Residential Segregation and Employment Outcomes of Rural Migrant Workers in China

Pengyu Zhu
Zhejiang University
Email address: 0011134@zju.edu.cn

Residential Segregation and Employment Outcomes of Rural Migrant Workers in China

Abstract

In China, many rural migrant workers live in urbanizing villages that are usually located at the peripheral areas of cities. Although this residential segregation is related to some policies (e.g. *hukou* system), it is largely by choice. Living in these urbanizing villages could incur both negative spatial mismatch effects and positive spillover effects. Through a survey across four mega-regions in China that are currently experiencing the most rapid urbanization, we collect some unique information on rural migrant workers' attitudes towards living in urbanizing villages, and therefore are able to address the self-selection bias that has broadly existed in most previous studies on residential segregation and spatial mismatch. The models show that the net effect of residential segregation in urbanizing villages on migrant workers' employment outcomes (both employment propensity and wage) appears to be positive, suggesting the spillover effects override the spatial mismatch effects. Current policy proposals by government officials to demolish urbanizing villages should be accompanied by alternative policies to assist with housing migrant workers in appropriate locations that not only reduce spatial mismatch effects but also maintain positive spillover effects.

Keywords: Residential Segregation; Spatial Mismatch; Employment Outcomes; Rural Migrant Workers; Self-selection Bias

1. Introduction

In the past 50 years, urban development in the U.S. has resulted in the dual emergence of flexible firms (employment decentralization) and flexible workers (residential suburbanization). As a result, many scholars argue that the employment decentralization, coupled with the segregation in housing markets, impairs job accessibility of inner city minority groups. This is known as the "Spatial Mismatch Hypothesis" (SMH), first proposed by Kain (1968). Since then, SMH has been widely tested in many empirical studies. Ihlanfeldt and Sjoquist (1998) and Houston (2005) provided comprehensive reviews of these studies. Generally, SMH studies have focused on the labor market outcomes of the various racial/ethnic minorities who face residential segregation in inner city neighborhoods, have limited mobility, and thus are spatially isolated from decentralizing blue-collar jobs (e.g., Ihlanfeldt, 1997; Kasarda and Ting, 1996; Raphael, 1998).

In China, many rural-to-urban migrant workers living in the suburbs of major cities are also experiencing residential segregation, although the causes of their segregation are somewhat different from racial/ethnic minorities in the U.S. Research shows that the residential segregation of racial/ethnic minorities in the U.S is primarily due to housing market discrimination and land use regulations (e.g. exclusionary zoning), which make it difficult for them to move to suburbs where many jobs have relocated (e.g., Ihlanfeldt, 1997; Ihlanfeldt and Sjoquist, 1998). Another reason often cited is that racial/ethnic minorities and immigrants often choose to reside near co-ethnics to share resources and their common culture (Logan, Alba, and Zhang, 2002). In China, residential segregation of rural migrant workers is largely due to economic reasons. On

the one hand, rural migrant workers cannot afford to purchase or rent expensive commercial apartments. On the other hand, the household registration regulation (*hukou* system) disqualifies them from purchasing government-sponsored low-income housing in cities. Additionally, the law currently prohibits owners from renting out government-sponsored low-income housing; otherwise, the owners may face the risk of forfeiting their property to the government. As a result, many rural migrant workers choose to live in urbanizing villages where housing costs are low. Another important reason is that their residential segregation is by choice—that is, they want to live in places where people from their hometown/regions are concentrated in order to benefit from the social networks, for networking purposes involving jobs, marriage, etc. Indeed, the residential segregation of rural migrant workers in many Chinese cities is inseparable from the special urbanization path and the economic development model of China.

The rapid urbanization in many regions of China has changed the urban-rural landscapes and induced a massive influx of rural migrants from less developed parts of the country. One major change in the landscapes of many large Chinese cities has been the encroachment on rural settlements by the ever-expanding urban boundaries. Due to the high capital and time cost of moving native rural residents, local authorities are more willing to leave these villages in their *status quo* when facing rapid urbanization, and develop the farmland surrounding these villages for industrial or residential uses. With the continuous spatial expansion of cities, many previous villages now become organic parts of the cities, namely, "villages in cities" or "urbanizing villages." In the meantime, rapid urbanization has also induced a massive rural-urban migration. Song et al. (2008) mentioned that 70 million rural migrants were working and living in urban areas at the

end of 2000. These rural migrant workers typically work in low-skilled occupations, often serving as labor to help "build the city." Few of them can afford the expensive commercial apartments in cities. Instead, urbanizing villages that resulted from the fast urban expansion are able to provide cheap housing to accommodate them.

From an urban planning perspective, urbanizing villages are considered as special areas of a city and usually are not regulated by zoning or other planning regulations (especially at their early stage of incorporation into the city). Therefore, native residents in urbanizing villages are able to construct inexpensive housing units and rent/sell them to rural migrants who are looking for shelter while excluded, in an economic sense, from the expensive urban housing system. Most of these housing units are illegal because their builders have not obtained building permits issued by relevant city government agencies. In some cases, the entire village may collectively build (illegal) housing units and then distribute proceeds among its native residents from the sale of these units. These urbanizing villages are generally perceived as undesirable places by local authorities and are usually associated with unplanned land uses, insufficient infrastructures and deteriorating housing conditions, but undoubtedly, they play an important role in housing rural migrants. For example, within the city boundary of Guangzhou, a city with a total population of over 8 million, there were about one million inhabitants residing in the 277 urbanizing villages in 2000 (Zhang et al., 2003).

Given the critical role urbanizing villages play in housing rural migrant workers, it is important to understand how residential segregation in urbanizing villages could affect the employment outcomes of these workers. So far, none relevant individual level data have been provided by either local governments or central government. Using data

from a survey conducted in twelve cities across the four most rapidly urbanizing regions in China, this study is the first to examine how residential segregation affects the employment outcomes of rural migrant workers. Thus, it will have important policy implications for sustainable urban development in China. More specifically, two related questions are addressed:

1) How does living in urbanizing villages influence rural migrant workers' employment probabilities?

2) Given that they are employed, how does living in urbanizing villages influence rural migrant workers' wages?

Another major contribution of this study to the literature is that it addresses the endogeneity problem related to the joint location choice/job choice decision, using some unique survey data that contain information on migrant workers' attitudes towards living in urbanizing villages.

In the following sections, we first briefly review the literature as it relates to the theoretical framework that guides our model specification and interpretation. We also discuss the self-selection bias that is widely found in most empirical studies on residential segregation and spatial mismatch. We then turn to empirical analyses and conclusions.

