

Siddique, Abu Bakkar

Article

Identity-based earning discrimination among Chinese people

IZA Journal of Development and Migration

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Siddique, Abu Bakkar (2020) : Identity-based earning discrimination among Chinese people, IZA Journal of Development and Migration, ISSN 2520-1786, Sciendo, Warsaw, Vol. 11, Iss. 1, pp. 1-42,
<https://doi.org/10.2478/izajodm-2020-0005>

This Version is available at:

<https://hdl.handle.net/10419/298677>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Abu Bakkar Siddique*

Identity-based Earning Discrimination among Chinese People

Abstract

Hukou registration is an instrument to control nonplanned population and capital movements, which the Chinese Communist Party has been exploiting extensively since the 1950s. It requires that each Chinese citizen be classified as either an agricultural or nonagricultural *hukou* inheritor and be distinguished by their location with respect to an administrative unit. *Hukou* distribution used to be entirely determined by birth, but nowadays, Chinese citizens can self-select their *hukou* status based on their ability that causes selection bias in conventional wage decomposition by *hukou* types. To avoid this bias, I estimated *hukou*-based earning discrimination by matching Chinese individuals based on a rich set of individual-, family-, and society-level characteristics. By deploying a recent nationally representative dataset, this paper finds that significant earning discriminations exist against agricultural *hukou* people. I further investigated the impact of *hukou* adoption within work ownership, work and employer types, and labor contract conditions. I argue that earning difference by *hukou* is not due to rural-urban segregations; rather, it is systematic and institutionally enforced. This is because, contrary to self-employment and no labor contract conditions, discrimination exists only when others employ them and where a labor contract condition is enforced. Moreover, they face discrimination only when they work for the Chinese government, not when they work for private firms, and they face higher discrimination in nonagriculture-related professions compared to agriculture-related professions.

Current version: February 07, 2020

Keywords: *Hukou*, labor market, earning discrimination, development, migration, matching estimator, China

JEL codes: C21, C31, I38, J7, J71, J15, J31, N35

Corresponding author: Abu Bakkar Siddique
anirbanju36@gmail.com

Schar School of Policy and Government, George Mason University, 3351 Fairfax Dr, Arlington, VA 22201, USA.
E-mail: anirbanju36@gmail.com. Cell: +15712745328

© The Author(s). 2020. Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

 Cite as: Siddique. *IZA Journal of Development and Migration* (2020) 11:5.

<https://doi.org/10.2478/izajodm-2020-0005>

1 Introduction

China's transition from the planned economic regime to a liberal economic regime has marked a remarkable welfare improvement for Chinese people. One of the key features of this transition process is the gradual dismantling of traditional institutions. The transition is not complete yet because the traditional institutions that historically regulated the flow of capital and labor are still playing a significant role in the Chinese society (Lardy, 1998; London, 2014). One such institution is the *hukou* or household registration system.¹ The system was designed to strictly limit people's mobility within the Chinese mainland (Song, 2014). It was considered a necessary component of the centrally planned economy because a strict central plan requires the ability to allocate resources not only at the enterprise and sectoral levels but also across geographic locations (Liu, 2005). However, now it has been identified as a major obstacle to China's quest to become a modern developed economy (Chan, 2009). It is the most pressing area for policy reform. Many have also identified the *hukou* system as the direct cause of extreme economic and social inequality in China (Afridi et al., 2015; Liu, 2005; Wu and Treiman, 2007). Other significant negative consequences include differential accessibility to public services and welfare programs, different costs of living in cities, and discriminatory labor market treatment (Song, 2014). It is also responsible for large productivity loss (Au and Henderson, 2006), middle-income trap (Zhang et al., 2013), and nonoptimal allocation of resources (Whalley and Zhang, 2011).

The majority of the existing research has attempted to partition the wage differential between the two *hukou* groups into components caused by two factors: a difference in productivity and an unexplained component that is often referred to as discriminations (Blinder, 1973; Oaxaca, 1973).² Most papers mainly compared wages between urban residents and rural–urban migrants and estimated the different magnitudes of the unexplained wage gap. For example, Song (2016) observed discrimination against rural–urban migrants in state-owned enterprises (SOEs) and private firms within urban China using the 2008 wave of the Rural-Urban Migration in China survey. By applying regression and decomposition technique, he found that urban *hukou* holders earn about 50% more than rural *hukou* holders do in the SOEs, but only 5% more in the private sector.³ While this literature is full of merits and a valuable resource to understand the *hukou*-based earning discrimination, their regression-based

1 *Hukou* has two main classifications: one by *hukou* type and other by *hukou* location. While the *hukou* type is separated into "agricultural" and "non-agricultural" *hukou*, the *hukou* by location is that each person is also categorized according to his/her place of *hukou* registration. The *hukou* types and locations define citizen's eligibility for public services in a specific locality. The place of registration is determined mainly by birth, and this is the individual's official and only "permanent" residence (Chan, 2009). Until 2003, a newborn baby's *hukou* type followed the mother's rather than the father's (Wang, 2004). Later, if parents differ in their *hukou* types, their children can freely choose *hukou* types either from father's or mother's *hukou* (Huang, 2012). To attract outstanding migrants, different cities/provinces have introduced different programs that allow people to convert their *hukou* types from agriculture to non-agriculture and transfer from one place to another place. Cities usually set a limit on how many *hukou* they want to offer each year, and criteria of alteration are entirely ability based. Thus, ordinary people are not usually able to alter their *hukou*. There are also cases where people do not want to convert their *hukou* because they are not willing to give up their contract land that is tied to their agricultural *hukou* types (Chen and Fan, 2016).

2 See Song (2014) for a review of existing literature.

3 Others such as Lee (2012) found 28% (using China's Urban Labor Survey in 2005), Gravemeyer et al. (2011) 52.9% (using Shenzhen Survey 2005 data from Shenzhen city), Gagnon et al. (2012) 40% (using Census 2005), and both Deng (2007) and Liu (2005) 60% (using China Household Income Project (CHIP) 2002 and Beijing sampled from the CHIP in 1995, respectively) of the average wage difference cannot be explained by observable characteristics. They called the unexplained portion of the wage gap as discrimination. Gagnon et al (2012) confirmed that *hukou*-based wage discrimination is only attributed to *hukou* type, rather than to *Hukou* location. Note that there is one research conducted by Dong and Bowles (2002) who claim that there is no discrimination against rural migrants.

empirical analyses, including Song (2016), suffer from several major limitations. First, no existing studies considered the fact that the *hukou* status may not be exogenous because outstanding rural–urban migrants can be granted urban *hukou*. So, the *hukou* status distributions may not be random because the status can be self-selected based on ability.⁴ Second, datasets that have been used in the existing studies do not allow controlling many important factors in their parametric estimation including ability, social and family background, health, attitude, ambition, work effort, networking, and so on that may influence labor market outcomes (Bowles et al., 2000, 2001a, 2001b; Taubman, 1976; Filippin and Paccagnella, 2012; Heineck and Anger, 2010; Heineck, 2011). Again, ability measures have multiple dimensions such as mathematical, educational, language, communication, and cognitive ability.⁵ Standard earning equation, which includes a person's age, schooling, and experience and parents' schooling, occupation, and income, cannot explain two-thirds to four-fifths of the variance of earnings (Bowles et al., 2001b). So, the existing literature is likely to suffer from omitted variable biases. Third, past empirical results were not also well representative of both types of *hukou* people and of China as a whole. Earlier surveys were residence based, and others were conducted in few cities in urban China where it was difficult to obtain a representative sample of agricultural *hukou* people except rural–urban migrants who lived in the urban areas. Even majorities of the rural–urban migrants were not well covered because they mostly lived in construction sites or workplace dormitories and hence may not be included in the residence-based samples.⁶ Thus, the residence-based surveys may be biased due to the omission of migrant people, and urban-focused surveys may be biased due to the exclusion of agricultural *hukou* people living in rural areas. Moreover, their datasets were relatively old, which may not reflect the recent state of the Chinese labor market.

This paper addresses all these gross limitations listed above and provides improved evidence of *hukou*-based earning discrimination using comprehensive nationally representative survey data. While existing literature examined only earning differences between urban *hukou* people and rural–urban migrants, this paper first estimates overall earning differences between urban *hukou* and rural *hukou* people and then verifies whether earning differences change if rural *hukou* people are migrants. We observe both wage and total income where total income includes both wage and nonwage earnings.⁷ I also estimated earning differences by *hukou* types within work ownership types, work types, employer types, and labor contract conditions. To my knowledge, I am not aware of other studies in the literature that studied *hukou*-based discrimination by applying similar approaches. I applied

4 Estimating *hukou*-based discrimination is different from estimating gender-based or ethnicity-based discriminations where the subjects cannot change their identity, which have less chance of self-selection bias. Therefore, models that are exclusively designed to estimate gender- and ethnicity-based discrimination such as Oaxaca–Blinder (OB) decomposition should not be directly applied to estimate *hukou*-based discrimination because results can be misleading.

5 For example, Chen and Hoy (2011) conducted face-to-face interviews with personnel management in 21 companies in Shanghai and found that they are concerned about the fact that people with rural *hukou* have local accents causing difficulties in communication, lack trust, and are highly mobile. These characteristics related to productivity may explain some of the remaining wage differentials.

6 For example, the CHIP data, which were used by Démurger et al. (2009) and Deng (2007), were collected from residential communities. Others such as Knight et al. (1999) used data from four cities in 1995, Gravemeyer et al. (2011) utilized survey from Shenzhen in 2005, Dong and Bowles (2002) used data from only two cities such as Dalian and Xiamen in 1998, and Lee (2012) used China Urban Labor Survey 2005, covered only urban areas.

7 According to China Family Panel Survey (CFPS) (2014) questionnaire, total income may include net income of agricultural production, profit from self-employment or operating private enterprises, various subsidies and alms from government, social donation, pension, wages of all family members, income from renting and selling properties, savings interests, and income from financial investments (Xie and Hu, 2014).

a nonparametric propensity score matching (PSM) strategy, which is a common program impact evaluation model stemming from Rosenbaum and Rubin's (1983) seminal work. The matching method should produce better results and offer the promise of causal inference with fewer assumptions (Guo and Fraser, 2015; Ho et al., 2007). Using nationally representative survey data, I found that *hukou*-based earning discrimination exists against agricultural *hukou* people. They only experience discrimination if others employ them; on the other hand, if they are self-employed, no statistical discrimination is discernable. Discrimination exists in both agriculture-related and nonagriculture-related professions; however, the worst discrimination exists in nonagriculture-related professions. They face earning disadvantages only in government jobs but not in private jobs. Agricultural *hukou* people experience higher wage discrimination where a labor contract is enforced compared to where a labor contract is not enforced. Within labor contract conditions and work types, if wage income dominates the total income, discrimination is evident; however, if nonwage income dominates the total income, discrimination is not apparent.

The findings confirm that earning difference by *hukou* is institutionally determined rather than simple rural–urban isolation.⁸ We can call it discrimination⁹ because this paper sheds further light on the extent of employer discrimination by *hukou* types by comparing the earnings of self-employed workers to their salary worker counterparts. If employer discriminations were a principal source of discrimination against agricultural *hukou* people, then we would expect the agricultural/nonagricultural *hukou* earnings ratios to be higher for the self-employed people compared to the people who are employed by others. No discrimination of this type is applicable to self-employed people.¹⁰ Similarly, by comparing workplace environments where labor contracts are applicable versus nonapplicable, we should expect agricultural/nonagricultural *hukou* earnings ratios to be higher for workplaces where labor contracts are not applicable compared to where labor contracts are applicable if there is no institutional discrimination. Theoretical and empirical evidence is ambiguous about whether discrimination should be higher or lower in public organizations compared to private organizations. Many argue that the public sector is nondiscriminatory because of the bureaucratic rules and regulations designed to ensure nondiscriminatory hiring and upgrading of public employees (Long, 1975). Alternatively, employer discrimination should be least in private sectors under competition because nondiscriminators can gain a competitive advantage by substituting lower cost minority workers for relatively higher cost mainstream, *ceteris paribus* (Arrow, 1973; Becker,

8 The earning difference by identity types in China is interesting because of its distinctive path dependency to state socialism. The socialist institutions such as the *hukou* system influence the welfare regimes very distinctively where the tools of the market economy are constrained by these socialist institutions (London, 2014). Chinese cities are experiencing labor shortage and recruiting illegal non-Chinese workers (Chan, 2010) while hundreds of millions of people stay in rural areas and earn very low incomes. Evidence also shows that there is surplus low-wage labor in the rural areas (Cai, 2007; Fields and Song, 2013; Knight et al., 2011; X. Zhang et al., 2011). This situation raises a simple question of why these rural workers choose not to migrate for higher paying jobs and thus had to work in the agricultural sector for considerably lower earnings. The reason is migration is costly and they face discrimination in the labor market due to their *hukou* status.

9 The labor market discrimination is defined as the valuation in the market place of personal characteristics of the worker that are unrelated to worker's productivity (Arrow, 1973). There are three main forms of labor market discriminations: wage discrimination, hiring discrimination, and premarket discrimination. While the wage discrimination is that two groups of workers with equal productivity receive different wages (Becker, 1957), the hiring discrimination is that they receive different access to the same jobs (Bertrand and Mullainathan, 2004). The premarket discrimination means comparable workers receive different levels of education and training that are likely to influence people's future labor market outcomes (Neal and Johnson, 1996).

10 See the work of Moore (1983) for further details. In the literature, people who are employed by others are called salary and wage workers. I used employed by others to be consistent with language of questionnaire.

1957).¹¹ However, we can expect that if higher discrimination exists in public employment, it is more institutionally designed compared to if higher discrimination exists in private employment, it is perhaps for low competition. There is also no theoretical foundation for discrimination in agriculture versus nonagriculture-related jobs. However, in China, nonagricultural sectors are more directed by the government than agricultural sectors; thus, if higher discrimination exists in nonagricultural professions compared to agricultural professions, it indicates institutional fault lines.

The next section presents the background of the *hukou* system transformation. Section 3 analyzes the problems with the parametric methods, proposes an alternative method, and describes data. Section 4 covers the empirical results, and section 5 concludes the paper.

2 Background of the *hukou* system and reforms

The modern *hukou* system, which was implemented in the late 1950s, is a modified version of the household registration system that has been a part of China for thousands of years. The purposes of the original and the current system are not different – restricting population movement and removing troublemakers, broadly maintaining social peace and order.¹² In the post reform, cheap labor has been a key factor of recent economic success for China and *hukou* is the instrument that helps the government maintain cheap labor for factories that are mostly owned by the government (Chan, 2009, 2010; Mackenzie, 2002; K. H. Zhang and Song, 2001). The local government officials believe that agricultural *hukou* people are the sources of cheap labor they need to grow their economies so that they can be promoted to higher positions. Thus, they oppress the migrant workers with lower wages and limited bargaining power. Another incentive of maintaining the *hukou* system is that if more people are settled in urban areas and if governments have to provide urban benefits such as education, health, housing, police, and others to all, it will be expensive as agricultural *hukou* people usually have a limited tax base. This is why the urban local governments will only accept outstanding migrants as their permanent residents.

To take advantage of the cheap labor, the Chinese Communist Party (CCP) began to lessen the strictness of the *hukou* system in 1978 when China first embraced the reforms of the market economy.¹³ Since then, China has been experiencing a dramatic change in the labor market. Workers, for the first time, could migrate to other cities or regions for employment purposes. In recent years, through a gradual reform process, the *hukou* system

11 Discriminatory firms should not be able to compete in the market for long run. On the other hand, competitive pressure and cost consciousness are generally less intense in the public organization, which implies that larger earning differential should be expected in public employment (Long, 1975).

12 Many others say that the primary logic of the *hukou* system was based on a division of labor. While, on the one hand, agricultural citizens are given access to arable land, non-agricultural citizens, on the other hand, rely on the state's allocation of resources and jobs and are expected to industrialize the nation (Chen and Fan, 2016). The reforms are since 1980s, which allowed people to migrate; now the "agricultural" and "non-agricultural" *hukou* distinction does not bear any necessary relationship (Cheng and Selden, 1994). There are more than hundreds of millions of Chinese working in urban industry while maintaining their "agricultural" *hukou* status (Fields and Song, 2013; Song, 2014).

13 The *hukou* system reforms can be divided into two broad phases: phase I (1958–1978) and phase II (1979–present). In phase I, the agricultural *hukou* holders were confined to agricultural works and they were not allowed to go to cities for working purposes (Chan, 2012). During this time, migration was limited including illegal immigration (H. X. Wu, 1994). The legal migration process was complex. During this prereform phase, the annual change rate was kept at a very low level of 0.15% to 0.20% (Y. L. Lu, 2002). In phase II, which started in the early 1980s, China implemented many policies that transferred the responsibility of managing the *hukou* system to local governments. Now local governments have full authority to set their own local *hukou* admission criteria (Song, 2014).

has been significantly relaxed by the local government authorities who were devolved the responsibility of managing the *hukou* system in the early 1980s (Cai, 2011; Chan, 2009; Song, 2014). In many cases, the cumulative effect of these reforms is not the abolition of the *hukou* system; it makes permanent migration of rural farmers to cities harder (Chan and Buckingham, 2008). The *hukou* system remains essentially potent and intact (Chan and Buckingham, 2008), and there is a unique labor market intervention that the CCP can play through the *hukou* system, which favors nonagricultural *hukou* over agricultural *hukou* inheritors.

