GCOE Discussion Paper Series

Global COE Program
Human Behavior and Socioeconomic Dynamics

Discussion Paper No. 285

Wage growth through job hopping in China

Kenn Ariga, Fumio Ohtake, Masaru Sasaki and Zheren Wu

November 2012

GCOE Secretariat
Graduate School of Economics
OSAKA UNIVERSITY
1-7 Machikaneyama, Toyonaka, Osaka, 560-0043, Japan

Wage growth through job hopping in China*

Kenn Ariga[†], Fumio Ohtake[‡], Masaru Sasaki[§] and Zheren Wu[¶]

November 2012

Abstract

This paper uses a unique survey of the Chinese youth to construct a panel data in which we keep track of geographical and job mobilities. Our estimation results deliver the following major findings.

(1) The sample individuals are highly mobile. Job quits and relocations are frequent and they are closely correlated. We find the job hopping to be highly productive as our estimates indicate each job quit generates more than .2 log increase in monthly wage.

.(2) The migrant disadvantage in urban labor market is compensated by their higher job mobility. After four jobs, the expected earnings differentials essentially disappear. We also find that migration and job mobility are highly selective processes. Our evidence indicates that the migrants are positively selected.

(3) Job and location mobilities are highly dependent upon family back ground and personal traits which we interpret as representing unobservable characteristics associated with risk taking, active and optimistic personality, as well as the implied economic incentives to migrate and keep searching for better jobs.

JEL Classification numbers: J31, J61 and J62

*Earlier versions of the paper were presented at Fudan University (China),Thailand Development Research Institute (Thailand), Kansai Labor Workshop, Trans Pacific Labor Seminar 2011 and Conference on Labor Economics at Awajishima (Japan). We thank the participants of these seminars and conferences for their valuable comments. We also wish to thank Masako Kurosawa for her valuable inputs at various stages of the research, and Liu Yang for her advises on Chinese official data . The research reported in the paper is supported by G-COE (Osaka University) 'Human Behavior and Socioeconomic Dynamics.'

†Corresponding author. Mailing address: Institute of Economic Research, Kyoto University Yoshida Honmachi, Sakyoku, Kyoto 6068501 Japan. E-mail: ariga@kier.kyoto-u.ac.jp

‡Institute of Social and Economic Research, Osaka University, Ibaraki Japan

§Graduate School of Economics, Osaka University, Osaka Japan

¶Department of Economics, Kinki University, Higashi-Osaka, Japan

1 Introduction

> " The only way to find a better job is to quit the one you had. Interviews took time away from work, and a new hire was expected to start right away. ... . The girls talked constantly of leaving. Workers are required to stay six months, and even then, permission to quit was not always granted. The factory held the first two months of every worker's pay; leaving without approval means losing that money and starting all over somewhere else.... Getting into the factory was easy. The hard part was getting out."
> (*Factory Girls* by L. Chan, Picador, 2008)

In a sense, there is nothing new in rural to urban migration in China: we know why it takes place, what it delivers, at least in the long run, and when it is likely to subsides. By now a truly voluminous empirical literature on migrations confirm some of our conventional wisdom on internal migration in China. They are motivated primarily by the economic incentives[1]: better work opportunities, better living conditions, better educations for the children, etc. Our own study shown below confirms the economic incentives as the most powerful inducement for the people to relocate, and change jobs. A simple logic suggests that the inducement is more powerful if the expected gains are larger, which we also confirm in the subsequent analysis. What makes it unique about internal migration in China is, however, its scale and its significance in an enormous economy in the midst of transformation.

This paper focuses on the joint processes of migration and school to work transition, using the survey we conducted in 2009 for the Chinese youth aged mostly between 18-30. Focusing upon the young workers enables us to observe the crucial process of career formation in the contemporary Chinese labor market. Our main objective of the analysis is to investigate the interactions between spacial and labor mobility including rural to urban migration in the context of school to work transition. For this purpose, we estimate a dual decision of workers: they choose location and jobs. We try to unearth multi-facetted interactions between location and occupation choices.

Moreover, by focusing on the youth, we show changing characteristics of internal migration in more recent years. In 1980s and 1990s, the rural to urban migration was dominated by the adult population. In the most recent years, however, the migrants become younger and more educated. As we will see below, the migrant workers in our sample are less attached to their birthplace, family occupations, and social network at home town. At the same time, our survey samples of high school education or less differ in many important ways from those with college education. Even after an explosive growth of college enrollment in the last decade, the college enrollment rate is still below 30%, whereas the senior high school enrollment rate is close to 80%.

[1] For example, using random samples from 2005 National Census data, Gagnon et al (2008) approximately 61% of migrants in urban areas listed economic reasons as their primary motives for migration.

Thus the great majority of the contemporary Chinese youth leave school and start their working life with high school education. High school graduates are no longer the elite few even among the cohort of migrant workers. According to 2009 *Survey of migrant workers in China* (National Bureau of Statistics 2009), among the young migrant workers from rural areas, 13.5% of them are senior high school graduates, whereas those with college education comprises 6.4%.

For these reasons, our survey is uniquely suited to estimate the impact of the regulation that potentially hinders the rural to urban migration. At the same time, our estimations reveal the major determinants of migration decision which selects those to migrate and others to remain in the rural area.

Our main findings are summarized below.

First of all, we find the individuals in our survey are highly mobile geographically as well as across jobs. Their mobility pays off handsomely in terms of employment status, pay, and job satisfaction. Each job change after (or simultaneously with) voluntary job separations yields roughly .25 increase in log of average monthly wage.

In line with the now voluminous literature on the economic effect of migration and job changes, we need to explicitly take account of the process through which only *some,* not all, of the rural residents decide to migrate (to urban areas), and, not all the sample individuals change jobs. An important finding corrected for such selectivity is that the migrants (to be defined later) are more mobile across jobs and across locations. As a result, the initial wage gap between those who finish school in rural area and those in the urban areas substantially narrows down when we compare their second or third jobs after finishing the school. After four jobs, or about 5 years of work experience, our estimations indicate the wage gap between the urban resident and migrants in our sample all but disappears.

On the other hand, not all the rural born samples migrate to urban areas after school. We will show that decisions on migration and relocation are heavily influenced by family back ground and individual attributes. Moreover, our estimations of the selection into migrants show that they are positively selected in terms of unobservable wage residuals. Thus, our results simultaneously confirm substantial disadvantage of rural migrants and, at the same time, highly successful career formation among those who migrate and continue to search for better jobs.

In what follows, section 2 reviews selectively the recent literature on the labor market in China, focusing on those directly relevant to the youth. In section 3, we offer a variety of statistics to provide a bird eye view of the data and sampled individuals in our survey. Section 4 reports our main results on wage. In 4.1, we summarize the major econometric issues at hand as we try to estimate the causal relation between earnings with migration and job changes. The substance of the analysis starts the OLS regressions of wage in 4.2. Our main results are shown in 4.3 and 4.4. In 4.3 we report the Heckman correction model of wage growth across jobs, using a recursive maximum likelihood estimation for four endogenous variables: *log wage (wage growth), work, relocation and migrate*. In 4.4, we take up selectivity issues arising from unobservability due to rejected wage offer, and migration to different labor markets. Section 5 concludes.

2 A Brief literature review

Given our focus on the interaction between migration and school to work transition, we need to place our analysis both in the context of spacial and labor mobilities. Thus our focus of this brief review is on the two issues centered upon the internal migration in contemporary China.

First we consider who among the rural born population will migrate and why. As such, the issue is fundamentally those of selectivity. Job mobility, on the other hand, is linked to the literature primarily through issues centered upon the wage gap between urban residents and rural migrants in the urban labor market.

2.1 Recent changes in internal migration in China

Our survey evidence show that a large share of the school leavers in our sample from rural areas migrate to industrialized provinces mostly on the eastern seaboard within the first few years after school. Earlier literature on internal migration in China focused primarily upon 'floating' migrants: i.e., those working temporarily in urban areas leaving the rest of family members at their homes in rural areas. According to Li at ILO (Li 2008), the migrant workers in urban areas exceeded 30 millions by the end of 1980s. By 2006, the estimated size of immigrants rose to 132 million. A popular perception of the migrant workers is that they are male farmers, uneducated and unskilled, engaged in low paying jobs in construction, service and manufacturing sectors. 2009 Survey of migrant workers in China (National Bureau of Statistics 2009) shows that the most recent data paints somewhat different pictures. First of all, their education attainment significantly improved, especially among the young cohort (aged between 16 and 29). The gender composition also shows clear change: among the young cohort, the share of female is now close to 50%, compared to less than 30% for those above age 40. The young migrants also differ in other dimensions. For example, the same survey reports that, whereas overall, 29.5% of the migrant workers were engaged in farming at least some time during the past year, only 10% did among the younger cohorts. Industry composition also differ significantly between the two groups. In the younger cohorts, 44% work in manufacturing and less than 10% in construction, whereas for the older cohorts, the shares of the two sectors are roughly comparable (31.5% in manufacturing and 27.8% in construction). Unlike the older cohorts, younger migrants moved to urban areas early on: among the cohorts of the age above 30 , the average age of the first immigration is 33.7 year, whereas those born in 1980s, the average age of the first migration is 21.1 years of age.

In short, the latest wave of migrant workers are not only younger, but better educated, migrated early (typically immediately after school), and less likely to retain farming work back home. It is thus fair to say that their migration decisions seem more permanent and determined to stay longer time in the urban area: the same survey reports that only 15% or less among the young migrants are determined to return home in the future. Our samples of the youth with mostly high school education are no longer atypical, selected few. Their migration decision seem far more permanent than those farm workers in 1980s and 90s who took up menial jobs in urban areas to supplement their income. The same 2009 survey also show that the migrating young cohorts are better educated than the comparable youth who remain in urban area.

Some of the recent research on migrants confirm these observations. Using three coordinated CHIP data sets in 2002, Xing (2010) find that permanent migrants are positively selected from rural population, especially in terms of education. On the other hand, Wu (2010), using a unique survey, finds that the middle level in human capital endowment are most likely to migrate.

As we will show later on, our own finding does indicate positive selection bias among the cohorts of the youth with less than college education. One of our sharper findings, however, is that the sorting occurs mainly at the level of family backgrounds and personal traits, rather than conventional variables representing human capital.

2.2 Discrimination against migrants in the urban labor market

In spite of these important changes in the composition of migrant workers, available evidence indicates strongly that the migrants are at disadvantage in urban labor market, in comparison with the urban residents.

By now it is well known that many factors can be potentially responsible for the apparent wedge in earnings between urban residents and migrant workers in the urban labor market in China.

First of all, the registration system of the country places rather stringent limitations on migrants to obtain permanent registrations in urban areas. This alone can account for more than one facets of the gap. To begin with, it effectively prohibits entry to some of jobs in government or state owned enterprises[2]. Moreover, it also excludes migrants from a variety of social benefits available to urban residents, most notably the migrant children are not allowed to attend local public schools without their registration (*hukou*) in the area. Finally, the urban registration (or the lack thereof) can be a source of statistical discrimination. To the extent that the urban residents and migrant workers differ in average quality or skills, the registration can be used as a form of screening device. Notice that these potential factors have different consequences.

The *hukou* system have underwent significant changes in the recent years. Most notably, with the effective abolishment of the agricultural and non agricultural residencies, the local government decides on whether or not to allow

[2] See Table 21 and our discussions in 4.4.3 that confirm this point.

permanent residency for the migrants[3]. Although the practices vary somewhat from one region to the other, it is fair to say that the difficulty of obtaining permanent residency in urban areas still remains a major factor in shaping the migrants' works and lives.

Aside from the lack of social securities and access to public schools in the urban area they reside, the migrants are far more likely to work without a formal labor contract. Thus they are more likely to have their wage unpaid, delayed, or retained for the first three to six months to prevent the early turnovers, as our quotation from Chan's book indicates[4].

Using two comparable household survey data sets for Shanghai in 1995, Meng and Zhang (2001) find evidence of discrimination against rural migrants in terms of both occupational attainment and earnings. They analyze the extent to which earnings differentials between rural migrants and urban residents are due to inter- or intra-occupational gaps and find that 82 percent of the hourly wage differential is due to unequal payment within occupation.

On the other hand, Demurger, S., M. Gurgand, L. Shi, and Y. Ximing (2008) use micro simulation and decompose the wage differentials into (1) sector allocation, (2) hourly wage, (3) working time, and (4) population structures. They find the main source of the gap is (4) and conclude that the gap is due to pre-market (education opportunities) rather than on-market sources. Another strand of literature tries to break down disadvantages into migration effect per se and the effect due to discriminations stemming from *hukou* system. Using Blinder-Oaxaca decomposition, Gagnon et al. (2009) find that 40 percent of the observed wage gap between rural and urban migrants might be due to *hukou* status.

Assuming at least some portion of the wage gap between urban residents and migrants is due to the gap in skills or human capital, one interesting question is the impact of the influx of migrant workers on the earnings of urban residents. Meng and Zhang examines the causal relationship between rural-urban migration and urban native workers' labour market outcomes in Chinese cities. They find that rural migrants in urban China have modest positive, or zero effects on the average employment and insignificant impact on earnings of urban workers. They conjecture that the reason for the lack of adverse effects is due partially to the labour market segregation between the migrants and urban natives, and partially due to the complementarity between the two groups of workers.

Xing (2010) compares the two types of migrants, those with urban registration (the permanent migrants) and the others without. Not surprisingly, he finds that the permanent residents are strongly positively selected (out of

[3]See Chan and Buckingham (2008) for the details. They also argue that as a result of localization on registration decisions, practices vary across regions. In some small cities, immigrants now have the possibility of obtaining permanent residency, in exchange for giving up the land they effectively own at their rural homes.

[4]See Meiyang (2007) for wage arrears and various discriminations against immigrant workers.

the rural born population), especially in terms of education. Knowing that the college education is one way to obtain urban registration, the distinction between discrimination and the difference due to education is not as clear cut as it may sound. Taken together, it seems difficult to draw any definitive conclusion on the source of the gap and its relation to the possible discrimination against the migrant workers.

On the other hand, there exists a broad consensus that at least part of the gaps are narrowed down over time as the migrant workers stay in urban area and settle down: i.e., migrant workers gradually assimilate to urban life and urban work places. The time needed for assimilation is yet another aspect of the wage gap between the rural migrants and urban resident in the urban labor market. Deng and Gustafsson (2006) show that the permanent rural migrants who received their urban *hukou* before age 25 are well integrated economically in their place of destination and they actually receive higher earnings than the local-born urban residents.

In line with the conventional wisdom, our regression results show that the immigrants earn significantly more than those who decided not to migrate. On the other hand, our results also confirm earlier findings that compared to native residents in urban areas, migrant workers do earn less and our best estimate indicates the presence of 10% gap. However, given the impacts of being migrant on relocations and job quits, our estimated model demonstrate that the initial wage gap between migrants and urban residents narrow down over time.

