China Economic Review 63 (2020) 101498

Contents lists available at ScienceDirect

China Economic Review

journal homepage: www.elsevier.com/locate/chieco

Return migration and *Hukou* registration constraints in Chinese cities

Jipeng Zhang[a,b,*], Jin Huang[a], Junhui Wang[a,b], Liang Guo[c]

[a] Research Institute of Economics and Management, Southwestern University of Finance and Economics, China
[b] Survey and Research Center for China Household Finance, Southwestern University of Finance and Economics, China
[c] Shanghai University of International Business and Economics, China

ARTICLE INFO

Keywords:
Return migration
Labor mobility
Hukou
Chinese economy

JEL classification:
O15
O18
P25
J00
C26

ABSTRACT

This paper documents the patterns of return migration and labor mobility constraints in China using two unique data: the 2017 China Household Finance Survey and a newly developed urban *Hukou* registration index. The size of return migrants is larger than that of migrants without local *Hukou* registration. Majority of return migrants move from more developed region back to their less developed home town where they have *Hukou* registration. Empirical results show that *Hukou* registration barrier, typically higher in more developed cities, leads to a higher probability of returning among low-skilled migrant workers, and such an effect only exits among migrants moving across provinces and migrants with rural *Hukou*.

1. Introduction

A vast literature on migration in China has been developed, however, very few papers have studied return migration using nationally representative data. More importantly, return migration is often mixed with *Hukou* transfer that leads to incorrect measure of return migration. This arises from the commonly used definition of migration in China that is measured by the separation of *Hukou* and living locations. The migrants who changed *Hukou* registration from one place to another are not counted as migrants in majority of the studies. This ignorance leads to significant under estimation of migrants and over estimation of return migrants. This study aims to fill this gap using the latest data from the 2017 China Household Finance Survey (CHFS). The uniqueness of this data is that it contains detailed information on both migration experience and *Hukou* registration changes over time, in addition to the detailed individual demographics and household finance information.

Return migration refers to the return of migrants from a working/residential location to the place where they are registered. Here we exclude the group who changed *Hukou* registration after migration. If a worker changes *Hukou* registration across prefectures, that worker would be mistakenly classified as return migrant because returning is usually measured relative to one's current *Hukou* registration location and ignoring Hukou transfer. This can lead to an overestimation of return migrants because the proportion of migrants with *Hukou* transfer is high. According to CHFS 2017, the ratio of cross-prefecture migration with *Hukou* transfer is 5.08%. Population with *Hukou* transfer are expected to have a faster increase. Starting from March 2014, China's Plan for New Urbanization (2014–2020) explicitly proposed that 100 million migrants should have their *Hukou* registered in the place they work.

* Corresponding author at: Research Institute of Economics and Management, Southwestern University of Finance and Economics, China.
E-mail address: jpzhang@swufe.edu.cn (J. Zhang).

https://doi.org/10.1016/j.chieco.2020.101498
Received 26 March 2019; Received in revised form 17 May 2020; Accepted 20 May 2020
Available online 02 July 2020
1043-951X/

The definition of migration relies on the administrative units. In this paper, we define migrants as those who move across prefecture boundary, but in the literature a common definition is based on township. For our research purpose, prefecture is a more appropriate defining units since household registration is usually managed by prefecture-level government. Throughout the paper we might use city referring to the urban area of a prefecture. For example, we use city for the working destination of migrants since most migrants work in the urban sector of a prefecture.

Return migration could be good or bad depending on the causes of returning. If migrants return for economic and social benefits under the condition of free choice, return can be a good thing in improving welfare and better allocation of labor across space. But if the returning is caused by institutional barriers such as *Hukou* registration constraints, and migrants are "forced" to going back to their hometown, returning could lead to significant welfare loss. In China, most migrants return to their hometown more likely because of Hukou registration constraint that reduces migrants' willingness to settle down in their working location. In this sense, return migration is bad for the welfare of migrants.

With the dynamic changes in population and economic structure, the allocation of labor resources between regions is very important for aggregate efficiency and welfare. The spatial mobility of labor, especially return migration, has a profound impact on economic decision making of individuals such as consumption, saving and human capital investment, but also on understanding the changes in spatial resource allocation, economic efficiency, as well as the impact of *Hukou* registration reform. The research of modern urban and regional economics (including economic geography) based on spatial equilibrium models already establish the impact of industry and transportation, as well as housing market on labor mobility and welfare. The findings on return migration could have important implications for China's regional development, including urbanization, and public policy reforms involving employment, education, housing, pensions and health care.

This paper analyzes the latest trends of three groups of migrants: floating population (migrants in common sense), return migrants, and migrants with *Hukou* transfer. Our data can measure return migration accurately, since it contains detailed information on migration time and destinations, working experience and location, and the date and location of recent *Hukou* transfer across cities.

Empirically, we combine the CHFS data with a newly constructed China *Hukou* Registration Index (CHRI) (Zhang, Wang, & Lu, 2019) of Chinese cities to investigate the causal relationship between return migration and urban *Hukou* registration constraints. CHRI divides the local *Hukou* registration policies into employment, talent program, investment (including taxation), and home purchase and then constructs corresponding *Hukou* registration index for 120 cities in two stages (2000–2013 and 2014–2017). The *Hukou* reform, in terms of intensity and scope, has made significant progress after 2013, hence our study focuses on the period of 2014–2017.

Our paper contributes to the literature on return migration. In the 1990s, China began state-owned enterprise reform, about 40 million laid-off workers have appeared (Solinger, 2002), which increased possibility of migrants to return. Chunyu, Liang, and Wu (2013) found that nearly 30% of the people entering Guangdong from Sichuan have returned to Sichuan Province. However, only about 10% of those who flowed into Sichuan from other provinces returned to the original provinces. They believe it marks a new phase in China's population mobility. Subsequent data showed that between 2000 and 2005, 276,000 people returned to Sichuan from Guangdong and more return population during 2005–2010, contrast to the relatively small number of return population during 1995–2005. One important contribution of our paper different from Chunyu et al. (2013) is that our data has information on *Hukou* transfer so that we could distinguish the migrants who changed *Hukou* registration location from those who return back to their original *Hukou* registration place.

Another literature document factor influencing migrant workers' return. Stark (1996) suggests that the human and financial capital accumulated elsewhere can yield higher returns at home and the third possibility is that the cost of living in their hometowns is lower. Wang and Zhao (2013) showed that age, education, marriage, and family labor endowment have significant impact on both migration and return decisions. Shi and Yang (2012) show that the richer the family human capital, the easier for the labor force to choose rural employment or return to the original countryside, but when the value of the family human capital reaches a certain level, the rural labor tends to work outside. Wang and Fan (2006) find that the family factor is the main cause of labor return, and the return population has a low level of skill in the total outflow population. Bai and Song (2002) show that reuniting with their families, especially taking care of the elderly and children, can play a role. Zhao (2002) find that education, age, married and separation with spouses would increase the possibility of labor return. Policies also influence the return of migrants. Bai and Song (2002) found that one important factor affecting the return of entrepreneurs is that the local government has explicit policies to support their return. Shi and Xue (2015) finds that the lack of social security explains the return of migrant workers.

Some research studies return migration through Todaro model (Sheng & Sun, 2009) and push-pull theory (Jin, 2009). Return migrants might have negative effects, because it exacerbates the labor surplus in rural areas. Meanwhile, there could be positive effects such as it can bring back more capital, newer technologies and more advanced corporate management methods, thus promoting local economic development. Ma (2001, 2002) show that floating population from rural areas can learn advanced skills and corporate management methods in cities and it can effectively help them get non-agricultural work after returning home.

In previous literature, the data used are mostly from small scale of survey. In order to better understand the latest trend in return migration, such as its scale, dynamic and geographical distribution, and accurately estimate the economic and social impact of return migration, we utilize a new nationally representative data. Another advantage of our research is that we can incorporate another unique data of *Hukou* registration barriers in Chinese cities to conduct a causal analysis on *Hukou* registration constraints and return migration.