2. Costs and Benefits of Residential Segregation

To date, the literature on residential segregation in the U.S. can be summarized as dwelling on two main reasons to explain why and how it might affect the labor market outcomes of racial or ethnic minorities: 1) Spatial Mismatch Effects and 2) Spillover Effects. These two effects are conceptually opposite and often intertwined. They provide

a theoretical basis for analyzing the employment impact of residential segregation in China.

1) Spatial Mismatch Effects—the Costs of Residential Segregation

The "Spatial Mismatch Hypothesis" (Kain, 1968) states that job decentralization adversely impacts various racial/ethnic minorities living in inner city neighborhoods because of their residential segregation (for recent reviews, see Ihlanfeldt and Sjoquist 1998; Houston 2005). Traditional studies of spatial mismatch have relied on the fact that residential segregation of African-Americans was due to racial discrimination in housing markets in suburban areas. More recently, some studies have pointed out that it is residential immobility in the housing market, not segregation *per se*, that presents the mismatch problem for minorities in the face of job decentralization (Kasarda 1988; Cooke and Shumway 1991; Houston 2005). Others have emphasized two other reasons that further constrain the employment outcomes of minorities: 1) the inaccessibility of suburban job sites by public transportation (Sanchez 1999); 2) the limited car ownership of minorities and low-skilled workers—the "automobile mismatch", as suggested or implied by Taylor and Ong (1995); Houston (2001); Cervero, Sandoval and Landis (2002); Raphael and Rice (2002).

The spatial mismatch effects of urbanizing villages in major Chinese cities are different from the spatial mismatch observed in American cities. It includes two aspects. First, urbanizing villages are not necessarily close to blue-collar jobs for which their residents (i.e. rural migrant workers) are looking. With the process of urban expansion, existing urbanizing villages are often located in the peripheral areas of cities.[1] With most

[1] Due to land value appreciation associated with urban expansion, older urbanizing villages in inner city locations are gradually purchased and developed by real estate developers and government.

jobs still located in inner cities, the residential segregation of rural migrant workers who live in these urbanizing villages could hinder their access to jobs, thus incurring "locational mismatch".

Second, urbanizing villages are often located within various industrial parks, such as: clothing manufacturing parks, automobile manufacturing parks, electronic device manufacturing parks, biochemical manufacturing parks, etc. Some older industrial parks provide many blue-collar jobs (e.g., clothing and toy manufacturing parks). But with the transition of the Chinese economy, more and more industrial parks are high-technology parks (e.g., electronic device manufacturing parks, biochemical manufacturing parks) and are usually located in suburban areas. They require the workers to have at least an associate degree, if not a bachelor's degree, in relevant fields. It is very likely that urbanizing villages located in these suburban high-tech industrial parks may further disadvantage the employment outcomes of their residents, who are mostly low-skilled rural migrant workers. Therefore, urbanizing villages implicitly become a barrier for them to find appropriate jobs, adversely affecting their employment propensity and wages. This is called "skill mismatch", as described in Pastor and Marcelli (2000) and Parks (2004a, 2004b).

Despite a different process through which segregation may have occurred, rural migrant workers in China face similar restrictions or immobility in the housing market to racial/ethnic minority groups in the U.S. The spatial mismatch effects suggest that rural migrant workers who are segregated in urbanizing villages will fare worse than they otherwise would. In this view, the spatial mismatch effects would adversely affect the employment outcomes of rural migrant workers.

2) Spillover Effects—the Benefits of Residential Segregation

Besides of low housing costs, urbanizing villages also offer socioeconomic benefits. The spillover effects attract migrant workers to voluntarily choose to live in urbanizing villages. Urbanizing villages, as a residential community that incurs segregation, are to some extent similar to ghettos or ethnic enclaves in Europe or the U.S. The literature on the benefits of ghettos or ethnic enclaves can be traced back to the 1950s. For example, Wirth (1956) pointed out that Jewish ghettos protected its people from the hostile outside world and enforced good behavior, since community leaders can ostracize those who misbehaved. Glazer and Moynihan (1963) also found that black communities protected black-owned businesses from white competition. Handlin (1959), Glazer and Moynihan (1963), and Wilson (1987) all emphasized that the outflow of middle-class blacks from inner city ghettos might adversely affect poor blacks that were left behind. More recently, Cutler and Glaeser (1997) and Glaeser (2011) echoed the arguments that ghettos might have benefited poor blacks by keeping rich and poor blacks together.

Spillover effects from residential segregation include network effects (Liu 2009) and human capital externalities (see Cutler and Glaeser 1997; Borjas 1995; Edin et.al. 2003; Elliott and Sims 2001). In this paper, network effects refer to the fact that urbanizing villages act as opportunity-increasing networks by disseminating employment information to migrant workers. In addition, people from the same hometown or region are often concentrated in one urbanizing village to form their own "community", which in turn builds social networks with other benefits such as cultural identity, sense of belonging, and even marriage. Human capital externalities refer to the stock of human

capital from skilled and well-connected migrant workers living in urbanizing villages. These skilled members could be migrant workers who had arrived earlier. The newer migrants interact with these skilled members and thus gain benefits from living in the urbanizing villages. In this view, the spillover effects would favorably influence the employment outcomes of rural migrant workers living in urbanizing villages.

The above discussion about the theoretical framework on the costs and benefits of residential segregation suggests that the net impact of urbanizing villages on employment outcomes of rural migrant workers is uncertain. It depends on the magnitude of both the spatial mismatch effects and the spillover effects. In the following sections, we develop empirical models to identify the net impact.

3. Self-selection Bias in Residential Location Choice and Empirical Solutions

Most literature on spatial mismatch in the U.S. has focused on how the residential segregation of African-Americans in central cities has disadvantaged them as jobs have moved to suburban areas. Recently, a few studies have examined how residential segregation impacts the job market outcomes of Hispanics or immigrant populations in the U.S. (e.g. Ihlanfeldt, 1993; Aponte, 1996; Preston et. al, 1998; Pastor and Marcelli, 2000; Parks, 2004a and 2004b; Liu, 2009). However, the largest challenge for most empirical studies on spatial mismatch remains, the self-selection bias in residential location choice. Most research assumes that residential location choice is exogenous from employment outcomes—namely, residential segregation influences employment, but not vice versa. However, residential location choice is indeed endogenous and causality can operate in both directions. On the one hand, residential segregation could affect

employment outcomes; on the other hand, the labor market success (i.e. change from unemployment to employment or wage increase) might also influence the location decision of whether to live in a certain location (e.g. an urbanizing village in this case). The difficulties of proposing an empirical strategy, that takes into account the possible endogeneity problem associated with residential location choice, is probably the reason why few earlier studies have examined the causal effects of residential segregation on the economic outcomes of racial and ethnic minorities in the U.S. (Damm, 2006).