2.3 Migration and job mobility

Since school to work transition often entails both geographical and job mobility, it is natural to place our analysis in the context of migration impacts on the wage growth through job changes. To the extent that the migration is costly by itself but its impact on future wage path is positive, migration decision can be considered as a form of investment. Lehmer and Ludsteck (2008) use German social security data to analyze the heterogenous returns from migration in the short run and long run effect on wage growth. In particular, they find the immediate impact is larger for the low skilled, whereas the high skilled workers reap larger long term gains. In the context of contemporary China, the immediate gains are easy to grasp as the migration from rural to urban regions by itself lifts the over all wage structure they face. On the other hand, given the large empirical literature, as we have reviewed some above, on the gap between immigrant and residents in urban areas in China, it is not obvious that the migrants and urban residents face the same labor market. Thus the effect of migration is accentuated in the case of China because of the variety of disadvantages placed upon rural migrants in the urban labor market..

One such disadvantage is information. Many empirical studies on migration decisions find a significant positive impact on migration of having family members or friends and acquaintances in the migration destination. This impact is especially well known among the international immigration of ethnic Chinese population. Wu (2010) finds different self-selection between individuals who have moved as pioneers and migrants from households in which other members have already migrated. One way that followers benefit from the pioneers in the urban labor market is information on jobs.

Knight and Yueh (2004) finds that the migrants in the urban labor market is far more mobile than the urban residents. They consider the difference arising from the difference in the reservation wages between the two types. Thus, the different mobility is due to the segregation of the labor markets which, in turn, is responsible for the gap in the access to the 'good' jobs.

Kondo and Ou (2010), on the other hand, compares two types of permanent migrants and finds that rural to urban migrants are more mobile across jobs and they are more likely to move to better jobs by changing work units, whereas urban migrants are more likely to be promoted within a work-unit. Our own results are in line with the ones in Kondo and Ou in that the immigrants are more mobile both in terms of job changes and relocation. Thus, as we indicated above, immigrants tend to catch up resident workers in the urban labor markets in industrialized provinces.

2.4 Joint decisions on labor and geographical mobilities

Although a large variety of issues surrounding labor markets in China have been exposed to rigorous econometric estimations and testing, none, as far as we know, look into the interactions between location changes such as internal migration and job mobility. The youth is the crucial period in which people try different jobs, different life styles, and different locations before they settle down to a chosen occupation and residence. Focusing upon only one aspect is often valuable as it simplifies the analysis, whereas to the extent that the analysis misses the interactions, it is possible that such analysis suffers from mis-specifications.

For example, consider the issue of selectivity of migrants in the labor market. By definition, those are the workers who moved out of their family locations and face a different labor market than the one in the home place. Who migrates depends crucially on the comparisons of available jobs in these markets. Mobility depends also upon monetary and psychic cost of relocations. In that case, it is crucially important that we combine information on geographical and occupational mobilities. Our survey of the Chinese youth is designed specifically to address these issues, as we see below.

3 ISYC: The survey results overview and preliminary empirical analysis

3.1 Survey Overview

We conducted an Internet based Survey of the Youth in China (hereafter called as ISYC) aged between 16 and 31 with less than college education. The survey was conducted between February and March in 2009. The entire set of questionnaire translated into English or original Chinese versions are available upon request from the authors. As shown in the top row of *Table 3,* among the total of 3,336 sampled individuals, 1,972 are males, vocational school graduates comprise 45% of the total sample, and, about 30% of the samples are academic high school graduates. Roughly 20% of the sample are high school drop outs, and the remaining 5% are the middle school graduates. Except for the drop outs (proportionately more males than females), the compositions do not differ markedly across gender.

Figure 1 compares the shares of the sampled individuals according to the 31 provinces and specially designated cities in comparison with the overall population shares aso 2010. Densely populated and industrialized regions are somewhat over represented in our survey, (especially *Guangdong* province) in comparison with official population data. As the official data is based upon resident registration (*hukou*) system, it is likely to seriously under-report the actual population in residence for those heavily populated area. The problem is even more severe if we restrict our attention to the youth population. Hence it is unclear to what extent that the survey over-represents those in population centers. In any case, it is clear that our survey does well in terms of representing the wide spectrum of population geographically.

ISYC has two important focuses in the questionnaire. First, we have a detailed set of questions on the last year samples spent at the high school. In this part of the questionnaire, we asked types of school they attended, commuting methods, selected course and subjects, grades, as well as a host of questions on activities during the school. We also have a fairly comprehensive set of questions on individual attributes, friends and family background. The second focus of the survey is on relocation and job history. The latter is used to reconstruct a panel data jobs and relocations as we explain next.

3.2 Creating a panel data on job, wage, and relocation history

3.2.1 Panel construction

One important objective of the survey is to trace job and location changes as we anticipated large flows across jobs and internal migrations for the sample youth in China. We asked each respondent to answer locations of (1) birthplace, (2) the last school (mostly high school) they attended, and (3) all the residences where they stayed 6 months or longer *after the last school* they attended, and (4) the current residence, irrespective of the length they stayed. For each location, we asked (a) province, (b) administrative unit below province, and (c) whether or not the area is urban, suburb, or rural. We also asked the calendar year when they started and ended the stay. Our preliminary analysis indicated that the distinction between urban and suburban is often blurred so we decided to combine these two into urban area.

For job history, we asked to list all the jobs after the last school they attended which lasted at least 6months, and (if applicable) the current or the last job, irrespective of the length of the time they worked.

Thus, for location changes, we know the sequence and calendar years of changes. We have the same information for job history. When we combine them into a panel, however, we cannot tell the precise sequence of location *and* job changes if they occur in the same year. Moreover, as we ask them to list locations and jobs lasting 6 months or longer, each year can have a record in which up to three distinct locations or jobs are listed.

Our strategy to deal with this problem is to set up three artificial subperiods within each calendar year to accommodate these events. Again, within location or job history, we know the sequence, but, the timing recorded in terms of this sub unit cannot be used to infer the timing between location and job changes.

3.2.2 Wage information

For each job listed in the survey, we asked (1) types of employment (such as regular full time, fixed term, etc), (2) types of employer (government, governmental organizations, private firms, etc), (3) the method used to find the job (i.e., through introduction by a friend, direct application to job advertisement, etc), (4) the reason why you took up the job (multiple choice), (5) the reason why you quit the job (if applicable), and (6) monthly wage (in Yuan, or RMB).

For wage information, we asked the respondents to pick one from the 13 wage ranges (the lowest is less than 300 yuan, and the highest is more than 5,000 yuan). We used 250 yuan for the lowest, and 6,000 yuan for the highest, and for other 11 ranges, we use the center values to convert the answer to the multiple choice into wage rates in yuan. We then converted into log real wage using the national CPI index. Henceforth, we use this converted log real wage as out measure of the wage rate. One important caveat on wage is that we only have single observation of wage for each job, thus unable to keep track of wage changes, if any, within each job.

Given space limitation, we relegate summary statistics and a brief description of the data to Appendix A.

3.2.3 Migration

Go back to *Figure 1* wherein triangle marks indicate the sample shares of the survey in terms of the current residence, whereas the shares in terms of the birth place locations are shown by square marks. It is evident that sizable migrations have taken place. The biggest winner is *Guangdong* province with the net increase in 328, followed by *Beijing* (138) and *Shanghai* (117). These three regions account for 84% of the sum of the province level net increases (795). We select: *Beijin, Tianjin , Shanghai, Jiangsu, Zhejiang,* and *Guangdong* as the densely populated centers (PC) of industrialized provinces (and designated cities). Aside from small increases in *Xingjing* and *Tibet*, only these 6 regions show net increase in sample residence over the sample size in terms of birth place. It should be noted also that all 6 regions are in East coastal regions. Available data on migrant workers confirm the concentration of migrant workers in Eastern coastal regions: according to National Bureau of Statistics of China (2010), among the young cohorts of migrant workers (84.8 millions in total), 72.3% of them currently reside in one of Eastern regions.

In what follows, these 6 regions are called PC regions, and the rest are called Non PC regions. The average wages for each of 31 regions show that top 6 of them are all in these PC regions. We categorize the residence location using two indices. First, in each province, we categorize residential locations into 1: rural areas, 2: urban which include small cities and regional center cities. Then we classify each province and special cities into PC and Non PC regions as stated above.

Table 1 shows the distribution of current residence for 4 cohorts of samples according to their birth places. It is clear that across region mobility is predominantly from rural in non-PC provinces into urban areas in PC provinces. Among the sample individuals born in rural areas in non PC urban areas, only 22.8% remain in the location of the same category, whereas 97% of those born in urban areas of PC provinces currently live also in PC provinces.

Given the predominant flows from rural to urban, and from non PC to PC regions, we call a sample individual *immigrant if a person is born in place other than PC urban areas, but currently resides in PC urban region.* Using these conventions, we find that among 3,366 samples in our survey, 2,728 of them are born in regions other than PC urban. Among those potential candidates for migrants, roughly 50% (1,327) of them has been migrant, and 1,183 (43%) of them are currently migrants, living in PC urban regions. Not surprisingly, migrants are geographically more mobile than non-migrants. They are more mobile vis a vis those born in regions other than PC urban, and also against those who are born in PC urban regions. See *Table 2* . Being a migrant is not a permanent status. About 15% of them moved out of PC urban and reside elsewhere when the survey was taken. Still, by and large, return migrations is not a very common phenomenon. Again, this observation is supported by earlier finding that the migrations are more permanent in nature among the younger cohorts. The survey results also show that roughly a half of the first time migrations occur as they start their last school (most of them are high schools): they attended schools in PC urban regions. The next peak of immigration is at their third year after graduation and it is typically the first relocation. Again, these observations are consistent with the earlier findings based upon other nation wide statistics or surveys: the migration into industrialized Eastern regions occur early, within a few years after finishing schools.

3.3 Job changes and wage growth

3.3.1 Facts

Let us start with some basic tabulations of the data.

Table 3 shows the work experience: 66% currently have jobs, 19% are currently jobless but worked in the past, and the remaining 15% or so have never worked in the past. *Table 3* also decomposes the variable according to the education attainments. We observe again that vocational school graduates have the highest averages in terms of current work, and also the share of those with regular full time job is the highest. Middle school graduates come second, then the dropouts, and the academic high school graduates at the bottom. Among those currently employed, the impact of education attainment differ. At the bottom of *Table 3,* average monthly wages are shown. Academic and vocational school graduates are more or less comparable in average wages, and they are followed by the dropouts, and the middle school graduates. (corrections made)**.** Overall, the average real monthly wage is about 1200 RMB in 2005 prices[5], which is slightly below but quite comparable to the average monthly wage of the young immigrant workers, 1,328 RMB reported in National Bureau of Statistics of China (2010) . As we might have expected, at least for those with some high school education or better, we do observe sizable increase in wage over age.

Far more striking than this difference, however, is the impact of the number of jobs held in the past as shown in *Table 4*. Compared to the first jobs, the fifth job on average earns 88%, 97%, 68%, and 67% more for graduates of academic high school, vocational high school, high school drop outs, and middle school, respectively.

Other variables on job characteristics and job satisfactions suggest strongly that job hopping pays off handsomely. *Table 5* shows the reasons for quitting the job. As the sample individuals experience more jobs, their reasons for quitting change. In the first and second jobs, the most popular reason is "dead end job", whereas at their third or later jobs, the most popular reason is "found a better job." We also notice that the share of "wage too low" also increases as they experience more jobs. Similarly, to the question why you took the job, the most popular reason for the first and the second job is "because it is a type of job that I was looking for,", whereas in the third and fourth jobs, the most popular answer is "the job offers opportunity to learn and master professional skills."

Overall, these tabulations suggest that the individuals in our sample do rather well by changing jobs, in terms of wage they earn, types of jobs they land, and also in terms of personal satisfaction from the job.

Are frequent job changes and migrations related? Simple tabulations indicate indeed that they are related. For one thing, migrants hold more jobs after school, controlling for the age. At age 29, migrant have on average had more than 3 jobs since graduation, compared to about 2 jobs for non-immigrants. As a final piece of suggestive data, *Table 6* shows the correlation between relocations and job changes. The table shows that if a sample continues the same job from one period to the next, the probability of relocation is about 12%, whereas conditional on job change, relocation probability jumps up to 28%.

Reflecting the job mobility, most of job spells are short: among the com-

[5] If we limit to the current jobs held by the sample individuals, the average monthly wage is 1,740 RMB as of 2009.

pleted job spells, 55.7% of them end within one year, and less than 7% of jobs last more than 3 years. In short, the individuals in our survey are highly mobile across jobs and across locations. Stepping stone mobility indeed seems to be at work in China.

3.3.2 Stayers and movers

Having noted that job mobility on average pays off rather handsomely, a natural question to follow is who moves from job to job more frequently? More importantly, if the wage gains from job mobility is so large, why not everyone moves?

Below we show that a short answer to the question is simply that those who did not do well in the initial job tend to move more often, compensating for the potential wage loss in the initial match by job mobility. The wage growth due to job mobility indeed compensate for the lower initial wage. *Figure 2* shows the wage growth across jobs. Each connected line corresponds to the path of average log real wage across jobs. For example, the one labeled "two jobs" is for those who have had two jobs since school graduation up to the time of the survey. In order to control for the differences in age cohorts, we limit the samples to those with age 23 or older. Thus, they have at least 5 years of potential labor market experiences. The following points can be confirmed from these paths. First of all, those with larger number of jobs starts with lower average wage. The average slope of wage growth is also inversely related to the total number of jobs held. In short, those who start up with lower paid jobs are more likely to move and continue to do so. By the latest jobs, the initial differences all but disappear. For example, the log difference between the "two jobs" group and the most mobile ("five or more jobs") is .229 in the first jobs, which is more than offset by job mobility as the real log wage at their latest job exceeds the corresponding average of the "two jobs" group by .05 log points.

Figure 3 is a similar plot for those in PC urban labor markets. The migrants start with lower wage but they appear to catch up through job mobility. By their last jobs, the initial difference between those born in PC urban areas and migrants disappears. In sum, these figures suggest that the job mobility compensate for the initial disadvantages, either due to poor job match, or due to the disadvantage being an immigrant. On the other hand, both spatial and labor mobilities are endogenous choices and we need to see if and why migrants exihibit higher mobility. In the next section, we make these points more formally in regression analysis.

4 Wage growth, migration and job changes

In Section 3, we have shown a series of tabulations. They suggest the pivotal role played by job mobility in wage growth. We also noted that the job mobility is highly selective and the degree of the mobility is systematically correlated with other variables of interests, most notably, geographical mobility. In this section, we report three types of regression results in which we address these issues. In the first step, we run OLS and fixed effect regressions. We exploit these results to obtain our measure of the match specific component of wage. We use this measure in the second and third sets of regressions. In the second set of regressions, we estimate Heckman correction model of wage growth in which we incorporate selectivity of observed wage changes due to endogeneity of job to job quit decisions. In a similar vein, we also report the results of recursive maximum likelihood estimation on wage level that incorporates work, immigration, and relocation decisions. In the last set of regressions, we address two additional selectivity issues. One is the within job wage growth, and the second is immigration decision.

4.1 Econometric issues

To the extent those decisions to take up jobs, to change jobs, and to migrate from one place to another, ... are all endogenous and possibly entail selectivity biases, we need to pay due attentions to the following econometric issues in order to fully address these questions.

First, selectivity bias. Not all the sample individuals choose to work (or to be employed), the decision of which reflects a comparison of the net benefit from taking up a job for each worker. A decision to work and a wage offer contain a common unobserved shock which induces selection bias.

