistent inter-province inequality. We begin by studying the effects of quality of labor, industry composition, government transfers, and geographical factors. We then turn to the analysis of cross-province migration patterns and test whether inter-province migration is driven, at least in part, by wage differences and whether it, in turn, helps offset some of these differences. We conduct our analysis using the average wages adjusted by province-specific CPI, which we refer to as real wages.

---

<sup>10</sup>Some of these numbers are not reflecting actual moving of the population, but the annex of rural areas by cities. For statistical purposes, however, this is identical to actual migration.

## 4.1 Reasons for real-wage inequality

We begin with a measure of the quality of human capital that may explain differences in real wages. Since we do not have a direct measure of labor quality, we use two different measures of education level in the province to proxy for the quality of labor. If observed differences in wages are fully explained by differences in labor quality, there is no reason to believe that such differences should go away. We then consider industry composition in the provinces. Because labor productivity may be higher in some industries, industry composition will affect average wage in the province. While we would expect labor to move to the industries where it is more productive, structural changes in the economy take a long time to complete and thus inequality that is due to industry composition is expected to be persistent. Next, we test whether the labor supply elasticity, as measured by the share of agricultural population in the province, explains wage differences. If migration across provinces is constrained, we would expect wages to be higher in provinces with lower share of potentially available labor that is currently employed in agriculture. We also test whether some of the wage differentials across provinces are offset by transfers from the central government. Finally, we test whether access to export markets, as measured by total commercial ship berth capacity in the province, explains wage differences. Since access to the sea cannot be changed and port capacity can only be changed slowly, real wage differences that are due to such geographical factors are likely to be persistent.

The results of this analysis are reported in Table 1. The first two columns present our results with respect to measures of education level in each province. In the first column we measure education level as a number of college graduates per capita, and in the second column we use real government expenditures on education per capita. We can see that both measures indicate that higher levels of education are associated with higher average real wages. We can also see from the adjusted  $R^2$  that the first measure explains 26 percent of standard deviation in real wages across provinces, while the second measure explains 66 percent, although there is a possibility that there is some spurious correlation in case of government expenses on education — provinces that are wealthy for whatever reason are likely to have higher wages and more expenditure on education.<sup>11</sup> For this reason, and

---

<sup>11</sup>Note the sample difference between the two columns. They are due to the fact that education expenditure data are only available starting 1999.

because of sample limitation in the case of the second measure, we will use the first measure as a control variable in the tests that follow. Moreover, the first measure seems to be more relevant, as it measures the contemporaneous level of college education in the province.

Columns (3) and (4) of Table 1 use different measures of industrial composition — employment shares of secondary and tertiary industries (primary being a base category) and the share of agricultural population. We find that higher shares of primary and tertiary industries, which tend to have larger labor productivity, are associated with higher real wage and explain about 54 percent of cross-province wage differences. Similarly, a higher share of agricultural population is associated with lower wages and explains 44 percent of cross-province wage differences.

Column (5) estimates a correlation between real wages and transfers from federal to provincial budgets. We find that there is no such correlation; that is, Chinese government is neither creating nor smoothing cross-province wage differences.

Finally, results reported in column (6) of Table 1 demonstrate that the provinces with larger commercial ports, as measured by their commercial berth capacity, tend to have higher wages. This is likely to be partly driven by export platforms in these provinces, which partly are owned by foreign companies and tend to pay higher wages, at least to skilled workers (Hale and Long, 2008). Berth capacity, however, only explains a small portion of cross-province variation in real wages.

In Column (7) we include all the above controls together, even though many of them are highly correlated.<sup>12</sup> We can see that taken together they explain 84 percent of the cross-province real wage differences if measured by the adjusted  $R^2$ . In terms of the standard deviation of the residual in the last year of the sample, it is about one third of the standard deviation of real wages across provinces in the same year. If we extract a first principle component from all the explanatory variables, and include that instead, we find that about 64 percent of cross-province wage differences are explain by this first principle component (see Column 9 of Table 1). The standard deviation of residual in the last year of the sample in this case is about half of the standard deviation of real wage in the same year.

All the regressions described above and reported in Table 1 are based on the panel data and

---

<sup>12</sup>See Appendix Table A4 for correlations.

therefore are identified based on time-series as well as cross-province differences. To separate the two effects we first include province fixed effects in our regressions with all controls and with the first principle component (Columns (8) and (10) of Table 1). We find that only education measures are significantly correlated with within variation in real wages, with fixed effects absorbing all other cross-province differences. Overall, regressions with fixed effects do not perform much better than regressions without. This indicates that, as we discussed above, the results are identified off of cross-province rather than within differences. We confirm this once again by re-estimating all the regressions in Table 1 with either a linear trend or with year fixed effects and find that the results remain almost identical.<sup>13</sup>

## 4.2 Labor mobility

Before addressing the question of whether cross-province migration offsets some of the wage differences, we have to establish that cross-province migration is at least in part driven by wage differences. To this end, in Table 2 we test whether cross-province migration into urban areas between years  $t - 5$  and  $t$  is correlated with differences in real wages across provinces in  $t - 5$ , where  $t \in \{2000, 2010\}$  is the year of the census. Similarly, in Table 3 we test whether cross-province migration between years  $t - 5$  and  $t$  is correlated with differences in real wages in  $t - 10$  and differences in growth rates of real wage between  $t - 10$  and  $t - 5$ . In both specifications, migration does not affect wage differences because it occurs *after* the real wage is measured. We find that both the level and the growth rate of real wages have positive effect on migration into urban areas of a given province from other provinces.<sup>14</sup> These results hold even when we control for the explanatory variables that we found to be important in the first part of our analysis. These results are also robust to including a different intercept for the observations resulting from the later, 2010, census period.<sup>15</sup> Not surprisingly, province fixed effects absorb the effects of the wage level, but not of

---

<sup>13</sup>The results of these regressions are not reported in interest of space. The effect of linear trend is not significant in most regressions.

<sup>14</sup>For comparison, Tables A.5. and A.6. provide the same analysis for rural to urban migration *within* provinces. These results show that for the within-province migration growth rate of wages matters more than the level.

<sup>15</sup>Even though a later period indicator does enter significantly, we chose to report benchmark regressions for consistency. In the interest of space we do not report regressions with the second period dummy, but they are available upon request.

wage growth. Importantly, two-thirds to three-quarters of variation in migration flows is explained by levels of and changes in real wages alone.

Coefficients on other explanatory variables have expected signs — higher education and larger share of tertiary industries attract more migrants, while higher share of agricultural population lower the inflow of migrants. Interestingly, share of secondary industries, while associated with higher wages (See Table 1), has a negative effect on migration. This shows that once wage differences are accounted for, secondary industries are not particularly attractive to permanent migrants.<sup>16</sup> Berth capacity does not enter significantly in these regressions.