Recent research has made some progress in addressing this issue. One approach is to aggregate racial segregation to the city or MSA level, and use the inter-city variation, rather than intra-city variation, to avoid the problems of intra-city sorting of the population. These studies argued that intra-city comparison of the effects of segregation on outcomes are biased, and that labor market outcomes are less likely to affect inter-city migration compared with intra-city migration (see Cutler and Glaeser, 1997; Dustmann and Preston, 1998; Gabriel and Rosenthal, 1999; Bertrand, Luttmer and Mullainathan, 2000). For example, Cutler and Glaeser (1997) examined the effects of residential segregation at the metropolitan level on outcomes for blacks in schooling, employment, and single parenthood. They found blacks living in more segregated MSAs fare worse than those living in less segregated MSAs. However, this strategy still faces an endogeneity problem, not just because abler minorities are more likely to leave cities that are more segregated (as suggested by the authors themselves), but also because there might still exist unobserved individual attributes across MSAs that affect both the MSA-level segregation measure and individual labor market outcomes (Edin, Fredriksson and Aslund, 2003).

This leads to the second approach—an instrumental variable method. For example, to further address the endogeneity of location choice (i.e. increased MSA-level segregation leads to poor outcomes and vice versa), Cutler and Glaeser (1997) instrumented their segregation indices across MSAs by using local government finance, topographical features, and residence before adulthood.

The third approach is to focus on at-home youth (usually 16-21 years old) employment outcomes, based on the assumption that parents' decisions upon residential location did not necessarily take into account the future labor market outcomes of the at-home youths (see Ihlanfeldt and Sioquist, 1990; Ihlanfeldt, 1993; Stoll, 1999; Raphael, 1998; O'Regan and Quigley, 1998; Painter, Liu, Zhang, 2008). With the same rationale, some other research has used parental choices of neighborhoods to study the outcomes of the offspring (see Borjas, 1995).

The fourth approach is based on quasi-experimental data (see Edin, et.al. 2003; Damm 2006). For example, Edin et.al (2003) was based on the 1985-1991 Swedish spatial dispersal policy under which almost all refugee immigrants were randomly assigned initial locations of residence at the time of being granted asylum. Using initial ethnic enclave size as an instrument for future ethnic enclave size, they found that living in an ethnic enclave improves earnings for low-skilled refugees, but not for high-skilled refugees. Similarly, Damm (2006) addressed the problem of self-selection into ethnic enclaves by using a Danish spatial dispersal policy that randomly dispersed new refugees across locations conditional on six individual characteristics. They also found ethnic group size increases the employment probability and earnings of refugees.

The validity of exogeneity claims in some of the above research is still debated. Moreover, research that limited their sample to at-home youth was only able to study a very specific age group (16-21) of the population—the "marginal workers". Methodologically, this study differentiates itself from previous research by using a unique dataset that contains attitude information collected from a survey. With the attitude information, this study utilizes an instrumental variables (IV) approach to address the endogeneity problem associated with residential location choice.

4. Model Specification

4.1 Baseline Model

To determine the influence of residential segregation and socioeconomic characteristics on migrant workers' employment outcomes, our empirical analyses use the following econometrical model:

$$Y_i = \beta_0 + \beta_1 U_i + \beta_2 X_i + \beta_3 M_i + \beta_4 R_i + \varepsilon_i$$

U_i is a dummy variable representing the residential location of rural migrant worker *i*, classified by whether in urbanizing villages or not. X_i is a set of individual socioeconomic characteristics that have been found to be relevant for job prospects based on the literature. These variables include: age, gender, marital status, educational attainment, occupation, and so on. M_i is a group of factors related to this migrant worker's migration and employment history. All these variables are listed in Table 1. The economic region that migrant worker *i* currently lives on is included in the model as control variable R_i.

Y_i is the employment outcomes of rural migrant worker i, measured by employment propensity and wage. When estimating the effects of residential segregation (in urbanizing villages) on employment propensity, a probit specification is used. When estimating the effects on wages (log), an OLS specification is used.

4.2 Addressing Self-selection Bias

To address the self-selection bias associated with residential location choice (whether live in urbanizing villages), we use an instrumental variables (IV) approach. The instrumental variables include: sense of belonging to the current neighborhood, preference for associating with people from the hometown, general attitude towards the sanitation condition in urbanizing villages, and general attitude towards the safety condition in urbanizing villages. These are all subjective attitudinal factors that affect the likelihood to live in urbanizing villages, but do not directly affect the employment propensity and wage of the interviewee. The only possible impact of these instrumental variables on employment outcomes is indirect -- only through their effects on the decision to live in urbanizing villages. Thus, theoretically, they are viable instruments for the variable of "whether or not to live in an urbanizing village" (U_i).

The IV approach utilizes a two-stage regression. The two stages are listed as follows:

$$U_i = \gamma_0 + \gamma_1 X_i + \gamma_2 M_i + \gamma_3 R_i + \gamma_4 IV_i + \omega_i$$

$$Y_i = \beta_0 + \beta_1 \widehat{U}_i + \beta_2 X_i + \beta_3 M_i + \beta_4 R_i + \varepsilon_i$$

IV_i represents the four instrumental variables applied to address the endogeneity problem associated with residential location choice. All other variables are the same as introduced in the baseline model. The standard errors of the regression coefficients in the second

stage are adjusted for both the employment propensity model and the wage model to ensure accurate inferences on the statistical significance of coefficients.

Because of the self-selection bias, our baseline model suffers possible two-directional causality: 1) living in urbanizing village affects employment outcome and 2) employment outcome affects whether to live in an urbanizing village. Intuitively, migrant workers with better employment outcomes are less likely to live in urbanizing villages. When they can afford better housing in those commercial condominium complexes, they usually choose to move out from urbanizing villages. This negative reverse causality (negative impact of employment outcome on the likelihood to live in an urbanizing village) could be counteracting the impact of the urbanizing village on employment outcomes. If the IV approach is able to successfully attenuate the negative reverse causality, it is expected that it provides larger coefficient estimates for the endogenous variable (i.e. urbanizing village) than the baseline model.