The second problem arises because of the endogeneity of migration decision. Our preliminary analysis have shown that those born in rural areas on average earn higher wage by migrating to the urban areas, especially in densely populated industrialized provinces. The regression coefficient on migration suffers, however, from selectivity bias because migration (to urban areas) are not randomly assigned outcome, but they are voluntary decisions. If anything, those who expect to benefit more from migrations are more likely to migrate, thus biasing the impact of migration upward. According to our definition, one must be born in areas other than PC urban and choose to reside in PC urban. Thus the selectivity of immigrant refers to those potentially able to migrate, i.e., those born in areas other than PC urban.

By the same token, the OLS estimates on the impact of job changes on wage are also suspect because decision to change a job clearly is also endogenous and we do not observe offers which were not taken.

Moreover, given the significant positive correlation between job to job quits and relocations, it is likely that (near) simultaneity of these mobility decisions might as well indicate that job to job quits with or without relocation differ each other. Whether or not to take up a new job in a remote location probably depends upon the cost of relocation and the net benefit from the new wage offer.

4.1.1 Regression design

Our approach to the issues summarized above is to model the joint decision on work, quit, migrate and relocation choices all of which in turn feed into the wage regression.

quit-relocation We assume that relocation causes quit, but not the other way around. Although this is clearly untrue as some of workers decides to take up a new job which involves relocation, we believe this is rather inconsequential because what matters is the likely joint effect of quits accompanied by relocation, as opposed to those without relocation. Using the recorded change in residence, we define *relocation 1*, a dummy variable which is equal to one in the sub-period where the change in residence is recorded. As we do not know precise timing of relocation and quit other than the calendar year of the events, we use *relocation* which takes value 1 whenever *relocation* dummy variable is unity in the previous, current, or in the next sub period. Since our time measure that divides one year to three periods is purely artificial device to accommodate more than one job or location changes within one year, *relocation* can be used as the variable that can tell if relocation takes place on or around the time of job changes.

migration We assume migration decision is done independently from quit or relocation decision. Again, this cannot be literally true. However, migration status is relatively stable as we documented above and relatively small portion of relocation involves migration while they are currently working.

Additional specifications are the following.

Labor market and migration Given our definition of migrant, those born in PC urban areas cannot be immigrant. Hence we need to estimate the system for two sets of sub samples. In the first subset, we exclude those born in PC urban regions. Hence we can model the decision to immigrate. The estimated impact of migration compares the wage earned by migrants in PC urban regions with those currently in areas other than PC urban. In the second subset, our samples are limited to those currently in PC urban areas. Thus the estimated impact of migrant measures the net disadvantage in earnings of migrants vis a vis those born in the PC urban areas. By construction, those native residents in PC urban areas cannot choose to be a migrant, we treat migrant status as exogenous in this set of samples.

treating quit decisions As we noted, we only have one wage observation for each job. Under highly restrictive assumption that wage at each job does not change over time, we only need to consider the endogeneity of job to job quit decision in the wage growth regression if all the workers face the same outside wage offer, for each type of workers[6]. Otherwise, we have to assume that wage offers are censored in the sense that some

[6] Presumption is that those with low wage or unhappy at the current job decide to take up a new job, which offers the same wage for everyone within the same type, which are controlled by regressors in the quit regression.

of those continue to stay at the current offer rejected wage offers, which by construction we cannot observe.

4.1.2 A model of on the job search

To fix these ideas, consider a simple model of on the job search. A worker i currently employed at a firm k with wage W_t^{ik} receives an outside offer with probability λ_t. The Bellman equation is given by

$$\begin{aligned} & J(\eta^{ik}, \mu^i, \epsilon_t^{ik}, \lambda_t, Z^{ik}) \\ = & W_t^{ik} + \beta\lambda_t \int_{J'>J} \left[J'(\eta', \mu^i, \epsilon'_{t+1}, \lambda_{t+1}, Z') - J(\eta^{ik}, \mu^i, \epsilon_{t+1}^{ik}, \lambda_{t+1}, Z^{ik}) \right] dF(\eta', \mu^i, \epsilon'_{t+1}, \lambda_{t+1}, Z') \end{aligned}$$

wherein β is discount factor and $F(\eta', \mu^i, \epsilon'_{t+1}, \lambda_{t+1}, Z')$ is the joint probability distribution for the state variables. We assume the log of the wage rate (denoted in lower case) is given by a conventional Mincerian model of wage augmented by job mobility. We posit

$$w_t^{ik} = \widetilde{w}_t^i + \eta^{ik} + \epsilon_t^{ik} \tag{1}$$

$$\widetilde{w}_t^i = \mu^i + g(\exp_t^i) \tag{2}$$

wherein μ^i represents innate ability for individual i, η^{ik} match specific productivity, and g represent general human capital accumulated by previous work experience[7]. Wage offers are given by

$$w_t^{ik'} = \widetilde{w}_t^i + \eta^{ik'} + \epsilon_t^{ik'}$$

If there is no cost of quitting a job and taking up a new one, the myopic comparison of the current wage and the the wage offer is sufficient and we have

$$q_t = 1$$

if and only if

$$w_t^{ik'} > w_t^{ik}$$

or,

$$\eta^{ik'} + \epsilon_t^{ik'} > \eta^{ik} + \epsilon_t^{ik}$$

In general, however, we only know that quit occurs when

$$J'(\eta', \mu^i, \epsilon'_{t+1}, \lambda_{t+1}, Z') > J(\eta^{ik}, \mu^i, \epsilon_{t+1}^{ik}, \lambda_{t+1}, Z^{ik})$$

Although such a equation does not lend itself to any closed form solution, if we assume further that the flow utility is given by the wage and the multiplicative random utility (θ_t^i)

[7] We assume away accumulation of job specific human capital in the base case. See, however, 4.4.1 for out analysis of the impact of job specific human capital employing a separately estimated effect of job tenure on log real wage.

$$\log U_t^{ik} = w_t^{ik} + \theta_t^i$$

Then it is sensible to postulate a pair of behavioral equations:

$$q_t = \lambda_t prob(\Delta w_t^i + \Delta\theta_t^i > d^i) > 0) \tag{3}$$
$$\Delta w_t^i = \Delta\eta^{ikk'} + \Delta\epsilon_t^i \tag{4}$$

Thus the probability of quitting to a new job is the joint probability that an offer arrives and that is acceptable. In order to make some progress on this specification, we need to find proxies for two crucial sources of unknown labor productivity, η^{ik} and μ^i.

We do so by estimating a fixed effect regression for real log wage, and also by estimating a OLS regression using the same pooled panel of jobs and wages in our survey records. In view of (1), we posit

$$u_t^{iOLS} = \eta^{ik} + \mu^i + \epsilon_t^{ik},$$
$$u_t^{FE} = e_t^{ik} + \overline{\mu}^i$$

wherein the second equation decomposes the total residuals into fixed term and the remaining residuals. Then compute

$$\begin{aligned}\widetilde{\eta}^{ik} &= u_t^{iOLS} - \overline{\mu}^i \\ &= \eta^{ik} + (\mu^i - \widehat{\mu}^i) + \epsilon_t^{ik} - e_t^{ik} \\ &= \eta^{ik} + \nu_t^{ik}\end{aligned}$$

We consider the construct $\widetilde{\eta}^{ik}$ is a noisy signal of the unobservable true match specific productivity, η^{ik}[8]. We regress this variable over responses on questions regarding the reasons for taking up and (when applicable) quitting the job and other job characteristics.

In view of the fact that we only observe accepted outside wage offer, the model (3) and (4) is a censored regression model. Thus we can estimate the system using Heckman correction method wherein the wage change is observed only when the worker decides to quit for a new job. Unfortunately, the model may also suffer also from endogeneity of two other determinants in wage and quit equations: migration and relocation decisions. Therefore, our estimation strategy is to posit an augmented Heckman correction model. To estimate simultaneously the wage growth across job, and three decisions on

[8] We are not claiming that we identify the match specific productivity in that the observed log wage clearly contains censored values of match specific productivity as is evident from the model. Ideally a fully dynamic structural model estimation can overcome the identification problem [For a fully dynamic and structural model incorporating spatial and labor mobility, see Kennan and Walker (2011)]. Our more modest objective here is to have some proxy for the impact of the match specific productivity as the driver of wage growth and job mobility.

job change, migration and relocation, we assume that these decisions are fed into wage growth and quit equations (1) (2) together with other regressors. This estimation is done by a joint maximization of likelihoods for these equations wherein we allow correlations in the error terms across equations. In the main results reported below, we employ *cmp.ado* routine in *STATA*. In a similar manner, we also estimate the log wage equation given in (1) and (2), again employing recursive maximum likelihood. In this second estimation, Heckman correction is employed against selectivity bias arising from endogenous work decision as well. Note that in view of the model above, the impact of job change on wage cannot be identified separately from the wage impact by match specific productivity. That is why we need a separately estimation on log wage together with work, migration, and relocation decisions.

4.2 OLS and fixed effect regressions

Table 7 collects our first try[9]. The full details of the regressions are relegated to Appendix B, which is available upon request from the author. First thing we notice is a large and highly significant effect of *jno* (job number) variable. The estimated coefficients on *jno* suggest that at least each job change increases log wage by .1. Moreover, the impact estimates are larger in random or fixed effect regressions, than pooled OLS, indicating possible negative cross section correlations of log real wage with job hoppings. Fixing individuals, our regressions show that each job change brings about more than .2 log points increase. As we see in more details below, the large difference of the estimated coefficient on this variable indicates that sample individuals are highly heterogenous in terms of job mobility.

As expected, being an immigrant has a significant negative impact when compared to those born in PC urban labor market, whereas among those born in areas other than PC urban, the same variable has significant positive effect. The impact of cumulative experience on wage is highly significant and positive in pooled OLS but not in panel regressions, suggesting again that OLS is likely to suffer from endogeneity (and selection bias) of work decision.

As we outlined above, we use these regression results to construct our measures of match specific productivity: we take the difference between OLS residuals and individual fixed (random) effects to construct four measures of $\widehat{\eta}^{ik}$ for each subset of samples.By construction of our measures of $\widetilde{\eta}^{ik}$, we have

$$\begin{aligned} w_t^{ik} &\equiv \widehat{w}_t^{FE} + \widetilde{\eta}_t^{ik} + e_t^{ik} - \epsilon_t^{ik} \\ &\equiv \widehat{w}_t^{FE} + \widetilde{\eta}_t^{ik} + \nu_t^{ik} \end{aligned}$$

If we use $\widetilde{\eta}^{ik}$ directly in the next stages of regressions on wage, we are close to regressing an identity. To avoid this problem, we regress our measure of match specific productivity, $\widetilde{\eta}^{ik}$, on the following set of variables: job tenure, answers for multiple choice questions on why the respondents took the job,

[9] See Appendix Table A7 for the details of the variables used throughout the regressions results reported below.

answers for multiple choice questions on why the respondents eventually quit the job, year dummy, job number dummy, and a set of variables related to individual attributes and family background materials[10]. We denote the fitted value of $\widetilde{\eta}^{ik}$by $\widehat{\eta}^{ik}$.We only note here that the regression results explain about .3 to .4 of the total variances of each measure.

4.3 Main results

In this and next sub sections, we report our major regression results. In what follows, we only report the results using η^{FE2} for PC urban samples, and η^{FE4} for non PC urban born samples[11]. Consult *Table R* for how different sets of estimations are employed for our five main variables: $\Delta \log(wage)$, *work*, *quit*, *immigrate*, and *relocation*. Regressions for each of these dependent variables are either a part of the system of estimations (e.g., Recursive Maximum Likelihood Heckman correction, standard Heckman correction), or in single equation (OLS, fixed effect, or probit), *Table R* is provided for your reference to see where these results are found. To facilitate comparison among regression results for each dependent variable, we consolidate the results on each dependent variable from different estimations into one Table.

As we have shown above, given that both job mobility and wage level is driven by the match specific productivity, we cannot identify the effect of job to job quit in the wage equation. Therefore, we need to estimate two sets of regressions. In one, we use RML to estimate an expanded Heckman correction model wherein we estimate wage growth, job to job quit, relocation and migration. In the second set, we estimate a system comprised of log wage, work, relocation and migration.

As explained above, we assume that migration decision feeds into relocation and quit decisions, but not the other way around. Similarly, we assume that migration decision feeds into job to job quit decision. Given the relative stability of migrant status, we believe this is a sensible restriction. In essence, the model is an augmented Heckman correction model in that both quit and work decisions censor wage offers, and they are influenced by relocation and migration decisions. A brief explanation for each variables used in these regressions can be found in Appendix Table A7.

4.3.1 Recursive Maximum Likelihood Estimation of Heckman correction model for wage growth

Since the rejected job offers are not observable, we use Heckman correction model to account for the censored wage offers not accepted. Our first main results are in *Tables 8* and *9* wherein we have an augmented Heckman correction model for the changes in log wage after job changes in which we include *log wage growth, job quit, relocation and migrate* as endogenous variables and

[10] These regression results as well as full details of regressions reported below are in Appendix B which is available from the corresponding author upon request.

[11] See Appendix B for other cases employing alternative proxies for match specific productivity, as well as full regressions results. Appendix B is available upon request from the authors.

exploits the recursive structure, using a *STATA .ado* file called *cmp.ado*[12]. *Table 8* summarizes the results for wage growth regressions. First thing that we notice is the large and significant impact of our measure of the match specific productivity (and the one for the last job), $\widehat{\eta}_t^{ik}$ on wage growth. As expected, it has a very large positive impact whereas the one in the previous match has a large negative impact. In line with a simple search model, after controlling for the match specific productivity, Heckman correction terms (Inverse Mills Ratio at the bottom of the Table) are never significant and quantitatively very small except in the first equation for each sub samples without the match specific terms. In *Table 8* we also include results of OLS regressions. In all of these alternative estimation methods and specifications, $\widehat{\eta}_t^{ik}$ is always highly significant. Thus we conclude from these results that the improvement in match specific productivity is indeed the dominant driving force inducing job mobility and wage growth. Aside from the dominant impact of match specific productivity, other variables are never consistently significant under alternative specifications and they are quantitatively small.

In *Table 9*, we have the estimated impact of job change on log wage. Although the estimated impacts in these regressions show that the impacts do vary widely, their mean estimates are surprisingly close each other and centered around .27~.28.

Another piece of evidence consistent with our model is shown in Table *9,* where we tabulate means of estimated match specific productivity at each job number. With possible exceptions of fifth or later jobs, means of match specific productivity monotonically increase over job number. If recall our earlier characterizations of wage growth across jobs in *Figure 2*, these regression results allow us to interpret the pattern in reflection of the improvement of match specific productivity as they continue to move jobs. We come back to this point later on when we analyze wages paid at the current (thus ongoing) jobs.

4.3.2 Quit decision

Our key probit regression for job to job quits are consolidated in *Table 10* wherein we collect job quit probit regressions in RML Heckman correction model for wage growth and single equation probit under alternative specifications. These two types of regression results are qualitatively quite similar, although the estimated coefficients tend to be more statistically significant in RML Heckman correction models.