2. Background

China differs from other countries in that the state has strict restrictions over migration through *Hukou* registration, which is the most relevant for our understanding of migration and urbanization in China. Since 2014, the reform of *Hukou* registration system has been accelerated. The "National New Urbanization Plan 2014-2020" issued in March 2014 states clearly that years of employment, residence, and the urban social insurance participation should be used only for large cities as the *Hukou* registration standards for local *Hukou* registration. In July, the State Council released the "Opinions on Further Promoting the Reform of *Hukou* Registration System" to further clarify the restrictions on obtaining local *Hukou* in different tiers of cities. In September 2016, the General Office of the State Council has formulated the "Proposal to Promote the *Hukou* acquisition of 100 Million non-registered Population in Cities", explicitly showing that the restrictions should be relaxed on among groups such as new generation of migrant workers and those working and living in cities for more than five years. In mega-cities, legally stable employment, residence, participation in urban social insurance and continuous residence are the main basis for local *Hukou*, meanwhile, small and medium-sized cities are not allowed to impose *Hukou* registration restrictions by channels of purchasing houses, investment and paying taxes, etc. In December 2016, the State Council issued a circular on the publication of the National Population Development Plan (2016–2030) and required implementing differential *Hukou* registration policies in cities of different size, thus to promote the *Hukou* registration for rural migrants who have the ability to work and live in cities. Our analysis will be centered around *Hukou* reform and its impact on migration for the period after 2014.

Geographically, China's population mobility has very distinct regional characteristics. In the 1990s, the inter-provincial floating population concentrated in some coastal provinces, of which Guangdong is the most important inflow province, the Yangtze River Delta is the second largest, and Sichuan Province is the main outflow province. By 2005, although the number of major outflow provinces had increased, the population were still concentrated in only one or two inflow provinces. This was mainly due to the acceleration of regional industrial adjustment since the late 1980s, and the shift of manufacturing industry from inland provinces to the coastal region. Ma and Chen (2012), using the 2010 population census, show that interprovincial migrants are mainly concentrated in Guangdong, Zhejiang, Shanghai, Jiangsu, Beijing, Fujian, Tianjin, Shandong and other places. In 2010, the inter-provincial floating population in the eight regions accounted for 77% of the national interprovincial floating population. Our paper will contribute to the understanding on the other direction of migration flow across regions, that is, return migration.

Previous studies have documented the factors that affect interprovincial migration, including per capita GDP and total population of outflow provinces, per capita GDP and migration stocks of inflow provinces, income gap and relative poverty, return on education, and dialect (Cai & Du, 2002; Xing, Jia, & Li., 2013, and Liu, Xu, & Xiao, 2015). Sun et al. (2011) explored the impact of *Hukou* reform on rural labor who moved to cities in the short-term, that is, migrant workers, the results of difference-in-difference show that the impact of *Hukou* reform on the total rural labor mobility and migration to large and medium-sized cities rather than the capital cities in provinces is not significant. In this paper, we also incorporate these factors into the understanding of migrants' returning decision.

3. Stylized facts on return migration in China

3.1. Definition of return migrants

Based on the prefectures of residence and *Hukou* registration, working and migration experience, the sample can be divided into four groups: floating population (migrants), return migrants, migrants with *Hukou* transfer and natives. Floating population is the most commonly used measure for migrants who live in a place (a place or location in this paper refers to prefecture, unless otherwise stated) for six months or more and have no local *Hukou*. The other three groups have the common feature that residential and registration prefectures are the same. Among which, return migrants are those who worked out of *Hukou* prefecture for six months or more and returned back. Migrants with *Hukou* transfer are those who moved from their original *Hukou* registration prefecture to the current prefecture. Natives have never changed *Hukou*, nor worked in another prefecture for 6 months or more. Table 1 summarizes the definitions of different groups.

Return migrants and floating population may include those with *Hukou* transfer, that is, those who have return and migration experience after transferring *Hukou*. These people account for a small proportion of the sample (0.63 and 0.57% respectively) and are

Table 1
Different groups in total population by migration status.

Classification	Definition and measure	Group
Current residence different from *Hukou* registration	Living in a prefecture without local *Hukou* registration for 6 months or more.	Floating population
Current residence same as *Hukou* registration	Had the experience of living away from the current *Hukou* registration prefecture for 6 months or more (NOT including the experience of living in previous *Hukou* registration prefecture, if any).	Return migrants
	Changed *Hukou* registration from another prefecture to the current residence prefecture.	Migrants with *Hukou* transfer
	Never changed *Hukou* registration prefecture, nor left the prefecture for 6 months or more.	Natives

Table 2
Summary of 2017 CHFS sample by migration status.

	Worked out of *Hukou* registration prefecture and returned		Working out of *Hukou* registration prefecture and NOT returning		Never worked out of the current *Hukou* registration prefecture	
	Hukou transferred	*Hukou* in original prefecture	*Hukou* transferred	*Hukou* in original prefecture	*Hukou* transferred	*Hukou* in original prefecture
Consistence of registered and actual residences	Return migration					Native
Separation of registered and actual residences			Floating population			
Observations	340	13,757	759	10,912	5309	73,416
Percentage	0.33%	13.17%	0.73%	10.44%	5.08%	70.26%

Notes: (1) All statistics in this paper adjusted by the sampling weight; (2) All migration definition refers to cross-prefecture move.

Table 3
Sample composition by *Hukou* type and gender (%).

Group	*Hukou* Type		Gender	
	Agricultural	Non-agricultural	Female	Male
Return migrants	68.29	31.71	38.27	61.73
Floating population	77.63	22.37	44.4	55.60
Migrants with *Hukou* transfer	31.33	68.67	55.91	44.09
Natives	66	34	52.17	47.83
Total	65.89	34.11	49.62	50.38

grouped as return migrants and floating population.

3.2. Data source and sample composition

The China Household Finance Survey (CHFS) is conducted by Research and Survey Center for China Household Finance at the Southwest University of Finance and Economics. The sample is representative at both national and province level and covers all provinces, municipalities, and autonomous regions, except Xinjiang and Xizang. The survey collects detailed information about household expenditure, income, employment, housing and financial asset, debt, social security, insurance, and individual demographics. More details about the survey design can be found in Gan, Yin, Jia, Xu, and Ma (2013).

Our analysis restricts the sample to those above working age at the time of the survey, by excluding observations below 16 years old. Considering that migrating and returning time may occur before the survey year, the sample includes the elderly population.

Table 2 shows the composition of different groups in the sample of 2017 CHFS. The total number of observations for aged 16 and above is 104,493. There are 14,097, return migrants, accounting for 13.50% of the sample. The number of cross-prefecture floating population is 11,671, accounting for 11.17%. The proportion of return migration in the total sample is 2.33 percentage points higher (about 20% higher) than that of the floating population.

Migrants with *Hukou* transfer account for 5.08%, a considerable proportion. If migrants with *Hukou* transfer were mistakenly treated as return migrants, the scale of return migration would be overestimated significantly (nearly 30% higher). However, the meaning of return migration and migration with *Hukou* transfer are completely different. In total, migrants including the migrants with *Hukou* transfer account for 29.74% of total population, while 70.26% of the sample are natives who have never transferred their *Hukou*, nor worked outside their *Hukou* registration prefecture.

Table 3 shows the sample distribution by *Hukou* type and gender. Individuals with rural *Hukou* are 67,815, accounting for 65.89%. Among return migrants, rural *Hukou* migrants account for 68.29%, while among the floating population, rural migrants account for 77.63%. Among migrants with *Hukou* transfer, rural *Hukou* population accounts for only 31.33%, which means *Hukou* transfer is mainly from urban to urban. There is no significant gender difference in total account, but the ratio of male is significantly higher than that of female among both return migrants and floating population, 61.73% and 55.60% respectively. The proportion of female with *Hukou* transfer is around 56%.

3.3. Spatial distribution

Table 4 presents the spatial distribution of return migrants by comparing their *Hukou* registration prefecture (origin) and working prefecture before returning (migration destination). Overall, there is the pattern that migrants are retuning back more to less "developed" regions over time. For example, the proportion of return migrants who came back to (*Hukou* registration prefecture) provincial capital prefectures before 2000 is relatively high, around 36%. After 2000, this proportion dropped to 20%. Similar patterns exist among more developed urban clusters and higher tiers (firs- and second-tier) of cities. The proportion of return migration in first-tier cities was higher at earlier period, exceeding 8%, but only 1.80% after 2000; the proportion of second-tier cities has also dropped from 34% to 22%. From the regional perspective, the proportion of return migrants in eastern regions dropped from 42% to 28%, followed by the central from 24% up to 33%, and the proportion of southwest region increased from 17% to 22% significant. The proportion of return migrants whose household registration locations in northeast decreased, while northwest increased slightly. These patterns might reflect the regional convergence or the more important role of central and southwest in providing job opportunities. Less prestigious cities are becoming more attractive for migrants to return to their home town with *Hukou* registration.