5. Data and Sample

5.1 Survey

The empirical analysis uses first–hand data from a large-sample rural migrant survey conducted in 2012-2013 in twelve cities across the four most rapidly urbanizing regions of China--Pearl River Delta Region, Yangtze River Delta Region, Bohai Bay Area, and Chengdu-Chongqing Region. From each of the four urbanizing regions, three cities were randomly selected based on the population size: one megacity with a population over two million, one large city with a population between 500,000 and two million, and one small-medium-sized city with a population less than 500,000. The

sample of the twelve cities included: Guangzhou, Zhongshan and Dongguan in the Pearl River Delta Region; Ningbo, Jiangyin and Yueqing in the Yangtze River Delta Region; Jinan, Weifang and Yanjiao in the Bohai Bay Area; Chongqing, Chengdu and Nancong in the Chengdu-Chongqing Region. A sample of 200 rural migrants in each city (2400 migrants in total) was then randomly selected from the migrant registration list provided either by the local Public Security Bureau or by the local government migrant administrative agency. During the interview, if the sampled migrant had already moved away, the systematic random sampling was continued until the desired sample size (200 migrants in each city) was reached. Migrant unavailability in sampling ranged from 15 percent to 30 percent in different cities, and the turnover rate (i.e. percentage of replacement) usually was higher in larger cities. In this survey, rural migrants were defined as people living in cities but their *Hukou* (i.e. the household and individual registration system in China) were registered as rural at the time of survey. That is, they were not affiliated with the city and had moved from their *Hukou* registration place (often their hometown) to the city. The survey collected information about rural migrants' various aspects of livelihood, including: demography, employment, income, housing, social network, attitudes towards urbanizing villages, and so on.

5.2 Data

The survey resulted in 2226 effective observations of rural migrants who had migrated from 31 provinces and municipalities across China. All interviewees were within the legal working age (16-65). Among them, there were 51 (2.3 percent) rural migrants who were not employed at the time of the interview. Among those employed, the median monthly wage was 1756 Yuan. There were slightly more male (56.4 percent)

than female interviewees. Most interviewees aged between 21 and 40 (68.4 percent), were married (65.1 percent), and had middle school education and above (45.2 percent). On average, the interviewees had stayed in the current city for six years. The summary statistics for all variables are reported in Table 1.

[Table 1 about here]

Among all migrant workers in the sample, 56.6 percent lived in urbanizing villages at the time of the interview. For those living in urbanizing villages, 97.4 percent of them were employed, with a median monthly wage of 1745 Yuan. For those not living in urbanizing villages, 98.3 percent were employed, with a median monthly wage of 1775 Yuan. At first glance, the differences in the employment outcomes between these two groups are small, with those not living in urbanizing villages slightly better off. This lack of inter-group variation is probably why the Probit and OLS model results on the urbanizing village variable (reported in the next section) are not statistically significant. However, as discussed earlier, the negative reverse causality (negative impact of employment outcome on the likelihood to live in an urbanizing village) could be counteracting the impact of urbanizing village on employment outcome. This could result in the small differences in employment outcomes between the two groups to be observed. With the IV approach proposed to attenuate the negative reverse causality, a more accurate estimate on impact of urbanizing village on employment outcome could be expected.

6. RESULTS

6.1 Employment propensity

Our major interest here is whether living in an urbanizing village increased or decreased the likelihood of being employed and the wage. Model 1 in Table 2 presents the Probit Model results for employment propensity.

[Table 2 about here]

The Probit Model shows that the urbanizing village is not statistically significantly associated with the propensity to be employed. Among other factors, gender, education, the length (number of years) of interviewees' first non-agricultural job, and living with family members have statistically significant effects. It appears male migrant workers are more likely to find jobs than females. People with higher educational attainment enjoy relatively higher employment propensity. In addition, people have disadvantages in job-hunting if he/she experiences a longer period in his/her first non-agricultural job. The model also shows people living with family members in the current city tend to be less likely to be employed.

As discussed earlier, migrant workers are possibly self-selecting to live in urbanizing villages due to various reasons, including the likelihood of finding work. To examine the severity of this self-selection bias or endogeneity problem, three statistical tests are conducted: Wu-Hausman F test, Durbin-Wu-Hausman chi-sq test, and Difference-in-Sargan test. All tests suggest rejecting the null hypothesis that the specified endogenous regressor (i.e. urbanizing village) can be treated as exogenous. Therefore, in the next step, we try to address this endogeneity problem through a group of four instrumental variables: sense of belonging to the current neighborhood, preference for associating with people from the hometown, general attitude towards the sanitation condition in urbanizing villages, and general attitude towards the safety condition in

urbanizing villages. Model 2 in Table 2 presents the employment propensity results using Instrumental Variable Probit (IV-probit) method to address the endogeneity problem associated with the urbanizing village variable. The standard errors of the coefficient estimates in Model 2 have been adjusted to ensure accurate inferences on the significance of coefficients.

Model 2 shows that urbanizing village has a statistically significant positive impact on the employment propensity. In terms of the magnitude of this impact, living in an urbanizing village could cause the migrant worker to be 68 percent more likely to find a job, holding other factors constant. This is a very large benefit that urbanizing village confers on the economic outcomes of migrant workers. Among other factors, gender and living with family members still play a role in their employment propensity. In general, male migrant workers have an advantage over females in job hunting. People currently living with family members are less likely to be employed. The coefficient estimates for these two variables are similar to previous results in Model 1.

The discussion is in order on the first-stage regression results shown in Appendix 1, since weak or invalid instruments can result in measurement error in the endogenous regressor (Bound et al., 1995; Hall et al., 1996; Greene, 1997; Staiger and Stock, 1997; Zhu, 2011, 2012). Several tests are conducted on the relevance of these instruments. First, based on t-statistics, all four instrumental variables have significant impact on whether the migrant worker chooses to live in an urbanizing village. In general, a stronger sense of belonging to the current neighborhood, a stronger preference for associating with people from their hometown, and a more positive attitude towards the sanitation and safety conditions in urbanizing villages lead to a higher possibility for a migrant worker

to live in an urbanizing village. Second, F-statistic from the first stage regression also shows that these instruments are jointly statistically significant. The F-statistic is a well-accepted simple tool used to infer the weakness of instruments. A common rule of thumb suggests that if $F > 10$ then one can treat the instruments as sufficiently strong and the usual 2-stage regression output can be accepted (Stock, 2010; Zhu, 2011, 2012). Our F-statistic for the first-stage regression is 44.58. Third, the Bound–Jaeger–Baker F statistics (see Bound et al., 1995) and "partial R-squared" measures (see Shea, 1997) all suggest these instruments are relevant. For example, the partial R-squared of the instruments accounts for 36 percent of the total R-squared. Fourth, IV redundancy test (LM test of redundancy of specified instruments) shows that none of these instruments are redundant. The above tests suggest that the IV-probit model is able to address the endogeneity problem and provide more plausible coefficient estimates for the instrumented urbanizing village variable.