Like in wage growth regressions, our main variable representing match specific productivity is always highly significant and positive for the new job and negative for the previous job that they quit. Pseudo R^2 at the bottom for single regressions show the dominant explanatory power of these two vari-

[12] The cmp.ado file is prepared by D.M. Roodman and available at http://ideas.repec.org/c/boc/bocode/s456882.html

See also: Roodman, D. (2009) "Estimating fully observed recursive mixed-process models with cmp," Center for Global Development. Washington, DC, working paper 168.

ables as the measure jumps from less than .05 to .5 to .6., confirming again the decisive impact of the match specific productivity. Our main identifying variable for the quit decision, *varatio,* the ratio of job vacancies to job seekers is highly significant and positive[13]. The quit decision is highly sensitive on the job availability within the region. Another noteworthy point is the strong positive impact of relocation, indicating the close correlation between job and geographical mobility. The large impact of relocation on job quit decision is, somewhat tautological because, in some cases, young employees live in company dormitory, especially if they are migrant workers living away from their parents[14]. Thus relocation and job to job quit is a single event. But this is not all. In many cases, relocation precedes job quit, suggesting that they relocate themselves and start searching for jobs. Another noteworthy point is the impact of migrant variable. The coefficients are unstable but tends to be negatively significant in PC urban samples (the top half of the Table), whereas the sign is positive and often significant in the case of samples born in non PC urban regions. At least in PC urban labor market, being a migrant worker by itself tends to reduce labor mobility. Consequently, the fact that migrant workers do exhibit higher job mobility than the resident in PC urban labor market must be attributed to their higher spatial mobility which enhances job mobility greatly, as we see below.

Quit and gaps in career We have 2,623 records of job terminations, out of which 2,101 resulted in a next job, and the rest, 522, resulted into either unemployment, non participation, or short time job lasting less than six months, or any combination thereof, which we do not know. We only know that in those 522 cases, there exist lapses in between jobs lasting more than six months.

We found no evidence that these two types of job terminations made any substantive differences in subsequent career or wage growth. In particular, when we regress log real wage including a dummy which distinguishes tow types of terminations from the previous jobs, we found the dummy never statistically significant and the estimated coefficient is small in magnitude[15]. This is confirmed also in *Table 8* (*pjobloss*). The decision to quit a job either immediately into next job or something else does not seem to exhibit any substantive difference.

Our tentative conclusion from these two observations is that temporary career break does not matter very much. Other interpretations are certainly possible, however. Since our records shows job records only lasting six months

[13] Liu Yang allowed us to use her compiled data that she used for her own research (Liu forthcoming).

[14] According to 2009 survey of migrant workers conducted by government, 43% of young (19-24) migrant workers live in company dormitory, 15% in construction cite (where they work) or inside the production facility (of the employer), whereas only 36% live in rented apartments.

[15] Regressions results demonstrating these points are in Appendix B available upon requests from the corresponding author.

or longer, there is no way to tell to what extent these interruptions are real or representing simply short job spells. Given the fact that even among jobs lasting more than six months, the majority of them do not go beyond one year, we should expect that many of these apparent interruptions are not real. But if so, then, most of them do quit jobs to find another one within a very short period of interruption or none at all.

4.3.3 Recursive Maximum Likelihood Estimation of Heckman correction model for log wage

In the second set of RML Heckman correction model, we incorporate work decision in the estimation.

Our main results of log wage regressions are shown in *Table 11.* The results are remarkably robust under alternative specifications. Our main findings can be summarized as follows. First of all, as consistent with results in wage growth estimations, match specific productivity is highly significant and exerts quantitatively large impact on the log wage. The point estimates show that one standard deviation in $\widehat{\eta}^{ik}$ (.13-.19) changes the predicted real log wage anywhere between .08 to .29. The impact of *migrant* variable differs in two subset of samples. In comparison with non-immigrant born in regions other than PC urban, the impact is consistently positive and the estimated coefficient suggest that the immigrants earn 7 to 10% more. In regressions for samples currently in PC urban labor market, the impact is negative and significant, indicating that the immigrant earns 10 to 12% less compared to those born in PC urban. We come back to these points later on when we asses the results for immigration decision.

Another noteworthy point is the consistently positive impact of *firstjob* which is a dummy variable equal to one if the first job is a regular full time position.

Importance of family and individual background factors Aside from wage and job mobility, RML estimates probit model for three other endogenous variables, *work*, *relocation*, and (to become) *immigrant.* Before we summarize the regression results for these variables, we note here the importance of family and individual background as the key factor shaping the environment in which these decisions are made. The economic well being of the family has consistently negative impact on all of these variables. The impact is negative if the parents' annual income is higher (*rich*), and also if the sample individual is the only child. GPA in the last year at high school has consistently positive impact on all the variables. The sample individuals are more likely to work, immigrate and relocate if the father is dead. The willingness to take risk (*risk*) is also positive and highly significant in all three regressions. All in all, strong economic motivation, good education back ground, and willingness to take risk are the major inducement to work and mobility.

Work decision The results are collected in *Table 12* which includes results of single equation probit. First point we note is that the Heckman correction model estimated show consistently the positive selectivity bias (in work decision) as shown in significantly positive inverse Mills ratio (see the bottom of *Table 11*).

Not surprisingly, migrants are more likely to work. As we noted above, work decisions also depends highly on the economic needs of the family: samples are more likely to work if they are from not wealthy family, and if their fathers deceased [not shown in table] . On the other hand, positive impact of GPA, and negative impact of *nostudy* [not shown in table] variables are consistent with the positive selection results.

It is interesting but somewhat difficult to interpret that the impact of gender differences in two sub samples. In PC urban labor market, male is more likely to work, whereas in non PC urban born samples, females are more likely to work. One related point in the regression result is the negative impact of *familybusiness,* which indicates that they had to help family business during the last year in school. This might explain why male in non PC urban samples are less likely to work.

Relocation decision The results are collected in *Table 13.* The first thing we notice in the probit regressions for relocation is highly significant and quantitatively large impact of *migrant*[16], in comparison with non-immigrants or PC urban born samples. Moreover, impacts of family and individual background variables are qualitatively similar to the results for *migrant* probit regressions: relocations are more likely if they are female, willing to take risk, comes from not wealthy family, with many siblings. Impacts of variables related to high school experiences are also similar in many variables such as GPA.

On the other hand, some of the results are indicative that relocations are also related to personal troubles or failure to get along socially, which we do not find in immigrant probit models. For example, having some work experience during the high school has positive impact on immigrant, but negative on relocation. Having experiences of being bullied at school has positive impact on relocation.

4.4 Selectivity issues

As we have explained with some details, work decision, earnings, job and geographical mobilities are not only endogenous but we may also encounter potential selectivity issues. As we have seen above, our wage data is censored because work decision itself is endogenous and we do not observe rejected wage offers. On top of this, our wage data is one observation only for each

[16] Zhang (2010) argue that his estimation of a Cox proportional hazard model indicates that migrant workers are less likely to change jobs, which is consistent with our results in Table 14. Our results in Table 17 shows that migrant workers are far more likely to relocate and then find a new job. As Zhang uses a panel of workers living (working) in selected cities, geographical mobility across cities cannot be analyzed in his data.

job, thus we cannot directly observe within job wage changes. Finally, given uneven developments across different regions within China, the selection on where (which region) to work is also an important determinant of the wage and other employment conditions. Thus we have selectivity issues in this aspect as well in that we do not observe employment opportunities available in regions other than the one actually chosen by sample individuals. Thus decision to migrate induces selectivity bias. In what follows, we investigate the ramifications of these selectivity issues in some details.

4.4.1 Within job wage growth

The mirror image of the selectivity due to job quitting is the selectivity of retained jobs. As we repeatedly noted, we have no information on within job wage changes. For the sake of argument, let us assume that the wage reported for each job corresponds to the wage at the end of the tenure. If job termination is exogenous, our data on the length of tenure and the wage can be used to estimate the impact of tenure on wage, i.e., the estimate of within job wage growth. Among the completed job spells, as we have shown above, each job tends to end sooner if the wage rate is lower. Thus the survivor bias should favor the wage for jobs that lasts longer, thus generating survivor bias in the tenure effect estimation. With this caveat in mind, *Table 14* shows our regression results. We use two sub samples. First four columns show the results using the completed job spells, and the rest of the Table display the results using the current jobs (i.e., jobs retained as of the survey time). These latter estimates show the tenure impact on wage is highly significant and ranges between .046 to .052. On the other hand, the estimates for completed job spells are small and never statistically significant, which makes sense given the fact that they quit the job already. These two contrasting effects of tenure is consistent with the presence of survivor bias. Hence we conclude from these regressions that the within job wage growth is not higher that the estimate using the current (thus ongoing) jobs.

Therefore, we use the point estimate of the impact of tenure on log wage using the current job spells (the one with the check mark) to construct a hypothetical log wage as follows:

$$\log \widetilde{w}_t^{ik} = \log w_t^{ik} + .0222544 tenure_t^{ik}$$

As tenure is measured in terms of 1/3 of a year, the estimated impact translates to roughly 7% increase per year of the wage[17]. As we indicated above, we believe that this estimate is significantly upward biased reflecting the strong survivor bias as our samples in this regressions are taken from the current (and thus on going) jobs.

Our hypothesis is that first, the wage reported in all the completed jobs

[17] We consider this tenure effect way above the most likely magnitude. Our point here is to show that the impact of job to job quit is dominant factor even if we assume a very high within job wage growth.

refer to the wage paid at the end of the job[18].

The estimated mean impact of job to job quit on log real wage is shown at the bottom of *Table 9.* The patterns are qualitatively similar to the one we obtained in our main regression results reported above. Naturally, the estimated mean impacts are smaller but still large, ranging between .16 to .21.

4.4.2 Migration decision

We conducted a similar estimation using treatment model. The idea is that those who are born in regions other than PC urban have choice on migration. In *Table 15,* we find in all regressions the mean predicted value of log wage is higher for the migrants and after controlling for the endogeneity of migration decision, and the net impact is between .1 to .15 in log points. The Table also reports Heckman selection model in which we treat migration as the censoring device in the sense that the migrants face different labor markets. We also ran a mirror image Heckman model for non-migrants The difference in predicted log wages conditional upon migrant and non-migrant is shown in the table and the mean difference is somewhat larger than .15. In both treatment and censoring models, the estimated impacts of migration is positive and that they are positively selected.

We conclude from these estimates that migrants are indeed positively selected in the sense that those who immigrate do earn more on average, even after controlling for the selectivity due to endogeneity of immigration decision.

Probit regressions in both treatment and Heckman corrections model are qualitatively almost identical and the results are summarized as follows. We find significant positive impact on becoming a migrant by: *female*, willing to take risk (*risk*), comes from not wealthy family (*rich),* with many siblings. These family back ground variables are robust and suggest that those becoming migrant are raised in relatively poor family with many siblings. Although GPA has positive impact, we find those who answered did not study at all at home (*nostudy*) has strong positive impact. On the other hand, the self evaluation of over all high school life (*schlife*) has strong positive impact, and the experience of extended absence from school (for reasons other than illness) carries strong negative effect. These results indicate rather strongly that migrants are more forthcoming and willing to take risk, sociable, if not very smart or academically motivated strongly. Note also the impact of the success of the first job also has positive impact on becoming a migrant

4.4.3 Migrant workers in urban labor markets

In contrast, the impact on log wage by *migrant* variable in samples of PC urban residents carry strongly negative impacts. Although the estimated coefficients vary somewhat across different specifications, the results in *Table*

[18] We could assume alternatively that the reported wages are at the beginning of the job, or mean of the wage during the job. The substantive results are not sensitive to these specifications as in either way the tenure effect tends to favor the current job over the new job offers.

11 show the impact is statistically significant and around negative .1.

Some of the past studies indicated that the negative impact on wage by being a migrant worker is due to the fact that the migrants are more willing to take up (low paying) jobs which urban residents are not. Although our results indeed show that the estimated negative impact is somewhat smaller once we include our estimated match specific productivity, the impact of *migrant* variable remains negatively significant. The estimated impact is also somewhat smaller (but still remain significant) in the absolute magnitude in recursive maximum likelihood model that incorporates work decision. Thus we conclude from these results that indeed the migrants are penalized in their wage in PC urban labor market compared to their original residents.

One final question we ask ourselves is, so, what is our bottom line considering the impact of job and location changes? Do the immigrants catch up or overtake the urban residents in terms of wages they earn in urban labor markets?

Table 16 shows that our estimation results indicate that they do catch up: the estimated wage gap in the first job is .088 log points, but by the fifth or later job, the difference is reversed and the immigrants on average earns more than PC urban residents by .039 log point[19]. As we have seen above, the driving force of the job mobility is the match specific productivity. The wage growth by job mobility generates improvement in match specific productivity. To see this point, *Table 17* summarizes the Blinder-Oaxaca decomposition of log real wage difference between PC urban born and migrant samples in PC urban labor market[20]. Positive contributions in *Table 17* indicates that those factors favor native residents. Consistent with the results in *Table 16*, match specific productivity components (η) of the wage favors migrant workers, whereas the PC urban resident have a large advantage in endowment in individual fixed effect (μ) . The impact of employer types on the wage difference between the two groups suggest the source of possible discrimination of the migrant workers: residents are significantly more likely to be employed at government enterprises and they benefit strongly[21], whereas the net endowment impacts by foreign and private firms are negative on residents, thus favoring migrant samples. Most likely, the government related jobs are available only to the residents with urban registrations, whereas they do not matter in private or foreign firms. The total impacts of endowment, coefficients, and interactions are shown at the bottom of the table. Overall, coefficient terms are significantly favorable to the residents.

[19] Note, however, that standard deviations in parenthesis of Table 19 suggest that these differences are not statistically significant.

[20] Lee (2012) employs a similar decomposition method and finds roughly 10% wage differentials after controlling for the observable characteristics. Lee also finds that the discrimination impact is larger for male.

[21] A similar observation can be found in Gagnon et al (2009), in which they find the disadvantage applies to all the migrants, including those with urban *hukou* registration.

4.4.4 Return migrations

Among those born in non PC urban regions, 53% migrated to PC urban regions at least once. Among them, 21% (195) returned to their home province at least once. As of the last sample period, 143 or 73% of 195 remain in non PC urban regions, and the remaining 52 have returned to PC urban regions again. *Table 18* reports OLS regression on log wage and Heckman correction model for return migrants. The results show that return migrants seem just as successful as migrants are, in comparison with those born in non PC urban regions and remain in the regions. Selection model in *Table 19* confirms that return migrants are positively selected among those born in non PC urban regions, just as the migrant are overall. On the other hand, selection equation for return migrants differ in important ways from those for migrants. To begin with, the age obviously has a positive impact as returned migrants must have stayed some time in urban regions before the return. Positive evaluation of school life and GPA both have significant positive impacts on migrant, whereas the signs are reversed for returned migrants. On the other hand, the impact of having a regular full time first job has even stronger positive impact on returned migrant. We also note the impacts of rich and risk variables are both significant in two cases. It is perhaps fair to say that at least in our samples, return migration is just as successful on average as migrants, whereas we also detect some differences in the way some of individual attribute variables influence these two decisions. The latter may well reflect non-economic factors inducing them to return.

5 Conclusion

As far as we know, this is the first systematic investigations into the transition from work to school in contemporary China. We emphasized the interactions between geographical and occupational mobilities among the youth. We found that the youth migrating to urban areas after school are positively selected; they are more likely than the comparable urban resident youth to move geographically. They also expect to have large gains from job to job quits. The higher mobility pays off in terms of wage and job satisfactions. We find that job and geographical mobilities are highly correlated and they are also selective.