The last two columns in Table 4 show the distribution of working destination of return migrants, that is, from where they returned. The proportion of return migrants whose destination are provincial capital prefectures increased from 40% to 49%. Also, the proportions in the Yangtze River Delta and Pearl River Delta increased from 11% to 17% and from 12% to 22%, respectively while the Beijing-Tianjin-Hebei region slightly decreased from 11.16% to 9.29%. From the perspective of city tiers, the proportions of first- and second-tier cities have increased significantly, from 15% to 22% and 34% to 38% respectively. From a regional perspective, the proportion of destination in the eastern region has increased significantly from 47% to 62%, while the proportion of all other regions has shown a decrease. These patterns show that there is an increasing share of migrants returning form more developed regions. This changing landscape of living choices among migrants might be related to the high *Hukou* registration barriers in more

Table 4
Spatial distribution of return migration over Time (%).

Region	Return year	Household Registration Location (Where they returned)		Residential (or Working) Location (Where they returned from)	
		≤1999	2000–2017	≤1999	2000–2017
Provincial Capital Prefectures	Yes	36.13	19.93	40.33	49.27
	No	63.87	80.07	59.67	50.73
	Total	100	100	100	100
City Cluster	Yangtze River Delta	14.14	8.11	10.73	17.07
	Pearl River Delta	1.9	1.21	12.17	21.96
	Beijing-Tianjin-Hebei	11.45	5.05	11.16	9.29
	Others	72.51	85.63	65.75	51.47
	Total	100	100	100	100
Tier of City	First-tier City	8.2	1.8	15.2	22.43
	Second-tier City	34.48	21.63	33.69	37.59
	Third-tier City and Below	57.32	76.57	51.11	39.99
	Total	100	100	100	100
Five regions	Eastern area	41.75	27.71	47.1	61.56
	Central area	23.87	32.79	18.39	15.39
	Northeast	10.19	8.67	13.09	6.7
	Northwest	7.53	9.18	8.99	6.76
	Southwest	16.66	21.65	12.23	9.38
	Total	100	100	100	100

developed cities.

Table 5 shows whether return migrants came back to the rural or urban areas of their home prefecture, to their home county or other counties in the prefecture. The data shows that 84% of return migrants came back to their original village/town/street, among which slightly more than half are in rural area. Here, rural and urban areas are defined according to the standard of National Statistical Bureau. Around 11% of the return migrants came back to another county of their home prefecture, and 9.73% moved back to another village/town/street in their home county.

3.4. Migration temporariness

An important limitation of our study on return migration is that we only know the latest migration information of people who migrated. This section documents the temporariness of migration and return migration based on the time a migrant went to a destination prefecture and/or the time they returned to home prefecture, as shown in Figs. 1 and 2. For both migration duration (years to the current prefecture) and return duration (years since returning to home prefecture) of return migrants, we see a clear concentration in the short term (within five years), especially for migration duration. This means that majority of return migrants stayed for a short period as a migrant before returning back to their home prefecture.

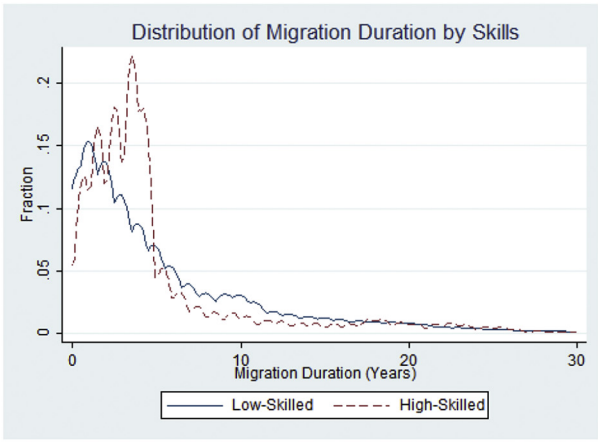
There is salient heterogeneity in migration and return duration in terms of skill, *Hukou* status, birth year (age or cohort), year of return, migration distance and health types. High-skilled migrants tend to be more concentrated in shorter span of migration, which suggests that they are more mobile. Several distinct peaks within 5-year migration stay seems to be related to schooling choices, which is also reflected in Panel e showing the stay of younger cohort is less stable. Rural, intra-provincial, healthier, and more recently returned migrants have relatively higher concentration in the 5-year or less duration, which implies that they are relatively more mobile or show more temporariness.

The return duration is more stable relative to the migration duration, suggesting that there is higher probability of settling in one's home prefecture. The high-skilled group is more concentrated in short-term period (within 5 years) after return, and it means that high-skill migrants are more likely to migrate again. Similar patterns exist for rural, intra-provincial, healthier, younger, and recently returned migrants, comparing to each of their counter group.

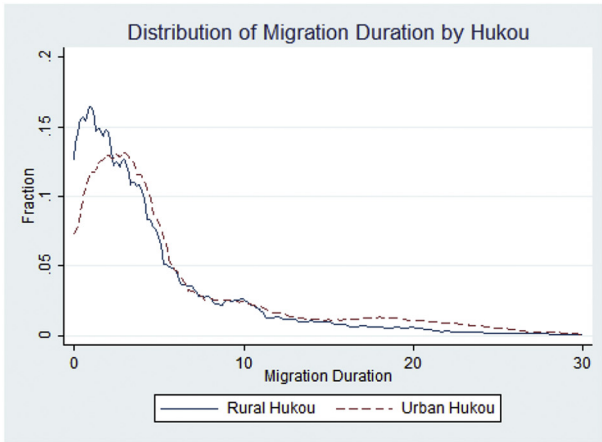
The pattern in Figs. 1 and 2 are only for the return migrants. An alternative way to show the migration temporariness is to compute the return ratio for all migrants, as shown in Fig. 3. We find that the return ratio is higher among urban-*Hukou* migrants compared with rural-*Hukou* migrants in the group of floating population with short-term migration duration (less than 10 years), but

Table 5
Destination of return migrants.

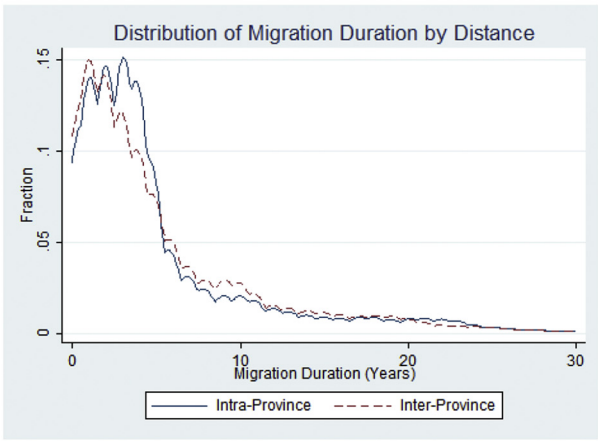
Hukou status	Within prefecture but cross county	Original town / street	Within county
Rural *hukou*	4.68%	84.08%	11.24%
Urban *hukou*	7.02%	84.41%	8.56%
Total	5.42%	84.19%	10.39%