Having established that migration is indeed driven by wage differences, we now turn to the main question of interest — does migration offset any of the wage differences? To this end, we estimate the effect of cross-province migration into urban areas between years  $t - 5$  and  $t$  on real wage in  $t + 1$ , while controlling for our explanatory variables in  $t + 1$ , where  $t \in 2000, 2010$  is a census year. The results are reported in Table 4.<sup>17</sup>

In column (1) we see that, far from offsetting wage differences, higher migration is associated with higher real wage. This effect is spurious, however. Because real wage differences across provinces are highly persistent over time, migration in  $t - 5$  to  $t$  may be endogenous even with respect to  $t + 1$  real wages. To rule out this possibility, we add real wage in  $t - 5$  as an additional control variable. As we can see from column (12), where real wage in  $t - 5$  is the only regressor, real wage is, indeed, highly persistent.<sup>18</sup>

We can see that once the real wage in  $t - 5$  control is included, migration has a much smaller positive effect on real wages (Column (2)) and no statistically significant effect in subsequent regressions that include  $t + 1$  variables that we found important in explaining real wage. Moreover, the reduction in the variance of the residual is minimal to nil (compare column (2) with column (12)). We conclude, therefore, that cross-province migration that occurred between 1995 and

---

<sup>16</sup>Note that our migration data are unlikely to include temporary migrant workers.

<sup>17</sup>For comparison, results of similar tests for rural to urban migration *within* provinces are reported in Table A.7. The results for the within-province migration are once again very similar to those discussed below for the cross-province migration.

<sup>18</sup>All the results we discuss here are not affected by the inclusion of an indicator for the later census sample. These regressions are not reported in interest of space but are available upon request.

2000 as well as between 2005 and 2010 did not produce any equalizing effects for real wages in subsequently years. This is not particularly surprising, since, as we discussed in the Introduction and in Section 2, the effects of hukou system on restricting migration are still substantial and therefore observed migration, although driven by real wage differences, is not sufficient to offset them.

## 5 Conclusion

Provincial statistics show that regional inequality in China has been persistent in the last two decades. We find that main sources of this persistence were structural and long-term factors such as labor quality, industrial composition, regional labor supply, and geographical location. We find that inter-provincial transfers did not offset wage inequality during our sample period, and neither did inter-province migration. These findings suggest that regional income inequality in China is not likely to go away in the near future.

While we believe that fully removing barriers to labor mobility will help reduce cross-province wage inequality, we are aware of the fact that urban infrastructure is unable to accommodate large inflows of new migrants. We therefore view the message of this paper as more positive than normative — given the gradual nature of structural changes and urban infrastructure development, we should expect that regional inequality in China will persist for quite some time.

A policy implication of this observation is that social tensions that arise due to such inequality may need to be addressed in the short run through redistribution. According to our analysis, such redistribution has not been present during our sample period (at least not in the measures we considered). Chinese government, however, recognizes inequality as an important problem, which is reflected in its most recent five-year plan.

## References

BAO, S., . O. B. BODVARSSON, J. W. HOU, AND Y. ZHAO (2011): “The Regulation of Migration in a Transition Economy: China’s Hukou System,” *Contemporary Economic Policy*, 29(4), 564–



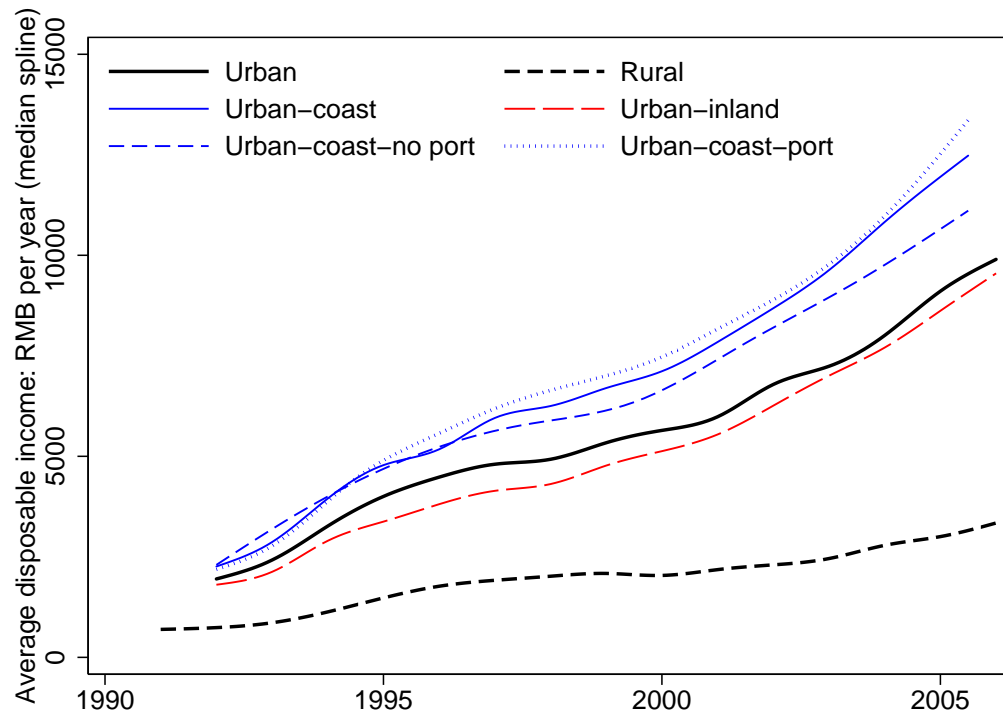
- CABALLERO, J., C. CANDELARIA, AND G. HALE (2009): "Bank Relationships and the Depth of the Current Economic Crisis," *Economic Letter* December 14, FRBSF.
- CANDELARIA, C., M. DALY, AND G. HALE (2009): "Interprovincial Inequality in China," *FRBSF Economic Letter*, 2009-13.
- CHAN, K. W. (2009): "The Chinese Hukou System at 50," *Eurasian Geography and Economics*, 50(2), 197–221.
- CHAN, K. W., AND W. BUCKINGHAM (2008): "Is China Abolishing the Hukou System?," *The China Quarterly*, 196, 582–606.
- CHAN, K. W., AND Y. HU (2003): "Urbanization in China in the 1990s: New Definition, Different Series, and Revised Trends," *The China Review*, 3(2), 49–71.
- CHAN, K. W., AND L. ZHANG (1999): "The Hukou System and Rural-Urban Migration in China: Processes and Changes," *The China Quarterly*, 160, 818–855.
- CHEN, J., AND B. FLEISHER (1996): "Regional Inequality and Economic Growth in China," *Journal of Comparative Economics*, 22, 141–164.
- FAN, C. C. (2005): "Interprovincial Migration, Population Redistribution, and Regional Development in China: 1990 and 2000 Census Comparisons," *The Professional Geographer*, 57(2), 295–311.
- FUJITA, M., AND D. HU (2001): "Regional disparity in China 1985–1994: The effects of globalization and economic liberalization," *The Annals of Regional Science*, 35(1), 3–37.
- HALE, G., AND C. LONG (2008): "Did Foreign Direct Investment Put an Upward Pressure on Wages in China?," *IMF Economic Review*, (3), 404–430.
- HUBER, P. (2006): "Regional Labor Market Developments in Transition," World Bank Policy Research Working Paper No. 3896.