6.2 Wage

Model 3 in Table 3 provides the OLS results on how socioeconomic factors and urbanizing village affects migrant workers' wage. After controlling for individual socioeconomic factors, occupations, and regional variables, the OLS model suggests that urbanizing village does not significantly affect the wage of migrant workers. Among socioeconomic factors, male migrant workers earn higher wages than females. Age and education background also have positive impact on the earnings. The length (number of years) of interviewees' first non-agricultural job helps to gain a higher wage in his/her current job. Migration history also matters. The number of the provinces the interviewee

had been to and the number of years he/she had stayed in current city both help gain a wage premium. As expected, occupation plays an important role in earnings. Migrant workers who are self-employed or private entrepreneurs and managers earn much higher wages than those in other occupations. Across different regions, migrant workers living in the Pearl River Delta Region and the Yangtze River Delta Region earn significantly higher wages than in the Bohai Sea Region and the Chengdu-Chongqing Region. This is in line with the regional median wage differences provided by the China Census.

[Table 3 about here]

Arguably, the above OLS model also suffers the same self-selection bias or endogeneity problem as mentioned in the previous section. Thus, three statistical tests (Wu-Hausman F test, Durbin-Wu-Hausman chi-sq test, and Difference-in-Sargan test) are conducted again and they all suggest that the specified regressor--urbanizing village--cannot be treated as exogenous. In next step, this endogeneity problem is addressed through those same four instrumental variables as used in the previous section.

Model 4 in Table 3 presents the wage results using a Two-Stage Least Squares (2SLS) method to address the endogeneity problem associated with the urbanizing village variable. It appears that living in urbanizing villages has a significantly positive impact on the wage of migrant workers. It is estimated that living in urbanizing villages on average has a wage premium of 13.9 percent for migrant workers, holding constant other socioeconomic factors, occupations and regional variables.[2]

In terms of the impact of other variables on wage, the findings are consistent between the OLS model and the 2SLS model, with some efficient estimates only slightly

[2] The percentage change in wage (from Y0 to Y1), for a discrete change in "urbanizing village" dummy variable (from 0 to 1), is calculated as 100(Y1-Y0)/Y0 = 100*(exp{b}-1). This formula is used here as well as in following discussions on other dummy variables.

changed in size. Male migrant workers on average earn 27.1 percent more than females. The age of the respondent has a non-linear effect: before a certain age, it has a positive effect on wage; beyond that age, the effect becomes negative. Based on our model, the cutoff age is 25 years old. This low cutoff age is possibly because migrant workers are mostly serving as low-skilled labor in cities. In terms of education and experience, migrant workers on average receive two percent higher wages with one extra year of education, and they gain one percent higher earnings if they spent one extra year in their first non-agricultural job. Migration history is also found to have significant impacts on wage, and the results are similar to previous ones in the OLS model-- the number of provinces the interviewee had been to helps to raise the wage by three percent, and the number of years the interviewee had stayed in current city increases the wage by one percent. Surprisingly, the average working hours per day has no significant impact. Again, this could be related to the fact that most of these migrant workers are in low-skilled jobs whose earnings are often not based on time spent. Those socially less independent migrant workers who left his/her hometown with fellows or friends and those who were currently living with family members also earn lower wages. As expected, occupation is the most important factor in affecting wages, with private entrepreneurs and managers earning 71.6 percent more than other occupations. Similar to OLS model, the 2SLS model also shows that migrant workers living in Pearl River Delta Region and Yangtze River Delta Region earn significantly higher wages (about 12-14 percent more) than those in Bohai Sea Region and Chengdu-Chongqing Region.

To provide evidence for the relevance of those instrument variables, results of the first-stage regression of the 2SLS model are shown in Appendix 2. The coefficient

estimates for how various factors affect the likelihood of living in urbanizing villages are close to those in Appendix 1. Similar to the previous section on employment propensity, several tests are also conducted based on the first-stage regression, including: t-statistics, F-statistic (36.2), the Bound–Jaeger–Baker F statistics, "partial R-squared" measures (22.6 percent), and IV redundancy test. These tests all support that the instrumental variables are sufficiently strong in the first-stage regression, suggesting that the 2SLS model is able to address the endogeneity problem and provide more plausible coefficient estimates for the instrumented urbanizing village variable. As corroboration we also apply the limited information maximum likelihood (LIML) estimation method in a two-stage context. The LIML estimation method usually performs better than 2SLS when instruments are weak (Stock 2010). The results of LIML are similar to those of 2SLS.

7. Conclusions

Through a survey across four mega-regions that are currently experiencing the most rapid urbanization in China, we collected some unique information on migrant workers' attitudes towards living in urbanizing villages. Therefore we were able to address the self-selection bias that has broadly existed in most previous studies on residential segregation and spatial mismatch. The IV-probit and 2SLS models show that the net effect of residential segregation in urbanizing villages on migrant workers' employment outcomes appears to be positive. Despite the fact that rural migrant workers in cities are not allowed to purchase government-sponsored low-income housing, their residential segregation in urbanizing villages is still largely by choice. They either do not want to spend too much on purchasing or renting the commercial housing in other

neighborhoods, or they prefer living in urbanizing villages because of the various benefits they enjoy from the social networks there.

People select locations for a reason. In a market where people are footloose, they choose places that can optimize their consumption of a bundle of goods and maximize their utility. In the U.S., earlier spatial mismatch hypothesis emphasizes that racial minorities suffer from discrimination in the housing market and are segregated in central city locations. The suburbanized employment opportunities have therefore hampered their employment prospects. In China, residential segregation in urbanizing villages experienced by many rural migrant workers is largely by choice. Urbanizing villages do not just provide rural migrant workers affordable housing, but also improve their employment outcomes. While negative spatial mismatch effects certainly exist in these urbanizing villages, as they are often not in proximity to appropriate job opportunities that satisfy migrant workers' needs, the positive spillover effects of urbanizing villages also plays an important role in improving migrant workers' employment propensity as well as income. Although this study does not intend to separately estimate the magnitude of these two opposite effects, the net effect of urbanizing villages on employment outcomes appears to be positive, suggesting the spillover effects override the spatial mismatch effects.