Different city governments initiated different point-based *hukou* conversion systems. Big cities such as Beijing, Shanghai, Guangzhou, Shenzhen, Chengdu, and others make it tough by grading an application according to a point system based on an applicant's education level, tax payments, birth control, and work experience (Sheehan, 2017). In addition to education, higher tax return, and work experience, some local governments require an urban capacity fee, investment, and home purchase. Some cities introduced blue-stamped *hukou*, a special type of *hukou* different from regular red-stamped *hukou*, targeting migrants who can afford to purchase a house in their cities and whose talents are urgently needed for the city (L. Wu, 2013). Blue-stamped *hukou* can be converted to regular nonagricultural *hukou* in 3–5 years depending on city policies. There were also many *hukou* selling programs led by local urban governments to increase their monetary returns (Chen and Fan, 2016). By 1993, more than 3 million nonagricultural *hukou* were sold in exchange for 25 billion yuan (Han, 1994). The price of the *hukou* type and location vary. For example, 10,000 and 15,000 yuan used to be charged to nonagricultural *hukou* (transfer) and agricultural *hukou* (conversion), respectively, to purchase local *hukou* in Xiamen in 1993 (Chen and Fan, 2016). On the other hand, some cities, usually less attractive, offered monetary incentives to attract people. For example, according to Cixi's policy in 2008, if local agricultural *hukou* people convert to nonagricultural *hukou* status and give up their contract lands, they could obtain 24,300 yuan so that they can fully cover their pension insurance premium and subsidies for social insurance, loans, and urban housing. A 4,000 yuan reward was also available for doing these voluntarily in the province of Zhejiang (Chen and Fan, 2016).¹⁴

There are many other examples of *hukou* reforms in China. Thus, the *hukou* conversion and transfers range from very difficult (Beijing, Shanghai, etc.) to less difficult (Sichuan, except in the city of Chengdu) to even monetary rewards (cities of Zhejiang), depending on the attractiveness of the cities and the purpose of the city governments. The burdensome process of applying for urban *hukou* and the low likelihood of finding affordable housing in a first-tier city have discouraged many low-income migrants from taking advantage of these reforms. These reforms helped only well-educated elites to obtain nonagricultural *hukou* in big cities. Those who have been able to convert their *hukou* have experienced a large increase in subjective well-being (Tani, 2017). However, evidence shows that, although there is an economic cost associated with, there are no subjective well-being costs for holding an agricultural *hukou* (Asadullah et al., 2018).

¹⁴ Here the government's main incentive was land acquisition.

3 Empirical strategy and data

3.1 Problem with the existing parametric empirical strategy

The simplest way to find the earning difference by *hukou* types is by estimating equation (i) using ordinary least square (OLS) regression, where Y_i is the annual wage/income for individual i . $hukou_i$ is a dummy variable – if a person holds agricultural *hukou*, the value is equal to 1 and if he holds nonagricultural *hukou*, the value is equal to 0. I expected a negative coefficient for β_1 , which captures the impact of holding an agricultural *hukou* status. X_i is a vector of observable characteristics that can potentially explain the earning difference. I partially overcame the conventional limitations of the wage equation using a comprehensive nationally representative data set of CFPS where I can control for every individual's ability, motivation, attitude, and others. I also controlled for dummies of the province, employer types, and religion. β_2 captures the effects of all these control variables. α_0 is a constant capturing the unexplained portion of the earning differences, and ε_i is the error term.

$$Y_i = \alpha_0 + \beta_1 hukou_i + \beta_2 X_i + \varepsilon_i \quad (i)$$

OLS regression would have been more accurate if we were estimating the impact of *hukou* registration in the pre-1980s when the assignment of individual *hukou* status used to be determined by birth and people were not allowed to change their *hukou* status. After 1978, people can change their *hukou* status, and thus, they are more likely to convert their agricultural *hukou* to nonagricultural *hukou* to avoid institutional discrimination. As mentioned earlier, the conversion process was very selective, and thus, only people with higher ability and income were able to convert their *hukou*. This selective process clearly generates selection biases. Therefore, it is not possible to answer the question: does agricultural *hukou* status cause lower earnings or is the negative correlation we observe because of the result of poorer individuals who are not able to convert their *hukou* status? In this paper, I am interested not only in association per se between *hukou* status and individual earnings but also what it reveals about underlying causation.

Earlier literature on discrimination studies and specifically *hukou*-based discrimination used the decomposition approaches developed by Oaxaca (1973) and Blinder (1973).¹⁵ However, saying the entire unexplained portion is discrimination may not be appropriate because of the limitations with the OB decomposition approaches: misspecification due to differences in the support of the empirical distributions of individual characteristics for *hukou* status differences. While this is a problem largely recognized in the program impact evaluation literature, it has not received much attention in the analysis of wage discrimination (Heckman et al., 1997; Ñopo, 2008). Dolton and Makepeace (1987) and Munro (1988) identified the limitations of

15 The wage differentials can be analyzed using equation (ii).

$$W^R - W^U = (\bar{X}^R - \bar{X}^U)\hat{\beta} + \bar{X}^R(\hat{\beta}^R - \hat{\beta}^U) \quad (ii)$$

where R and U superscripts refer to agricultural and non-agricultural *hukou* people, respectively. W stands for wage/income. $\hat{\beta}$ are the linear estimates of the parameters β from the wage equation. It generates the counterfactual: what would an agricultural *hukou* inheritor earn if compensation scheme for his/her individual characteristics aligned with that of a non-agricultural *hukou* inheritor? Based on this counterfactual, the OB decomposition separates the wage differential between an agricultural and non-agricultural *hukou* holder into two components: explained and unexplained portions of the wage gap. $(\bar{X}^U - \bar{X}^R)\hat{\beta}$ is the explained portion of the earning gap, which can be attributed to differences in the mean observable characteristics. $\bar{X}^R(\hat{\beta}^U - \hat{\beta}^R)$ is the unexplained portion of the earning gap, which is often attributed to discrimination or different returns on productivity characteristics (L. Lee, 2012; Meng, 1998; Ñopo, 2008).

the original OB model that the wage gap decomposition is only informative about the average unexplained earning differences but not about its distribution. To address the distributional limitation, Buchinsky (1994) proposed the estimation of quintile earning equations and Jenkins (1994) and Hansen and Wahlberg (1999) proposed using generalized Lorenz curves for both observed earnings and predicted counterfactual earnings. However, their proposition still ignores the problem of *hukou* status differences in common supports. There are combinations of conditions and characteristics for which it is possible to find some people in the labor force while others are not with their *hukou* status. Again, people with agricultural *hukou* may concentrate on certain occupations that require access to certain resources such as land for agriculture professions and nonagricultural *hukou* people are more likely to be in the managerial occupations that require long tenure. Thus, it is not accurate to compare earnings across *hukou* status.¹⁶

So, the OB decomposition fails to recognize the *hukou* status differences in the support and distribution of explained earning differences while estimating earning equations for all working Chinese without restricting the comparison only to those individuals with comparable characteristics. The OB decomposition without this restriction is implicitly based on an “out-of-support assumption”, which becomes necessary to assume that the linear estimators of the earning equations are also valid out of the supports of individual characteristics for which they are estimated. This assumption should overestimate the component of the gap attributable to differences in the rewards (Ñopo, 2008).

Alternatively, we could treat *hukou* adoption as endogenous and use an instrumental variable estimator (IV). However, we don't have a good instrument in the dataset. Moreover, IV estimation procedures impose arbitrarily a linear functional form assumption that coefficients on control variables are restricted to be the same for both *hukou* types (Heckman and Navarro-Lozano, 2004; Jalan and Ravallion, 2003). Hence, we don't pursue this approach. Another parametric solution that allows a full set of interaction effects via the Heckman's selection correction model that come at the cost of imposing strong distributional assumptions such as the unobserved determinants of wage/income and *hukou* adoption are jointly normally distributed, with zero mean, constant variance, and a covariance term (Main and Reilly, 1993). Therefore, in the next section, I deal with an alternative nonparametric matching strategy to remove the restrictive assumption.

3.2 PSM estimator

PSM can overcome the selection bias problem and provide valid estimates of average treatment effects (ATE) as well as average treatment effects for the treated (ATT) with fewer assumptions (Guo and Fraser, 2015; Ho et al., 2007).¹⁷ Although matching techniques concerned with the comparison of groups with similar characteristics have been of special

¹⁶ This situation is analogous to the case of comparing wage across gender while there are male- and female-dominated occupations as discussed by Deutsch et al. (2005) about occupational segregation by gender in Latin America during the 1990s.

¹⁷ Rosenbaum and Rubin (1983) defined PSM as the conditional probability of assignment to a treatment given a vector of observed covariates. It does not require any estimation of earning equations and hence no validity-out-of-the-support assumptions.

interest to experimental design and statistics, they have also been widely used for nonexperimental designs (Ñopo, 2008). Rosenbaum and Rubin (1983) defined the ATE (n) in a counterfactual framework as

$$n = Y_i^R - Y_i^U \quad (\text{iii})$$

where Y_i^R and Y_i^U denote the earning of individual i with rural and urban *hukou*, respectively. In estimating the impact in equation (iii), a problem that arises is due to the fact that either Y_i^R or Y_i^U is observed, but not both for everyone. What is generally observed can be written as

$$Y_i = D_i Y_{iR} + (1 - D_i) Y_{iU}, \quad D = 0, 1 \quad (\text{iv})$$

Accordingly, we can rewrite the expression n as follows:

$$n = P \cdot [E(Y_R|D = 1) - E(Y_U|D = 1)] + (1 - P) \cdot [E(Y_U|D = 0) - E(Y_U|D = 0)] \quad (\text{v})$$

where P is the probability of observing an individual Chinese with $D = 1$ and n is the ATE. Equation (v) says that the impact of *hukou* system for the entire sample is the weighted average of the effect of *hukou* adoption in the two groups of individuals, those with agricultural *hukou* status or treated (the first term) and those with nonagricultural *hukou* status or controls (the second term), each weighted by its relative frequency. The main problem of arguing discrimination as causal inference stems from the fact that the unobserved counterfactuals, $E(Y_R|D = 1)$ and $E(Y_U|D = 1)$, cannot be estimated (Becerril and Abdulai, 2010; Heckman et al., 1998; Mendola, 2007; Smith and Todd, 2005). In an ideal case, experimental data would provide us with information on the counterfactuals that should solve the problem of causal inference. This is not such an ideal case, and available data do not provide information on the counterfactual situation, thus creating a problem of missing data (Blundell and Dias, 2000). This missing data situation requires estimating the direct consequence of the *hukou* registration system from the variation of income across individuals using statistical matching techniques. The current paper attempts to address this central problem of missing information on the counterfactuals by using the PSM method that summarizes the characteristics of everyone into a single index variable that determines the *hukou* registration status. Then, it uses this propensity score to match similar individuals (Rosenbaum and Rubin, 1983) to isolate the *hukou* registration effect from other socioeconomic determinants of individual-level income. The PSM can be expressed by

$$p(X) = \Pr[D = 1|X] = E[D|X]; \quad P(X) = F\{h(X_i)\} \quad (\text{vi})$$

Here, $F\{n\}$ can be logistic cumulative distribution and X is a vector that determines the *hukou* status. There are several assumptions for estimating the *hukou* effect on earnings using PSM. The first is the conditional independence assumption that states that *hukou* status inheritance is random and uncorrelated with income, once we control for X (Mendola, 2007). The second assumption is that ATE is only defined within the region of common support. This implies that the propensity score must be within 0 and 1; although this is done at the cost of reducing sample size, the quality of matches improves significantly by excluding the tails of the distribution of $p(X)$. Third, we assume that the selection to urban *hukou* was observable since the administrative offices that grant urban *hukou* to rural Chinese people based on

qualifications such as education, income, and others are observable individual characteristics. Officers do not have an opportunity to consider any unobservable characteristics that may favor someone to get an urban *hukou*. These assumptions ensure that people with the same X values have an equal possibility of being in both *hukou* groups (Heckman et al., 1997). When the propensity score is calculated, it should capture the similarities of individual Chinese to match each agricultural *hukou* adopter with his/her closest nonagricultural *hukou* adopter. There is a need to check if the matching procedure can balance the distribution of the relevant variables in the agricultural and nonagricultural *hukou* groups.¹⁸ The second step is to compute the differences of each pair of matched units, and then the ATE is obtained as the average of all these differences.¹⁹

Nearest neighbor matching (NNM) is the most straightforward matching method that identifies for each individual the closest twin in the opposite *hukou* status; then, it estimates the *hukou* status effect as the average difference of individual's earning between each pair of matched individuals.²⁰ NNM faces the risk of bad matches if the closest neighbor is far away. An imposition of a tolerance level on the maximum propensity score distance (caliper) can help to avoid the risk. The radius matching means that those individuals from the nonagricultural *hukou* group are chosen as matching partner for an agricultural *hukou* holder that lies within the chosen radius. It is often difficult to know a priori what choice for the tolerance level is reasonable (Smith and Todd, 2005). Caliper matching helps avoid bad matches, which is a crucial part of the successful application of the PSM approach.

3.3 Data and descriptive analysis

To obtain a generalizable result that has policy relevance should depend on whether sample data and their sampling procedures are well representative (Brunswik, 1956; Cappelen et al., 2015; Hogarth, 2005; Kruskal and Mosteller, 1979). This is more relevant for research in China where data can be manipulated to support authoritarian regimes as well as for its large geo-boundary and economic and regional fragmentation (Jian et al., 1996; Poncet, 2005; Wallace, 2016). This paper uses the survey wave of 2014 of the CFPS, which is the largest and most comprehensive social survey in China (Xie and Hu, 2014, p. 26) and has not been used in any earlier studies of *hukou*-based earning discrimination.²¹ The survey obtained its nationally representative

18 Caliendo and Kopeinig (2008) suggested that the basic idea is to compare the situation before and after matching and then check if there are any remaining differences after conditioning in the propensity score. After matching, there should be no systematic differences in the distribution of covariates. However, there still can be some "hidden bias" if there are unobserved variables that simultaneously determine the adoption of *hukou* and earnings. Thus, the use of a sensitivity analysis is necessary to address the issue (Rosenbaum, 2002).

19 Guo and Fraser (2015, pp. 48–50) emphasized the importance of estimating appropriate treatment effects using appropriate methods suitable for research questions. In this paper, the estimation of ATE is more appropriate because I am interested in the effect of *hukou* status at the population level where every individual Chinese is either agricultural or nonagricultural *hukou* Chinese. This research question is fundamentally different from medical research where ATT is more common because of specific medical intervention to a few people. Moreover, ATE is more efficient when there is sufficient overlap (common support) in the distribution of the estimated propensity scores (Hirano et al., 2003; Rubin and Thomas, 1992).

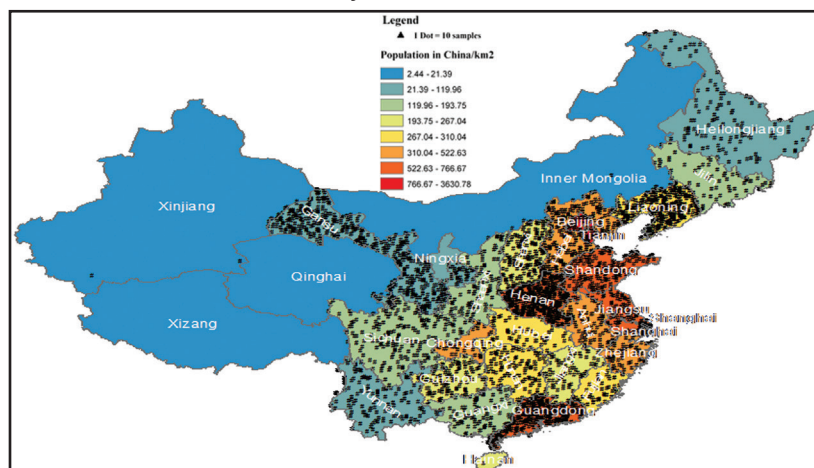
20 NNM can be applied with a replacement or without a replacement. Matching with a replacement involves a trade-off between bias and variance. If we allow replacement, the average quality of matching will increase, and the bias will decrease. The caliper/radius matching works in the same direction as allowing for replacement where bad matches are avoided and hence the matching quality increases. However, the variance of the estimates increases if there are fewer matches.

