ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Cai, Shu; Zimmermann, Klaus F.

Working Paper

Social assimilation and labor market outcomes of migrants in China

ZEF Discussion Papers on Development Policy, No. 304

Provided in Cooperation with:
Zentrum für Entwicklungsforschung / Center for Development Research (ZEF), University of Bonn

Suggested Citation: Cai, Shu; Zimmermann, Klaus F. (2020) : Social assimilation and labor market outcomes of migrants in China, ZEF Discussion Papers on Development Policy, No. 304, University of Bonn, Center for Development Research (ZEF), Bonn

This Version is available at:
https://hdl.handle.net/10419/229982

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.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

zef
Center for
Development Research
University of Bonn

ZEF Discussion Papers on
Development Policy No. 304

Shu Cai and Klaus F. Zimmermann

Social Assimilation and Labor Market Outcomes of Migrants in China

Bonn, December 2020

The **Center for Development Research (ZEF)** was established in 1995 as an international, interdisciplinary research institute at the University of Bonn. Research and teaching at ZEF address political, economic and ecological development problems. ZEF closely cooperates with national and international partners in research and development organizations. For information, see: www.zef.de.

ZEF Discussion Papers on Development Policy are intended to stimulate discussion among researchers, practitioners and policy makers on current and emerging development issues. Each paper has been exposed to an internal discussion within the Center for Development Research (ZEF) and an external review. The papers mostly reflect work in progress. The Editorial Committee of the ZEF DISCUSSION PAPERS ON DEVELOPMENT POLICY includes Joachim von Braun (Chair), Christian Borgemeister, and Eva Youkhana. Alisher Mirzabaev is the Managing Editor of the series.

Shu Cai, Klaus F. Zimmermann, Social Assimilation and Labor Market Outcomes of Migrants in China, ZEF Discussion Papers on Development Policy No. 304, Center for Development Research, Bonn, December 2020, pp. 27.

ISSN: 1436-9931

Published by:
Zentrum für Entwicklungsforschung (ZEF)
Center for Development Research
Genscherallee 3
D – 53113 Bonn
Germany
Phone: +49-228-73-1861
Fax: +49-228-73-1869
E-Mail: zef@uni-bonn.de
www.zef.de

The authors:
Shu Cai, Jinan University and Global Labor Organization. Contact: shucai.ccer@gmail.com
Klaus F. Zimmermann, UNU–MERIT, Maastricht University, CEPR and Global Labor Organization. Contact: klaus.f.zimmermann@gmail.com

Abstract

Previous research has found identity to be relevant for international migration, but has neglected internal mobility as in the case of the Great Chinese Migration. However, the context of the identities of migrants and their adaption in the migration process is likely to be quite different. The gap is closed by examining social assimilation and the effect on the labor market outcomes of migrants in China, the country with the largest record of internal mobility. Using instrumental variable estimation, the study finds that identifying as local residents significantly increases migrants' hourly wages and reduces hours worked, although their monthly earnings remained barely changed. Further findings suggest that migrants with strong local identity are more likely to use local networks in job search, and to obtain jobs with higher average wages and lower average hours worked per day.

Keywords: Social assimilation, identity, labor market, migration

JEL codes: J22, J31, J61, Z13

1. Introduction

As a form of reallocating labor resources across regions according to wage differentials, migration is an important engine of increasing aggregate productivity and promoting economic growth (Banerjee and Duflo, 2005; Tombe and Zhu, 2019; Bryan and Morten, 2019). While economic assimilation of immigrants has received considerable attention from economists (Chiswick, 1978; Borjas, 1985; Lubotsky, 2007; Abramitzky et al., 2012), research on the social assimilation of migrants in their host places, however, has focused on international mobility. With the identity literature pioneered by Akerlof and Kranton (2000), the self-identity of migrants and their adaption to identify with host areas in the process of migration was recognized to be an important factor for economic decisions and labor market outcomes (Battu et al., 2007; Constant and Zimmermann, 2008, 2011; Constant, Gataullina and Zimmermann, 2009).