- JIAN, T., J. D. SACHS, AND A. M. WARNER (1996): "Trends in Regional Inequality in China," *China Economic Review*, 7(1), 1–21.
- KANBUR, R., AND X. ZHANG (1999): "Which Regional Inequality? The Evolution of Rural-Urban and Inland-Coastal Inequality in China from 1983 to 1995," *Journal of Comparative Economics*, 27(4).
- (2005): "Fifty Years of Regional Inequality in China: A Journey through Central Planning, Reform, and Openness," *Review of Development Economics*, 9(1), 87–106.
- LAVELY, W. (2001): "First Impressions from the 2000 Census of China," *Population and Development Review*, 27(4), 755–769.
- MAGRINI, S. (2007): "Regional (di)convergence," in *Handbook of Regional and Urban Economics*, vol. 4, pp. 2741–2796. Elsevier B.V.
- MENG, X., R. GREGORY, AND Y. WANT (2005): "Poverty, Inequality, and Growth in Urban China, 1986-2000," *Journal of Comparative Economics*, 33(4), 710–729.
- RICE, P., AND A. VENABLES (2003): "Equilibrium Regional Disparities: Theory and British Evidence," *Regional Studies*.
- WAN, G., M. LU, AND Z. CHEN (2007): "Globalization and Regional Income Inequality: Empirical Evidence from within China," *The Review of Income and Wealth*, 53(1), 35–59.
- WANG, F.-L. (2004): "Reformed Migration Control and New Targeted People: China's Hukou System in the 2000s," *The China Quarterly*, 177, 115–132.
- (2005): "China's Household Registration (Hukou) System: Discrimination and Reform," Discussion paper, Congressional Statement, available at <http://www.cecc.gov/pages/roundtables/090205/Wang.php>.
- WHALLEY, J., AND S. ZHANG (2007): "A Numerical Simulation Analysis of (Hukou) Labour Mobility Restrictions in China," *Journal of Development Economics*, 83(2), 392–410.

XIA, Q., L. SONG, L. SHI, AND S. APPLETON (2013): “The Effects of the State Sector on Wage Inequality in Urban China: 19882007,” IZA Discussion Paper No. 7142.

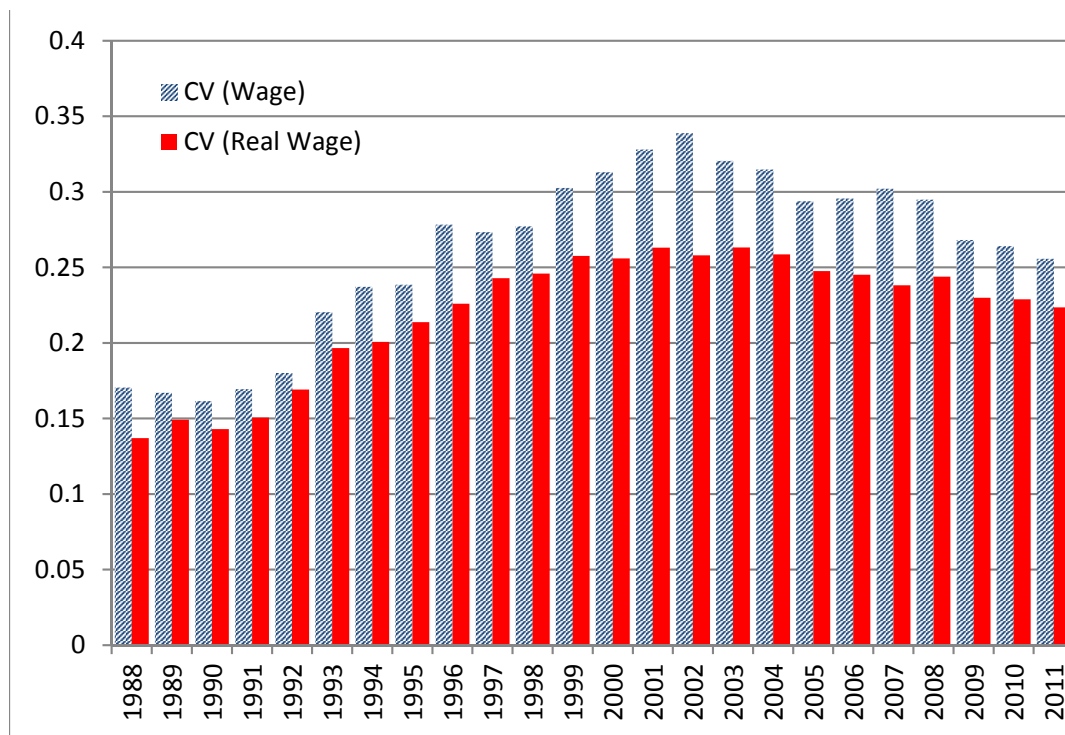
YAO, S., AND Z. ZHANG (2001): “On Regional Inequality and Diverging Clubs: A Case Study of Contemporary China,” *Journal of Comparative Economics*, 29(3), 466–484.

Figure 1: Disposable income in different province groups



Source: CEIC China Database, authors' calculations

Figure 2: Inequality in nominal and real wage



Source: CEIC China Database, authors' calculations

Table 1: What are the reasons for wage inequality?