Migrant workers are the key labor force for many low-skilled jobs that are often essential for a city to prosper. In this regard, urbanizing villages play an important role in promoting sustainable economic development for many cities in China. Currently, officials of many cities are in the process of proposing to demolish existing urbanizing villages, apparently without fully considering the consequences. Their proposals could

further increase the cost of living faced by migrant workers and possibly drive them to live further out into rural areas where public transportation is even worse. Planners and policy makers should consider alternative policies such as affordable housing projects to specifically accommodate migrant workers at appropriate locations (close to blue-collar jobs, transit friendly, etc.) as well as to provide a space for the spillover effects to bloom. These policies would reduce the spatial mismatch effects while maintaining rich network effects and human capital externalities, thus enhancing migrant workers' employment outcomes and facilitating their socioeconomic assimilation into cities.

Aponte, R. (1996). Urban employment and the mismatch dilemma: Accounting for the immigration exception. *Social Problems*, 268-283.

Bertrand, Marianne, Erzo Luttmer, and Sendhil Mullainathan, "Network Effects and Welfare Cultures," *Quarterly Journal of Economics*, CXV (2000), 1019–1055.

Borjas, G. (1995). Ethnicity, Neighborhoods, and Human-Capital Externalities. *American Economic Review*, LXXXV, 365–390.

Bound J, Jaeger DA, Baker RM (1995) Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association 90*(430):443-450.

Cervero, R., Sandoval, O., & Landis, J. (2002). Transportation as a Stimulus of Welfare-to-Work Private versus Public Mobility. *Journal of Planning Education and Research*, 22(1), 50-63.

Cooke, T., & Shumway, M. (1991). Developing the spatial mismatch hypothesis: Problems of accessibility to employment for low-wage central city labor. *Urban Geography*, 12:310-23.

Cutler, D. & Glaeser, E. (1997). "Are Ghettos Good or Bad?" *Quarterly Journal of Economics*, CXII, 827–872.

Damm, A. P. (2009). Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence. *Journal of Labor Economics*, 27(2), 281-314.

Dustmann, Christian, and I. Preston, "Attitudes to Ethnic Minorities, Ethnic Context and Location Decision," CEPR Discussion Paper No. 1942, 1998.

Edin, P. A., Fredriksson, P., & Aslund, O. (2003). Ethnic enclaves and the economic success of immigrants—Evidence from a natural experiment. *The Quarterly Journal of Economics*, 118(1), 329-357.

Elliott, J. R., & Sims, M. (2001). Ghettos and barrios: The impact of neighborhood poverty and race on job matching among blacks and Latinos. *Social Problems*, 48(3), 341-361.

Gabriel, Stuart, and Stuart Rosenthal, "Location and the Effect of Demographic Traits on Earnings," *Regional Science and Urban Economics*, XXIX (1999), 445–461.

Glazer, N., & Moynihan, D. P. (1963). *Beyond the melting pot: The Negroes, Puerto Ricans, Jews, Italians and Irish of New York City*.

Glaeser, E. (2011). *Triumph of the city: How our greatest invention makes US richer, smarter, greener, healthier and happier*. Macmillan.

Greene WH (1997) *Econometric Analysis, 3rd edition*. Prentice Hall, Upper Saddle River, NJ

Handlin, O. (Ed.). (1959). *Immigration as a factor in American history*. Prentice-Hall.

Houston, D. S. (2005). Methods to test the spatial mismatch hypothesis. Economic Geography, 81, 407–434.

Houston, D. S. (2001). Testing the spatial mismatch hypothesis in the United Kingdom using evidence from firm relocations. In *Transport planning, logistics, and spatial mismatch*, ed. D. Pitfield, 134-51. European Research in Regional Science 11. London: Pion.

Ihlanfeldt, K.R. (1993). "Intra-urban job accessibility and hispanic youth employment rates. Journal of Urban Economics. 33, 254-271.

Ihlanfeldt, K.R. (1997). Information on the Spatial Distribution of Job Opportunities within Metropolitan Areas. *Journal of Urban Economics*, 41:218–42.

Ihlanfeldt, K.R., and D. Sjoquist, (1990). "Job Accessibility and Racial Differences in Youth Unemployment Rates," American Economic Review, LXXX, 267–76.

Ihlanfeldt, K. R., & Sjoquist, D. L. (1998). The spatial mismatch hypothesis: A review of recent studies and their implications for welfare reform. Housing Policy Debate, 9, 849–892.

Kain, J. F. 1968. Housing segregation, Negro employment, and metropolitan decentralization. The Quarterly Journal of Economics 82 (2): 175-97.

Kasarda, J. D. (1988). Jobs, migration, and emerging urban mismatches. In M. G. H. McGeary & L. E. Lynn (Eds.), *Urban Change and Poverty*. Washington, DC: National Academy Press. pp. 148–198.

Karsarda, J.D. and Ting, K. (1996). A Joblessness and Poverty in America's Central Cities: Causes and Policy Prescriptions. *Housing Policy Debate*, 7(2): 387-419.

Liu, C.Y. (2009). Ethnic Enclave Residence, Employment, and Commuting of Latino Workers. *Journal of Policy Analysis and Management*, 28(4): 600–625.

Logan, J. R., Alba, R. D., & Zhang, W. Q. (2002). Immigrant enclaves and ethnic communities in New York and Los Angeles. *American Sociological Review*, 67, 299–322.

O'Regan, K. M., and Quigley, J. M. 1998. Where youth live: Economic effects of urban space on employment prospects. *Urban Studies*, 35:1187-1205.

Painter, G., Liu, C. Y., & Zhuang, D. (2007). Immigrants and the spatial mismatch hypothesis: Employment outcomes among immigrant youth in Los Angeles. *Urban Studies*, 44, 2627–2649.

Parks, V. (2004a). Access to work: The effects of spatial and social accessibility on unemployment for native-born black and immigrant women in Los Angeles. Economic Geography, 80, 141–172.

Parks, V. (2004b). The gendered connection between ethnic residential and labor-market segregation in Los Angeles. Urban Geography, 25, 589–630.

Pastor, M., & Marcelli, E. (2000). Social, spatial and skill mismatch among immigrants and native-born workers in Los Angeles. San Diego, CA: Center for Comparative Immigration Studies, University of California.