21 Institute of Social Science Survey, Peking University, 2015, "China Family Panel Studies (CFPS)", <https://doi.org/10.18170/DVN/45LCSO>, Peking University Open Research Data Platform, V31.

sample by integrating rural and urban areas and was collected from 25 provinces that represent 94.5% of the total population in Mainland China. Half of the observations came from the five large provinces of Shanghai, Liaoning, Henan, Gansu, and Guangdong with oversampling. For the remaining 20 provinces, one independent sampling frame was used (Xie and Hu, 2014). Wide-ranging information was collected through computer-assisted person-to-person interviews of all family members, and methods learned from the most influential survey projects in the world and their experiences (Xie and Hu, 2014) were utilized.²² Map 1 displays sample distribution across Chinese provinces. While the background color codes represent the Chinese population per square kilometer (blue to yellow to red colors shows low to medium to high population density, respectively), every black triangle dot represents 10 sample individuals. It clearly shows that more sample individuals were interviewed in highly dense areas. The eastern part of China is highly populated, and thereby, more interviews were taken place there.

Table 1 reports descriptive statistics as a mean test by *hukou* status for almost 35,000 individuals. I selected socioeconomic determinants of *hukou* adoption and earnings from a long list of variables in the original survey data on the basis of theories. About 73% of individuals in the dataset hold agricultural *hukou* and others hold nonagricultural *hukou*. However, only 43% of individuals are engaged in agriculture-related professions, which means almost half of the agricultural *hukou* inheritors are not engaged in agriculture-related professions. The purpose of this mean test is to check differences in baseline characteristics by *hukou* status. Any statistically different characters support the importance of individuals for *hukou* change and differences in earnings. Characteristics that influence earnings and *hukou* adoption consist of seven major groups: (i) demographic characteristics, e.g., gender, age, age squared, body mass index (BMI), and health; (ii) ability, e.g., schooling, training, math test, language test, and memorization; (iii) family, e.g., marriage, time spent on household work, and watching TV; (iv) socioeconomic background, e.g., CCP membership, unfair treatment, and social status; (v) migration status; (vi) ambitions, e.g., confidence about the future and interpersonal networking; and (vii) attitude, e.g., ideal number of children, helpful people, trustworthy people, and neighbors.

Map 1 Sample distributions. Notes: The author created this map using ArcMap 6.10 software. China's provincial shapefile was collected from ArcGIS online and sample data from CFPS 2014. CFPS, China Family Panel Studies.



²² See Xie and Hu (2014) for a more detail discussion about the representativeness of this dataset.

Table 1 Characteristics of agricultural and nonagricultural *hukou*: summary statistics

Variables	Mean-difference	Agricultural <i>hukou</i> = 1			Nonagricultural <i>hukou</i> = 0		
		Mean	St. Dev.	Min/Max	Mean	St. Dev.	Min/Max
Demographic characteristics							
Gender (1 = male, 0 = otherwise)	0.01	0.50	0.499	0/1	0.50	0.50	0/1
Age (years)	3.07*	44.76	17.39	16/104	47.83	17.42	16/96
BMI	-0.05*	1.39	0.22	0.67/3.2	1.34	0.22	0.71/3.34
Self-rated health status (1 to 5, 1 = excellent, 5 = poor)	0.05*	2.94	1.29	1/5	2.99	1.15	1/5
Ability							
Education (years of schooling)	3.51*	6.5	4.46	0/19	10.01	4.57	0/22
Nondegree training (no. of times)	0.21*	0.10	0.74	0/50	0.32	1.53	0/48
Math test score (math ability)	3.33*	6.92	5.71	0/24	10.25	6.58	0/24
Word test score (language ability)	6.46*	14.92	11.35	0/34	21.38	10.45	0/34
Memory test (1 to 5, 1 = better, 5 = worse)	-0.46*	2.69	1.31	1/5	2.24	1.17	1/5
Currently attending school (yes = 1, no = 0)	-1.52	0.07	0.26	0/1	0.07	0.26	0/1
Migration status							
Migrant (migrant = 1, nonmigrant = 0)	0.281*	0.10	0.30	0/1	0.38	0.49	0/1
Ambitions							
Future confidence (0 to 5, 0 = low, 5 = high)	-0.06*	4.05	1.03	1/5	3.99	0.98	1/5
Networking/relation (0 to 10, 0 = low, 10 = high)	0.03	7.24	1.85	1/10	7.27	1.72	1/10
Attitude							
Ideal number of children	-0.21*	2.07	0.80	1/10	1.87	0.68	1/10
People are helpful/selfish (helpful = 1, selfish = 0)	0.00	0.68	0.47	0/1	0.68	0.47	0/1
People are trustworthy (yes = 1, otherwise = 0)	0.04	0.52	0.50	0/1	0.57	0.50	0/1
Trust your neighbor (0 to 10, 0 = distrustful, 10 = very trustworthy)	-0.29	6.71	2.20	0/10	6.42	2.09	0/10
Family background							
Marital status (married = 1, otherwise = 0)	-0.01*	0.24	0.43	0/1	0.23	0.42	0/1
Time spent on household works (h)	-0.44*	2.14	2.01	0/24	1.71	1.66	0/21
Time spent on watching TV (h)	2.39*	10.92	9.74	0/154	13.31	11.49	0/112
Family's social status (1 to 5, 1 = very low, 5 = very high)	-0.14*	3.16	0.97	1/5	3.03	0.87	1/5
Work and employment							
Currently employment (yes = 1, no = 0)	-0.20*	0.79	0.40	0/1	0.59	0.49	0/1
Whether retired or not (yes = 1, no = 0)	0.23*	0.01	0.09	0/1	0.25	0.43	0/1
Occupation							
Sector of works (agriculture = 1, otherwise = 0)	-0.45*	0.54	0.5	0/1	0.09	0.29	0/1

(Continued)

Table 1 Continued

Variables	Mean-difference	Agricultural <i>hukou</i> = 1			Nonagricultural <i>hukou</i> = 0		
		Mean	St. Dev.	Min/Max	Mean	St. Dev.	Min/Max
Other socioeconomic background							
Whether CCP member (yes = 1, otherwise = 0)	0.10*	0.037	0.19	0/1	0.13	0.34	0/1
Faced unfair treatment by government (yes = 1, otherwise = 0)	-0.03*	0.12	0.32	0/1	0.08	0.28	0/1
Faced unfair treatment due to inequality of personal wealth (yes = 1, otherwise = 0)	-0.04*	0.13	0.34	0/1	0.09	0.28	0/1
Social status in the local area (1 to 5, 1 = very low, 5 = very high)	-0.14*	2.96	1.02	1/5	2.82	0.93	1/5
Number of deceased siblings	-0.01	0.2	0.58	0/7	0.19	0.59	0/10
Relative income in the local area	-0.04*	2.53	1.02	1/5	2.49	0.94	1/5

Notes: This is equivalent to bivariate regression as *hukou* is an independent variable. * $p < 0.05$. St. Dev., standard deviation; Min, minimum; Max, maximum; BMI, body mass index; CCP, Chinese Communist Party.

Although there is an equal number of males and females in the agricultural and non-agricultural *hukou* groups, their other demographic characteristics are different. While non-agricultural *hukou* people are older and have poorer self-reported health status, they have a lower BMI. For most categories of ability measures, nonagricultural *hukou* people stand higher than agricultural *hukou* people. Nonagricultural *hukou* adopters have a higher level of education and mathematical and language abilities. However, agricultural *hukou* people have a higher memorizing ability. While word and math tests are scores of respondents' cognitive ability based on the cognitive tests, memory test is the score of measuring an individual's ability to remember important things that happen to him/her within a week (Xie and Hu, 2014). There is no significant difference between the number of current school-going people by *hukou* types.

More nonagricultural *hukou* people are in migrant status than agricultural *hukou* people, and the difference is 28%, which is substantial. For the ambition group, while networking is not different by *hukou* groups, agricultural people reported that they are more confident about their future life than nonagricultural *hukou* people. By attitude, while the size of an ideal number of children is larger for rural *hukou* people, there is no significant difference between the two *hukou* groups whether they consider people are helpful/selfish and trustworthy and whether they trust their neighbors.

In terms of family background, a higher percentage of agricultural *hukou* people are in a marital relationship and they reported that they have a higher social status in their locality. While agricultural *hukou* people spend more time doing household works, nonagricultural people spend more time watching television. Considering socioeconomic background, higher CCP membership belongs to nonagricultural *hukou* and agricultural *hukou* holders reported that they face more unfair treatment by the government due to their low level of personal wealth. In terms of employment, agricultural *hukou* holders are more employed by 20% and in agricultural sectors by 45% than nonagricultural *hukou* holders. However, nonagricultural *hukou* people are more in retirement by 23%.

4 Empirical result analysis

This section presents the empirical results on the discriminatory consequences of the *hukou* system on the earnings of the Chinese citizens. In a counterfactual framework, the question would be as follows: how would picking out an individual at random in our sample and, going back in history, changing his/her *hukou* status alter his/her current earnings? I also investigated the impact of *hukou* within work ownership, employer types, work types, and labor contract conditions. I was interested in causal effect; thus, I selected PSM as the most appropriate estimation techniques here, however, I started with OLS multivariate estimation as baseline results and then moved to PSM estimation. Before we move to results, we should understand the dependent variable. We observe wages as well as income where income adds up wage earning and capital earning. Note that, contrary to standard wage/income function which is often in logarithmic form, we estimate both the parametric and nonparametric models on raw wages/incomes. This is because log-normal wage equation suffers from at least three problems: (i) the bias created by the logarithmic transformation, (ii) the failure of the assumption that all error terms have equal variance (homoskedasticity), and (iii) the sensitivity of research results to zero-valued wage/income (Burger et al., 2009; Flowerdew and Aitkin, 1982). Table 2 reports a straightforward comparison of annual wage and income by *hukou* status. It also separates the positive earners from the zero earners. The average gross annual wage and income of agricultural *hukou* holders are less by 6,251 and 6,914 yuan, respectively. These amounts are almost half of the wage and income of nonagricultural *hukou* holders. The difference is higher in incomes than in wages. The wage gap increases when comparing only positive wages after excluding zero earners but declines when comparing income gap among positive income earners. This indicates that agricultural *hukou* holders earn more from nonwage sources than nonagricultural *hukou* holders.

4.1 OLS results

4.1.1 Wage analysis

Table 3 presents the OLS regression results on the original wage. It shows that agricultural *hukou* holders are significantly disadvantaged compared to nonagricultural *hukou* holders. Therefore, *hukou*-based earning discrimination exists in the Chinese society. Looking at the details of the results, regressing *hukou* registration on wage in column (3.1) while controlling

Table 2 Wage and capital earnings by *hukou* inheritors

Variables	Mean difference	Agricultural <i>hukou</i> = 1			Nonagricultural <i>hukou</i> = 0		
		Mean	St. Dev.	Min/Max	Mean	St. Dev.	Min/Max
Wage	6,251.35*	7,327.07	15,873.77	0/442,000	13,578.42	25,530.21	0/408,400
Wage excluding zero earners	8,635.96*	24,322.81	20,569.24	1/44,2000	32,958.77	30,714.97	8/408,400
Income	6,914.38*	7,489.47	15,930.62	0/442,000	14,403.85	25,354.24	0/408,400
Income excluding zero earners	4,918.77*	15,563.41	20,043.15	1/442,000	20,482.17	28,100.46	1/408,400

Notes: * $p < 0.05$. St. Dev., standard deviation; Min, minimum; Max, maximum.

Table 3 Impact of Hukou status on wages: OLS estimates

	Model 3.1	Model 3.2	Model 3.3	Model 3.4	Model 3.5
Hukou status (agriculture = 1, non-agriculture = 0)					
Hukou status	-5,438.114*** (269.984)	-3,185.901*** (274.530)	-3,227.510*** (279.679)	-4,761.747*** (320.599)	-4,720.358*** (370.527)
Demographic characteristics					
Gender (men = 1)	6,428.181*** (206.449)	5,900.903*** (209.960)	6,010.022*** (214.962)	4,514.904*** (227.875)	4,611.116*** (228.265)
Age (years)	430.630*** (28.660)	-0.920 (34.092)	0.055 (34.619)	27.481 (40.876)	50.419 (40.998)
Age (years) squared	-7.424*** (0.292)	-2.609*** (0.319)	-2.678*** (0.326)	-2.191*** (0.383)	-2.379*** (0.384)
BMI	-913.519** (461.287)	117.785 (464.174)	155.132 (472.072)	-831.625* (472.471)	-736.267 (471.835)
Self-rated health status	-466.469*** (87.732)	-315.300*** (88.775)	-274.098*** (92.335)	-252.159*** (93.957)	-250.935*** (93.802)
Abilities					
Years of schooling		465.632*** (31.503)	469.460*** (32.128)	432.130*** (32.155)	425.256*** (32.015)
Nondegree training		2,265.070*** (346.046)	2,263.622*** (351.941)	2,060.109*** (352.226)	2,053.600*** (350.637)
Language ability (test score)		-66.088*** (14.505)	-66.511*** (14.645)	-25.768* (14.559)	-29.657** (14.546)
Mathematical ability (test score)		291.339*** (34.122)	293.906*** (34.334)	255.777*** (33.790)	251.335*** (33.678)
Memorizing ability		-216.803*** (80.333)	-200.820** (81.378)	-269.374*** (82.082)	-264.113*** (82.006)
Currently attending school		18,651.163** (7,435.794)	18,432.112** (7,505.840)	17,103.573** (7,546.310)	17,552.734** (7,513.705)
Ambitions					
Confidence about the future			-123.459 (90.951)	-77.527 (93.802)	-95.008 (93.734)
Interpersonal networking			262.077*** (55.514)	238.595*** (56.684)	233.468*** (56.682)
Attitude					
The ideal number of children			111.526 (128.507)	-28.739 (129.044)	-10.891 (128.888)
People are helpful/selfish			-46.385 (242.597)	-85.796 (244.075)	-72.969 (244.443)
People are trustworthy			-51.898 (232.388)	-143.117 (232.768)	-141.040 (232.414)
Trust your neighbour			22.048 (52.289)	11.868 (52.992)	15.381 (53.015)

(Continued)

Table 3 Continued

	Model 3.1	Model 3.2	Model 3.3	Model 3.4	Model 3.5
Family background					
Marriage (married = 1)				1,129.312*** (350.500)	1,225.275*** (349.948)
Time spend on household works				-824.048*** (46.924)	-808.497*** (46.819)
Time spend for watching TV				-101.935*** (9.755)	-100.011*** (9.738)
Family's social status				38.701 (117.766)	64.993 (117.466)
Other socioeconomic background					
CCP member (yes = 1)				1,300.839*** (490.545)	1,311.241*** (490.171)
Faced unfair treatment by the government				-301.924 (341.165)	-316.769 (341.359)
Faced unfair treatment due to inequality of personal wealth				-238.444 (303.369)	-302.730 (302.937)
Social status in the local area				303.015*** (117.461)	330.761*** (117.131)
Retirement (Yes = 1)				-8,929.176*** (421.436)	-9,039.505*** (420.279)
Migration status					
Hukou* migration					1,861.831** (746.307)
Migrant					1,528.756*** (572.376)
Province dummy					Yes
Religion dummy					Yes
Constant	21,722.64*** (2,743.592)	21,224.93*** (2,794.445)	19,639.16*** (2,924.843)	22,888.94*** (3,045.005)	21,195.53*** (3,047.077)
Adjusted R ²	0.15	0.21	0.21	0.23	0.23
Number of observations	30,622	28,936	28,481	27,211	27,206

Notes: The dependent variable is wage. Standard errors are robust (in parenthesis). * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

for demographic characteristics shows that agricultural *hukou* holders earn 5,438 yuan less than nonagricultural *hukou* holders. This amount is almost 1,000 yuan less than 6,251 yuan in the bivariate estimate as is shown in Table 2. All the control variables have expected direction of effects, and they are statistically significant in column (3.1). Male Chinese earns more than female Chinese by a large margin, which is consistent with recent wage estimates in China (Asadullah and Xiao, 2019). Older people earn more, but it is not consistent while adding more controls in the model. However, we find a significant and inverse U-shaped relationship between age and earnings.²³ Higher BMI reduces wages for Chinese people but is not statistically significant consistently. Health matters for wage earnings (Schultz, 2002). We find that, based on self-reported health status, people with poorer health earn significantly less, which is also consistent with other Chinese estimates (Chuanchuan, 2011).

All the model specifications in Table 3 control for Chinese provinces and major religions. There are observations from 25 provinces in China and controlling them should account for regional disparities including living standard, cost of living, regional economic development, and so on, which are essential, while regional inequality in China is one of the highest in the world (Fleisher et al., 2010; Zhang and Kanbur, 2005). The reference province is Beijing; people in most provinces such as Shandong, Hunan, Gansu, and so on earn significantly less compared to people in Beijing. On the other hand, in provinces such as Shanghai, Zhejiang, and Inner Mongolia, people's earnings are not significantly different from the earnings of people in Beijing (note that the coefficients of regional dummies are not reported here). For the religion dummy, there are seven different religious groups such as Buddha, Taoist deity, Muslim, Catholic, Jesus Christ, Ancestor, and others without religious beliefs. Buddha is the reference religion; it shows that there are no significant earning differences based on religious identity except Muslim people who earn significantly less than Buddhist people. This is consistent with earlier evidence that Muslims face systematic discrimination in their social, economic, and political rights from Han Chinese who are predominantly non-Muslims (Chuah, 2004; J. N. Smith, 2002).