|                              | (1)                 | (2)                 | (3)                 | (4)                  | (5)              | (6)                 | (7)                  | (8)                | (9)                 | (10)              |
|------------------------------|---------------------|---------------------|---------------------|----------------------|------------------|---------------------|----------------------|--------------------|---------------------|-------------------|
| COLLEGE                      | 0.129***<br>(0.029) |                     |                     |                      |                  |                     | -0.111***<br>(0.034) | -0.016<br>(0.025)  |                     |                   |
| EDUCATION                    |                     | 0.383***<br>(0.026) |                     |                      |                  |                     | 0.397***<br>(0.061)  | 0.173**<br>(0.083) |                     |                   |
| SECONDARY                    |                     |                     | 0.245***<br>(0.049) |                      |                  |                     | 0.174***<br>(0.051)  | -0.152*<br>(0.079) |                     |                   |
| TERTIARY                     |                     |                     | 0.309***<br>(0.057) |                      |                  |                     | 0.204**<br>(0.083)   | -0.093<br>(0.081)  |                     |                   |
| AGRICULTURAL                 |                     |                     |                     | -0.662***<br>(0.094) |                  |                     | 0.119<br>(0.076)     | 0.008<br>(0.232)   |                     |                   |
| SUBSIDY                      |                     |                     |                     |                      | 0.016<br>(0.063) |                     | -0.075**<br>(0.029)  | -0.030<br>(0.032)  |                     |                   |
| BERTH CAPACITY               |                     |                     |                     |                      |                  | 0.374***<br>(0.123) | 0.070<br>(0.058)     | -0.115<br>(0.093)  |                     |                   |
| PCA                          |                     |                     |                     |                      |                  |                     |                      |                    | 9.060***<br>(0.710) | -1.408<br>(2.424) |
| Province Fixed Effects       | No                  | No                  | No                  | No                   | No               | No                  | No                   | Yes                | No                  | Yes               |
| Adj. R <sup>2</sup>          | 0.264               | 0.662               | 0.542               | 0.436                | -0.001           | 0.294               | 0.839                | 0.951              | 0.636               | 0.936             |
| N Obs.                       | 532                 | 364                 | 504                 | 392                  | 336              | 532                 | 336                  | 336                | 336                 | 336               |
| First data year              | 1993                | 1999                | 1993                | 1997                 | 1999             | 1993                | 1999                 | 1999               | 1999                | 1999              |
| Last data year = $\tau$      | 2011                | 2011                | 2010                | 2010                 | 2010             | 2011                | 2010                 | 2010               | 2010                | 2010              |
| S.D. of residual ( $\tau$ )  | 18.8                | 14.6                | 11.0                | 12.0                 | 21.8             | 19.1                | 7.5                  | 6.9                | 10.0                | 7.9               |
| S.D. of real wage ( $\tau$ ) | 20.9                | 20.9                | 21.5                | 21.5                 | 21.5             | 20.9                | 21.5                 | 21.5               | 21.5                | 21.5              |

Dependent variable: Average real wage, in percent deviation from China mean. 28 provinces.

Robust standard errors clustered on province are in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

PCA = 0.426\*COLLEGE + 0.462\*EDUCATION + 0.352\*SECONDARY + 0.432\*TERTIARY - 0.455\*AGRICULTURAL + 0.148\*SUBSIDY + 0.257\*BERTH

Table 2: Is migration explained by inequality in wages? Inter-provincial results

|                             | (1)                  | (2)                  | (3)                  | (4)                  | (5)                  | (6)                  | (7)                 | (8)                    | (9)                   |
|-----------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------------------|------------------------|-----------------------|
| REAL WAGE $_{t-5}$          | 20.686***<br>(3.223) | 17.100***<br>(2.202) | 15.769***<br>(1.834) | 14.538***<br>(1.686) | 21.178***<br>(5.071) | 15.517***<br>(2.294) | 7.418<br>(7.519)    | 14.034***<br>(2.307)   | 7.049<br>(10.681)     |
| COLLEGE $_{t-5}$            |                      | 1.569***<br>(0.284)  |                      |                      |                      | 0.479<br>(0.779)     | -6.462**<br>(2.479) |                        |                       |
| SECONDARY $_{t-5}$          |                      |                      | -2.464***<br>(0.885) |                      |                      | -4.293***<br>(0.932) | 2.106<br>(3.073)    |                        |                       |
| TERTIARY $_{t-5}$           |                      |                      | 9.539***<br>(1.012)  |                      |                      | 5.086*<br>(2.625)    | -4.111<br>(7.783)   |                        |                       |
| AGRICULTURAL $_{t-2.5}$     |                      |                      |                      | -9.526***<br>(2.072) |                      | -6.572***<br>(2.338) | 9.188<br>(14.929)   |                        |                       |
| BERTH CAPACITY $_{t-5}$     |                      |                      |                      |                      | -0.660<br>(3.303)    | 1.564<br>(1.846)     | 9.154<br>(5.629)    |                        |                       |
| PCA                         |                      |                      |                      |                      |                      |                      |                     | 105.967***<br>(23.714) | -236.048<br>(255.327) |
| Province Fixed Effects      | No                   | No                   | No                   | No                   | No                   | No                   | Yes                 | No                     | Yes                   |
| Adj. R <sup>2</sup>         | 0.687                | 0.736                | 0.827                | 0.787                | 0.682                | 0.849                | 0.891               | 0.745                  | 0.776                 |
| S.D. of residual in 2010    | 323.8                | 284.3                | 226.0                | 256.1                | 320.2                | 213.6                | 106.7               | 287.5                  | 146.2                 |
| S.D. of immigration in 2010 | 660.8                | 660.8                | 660.8                | 660.8                | 660.8                | 660.8                | 660.8               | 660.8                  | 660.8                 |

Dependent variable: Migration to Urban area of province  $i$  from outside province  $i$  in 1995–2000 or cross-province migration in 2005–2010, in bp of total province population in 1997 or 2007, respectively.

28 provinces, 56 observations. Robust standard errors clustered on province are in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

$t =$  Census year: 2000 or 2010.

PCA =  $0.473*(COLLEGE_{t-5}) + 0.451*(SECONDARY_{t-5}) + 0.468*(TERTIARY_{t-5}) - 0.495*(AGRICULTURAL_{t-2.5}) + 0.330*(BERTH_{t-5})$

Table 3: Is migration explained by inequality in wage growth? Inter-provincial results