Preston, V., McLafferty, S., & Liu, X. F. (1998). Geographical barriers to employment for American-born and immigrant workers. *Urban Studies*, 35(3), 529-545.

Raphael, S. 1998. The spatial mismatch hypothesis and black youth joblessness: Evidence from the San Francisco Bay area. *Journal of Urban Economics*, 43:79-111.

Raphael, S., & Rice, L. (2002). Car ownership, employment, and earnings. *Journal of Urban Economics*, 52, 109–130.

Sanchez, T. W. (1999). The connection between public transit and employment: The cases of Portland and Atlanta. *Journal of the American Planning Association*, 65, 284–296.

Song, Y, Zenou, Y. and Ding, C. (2008). Let's Not Throw the Baby Out with the Bath Water: The Role of Urban Villages in Housing Rural Migrants in China. *Urban Studies*, 45(2): 313–330.

Staiger D, Stock JH (1997) Instrumental variables regression with weak instruments. *Econometrica* 65(3):557-586.

Stoll, M. A. (1999). Spatial mismatch, discrimination, and male youth employment in the Washington, DC area: Implications for residential mobility policies. Journal of Policy Analysis and Management, 18, 77–98.

Stock JH (2010) The other transformation in econometric practice: robust tools for inference. Journal of Economic Perspectives 24(2):83-94.

Taylor, B. D., & Ong, P. M. (1995). Spatial mismatch or automobile mismatch: An examination of race, residence and commuting in US metropolitan areas. *Urban Studies*, 32,1453–1473.

Wilson, William Julius. (1987). *The Truly Disadvantaged: The Inner City, the Underclass and Public Policy*. Chicago: University of Chicago Press.

Wirth, L. (1956). *The Ghetto*. Chicago, IL: University of Chicago Press.

Zhang, L., Zhao, S. X. B. and Tian, J. P. (2003) Selfhelp in housing and Chengzhongcun in China's urbanization. *International Journal of Urban and Regional Research*, 27, 912–937.

Zhu, P. (2012) "Are telecommuting and personal travel complements or substitutes?" *The Annals of Regional Science* 48(2) 619-639

Zhu, P. (2013). Telecommuting, Household Commute and Location Choice. *Urban Studies*. DOI: 10.1177/0042098012474520.

Table 1 Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Employment Propensity	2226	0.98	0.15	0	1
Wage	2175	1756	2469	0	60000
Living in the urbanizing village (yes=1,no=0)	2203	0.57	0.50	0	1
Socioeconomic Characteristics					
gender	2226	0.56	0.50	0	1
age	2220	32.37	9.95	16	65
age square	2220	1146.7	714.0	256	4225
marital	2226	0.65	0.48	0	1
education	2226	8.84	3.46	0	18
Interviewee was cadre before he/she left for city	2226	0.03	0.17	0	1
Interviewee is Communist Party member	2226	0.06	0.23	0	1
Migration and Employment History					
Length (years) of his/her first non-agricultural job	2220	8.38	6.22	0	41
Number of provinces that he/she had been to	2224	1.20	1.39	0	20
Number of non-agricultural jobs taken so far	2215	2.00	1.72	0	30
Number of years he/she stayed in current city	2223	5.99	5.02	1	34
Average working hours per day	2173	9.78	2.47	0.5	22
Having relatives or friends in current city	2200	0.54	0.50	0	1
Left hometown with family members	2226	0.34	0.47	0	1
Left hometown with fellows or friends	2226	0.17	0.38	0	1
Living with family members in current city	2226	0.60	0.49	0	1
Occupations					
Construction, Manufacturing, Mining Worker	2173	0.33	0.47	0	1
Self-employed	2173	0.18	0.38	0	1
Clerk (Clerical or Administrative Support)	2173	0.11	0.32	0	1
Private entrepreneur and Manager	2173	0.08	0.26	0	1
Service worker	2173	0.30	0.46	0	1
Institutional official	2173	0.01	0.08	0	1
Economic Regions					
Pearl River Delta Region	2214	0.25	0.43	0	1
Yangtze River Delta region	2214	0.27	0.44	0	1
The Bohai Sea region	2214	0.25	0.43	0	1
Chengyu Region	2214	0.23	0.42	0	1

Notes: 1. Employment propensity is a dummy variable indicates whether interviewee is employed.

2. Wage indicates the employee's monthly income.

3. Education indicates the years of education the interviewee accapted.

4. Self-employed: conducting the business for profit with less than 8 employees.

Table 2 Probit and IVprobit Regression Results for Employment Propensity

Variables	Employment Propensity			
	Model 1 (Probit)		Model 2 (IV-probit)	
Urbanizing village	-0.10	(0.15)	0.68**	(0.33)
Socioeconomic Characteristics				
Gender	0.58***	(0.16)	0.52***	(0.19)
Age	0.05	(0.06)	0.02	(0.07)
Age-squared	0	(0.00)	0	(0.00)
Marital	-0.02	(0.24)	-0.09	(0.31)
Education	0.04*	(0.02)	0.03	(0.03)
Interviewee was cadre before he/she left for city	-0.45	(0.35)	-0.07	(0.48)
Interviewee is Communist Party member	-0.01	(0.34)	-0.21	(0.37)
Migration and Employment History				
Length (years) of his/her first non-agricultural job	-0.03**	(0.02)	-0.02	(0.02)
Number of provinces that he/she had been to	0.11	(0.08)	0.09	(0.09)
Number of non-agricultural jobs taken so far	0.06	(0.06)	0.03	(0.06)
Number of years he/she stayed in current city	0.02	(0.02)	0.03	(0.03)
Having relatives or friends in current city	-0.02	(0.14)	-0.16	(0.17)
Left hometown with family members	-0.10	(0.16)	0.02	(0.18)
Left hometown with fellows or friends	0.14	(0.24)	0.22	(0.30)
Living with family members in current city	-.40**	(0.20)	-0.49**	(0.25)
Economic Regions				
Pearl River Delta Region	0.20	(0.22)	0.18	(0.29)
Yangtze River Delta region	-0.24	(0.19)	-0.28	(0.22)
The Bohai Sea region	0.25	(0.25)	0.35	(0.32)
Intercept	0.75	(0.98)	0.91	(1.27)
Number of obs	2139		2139	
Prob > chi2	0.00		0.02	
Pseudo R2	0.13		-	

Note: standard errors in parentheses. ***p<0.01, **p<0.05,*p<0.1

For regional dummy variables, the reference is "Chengdu-Chongqing Region".