An individual's ability can potentially determine both of his/her income and *hukou* status. Abilities often measured as schooling and different forms of ability test are widely proven important determinants of earnings (J. Angrist, 1998). It is also true for *hukou* status since China's most recent attempt to liberalize *hukou* has allowed rural *hukou* holders to apply for first-tier urban *hukou* based on a points system that considers educational attainment among other factors (A. Chen, 2018). Column (3.2) adds the most important list of control variables that measure ability such as years of schooling, training, cognitive skills, and memorizing capacity. These additional controls drop the coefficient of *hukou* by approximately 2200 yuan, but *hukou* status stays highly statistically significant. People with more years of schooling and higher training earn more on average that is well documented in the literature too (J. Angrist, 1998). Note that nondegree training rewards relatively more than education in the school since the coefficient of training is almost five times larger than the coefficient of years of schooling. This is because training is often directly related to the skills required for job performance; on the other hand, education is often very general and may not directly help to perform jobs well (Eck, 1993). Moreover, controlling for language and mathematical ability separately, which are part of

²³ Note that we do not have good measure of experience in the CFPS 2014. That is why we control for age and age squared, which is a close substitute to experience. The result is consistent with most recent literature (Asadullah and Xiao, 2019; Mishra and Smyth, 2013, 2015).

the education in school, may underestimate the true returns to education (Asadullah and Xiao, 2019, p. 90). A person currently attending school may also earn significantly more compared to a person not attending school. This is because while a person is employed and simultaneously enrolled in school, the educational program is likely to be related to job requirements. Highly mathematically able people earn more; however, people with better language and memorizing ability earn relatively less. The word and memorizing abilities do not have an expected direction of effects. The reasons might be that people with better language ability learned literature that did not help them earn much or that the test did not capture the true language ability. Similarly, memorizing ability measure was self-reported; thus, it might be the case that regional culture affected their self-reported statement as we see in Table 1 that agricultural *hukou* inheritors reported higher memorizing ability than nonagricultural *hukou* inheritors.²⁴

Column (3.3) includes controls for ambition and attitude-related variables; however, this additional controls increase the coefficient of *hukou* slightly. Measuring the actual ambition level for an individual is difficult. However, our dataset contains some variables that directly and indirectly represent how an individual is strongly ambitious by asking questions: how confident are you in your future and how good are you in maintaining interpersonal relations or networking. Controlling these variables allowed me to capture many personality characteristics that should matter for earning differences as well as whether to change *hukou* status. The results show that being good at networking or maintaining interpersonal relationships help Chinese individuals earning higher wages, but a higher level of confidence with future life affects them negatively. Theoretically, it does make sense that higher networking ability should be positively associated with earning (Pietro, 2007), but it does not make sense that higher confidence with the future is negatively associated with wages.²⁵ Instead, it may be that some people are temporarily in unexpected conditions and currently they are earning low wages, but they are confident to recover in the future is dominating the effect. Attitude is another important part of one's personality. Evidence about the impact of an individual's personality traits or attitudes on earnings in the literature is heterogeneous depending on the dimensions of attitudes (Heineck, 2011; Heineck and Anger, 2010; Semykina and Linz, 2007). For example, Heineck (2011) found among British people a positive relationship between openness to experience and wages but a negative linear relationship between wages and agreeableness. I controlled for variables that represent different dimensions of an individual's attitudes in column (3.3), but they are not statistically significant. They are mainly opinion type questions that have a substantial connotation of personality characteristics such as asking an ideal number of children; opinion whether people are helpful, selfish, or trustworthy' and acknowledgment whether they trust their neighbors or not. It turns out that these individual attitudinal characteristics are not strong predictors of Chinese people's wage earning.

Earlier literature claims a causal impact of family and socioeconomic background on an individual's earning and other achievements (Datcher, 1982; Ermisch and Francesconi, 2001; Loury, 1977; Meghir and Palme, 2005). I added controls for individual family and socioeconomic

²⁴ We did not control for occupation as it may disturb the coefficient of *hukou* if the occupation is an intermediary variable between *hukou* and earning (Angrist and Pischke, 2008). In another word, if an individual engages in agricultural farming because his/her *hukou* status requires, I should not control for occupation in the regression analysis, as it may be a bad control.

²⁵ Because empirical evidence such as Filippin and Paccagnella (2012) have shown that a small difference in self-confidence can lead to gaps in human capital and economic outcomes.

background differences, represented by nine different variables covering various aspects of family and socioeconomic background, in column (3.4). Most of them are statistically significant predictors of earning. These controls increase the coefficient of *hukou* by a large margin of more than 1,500 yuan. Marriage has a wage premium – empirical evidence shows that married workers tend to be in higher paying job grades; they receive higher performance ratings than single men; thus, married men are more likely to be promoted (Korenman and Neumark, 1991). On the other hand, divorce and separation have adverse economic consequences particularly on women (Duncan and Hoffman, 1985). Therefore, I controlled for marital status, and it shows that being in a marital relationship accounts for a substantial wage difference, more than a 1,100 yuan that is approximately 17% of the average wage gap between the two *hukou* groups. On the other hand, spending more time on household works and watching TV are associated with lower earnings. This is in connection with the pervasiveness of income–leisure tradeoffs in the literature (Battalio et al., 1981; Bielby and Bielby, 1988).

CCP members earn more than non-CCP members. Becoming a CCP member can bring a lot of advantages in the Chinese community and thereby should influence an individual's wage earnings.²⁶ The dataset allows me to control for this important characteristic of Chinese individuals, and it shows that CCP membership accounts for a substantial amount of earning differences, more than 1,300 yuan. Classical economists well understood that individuals are motivated at least partly by concerns about relative position. A large volume of empirical research demonstrates the relationship between relative position and well-being (Frey and Stutzer, 2002).²⁷ We control for the relative social status of family and individual him/herself. While the higher relative social status of an individual is positively associated with wages, the higher relative status of a family does not influence wages. If a person reported that he/she faced unfair treatment due to inequality of his/her personal wealth or experienced any unfair treatment from the Chinese government did not matter for differences in their wage earnings. The dataset also allows me to control who is in retirement. Retired people should earn a pension, not wages, but there are some people who took another job after retirement from their main profession. Results show that people who are retired earn less than those who are not retired.

Migrant Chinese earn more than the local Chinese residents (column 3.5). Note that these migrants can be from any possible four types of migrants: rural–urban, urban–urban, urban–rural, and rural–rural. The dataset shows that only 17% of people are in migrant status: 7% of them carry rural *hukou*, while the other 10% carry urban *hukou*. It is expected that rural–urban migrants are vulnerable and earn less than the local resident, which many earlier literature confirmed (Frijters et al., 2009; Gagnon et al., 2012; Lee, 2012). However, we cannot confirm this hypothesis from these findings as more migrants in this analysis hold an urban *hukou* and do not face considerable discrimination while living in a different city. Migrants who are carrying a rural *hukou* can be in either another rural or urban area. I considered interaction of migration status with *hukou* in column (3.5), which checks if an agricultural *hukou* holder becomes migrant and the coefficient of the interaction term is positive. This means migrants

26 Bian et al. (2001) found that party membership is positively associated with mobility into positions of political and managerial authority during the post-1978 reform era. Walder (1995) further explained that when a member of CCP has both educational and political credentials, he/she could have administrative posts with high prestige, considerable authority, and clear material privileges.

27 Luttmer (2005) found that, controlling for an individual's own income, higher earnings of neighbors are associated with lower levels of self-reported happiness. McBride (2001), Ferrer-i-Carbonell (2005), and Blanchflower and Oswald (2004) found tantalizing evidence that relative income affects subjective well-being.

earn more on average than local residents. This is interesting because if migrants earn more, then why more people do not migrate. The possible answer is migration in China is associated with many explicit and implicit costs (see Zhao, 1999, p. 777).

4.1.2 *Income analysis*

Another indicator of living standard is income where income can be either wage or nonwage such as capita earnings. Table 4 reports the OLS estimates of the effect of *hukou* registration on income. I used the same specifications here as was in Table 3. The overall results from Table 4 show that the agricultural *hukou* holders earn significantly lower incomes compared to non-agricultural *hukou* holders. This is equivalent to what we have seen in Table 3. As we expected, the magnitude of the discrimination is larger while estimating the effect on income: 5,640 yuan compared to 4,720 yuan in the full model. Along with *hukou* status, other stronger predictors of negative income of individual Chinese are retirement, being female, spending more time on household works and watching TV, poor health, and so on. On the other hand, positive income predictors are male gender, CCP membership, nondegree training, education, mathematical ability, migration, networking ability, being in marital relationship, and others. Note that in the income analysis we added additional control of whether individuals are employed or not since earnings from formal employment should differ from non-formal employment. It shows that if an individual is formally employed earns significantly higher by more than 6,500 yuan than if an individual is not employed. However, this strong control did not impact on the coefficient of *hukou*.

This is certainly good news for the regression result because this estimation has reasonable sets of controls to potentially avoid omitted biases. Econometricians tend to address this strong association as causal when adding and dropping some variables do not change the conclusion for key interested coefficient (J. D. Angrist and Pischke, 2008). One may be concerned about the low adjusted *R*-squared value of 23 for wage and 24 for income; note that other related literature that studied the impact of *hukou* also found low *R*-squared value ranging approximately from 15 to 25 (Q. Deng, 2007; Liu, 2005; Song, 2016). Note that the log-normal equation has an improved *R*-squared with a similar conclusion; however, we did not report this result as we believe that this result is less biased as discussed early in this paper. We cannot fully rely on this estimation because of selection biases as explained before. Thus, I move to PSM nonparametric estimation in the next section to estimate causality.

4.2 **Matching results**

4.2.1 *Wage and income analyses for all samples*

We need to specify the propensity scores of adopting a *hukou* before estimating the impact of *hukou* registration on earnings nonparametrically. I use a logit model, which is the most preferred model (Rosenbaum, 1986; H. L. Smith, 1997), to predict the probability to adopt a *hukou* type and include different individual-, family-, and society-level characteristics as regressors. There is differentiated advice in the literature regarding the inclusion of control variables in the PSM. While Rubin and Thomas (1996) suggested against “trimming” models in the name of parsimony, Sianesi (2004) and Smith and Todd (2005) recommended that the selection of

Table 4 Impact of hukou status on income: OLS estimates

	Model 4.1	Model 4.2	Model 4.3	Model 4.4	Model 4.5
Hukou status (agriculture = 1, non-agriculture = 0)					
Hukou status	-5,958.411*** (269.550)	-3,653.200*** (274.223)	-3,694.116*** (279.392)	-5,554.602*** (319.534)	-5,640.891*** (367.701)
Demographic characteristics					
Gender (men = 1)	6,405.167*** (205.890)	5,839.734*** (209.381)	5,949.340*** (214.377)	3,616.396*** (229.989)	3,711.284*** (230.173)
Age (years)	382.852*** (28.413)	-26.921 (33.849)	-26.742 (34.376)	-205.558*** (40.908)	-183.705*** (40.989)
Age (years) squared	-6.776*** (0.288)	-2.125*** (0.316)	-2.182*** (0.322)	0.694* (0.386)	0.529 (0.386)
BMI	-1,041.924** (459.608)	-52.315 (462.531)	-6.129 (470.449)	-1,053.251** (468.708)	-948.092** (467.905)
Self-rated health status	-470.691*** (87.550)	-315.736*** (88.580)	-270.613*** (92.132)	-96.887 (93.485)	-91.632 (93.323)
Abilities					
Years of schooling		476.476*** (31.406)	479.601*** (32.034)	423.711*** (31.731)	416.443*** (31.583)
Nondegree training		2,227.160*** (343.694)	2,225.981*** (349.543)	1,933.577*** (344.131)	1,925.112*** (342.590)
Language ability (test score)		-58.194*** (14.460)	-58.767*** (14.600)	-17.623 (14.389)	-22.079 (14.374)
Mathematical ability (test score)		284.402*** (33.990)	287.254*** (34.202)	240.450*** (33.300)	235.525*** (33.176)
Memorizing ability		-238.198*** (80.080)	-221.034*** (81.136)	-287.082*** (81.318)	-280.665*** (81.215)
Currently attending school		18,230.432** (7,403.289)	18,031.170** (7,475.821)	16,648.559** (7,430.966)	17,080.562** (7,393.865)
Ambitions					
Confidence about the future			-100.613 (90.650)	-127.343 (93.011)	-147.581 (92.909)
Interpersonal networking			266.600*** (55.340)	222.314*** (56.081)	215.905*** (56.059)
Attitude					
The ideal number of children			90.663 (127.982)	-70.074 (128.085)	-49.033 (127.899)
People are helpful/selfish			-59.193 (242.143)	15.385 (241.684)	30.712 (242.024)
People are trustworthy			-70.856 (231.839)	-237.843 (230.746)	-233.819 (230.349)
Trust your neighbor			23.572 (52.140)	18.075 (52.532)	22.091 (52.542)

(Continued)

Table 4 Continued

	Model 4.1	Model 4.2	Model 4.3	Model 4.4	Model 4.5
Family background					
Marriage (married = 1)				1,255.044*** (347.279)	1,361.699*** (346.628)
Time spend on household works				-767.354*** (46.194)	-749.054*** (46.081)
Time spend for watching TV				-82.188*** (9.648)	-79.758*** (9.630)
Family's social status				59.314 (116.302)	89.635 (115.969)
Other socioeconomic background					
CCP member (yes = 1)				1,191.020** (485.494)	1,199.951** (485.026)
Unfair treatment by the government				-453.871 (338.964)	-471.574 (339.152)
Unfair treatment due to inequality of personal wealth				-303.329 (302.872)	-376.583 (302.351)
Social status in the local area				195.777* (115.985)	226.894** (115.623)
Currently employed (yes = 1)				6,504.845*** (218.820)	6,615.103*** (218.986)
Retirement (yes = 1)				-5,143.453*** (407.785)	-5,230.073*** (406.947)
Migration status					
Hukou* migration					2,457.494*** (736.343)
Migrant					1,400.082*** (563.753)
Province dummy		Yes	Yes	Yes	Yes
Religion dummy		Yes	Yes	Yes	Yes
Constant	23,747.39*** (2,721.452)	22,502.46*** (2,773.622)	20,790.11*** (2,904.484)	23,285.62*** (3,013.113)	21,465.60*** (3,015.384)
Adjusted R ²	0.15	0.21	0.21	0.24	0.24
Number of observations	30,597	28,920	28,465	27,195	27,190

Notes: The dependent variable is income (wage + non-wage earnings). Standard errors are robust (in parenthesis). * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. OLS, ordinary least square; CCP, Chinese Communist Party.

covariates should be grounded on the theory that relates covariates to outcomes and treatment. Note that omitted variable bias produces inaccurate propensity scores (Baser, 2006, p. 379). Inclusion of variables that are weakly related to treatment (*hukou* adoption) usually reduces bias more than it increases variance when using matching, so under most conditions, these variables should be included (Heckman et al., 1997; Rubin and Thomas, 2000). Mendola (2007) had deployed the largest set of variables to estimate the effects of technology adoption on household well-being so that it makes less likely that unobservable characteristics remain out of the matching process. Both perspectives have merits although higher supports are for the inclusion of more variables even though they are weakly related.

In the logit model, I aimed at making agricultural and nonagricultural *hukou* inheritors more comparable based on scores that are built on several criteria. I took both the above perspectives carefully into consideration by selecting as much as covariates available in the dataset that have a theoretical connection to wage/income, *hukou* adoption or both. In Table 1 in section 3.3, we saw that most variables differ between agricultural and nonagricultural *hukou* people suggesting the absolute or relative subsistence or stoppage pressure of adopting nonagricultural *hukou* (or retaining nonagricultural *hukou*). Note that I only considered those demographics, ability, family, socioeconomic, and other variables that have theoretical connection to *hukou* status and earnings, and I saw that most of them are significantly related to earnings in section 4.1. I applied identical specification for logit model except that I dropped age squared (specification 1) and added employment variable while estimating impact of *hukou* adoption on income for the same reason explained in section 4.2 (specification 2). There is no science about what is the best model for propensity estimation. However, Rosenbaum and Rubin (1984) recommended an iterative approach for achieving covariate balance, and Diamond and Sekhon (2012) suggested that it should aim at maximizing the balance of covariates without limits. So, I have a third specification that includes the sector of works, relative income in the local area, and others (specification 3). They are important variables in the matching process because they differ by *hukou* types. So, specifications 1 and 2 are more parsimonious than specification 3, but they are useful in order to check the consistency of the causal impact (Smith and Todd, 2005). Note that I tried with several other specifications by adding and dropping some variables, but I choose these specifications as they are grounded on theories and maximize the balance of covariates (Diamond and Sekhon, 2012). Note that other specifications did not change the conclusion.