|   | (1)                   | (2)                  | (3)                  | (4)                  | (5)                   | (6)                  | (7)                 | (8)                   | (9)                   |
|---|-----------------------|----------------------|----------------------|----------------------|-----------------------|----------------------|---------------------|-----------------------|-----------------------|
| REAL WAGE <sub>t-5</sub> /REAL WAGE <sub>t-10</sub> | 52.365***<br>(11.746) | 47.656***<br>(9.914) | 37.876***<br>(7.203) | 44.616***<br>(8.026) | 58.340***<br>(15.273) | 43.176***<br>(8.657) | 29.667*<br>(15.594) | 42.279***<br>(9.520)  | 42.444***<br>(13.457) |
| REAL WAGE <sub>t-10</sub>                           | 22.353***<br>(3.207)  | 18.497***<br>(1.920) | 17.321***<br>(1.747) | 15.956***<br>(1.533) | 25.098***<br>(5.281)  | 18.132***<br>(2.335) | 12.053<br>(8.359)   | 15.898***<br>(2.117)  | 19.672***<br>(5.768)  |
| COLLEGE <sub>t-5</sub>                              |                       | 1.748***<br>(0.263)  |                      |                      |                       | 0.822<br>(0.629)     | -4.241**<br>(1.833) |                       |                       |
| SECONDARY <sub>t-5</sub>                            |                       |                      | -2.112**<br>(0.795)  |                      |                       | -3.177***<br>(0.812) | 0.809<br>(3.331)    |                       |                       |
| TERTIARY <sub>t-5</sub>                             |                       |                      | 9.274***<br>(1.055)  |                      |                       | 4.749*<br>(2.516)    | -2.114<br>(4.690)   |                       |                       |
| AGRICULTURAL <sub>t-2.5</sub>                       |                       |                      |                      | -8.899***<br>(2.316) |                       | -4.507**<br>(2.042)  | 4.031<br>(10.080)   |                       |                       |
| BERTH CAPACITY <sub>t-5</sub>                       |                       |                      |                      |                      | -3.263<br>(3.469)     | -0.736<br>(1.616)    | 1.010<br>(2.653)    |                       |                       |
| PCA   |                       |                      |                      |                      |                       |                      |                     | 99.276***<br>(27.022) | -223.414<br>(163.428) |
| Province Fixed Effects                              | No                    | No                   | No                   | No                   | No                    | No                   | Yes                 | No                    | Yes                   |
| Adj. R <sup>2</sup>                                 | 0.746                 | 0.817                | 0.883                | 0.832                | 0.755                 | 0.894                | 0.906               | 0.803                 | 0.894                 |
| S.D. of residual in 2010                            | 328.0                 | 267.1                | 209.5                | 260.3                | 311.4                 | 188.6                | 103.0               | 284.7                 | 119.4                 |
| S.D. of immigration in 2010                         | 660.8                 | 660.8                | 660.8                | 660.8                | 660.8                 | 660.8                | 660.8               | 660.8                 | 660.8                 |

Dependent variable: Migration to Urban area of province  $i$  from outside province  $i$  in 1995–2000 or cross-province migration in 2005–2010, in bp of total province population in 1997 or 2007, respectively.

28 provinces, 56 observations. Robust standard errors clustered on province are in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

$t =$  Census year: 2000 or 2010.

PCA =  $0.473*(COLLEGE_{t-5}) + 0.451*(SECONDARY_{t-5}) + 0.468*(TERTIARY_{t-5}) - 0.495*(AGRICULTURAL_{t-2.5}) + 0.330*(BERTH_{t-5})$



Table 4: Is migration reducing wage inequality?

|                           | (1)                  | (2)                 | (3)                 | (4)                 | (5)                 | (6)                 | (7)                 | (8)                 | (9)               | (10)                | (11)                | (12)                |
|---------------------------|----------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|-------------------|---------------------|---------------------|---------------------|
| IN-MIGRATION $\alpha$     | 34.397***<br>(4.003) | 5.247***<br>(1.820) | 3.603<br>(2.431)    | 1.109<br>(3.080)    | 2.267<br>(4.438)    | 0.072<br>(2.955)    | 5.367**<br>(2.012)  | 1.898<br>(3.688)    | 3.383<br>(8.249)  | 0.725<br>(2.675)    | -10.334<br>(16.819) |                     |
| REAL WAGE $_{t-5}$        |                      | 0.871***<br>(0.062) | 0.851***<br>(0.045) | 0.777***<br>(0.049) | 0.756***<br>(0.079) | 0.877***<br>(0.057) | 0.905***<br>(0.070) | 0.714***<br>(0.111) | 0.436<br>(0.449)  | 0.790***<br>(0.046) | 0.538<br>(0.522)    | 0.979***<br>(0.074) |
| COLLEGE $_{t+1}$          |                      |                     | 0.038<br>(0.024)    |                     |                     |                     |                     | -0.031<br>(0.038)   | 0.017<br>(0.076)  |                     |                     |                     |
| EDUCATION $_{t+1}$        |                      |                     |                     | 0.129***<br>(0.036) |                     |                     |                     | 0.116*<br>(0.060)   | 0.161<br>(0.174)  |                     |                     |                     |
| SECONDARY $_{t+1}$        |                      |                     |                     |                     | 0.079**<br>(0.033)  |                     |                     | 0.084<br>(0.050)    | -0.262<br>(0.263) |                     |                     |                     |
| TERTIARY $_{t+1}$         |                      |                     |                     |                     | 0.111<br>(0.067)    |                     |                     | 0.062<br>(0.082)    | -0.098<br>(0.257) |                     |                     |                     |
| AGRICULTURAL $_{t+1}$     |                      |                     |                     |                     |                     | -0.148**<br>(0.054) |                     | -0.012<br>(0.101)   | -0.098<br>(0.468) |                     |                     |                     |
| BERTH CAPACITY $_{t+1}$   |                      |                     |                     |                     |                     |                     | -0.033<br>(0.045)   | -0.040<br>(0.048)   | -0.173<br>(0.177) |                     |                     |                     |
| PCA                       |                      |                     |                     |                     |                     |                     |                     |                     |                   | 2.608***<br>(0.797) | -4.114<br>(8.103)   |                     |
| Province Fixed Effects    | No                   | No                  | No                  | No                  | No                  | No                  | No                  | No                  | Yes               | No                  | Yes                 | No                  |
| Adj. R <sup>2</sup>       | 0.644                | 0.852               | 0.861               | 0.881               | 0.866               | 0.860               | 0.852               | 0.876               | 0.832             | 0.870               | 0.801               | 0.850               |
| S.D. of residual in 2010  | 8.6                  | 8.5                 | 7.9                 | 7.7                 | 7.8                 | 8.0                 | 8.4                 | 7.5                 | 5.7               | 7.7                 | 7.0                 | 9.0                 |
| S.D. of real wage in 2010 | 20.9                 | 20.9                | 20.9                | 20.9                | 20.9                | 20.9                | 20.9                | 20.9                | 20.9              | 20.9                | 20.9                | 20.9                |

Dependent variable: Average real wage, in percent deviation from China mean, in year  $t + 1$ .

28 provinces, 56 observations. Robust standard errors clustered on province are in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

$t =$  Census year: 2000 or 2010.

PCA =  $0.423*(COLLEGE_{t+1}) + 0.455*(EDUCATION_{t+1}) + 0.374*(SECONDARY_{t+1}) + 0.437*(TERTIARY_{t+1}) - 0.460*(AGRICULTURAL_{t+1}) + 0.266*(BERTH_{t+1})$

$\alpha$  Coefficients and standard errors have been multiplied by 1000 for presentation purposes

No 2011 data available for SECONDARY, TERTIARY, and AGRICULTURAL variables; 2010 data used instead

Table A.1: Variable Definitions

**NOMINAL WAGE** Average annual wage in yuan per person for staff and workers in enterprises, institutions, and government agencies, which reflects the general level of wage income.