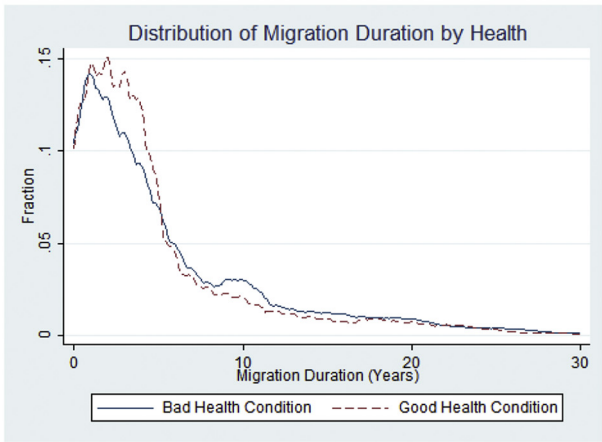
a



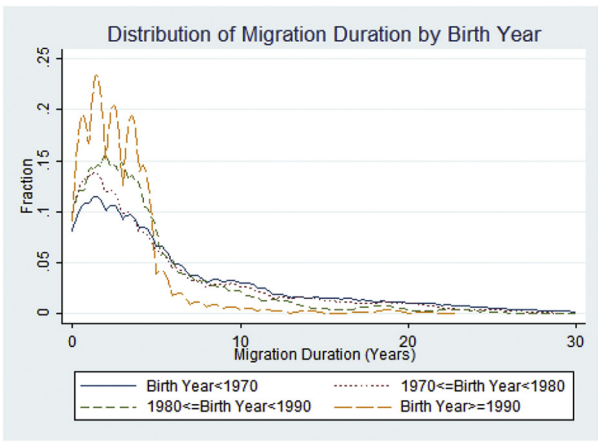
b



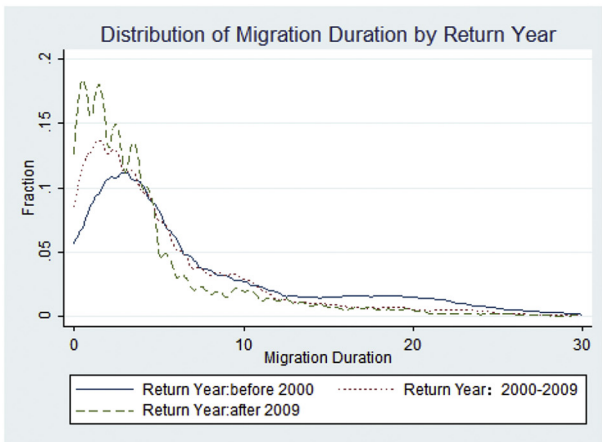
c



d



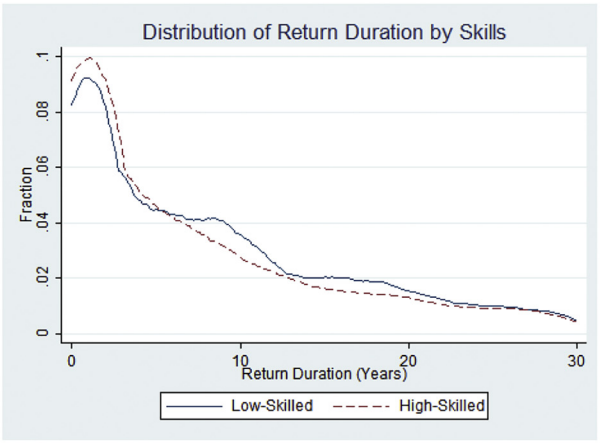
e



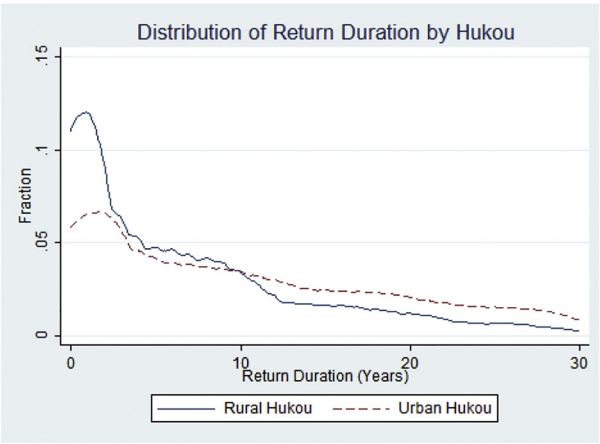
f

Fig. 1. Distribution of Migration Duration of Different Groups.

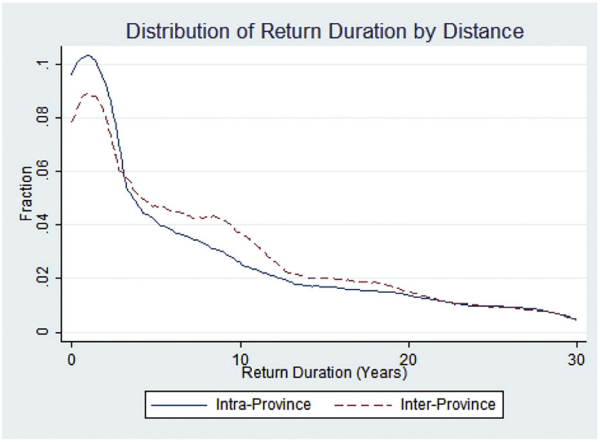
for the groups with migration duration longer than 10 years, the return ratio of rural-*Hukou* migrants becomes higher. This implies that rural migrants are less mobile in the short team and are less likely to settle in migration destination in the longer term, which could be attributed to the impact of *Hukou*.



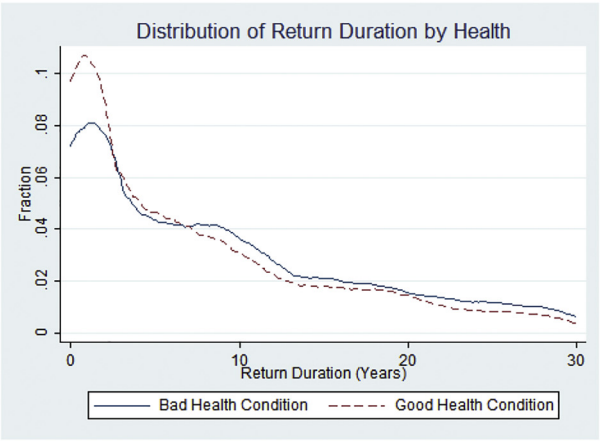
a



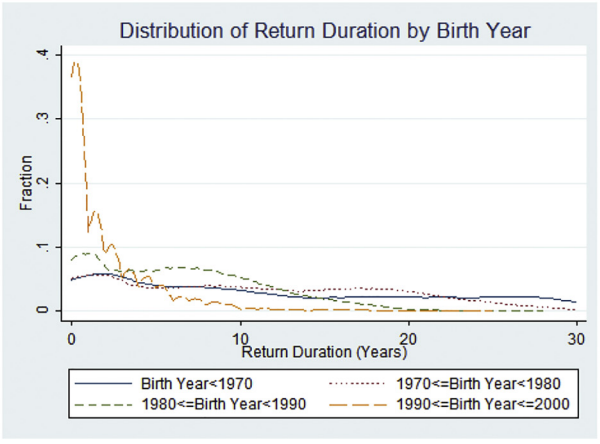
b



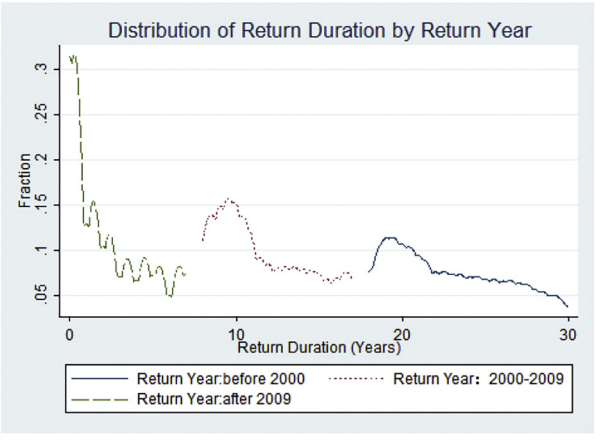
c



d



e



f

Fig. 2. Distribution of Return Duration of Different Groups.

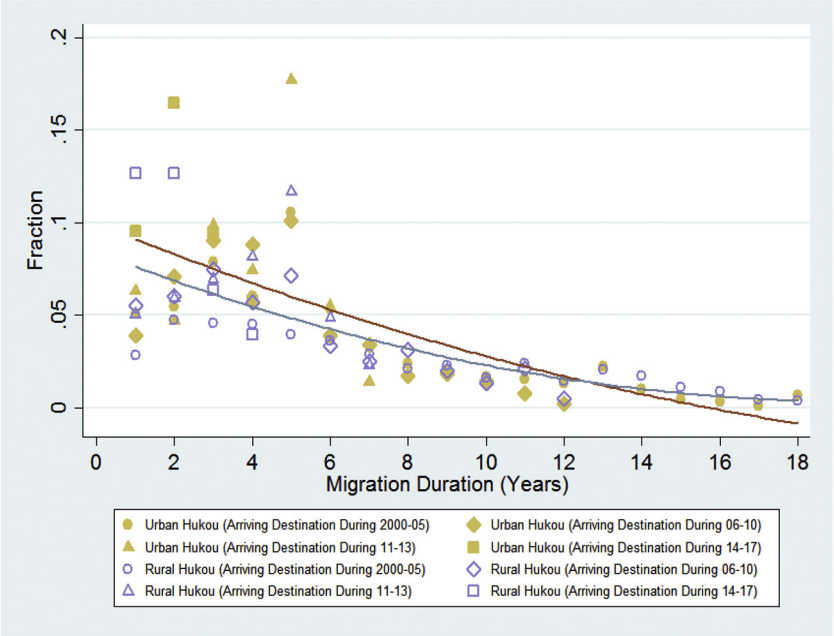


Fig. 3. Returning Fraction by Migration Duration for Rural and Urban Migrants.