The results of the logit estimate of propensity scores are reported in Table A1 in the appendix. Most of the variables included have the expected signs, and they are mostly statistically significant. Migration, a response to income and nonincome differentials, is probably the most important variable that influences individual Chinese citizens whether to change their agricultural *hukou* to nonagricultural *hukou* (Mundlak, 1979). Migrants have already overcome parts of the psychic and nonpsychic costs of *hukou* conversion (Sjaastad, 1962). The result shows that migrants are less likely to keep their agricultural *hukou*, which is consistent with earlier studies in China (Zhao, 1999). All demographic characteristics are statistically significant except for self-reported health status. Females, younger people, and low BMI people tend to adopt more nonagricultural *hukou*. Younger people tend to migrate more to urban areas, and thus, they adopt more nonagricultural *hukou* (Zhao, 1999). More females may change their status through marriages as marriages let them to relocate to their husband's hometown

and take the husband's *hukou* status. Similarly, years of schooling, training, and math and word test scores that measure different types of abilities lead people to adopt nonagricultural *hukou* except memorizing ability. Schooling plays the most influencing role since it provides information advantage of job searches in different locations (Schwartz, 1973) and it reduces the psychic cost of migration (Sjaastad, 1962). This is consistent with what we have learned early in this paper that most nonagricultural *hukous* are rewarded based on different types of ability such as education.

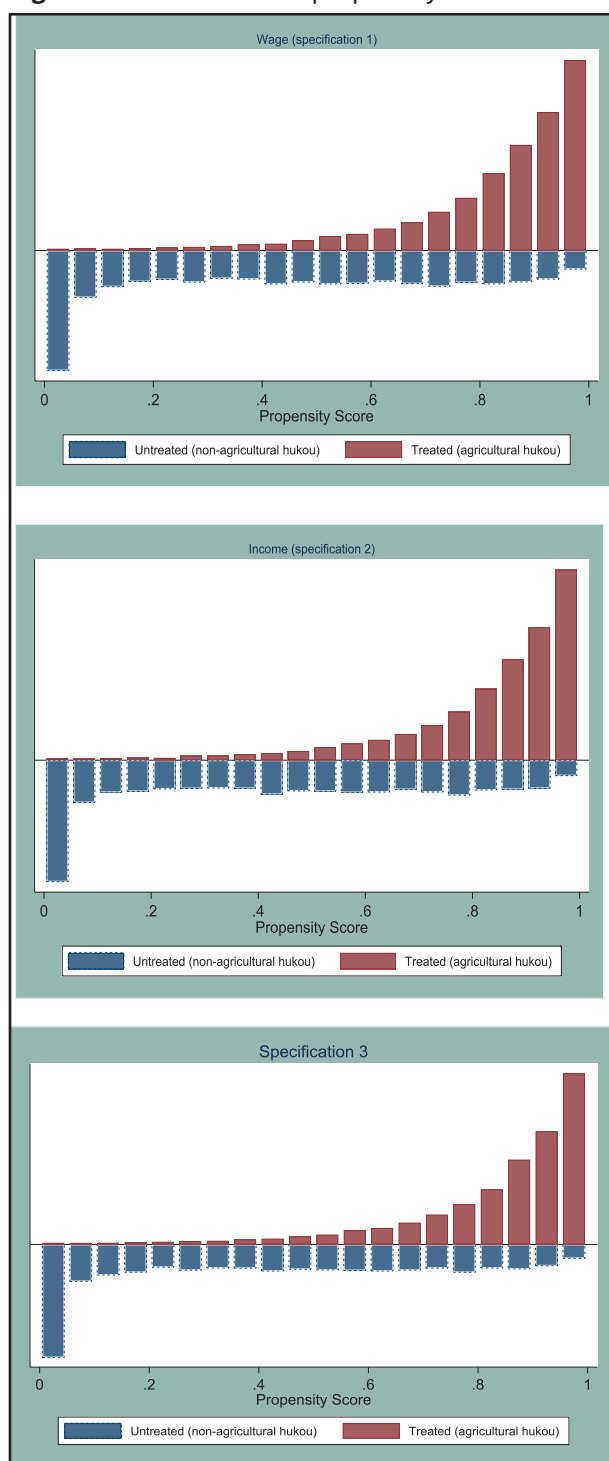
While higher confident people keep their agricultural *hukou*, people with better networks adopt nonagricultural *hukou*. People who trust their neighbors and consider a larger number of children as ideal tend to keep their agricultural *hukou*. For people who want more than one child, keeping their agricultural *hukou* was a big incentive, as in urban areas, the one-child policy was implemented strictly, while it was relaxed in rural areas (J. Zhang, 2017). That is why we see that the coefficient of the ideal number of children is relatively large and positive. Other attitudinal variables are not statistically significant. People who spend more time on household work tend to adopt agricultural *hukou*, but people who spend more time watching TV adopt nonagricultural *hukou*. Both individual and family social status are positively associated with agricultural *hukou* since relative status is probably more recognized in rural areas than in urban areas (Reiss, 1959). On the other hand, having a membership of CCP lead people to adopt nonagricultural *hukou* since it is much easier for CCP member than for non-CCP member to convert their status (Walder, 1995). Similarly, people who are in retirement tend to adopt more nonagricultural *hukou* since it is associated with higher pension and other benefits than agricultural *hukou* (London, 2014). Other socioeconomic features such as employment status, number of deceased siblings, and current school enrolment are also significantly associated with *hukou* types.

We should be less worried about which variables should be in the model of PSM when it satisfies some properties/tests such as balance check, bias reduction, and density of propensity score before and after matching (Diamond and Sekhon, 2012). The propensity scores only serve as a device to balance the observed distribution of covariates across agricultural and nonagricultural *hukou* groups. Therefore, the resultant balance assesses the reliability of the PSM application (Diamond and Sekhon, 2012; W. S. Lee, 2008). The common support conditions are imposed, and the balancing property is set and satisfied in all estimations at 1% level. The matching procedure is performed in the region of common support (Leuven and Sianesi, 2018). Figure 1 shows the distribution of the propensity scores and the region of common supports for all three specifications. It clearly reveals the significance of good matching, as well as the imposition of the common support condition to avoid bad matches.

While good balance is required for successful matching estimation, there is no guarantee that matching improves balance. The PSM application may make balance worse, even if covariates are distributed ellipsoidally because in a given finite sample, there may be departures from an ellipsoidal distribution (Sekhon, 2011). Therefore, to warrant the validity of the PSM analysis, it is pertinent to check the balance of covariates and the distribution of propensity score for agricultural and nonagricultural *hukou* before and after matching. Figure 2 reports the standardized percentage bias across selected covariates by graphs and a histogram.²⁸ It shows

²⁸ Both graphs and the histogram in Figure 2 are based on specification 1. They are identical under specifications 2 and 3.

Figure 1 Distribution of propensity scores and common support.



that matching has clearly improved the balance of covariates and minimized the bias. The bias of matched covariates is all within an acceptable threshold, while unmatched covariates have a significantly larger bias. Similarly, the kernel density plots in Figure 3 show that there is a large difference in the density of the two groups before matching and the difference is disappeared after matching in all three specifications. The matched samples are almost indistinguishable between agricultural and nonagricultural *hukou* groups, and they overlap entirely. This is

Figure 2 Plots of standardized bias reduction after matching.

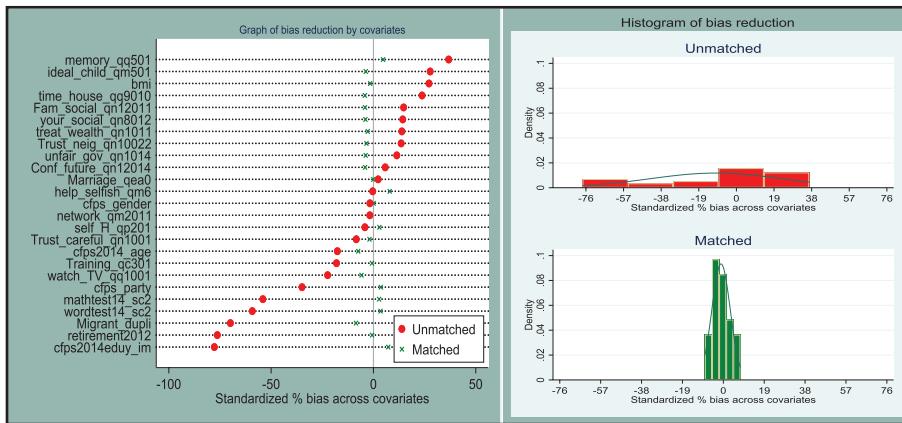
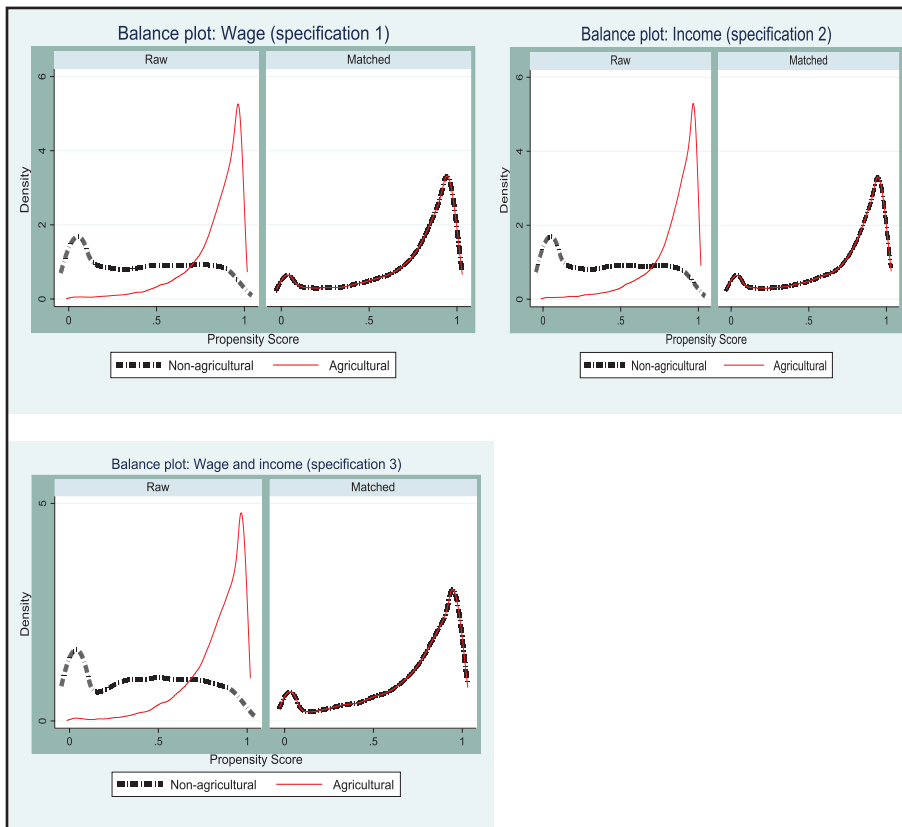


Figure 3 Density of the propensity scores before and after matching for all samples.



certainly good news for this estimation as these tests indicate a good balance of covariates between the two *hukou* groups.²⁹

I estimated the impact of *hukou* registration on wage and income by two different methods: NNM and caliper/radius matching. The robustness of the effect was verified using different number of neighbors to match (1 and 3) and radius (caliper of 0.01 and 0.05), and I found that the effect is not sensitive in terms of statistical significance, although it changes the size of

29 For specification 1 for whole-sample analysis, the total number of observation was 27,212 (model 5.1 in Table 5). Out of these, on-support untreated numbers were 7,785 and treated numbers were 4,488 and off-support untreated numbers were 0 and treated numbers were 14,939. Similarly, for specification 2, the total number of observation was 27,196. Out of these, on-support untreated numbers were 7,777 and treated number were 4,484 and off-support untreated numbers were 0 and treated numbers were 14,935. For common support graph, psgraph, and graph and histogram, pstest command were used in Stata.

Table 5 Impact of *hukou* status on earnings: matching estimates

Dependent variable		Model 5.1	Model 5.2	Model 5.3	Model 5.4	Model 5.5	Model 5.6
		Specifications 1 and 2		Specification 3		Specification 3	
		NNM 1	NNM 3	NNM 1	NNM 3	Calip 0.01	Calip 0.05
Wage	ATE	-3,814.18*** (357.02)	-3,561.34*** (330.72)	-4,062.46*** (420.31)	-4,285.88*** (362.24)	-4,062.46*** (420.31)	-4,062.46*** (420.31)
	Observation	27,212	27,212	24,890	24,890	24,890	24,890
Income	ATE	-4,606.76*** (396.00)	-4,348.29*** (337.49)	-4,743.90*** (449.87)	-4,703.81*** (379.67)	-4,743.90*** (449.87)	-4,743.89*** (449.87)
	Observation	27,196	27,196	24,875	24,875	24,875	24,875
Balancing property satisfied		Yes	Yes	Yes	Yes	Yes	Yes
Common support imposed		Yes	Yes	Yes	Yes	Yes	Yes
Treatment model		Logit	Logit	Logit	Logit	Logit	Logit

Notes: Standard errors are Abadie–Imbens' robust standard errors (in parenthesis). * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. While specification 1 was used for wages, specification 2 was used for income in models (5.1) and (5.2). For all others, specification 3 was used. NNM, nearest neighbor matching; ATE, average treatment effects.

earning differences slightly.³⁰ The results are presented in Table 5. Overall, matching estimates show that agricultural *hukou* has a robust negative effect on individual earnings. Earning difference by *hukou* status is slightly larger under specification 3 where a larger number of covariates were added to match (models 5.3 to 5.6). The NNM estimates for specifications 1 and 2 suggest that the impact of agricultural *hukou* adoption is about 3,561 to 3,814 yuan less in wage and 4,348 to 4,606 yuan less in income depending on the number of neighbors matched (models 5.1 and 5.2). The earning disadvantage increases for agricultural *hukou* inheritors while applying specification 3 with NNM (1) to 4,062 yuan in wage and 4,743 yuan in income. The amount is slightly lower for NNM (3) matching for income but larger for wage. The results for caliper of 0.01 and 0.05 are identical to NNM (1); this means there is a good match within the radius of 0.01. This confirms that there were not bad matches with the NNM method of matching. In other words, this nonparametric estimation was done between two very comparable people who held a different *hukou* status as the result is not sensitive to the choice of the radius.

We also see that the earning difference in income was larger than wage, which is consistent with the earlier OLS estimates in Tables 3 and 4. This is the average difference in earnings between a similar pair of individuals who belong to different *hukou* registration estimated as ATE. The PSM estimate for differences in wage and income is relatively conservative compared to the bivariate estimate in Table 2. While the bivariate estimate was 6,251 and 6,914 yuan in wage and income, respectively, the PSM estimate was approximately 4,000 to 4,300 yuan in wage and 4,700 yuan in income depending on the methods of matching. So, the difference is

30 I used *teffects psmatch* command in Stata as it has an important advantage over *psmatch2* command written by Leuven and Sianesi (2018) since propensity scores are estimated rather than known when calculating standard errors. This allowed me to calculate correct robust standard errors of Abadie and Imbens (2012), which offer an important correction to the standard errors of a sample mean when missing data are imputed using the "hot deck". Abadie and Imbens (2012) derived a method to estimate the standard errors of the estimator that matches on estimated treatment probabilities, and this method is implemented in *teffects psmatch*.

approximately 2,000 yuan between PSM and bivariate estimates. However, the PSM estimate is smaller than the OLS parametric estimate for income, although the estimate for wage is closer in the full model in Tables 3 and 4. Therefore, the OLS estimate overestimated the true effect of *hukou* registration for income.

4.2.2 Wage and income analyses by different groups and conditions

The CFPS data allowed me to investigate the earning disadvantages resulting from *hukou* registration types within work ownership, types of work, types of employers, and labor contract conditions. This further disaggregation provides many stimulating understanding of the consequence of the *hukou* registration system on the earnings of Chinese individuals. There are adequate representations of people of each *hukou* type in each group of work ownership, types of work, types of employers, and labor contract conditions as shown in Table 6 as cross-tabulation by *hukou* types. I checked the conditions of the balance of the pre- and postmatching for each group separately while estimating the impact of *hukou* within different groups. Figure 4 presents the kernel density plots for all four different groups: panel A for work ownership, panel B for work types, panel C for employer types, and panel D for labor contract conditions. Figure 4 confirms that there are good balances between agricultural and nonagricultural *hukou* people across covariates after matching. The agricultural *hukou* group is almost indistinguishable from the nonagricultural *hukou* group after matching, while there was a big variance before matching for each subsample. So, by matching based on the estimated propensity score, the comparability improves evidently within different groups by *hukou* types, which is the same as what we see in Figure 3.