In this study, we therefore close the evidence gap on internal labor mobility, which faces different challenges than international migration. Specifically, we investigate the impact of adaption to local identity in host areas on labor market performance among migrants who worked as employees, by using data on China, the country with the largest recent migration experience but no research on identity. The Dynamic Monitoring Survey of Migrant Population of China used was conducted among a representative population of migrants in eight prefectures in China in 2013, and collected detailed information on migrants' identity and social assimilation to host areas and their labor market outcomes, thus providing us a unique opportunity to examine this question.

To address the endogeneity problem, we explore exogenous variation in assimilation caused by linguistic distance between the original province of migrants and their host county. In addition, we control for individual sociodemographic characteristics and fixed effects of the original province and the host county. The instrumental variable estimates suggest that adaption to local identity significantly increases hourly wage and reduces the average hours worked per day and the likelihood of overworking, keeping the monthly wage unaffected.

To account for the communication effect of language on migrants' labor market outcomes, we control for individual language skills in the regressions. The results indicate that the capability of speaking or understanding a local dialect are not significantly correlated with migrants' labor market outcomes. This might not be surprising, given that most migrants are able to communicate at the workplace by using Mandarin Chinese. The instrumental variable estimates presented are robust to taking account of the potential communication effect of language. This should reduce concerns regarding possible violation of the exclusion restriction for identification.

In order to examine the potential role of social networks through which social assimilation may affect migrants' labor market outcomes, we investigate the impacts of assimilation on

migrants' access to networks of local people and the utilization of local networks in their job search. The results suggest that adaption to local identity significantly raises migrants' probability of interacting with local people and having local neighbors. In addition, socially assimilated migrants are also more likely to find a job with help from local networks. These results highlight the importance of social networks for explaining the advantages of adaption to local identity. Interestingly, we find the labor market impacts of social assimilation still to be robust even if we control for occupation, industry, and work unit, indicating that there might be substantial differences in the quality of jobs even within occupation, industry, and type of work unit.

Akerlof and Kranton (2000) provided a theoretical framework to consider how identity affects economic outcomes. In the model, they extend the utility function by including a person's sense of self, namely identity. In order to achieve a better self-image, individuals may choose a seemingly sub-optimal choice, but their overall utility is maximized. For instance, people might be willing to "pay" an income penalty in choosing an occupation in order to reinforce their identity.[1] Battu et al. (2007) developed a model in which non-white individuals determine the level of adaptation to white culture by balancing peer pressure from same-race friends and beneficial effect of the high quality jobs through whites' social networks which do not suffer from discrimination.

Using a tangible measure of ethnic identity, i.e., the ethnosizer, constructed from individual information on language, culture, social interaction, history of migration, and self-identification, Constant and Zimmermann (2008) classified international migrants concerning their ethnic identity as assimilated, separated, marginalized (belonging to nowhere) and integrated (being close to the culture of host and sending country cultures). Different to this, internal mobility in China seems to focus much more on the alternatives to assimilate or to keep the origin region identity.

This study builds on a small but growing pool of literature that investigates the relationship between migrants' identity and their labor market outcomes in host places with a focus on international migrants. Constant and Zimmermann (2008, 2009) found for immigrants in Germany that assimilated men and women are more likely to work, and women who exhibit the integrated identity are more likely to work than women who are German assimilated, but this does not hold for men. They also received no significant relationship between ethnic identity and the earnings of men or women. Casey and Dustmann (2010) confirmed a positive association between German identity and employment for females but not for males using German panel data. They also provided evidence for a positive association between home country identity and employment for only the males among second generation immigrants. Battu and Zenou (2010) presented evidence for an employment penalty associated with oppositional identity among ethnic minorities in the UK. Exploiting genetic distance between

[1] See Alesina and Ferrara (2005) for more discussion on the endogenous formation of ethnic identity.

immigrants' home and host countries as instruments for immigrants' identity, Islam and Raschky (2015) documented that immigrants' ethnic identity does not have a strong influence on their labor market outcomes in the host country with Canadian survey data. In the study of Gorinas (2014) the employment of immigrants in Denmark did not systematically associate with measures of ethnic identity, yet were significantly related to openness to majority norms, particularly for first-generation immigrant women.[2]