**CPI** Consumer price index (1988 = 100).

**REAL WAGE** Average annual wage in yuan per person for staff and workers in enterprises, institutions, and government agencies deflated by the consumer price index (1988 = 100).

**COLLEGE** Number of college graduates of regular institutions of higher learning scaled by population.

**EDUCATION** Expenditure of provincial government on education scaled by population.

**PRIMARY** Share of employees involved in the production of raw goods (primary sector).

**SECONDARY** Share of employees involved in the manufacture of goods (secondary sector).

**TERTIARY** Share of employees in the services sector (tertiary sector).

**AGRICULTURAL** Share of agricultural population.

**BERTH CAPACITY** Number of berths in major coastal ports (10,000 ton class).

**SUBSIDY** National government subsidy to provincial government per capita.

**INMIGRATION** Number of migrants moving into an urban area of a province from an urban or rural area in another province, in basis points of population from 2.5 years before.

**WITHINMIGRATION** Number of migrants moving into an urban area of a given province from a rural area in that province, in basis points of population from 2.5 years before.

Table A.2: Summary Statistics

|                               | First Data Year (of PD) | Last Data Year (of PD) | Mean (level) | Std. Dev. (level) | Mean (PD) | Std. Dev.(PD) |
|-------------------------------|-------------------------|------------------------|--------------|-------------------|-----------|---------------|
| NOMINAL WAGE                  | 1993                    | 2011                   | 16250.77     | 12488.94          | 1.49      | 29.35         |
| CPI                           | 1993                    | 2011                   | 267.18       | 46.21             | -         | -             |
| REAL WAGE                     | 1993                    | 2011                   | 5547.11      | 3490.30           | -6.29     | 22.30         |
| COLLEGE PER CAPITA            | 1993                    | 2011                   | 0.00         | 0.00              | 13.57     | 88.05         |
| EDUCATION PER CAPITA          | 1999                    | 2011                   | 140.68       | 123.31            | 10.81     | 48.46         |
| SECONDARY                     | 1993                    | 2010                   | 0.23         | 0.10              | -5.26     | 42.11         |
| TERTIARY                      | 1993                    | 2010                   | 0.30         | 0.09              | 4.88      | 28.05         |
| AGRICULTURAL                  | 1997                    | 2010                   | 0.67         | 0.16              | -5.20     | 23.09         |
| SUBSIDY PER CAPITA            | 1999                    | 2010                   | 451.24       | 339.03            | 19.05     | 57.67         |
| BERTH CAPACITY                | 1993                    | 2011                   | 16.76        | 32.01             | -         | -             |
| INMIGRATION (2000 CENSUS)     | -                       | -                      | 0.03         | 0.04              | -         | -             |
| WITHINMIGRATION (2000 CENSUS) | -                       | -                      | 0.03         | 0.01              | -         | -             |
| INMIGRATION (2010 CENSUS)     | -                       | -                      | 0.04         | 0.07              | -         | -             |
| WITHINMIGRATION (2010 CENSUS) | -                       | -                      | 0.06         | 0.02              | -         | -             |

See Table A.1 for variable definitions

PD = Percent deviation from country average in each year

Migration data reported as percent of share of population from 2.5 years before

Table A.3a: Migration Statistics: Cross Province and Intra-Province, 2000 Census

| Province          | Immigration       |           | Within-migration  |           |
|-------------------|-------------------|-----------|-------------------|-----------|
|                   | Gross (Thousands) | Share (%) | Gross (Thousands) | Share (%) |
| Anhui             | 174               | 0.28      | 1143              | 1.86      |
| Beijing           | 1592              | 12.84     | 227               | 1.83      |
| Chongqing/Sichuan | 448               | 0.40      | 3031              | 2.68      |
| Fujian            | 905               | 2.76      | 1413              | 4.30      |
| Gansu             | 182               | 0.73      | 468               | 1.88      |
| Guangdong         | 8534              | 12.10     | 3863              | 5.48      |
| Guangxi           | 236               | 0.51      | 1116              | 2.41      |
| Guizhou           | 202               | 0.56      | 681               | 1.89      |
| Hainan            | 183               | 2.46      | 191               | 2.57      |
| Hebei             | 497               | 0.76      | 1415              | 2.17      |
| Heilongjiang      | 238               | 0.63      | 613               | 1.63      |
| Henan             | 326               | 0.35      | 1576              | 1.71      |
| Hunan             | 273               | 0.42      | 1429              | 2.21      |
| Inner Mongolia    | 228               | 0.98      | 797               | 3.43      |
| Jiangsu           | 1265              | 1.77      | 2168              | 3.03      |
| Jiangxi           | 163               | 0.39      | 922               | 2.22      |
| Jilin             | 206               | 0.79      | 413               | 1.57      |
| Liaoning          | 600               | 1.45      | 742               | 1.79      |
| Ningxia           | 76                | 1.43      | 141               | 2.67      |
| Qinghai           | 70                | 1.42      | 90                | 1.82      |
| Shaanxi           | 366               | 1.03      | 714               | 2.00      |
| Shandong          | 667               | 0.76      | 2860              | 3.26      |
| Shanghai          | 1947              | 13.08     | 344               | 2.31      |
| Shanxi            | 225               | 0.72      | 734               | 2.34      |
| Tianjin           | 442               | 4.64      | 148               | 1.55      |
| Xinjiang          | 551               | 3.21      | 324               | 1.89      |
| Yunnan            | 627               | 1.53      | 1042              | 2.54      |
| Zhejiang          | 1892              | 4.27      | 2236              | 5.04      |
| Total             | 23115             | 1.99      | 30841             | 2.66      |