4. *Hukou* registration constraints and return migration

The summary analysis indicates that migrants in more developed cities are returning back to their less developed hometown and migration has salient feature of temporariness. One possible explanation is that the high *Hukou* registration barriers in developed cities deny migrants' full access to public services in the destination they worked and hence push them back to their hometown where they have *Hukou* registration. Our empirical analysis attempts to establish the causal relationship between *Hukou* registration constraints in Chinese cities and return migration.

4.1. An illustrative model of return migration

To illustrate the decision process of return migration and the role of the *Hukou* registration barrier, we consider an individual i of age a working in a destination city d with Hukou registration in an origin place o, following Dustmann and Gorlach (2016). At time t, the migrant decides whether to go home or to work at their current destination d for another period. To simplify the notation, we omit the i and t and use only a to capture the dynamic evolution of state vector $\Omega_a = \{a, A, H\}$, where A refers to asset and H is human capital. The decision of return in each period will depend on the maximization of lifetime optimization and, in each period, it is simply the comparison of value functions in two locations conditional on the current state.

The value function is $V(\Omega_a) = max\{V^d(\Omega_a), V^o(\Omega_a)\}$, and the conditional value function under the choice of staying or return is the following

$$V^l(\Omega_a) = max_c^l u(\delta, c) + V(\Omega_{a+1}) s.t. A_{a+1} \leq (1+r)A_a + w^l(H_a) - C_a - e_0^l \quad (1)$$

where $l = \{o, d\}$, δ refers to the destination preference. The utility preference at origin will be normalized to 1, hence $\delta < 1$ means disutility at destination. Period income $w^l(H_a) = \alpha_0{}^l H_a{}^{\alpha_1{}^l}$ and $H_a = H_{a-1} + h_a{}^l$. In this environment, human capital accumulation, its productivity and rental rate all depend on location. $e_0{}^l$ is the housing or living cost that is equal to zero in origin assuming that local residents have endowed housing in their *Hukou* registration place. This cost should be understood more generally and can incorporate other costs for migrants without local *Hukou* such as denying local social benefits, childrens' entry barrier to schools, and labor market discrimination.

Now we highlight the impact of *Hukou* on returning decision. When there is only wage difference in a destination city and an origin place, an individual will move to a destination city with higher wage since most migrants in China are from less developed regions. Migrant would never return to their hometown if the wage differences are permanently higher in the destination. But higher consumption preference δ at the origin place could change that comparison. A migrant in a destination with high registration *Hukou* barrier might have stronger preference for consumption at hometown because he/she is less likely to become a destination resident. Consumption preference is also related to family and especially children and parents left behind in the origin place. Migrants in cities with higher *Hukou* registration barriers are less likely to put their kids to schools and bring their parents with them, which leads to stronger home preference and higher probability of return.

There are also other possibilities that affect returning decision of migrants through *Hukou* channel. For example, the higher living

cost in destination cities with higher *Hukou* registration requirement. Moreover, the public housing system disqualify migrants for affordable residence. Also, human capital accumulation might be faster in higher ranked cities when a person is young but return to human capital might be increasing over time in his or her origin place or simply the rental rate of human capital in one's hometown increases. The hometown preferences might increase over time because of family needs and aging and leads to higher probability of returning. All these other factors including living cost, human capital accumulation, and age and time related changes need to and will be controlled for in the regression analysis.

To summarize, the main objective of empirical analysis is to estimate the impact of *Hukou* registration barrier, but there are other relevant choice-specific factors, and individual-specific variables driving the decision of migration and return migration. Choice-specific factors include the employment opportunities in the province, overall wage level, housing price and/or living cost, regional dummies (northwest, southwest, middle, northeast, southeast, or middle-east), and other possible amenities. Individual characteristics include gender, education, marriage, occupation, previous moving experiences, family background; individual-choice specific factors can simply be the interaction items between individual and choice characteristics, or variables depending on both such as the distance of the chosen destination to ones' home province and current location, home-province dummy of destination, and other measures that capture the different effects of individual characteristics on choices, like skill-location matching efficiency. Whenever possible, we also control for these other explanatory variables if they are relatively exogenous.

4.2. The Hukou registration index data and migrants sample

Empirical analysis on the return decision of migrants utilizes a newly developed data, China *Hukou* Registration Index (CHRI) (Zhang, Wang, & Lu, 2019), that quantifies the changes of *Hukou* registration policies in 120 cities from 2000 to 2016. CHRI consists of five *Hukou* registration stringency measures: a comprehensive index and four subcategories including talent recruitment index, general employment index, investment index, and home purchase index, constructed by the following steps. 1) Classify the policy documents into different categories. 2) Read each document carefully and take out the text information according to a unified format into an Excel file by category. 3) Extract the quantitative information in the text information. 4) For the qualitative information regarding *Hukou* registration requirements, construct dummy variables and count the total number of conditions. 5) Compute a *Hukou* registration index for all policy categories of each city in both 2000–2013 and 2014–2016. Eventually, the study compiles around 1000 policy documents from prefectural, provincial and national governments. In this paper, we focus on the employment index to analyze the impact of *Hukou* registration stringency, since employment is the most relevant measure for migrants' location choice in China.

China's *Hukou* registration reform has made significant progress starting from small and medium-sized cities during the 2000s. The intensity and scope of the reform have further increased since 2013. As documented in Zhang, Wang and Lu, there are systematic differences for the registration barriers/thresholds before and after 2013. Therefore, the regression analysis focuses on the period 2014–2017. In order to match with registration index data, we utilize the time of return decision and the migration status in 2017 to exclude those return migrants who were back before 2014 so the return decisions occurred only in 2014–2017. We also exclude a small proportion of migrants who moved to a destination city before 2000 since those migrants have stayed for more than 17 years in the destination cities and their returning decisions are not likely to be affected by the recent *Hukou* reform. Migrants who were not employed before the return are excluded as well because the focus of this paper is on migrant workers. After these data cleaning procedure, there are 9207 observations in 107 cities. The descriptive statistics of the data are shown in Table 6.

4.3. Empirical analysis

We estimate the following linear probability model:

$$return_{ij} = \alpha + \beta\ threshold_j + \lambda X_{ij} + \gamma Z_j + \varphi_t + \varnothing_h + \varepsilon_{ij} \tag{2}$$

where $return_{ij}$ is binary variable and indicates whether an individual i in city j returns. $threshold_j$ is city j's *Hukou* registration stringency index.X_{ij} is a row vector of explanatory variables including characteristics of individuals and families: education level, age, gender, health, household income (in logarithm), and family size. In some specifications, we also control for the origin prefecture (*Hukou* registration place) fixed effects of individual migrants. $Z$is the characteristics of destination cities, including: wage level, GDP (in logarithm), urban population size, tertiary industry ratio, foreign direct investment, housing price; medical public service (using the number of beds in hospitals and health centers per person as proxy), educational public service (using the number of full-time teachers per student in primary school as a proxy variable), and unemployment rate in cities. φ_t is the year (moving to destination city) fixed effect. $\varnothing_h$ represents *Hukou* registration city fixed effect. $\varepsilon_{ij}$is the random error term.

The coefficient β of $threshold_j$ is our main interest. The OLS estimate could be biased because of selection bias since individuals self-select into a city for employment, schooling, and other amenities. Also, unobserved factors can be correlated with both the threshold measure and migrants' returning decision. To deal with the selection bias, we start with controlling for the most relevant factors that affect individuals' returning decision. Later, we consider to use instrumental variable to deal with the endogeneity problem caused by omitted variables.

Table 7 shows that the threshold's coefficient is negative without controlling for personal, family and cities characteristics. This implies that people are less likely to return from the cities that have higher barriers of *Hukou* registration. Such a correlation is counter intuitive. This also demonstrates clearly the endogeneity problem of *Hukou* registration barrier. However, after controlling for individual and city characteristics, time fixed effect and origin prefecture (*Hukou* registration place), the selection problem is

Table 6
Descriptive statistics and t-statistics of difference.