Our study contributes to the extant literature in several aspects. First, while most previous literature studies identity and migratory performance for international migrants in developed countries, we focus on internal migration within China, the largest developing country in the world. Second, the extant studies primarily gauge the likelihood of employment or earnings as labor market outcomes in host places. We complement the literature by examining the quality of jobs, including working hours, likelihood of overworking, hourly wage, as well as monthly earnings. Third, evidence from most previous studies only indicate the association between immigrants' identity and labor market outcomes. In this study, we provide causal evidence on the relationships by using exogenous variation in social assimilation caused by cultural difference between the original and host places. Fourth, we highlight the role of native networks on explaining the labor market advantages of holding identity committed to the host place. While networks have been emphasized theoretically as an important channel through which ethnic identity may affect immigrants' labor market outcomes (Batttu et al., 2007; Verdier and Zenou, 2017), they are seldom examined in empirical studies.

The rest of the paper is organized as follows. Section 2 introduces the background of the study. Section 3 describes the data and measurements. Section 4 lays out the empirical specification and estimation strategy. Section 5 presents the empirical results. The final section concludes.

[2] Related studies include immigration in Sweden (Edin et al., 2003; Nekby, 2010), Canada (Pendakur and Pendakur, 2006), France (Delaporte, 2019), and Europe (Bisin et al., 2011).

2. Background of the Study

2.1 The Great Chinese Migration

China has witnessed a massive flow of migration from the interior to the coast, or from poor rural areas to more developed urban areas. According to the National Bureau of Statistics of China, an estimated 245 million of people had migrated outside of their home township over six months in 2013, about 18% of its total population in that year. This is about ten times the size of immigration from Europe to the U.S. during the Age of Mass Migration (Sequeira et al., 2020). Within the great flow, a substantial amount is job-related migration caused by the large wage differentials across regions. This is partly due to a decreased marginal labor productivity in the agricultural sector as a result of abundant laborers, and partly due to the accelerated development of manufacturing and construction industries in urban areas, mainly in coastal cities, after China became a member of the World Trade Organization (WTO).

The great migration is accompanied by difficulties of social assimilation due to the vast cultural differences between the places of origin and destination and institutional restrictions such as the hukou (the local household registration) system. China has a vast land spanned by many degrees of latitude and longitude with varied climate zones and complicated terrain. It also has a large population size with diverse cultures (Talhelm et al., 2014). Migrants thus may face great challenges due to large differences concerning language, customs, attitudes, eating habits, and other lifestyle. Regarding institutional issues, the conversion to local urban hukou and the related social benefits (such as pension, education, medical insurance, permission to purchase housing and vehicle) is still quite restrictive for non-hukou migrants, even though non-hukou migration has been tolerated. These cultural and institutional barriers hinder migrants from adapting to local identity, which in turn may affect their labor market performance.

2.2. Dialects in China

China is unique in its language which has a unified writing system whereas its spoken language varies substantially across regions. The geographic variation of dialect is the result of historical interactions across regions and linguistic evolutionary processes involving mass migration flows, military borders, and political events. Thus, the similarity of dialect between regions may be informative about these historical interactions and indicate similarity in cultural identity (Falck et al., 2012; Suedekum, 2018).[3]

[3] Although culture is a broader concept than language, which includes other domains such as traditions, habit, and beliefs, it is well accepted that language is an important and clear indicator of culture (Herrmann-Pillath et al., 2014). Comparing with dialect, ethnicity and religion are much more homogeneous in China.

Differences in dialect may affect social assimilation of migrants for serval reasons. Individuals may bear psychological costs when interacting with people speaking different dialects, or they may be discriminated. All this may hinder their social assimilation in the host society.[4] It is worth to mention that in addition to the cultural effect of dialect on social assimilation, there might be communication effects on their labor market outcomes as well. This should be less of a concern in our study setting, however. Given the popularization of Mandarin Chinese (i.e., putonghua), people can easily communicate with each other.[5] In addition, we use linguistic distance to capture the cultural effect of language, rather than language skills of the individuals. Lastly, we check the robustness of results by controlling for individual skills of local dialects. Thus, we identify variations in social assimilation caused by dialect differences reflecting persistent cultural differences across regions that developed over time, on the top of communication effects on labor market outcomes.