Setting aside the obvious and inevitable limitations of analyses based upon one shot surveys, this paper leaves many future research agendas. As we noted repeatedly, our survey collected single wage rate for each job. Thus our analysis on within job wage growth is highly speculative. Circumstantial evidence suggests that our findings cannot be applied directly to different segments of the Chinese youth, especially the college graduates as they are largely free to relocate themselves and move their resident registrations.

As we demonstrated in the paper, the labor market itself is in the midst of transformation. It is unclear to what extent and how long this wave of internal migration continues. Our characterization of the school to work transition in China may well be highly dependent on the current wave of massive immigration, reflecting enormous regional disparities within China.

References

[1] Ariga, K., Ohtake, F., Kurosawa, M., and Sasaki, M. (2009), 'How do high school graduates in Japan compete for regular, full time jobs? An empirical analysis based upon an internet survey of the youth,' *Japanese Economic Review, forthcoming*

[2] Chang, L.T., *Factory Girls,* Picador, 2008

[3] Chan, K.W. and W. Buckingham (2008), ' Is China Abolishing *Hukou* System?' *The China Quarterly*, 195: 582-606.

[4] deBrauw,A., and J. Giles (2008). "Migrant labor markets and the welfare of rural households in the developing world: Evidence from China,"

[5] Demurger, S., M. Gurgand, L. Shi, and Y. Ximing (2009), "Migrants as second-class workers in urban China? A decomposition analysis," *Journal of Comparative Economics* 37: 610-628

[6] Gagnon, J., T. Xenogiani, and C. Xing (2009), "Are all migrants really worse off in urban labour markets: new empirical evidence from China." OECD Development Centre Working Paper No. 278.

[7] Knight, J., and L. Yueh (2004), " Job mobility of residents and migrants in urban China," *Journal of Comparative Economics* 32: 637-660

[8] Kondo, A. and Ou, D. (2010), ' In search of a better life: The occupational attainment of rural and urban migrants in China,' ISER Discussion Paper793

[9] Lee, L. (2012),'Decomposing wage differentials between migrant workers and urban workers in urban China's labor markets,' *China Economic Review* 23: 461-70

[10] Lehmer, F. and J. Ludsteck (2008), "The returns to job mobility and inter-regional migration," IAB Discussion Paper 200806, Institute fur Arbeitsmarkt- und Berufsforschung (IAB), Nurnberg [Institute for Employment Research, Nuremberg, Germany].

[11] Li, S. (2008), 'Rural migrant workers in China: Scenario, challenges, and public policy,' ILO working paper 89

[12] Kennan, J. and J.R. Walker (2011), 'The Effect of Expected Income on Individual Migration Decisions,' *Econometrica* 71(1) : 211–251

[13] Matsumoto, M. and S. Elder (2010), ' Characterizing the school-to-work transitions of young men and women: Evidence from the ILO School-to-work transition survey,' ILO Employment Sector, Employment Working Paper 51

[14] Meiyang, W. (2007?) "Wage arrears for and discrimination against migrant workers in China's urban labor market,' Institute of Population and Labor Economics, China Working Paper 67

[15] Meng, X., and J. Zhang (2001), "The two tier labor market in urban China," *Journal of Comparative Economics* 29: 4845-504

[16] Meng,X.,K. Shen and S. Xue(2010), 'Economic Reform, Education Expansion, and Earnings Inequality for Urban Males in China, 1988-2007,' IZA Discussion Paper 4919

[17] National Bureau of Statistics of China (2010), *Monitoring report for migrant workers in 2009*

[18] Qu. Zhaopeng, and Zhong Zao (2011), "Evolution of the Chinese Rural-Urban Migrant Labor Market from 2002 to 2007," IZA discussion paper 5421

[19] Wu, Zheren (2010), 'Self-selection and Earnings of Migrants: Evidence from Rural China,' *Asian Economic Journal* 24(1): 23-44

[20] Zhang, H. (2010), 'The Hukou system's constraints on migrant workers' job mobility in Chinese cities,' *China Economic Review* 21: 51-64

6 Appendix Summary Statistics

6.1 Cross section characteristics

Appendix A additional *Tables A1 through A6* collect cross section characteristics of the sample individuals in terms of age, gender, education attainments, regions of the birth and the current residence, martial status, and labor force participation. The last table, *Table A7,* lists variables used in the regression analyses with a brief description of each variable.

Table A1 shows age composition. The sample individuals are more or less evenly distributed among age 21 and 27, with roughly 300-350 samples in each age bracket, but, significantly smaller shares for those with age 20 or younger and also 28 or older.

Table A2 summarizes the job placement services provided by the school and the participation to these activities. Naturally, vocational schools offer more job placement services than academic high schools. The participation rates do not differ much between the two types. In either types of schools, more than 60% of the sample attended all of these events organized by school. Among graduates of academic high schools, about one third of the sample said that the school provided none of the placement services we listed (or do not remember even if they did). As we have noticed in the Japanese survey in Ariga et al (2011), the graduates of academic high school seem to have problem in transition to work. *Table A3* shows education attainments and work experience. 25% of academic high school graduates never worked since graduation, as opposed to 7% for vocational schools, 13% for middle school only, and even for high school dropouts, the share is only 15%.

6.2 Families and social status

More than 90% of the respondents have both parents alive and married. Their joint income distributions are in *Table A4*. About a quarter of the sample report that their parents make more than 40 thousand RMB per year (about a half million yen, or 4,000US$). About 6% of the parents (either one of them or both) are ranking government officials. 13% of them are lower level government officials or employees of state owned firms. Single largest category of work is self employment, comprising roughly one third. It is a fair guess that the majority of them are farmers.

Table A5 shows the parents' ages at the birth of the sample workers. The mean age is 26 for mothers, and 28 for fathers. *Table A6* shows the distribution of the number of siblings. Somewhat surprisingly, about 14% of them have 3 or more siblings. 30% of them are the only child, and 38% have only one sibling. Given the fact that 63% of the sample individuals were born in rural areas, these results are consistent with the common perception that the birth control (so called *only child*) policy had not been strictly adhered to in the rural area.

As of the time survey was conducted in early 2009, about 40% of our samples live alone, with the rest living with one or both parents. Among the respondents at or above age 25, roughly a half of them live alone, indicating the typical age at which the youth become independent. About 30% of female respondents are married or divorced, whereas for male the share is 25%. In both sexes, 60% of the married samples have children, and about 9% of them do not live with their children.