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
Dependent variable						
Return	If return = 1; Otherwise = 0	9207	0.28	0.45	0	1
Main explanatory variable						
Hukou registration threshold	*Hukou* registration stringency index for employment	107	0.25	0.12	0.02	0.74
Control variables for individual						
Education	Years of education	9207	10.81	3.98	0	22
Age	Age	9207	36.35	12.65	17	68
Gender	Male = 1; Female = 0	9207	0.56	0.5	0	1
Marriage	Married = 1; Otherwise = 0	9207	0.71	0.45	0	1
Urban *Hukou*	Urban *Hukou* = 1; Rural *Hukou* = 0	9207	0.23	0.42	0	1
Health	Health degree: very good = 1, very bad = 5	9207	2.16	0.89	1	5
Income	Household income(yuan)	9207	127,570	271,791.3	0.175	5,000,000
Family size	Household population	9207	4.05	1.74	1	15
Num_child	Number of children (under 16 years old)	9207	0.75	0.89	0	6
Control variables at prefecture level						
Wage	Urban wage level (yuan)	107	52,779	11,728	32,571	103,400
gdp per capia	Unit: yuan	107	74,405	34,715	16,598	196,461
Population	Resident population (10 thousand)	107	319	399	45	2350
Tertiary_ratio	Tertiary industry ratio (%)	107	46.02	11.12	23.34	77.82
FDI per capia	Unit:10 thousand	107	3985	3799	46	17,744
Housing price	Unit:yuan/m^2	107	5983	3684	2326	24,040
Education_city	Number of full-time teachers per student in primary school	107	0.06	0.01	0.03	0.11
Medical service	Number of beds of hospitals per 10 thousand people	107	53.61	20.96	17.1	135.81
Unemployment rate	Unemployment rate (%)	107	4.42	2.13	0.38	11.26

Variables	Variable description	Floating population		Return migrants		Mean difference
		Sample size	Mean	Sample size	Mean	
Education	Years of education	6660	10.72	2547	11.05	−0.33***
Age	Age	6660	37.30	2547	33.85	3.45***
Gender	Male = 1;Female = 0	6660	0.53	2547	0.61	−0.08***
Marriage	Married = 1 ; Otherwise = 0	6660	0.75	2547	0.59	0.16***
Health	Health degree: very good = 1, very bad = 5	6660	2.13	2547	2.22	−0.09***
Income	Household income(yuan)	6660	137,565.4	2547	101,433.5	36,131.9***
Family size	Household population	6660	3.94	2547	4.32	−0.38***
Num_child	Number of children (under 16 years old)	6660	0.77	2547	0.71	0.06***

alleviated and the coefficient becomes positive and significant. Overall, higher household registration barrier in migrants' destination cities makes migrants more likely return to their *Hukou* registration place.

Apart from the *Hukou* registration constraints, other individual characteristics and city level variables also affect the returning decision. Specifically, education level has significantly negative effect on returning decision, which is reasonable because people with higher education are more capable to get a steady job in a city, and therefore more likely to stay in a city. The gender difference is also significant and man is more likely return to original *Hukou* city. The reason might be that in Chinese culture, the economic and family burden for men to settle in a city is higher than women and the high living cost force men to return instead of settling down. People who get married are more likely to stay in the working cities, which might be caused by the higher economic incentive to support family. People with urban *Hukou* is less likely to return back to household registration locations than these with a rural *Hukou*, which highlights the higher cost of staying in cities for rural migrants. Household income and household size are both negatively correlated with the probability of returning, both of which can reflect the economic impact on returning decisions. Having a child in a household will significantly increase the probability of returning, probably because of limited access to public education in cities for migrants' kids, which force adults to return to take care of the children for education purpose.

After controlling for various individual and city characteristics, origin prefecture and migration year fixed effects, the *Hukou* stringency coefficient becomes reasonable, but the endogenous problem of *Hukou* barrier can still arise from omitted variables. In order to make the results more convincing, we use the city's per capita grain output in 1990 as an instrument variable for the household registration barrier in Chinese cities. The rationale behind our choice of instrumental variable has empirical support. Cai, Du, and Wang (2001) found that there is a significant positive correlation between the planned migration population in each city in 1952–1998 and the per capita annual food production in previous year. This is because in the era of high commodity circulation costs, the grain output of a city determines the population capacity of the city and also *Hukou* registration stringency. Since the production of grain in 1990 contained exogenous determinants such as geological conditions, the urban household registration policy determined in the 1990s will be partially extended to the period after 2000 because of the persistency of *Hukou* institution and the difficulty of *Hukou* reform. Moreover, the grain out in the 1990s is not likely affecting the returning decisions of migrants in a city in

Table 7
Explanatory factors of return migration.

Dept. Var.	(1)	(2)	(3)	(4)	(5)
	If Return (0–1)				
	OLS				
Threshold	−0.207***	−0.185***	−0.145***	0.176**	0.164**
	(0.0383)	(0.0346)	(0.0334)	(0.0744)	(0.0761)
Education			−0.00625***	−0.00655***	−0.00649***
			(0.00165)	(0.00161)	(0.00162)
Age			−0.00292	−0.00256	−0.00274
			(0.00292)	(0.00289)	(0.00290)
Age^2			0.00196	0.00145	0.00172
			(0.00336)	(0.00333)	(0.00334)
Gender			0.0553***	0.0545***	0.0539***
			(0.00756)	(0.00750)	(0.00753)
Marriage			−0.0442***	−0.0442***	−0.0442***
			(0.0161)	(0.0161)	(0.0160)
Urban Hukou			0.0410***	0.0418***	0.0422***
			(0.0143)	(0.0141)	(0.0140)
Health			0.0271***	0.0275***	0.0274***
			(0.00578)	(0.00571)	(0.00572)
Family income(ln)			−0.0300***	−0.0294***	−0.0294***
			(0.00392)	(0.00391)	(0.00391)
Family size			0.0449***	0.0443***	0.0440***
			(0.00531)	(0.00524)	(0.00524)
Num. of child			−0.0482***	−0.0471***	−0.0467***
			(0.00916)	(0.00911)	(0.00901)
Wage(ln)				−0.0311	0.0179
				(0.0657)	(0.0700)
gdppc(ln)				0.00231	−0.00419
				(0.0259)	(0.0313)
Population				−0.0626***	−0.0645***
				(0.0200)	(0.0208)
Tertiary ratio				0.000850	0.000650
				(0.00104)	(0.00110)
Housing price(ln)				**−0.0547****	**−0.0827*****
				(0.0219)	**(0.0245)**
Education_city					−0.550
					(0.944)
Medical service					0.000628
					(0.000439)
Unemployment rate					−0.00556*
					(0.00312)
Constant	0.359***	0.315	0.541***	1.233*	1.055*
	(0.0209)	(0.199)	(0.207)	(0.632)	(0.614)
Origin prefecture FE		Y	Y	Y	Y
Year of arrival FE		Y	Y	Y	Y
N	9207	9207	9207	9207	9207
R-squared	0.006	0.126	0.159	0.163	0.164

Note: robust standard errors are in parentheses and clustered at city level.
*** $p < .01$.
** $p < .05$.
* $p < .1$.

recent years.

The exclusion restriction assumes that per capita grain production should not be correlated with unobserved factors that affect migration decision. The per capita grain output in 1990 could affect cities' *Hukou* stringency via other channels that might affect migrants' returning decision today, which would violate the exclusion restriction. Per capital grain output could be related to the productivity level of a city that should be a factor in affecting both *Hukou* migration decisions. We add a set of proxies of destination productivity to alleviate this channel of influence. The grain productivity could also affect the cultural and political environment that could affect both people's and local governments' openness toward migrants and contributing to the returning probability of migrants. This impact we believe would be more directly reflected in the *Hukou* registration stringency and already have been controlled for. In summary, the grain production in 1990s seems to be a reasonably good instrumental variable for *Hukou* registration barriers in Chinese cities today when investigating migrants' location choice.

Table 8 reports the regression results of two stage least squares. In the first stage regression, the coefficient of the per capita grain output is significantly negative, meaning that the cities with higher grain production in 1990 have lower *Hukou* registration barrier,

Table 8
Estimation using grain output per capita in 1990 as IV.