[4] Dialect similarity may affect social assimilation through the sorting of migration as well. That is, people may be more likely to move to an area with a similar dialect than to an area with a very different dialect.

[5] According to the Putonghua Popularization Survey conducted by the State Language Commission in 2010, approximately 70% of the Chinese population can speak putonghua compared to about 50% by the end of last century.

3. Data and Measurements

The data used in this study are from the Dynamic Monitoring Survey of Migrant Population of China in 2013. Starting in 2009, the National Health and Family Planning Commission of China (NHFPC) conducted a nation-wide survey of the migrant population on an annual basis. In 2013, the survey included a special module on social assimilation in eight prefectures: four from Eastern China, including Songjiang district in Shanghai, Suzhou and Wuxi prefectures in Jiangsu Province, and Quanzhou prefecture from Fujian Province; two from Middle China, including Wuhan prefecture from Hubei Province and Changsha prefecture from Hunan Province; and two from Western China, including Xi'an and Xianyang prefectures from Shan'xi Province. Figure 1 illustrates the geographic location of the eight prefectures.

The full population on which the sampling is based includes all migrants aged 15 - 59 (inclusively). A person is considered to be a migrant if he or she lived in a county for at least one month, whereas his or her hukou was registered outside the county he or she lived at the time of the survey. The survey uses a multi-step Probability Proportionate to Size Sampling (PPS) method for the sampling. In the first step, the survey selected the township according to the PPS method within each prefecture. For each of the selected townships, the survey then chose the sampling unit, namely, villages or communities, using the PPS method. In the last step, the survey randomly chose 20 migrants in each sampling unit.

The designed sample size of the eight prefectures is as follows: Songjiang (2,000), Suzhou (4,000), Wuxi (2,000), Quanzhou (2,000), Wuhan (2,000), Changsha (1,880), Xi'an (2,000), and Xianyang (1,000). In the data set, the sample size of Suzhou and Wuhan is 3,999 and 1,999, respectively, whereas the sample sizes of the other prefectures are equal to the designed sample size. There are in total 16,878 migrants, from 8 prefectures, 68 counties, and 844 villages or communities in the data set.

The survey collected detailed information on migrants' demographic and social characteristics, migration experience, employment, income, and so on. In particular, related to our main outcome of interest, the survey contains information on labor market performance of the respondents, including monthly income and work time (average days per week, average hours per day). It also contains information on other labor market characteristics, which include employment type (employee, employer, self-employed, and others), occupation, industry, and types of work unit.[6] Given the difficulties to separate earnings between those from labor inputs and those from capital inputs for employer and the self-employed, we restricted the sample to only employees. For respondents from the eight prefectures, the survey also asked the question "Which of the following types of identity do

[6] The types of work unit include land contractor, state organization, state-owned and state-holding enterprise, collective enterprise, individual business, private enterprise, Hong Kong, Macao and Taiwan Enterprise, Japanese and Korean enterprise, European and American enterprise, Chinese-Foreign equity joint ventures, others, and not applicable.

you think you belong to?" The answers to the question include "local citizen", "new local citizen", "the citizen of your hometown", and "do not know". Only three percent of the respondents said they do not know, which includes all individuals who either cannot affiliate with either host and home areas or affiliate more or less equally with both. Since the number of those who cannot decide is surprisingly small, we drop those respondents from our sample.[7] We therefore can measure social assimilation by a dummy, which equals 1 if respondents say they feel they are local citizens or new local citizens, and 0 if they feel they are citizens of their hometown. Some 45% of our sample of migrants who are employees are affiliating with the host region.