Table 3 OLSt and 2SLS Regression Results for Wage

Variables	Wage			
	Model 3(OLS)		Model 4(2SLS)	
Urbanizing village	0.02	(0.02)	0.13**	(0.05)
Socioeconomic Characteristics				
Gender	0.20***	(0.02)	0.24***	(0.03)
Age	0.04***	(0.01)	0.05***	(0.01)
Age-squared	-.001***	(0.00)	-.001***	(0.00)
Marital	0.04	(0.03)	0.02	(0.04)
Education	0.03***	(0.00)	0.02***	(0.01)
Interviewee was cadre before he/she left for city	-0.02	(0.07)	-0.03	(0.08)
Interviewee is Communist Party member	0.04	(0.05)	0.02	(0.06)
Migration and Employment History				
Length (years) of his/her first non-agricultural job	0.01***	(0.00)	0.01*	(0.00)
Number of provinces that he/she had been to	0.02***	(0.01)	0.03***	(0.01)
Number of non-agricultural jobs taken so far	0	(0.01)	0.01	(0.01)
Number of years he/she stayed in current city	0.01**	(0.00)	0.01***	(0.00)
Average working hours per day	0.01	(0.00)	-0.01	(0.01)
Having relatives or friends in current city	-0.01	(0.02)	-0.01	(0.03)
Left hometown with family members	-0.04	(0.03)	-0.02	(0.03)
Left hometown with fellows or friends	-0.03	(0.03)	-0.07**	(0.04)
Living with family members in current city	-0.02	(0.03)	-0.07**	(0.03)
Occupations				
Construction, Manufacturing, Mining Worker	0.14	(0.14)	0.10	(0.19)
Self-employed	0.32**	(0.14)	0.29	(0.19)
Clerk (Clerical or Administrative Support)	0.16	(0.14)	0.10	(0.19)
Private entrepreneur and Manager	0.50***	(0.14)	0.54***	(0.20)
Service worker	0.04	(0.14)	0.02	(0.19)
Economic Regions				
Pearl River Delta Region	0.17***	(0.03)	0.13***	(0.04)
Yangtze River Delta region	0.16***	(0.03)	0.11***	(0.04)
The Bohai Sea region	0.06*	(0.03)	0.01	(0.04)
Intercept	5.81***	(0.21)	5.89***	(0.27)
Number of obs	2079		2079	
Prob > F / Prob > chi2	0.00		0.00	
R-squared	0.24		0.26	

Note: standard errors in parentheses. ***p<0.01, **p<0.05,*p<0.1

For occupation dummy variables, the reference is "institutional official"

For regional dummy variables, the reference is "Chengdu-Chongqing Region".

Appendix 1 First Stage Regression Results for Ivprobit Model

Variables	Urbanizing village	
Socioeconomic Characteristics		
Gender	0.05***	(0.02)
Age	0	(0.01)
Age-squared	0	(0.00)
Marital	-0.02	(0.03)
Education	0	(0.00)
Interviewee was cadre before he/she left for city	-0.04	(0.05)
Interviewee is Communist Party member	-0.06	(0.04)
Migration and Employment History		
Length (years) of his/her first non-agricultural job	-.004*	(0.00)
Number of provinces that he/she had been to	0.01*	(0.01)
Number of non-agricultural jobs taken so far	0.02***	(0.01)
Number of years he/she stayed in current city	0	(0.00)
Having relatives or friends in current city	0	(0.02)
Left hometown with family members	0	(0.02)
Left hometown with fellows or friends	0.04*	(0.02)
Living with family members in current city	0.07***	(0.02)
Economic Regions		
Pearl River Delta Region	0.14***	(0.03)
Yangtze River Delta Region	0.09***	(0.03)
The Bohai Sea Region	0.05*	(0.03)
Instrumental Variables		
Sense of belonging to the current neighborhood	0.22***	(0.02)
Preference for associating with people from hometown	0.05***	(0.02)
Attitude towards the sanitation condition in urbanizing villages	0.22***	(0.02)
Attitude towards the safety condition in urbanizing villages	0.21***	(0.02)
Intercept	0.36***	(0.12)
Adj R-squared	0.43	
F	44.58	
Prob > F	0.00	

Note: standard errors in parentheses. ***p<0.01, **p<0.05,*p<0.1

For regional dummy variables, the reference is "Chengdu-Chongqing Region".

Appendix 2 First Stage Regression Results for 2SLS Model

Variebles	Urbanizing village	
Socioeconomic Characteristics		
Gender	0.05**	(0.05)
Age	0	(0.00)
Age-squared	0	(0.00)
Marital	-0.02	-(0.02)
Education	0	(0.00)
Interviewee was cadre before he/she left for city	-0.05	-(0.05)
Interviewee is Communist Party member	-0.04	-(0.04)
Migration and Employment History		
Length (years) of his/her first non-agricultural job	-0.005**	(0.00)
Number of provinces that he/she had been to	0.02**	(0.02)
Number of non-agricultural jobs taken so far	0.02***	(0.02)
Number of years he/she stayed in current city	0	(0.00)
Average working hours per day	0.01***	(0.01)
Having relatives or friends in current city	-0.01	-(0.01)
Left hometown with family members	-0.01	-(0.01)
Left hometown with fellows or friends	0.04*	(0.04)
Living with family members in current city	0.07***	(0.07)
Occupations		
Construction, Manufacturing, Mining Worker	-0.04	-(0.04)
Self-employed	-0.12	-(0.12)
Clerk (Clerical or Administrative Support)	-0.10	-(0.10)
Private entrepreneur and Manager	-0.16	-(0.16)
Service worker	-0.15	-(0.15)
Economic Regions		
Pearl River Delta Region	0.13***	(0.13)
Yangtze River Delta Region	0.08***	(0.08)
The Bohai Sea Region	0.03	(0.03)
Instrumental Variables		
Sense of belonging to the current neighborhood	0.22***	(0.22)
Preference for associating with people from hometown	0.04*	(0.04)
Attitude towards the sanitation condition in urbanizing villages	0.21***	(0.21)
Attitude towards the safety condition in urbanizing villages	0.21***	(0.21)
Intercept	0.32*	(0.32)
Adj R-squared	0.44	
F	36.22	
Prob > F	0	

Note: standard errors in parentheses. ***p<0.01, **p<0.05,*p<0.1

For occupation dummy variables, the reference is "institutional official"

For regional dummy variables, the reference is "Chengdu-Chongqing Region".