	(1)	(2)	(3)	(4)
	Threshold	Return	Threshold	Return
	2SLS			
	First stage	Second stage	First stage	Second stage
Instrument variable				
IV (grain output)	−0.120***		−0.122***	
	(0.011)		(0.011)	
First-stage F	69.62		72.83	
Hukou registration constraints effects				
Threshold		1.017*		1.029*
		(0.585)		(0.582)
Education	0.0000723	−0.00686***	−0.00016	−0.0066***
	(0.00026)	(0.0016)	(0.00024)	(0.0016)
Age	0.0000280	−0.000746	0.0000194	−0.00074
	(0.00009)	(0.0005)	(0.00008)	(0.0005)
Gender	0.000053	0.0350***	0.00016	0.0348***
	(0.0016)	(0.0097)	(0.0015)	(0.0097)
Marriage	−0.00167	−0.0389**	−0.00138	−0.039**
	(0.0025)	(0.0154)	(0.0023)	(0.015)
Rural *Hukou*	−0.00121	0.0391***	−0.00283	0.041**
	(0.002)	(0.0135)	(0.002)	(0.0135)
Health	0.000941	0.0193***	0.000760	0.0196***
	(0.000949)	(0.00585)	(0.00088)	(0.0058)
Income	0.000905	−0.0190***	0.00115**	−0.0195***
	(0.0006)	(0.0038)	(0.00057)	(0.0038)
Family size	−0.000307	−0.0137***	−0.000587	−0.0134***
	(0.000475)	(0.0029)	(0.00044)	(0.0029)
Num. of child	0.00138	0.0133**	0.00179*	0.0127**
	(0.00100)	(0.00620)	(0.0009)	(0.0062)
Wage	0.0146	−0.113*	−0.0845***	0.0286
	(0.0108)	(0.066)	(0.011)	(0.0903)
gdp	0.0208***	−0.00868	0.0532***	−0.0568
	(0.0028)	(0.0232)	(0.0030)	(0.0397)
Population	0.00019***	−0.000311**	0.00006***	−0.00014**
	(0.0000078)	(0.00013)	(0.0000086)	(0.00007)
Tertiary_ratio	−0.0019***	0.00148	−0.0005***	−0.00053
	(0.00014)	(0.0011)	(0.00014)	(0.0009)
FDI	−0.00114	0.0177*	0.00580***	0.0099
	(0.0014)	(0.00907)	(0.00134)	(0.0078)
Housing price	0.0063***	−0.012***	0.0065***	−0.012***
	(0.0002)	(0.0044)	(0.0002)	(0.0046)
Education service			−0.947***	0.43
			(0.0937)	(0.763)
Medical service			−0.00097**	0.00149**
			(0.00006)	(0.00067)
Unemployment rate			1.098***	−1.213*
			(0.0408)	(0.716)
Constant	−0.248**	2.288***	0.285**	1.582**
	(0.121)	(0.762)	(0.115)	(0.770)
Huji registration FE	Y	Y	Y	Y
Arrival year FE	Y	Y	Y	Y
N	7241	7241	7241	7241

Note: Standard errors in parentheses; t-statistics are in parentheses and standard errors clustered at city level.
*** p < .01.
** p < .05.
* p < .1

consistent with the theoretical prediction. Statistically, the instrumental variable is a strong predictor of *Hukou* registration barrier today, with an F value of 72.83.

The second-stage regression shows that *Hukou* registration barrier leads to more migrants return to their *Hukou* registration place. One standard deviation increase of *Hukou* registration stringency leads to 10% more likely of returning. The estimated effect is much larger than that from an OLS regression with city, individual, and fixed effects controls. Put it differently, if first-tier cities remove *Hukou* registration barrier, migrants in those places are 50% less likely to return, given everything else equal.

Table 9
Returning decision of migrants by skills.

	(1)	(2)	(3)	(4)
	2SLS			
	First stage	Second stage	First stage	Second stage
	High		Low	
IV (grain output)	−0.150***		−0.125***	
	(−6.97)		(−9.75)	
Threshold		0.184		1.500**
		(0.2)		(2.21)
Personal characteristics	Y	Y	Y	Y
Household characteristics	Y	Y	Y	Y
Urban characteristics	Y	Y	Y	Y
Time of arrival -FE	Y	Y	Y	Y
Origin prefecture -FE	Y	Y	Y	Y
N	2018	2018	5223	5223
F value	25.40		57.72	

Note: t-statistics are in parentheses.
*** p < .01.
** p < .05.
* p < .1.

4.4. Heterogeneity analysis and alternative instrumental variable

Due to the skill bias of urban development policy, the impact of household registration stringency index on migrants with different skill levels may be different. Following the existing literature, workers with college degree or above are defined as high-skilled group.

In the specifications after controlling for individual and city characteristics and the IV estimation, the estimated effects of *Hukou* stringency measure on the returning decision of high-skilled migrants are not significant, as seen in the column 2 of Table 9. For the low-skilled group, the impact is positive and significant, and the estimated effects are 50% larger than the results in Table 8 where the high and low skilled workers are grouped together. These findings can be explained by the fact that in recent years, major cities have relaxed the *Hukou* registration system to attract high-skilled workers, but not for the low-skilled workers.

Migration distance and provincial boundary can also be important factors affecting migrants' returning decision, because cross-province migration involves long distance move that comes with higher economic and psychological cost. Also, migrants from a different prefecture but in the same province are better treated than migrants from another province. Hence, we compare the

Table 10
Returning decision of migrants by distance.

	(1)	(2)	(3)	(4)
	2SLS			
	First stage	Second stage	First stage	Second stage
	Inter-Province		Intra-Province	
IV (grain output)	−0.160***		−0.0652***	
	(−10.79)		(−4.59)	
Threshold		1.478**		−0.814
		(2.39)		(−0.37)
Personal characteristics	Y	Y	Y	Y
Household characteristics	Y	Y	Y	Y
Urban characteristics	Y	Y	Y	Y
Time of arrival -FE	Y	Y	Y	Y
Original prefecture-FE	Y	Y	Y	Y
N	3965	3965	3276	3276
F value	42.26		133.09	

Note: t-statistics are in parentheses.
*** p < .01.
** p < .05.
* p < .1.

Table 11
Robustness checks using alternative IV and different age restrictions.

	(1)	(2)	(3)	(4)	(5)
	Keep samples with 23 years old and above	Keep comparable samples between 24 and 60 years old	Keep comparable samples between 24 and 50 years old	Delete those moved before 2000 and age 23 and below	Delete those moved before 2000, keep age between 24 and 65
Basic regression	0.199***	0.191***	0.164**	0.206***	0.210***
	(0.071)	(0.071)	(0.070)	(0.078)	(0.078)
IV estimation (Bartik Shock_lagged)	1.115**	1.247***	1.198***	1.136**	1.124**
	(0.449)	(0.456)	(0.448)	(0.483)	(0.469)
Personal characteristics	Y	Y	Y	Y	Y
Household characteristics	Y	Y	Y	Y	Y
Urban characteristics	Y	Y	Y	Y	Y
Year of arrival-FE	Y	Y	Y	Y	Y
Original registration prefecture-FE	Y	Y	Y	Y	Y
N	7952	7611	6629	6912	6814

*** p < .01.** p < .05.* p < .1.

difference between cross-province migration and within-province migration. As shown in Table 10 columns 4, the estimated effects of *Hukou* registration barrier on the returning decision of migrants within province are not significant, while the coefficient of *Hukou* stringency for cross province migrants is positive and significant in the IV estimation of column 2. This means that the *Hukou* registration stringency only affects the return decision among migrants moving out-of-province. This is reasonable because in a province, the cost of migration is much smaller than cross-province migration, especially in terms of the impact of *Hukou*, since most provinces provide better public services for within-province migrants than for migrants from another province.

We conducted additional robustness check on the instrumental variable, age restriction, *Hukou* status and other categories of registration indices. The results shown in Table 11, using alternative instrumental variable–Bartik shock, have similar findings as those using historical grain production. The results are also robust to the change of age restrictions on the sample population. Bartik shock is a commonly used instrumental variable in labor and urban economics (Diamond, 2016). It uses the employment structure (proportion) of local industries in 2005 as the weights to compute the average employment growth of all industries (excluding local growth) nationally between 2005 and 2010. The rationale behind this instrumental variable is the following. One-period-lagged Bartik shock can proxy the unobservable productivity shock in local economy and has a direct impact on the employment demand and other unobservable urban amenities, hence it correlates to high threshold for urban settlement. Meanwhile it satisfies the exclusion restriction in the sense that lagged Bartick shock has no correlation with local productivity and amenity changes that might affect migrants' returning decision.

Housing price has important impact on the returning decision of migrants, but it is endogenous because of correlation with unobserved productivity shock that might cause downward biases on returning probability. We tried to use grain output in 1990 and lagged Huji registration constraint as instrumental variables by treating both housing price and Huji registration index as endogenous. We do find that higher housing price leads to higher probability of returning, and such an effect is decreasing with housing price. This implies that the impact of housing price on retuning decision is smaller in more expensive cities. One interpretation for this finding is that migrants in higher tier cities where housing price is more expensive are not likely to bus a housing unit because of affordability, instead they might value the working opportunity and consumption externality more in those locations.

The endogeneity problem of hosing price does not affect our main estimates on Hukou registration index, same for heterogeneity analysis. For *Hukou* status in Table 12, we find that *Hukou* registration constraints have a positive impact on the retuning probability of rural migrants, but there is no significant impact on urban *Hukou* migrants. This suggests that urban migrants who already entitled to social benefits in urban community are less responsive to *Hukou* registration barriers.