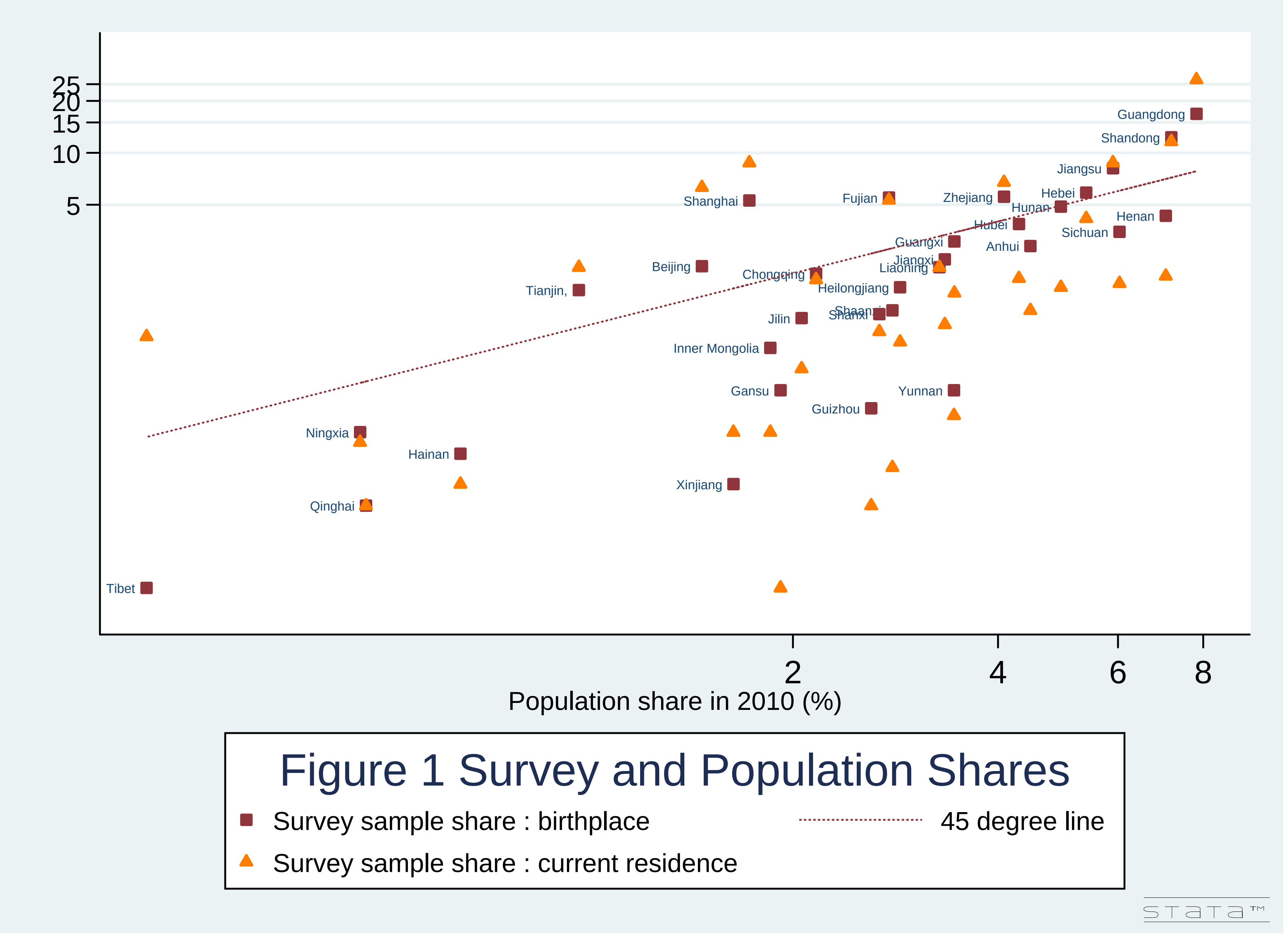

Figure 1 Survey and Population Shares

Figure 2 Wage growth across jobs

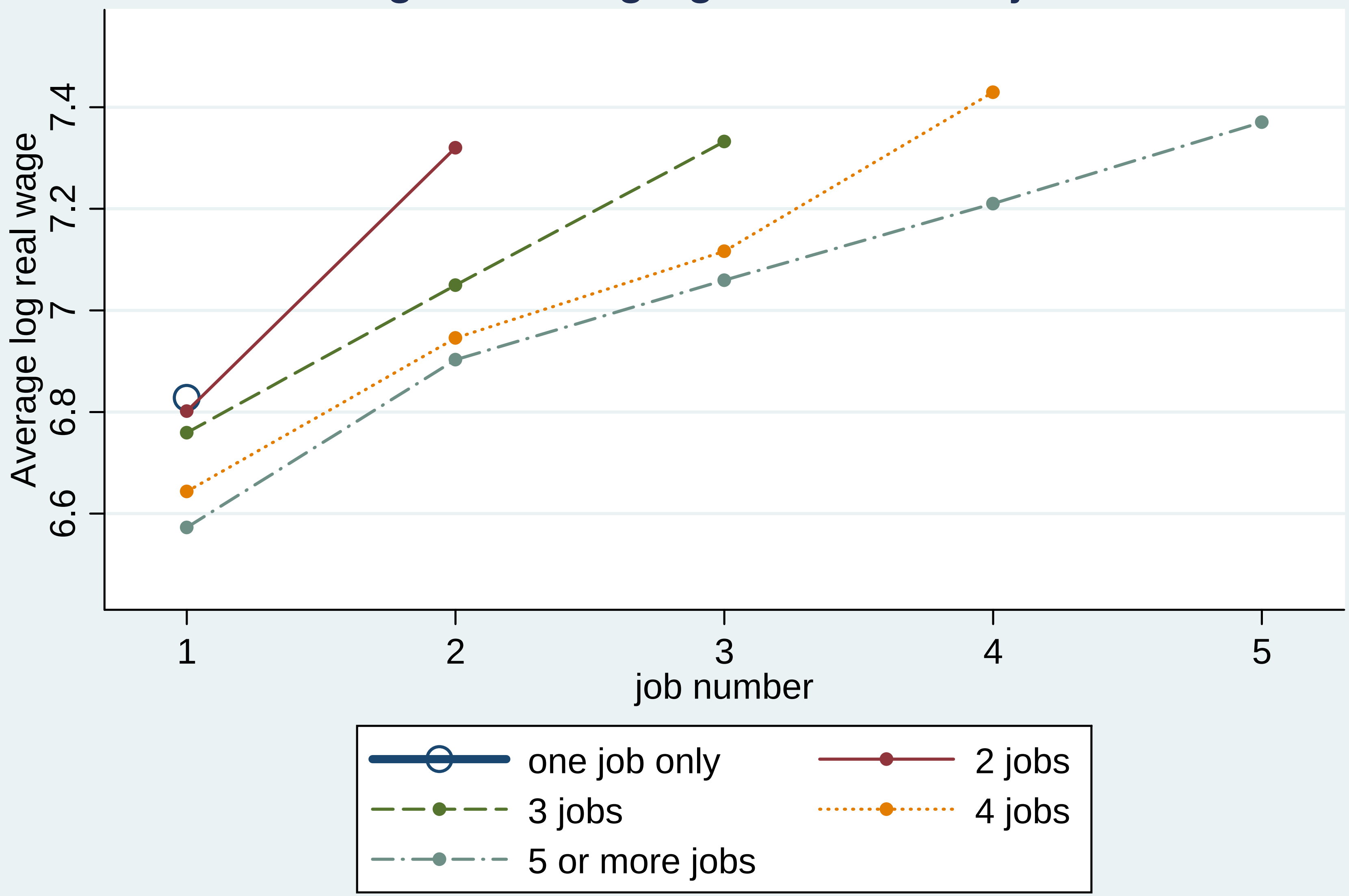

Figure 3 Wage growth of migrants and PC urban borns

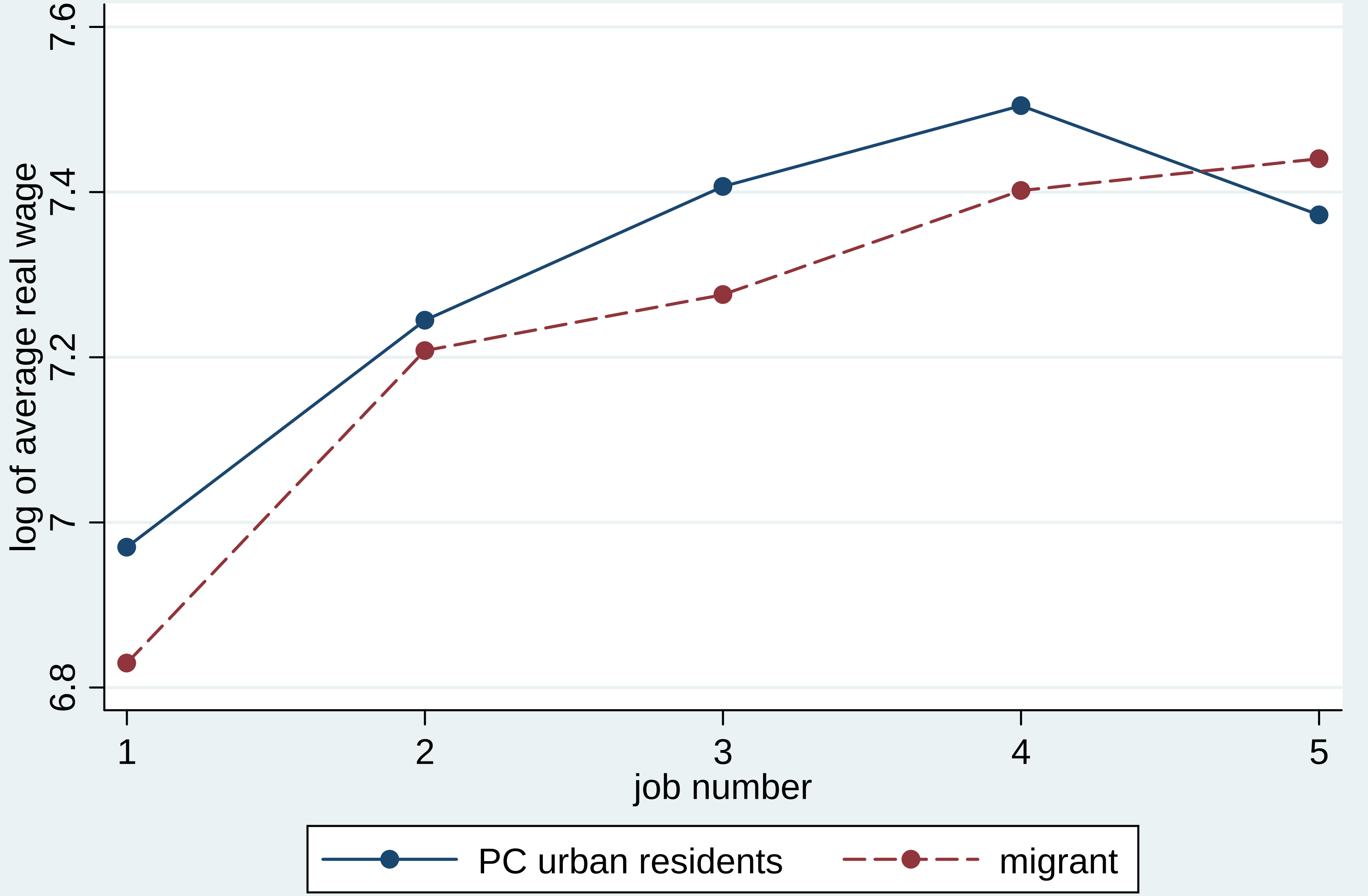

Table 1 Birthplace and current residence locations

	Current Pculocation				
Pculocation at birth	NonPC rural	NonPC urban	PC rural	PC urban	Total
NonPC rural	15.5	47.3	1.6	35.6	100
NonPC urban	1.1	72.7	0.0	26.2	100
PC rural	0.5	2.5	28.0	69.1	100
PC urban	0.0	2.0	0.6	97.3	100
Total	6.2	32.7	7.5	53.6	100

Table 2 Geographical mobility of migrant workers

Total number of locations lived	Among current residents in pc-urban regions		
	non-migrant	migrant	Total
1	70.74	29.26	100
2	47.11	52.89	100
3	51.12	48.88	100
4	41.65	58.35	100
5 or more	39.69	60.31	100
Total	54.96	45.04	100
	Among those born in non pc urban regions		
1	50.85	49.15	100
2	24.31	75.69	100
3	25.05	74.95	100
4	15.93	84.07	100
5 or more	17.71	82.29	100
Total	31.22	68.78	100

Table 3 Current Employment Status and wage

	middle school	academic high school	vocational high school	HS dropouts	Total
Sample Size	164	1,022	1,515	665	3,366
never worked	12.8	25.8	7.4	14.6	14.7
worked before, but not now	19.5	20.1	17.2	22.6	19.2
currently have a job	67.7	54.1	75.5	62.9	66.1
regular full time	59.2	45.6	63.3	50.2	52.2
Average wage (RMB per month) (current job)	966	1217	1227	1180	1204

Table 4 Across job wage growth

	education attainment				
Job #	middle	academic	vocational	dropout	total
First job	722	884	925	891	897
Second job	1049	1282	1358	1289	1308
Third job	1258	1547	1429	1418	1447
Fourht Job	1183	1669	1612	1512	1577
Fifth or later jobs	1316	1753	1556	1567	1561
Net % gain (%)					
Total	963	1145	1179	1133	1148

Figures are average monthly real wage in 2005 RMB prices

Table 5 Job characteristics and reasons for quitting the job

Reasons for quitting the job	Job history				
	1st job	2nd job	3rd job	4th or later	total
cannot get along with colleagues	1.8	2.3	2.3	0.0	2.0
the job not interesting	14.8	10.7	3.7	4.8	12.3
the job different from what I expected	8.5	6.4	5.3	9.7	7.6
too busy at work	6.6	7.3	15.7	5.2	7.7
dead end job, no future	30.5	29.8	21.6	11.4	28.9
wage too low	18.2	13.1	10.8	22.8	16.2
found a better job	15.2	22.9	33.9	38.6	20.0
others	4.3	7.4	6.6	7.6	5.5

Types of employers	Job history				
	1st job	2nd job	3rd job	4th or later	total
central and local governments	2.4	1.4	0.9	0.8	1.9
Institutions (e.g., schools, hospitals)	4.8	4.1	6.4	1.4	4.7
state enterprise	9.7	8.7	6.8	10.2	9.1
foreign firms	16.2	18.4	14.1	16.0	16.6
domestic private firms	47.4	54.8	55.4	61.7	50.7
self employment	15.8	11.2	13.7	9.1	14.1
helping family business	0.9	0.4	0.7	0.0	0.7
Others	2.9	0.9	2.1	0.8	2.2

Reasons for taking the job	Job history				
	1st job	2nd job	3rd job	4th or later	total
type of job that I was looking for	31.4	19.3	14.6	17.1	25.9
company well known	7.3	9.0	13.4	6.6	8.4
the pay is good	21.4	10.2	12.7	23.1	17.5
little overtime, many holidays	13.1	13.4	12.2	12.4	13.1
school education useful in the job	4.5	7.1	5.0	1.7	5.2
opportunity to learn professional skills	22.3	41.0	42.2	39.1	30.0

Table 6 Job changes and relocations

(1) Job change=0		
relocation	Frequency	Percent
0	25778	88.02
1	3509	11.98
Total	29287	100
(2)Job change=1		
relocation	Frequency	Percent
0	3326	72.16
1	1283	27.84
Total	4609	100

Table 7 OLS regressions on log real wage

Regression #	1	2	3	4	5	6	7	8
Samples	PC urban labor market				Samples born in non PC urban regions			
Regression type	OLS		Fixed Effects	Random Effects	OLS		Fixed Effects	Random Effects
migrant	-0.151	-0.143		-0.095	0.072	0.086	0.082	0.085
	(0.013)***	(0.015)***		(0.043)**	(0.010)***	(0.011)***	(0.045)*	(0.039)**
relocation	0.026	0.035	0.012	0.013	0.000	0.015	0.016	0.015
	(0.012)**	(0.012)***	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)*	(0.009)*
first job	0.131	0.123		0.082	0.069	0.070		0.116
	(0.020)***	(0.020)***		(0.054)	(0.016)***	(0.015)***		(0.046)**
cumexperience	0.016	0.017	-0.009	-0.002	0.016	0.015	-0.001	-0.001
	(0.001)***	(0.001)***	(0.006)	(0.004)	(0.001)***	(0.001)***	(0.003)	(0.002)
Job number	0.100	0.105	0.228	0.224	0.102	0.110	0.256	0.249
	(0.008)***	(0.008)***	(0.017)***	(0.017)***	(0.007)***	(0.007)***	(0.016)***	(0.015)***
Constant	6.096	6.026	6.150	5.624	5.992	5.937	6.212	5.676
	(0.130)***	(0.149)***	(0.339)***	(0.329)***	(0.103)***	(0.116)***	(0.215)***	(0.228)***
Other regressors included*								
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School characteristics and school life	Yes	Yes		Yes	Yes	Yes		Yes
Individual Attributes		Yes		Yes		Yes		Yes
Family Background		Yes		Yes		Yes		Yes
Observations	6,907	6,564	6,564	6,564	10,482	9,911	9,911	9,911
R-squared	0.254	0.296	0.435		0.248	0.285	0.478	
Number of samplenum			810	810			1,171	1,171
	Robust standard errors in pare			*** p<0.01, ** p<0.05, * p<0.1				
Match specific productivity	η^{FE1}	η^{RE1}	η^{FE2}	η^{RE2}	η^{FE3}	η^{RE3}	η^{FE4}	η^{RE4}
main results			yes				yes	
Computed as the difference in resdisuals in	#3--#1	#4--#1	#3-#2	#4-#2	#7-#5	#8-#5	#7-#6	#8-#6

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 8 Heckman correction model for wage growth

	PC Urban Labor Market			Non PC Urban born samples		
	1	2	3	4	5	6
Estimation Methods	RML Heckma	RML Heckma	Panel OLS	RML Heckma	RML Heckma	Panel OLS
Match specific productivity		η^{FE2}			η^{FE4}	
η(new job)		1.962	2.087		2.024	2.909
		(0.362)***	(0.288)***		(0.346)***	(0.387)***
η(last job)		-1.506	-1.655		-1.549	-1.419
		(0.368)***	(0.282)***		(0.338)***	(0.340)***
pjobloss	-0.019	0.009	0.018	0.010	0.035	0.038
	(0.042)	(0.048)	(0.050)	(0.035)	(0.041)	(0.042)
immigrant	0.000	-0.014		-0.021	-0.005	
	(0.039)	(0.038)		(0.110)	(0.072)	
relocation	0.064	-0.021		-0.059	-0.020	
	(0.156)	(0.140)		(0.484)	(0.161)	
first job	-0.034	0.052		-0.024	0.012	
	(0.037)	(0.040)		(0.032)	(0.033)	
cumexp	-0.017	0.002	-0.001	-0.014	0.006	0.021
	(0.004)***	(0.006)	(0.006)	(0.003)***	(0.005)	(0.007)***
Constant	0.853	0.731	0.729	0.486	-0.212	-0.529
	(0.420)**	(0.561)	(0.585)	(0.466)	(0.480)	(0.497)
Other regressors in the regressions	School type dummy, job characteristics (current and last jobs), year dummy					
Adj. R2			0.23			0.2102
Inverse Mills Ratio	0.545*	-0.011		-0.273	0.001	
	(0.294)	(0.033)		(0.155)	(0.028)	
Observations	9,902	5,775	5,775	15,173	8,625	

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

Table 9 Estimated mean impact of job to job quit on log wage

		PC Urban Labor Market		Non PC Urban born samples	
	Equation # in Table 8	1	2	4	5
	Match specific productivity		η^{FE2}		η^{FE4}
Tenure effect on wage ignored	*Estimated wage growth conditional on quit>0(standad deviation)*	0.087	0.282	0.185	0.281
		(0.219)	(0.198)	(0.175)	(0.206)
	E(η)				
	first job		0.044		0.015
	second job		0.131		0.126
	third job		0.176		0.170
	fourth job		0.185		0.209
	fifth or later		0.163		0.293
Tenure effect on wage included	Estimated wage growth conditional on quit>0(standad deviation)	0.186	0.178	0.203	0.212
		(0.171)	(0.196)	(0.169)	(0.184)
	E(η)				
	first job		0.003		-0.008
	second job		0.088		0.101
	third job		0.157		0.160
	fourth job		0.211		0.229
	fifth or later		0.171		0.345

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

Table 10 Job quit decision

	PC Urban				Non PC Urban born			
Equation # in Table 8		1		2		4		5
Estimation	Probit	Heckman RML	Probit	Heckman RML	Probit	Heckman RML	Probit	Heckman RML
η^{ij}			16.241	14.333			18.741	18.971
			(0.518)***	(0.550)***			(0.490)***	(0.679)***
η^{ij}_{j-1}			-14.958	-13.257			-16.707	-16.844
			(0.511)***	(0.526)***			(0.460)***	(0.621)***
varatio	0.017	0.137	0.061	0.112	0.059	0.115	0.387	0.342
	(0.048)	(0.054)**	(0.098)	(0.02)***	(0.047)	(0.058)**	(0.105)***	(0.108)***
migrant	0.018	-0.138	-0.034	-0.201	0.077	0.097	0.114	0.368
	(0.035)	(0.044)***	(0.073)	(0.066)***	(0.063)	(0.034)***	(0.061)*	(0.159)**
relocation3	0.130	0.491	0.294	1.337	0.284	0.435	0.261	0.785
	(0.035)***	(0.193)**	(0.064)***	(0.131)***	(0.395)	(0.034)***	(0.061)***	(0.216)***
first job	-0.025	-0.014	-0.037	-0.001	-0.026	0.027	0.049	0.021
	(0.032)	(0.037)	(0.064)	(0.066)	(0.026)	(0.034)	(0.061)	(0.062)
risk	0.010	0.019	0.010	0.007	0.004	0.007	0.015	0.007
	(0.007)	(0.009)**	(0.015)	(0.015)	(0.006)	(0.008)	(0.015)	(0.015)
age	-0.073	0.021	0.041	0.040	-0.084	0.005	0.071	0.075
	(0.007)***	(0.009)**	(0.015)***	(0.015)***	(0.006)***	(0.007)	(0.014)***	(0.015)***
Constant	0.352	-2.321	-2.662	-2.672	0.549	-1.991	-3.908	-4.030
	(0.173)**	(0.229)***	(0.379)***	(0.386)***	(0.135)***	(0.178)***	(0.345)***	(0.363)***
Other regressors in the regressions	Family background variables							
Pseudo Rsq	0.022		0.513		0.030		0.618	
Observations	10340	9,902	5791	5,775	16437	15,173	8643	8,625

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

Table 11 Recursive maximum likelihood estimation: Log real wage

	PC Urban labor market			Non PC Urban born		
Equation #	1	2	3	4	5	6
Estimation Method	OLS	Recursive maximum		OLS	Recursive maximum likelihood (RML)	
System of Equations		logwage, work, relocation			log wage, work, immigrant, relocation	
Match specific productivity			η^{FE2}			η^{FE4}
η			1.301			1.186
			(0.070)***			(0.072)***
first job	0.098	0.115	0.121	0.100	0.094	0.096
	(0.051)*	(0.019)***	(0.018)***	(0.043)**	(0.016)***	(0.016)***
migrant	-0.122	-0.102	-0.039	0.072	0.074	0.103
	(0.038)***	(0.013)***	(0.013)***	(0.037)**	(0.023)***	(0.023)***
male	0.023	0.031	0.015	0.075	0.077	0.103
	(0.031)	(0.012)***	(0.011)	(0.026)***	(0.010)***	(0.010)***
Acedmic HS	-0.084	0.023	0.031	-0.033	0.036	0.045
	(0.050)*	(0.018)	(0.018)*	(0.039)	(0.014)**	(0.014)***
Vocational HS	-0.027	0.050	0.034	0.024	0.080	0.067
	(0.045)	(0.017)***	(0.016)**	(0.037)	(0.013)***	(0.013)***
Middle School	-0.231	0.036	0.054	-0.159	0.041	0.052
	(0.070)***	(0.033)	(0.032)*	(0.061)***	(0.026)	(0.025)**
cumexperience	-0.002	0.014	0.035	-0.002	0.013	0.031
	(0.004)	(0.001)***	(0.002)***	(0.002)	(0.001)***	(0.001)***
Other regressors in the regressions	school type, job characteristics, job number dummy, year dummy					
Constant	7.249	6.318	5.700	6.796	5.404	5.075
	(0.256)***	(0.130)***	(0.133)***	(0.191)***	(0.213)***	(0.213)***
Inverse Mills Ratio		0.532	0.533		0.586	0.581
		(0.144)	(0.146)		(0.158)	(0.148)
Observations	6,907	10,322	10,322	10,482	16,434	16,434
Number of samplenum	851			1,238		

Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

Table 12 Work decision

	PC Urban labor market						Non PC Urban Born					
Equation Number in Table 11	2	3					5	6				
Estimation Methods	RML logw	RML logw	Probit	Probit	Probit	Probit	RML logw	RML logw	Probit	Probit	Probit	Probit
migrant	0.128	0.145	0.150	0.149	0.126		0.384	0.389	0.247	0.252	0.230	0.276
	(0.033)***	(0.033)***	(0.032)***	(0.035)***	(0.036)***		(0.080)***	(0.082)***	(0.022)***	(0.023)***	(0.023)***	(0.024)***
age	-0.043	-0.044	-0.050	-0.045	-0.022	-0.021	-0.048	-0.049	-0.056	-0.048	-0.016	-0.014
	(0.005)***	(0.006)***	(0.005)***	(0.006)***	(0.007)***	(0.007)***	(0.004)***	(0.004)***	(0.004)***	(0.005)***	(0.005)***	(0.005)**
male	0.207	0.186	0.181	0.197	0.206	0.209	-0.096	-0.102	0.022	-0.014	-0.026	-0.034
	(0.032)***	(0.032)***	(0.030)***	(0.033)***	(0.034)***	(0.035)***	(0.027)***	(0.028)***	(0.024)	(0.027)	(0.027)	(0.027)
schlife	0.019	0.014	0.021	0.018	0.019	0.018	0.012	0.016	0.024	0.023	0.027	0.025
	(0.015)	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.011)	(0.012)	(0.013)*	(0.013)*	(0.013)**	(0.013)*
bullied	-0.068	-0.072	-0.099	-0.124	-0.130	-0.129	0.063	0.055	-0.004	-0.006	-0.013	-0.011
	(0.027)**	(0.027)***	(0.029)***	(0.030)***	(0.031)***	(0.031)***	(0.021)***	(0.021)***	(0.023)	(0.024)	(0.024)	(0.024)
hutoko	-0.103	-0.098	-0.098	-0.113	-0.093	-0.128	-0.032	-0.034	-0.028	-0.030	-0.028	-0.053
	(0.026)***	(0.027)***	(0.030)***	(0.030)***	(0.031)***	(0.031)***	(0.021)	(0.021)*	(0.023)	(0.023)	(0.024)	(0.024)**
GPA	0.038	0.112	0.075	0.079	0.075	0.082	0.011	0.059	0.016	0.021	0.020	0.023
	(0.017)**	(0.017)***	(0.018)***	(0.018)***	(0.019)***	(0.019)***	(0.012)	(0.013)***	(0.013)	(0.014)	(0.014)	(0.014)*
parttimejob	0.065	0.074	0.047	0.035	0.025	0.016	0.002	0.008	-0.020	-0.013	-0.007	-0.019
	(0.031)**	(0.032)**	(0.035)	(0.036)	(0.037)	(0.037)	(0.024)	(0.024)	(0.027)	(0.027)	(0.028)	(0.028)
familybusiness	-0.035	-0.047	-0.105	-0.115	-0.111	-0.106	-0.054	-0.059	-0.127	-0.133	-0.127	-0.115
	(0.030)	(0.030)	(0.033)***	(0.034)***	(0.035)***	(0.035)***	(0.022)**	(0.023)***	(0.024)***	(0.025)***	(0.026)***	(0.026)***
rich	-0.024	-0.022		-0.007	-0.002	-0.004	-0.034	-0.034		-0.015	-0.015	-0.014
	(0.007)***	(0.007)***		(0.008)	(0.008)	(0.008)	(0.006)***	(0.006)***		(0.006)**	(0.007)**	(0.007)**
Constant	1.095	0.770	0.808	0.748	-1.634	-1.607	1.384	1.208	0.936	0.750	-1.799	-1.988
	(0.162)***	(0.164)***	(0.155)***	(0.178)***	(0.225)***	(0.227)***	(0.128)***	(0.129)***	(0.123)***	(0.140)***	(0.175)***	(0.178)***
Other regressors in the regressions	School characterisitcs											
	Family background			Family background			Family background			Family background		
					year dummy						year dummy	
						bithplace D						bithplace D
Pseudo Rsq.			0.0348	0.0436	0.0994	0.1017			0.0354	0.0393	0.094	0.0963
Observations	10,322	10,322	9,753	9,753	9,753	9,753	16,434	16,434	15,365	15,365	15,365	15365

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

Table 13 Relocation

	PC urban					Non PC urban born				
Corresponding Equation# in	2	3				5	6			
Estimation Methods	RML logwage	RML logwage	Probit	Probit	Probit	RML logwage	RML logwage	Probit	Probit	Probit
first job	-0.064	-0.065	-0.073	-0.072	-0.079	0.017	0.017	-0.038	-0.038	-0.035
	(0.029)**	(0.029)**	(0.029)**	(0.029)**	(0.029)***	(0.023)	(0.023)	(0.023)*	(0.023)	(0.023)
migrant	0.301	0.302	0.348	0.298		0.539	0.531	0.236	0.216	0.430
	(0.036)***	(0.036)***	(0.034)***	(0.036)***		(0.078)***	(0.078)***	(0.027)***	(0.027)***	(0.033)***
age	-0.013	-0.012	-0.000	-0.003	-0.010	-0.020	-0.020	-0.004	-0.005	-0.012
	(0.006)**	(0.006)*	(0.006)	(0.006)	(0.006)	(0.005)***	(0.005)***	(0.005)	(0.005)	(0.005)**
male	-0.039	-0.040	-0.072	-0.046	-0.046	-0.167	-0.167	-0.152	-0.154	-0.167
	(0.033)	(0.033)	(0.030)**	(0.033)	(0.034)	(0.028)***	(0.028)***	(0.025)***	(0.027)***	(0.028)***
schlife	-0.031	-0.031	-0.029	-0.028	-0.030	-0.017	-0.017	-0.010	-0.006	-0.009
	(0.017)*	(0.017)*	(0.016)*	(0.017)*	(0.017)*	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
bullied	0.114	0.114	0.116	0.110	0.107	0.118	0.116	0.102	0.097	0.098
	(0.031)***	(0.030)***	(0.030)***	(0.030)***	(0.031)***	(0.025)***	(0.025)***	(0.024)***	(0.024)***	(0.025)***
hutoko	0.012	0.012	0.023	0.013	-0.010	-0.028	-0.028	0.004	-0.002	-0.045
	(0.030)	(0.030)	(0.030)	(0.030)	(0.031)	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)*
GPA	0.068	0.078	0.064	0.074	0.072	0.030	0.036	0.025	0.029	0.028
	(0.018)***	(0.018)***	(0.018)***	(0.018)***	(0.018)***	(0.014)**	(0.014)**	(0.014)*	(0.014)**	(0.014)**
parttimejob	-0.089	-0.089	-0.135	-0.121	-0.101	-0.045	-0.044	-0.058	-0.050	-0.043
	(0.036)**	(0.036)**	(0.035)***	(0.036)***	(0.036)***	(0.029)	(0.029)	(0.028)**	(0.028)*	(0.029)
familybusiness	0.037	0.036	0.069	0.053	0.042	0.036	0.035	0.053	0.041	0.046
	(0.033)	(0.033)	(0.032)**	(0.033)	(0.033)	(0.026)	(0.026)	(0.025)**	(0.026)	(0.026)*
rich	-0.022	-0.022		-0.021	-0.019	-0.037	-0.037		-0.031	-0.030
	(0.008)***	(0.008)***		(0.008)***	(0.008)**	(0.007)***	(0.007)***		(0.006)***	(0.006)***
risk	0.014	0.014		0.014	0.016	0.023	0.023		0.024	0.029
	(0.007)**	(0.007)**		(0.007)**	(0.007)**	(0.006)***	(0.006)***		(0.006)***	(0.006)***
Constant	-0.090	-0.155	-4.948	-4.806	0.387	-1.150	-1.183	-4.068	-4.212	0.665
	(0.816)	(0.815)	(0.449)***	(0.463)***	(0.833)	(0.730)	(0.729)	(0.320)***	(0.331)***	(0.646)
Other regressors in the regressions	School characteristics, Individual attributes									
				family background					family background	
					year dummy					year dummy
Pseudo Rsq			0.0425	0.048	0.0563			0.0346	0.0387	0.0532
Observations	10,322	10,322	10,250	10,250	10,250	16,434	16,434	16,350	16,350	16,350

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

Table 14 Impact of Tenure at job on Wage

Sample	Completed job spells				Current (Incomplete) Job Spells									
					PC Urban					Non PC Urban born				
									√					
migrant		-0.095	0.082	0.085	-0.040	0.049	0.063	-0.011	0.006	0.159	0.168	0.169	0.167	0.173
		(0.043)**	(0.045)*	(0.039)**	(0.049)	(0.044)	(0.045)	(0.044)	(0.046)	(0.038)***	(0.035)***	(0.035)***	(0.035)***	(0.035)***
relocation	0.012	0.013	0.016	0.015	-0.050	-0.076	-0.091	-0.080	-0.096	0.002	-0.018	-0.023	-0.018	-0.025
	(0.010)	(0.010)	(0.009)*	(0.009)*	(0.052)	(0.050)	(0.051)*	(0.051)	(0.052)*	(0.042)	(0.040)	(0.039)	(0.040)	(0.039)
cumexperience	-0.007	-0.001	-0.002	-0.000	0.004	-0.007	-0.008	-0.003	0.002	0.008	0.002	0.002	0.002	0.006
	(0.010)	(0.009)	(0.006)	(0.006)	(0.004)	(0.004)*	(0.004)**	(0.004)	(0.003)	(0.003)***	(0.003)	(0.003)	(0.003)	(0.002)**
tenure	-0.002	-0.002	0.000	-0.000	0.017	0.030	0.031	0.030	0.022	0.010	0.021	0.020	0.020	0.014
	(0.009)	(0.009)	(0.006)	(0.006)	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.005)***	(0.004)**	(0.004)***	(0.004)***	(0.004)***	(0.004)***
jno	0.222	0.217	0.256	0.249	0.003	0.078	0.085	0.066		-0.008	0.059	0.061	0.051	
	(0.041)***	(0.040)***	(0.032)***	(0.031)***	(0.027)	(0.025)***	(0.025)***	(0.025)***		(0.022)	(0.021)***	(0.021)***	(0.021)**	
Job characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual and School	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Uih	No	No	No	No	No	yes	yes	yes	yes	yes	yes	yes	yes	yes
Cross Product of Job	No	No	No	No	No	No	yes	No	yes	No	No	yes	No	yes
Constant	6.158	5.631	6.211	5.677	6.730	6.615	6.437	6.499	6.586	6.754	6.585	6.438	6.514	6.538
	(0.339)***	(0.328)***	(0.224)***	(0.235)***	(0.219)***	(0.232)***	(0.247)***	(0.231)***	(0.237)***	(0.190)***	(0.165)***	(0.176)***	(0.165)***	(0.165)***
Observations	6,564	6,564	9,911	9,911	679	593	593	593	593	1,008	954	954	954	954
R-squared	0.435		0.478		0.135	0.324	0.358	0.301	0.331	0.144	0.274	0.292	0.259	0.272

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

Table 15 Probit Regression on migrant

Corresponding Equation# in Table 11	5				6
Non PC Urban born	RML logwage	Treatment Model	Heckman	Heckman	RML logwage
Dependent Variable	migrant	migrant	migrant	non-migrant	migrant
Estimated impact on log wage	0.094	0.103	E(*y|imig*)-E(*y|nonimig*)=0.1650		0.096
	(0.016)***	(0.026)***			(0.016)***
first job	0.287	0.148	0.226	-0.186	0.290
	(0.023)***	(0.029)***	(0.025)***	(0.026)***	(0.023)***
age	-0.010	-0.007	-0.008	-0.005	-0.010
	(0.005)*	(0.007)	(0.006)	(0.006)	(0.005)*
male	-0.351	-0.293	-0.328	0.322	-0.350
	(0.025)***	(0.031)***	(0.028)***	(0.029)***	(0.025)***
schlife	0.093	0.121	0.113	-0.103	0.095
	(0.013)***	(0.017)***	(0.015)***	(0.015)***	(0.013)***
absence	-0.286	-0.190	-0.156	0.290	-0.282
	(0.042)***	(0.052)***	(0.047)***	(0.049)***	(0.042)***
GPA	0.013	0.055	0.037	-0.022	0.010
	(0.014)	(0.018)***	(0.016)**	(0.017)	(0.014)
parttimejob	0.103	0.145	0.108	-0.153	0.104
	(0.028)***	(0.035)***	(0.031)***	(0.032)***	(0.028)***
familybusiness	-0.033	-0.006	-0.056	-0.016	-0.039
	(0.026)	(0.032)	(0.028)**	(0.029)	(0.026)
rich	-0.049	-0.062	-0.056	0.054	-0.049
	(0.007)***	(0.008)***	(0.007)***	(0.008)***	(0.007)***
risk	0.015	0.036	0.029	-0.024	0.015
	(0.005)***	(0.007)***	(0.006)***	(0.006)***	(0.005)***
Constant	0.902	0.323	-1.347	-2.046	0.886
	(0.177)***	(0.310)	(0.228)***	(0.222)***	(0.177)***
Additional RHS	School Characteristics, Indibidual attributes and family background				
Observations	15,936	9,911	13,308	11,968	16,434

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

Table 16 Job Mobility and Predicted Log Wage*

	Mean of Predicted Log Real Wage		
Job number	samples born in PC uran	migrant	Total
First	6.793(.237)	6.705(.207)	6.732(.220)
Second	7.034(.209)	6.904(.245)	6.943(.242)
Third	7.113(.210)	6.996(.221)	7.027(.224)
Fourth	7.165(.241)	7.067(.195)	7.093(.212)
Fifith or later	7.168(.288)	7.207(.193)	7.196(.223)
Total	6.988(.264)	6.893(.263)	6.920(.267)

*The predictions are based upon the samples and RML estimate shown in equation(3) of Table 13 . Standard deviations in parenthesis