In addition to the survey data, we also use linguistic data to construct the linguistic distance between the original place and the current residential place of the migrants. The linguistic data on local dialects is from the Chinese Dialect Dictionary (Xu and Ichiro, 1999), which is based on a detailed census done by massive on-site investigation in 1983-1987.[8] It identifies the main Chinese dialects and draws a dialect tree constituted by ten dialectal super-groups, 20 dialectal groups, and 105 dialectal sub-groups, according to similarity of phonological and grammatical attributes, such as pronunciation. The dictionary also classifies every county in China into a dialectal sub-group. Using the linguistic atlas of China, we construct the linguistic distance between county-pairs to measure their dialect similarity. Specially, following Spolaore and Wacziarg (2009), we coded the distance as 0 if the dialects of the two counties belong to the same dialectal sub-group, 1 if they belong to the same dialectal group but different sub-groups, 2 if they belong to the same dialectal super-group but different dialectal groups, and 3 if they belong to different dialectal super-groups. With the county-level matrix of dialect similarity in hand, we then construct a measure of linguistic distance between the residential county of migrants and the province they coming from, using the population weighted dialect distance between the residential county and each county in the original province.[9]

Table 1 provides summary statistics for labor market outcomes and demographic characteristics. The number of observations is 9,790, and 45% of the sample feel they belong to local citizens. Column (1) reports the sample mean of migrants who feel that they belong to the local citizens, whereas Column (2) reports the sample mean of migrants who do not feel they belong to the local citizens. The last two columns report the difference between two groups and the p-value of the hypothesis that the difference is equal to 0. For the labor market outcomes, the monthly income in assimilated group is 179 *yuan* higher than the un-

[7] This excludes what the identity literature has called "integration" (affiliation with both host and origin areas) and "marginalization" (cannot affiliate with both) shown to be relevant for international migrants. See Constant, Gataullina and Zimmermann (2009), for instance.

[8] This data has also been used by Liu et al. (2015). We acknowledge them for making a digital version of the data available to us.

[9] We only know the original province of the migrants. The weights are constructed by using data from population census in 2000.

assimilated group, and the difference is statistically significant. The difference of hourly wage between the two groups is 1.59 *yuan* per hour, which is significantly different from 0. The next three rows show that socially assimilated migrants have significantly less working time, including average working days per week, average hours per day and average hours per week, than the un-assimilated migrants. On average, socially assimilated migrants work 0.34 hours or about 20 minutes less every day. The next three rows summarize the situations of overwork in both groups. It is worth noting that overwork is common among migrants, which can be illustrated by the high averages. The un-assimilated migrants overwork much more than socially assimilated ones, and the differences are statistically significant. Regarding the demographic characteristics, the age gap between the two groups is not large, but statistically significant. There is no significant difference in gender. On average, married migrants are more committed to local identity than unmarried ones. The last five rows summarize the education level of migrants. Generally, the assimilated migrants are more likely to have a higher educational level. We study further details about the relationship between social assimilation and labor market outcomes below in the regression analyses.

4. Empirical Strategy

In the empirical analysis, we estimate the following equation for the determinants of migrants' labor market outcomes:

$$y_{icp} = \alpha_0 + \alpha_1 assimilation_{icp} + \alpha X_{icp} + \delta_c + \delta_p + \varepsilon_{icp} \qquad (1)$$

where i represents individuals, c represents the current residential county of individual i, p represents the home province of the individual. y_{icp} are the labor market outcomes of individual i, including monthly income, hourly wage, working time, and so on. The key explanatory variable is social assimilation, which is a dummy and equal to 1 if the individual feels that he or she belongs to the group of locals (local citizen or new local citizen), and 0 otherwise. X_{icp} is the vector of control variables, including age, age squared, dummy of male, marital status (including dummies of married first time, married second or more times, divorced, widowed), education categories (including dummies of education level of middle school, education level of high school, education level of college, and education level above college). δ_c and δ_p are the fixed effects of the county of current residence and the province of the original place of the respondents, respectively. ε_{icp} is the error term, which is clustered by sampling unit, namely communities. The parameter α_1 is our main interest, indicating how social assimilation of the migrants affects their labor market outcomes.