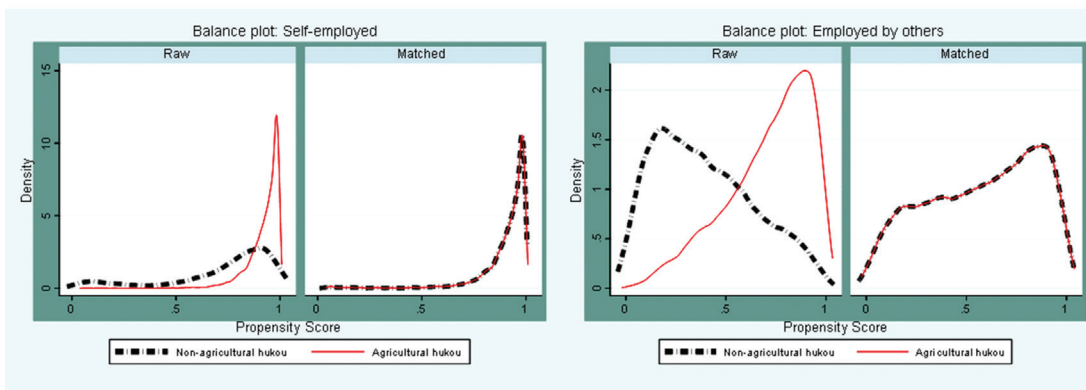
While Table 7 reports the estimated causal impact of *hukou* registration within types of work ownership and types of work, Table 8 reports the impact within types of employers and labor contract conditions. We find convincing evidence of *hukou*-based earning discrimination against agricultural *hukou* people. Considering work ownership, in the survey, people were asked to answer, “do you work for yourself/family or are you employed by others/organizations/units/companies?”. Their answer differentiated them as either self-employed if they answered, “work for myself/family” or employed by others if they answered, “employed by others/organizations/units/companies”. As it is expected, earning discriminations should exist only among the people who are employed by others but certainly not among those who are self-employed. This is because those discriminated against in the labor market should

Table 6 Cross-tabulation of groups samples by *hukou* registration types

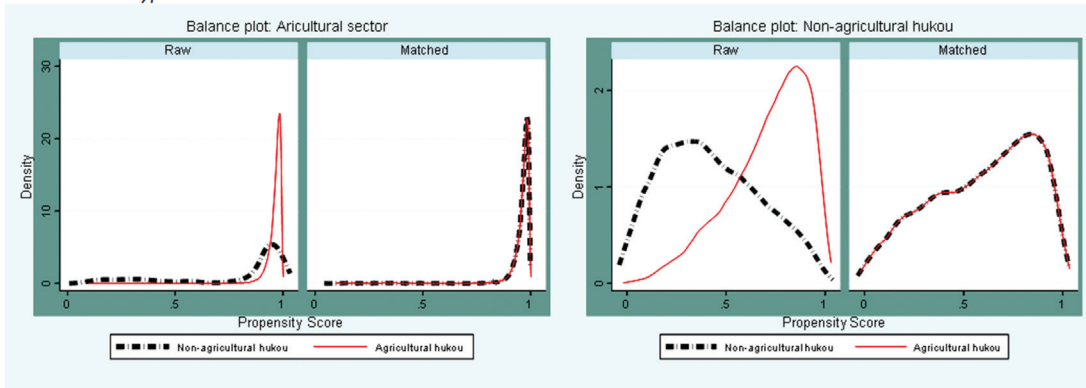
Categories	Position/sector	Nonagricultural <i>hukou</i>	Agricultural <i>hukou</i>	Total
Work ownership	Self-employed	1,311	12,878	14,189
	Employed by others	4,433	7,799	12,232
Work types	Agricultural	540	11,207	11,747
	Nonagricultural	5,218	9,507	14,725
Employer types	Government	1,960	1,015	2,975
	Private	1,876	5,191	7,067
Labor contract conditions	Applicable	4,226	6,331	10,557
	Not applicable	4,727	16,934	21,661

Figure 4 Density of the propensity scores before and after matching with disaggregation.

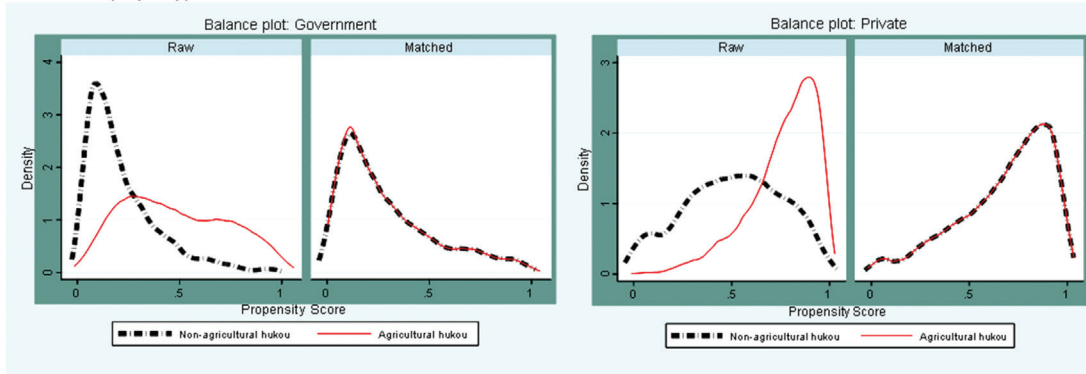
Panel A: Work ownership



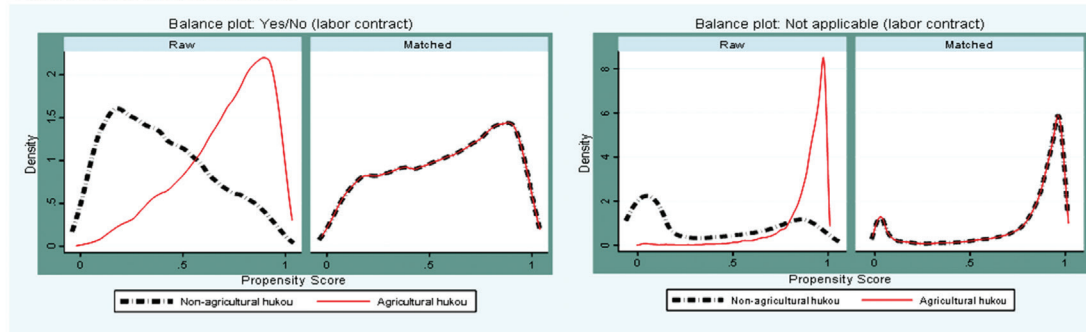
Panel B: Work types



Panel C: Employer types



Panel D: Labor contract condition



have a greater incentive to enter self-employment as discrimination lowers the expected wage in the labor market and thus lowers the opportunity cost of self-employment (Coate and Tennyson, 1992). The results show the expected effect of *hukou* registration as only people who are employed by others with an agricultural *hukou*

Table 7 Impact of *hukou* status on earnings across work ownership and types of works: matching estimates

	Model 7.1	Model 7.2	Model 7.3	Model 7.4	Model 7.5	Model 7.6	Model 7.7	Model 7.8
	Work ownership				Type of work			
	Self-employed		Employed by others		Agricultural		Nonagricultural	
	NNM 1	Caliper 0.05	NNM 1	Caliper 0.05	NNM 1	Caliper 0.1	NNM 1	Caliper 0.05
Wage (ATE)	560.82 (081.24)	560.82 (581.24)	-1796.77** (800.94)	-1796.77** (800.94)	-2218.81** (977.53)	-2218.81** (977.53)	-2924.92*** (625.92)	-2924.92*** (625.92)
Observation	11,912	11,912	8,421	8,421	9,884	9,884	10,481	10,481
Income (ATE)	302.51 (609.25)	302.51 (609.25)	-2056.72*** (783.93)	-2056.72*** (783.93)	-2295.04 (1653.54)	-2295.04 (1653.54)	-2714.25*** (630.09)	-2714.25*** (630.09)
Obs.	11,906	11,906	8,415	8,415	9,879	9,879	10,474	10,474
Balancing property satisfied	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Common support imposed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment model	Logit	Logit	Logit	Logit	Logit	Logit	Logit	Logit

Notes: Standard errors are Abadie–Imbens’ robust standard errors (in parenthesis). * $p < 0.1$ ** $p < 0.05$, and *** $p < 0.01$. Only specification 3 was deployed here. For the result of model 7.6, I increased the radius to 0.1 because there was not enough match within the radius of 0.05. NNM, nearest neighbor matching; ATE, average treatment effects.

people earn approximately 1,796 yuan less in wage and 2,056 yuan less in income compared to people who are employed by others with a nonagricultural *hukou* (models 7.3 and 7.4). On the other hand, people with agricultural *hukou* earn more if they are self-employed compared to self-employed non-agricultural *hukou* people, although the results are not statistically significant (models 7.1 and 7.2). Note that results are not sensitive to the choice of matching method and there is a good match within the radius of 0.05.

Moore (1983) predicted that self-employment was a method to avoid racist (or sexist) employment practices and should result in a higher black/white (or female/male) earnings ratio among the self-employed workers than their wage and salary counterparts. Therefore, taken together (results from models 7.1 to 7.4), higher earning for agricultural/non-agricultural *hukou* people while self-employed and lower earning for agricultural/non-agricultural *hukou* people while employed by others confirm Moore’s prediction about employer discrimination. In the Chinese context, an alternative interpretation is that the agricultural *hukou* people earn more because there are unique access and benefits that are tied to rural *hukou* including farming contract land, housing land, and compensation for land requisition (Lu and Song, 2006). They are considered valuable assets to which urban *hukou* people do not have access. Thus, when rural *hukou* people engage in independent business or self-employment, they do better than their urban *hukou* counterparts because of their initial endowment resources. Overall, earning differences within work ownership provide strong evidence of employer discrimination and self-employed people are able to avoid it.

In models 7.5 to 7.8, I also found suggestive evidence of *hukou*-based discrimination against agricultural *hukou* people across agricultural and nonagricultural works. People

Table 8 Impact of *hukou* status on earnings across employer types and labor contract conditions: matching estimates

	Model 8.1	Model 8.2	Model 8.3	Model 8.4	Model 8.5	Model 8.6	Model 8.7	Model 8.8
	Types of employer				Sign labor contract			
	Public		Private		Yes/No (applicable)		Not applicable	
	NNM 1	Caliper 0.05	NNM 1	Caliper 0.05	NNM 1	Calip 0.05	NNM 1	Caliper 0.05
Wage (ATE)	-4351.77*** (1375.46)	-4351.77*** (1375.46)	-807.17 (817.79)	-807.17 (817.79)	-2604.17*** (810.65)	-2604.17*** (810.65)	527.86*** (161.46)	527.86*** (161.46)
Observation	2,331	2,331	6,099	6,099	8,423	8,423	16,460	16,460
Income (ATE)	-3839.77*** (1335.82)	-3839.77*** (1335.82)	-1167.61 (794.03)	-1167.61 (794.03)	-2196.11*** (777.48)	-2196.11*** (777.48)	271.6 (165.13)	271.6 (165.13)
Observation	2,330	2,330	6,094	6,094	8,417	8,417	16,451	16,451
Balancing property satisfied	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Common support imposed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment model	Logit	Logit	Logit	Logit	Logit	Logit	Logit	Logit

Notes: Standard errors are Abadie-Imbens' robust standard errors (in parenthesis). * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Only specification 3 was deployed here. NNM, nearest neighbor matching; ATE, average treatment effects.

were asked: “Is your job an agricultural job or a nonagricultural job?”. From their answers, I can further investigate *hukou* registration impact by types of work. As explained in the questionnaire, an agricultural job includes works related to forestry, stock farming, fishing, and other sideline productions. Anecdotally, we can expect that agricultural *hukou* holders should face less or no earning discrimination in agricultural jobs as opposed to nonagricultural jobs. The results in Table 7 show that, while agricultural *hukou* holders face earning disadvantage in both agriculture and nonagriculture-related professions, the size of earning disadvantage is higher in nonagriculture-related professions than in agriculture-related professions. In the nonagriculture-related jobs, the earning difference is 2,924 yuan in wage and 2,714 yuan in income, and in the agriculture-related jobs, the earning difference is approximately 2,218 yuan in wage and 2,295 yuan in income. Note that result for income is not statistically significant, which gives an indication that nonwage earning dominates the impact here. This further indicates the institutional discrimination, and agricultural *hukou* people engage more in nonwage earning to avoid this institutional discrimination. However, whoever with an agricultural *hukou* stays in waged job cannot avoid discrimination. In other words, nonagricultural *hukou* people are a privileged group not only in the non-agriculture-related professions but also in agriculture-related professions. Thus, overall results in Table 7 tell us that earning discrimination is not a random incident rather it occurs systematically to agricultural *hukou* inheritors as we understand the issue that they face discrimination when others employ them but not when they are self-employed, and it exists in both agricultural and nonagricultural professions.

Table 8 reports estimated the impact of *hukou* status on earnings within employer types and conditions of a labor contract. By employer types, I checked whether agricultural *hukou* people face discrimination in public or private organizations or both. Public sector jobs include government, political party (CCP), people’s organization, and all forms of SOE. Private organizations include only private enterprises and do not include international organizations, residential communities, charities, etc. The result shows that the earning disadvantage for agricultural *hukou* people exists only in public organizations and there is no statistical discrimination in private enterprises for the same *hukou* people. If rural *hukou* people work for a government organization including SOEs, they earn approximately 4,351 yuan less in wage and 3,839 yuan less in income than if urban *hukou* people with similar characteristics work for the same government organizations (models 8.1 and 8.2). The earning gap by *hukou* status is 807 yuan in wage and 1,167 yuan in income when they work in private enterprises, but this earning difference is not statistically significant (models 8.3 and 8.4). These are not surprising results because CCP has been continuously pursuing economic policies since Chinese economic reforms started, which are extremely biased toward urban people. These urban-biased policies are even strongly enforced in SOEs where people with rural *hukou* have minimal access to any top positions (Yang, 1999, 2002). Therefore, it seems that there is severe earning discrimination in government jobs and privatization reforms have potentially reduced the earning gaps between the two *hukou* groups through competition. The finding is consistent with some of the existing Chinese literature; for example, Song (2016) concluded that for observationally equivalent workers, agricultural *hukou* people earn 50% less than nonagricultural do in the SOE, but only 5% less in the private sector. It is also supportive of the prediction of Becker (1957) and Arrow (1973) that the employer discrimination is less intense under competition because non-discriminators can gain a competitive advantage by substituting lower cost agricultural *hukou*

people for relatively higher cost nonagricultural *hukou* people, *ceteris paribus*. It differs from the hypothesis of Long (1975) that the public sector is nondiscriminatory.

Another important condition where impacts of *hukou* status should differ and provide evidence of institutionally designed discrimination is the signing of a labor contract. Signing a labor contract represents particular types of workplace environments that are expected to be meaningfully different from the workplace environments where signing a labor contract does not apply. In the CFPS survey, people were asked: “Do you sign a labor contract for this job?”. They responded as “yes/no” as well as “not applicable”. The labor contract condition is a contract between labor and owner/management governing wages and benefits and working conditions that may or may not favor employees and employers. On the other hand, the “not applicable” option implies different workplace environments such as informal jobs, self-employment, an entrepreneurial venture, and a person running his/her own business. There is empirical evidence that a better understanding of a labor contract is associated with improved workers’ satisfaction. However, a labor contract comes with reduced wages and sometimes coexisted with right violations against workers (Z. Cheng et al., 2014). It is expected that earning discrimination should exist only where labor contract is applicable because of biased institutions, legal systems, contracts, and others, which may favor high-class urban *hukou* people and discriminate against low-class rural *hukou* people. On the other hand, in a workplace where a labor contract is not applicable, people tend to be more independent and have fewer constraints, and thus less chance of earning discrimination by *hukou* status. Results in Table 8 are consistent with our expectation that if a labor contract is applicable, agricultural *hukou* people earn significantly less than nonagricultural *hukou* people. The ATE of *hukou* status is approximately 2,604 yuan in wages and 2,196 yuan in incomes (models 8.5 and 8.6). On the other hand, if a labor contract condition is not applicable, agricultural *hukou* people indeed earn more than nonagricultural *hukou* people in wage although they are equal in income (models 8.7 and 8.8). Note that this is not the impact of a labor contract but rather the impact of different employment and workplace environments where people are discriminated against based on their *hukou* identity. Thus, we observe further signs of institutionally enforced discrimination where people with nonagricultural *hukou* are offered better wages at the expense of agricultural *hukou* people’s wages.

5 Conclusions

This study focuses on estimating earning discrimination against agricultural *hukou* people in China. Chinese citizens with agricultural *hukou* are found to be treated unfavorably in terms of both wage and nonwage earnings. Both parametric and nonparametric estimations reveal that the *hukou*-based earning gap is overwhelmingly generated not only between the two *hukou* types overall but also within sectors, job categories, employer categories, and labor contract conditions. The results in Tables 7 and 8 provide strong evidence that the earning difference between the two *hukou* groups is earning discrimination by *hukou*-based identity. To believe this is earning discrimination, results in Tables 7 and 8 for subsamples are more compelling than the results in Table 5 for all samples. Evidence of earning differences within work ownership, work types, employer types, and labor contract conditions gives a clear signal of

systematic discrimination. Such disaggregation should eliminate the doubts that the earning disadvantage for people with agricultural *hukou* is due to rural–urban segregations rather than *hukou* identity. Alternatively, if the earning disadvantage for agricultural *hukou* people is due to rural–urban segregation, people should earn equal within the same occupation and people should change their status and join the high-earning groups. However, there is a systematic barrier, and this barrier is institutionally cultivated in China. The full answer is certainly more than just a *hukou* system, but *hukou* is an important one.