Table 17 Blinder-Oaxaca Decomposition of Wage Differences

η					η_{FE1}			η_{FE2}	
	Endowments	Coefficients	Interaction	Endowments	Coefficients	Interaction	Endowments	Coefficients	Interaction
μ	not included in regression			0.134			0.136		
				(0.012)***			(0.012)***		
η							-0.064		
							(0.007)***		
male	-0.004	-0.100	0.011				0.000		
	(0.001)***	(0.013)***	(0.003)***				(0.000)		
first job	-0.004	0.060	-0.005				0.002	0.027	
	(0.002)***	(0.025)**	(0.002)**				(0.001)**	(0.012)**	
middle school		-0.008		§					
		(0.003)***							
academic HS	0.002						0.001		
	(0.001)**						(0.001)**		
vocational HS		-0.066	-0.013				-0.000		
		(0.018)***	(0.004)***				(0.001)		
cumexperience					-0.020		-0.004		
					(0.008)**		(0.005)		
Regular Fulltime		-0.103	0.005	-0.004			-0.004	-0.026	
		(0.026)***	(0.002)**	(0.001)**			(0.002)**	(0.012)**	
Employer type: gov't				0.004			0.004		
				(0.002)**			(0.002)**		
gov. orgnizations	0.004	0.008		0.006			0.006		
	(0.002)**	(0.004)**		(0.002)***			(0.002)***		
gov't. enterprises	0.050	0.015	0.021	0.088			0.080		
	(0.008)***	(0.006)**	(0.009)**	(0.009)***			(0.008)***		
foreign firms	-0.042			-0.051			-0.048		
	(0.008)***			(0.008)***			(0.008)***		
private firms	-0.031			-0.047			-0.045		
	(0.007)***			(0.008)***			(0.008)***		
self employed					0.010		0.002		
					(0.005)**		(0.003)		
others					insig		-0.000		
							(0.000)		
unemploymentrate		0.370	0.039	0.015			0.014		
		(0.042)***	(0.005)***	(0.002)***			(0.002)***		
Constant		0.382							
		(0.215)*							
Total	-0.038	0.111	0.049	0.126	0.001	-0.001	0.067	0.061	-0.001
	(0.008)***	(0.013)***	(0.010)***	(0.013)***	(0.006)	(0.004)	(0.013)***	(0.007)***	(0.005)
PC Urban born		7.117			7.131			7.131	
		(0.012)***			(0.012)***			(0.012)***	
Migrants		6.994			7.004			7.004	
		(0.007)***			(0.007)***			(0.007)***	
Difference		0.122			0.127			0.127	
		(0.014)***			(0.014)***			(0.014)***	
Observations		6,907			6,564			6,564	

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

§ The estimates and standard errors are not shown unless siginificant at 5% level

Table 18 Comparing Return migrant and migrant samples

	OLS				Heckman Correction model			
	Migrants		Returned migrants		Migrants		Returned migrants	
Match specific productivity		η^{FE2}		η^{FE2}		η^{FE2}		η^{FE2}
η		1.487		1.507		1.201		0.691
		(0.069)***		(0.069)***		(0.094)***		(0.336)**
migrant	0.068	0.046						
	(0.015)***	(0.014)***						
returnmigrant			0.031	0.046				
			(0.026)	(0.025)*				
male	0.073	0.103	0.068	0.100	0.049	0.080	0.025	0.021
	(0.023)***	(0.022)***	(0.023)***	(0.022)***	(0.014)***	(0.014)***	(0.057)	(0.056)
first job	0.108	0.105	0.103	0.100	0.093	0.104	-0.121	-0.108
	(0.027)***	(0.026)***	(0.027)***	(0.026)***	(0.023)***	(0.023)***	(0.102)	(0.101)
middle school	-0.158	-0.136	-0.165	-0.139	-0.158	-0.136	-0.165	-0.139
	(0.054)***	(0.053)**	(0.054)***	(0.053)***	(0.054)***	(0.053)**	(0.054)***	(0.053)***
academic HS	-0.032	-0.024	-0.032	-0.024	-0.030	-0.020	0.025	0.081
	(0.032)	(0.032)	(0.032)	(0.032)	(0.019)	(0.019)	(0.079)	(0.082)
vocational HS	0.026	0.015	0.025	0.015	0.030	0.010	-0.150	-0.107
	(0.031)	(0.031)	(0.031)	(0.031)	(0.017)*	(0.017)	(0.068)**	(0.071)
cumexperience	-0.000	0.021	-0.000	0.021	0.015	0.032	0.028	0.036
	(0.001)	(0.002)***	(0.001)	(0.002)***	(0.002)***	(0.002)***	(0.007)***	(0.008)***
Constant	6.039	6.051	6.107	6.103	6.039	6.051	6.107	6.103
	(0.097)***	(0.092)***	(0.096)***	(0.090)***	(0.097)***	(0.092)***	(0.096)***	(0.090)***
Observations	10482	10482	10482	10482	10482	10482	10482	10482
*E(y|observed)-E(y|unobserved)***					0.1647	0.1647	0.1574	0.1571

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

See Appendix Table A7 for varaibles in the regressors. Full regression results are available upon request from the corresponding author

Table 19 Heckman Correction model Probit for Migrant and Returned Migrant

Heckman Correction modl: Selection probit for migrant or returned migrant										
	migrant					returnmigrant				
age	-0.009	-0.008	-0.009	-0.009	-0.009	0.041	0.041	0.041	0.041	0.041
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.013)***	(0.013)***	(0.013)***	(0.013)***	(0.013)***
male	-0.329	-0.328	-0.329	-0.328	-0.328	-0.119	-0.116	-0.118	-0.120	-0.120
	(0.028)***	(0.028)***	(0.028)***	(0.028)***	(0.028)***	(0.066)*	(0.066)*	(0.066)*	(0.066)*	(0.066)*
first job	0.227	0.226	0.227	0.226	0.226	0.642	0.642	0.642	0.642	0.642
	(0.025)***	(0.025)***	(0.025)***	(0.025)***	(0.025)***	(0.069)***	(0.069)***	(0.069)***	(0.069)***	(0.069)***
schlife	0.110	0.113	0.110	0.112	0.111	-0.127	-0.127	-0.127	-0.127	-0.127
	(0.015)***	(0.015)***	(0.015)***	(0.015)***	(0.015)***	(0.031)***	(0.031)***	(0.031)***	(0.031)***	(0.031)***
absence	-0.150	-0.156	-0.149	-0.154	-0.151	-0.134	-0.129	-0.133	-0.132	-0.131
	(0.046)***	(0.047)***	(0.046)***	(0.047)***	(0.047)***	(0.100)	(0.100)	(0.100)	(0.101)	(0.101)
GPA	0.039	0.037	0.034	0.042	0.044	-0.120	-0.119	-0.121	-0.114	-0.117
	(0.016)**	(0.016)**	(0.016)**	(0.016)***	(0.016)***	(0.035)***	(0.035)***	(0.035)***	(0.036)***	(0.036)***
parttime job	0.115	0.108	0.116	0.110	0.114	-0.102	-0.101	-0.101	-0.104	-0.101
	(0.031)***	(0.031)***	(0.031)***	(0.031)***	(0.031)***	(0.066)	(0.065)	(0.066)	(0.066)	(0.066)
family business	-0.052	-0.056	-0.051	-0.055	-0.053	0.147	0.150	0.148	0.149	0.149
	(0.028)*	(0.028)**	(0.028)*	(0.028)*	(0.028)*	(0.061)**	(0.061)**	(0.061)**	(0.061)**	(0.061)**
risk	0.029	0.029	0.029	0.029	0.029	0.070	0.069	0.070	0.069	0.069
	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.014)***	(0.014)***	(0.014)***	(0.014)***	(0.014)***
rich	-0.058	-0.056	-0.058	-0.057	-0.058	-0.096	-0.096	-0.096	-0.096	-0.096
	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.015)***	(0.015)***	(0.015)***	(0.015)***	(0.015)***
Constant	-1.323	-1.347	-1.307	-1.357	-1.351	-4.467	-4.467	-4.466	-4.485	-4.482
	(0.228)***	(0.228)***	(0.228)***	(0.228)***	(0.228)***	(0.454)***	(0.454)***	(0.454)***	(0.456)***	(0.455)***
Observations	13308	13308	13308	13308	13308	15072	15072	15072	15072	15072
Inverse Mills Ratio	0.065	0.021	0.068	0.035	0.056	0.310	0.322	0.315	0.288	0.289
	(0.024)	(0.022)	(0.024)	(0.024)	(0.024)	(0.090)	(0.090)	(0.090)	(0.092)	(0.092)

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A1 Age Composition

age	Female	Male	Total
16	0	1	1
17	2	4	6
18	17	21	38
19	58	69	127
20	106	132	238
21	130	206	336
22	161	193	354
23	134	206	340
24	140	182	322
25	118	173	291
26	144	184	328
27	132	194	326
28	101	136	237
29	93	158	251
30	55	101	156
31	3	12	15
Total	1,394	1,972	3,366

Table A2 Job placement services at high school

Job placement services	Academic	Vocational	total
0	337	245	582
	33	16	23
1	325	307	632
	32	20	25
2	205	379	584
	20	25	23
3	84	285	369
	8	19	15
4	35	160	195
	3	11	8
5	36	139	175
	4	9	7
Total	1022	1515	2537
	100	100	100

Table A3 Education atainements and work experience

	Ever worked?		
	yes	no	
Middle school1	143	21	164
	87.2	12.8	100
ademic Highsch	758	264	1,022
	74.17	25.83	100
cational Highsch	1,403	112	1,515
	92.61	7.39	100
gh school dropo	568	97	665
	85.41	14.59	100
Total	2,872	494	3,366
	85.32	14.68	100

Table A4 Parents' joint income			
	Freq.	Percent	Cum.
Less than 10,000RMB	639	19.11	19.11
10,000-20,000	848	25.37	44.48
20,000-30,000	636	19.02	63.51
30,000-40,000	384	11.49	74.99
40,000-50,000	230	6.88	81.87
50,000-60,000	175	5.23	87.11
60,000～80,000	124	3.71	90.82
80,000～100,000	116	3.47	94.29
100,000～150,000	111	3.32	97.61
150,000 or more	80	2.39	100
Total	3,343	100	

Table A5 Parents' ages at birth					
Variable	Obs	Mean	Std. Dev.	Min	Max
Father's age	3230	27.968	4.811	16	78
Mother's age	3266	25.914	4.453	14	78

Table A6 Number of borthers and sisters			
number of siblings	Freq.	Percent	Cum.
0	996	29.59	29.59
1	1,266	37.61	67.2
2	646	19.19	86.39
3	242	7.19	93.58
4	104	3.09	96.67
5	45	1.34	98.01
6	28	0.83	98.84
7	10	0.3	99.14
8	4	0.12	99.26
9	5	0.15	99.41
10	6	0.18	99.58
11	3	0.09	99.67
12	5	0.15	99.82
13	2	0.06	99.88
14	2	0.06	99.94
15	1	0.03	99.97
18	1	0.03	100
Total	3,366	100	

Appendix Table A7 Definition of variables used in the paper

Variable category	Name	Variable type	Biref explanation
Job and job search Characteristics	*logavw*	*V*	log of monthly wage converted from answers in 12 wage ranges
	logavw1	*V*	*logavw* in the previous period (measured in *t* defined below)
	logavwq	*V*	censored *logavw (* set missing unless the samples changed job in the previous period)
	logavnq	*V*	censored *logavw (* set missing if the samples changed job in the previous period)
	cumexperience	*V*	cumulative work experience since school graduation measured in *t*
	kinzoku	*V*	cumulative work experience at the current employer measured in *t*
	RegFulltime	*V*	Dummy for regular full time job
	Ftype1- Ftype7	*V*	employer type dummies
	route1- route5	*V*	dummies for the routes used to learn about the current job
	first job	*C*	dummy=1 if the first job after school was regular full time
	jno	*V*	number of jobs lasted 6 months or longer including the current job
	quit	*V*	dummy =1 if quitted a job in the previous period and start a new job in this period
	whytake	*V*	reason for takin up the current job (multiple choice)
	whyleave	*V*	reason for quitting the job (multiple choice)
	pjobloss	*V*	dummy==1 if the previous job was terminated by employer
esidence and relocations	*Lno*	*V*	Number of locations lived after school which lasted 6 months or longer, including the current
	pculocation pcuD1-pcuD4	*V*	location type variable (=1,2,3, or 4). See the main text (*pcuD1-pcuD4* are dummies)
	pculocation_b pcubD1- pcubD	*V*	location type variable for the birthplace
	pculocation_g pcugD1- pcugD	*V*	location type variable for the last school attended
	migrant	*V*	Dummy =1 if born in pculocation 1 to 3 but resides in 4
	relocation	*V*	dummy =1 if *Lno* (t) > *Lno* (t-1)
	relocation3	*V*	dummy=1 if *relocation* (t),(t-1), or (t+1)=1
School types, charactersitics and school life	*edu1 edu2 edu3*	*C*	school types (=1 middle school, =2 academic HS, =3 for vocational HS)
	public_imp public_com	*C*	High school types (_imp= targeted public HS, _com=other public HS. Default is private HS)
	sinroall	*C*	Number of job placement actitivities at HS
	GPA	*C*	Average score during the last year at HS
	BFGF	*C*	Dummy=1 if he (she) had a girl(boy) friend at HS
	schlife	*C*	Self evaluation of HS life (=5 highest,=1 lowest)
	bullied	*C*	dummy=1 if ever being bullied at any school attended
	familybusiness	*C*	helped family business (farm) during the school
	course1 course2	*C*	curriculum types dummies. _1 for college advancement, _2 for mixed, _3 for others
	tugaku1 tugaku2	*C*	living arrangements during HS _1 if commuted from home, _2 if stayed at dormitory, _3 othe
	leader1 leader2	*C*	dummy =1 if elected as a class leader (_1) or a member of student council at HS
	nostudy	*C*	dummy=1 if did very liitle study outside school
	lotstudy	*C*	dummy =1 if studied a lot outside school
	absence	*C*	dummy=1 if been absent or late to school frequently
	absence	*C*	dummy=1 if been absent for extended periods of time from school (reasons other than sicknes
	nofriend	*C*	dummy=1 if very few friends at school
	talkfriend	*C*	dummy=1 if talked with many friends at school
	sports	*C*	dummy=1 if spent fair amount of time in sports during HS
	extra curricular	*C*	dummy=1 if participated in extra curricular activities
	parttimejob	*C*	dummy=1 if part time jobs during school
Inidividual attributes	*male*	*C*	
	age	*V*	measured in years (not *t*)
	risk	*C*	dummy=1 if the sample agreed to a proverb praising risk taking attitude
	married	*C*	married
	femalemarried	*C*	female and married
Family background	*onlychild*	*C*	only child of the parent
	brothers	*C*	number of siblings
	firstborn	*C*	eldest sibling
	rich	*C*	The family was well to do when the sample individual was age 15
	nw_gov1	*C*	Parent(s) worked for (local) governement (managerial position)
	nw_gov2	*C*	Parent(s) worked for (local) governement
	nw_manage	*C*	Parent(s) have managerial job(s)
	nw_village	*C*	Parent(s) worked for viallge office
	nw_ins	*C*	Parent(s) worked for (local) governemental organizations
	birthorder	*C*	birth order among the siblings
	died_f	*C*	(biological) Father is dead
	died_m	*C*	(biological) Mother is dead
	divorce	*C*	Parents divorced
province level data and time dummies	*unemployment*	*V*	Provincial unemployment rate
	lnmeanwage	*V*	Provincial log average wage
	year3- year14 by3- by14	*V*	Year dummy (1996=year1)
	t	*V*	Number of periods after school graduation (3 periods=1 year) see the main text.
	varatio	*V*	The ratio of posted vacancies to job applicants in each province