In addition, we also run the regressions using other categories of registration indices and the results (available from the authors) are consistent with the main findings of the paper. One exception is that home purchase index has no significant impact on returning probability of migrants. This might be explained by two facts. Migrants in the empirical sample have no local Hukou, but people who bought a home might have changed Hukou to local status and thus not in the sample. Another possibility is that we control a set of city-characteristics that could capture the impact of home purchase index.

5. Conclusion

Using a nationally representative micro-data (CHFS 2017), this paper studies the latest trend of labor mobility and return migration in China. We document that the size of return migrants is much larger than that of migrants without local *Hukou* registration. Return migrants are those returning from a residential prefecture to their Hukou registration prefecture, not including the group who changed *Hukou* registration after migration. According to CHFS 2017, the ratio of cross-prefecture migration with *Hukou* transfer is 5.93%, moreover population with *Hukou* transfer are expected to have a faster increase as Chinese government is pushing *Hukou* registration reform.

Table 12
Robustness check on housing price endogeneity and heterogeneity analysis.

	(1)	(2)
	Dependent Variable: Huji Index	Dependent Variable: Return
	2SLS	
	First stage	Second stage
Instrument variable		
IV (grain output in 1990)	−0.025	
	(0.020)	
IV (threshold lagged)	0.340***	
	(0.028)	
First-stage F	828.1	
Hukou registration constraint effects		
Threshold		0.563**
		(0.277)
ln (housing price)	−0.994***	2.433***
	(0.139)	(0.568)
ln(housing price) square	0.056***	−0.137***
	(0.007)	(0.031)
N	9207	9207
R-squared	0.87	0.165

Heterogeneity Analysis on Return Probability (Second Stage)

	(3)	(4)	(5)	(6)	(7)	(8)
	2SLS					
	Skill	Distance			*Hukou*	
	Low	High	Inter-Province	Intra-Province	Urban *Hukou*	Rural *Hukou*
Hukou registration constraint effects						
Threshold	0.414	0.636	0.601**	−0.871	−0.358	0.691**
	(0.327)	(0.388)	(0.256)	(0.834)	(0.556)	(0.280)
ln(housing price)	1.824**	2.984***	3.129***	0.280	3.237***	2.097***
	(0.711)	(0.734)	(0.607)	(1.266)	(1.092)	(0.609)
ln(housing price) square	−0.103***	−0.167***	−0.175***	−0.022	−0.179***	−0.119***
	(0.039)	(0.040)	(0.033)	(0.067)	(0.060)	(0.033)
Personal characteristics	Y	Y	Y	Y	Y	Y
Household characteristics	Y	Y	Y	Y	Y	Y
Urban characteristics	Y	Y	Y	Y	Y	Y
Origin prefecture FE	Y	Y	Y	Y	Y	Y
Arrival year FE	Y	Y	Y	Y	Y	Y
N	4863	4344	5565	3642	2141	7066
R-squared	0.156	0.224	0.151	0.233	0.268	0.154

Note: all control variables and fix effects are the same with Table 7.
*** p < .01.** p < .05.* p < .1.

The residence of cross-prefecture migrants is mainly in provincial capital cities, first-tier and second-tier cities, as well as three major economic zones (the Yangtze River Delta, the Pearl River Delta, and the Beijing-Tianjin-Hebei region) while the destination of the return migration is mainly in less developed areas including the third-tier cities and below. Although majority of migrants have agricultural *Hukou*, among return migration, the proportion of non-agricultural household registration population is about 31%. Young people account for a relatively high proportion, but the return population has an aging trend.

Empirically, we show that urban *Hukou* registration stringency affects the return decision of migrants. Higher household registration barrier leads to higher probability of return among the low-skilled migrant workers, but the impact on the high-skilled group is not significant. Other personal and family factors also affect return migration, such as gender, age, education, health status, family size, and number of children. Urban characteristics such as industrial structure, living cost, wage, and population size are also important factors affecting return migration. Heterogeneity analysis shows that the impact of *Hukou* registration barrier on returning decision only exist among migrants who moved away from home province and migrants with rural *Hukou*.

Our analysis only establishes the evidence on one possible explanation of return migration. We cannot rule out other facts that might also contribute to the return of migrants, such as relative improvement of job opportunities in less developed regions and the life-cycle choice of returning home for “retirement” from city jobs as migrants. But, the policy implications of our findings clearly

point to the *Hukou* reform that should focus more on providing better public services for low-skilled workers, those who move across provinces and rural migrants.

Last but not least, we want to point out several important caveats of the paper. First, return migration in this paper is "temporary" since in the data we only know the recent move of migrants not the complete migration history. Second, there is a potential reverse causality problem. That is, city leaders may adjust the stringency of their local *Hukou* systems based on the observed number of return migrants leaving their cities. To completely solve this problem, one needs estimate a structural model by endogenizing the determination of *Hukou* stringency. We leave these for future work with better data, better modeling and estimation of return migration.

Acknowledgements

This research is supported by the Fundamental Research Funds for the Central Universities (JBK1902038 and JBK1805007).

References

Bai, N., & Song, H. (2002). *Return home of go to the City?A study of return migration in rural China.* Beijing: China Finance Economics Press.
Cai, F., & Du, Y. (2002). The double cause of migration and its policy implications- test the relative poverty hypothesis. *Chinese Journal of Population Science, 4*, 1–7.
Cai, F., Du, Y., & Wang (2001). Household registration system and labor market protection. *Economic Research, 12*, 41–49.
Chunyu, M., Liang, Z., & Wu, Y. (2013). Interprovincial return migration in China: Individual and contextual determinants in Sichuan province in the 1990s. *Environment and Planning A, 45*, 2939–2958.
Diamond, R. (2016). The determinants and welfare implications of US Workers' diverging location choices by skill: 1980-2000. *American Economic Review, 106*(3), 479–524.
Gan, L., Yin, Z., Jia, N., Xu, S., & Ma, S. (2013). *Data you need to know about China.* Berlin Heidelberg: Springer.
Jin, S. (2009). A push-pull model analysis of migration decision making for rural migrants. *Statistics & Decisions, 9*, 64–66.
Liu, Y. Y., Xu, X. X., & Xiao (2015). The pattern of labor cross-dialects migration. *Economic Research Journal, 10*, 134–146.
Ma, Z. (2001). Urban labor-force experience as a determinant of rural occupation change: Evidence from recent urban-rural return migration in China. *Environment and Planning A, 33*(2), 237–256.
Ma, Z. (2002). Social-capital mobilization and income returns to entrepreneurship: The case of return migration in rural China. *Environment and Planning A, 34*(10), 1763–1784.
Ma, H. Q., & Chen, Z. C. (2012). Patterns of Inter-Provincial Migration in China: Evidence from the Sixth Population Census. *Population Research, 36*(6), 87–99.
Sheng, Y. N., & Sun, M. (2009). An economic analysis of return migration of peasant workers. *Population Research, 33*(6), 102–109.
Shi, Z. L., & Xue, W. L. (2015). Urban and rural long-term security and remigration decision of migrant workers. *China Population, Resources and Environment, 25*(3), 143–152.
Shi, Z. L., & Yang, Y. Y. (2012). Household endowment, household decision and return of rural migrant labor. *Sociological Studies, 2012*(3), 157–245.
Solinger, D. J. (2002). Labour market reform and the plight of the laid-off proletariat[J]. *China Quarterly, 170*(170), 304–326.
Stark, O. (1996). *On the microeconomics of return migration[M]//trade and development.* London: Palgrave Macmillan32–41.
Sun, W. K., Bai, C. E., & Xie, P. C. (2011). The Impact of Household Registration Reform on Rural Labor Mobility in China. *Economic Research, 1*, 28–41.
Wang, W., & Fan, C. (2006). Success or failure: Selectivity and reasons of return migration in Sichuan and Anhui, China. *Environment and Planning A, 38*, 939–958.
Wang, Z. C., & Zhao, Z. (2013). Dynamic choice of migration patterns for migrant workers: Going out, return or remigration. *Management World, 2013*(1), 78–88.
Xing, C. B., Jia, S. Y., & Li. (2013). Regional distribution of the return to education for rural-to-urban migrants and its impact on migration. *Economic Research Journal, 11*, 114–126.
Zhang, J. P., Wang, R., & Lu, C. (2019). A quantitative analysis on *Hukou* reform in Chinese cities: 2000-2016. *Growth and Change, 50*(1), 201–221.
Zhao, Y. (2002). Causes and consequences of return migration: Recent evidence from China. *Journal of Comparative Economics, 30*(2), 376–394.