Census migration data from 1995-2000

Table A.3b: Migration Statistics: Cross Province and Intra-Province, 2010 Census

| Province          | Immigration       |           | Within-migration  |           |
|-------------------|-------------------|-----------|-------------------|-----------|
|                   | Gross (Thousands) | Share (%) | Gross (Thousands) | Share (%) |
| Anhui             | 310               | 0.51      | 2856              | 4.67      |
| Beijing           | 4017              | 24.60     | 1759              | 10.77     |
| Chongqing/Sichuan | 1018              | 0.93      | 6716              | 6.14      |
| Fujian            | 1883              | 5.21      | 2799              | 7.75      |
| Gansu             | 176               | 0.69      | 1193              | 4.68      |
| Guangdong         | 13841             | 14.33     | 8134              | 8.42      |
| Guangxi           | 400               | 0.84      | 2337              | 4.90      |
| Guizhou           | 245               | 0.68      | 1277              | 3.52      |
| Hainan            | 266               | 3.15      | 501               | 5.92      |
| Hebei             | 479               | 0.69      | 2745              | 3.95      |
| Heilongjiang      | 256               | 0.67      | 2519              | 6.59      |
| Henan             | 303               | 0.32      | 4168              | 4.45      |
| Hunan             | 355               | 0.56      | 3128              | 4.92      |
| Inner Mongolia    | 542               | 2.23      | 2253              | 9.28      |
| Jiangsu           | 3182              | 4.12      | 5398              | 6.99      |
| Jiangxi           | 202               | 0.46      | 1576              | 3.61      |
| Jilin             | 258               | 0.95      | 2271              | 8.32      |
| Liaoning          | 1060              | 2.47      | 4356              | 10.14     |
| Ningxia           | 176               | 2.88      | 476               | 7.80      |
| Qinghai           | 140               | 2.54      | 298               | 5.39      |
| Shaanxi           | 464               | 1.25      | 1854              | 5.00      |
| Shandong          | 1163              | 1.24      | 5781              | 6.17      |
| Shanghai          | 4217              | 20.44     | 1810              | 8.77      |
| Shanxi            | 341               | 1.01      | 1957              | 5.77      |
| Tianjin           | 2004              | 17.97     | 1177              | 10.56     |
| Xinjiang          | 643               | 3.07      | 1052              | 5.02      |
| Yunnan            | 369               | 0.82      | 1577              | 3.49      |
| Zhejiang          | 4237              | 8.22      | 3400              | 6.60      |
| Total             | 42549             | 3.42      | 75366             | 6.06      |

Census migration data from 2005-2010

Table A.4: Correlation matrix for explanatory variables

|                      | educ.pc.pd | red.pc.pd | shempPri.pd | shempsec.pd | shempter.pd | shpopagr.pd | rgovtsub.pc.pd | berth |
|----------------------|------------|-----------|-------------|-------------|-------------|-------------|----------------|-------|
| COLLEGE PER CAPITA   | 1.000      |           |             |             |             |             |                |       |
| EDUCATION PER CAPITA | 0.778      | 1.000     |             |             |             |             |                |       |
| PRIMARY              | -0.744     | -0.808    | 1.000       |             |             |             |                |       |
| SECONDARY            | 0.576      | 0.616     | -0.832      | 1.000       |             |             |                |       |
| TERTIARY             | 0.781      | 0.812     | -0.818      | 0.416       | 1.000       |             |                |       |
| AGRICULTURAL         | -0.726     | -0.838    | 0.812       | -0.530      | -0.838      | 1.000       |                |       |
| SUBSIDY PER CAPITA   | 0.234      | 0.384     | -0.024      | -0.057      | 0.220       | -0.324      | 1.000          |       |
| BERTH CAPACITY       | 0.229      | 0.383     | -0.562      | 0.589       | 0.284       | -0.475      | -0.155         | 1.000 |
| <i>N</i>             | 336        |           |             |             |             |             |                |       |

Table A.5: Is migration explained by inequality in wages? Intra-provincial results

|                                 | (1)                 | (2)                 | (3)                | (4)                  | (5)                | (6)                | (7)                  | (8)                | (9)                   |
|---------------------------------|---------------------|---------------------|--------------------|----------------------|--------------------|--------------------|----------------------|--------------------|-----------------------|
| REAL WAGE $_{t-5}$              | 3.595***<br>(0.696) | 4.328***<br>(1.042) | 3.485**<br>(1.401) | 1.981*<br>(1.134)    | 1.628<br>(1.633)   | 0.429<br>(1.888)   | -7.472<br>(6.974)    | 2.661*<br>(1.338)  | -10.025<br>(10.555)   |
| COLLEGE $_{t-5}$                |                     | -0.321<br>(0.248)   |                    |                      |                    | -1.204*<br>(0.639) | -3.192<br>(3.831)    |                    |                       |
| SECONDARY $_{t-5}$              |                     |                     | -1.406<br>(0.831)  |                      |                    | -1.639<br>(1.468)  | -5.509<br>(4.994)    |                    |                       |
| TERTIARY $_{t-5}$               |                     |                     | 2.303**<br>(0.895) |                      |                    | 4.212*<br>(2.263)  | -1.419<br>(9.725)    |                    |                       |
| AGRICULTURAL $_{t-2.5}$         |                     |                     |                    | -2.501***<br>(0.762) |                    | -2.919<br>(2.315)  | 13.222<br>(22.070)   |                    |                       |
| BERTH CAPACITY $_{t-5}$         |                     |                     |                    |                      | 2.639**<br>(1.243) | 3.481**<br>(1.476) | 14.051***<br>(4.609) |                    |                       |
| PCA                             |                     |                     |                    |                      |                    |                    |                      | 14.878<br>(16.521) | -340.074<br>(251.441) |
| Province Fixed Effects          | No                  | No                  | No                 | No                   | No                 | No                 | Yes                  | No                 | Yes                   |
| Adj. R <sup>2</sup>             | 0.074               | 0.067               | 0.095              | 0.089                | 0.101              | 0.217              | 0.319                | 0.062              | -0.269                |
| S.D. of residual in 2010        | 173.4               | 179.6               | 168.6              | 148.8                | 182.1              | 172.5              | 109.2                | 166.8              | 115.2                 |
| S.D. of withinmigration in 2010 | 216.5               | 216.5               | 216.5              | 216.5                | 216.5              | 216.5              | 216.5                | 216.5              | 216.5                 |

Dependent variable: Migration to Urban area of province  $i$  from Rural area of province  $i$  in 1995–2000 or migration within province in 2005–2010, in bp of total province population in 1997 or 2007, respectively.

28 provinces, 56 observations. Robust standard errors clustered on province are in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

$t =$  Census year: 2000 or 2010.

PCA =  $0.473*(COLLEGE_{t-5}) + 0.451*(SECONDARY_{t-5}) + 0.468*(TERTIARY_{t-5}) - 0.495*(AGRICULTURAL_{t-2.5}) + 0.330*(BERTH_{t-5})$

Table A.6: Is migration explained by inequality in wage growth? Intra-provincial results