The consequence of this earning discrimination in the long run against a particular group of people is more than just widening income and social inequality. It is an unfair system and a major barrier toward a modern welfare state. Earning, either wage or nonwage, is the foremost spine of maintaining the standard of living. While a group of people can earn high, at the same time, another group cannot in the same society and in the same occupation; the sufferings of the disadvantaged group multiply through market forces such as the cost of living. The only exception is if some form of social protection programs protects these low-earning people, but evidence shows that agricultural *hukou* holders have lower access to social security programs, particularly when they are not in their *hukou* location (Nielsen et al., 2005; Xu et al., 2011). This finding has many policy implications. Considering that the *hukou* system has many advantages in the Chinese society, if the Chinese government does not want to abolish this old system, it should focus on reducing *hukou*-based discrimination in the labor market, particularly in the government organization. City government should stop granting urban *hukou* only to the super smart and rich people. Rather, it should be accessible on the basis of needs; otherwise, human capital divisions between rural and urban areas will worsen.

Although this paper provides improved evidence of earning discrimination against agricultural *hukou* people, it is not without limitations. This is self-reported data by individual adults; thus, it does not capture the underground economic activities (Becker, 1968). It may be that rural and urban *hukou* people engage in underground economy differently that may lead them reporting lower wages/incomes. There is also a possibility that people may hide their identity to avoid legal restriction or discrimination (Deng and Cordilia, 1999; Dutton, 1997; Wang, 2004), and thus, they are less likely to appear on the social survey.

Acknowledgments

I am grateful to anonymous reviewers for their helpful comments and suggestions. I am also grateful to Professor John Earle and Naoru Koizumi for their helpful feedback. I am indebted to Matthew Lundgren for addressing English grammars.

Availability of data and material

The dataset was used in this study is available in the Peking University Open Research Data Platform, V31 repository at <https://doi.org/10.18170/DVN/45LCSO>. The dataset can also be available from the author on reasonable request.

Competing interests

The author declares that he has no competing interests

Funding

There was no financial support for this research project

Authors' contributions

All the work in this paper was performed by Abu Bakkar Siddique

References

- Abadie, A.; G. W. Imbens (2012): A Martingale Representation for Matching Estimators. *Journal of the American Statistical Association* 107(498), 833-843.
- Afridi, F.; S. X. Li; Y. Ren (2015): Social Identity and Inequality: The Impact of China's Hukou System. *Journal of Public Economics* 123, 17-29.
- Angrist, J. (1998): Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants. *Econometrica* 66(2), 249-288.
- Angrist, J. D.; J.-S. Pischke (2008): *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Arrow, K. J. (1973): The Theory of Discrimination, in: Ashenfelter, O.; A. Rees (eds.), *Discrimination in Labor Markets*, Princeton University Press, 3-33.
- Asadullah, M. N.; S. Xiao (2019): Labor Market Returns to Education and English Language Skills in the People's Republic of China: An Update. *Asian Development Review* 36(1), 80-111. doi:10.1162/adev_a_00124.
- Asadullah, M. N.; S. Xiao; E. Yeoh (2018): Subjective Well-Being in China, 2005–2010: The Role of Relative Income, Gender, and Location. *China Economic Review* 48, 83–101. doi:10.1016/j.chieco.2015.12.010.
- Au, C.-C.; J. V. Henderson (2006): How Migration Restrictions Limit Agglomeration and Productivity in China. *Journal of Development Economics* 80(2), 350-388. doi:10.1016/j.jdeveco.2005.04.002.
- Baser, O. (2006): Too Much Ado about Propensity Score Models? Comparing Methods of Propensity Score Matching. *Value in Health* 9(6), 377-385. doi:10.1111/j.1524-4733.2006.00130.x.
- Battalio, R. C.; L. Green; J. H. Kagel (1981): Income-Leisure Tradeoffs of Animal Workers. *The American Economic Review* 71(4), 621-632.
- Becerril, J.; A. Abdulai (2010): The Impact of Improved Maize Varieties on Poverty in Mexico: A Propensity Score-Matching Approach. *World Development* 38(7), 1024-1035. doi:10.1016/j.worlddev.2009.11.017.
- Becker, G. S. (1957): *The Economics of Discrimination*. University of Chicago, Chicago.
- Becker, G. S. (1968): Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76(2), 169-217.
- Bertrand, M.; S. Mullainathan (2004): Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review* 94(4), 991-1013. doi:10.1257/0002828042002561.
- Bian, Y.; X. Shu; J. R. Logan (2001): Communist Party Membership and Regime Dynamics in China. *Social Forces* 79(3), 805-841. doi:10.1353/sof.2001.0006.
- Bielby, D. D.; W. T. Bielby (1988): She Works Hard for the Money: Household Responsibilities and the Allocation of Work Effort. *American Journal of Sociology* 93(5), 1031-1059. doi:10.1086/228863.
- Blanchflower, D. G.; A. J. Oswald (2004): Well-Being Over Time in Britain and the USA. *Journal of Public Economics* 88(7), 1359-1386. doi:10.1016/S0047-2727(02)00168-8.
- Blinder, A. S. (1973): Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources* 8(4), 436-455. JSTOR. doi:10.2307/144855.
- Blundell, R.; M. C. Dias (2000): Evaluation Methods for Non-Experimental Data. *Fiscal Studies* 21(4), 427-468. doi:10.1111/j.1475-5890.2000.tb00031.x.
- Bowles, S.; H. Gintis; M. Osborne (2000): The Determinants of Earnings: Skills, Preferences, and Schooling. Working Paper No. 2000-07.
- Bowles, S.; H. Gintis; M. Osborne (2001a): Incentive-Enhancing Preferences: Personality, Behavior, and Earnings. *The American Economic Review*; Nashville 91(2), 155-158.
- Bowles, S.; H. Gintis; M. Osborne (2001b): The Determinants of Earnings: A Behavioral Approach. *Journal of Economic Literature* 39(4), 1137-1176. doi:10.1257/jel.39.4.1137.
- Brunswik, E. (1956): *Perception and the Representative Design of Psychological Experiments*. University of California Press.
- Buchinsky, M. (1994): Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression. *Econometrica* 62(2), 405-458.
- Burger, M., van Oort, F., & Linders, G.-J. (2009): On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation. *Spatial Economic Analysis* 4(2), 167-190. doi:10.1080/17421770902834327.
- Cai, F. (2007): The Myth of Surplus Labor Force in Rural China. *Chinese Journal of Population Science* 2(2), 7.
- Cai, F. (2011): Hukou System Reform and Unification of Rural-Urban Social Welfare. *China & World Economy* 19(3), 33-48. doi:10.1111/j.1749-124X.2011.01241.x.
- Caliendo, M.; S. Kopeinig (2008): Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys* 22(1), 31-72. doi:10.1111/j.1467-6419.2007.00527.x.

- Cappelen, A. W.; K. Nygaard; E. Ø. Sørensen; B. Tungodden (2015): Social Preferences in the Lab: A Comparison of Students and a Representative Population. *The Scandinavian Journal of Economics* 117(4), 1306-1326. doi:10.1111/sjoe.12114.
- Chan, K. W. (2009): The Chinese Hukou System at 50. *Eurasian Geography and Economics* 50(2), 197-221. doi:10.2747/1539-7216.50.2.197.
- Chan, K. W. (2010): The Household Registration System and Migrant Labor in China: Notes on a Debate. *Population and Development Review* 36(2), 357-364. doi:10.1111/j.1728-4457.2010.00333.x.
- Chan, K. W. (2012): Crossing the 50 Percent Population Rubicon: Can China Urbanize to Prosperity? *Eurasian Geography and Economics* 53(1), 63-86. doi:10.2747/1539-7216.53.1.63.
- Chan, K. W.; W. Buckingham (2008): Is China Abolishing the Hukou System?*. *The China Quarterly* 195, 582-606. doi:10.1017/S0305741008000787.
- Chen, A. (2018, May 29): Bridging the Gap: Why China's Hukou System Needs Reform. *Brown Political Review*. <http://brownpoliticalreview.org/2018/05/bridging-gap-chinas-hukou-system-needs-reform/>.
- Chen, C.; C. C. Fan (2016): China's Hukou Puzzle: Why Don't Rural Migrants Want Urban Hukou? *China Review* 16(3), 9-39.
- Chen, Y.; C. Hoy (2011): Explaining Migrants' Economic Vulnerability in Urban China. *Asian Population Studies* 7(2), 123-136. doi:10.1080/17441730.2011.576816.
- Cheng, T.; M. Selden (1994): The Origins and Social Consequences of China's Hukou System*. *The China Quarterly* 139, 644-668. doi:10.1017/S0305741000043083.
- Cheng, Z.; H. Wang; Y. Chen (2014): Labor Contract, Trade Union Membership, and Workplace Relations: A Study of Migrant Workers in Guangdong Province, China, in: Hao, Z.; S. Chen (eds.), *Social Issues in China: Gender, Ethnicity, Labor, and the Environment*. Springer 183-206. doi:10.1007/978-1-4614-2224-2_10.
- Chuah, O. (2004): Muslims in China: The Social and Economic Situation of the Hui Chinese. *Journal of Muslim Minority Affairs* 24(1), 155-162. doi:10.1080/1360200042000212133.
- Chuanchuan, Z. (2011): Empirical Analysis on Impact of Health Change on Labor Supply and Income. *Economic Review* 4.
- Coate, S.; S. Tennyson (1992): Labor Market Discrimination, Imperfect Information and Self Employment. *Oxford Economic Papers* 44(2), 272-288. doi:10.1093/oxfordjournals.oep.a042046.
- Datcher, L. (1982): Effects of Community and Family Background on Achievement. *The Review of Economics and Statistics* 64(1), 32-41. JSTOR. doi:10.2307/1937940.
- Démurger, S.; M. Gurgand; S. Li; X. Yue (2009): Migrants as Second-Class Workers in Urban China? A Decomposition Analysis. *Journal of Comparative Economics* 37(4), 610-628. doi:10.1016/j.jce.2009.04.008.
- Deng, Q. (2007): Earnings Differential Between Urban Residents and Rural Migrants: Evidence from Oaxaca-Blinder and Quantile Regression Decompositions. *Chinese Journal of Population Science*, 2.
- Deng, X.; A. Cordilia (1999): To Get Rich is Glorious: Rising Expectations, Declining Control, and Escalating Crime in Contemporary China. *International Journal of Offender Therapy and Comparative Criminology* 43(2), 211-229. doi:10.1177/0306624x99432008.
- Deutsch, R.; A. Morrison; H. Nopo; C. Piras (2005): Working within Confines: Occupational Segregation by Sex for Three Latin American Countries. *The IUP Journal of Applied Economics* IV(3), 50-59.
- Diamond, A.; J. S. Sekhon (2012): Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies. *The Review of Economics and Statistics* 95(3), 932-945. doi:10.1162/REST_a_00318.
- Dolton, P. J.; G. H. Makepeace (1987): Marital Status, Child Rearing and Earnings Differentials in the Graduate Labour Market. *The Economic Journal* 97(388), 897-922. JSTOR. doi:10.2307/2233079.
- Dong, X.; P. Bowles (2002): Segmentation and Discrimination in China's Emerging Industrial Labor Market. *China Economic Review* 13(2), 170-196. doi:10.1016/S1043-951X(02)00057-3.
- Duncan, G. J.; S. D. Hoffman (1985): A Reconsideration of the Economic Consequences of Marital Dissolution. *Demography* 22(4), 485-497. doi:10.2307/2061584.
- Dutton, M. (1997): The Basic Character of Crime in Contemporary China. *The China Quarterly* 149, 160-177. doi:10.1017/S030574100004368X.
- Eck, A. (1993): Job-Related Education and Training: Their Impact on Earnings. *Monthly Labor Review* 116(10), 21-38.
- Ermisch, J.; M. Francesconi (2001): Family Matters: Impacts of Family Background on Educational Attainments. *Economica* 68(270), 137-156. doi:10.1111/1468-0335.00239.
- Ferrer-i-Carbonell, A. (2005): Income and Well-Being: An Empirical Analysis of the Comparison Income Effect. *Journal of Public Economics* 89(5), 997-1019. doi:10.1016/j.jpubeco.2004.06.003.
- Fields, G.; Y. Song (2013): A Theoretical Model of the Chinese Labor Market (SSRN Scholarly Paper ID 2238311): Social Science Research Network. <https://papers.ssrn.com/abstract=2238311>.

- Filippin, A.; M. Paccagnella (2012): Family Background, Self-Confidence and Economic Outcomes. *Economics of Education Review* 31(5), 824-834. doi:10.1016/j.econedurev.2012.06.002.
- Fleisher, B.; H. Li; M. Q. Zhao (2010): Human Capital, Economic Growth, and Regional Inequality in China. *Journal of Development Economics* 92(2), 215-231. doi:10.1016/j.jdeveco.2009.01.010.
- Flowerdew, R.; M. Aitkin (1982): A Method of Fitting the Gravity Model Based on the Poisson Distribution*. *Journal of Regional Science* 22(2), 191-202. doi:10.1111/j.1467-9787.1982.tb00744.x.
- Frey, B. S.; A. Stutzer (2002): What Can Economists Learn from Happiness Research? *Journal of Economic Literature* 40(2), 402-435. doi:10.1257/002205102320161320.
- Frijters, P.; L. Lee; X. Meng (2009): Jobs, Working Hours, and Remuneration Packages for Migrants and Urban Residents. <http://Rse.Anu.Edu.Au/Rumici/Pdf/Wage-MengV7-25April2009.Pdf>. <https://openresearch-repository.anu.edu.au/handle/1885/9293>.
- Gagnon, J. J.; T. Xenogiani; C. Xing (2012): Are All Migrants Really Worse Off in Urban Labour Markets? New Empirical Evidence from China (SSRN Scholarly Paper ID 1981224). Social Science Research Network. <https://papers.ssrn.com/abstract=1981224>.
- Gravemeyer, S.; T. Gries; J. Xue (2011): Income Determination and Income Discrimination in Shenzhen. *Urban Studies* 48(7), 1457-1475. doi:10.1177/0042098010375317.
- Guo, S.; M. W. Fraser (2015): *Propensity Score Analysis*. Sage.
- Han, J. (1994): *Dangdai nongcun jingji xingshi toudi yu jindai gaige silu (Contemporary Rural Economic Conditions and Reforms in Modern Times)*. 1, 9-16.
- Hansen, J.; R. Wahlberg (1999): Endogenous Schooling and the Distribution of the Gender Wage Gap (SSRN Scholarly Paper ID 217753). Social Science Research Network. <https://papers.ssrn.com/abstract=217753>.
- Heckman, J.; H. Ichimura; J. Smith; P. Todd (1998): Characterizing Selection Bias Using Experimental Data. *Econometrica* 66(5), 1017-1098.
- Heckman, J.; H. Ichimura; P. E. Todd (1997): Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies* 64(4), 605-654. doi:10.2307/2971733.
- Heckman, J.; S. Navarro-Lozano (2004): Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models. *The Review of Economics and Statistics* 86(1), 30-57. doi:10.1162/003465304323023660.
- Heineck, G. (2011): Does it Pay to Be Nice? Personality and Earnings in the United Kingdom. *ILR Review* 64(5), 1020-1038. doi:10.1177/001979391106400509.
- Heineck, G.; S. Anger (2010): The Returns to Cognitive Abilities and Personality Traits in Germany. *Labour Economics* 17(3), 535-546. doi:10.1016/j.labeco.2009.06.001.
- Hirano, K.; G. W. Imbens; G. Ridder (2003): Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica* 71(4), 1161-1189. doi:10.1111/1468-0262.00442.
- Ho, D. E.; K. Imai; G. King; E. A. Stuart (2007): Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis* 15(3), 199-236. doi:10.1093/pan/15(3).199.
- Hogarth, R. M. (2005): The Challenge of Representative Design in Psychology and Economics. *Journal of Economic Methodology* 12(2), 253-263. doi:10.1080/13501780500086172.
- Huang, Y. (2012): "Jumping Out of the Agricultural Gate" (tiaochu nongmen): Social Mobility and Gendered Intra-household Resource Distribution among Children in a Central Chinese Village, 1950-2012. *China Perspectives*; Wanchai, 4, 25-33.
- Jalan, J.; M. Ravallion (2003): Does Piped Water Reduce Diarrhea for Children in Rural India? *Journal of Econometrics* 112(1), 153-173.
- Jenkins, S. P. (1994): Earnings Discrimination Measurement: A Distributional Approach. *Journal of Econometrics* 61(1), 81-102. doi:10.1016/0304-4076(94)90078-7.
- Jian, T.; J. D. Sachs; A. M. Warner (1996): Trends in Regional Inequality in China. *China Economic Review* 7(1), 1-21. doi:10.1016/S1043-951X(96)90017-6.
- Knight, J.; Q. Deng; S. Li (2011): The Puzzle of Migrant Labour Shortage and Rural Labour Surplus in China. *China Economic Review* 22(4), 585-600. doi:10.1016/j.chieco.2011.01.006.
- Knight, J.; L. Song; J. Huaibin (1999): Chinese Rural Migrants in Urban Enterprises: Three Perspectives. *The Journal of Development Studies* 35(3), 73-104. doi:10.1080/00220389908422574.
- Korenman, S.; D. Neumark (1991): Does Marriage Really Make Men More Productive? *The Journal of Human Resources* 26(2), 282-307. JSTOR. doi:10.2307/145924.
- Kruskal, W.; F. Mosteller (1979): Representative Sampling, I: Non-Scientific Literature. *International Statistical Review/Revue Internationale de Statistique* 47(1), 13-24. doi:10.2307/1403202.
- Lardy, N. R. (1998): *China's Unfinished Economic Revolution*. Brookings Institution Press.
- Lee, L. (2012): Decomposing wage Differentials between Migrant Workers and Urban Workers in Urban China's Labor Markets. *China Economic Review* 23(2), 461-470. doi:10.1016/j.chieco.2012.03.004.