To address the endogeneity problem, we use linguistic distance between the residential county of the respondents and their original province as an instrumental variable for their social assimilation. Previous studies have shown that linguistic distance is an important determinant of social assimilation (Konya, 2007; Fouka, 2020; Ginsburgh and Weber, 2020). We thus assume that social assimilation is a function of the following determinants:

$$assimilation_{icp} = \beta_0 + \beta_1 linguistic\ distance_{icp} + \beta X_{icp} + \delta_c + \delta_p + \tau_{icp}. \qquad (2)$$

The variable *linguistic distance*$_{icp}$ is the instrumental variable which measures the distance between the language spoken at the residential county c of migrant i at the time of the survey and the language spoken at his or her original province p. We also control for individual characteristics X_{icp} that may affect social assimilation, which are the same as in equation (1). In addition, we control for the fixed effects of the current residential county and the fixed effects of the original province, which absorb those determinants of social assimilation that are common to all migrants at the same destination county or from the same original province. In particular, we control for factors such as economic conditions and public policy of the destination counties via the fixed effects δ_c and those of the original provinces via the fixed effects δ_p. The exclusion condition of our instrumental variable estimation is that, conditional on the individual characteristics X_{icp} and the fixed effects of residential county and original province, the linguistic distance between the host county and the home province affects migrants' labor market outcomes only through their social assimilation.

5. Results

5.1 Main Results

Table 2 reports the Ordinary Least Squared (OLS) regression results of a variety of labor market outcomes on the assimilation dummy variable as specified in equation (1). Column (1) shows that the correlation between monthly income and feeling local is positive, but not statistically significant. Column (2) shows that assimilated migrants have higher hourly wages. On average, the hourly wage of the socially assimilated group is 0.66 *yuan* higher than of the un-assimilated group. Columns (3) to (5) report a negative and significant relationship between working time and assimilation. On average, the feeling of belonging to the local citizens is correlated with 1.48 hours decrease in working hours every week. Columns (6) to (8) show a negative and significant association between overwork and social assimilation. Specifically, the feeling of belonging to the local citizens is associated with a four-percentage point decrease in the probability of working over 40 hours per week.

Panel A of Table 3 reports the two-stage least squares (2SLS) regression estimates of the impact of social assimilation on the labor market outcomes from equations (1) and (2). Column (1) displays a strong first-stage relationship between assimilation and linguistic distance in our sample. The greater the linguistic distance between origin and destination, the lower the likelihood of holding identity committed to the host place. If a migrant who originally moved within the same dialect sub-group, now chooses to move outside the dialect sub-group, the probability of assimilation would be 10 percentage points lower. The first-stage Kleibergen-Paap F-statistic of the instrumental variable is 52.3, which is by far greater than 10 and suggests no weak instrumental variable problem. Column (2) shows that social assimilation increases migrant's monthly income, though the coefficient is economically and statistically insignificant. Column (3) shows that social assimilation can increase migrants' hourly wages, which is marginally significant (t=1.57). On average, feeling assimilated can increase migrants' hourly wages by 3.25 *yuan*, which represents about a 24-percentage point increase above the 13-*yuan* baseline hourly wage.

In Columns (4) to (6), we consider whether assimilation would change migrants' working time. In Column (6), we find that assimilation significantly reduces average working time per week by about nine hours. Columns (7) to (9) show that social assimilation reduces the probability of overworking for migrants. The impact on overwork for over five days per week is -0.22, which is statistically significant. It is worth noting that social assimilation reduces the probability of overworking beyond the regular eight hours a day by 44 percentage points, which is a large magnitude compared with the average rate of overworking, namely 48

percentage points. In other words, social assimilation almost solely eliminates the likelihood of migrants' daily overwork. Column (9) indicates that social assimilation reduces the probability of weekly overwork for more than 40 hours by 26 percentage points.