|   | (1)                  | (2)                  | (3)                  | (4)                  | (5)                  | (6)                  | (7)                  | (8)                  | (9)                    |
|---|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|------------------------|
| REAL WAGE <sub>t-5</sub> /REAL WAGE <sub>t-10</sub> | 52.201***<br>(4.671) | 53.212***<br>(4.993) | 53.153***<br>(4.689) | 50.868***<br>(4.756) | 52.083***<br>(4.703) | 52.942***<br>(5.089) | 51.971***<br>(9.895) | 53.231***<br>(5.246) | 64.117***<br>(8.395)   |
| REAL WAGE <sub>t-10</sub>                           | 4.104***<br>(0.741)  | 4.931***<br>(1.064)  | 5.402***<br>(1.191)  | 3.003***<br>(0.944)  | 4.049***<br>(1.210)  | 4.077***<br>(1.656)  | 7.145<br>(5.040)     | 4.762***<br>(1.347)  | 13.076***<br>(3.890)   |
| COLLEGE <sub>t-5</sub>                              |                      | -0.375*<br>(0.217)   |                      |                      |                      | -0.837**             | -1.459<br>(1.480)    |                      |                        |
| SECONDARY <sub>t-5</sub>                            |                      |                      | -1.272**<br>(0.532)  |                      |                      | (0.403)              | (1.480)              |                      |                        |
| TERTIARY <sub>t-5</sub>                             |                      |                      | 0.467<br>(0.748)     |                      |                      | -1.264*<br>(0.690)   | -3.340<br>(2.789)    |                      |                        |
| AGRICULTURAL <sub>t-2.5</sub>                       |                      |                      |                      | -1.532<br>(1.015)    |                      | (1.449)              | (3.312)              |                      |                        |
| BERTH CAPACITY <sub>t-5</sub>                       |                      |                      |                      |                      | 0.065<br>(1.170)     | -0.970<br>(1.546)    | 2.005<br>(2.785)     |                      |                        |
| PCA   |                      |                      |                      |                      |                      |                      |                      | -10.129<br>(16.023)  | -217.897**<br>(82.706) |
| Province Fixed Effects                              | No                   | No                   | No                   | No                   | No                   | No                   | Yes                  | No                   | Yes                    |
| Adj. R <sup>2</sup>                                 | 0.680                | 0.690                | 0.696                | 0.686                | 0.674                | 0.767                | 0.860                | 0.677                | 0.821                  |
| S.D. of residual in 2010                            | 164.6                | 171.5                | 162.3                | 150.2                | 164.6                | 137.7                | 60.1                 | 168.9                | 74.1                   |
| S.D. of withinmigration in 2010                     | 216.5                | 216.5                | 216.5                | 216.5                | 216.5                | 216.5                | 216.5                | 216.5                | 216.5                  |

Dependent variable: Migration to Urban area of province  $i$  from Rural area of province  $i$  in 1995–2000 or migration within province in 2005–2010, in bp of total province population in 1997 or 2007, respectively.

28 provinces, 56 observations. Robust standard errors clustered on province are in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

$t =$  Census year: 2000 or 2010.

PCA =  $0.473*(COLLEGE_{t-5}) + 0.451*(SECONDARY_{t-5}) + 0.468*(TERTIARY_{t-5}) - 0.495*(AGRICULTURAL_{t-2.5}) + 0.330*(BERTH_{t-5})$



Table A.7: Does migration reduce wage inequality? Intra-provincial results

|                           | (1)                  | (2)                 | (3)                 | (4)                 | (5)                 | (6)                  | (7)                 | (8)                 | (9)                | (10)                | (11)               | (12)                |
|---------------------------|----------------------|---------------------|---------------------|---------------------|---------------------|----------------------|---------------------|---------------------|--------------------|---------------------|--------------------|---------------------|
| WITHIN-MIGRATION $\alpha$ | 28.489***<br>(7.115) | 4.033<br>(3.850)    | 6.533*<br>(3.683)   | 4.467<br>(3.327)    | 3.583<br>(4.046)    | 1.976<br>(4.088)     | 6.000<br>(4.038)    | 5.146<br>(3.816)    | 10.574<br>(10.455) | 3.371<br>(3.899)    | -0.784<br>(10.532) |                     |
| REAL WAGE $_{t-5}$        |                      | 0.965***<br>(0.077) | 0.891***<br>(0.057) | 0.778***<br>(0.058) | 0.783***<br>(0.065) | 0.874***<br>(0.041)  | 1.008***<br>(0.069) | 0.748***<br>(0.085) | 0.604<br>(0.433)   | 0.790***<br>(0.043) | 0.493<br>(0.673)   | 0.979***<br>(0.074) |
| COLLEGE $_{t+1}$          |                      |                     | 0.047**<br>(0.022)  |                     |                     |                      |                     | -0.023<br>(0.039)   | 0.048<br>(0.103)   |                     |                    |                     |
| EDUCATION $_{t+1}$        |                      |                     |                     | 0.133***<br>(0.031) |                     |                      |                     | 0.114*<br>(0.060)   | 0.110<br>(0.198)   |                     |                    |                     |
| SECONDARY $_{t+1}$        |                      |                     |                     |                     | 0.074**<br>(0.029)  |                      |                     | 0.076<br>(0.051)    | -0.234<br>(0.248)  |                     |                    |                     |
| TERTIARY $_{t+1}$         |                      |                     |                     |                     | 0.127***<br>(0.045) |                      |                     | 0.059<br>(0.076)    | -0.015<br>(0.258)  |                     |                    |                     |
| AGRICULTURAL $_{t+1}$     |                      |                     |                     |                     |                     | -0.144***<br>(0.044) |                     | -0.008<br>(0.106)   | -0.175<br>(0.435)  |                     |                    |                     |
| BERTH CAPACITY $_{t+1}$   |                      |                     |                     |                     |                     |                      | -0.046<br>(0.048)   | -0.047<br>(0.050)   | -0.223<br>(0.198)  |                     |                    |                     |
| PCA                       |                      |                     |                     |                     |                     |                      |                     |                     |                    | 2.643***<br>(0.684) | -1.317<br>(6.377)  |                     |
| Province Fixed Effects    | No                   | No                  | No                  | No                  | No                  | No                   | No                  | No                  | Yes                | No                  | Yes                | No                  |
| Adj. R <sup>2</sup>       | 0.086                | 0.850               | 0.864               | 0.883               | 0.867               | 0.860                | 0.851               | 0.878               | 0.842              | 0.871               | 0.791              | 0.850               |
| S.D. of residual in 2010  | 17.0                 | 8.9                 | 8.0                 | 7.8                 | 7.8                 | 8.1                  | 8.7                 | 7.6                 | 5.6                | 7.7                 | 7.1                | 9.0                 |
| S.D. of real wage in 2010 | 20.9                 | 20.9                | 20.9                | 20.9                | 20.9                | 20.9                 | 20.9                | 20.9                | 20.9               | 20.9                | 20.9               | 20.9                |

Dependent variable: Average real wage, in percent deviation from China mean, in year  $t + 1$ .

28 provinces, 56 observations. Robust standard errors clustered on province are in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

$t =$  Census year: 2000 or 2010.

PCA =  $0.423*(COLLEGE_{t+1}) + 0.455*(EDUCATION_{t+1}) + 0.374*(SECONDARY_{t+1}) + 0.437*(TERTIARY_{t+1}) - 0.460*(AGRICULTURAL_{t+1}) + 0.266*(BERTH_{t+1})$

$\alpha$  Coefficients and standard errors have been multiplied by 1000 for presentation purposes

No 2011 data available for SECONDARY, TERTIARY, and AGRICULTURAL variables; 2010 data used instead