- Lee, W. S. (2008): Propensity Score Matching and Variations on the Balancing Test. In Third Conference on Policy Evaluation. ZEW, Mannheim, Germany, 27-28.
- Leuven, E.; B. Sianesi (2018): PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. <https://econpapers.repec.org/software/bocbocode/s432001.htm>.
- Liu, Z. (2005): Institution and Inequality: The Hukou System in China. *Journal of Comparative Economics* 33(1), 133-157. doi:10.1016/j.jce.2004.11.001.
- London, J. D. (2014): Welfare Regimes in China and Vietnam. *Journal of Contemporary Asia*, 44(1), 84-107. doi:10.1080/00472336.2013.822988.
- Long, J. E. (1975): Public-Private Sectoral Differences in Employment Discrimination. *Southern Economic Journal*, 42(1), 89-96. JSTOR. doi:10.2307/1056567.
- Loury, G. (1977): A Dynamic Theory of Racial Income Differences. *Women, Minorities, and Employment Discrimination*, 153, 86-153.
- Lu, Y. L. (2002): Structure and Change: The Household Registration System in China after 1949. *Peking University Journal* 39(2), 123-130.
- Lu, Z.; S. Song (2006): Rural–Urban Migration and Wage Determination: The Case of Tianjin, China. *China Economic Review* 17(3), 337-345. doi:10.1016/j.chieco.2006.04.007.
- Luttmer, E. F. P. (2005): Neighbors as Negatives: Relative Earnings and Well-Being. *The Quarterly Journal of Economics* 120(3), 963-1002. doi:10.1093/qje/120.3.963.
- Mackenzie, P. W. (2002): Strangers in the City: The Hukou and Urban Citizenship in China. *Journal of International Affairs* 56(1), 305-319.
- Main, B. G. M.; B. Reilly (1993): The Employer Size-Wage Gap: Evidence for Britain. *Economica* 60(238), 125-142. JSTOR. doi:10.2307/2554585.
- McBride, M. (2001): Relative-Income Effects on Subjective Well-Being in the Cross-Section. *Journal of Economic Behavior & Organization* 45(3), 251-278. doi:10.1016/S0167-2681(01)00145-7.
- Meghir, C.; M. Palme (2005): Educational Reform, Ability, and Family Background. *The American Economic Review*; Nashville 95(1), 414-424.
- Mendola, M. (2007): Agricultural Technology Adoption and Poverty Reduction: A Propensity-Score Matching Analysis for Rural Bangladesh. *Food Policy* 32(3), 372-393. doi:10.1016/j.foodpol.2006.07.003.
- Meng, X. (1998): Male–Female Wage Determination and Gender Wage Discrimination in China's Rural Industrial Sector. *Labour Economics* 5(1), 67-89. doi:10.1016/S0927-5371(97)00028-6.
- Mishra, V.; R. Smyth (2013): Economic Returns to Schooling for China's Korean Minority. *Journal of Asian Economics* 24, 89-102. doi:10.1016/j.asieco.2012.09.001.
- Mishra, V.; R. Smyth (2015): Estimating Returns to Schooling in Urban China Using Conventional and Heteroskedasticity-Based Instruments. *Economic Modelling* 47, 166-173. doi:10.1016/j.econmod.2015.02.002.
- Moore, R. L. (1983): Employer Discrimination: Evidence From Self-Employed Workers. *The Review of Economics and Statistics* 65(3), 496-501. JSTOR. doi:10.2307/1924197.
- Mundlak, Y. (1979): Intersectoral Factor Mobility and Agricultural Growth. International Food Policy Research Institute.
- Munro, A. (1988): The Measurement of Racial and Other Forms of Discrimination. University of Stirling Department of Economics.
- Neal, D. A.; W. R. Johnson (1996): The Role of Pre-market Factors in Black-White Wage Differences. *Journal of Political Economy* 104(5), 869-895.
- Nielsen, I.; C. Nyland; R. Smyth; M. Zhang; C. J. Zhu (2005): Which Rural Migrants Receive Social Insurance in Chinese Cities? Evidence from Jiangsu Survey Data. *Global Social Policy* 5(3), 353-381. doi:10.1177/1468018105057416.
- Ñopo, H. (2008): Matching as a Tool to Decompose Wage Gaps. *The Review of Economics and Statistics* 90(2), 290-299. doi:10.1162/rest.90.2.290.
- Oaxaca, R. (1973): Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review* 14(3), 693-709. JSTOR. doi:10.2307/2525981.
- Pietro, G. D. (2007): The Effect of Computer use on Earnings in Italy. *Empirical Economics* 33(2), 245-262. doi:10.1007/s00181-006-0101-6.
- Poncet, S. (2005): A Fragmented China: Measure and Determinants of Chinese Domestic Market Disintegration. *Review of International Economics* 13(3), 409-430. doi:10.1111/j.1467-9396.2005.00514.x.
- Reiss, A. J. (1959): Rural-Urban and Status Differences in Interpersonal Contacts. *American Journal of Sociology* 65(2), 182-195. doi:10.1086/222660.
- Rosenbaum, P. R. (1986): Dropping out of High School in the United States: An Observational Study. *Journal of Educational Statistics* 11(3), 207-224. doi:10.3102/10769986011003207.

- Rosenbaum, P. R. (2002): Overt Bias in Observational Studies, in: P. R. Rosenbaum (ed.), *Observational Studies*. Springer, 71-104 doi:10.1007/978-1-4757-3692-2_3.
- Rosenbaum, P. R.; D. B. Rubin (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70(1), 41-55. doi:10.1093/biomet/70.1.41.
- Rosenbaum, P. R.; D. B. Rubin (1984): Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association* 79(387), 516-524.
- Rubin, D. B.; N. Thomas (1992): Characterizing the Effect of Matching Using Linear Propensity Score Methods with Normal Distributions. *Biometrika* 79(4), 797-809. doi:10.1093/biomet/79.4.797.
- Rubin, D. B.; N. Thomas (1996): Matching Using Estimated Propensity Scores: Relating Theory to Practice. *Biometrics* 52(1), 249-264. JSTOR. doi:10.2307/2533160.
- Rubin, D. B.; N. Thomas (2000): Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates. *Journal of the American Statistical Association* 95(450), 573-585. doi:10.1080/01621459.2000.10474233.
- Schultz, T. P. (2002): Wage Gains Associated with Height as a Form of Health Human Capital. *The American Economic Review* 92(2), 349-353. JSTOR.
- Schwartz, A. (1973): Interpreting the Effect of Distance on Migration. *Journal of Political Economy* 81(5), 1153-1169.
- Sekhon, J. S. (2011): Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R. *Journal of Statistical Software* 042(i07). <https://ideas.repec.org/a/jss/jstsof/v042i07.html>.
- Semykina, A.; S. J. Linz, (2007): Gender Differences in Personality and Earnings: Evidence from Russia. *Journal of Economic Psychology* 28(3), 387-410. doi:10.1016/j.joep.2006.05.004.
- Sheehan, S. (2017): China's Hukou Reforms and the Urbanization Challenge. *The Diplomat*, 22.
- Sianesi, B. (2004): An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s. *The Review of Economics and Statistics* 86(1), 133-155. doi:10.1162/003465304323023723.
- Sjaastad, L. A. (1962): The Costs and Returns of Human Migration. *Journal of Political Economy* 70(5, Part 2), 80-93.
- Smith, H. L. (1997): Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies. *Sociological Methodology* 27(1), 325-353.
- Smith, J. N. (2002): "Making Culture Matter": Symbolic, Spatial and Social Boundaries between Uyghurs and Han Chinese. *Asian Ethnicity* 3(2), 153-174. doi:10.1080/14631360220132718.
- Smith, A. J.; E. P. Todd (2005): Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators? *Journal of Econometrics* 125(1), 305-353. doi:10.1016/j.jeconom.2004.04.011.
- Song, Y. (2014): What should Economists know about the Current Chinese Hukou System? *China Economic Review* 29, 200-212. doi:10.1016/j.chieco.2014.04.012.
- Song, Y. (2016): Hukou-Based Labour Market Discrimination and Ownership Structure in Urban China. *Urban Studies* 53(8), 1657-1673. doi:10.1177/0042098015576861.
- Tani, M. (2017): Hukou Changes and Subjective Well-Being in China. *Social Indicators Research* 132(1), 47-61. doi:10.1007/s11205-016-1247-z.
- Taubman, P. (1976): The Determinants of Earnings: Genetics, Family, and Other Environments: A Study of White Male Twins. *The American Economic Review* 66(5), 858-870.
- Walder, A. G. (1995): Career Mobility and the Communist Political Order. *American Sociological Review* 60(3), 309-328. JSTOR. doi:10.2307/2096416.
- Wallace, J. L. (2016): Juking the Stats? Authoritarian Information Problems in China. *British Journal of Political Science* 46(1), 11-29. doi:10.1017/S0007123414000106.
- Wang, F.-L. (2004): Reformed Migration Control and New Targeted People: China's Hukou System in the 2000s. *The China Quarterly* 177, 115-132. doi:10.1017/S0305741004000074.
- Whalley, J.; S. Zhang (2011): A Numerical Simulation Analysis of (hukou) Labour Mobility Restrictions in China. In *China's Integration into the World Economy* (Vol. 1-0, pp. 295-324). World Scientific. doi:10.1142/9789814304795_0011.
- Wu, H. X. (1994): Rural to Urban Migration in the People's Republic of China*. *The China Quarterly* 139, 669-698. doi:10.1017/S0305741000043095.
- Wu, L. (2013): Decentralization and Hukou Reforms in China. *Policy and Society* 32(1), 33-42. doi:10.1016/j.polsoc.2013.01.002.
- Wu, X.; D. J. Treiman (2007): Inequality and Equality Under Chinese Socialism: The Hukou System and Inter-generational Occupational Mobility. *American Journal of Sociology* 113(2), 415-445. doi:10.1086/518905.
- Xie, Y.; J. Hu (2014): An Introduction to the China Family Panel Studies (CFPS). *Chinese Sociological Review* 47(1), 3-29. doi:10.2753/CSA2162-0555470101.2014.11082908.

- Xu, Q.; X. Guan; F. Yao (2011): Welfare Program Participation Among Rural-to-Urban Migrant Workers in China. *International Journal of Social Welfare* 20(1), 10-21. doi:10.1111/j.1468-2397.2009.00713.x.
- Yang, D. T. (1999): Urban-Based Policies and Rising Income Inequality in China. *The American Economic Review*; Nashville 89(2), 306-310.
- Yang, D. T. (2002): What has Caused Regional Inequality in China? *China Economic Review* 13(4), 331-334.
- Zhang, J. (2017): The Evolution of China's One-Child Policy and Its Effects on Family Outcomes. *Journal of Economic Perspectives* 31(1), 141-160. doi:10.1257/jep.31.1.141.
- Zhang, X.; R. Kanbur (2005): Spatial Inequality in Education and Health Care in China. *China Economic Review* 16(2), 189-204. doi:10.1016/j.chieco.2005.02.002.
- Zhang, K. H.; S. Song (2001): Promoting Exports: The Role of inward FDI in China. *China Economic Review* 4(11), 385-396.
- Zhang, X.; J. Yang; S. Wang (2011): China has Reached the Lewis Turning Point. *China Economic Review* 22(4), 542-554. doi:10.1016/j.chieco.2011.07.002.
- Zhang, L.; H. Yi; R. Luo; C. Liu; S. Rozelle (2013): The Human Capital Roots of the Middle Income Trap: The Case of China. *Agricultural Economics* 44(s1), 151-162. doi:10.1111/agec.12059.
- Zhao, Y. (1999): Labor Migration and Earnings Differences: The Case of Rural China. *Economic Development and Cultural Change* 47(4), 767-782.

Appendix A

Table A1 Estimation of propensity score (logit model)

	Specification 1	Specification 2	Specification 3
	<i>Hukou</i> (agricultural = 1)	<i>Hukou</i> (agricultural = 1)	<i>Hukou</i> (agricultural = 1)
Migration status			
Migrant (yes = 1)	-1.401*** (0.043)	-1.364*** (0.043)	-1.391*** (0.046)
Demographic characteristics			
Gender (men = 1)	0.598*** (0.040)	0.493*** (0.041)	0.510*** (0.043)
Age (years)	-0.043*** (0.002)	-0.040*** (0.002)	-0.041*** (0.002)
BMI	1.062*** (0.091)	1.092*** (0.091)	1.046*** (0.097)
Self-rated health status	-0.029* (0.016)	-0.014 (0.016)	-0.018 (0.017)
Abilities			
Years of schooling	-0.218*** (0.007)	-0.223*** (0.007)	-0.228*** (0.007)
Nondegree training	-0.093*** (0.028)	-0.110*** (0.031)	-0.110*** (0.033)
Language ability (test score)	-0.011*** (0.003)	-0.012*** (0.003)	-0.012*** (0.003)
Mathematical ability (test score)	-0.022*** (0.004)	-0.022*** (0.004)	-0.020*** (0.005)
Memorizing ability	0.153*** (0.015)	0.152*** (0.015)	0.153*** (0.016)
Ambitions			
Confidence about future	0.086*** (0.019)	0.080*** (0.019)	0.071*** (0.020)
Networking/relation	-0.030*** (0.011)	-0.034*** (0.011)	-0.033*** (0.012)
Attitudes			
The ideal number of children	0.343*** (0.032)	0.344*** (0.032)	0.341*** (0.035)
People are helpful/selfish	-0.014 (0.040)	0.002 (0.040)	-0.014 (0.043)
People are trustworthy	-0.016 (0.039)	-0.028 (0.039)	-0.032 (0.041)
Trust your neighbor	0.077*** (0.009)	0.078*** (0.009)	0.074*** (0.009)
Family background			
Marriage (married = 1)	-0.091* (0.049)	-0.000 (0.049)	-0.110** (0.055)
Time spend for household works	0.119*** (0.011)	0.126*** (0.011)	0.132*** (0.012)
Time spend for watching TV	-0.011*** (0.002)	-0.009*** (0.002)	-0.010*** (0.002)
Family's social status	0.109*** (0.025)	0.113*** (0.025)	0.129*** (0.026)
Other socioeconomic background			
CCP member (yes = 1)	-0.307*** (0.064)	-0.315*** (0.065)	-0.346*** (0.066)
Unfair treatment by the government	0.116* (0.063)	0.090 (0.064)	0.077 (0.067)
Unfair treatment due to inequality of personal wealth	0.241*** (0.061)	0.230*** (0.061)	0.237*** (0.065)
Social status in the local area	0.110*** (0.023)	0.095*** (0.023)	0.071*** (0.026)
Retirement (yes = 1)	-3.386*** (0.100)	-3.144*** (0.103)	-3.197*** (0.107)
Currently employed (yes = 1)		0.678*** (0.045)	0.736*** (0.051)
Hospitalization (yes = 1)			0.138** (0.064)

(Continued)

Table A1 Continue

	Specification 1	Specification 2	Specification 3
	<i>Hukou</i> (agricultural = 1)	<i>Hukou</i> (agricultural = 1)	<i>Hukou</i> (agricultural = 1)
Number of deceased siblings			0.218*** (0.042)
Relative income			0.033 (0.023)
Currently attending school			0.403*** (0.136)
Constant	1.667*** (0.212)	1.030*** (0.217)	1.082*** (0.231)
Pseudo R^2	0.352	0.359	0.374
Number of observations	27,212	27,212	24,890

Notes: Standard errors are robust standard errors (in parenthesis). Dependent variable is *hukou* status: agricultural *hukou* = 1, otherwise = 0. Coefficient should be interpreted as odds ratio. * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. BMI, body mass index; CCP, Chinese Communist Party.