To address the concern that the instrumental variable, namely, the linguistic distance between original and destination provinces, may affect labor market outcomes through the communication impact of language, we further control for two dummies indicating whether the migrants can speak the local language and whether they can understand it in equations (1) and (2). The results are reported in Panel B of Table 3. As shown, the coefficients of the two dummies are not significantly different from 0 in regressions of all labor market outcomes, except that migrants who speak the local language are more likely to be worked over 40 hours a week, whereas if they can understand the local language this dummy is not correlated with the likelihood of working more than 40 hours a week. This might not be surprising, given that most migrants may have no difficulty to communicate at the workplace by using Mandarin Chinese. By controlling for the communication effect of language, the impact of social assimilation is similar to those in Panel A of Table 3. Social assimilation reduces hours worked per day or per week. It also reduces the likelihood of overworking. If anything, the results are even stronger than the benchmark estimates. Overall, these results suggest that the communication effect of language is unlikely to seriously affect conclusions from our instrumental variable estimates.

5.2 Mechanisms

To investigate the mechanisms of how social assimilation affects labor market outcomes, we consider the effects of assimilation on migrants' networks and residence choice. In addition, we explore the connections migrants use during job search, which are important for obtaining higher-quality jobs.

We first investigate the effect of social assimilation on migrants' network and neighborhood choice. Columns (1) and (2) in Table 4 show that social assimilation significantly increases the probability of migrants interacting with locals, whereas it reduces the probability of interacting with people from their own original place of living. Social assimilation also lowers the chance of participating in the activities of ethnic organizations, though the result is not statistically significant. Columns (4) to (6) report the estimated effects of social assimilation on migrants' residence choice. We find that social assimilation significantly increases the probability of having local neighbors by 21 percentage points and reduces the probability of having non-local neighbors by 37 percentage points. These results indicate that socially assimilated migrants are more likely to interact with local citizens and less likely to interact with people from their hometown ethnic group. They are also more likely to live in a community with mostly local citizens and less likely to live with non-local citizens.

Personal relationships, measured here using data on daily interaction and residence choice, are a key factor in job search. To examine the role of networks, Table 5 examines the impacts of assimilation on migrants' job search. We find that social assimilation significantly increases the probability of finding jobs through local people by 17 percentage points. They are also more likely to find a job self-dependently. However, the socially assimilated migrants are less likely to find a job through family members, relatives or friends. It is interesting that social assimilation does not affect the probability of finding jobs through government-related channels.[10] Social assimilation significantly increases the probability of finding jobs through local connections and decreases the probability of finding jobs through family and friends, implying that networks with local citizens are an important channel through which migrants may obtain a high-quality job.

To explore to what extent the benefits of social assimilation for labor market outcomes are through job choice, we further control for occupation, industry, and work unit of the migrants in equation (1) above and examine the following regressions:

$$y_{icp} = \alpha_0 + \alpha_1 assimilation_{icp} + \alpha X_{icp} + \gamma D_{icp} + \delta_c + \delta_p + \varepsilon_{icp} \qquad (3)$$

where D_{icp} is a vector of dummies indicating types of occupation, industry, and work unit of migrant *i*. Table 6 reports the IV estimates of equation (3) by using linguistic distance between original and destination provinces as an instrumental variable for the feeling to belong to the local citizens. The coefficients on work time and propensity of overwork are still negative and significant, although the magnitudes are smaller in absolute values than those in Panel A of Table 3. These results indicate that choices of occupation, industry, and type of work unit can partially explain the impact of social assimilation on reducing work time, likely through the beneficial effects of networking with local citizens. The significantly higher hourly wage and the lower likelihood of overwork indicate there might be substantial differences in the quality of jobs even within the same occupation, industry, and type of work unit.

5.3 Extensions

The previous results show that social assimilation can reduce the work time of migrants and increase their hourly wages, but how does this associate with other measures of local integration? Table 7 reports instrumental variable estimates of the impact of social assimilation on migrants' subjective statements regarding their attitudes and perceptions concerning the citizens at the destination area. Socially assimilated migrants are significantly more willing to be neighbors of local citizens. They are also more willing to be friends with local citizens and are more willing to be involved in local communities, though the impacts are statistically insignificant. Column (4) shows that social assimilation increases migrants' sense

[10] Similar findings are obtained in the standard international migration literature.

of belonging to the city they are were living in. Column (5) indicates that assimilated migrants are more willing for themselves or their relatives to marry local citizens.

Columns (6) to (8) represent the migrants' perception of local citizens' attitudes toward migrants. The socially assimilated migrants are more likely to feel that local citizens are willing to accept migrant as one of them (statistically significant). But there is no significant impact on the feeling that local citizens are willing to be their neighbors and that local citizens like and respect them.

Overall, the results indicate that socially assimilated migrants are significantly more willing to be neighbors of local citizens, feel they belong to the host place, are more willing to marry local citizens and feel that they are accepted by local citizens.

5.4 Heterogeneity of Social Assimilation Effects

In this section, we examine the heterogeneity of the impacts of social assimilation on labor market outcomes of migrants by examining the following regression equation:

$$y_{icp} = \alpha_0 + \alpha_1 assimilation_{icp} + \alpha_2 assimilation_{icp} \times d_{icp} + \alpha X_{icp} + \delta_c + \delta_p + \varepsilon_{icp} (4)$$

where d_{icp} is a dummy and equal to 1 if migrant i is male and 0 otherwise. The other regressors are defined as before. The parameter α_1 indicates the impact of social assimilation on labor market outcomes among female migrants, and α_2 indicates the gender difference in the effect of social assimilation. Thus, $\alpha_1 + \alpha_2$ indicates the impact of social assimilation on labor market outcomes among male migrants. Given that both the social assimilation variable and its interaction with the dummy variable d_{icp} are endogenous, we use language distance and the interaction term of the language distance with d_{icp} as an instrumental variable. To test for weak instruments, we report the cluster-robust Kleibergen-Paap (KP) F-statistic. We also report p-values on testing whether the estimator of $\alpha_1 + \alpha_2$ is different from 0.

In addition, we examine the heterogeneity of the impacts along age and the education level of migrants. Specifically, we define a dummy indicating that migrants are young if they are 30 years old or younger than 30 years old, and define that a migrant is highly educated if he or she has completed college education or above.

Table 8 report results on the heterogeneity analysis. Panel A shows that the average hours worked per day significantly decrease among female migrants as they are locally assimilated. Their probability of working over eight hours per day declines significantly as well. However, there is essentially no change in their hourly wage as a result of social assimilation. For male migrants, social assimilation increases the hourly wage by 5.58 *yuan* (see the estimator of $\alpha_1 + \alpha_2$), which is significantly different from 0. Social assimilation also significantly decreases daily work hours of male migrants, but the magnitude is about half of the impact among female migrants. The results indicate that assimilation mainly reduces work time among

female migrants, while socially assimilated male migrants tend to have jobs with higher hourly wages and less work time.

Panel B examines the heterogeneity of assimilation impacts between young and old migrants as defined earlier. As shown, for most of the outcome variables, the coefficients of the interaction term are small and not significantly different from 0, indicating there is no notable difference in the impact between young and old migrants. These are confirmed by the similarity of the estimated effects on labor market outcomes of the two groups. Specifically, the results suggest, for both young and old migrants, that social assimilation increased the hourly wage and significantly reduced the work hours and likelihood of being overworked.

Panel C presents the results of the heterogeneity analysis by education level of migrants. We define that a migrant is highly-educated if he or she has completed college education or above, which holds for about 13 percent of the sample. As shown, social assimilation significantly raises monthly earnings and hourly wages of the less educated migrants, whereas it also significantly reduces their work time. The coefficients of the interaction term indicate that the impacts are differ significantly between the less- and the highly- educated migrants. For the highly-educated migrants, social assimilation has no impacts on their work time, whereas their hourly wage and monthly earnings reduce significantly as they assimilate locally.

6. Conclusion

The paper examines the impact of social assimilation of migrants on their labor market outcomes in China. Exploring variation in assimilation caused by linguistic distance between the original and current place of living, we find that social assimilation increases migrants' hourly wages, and reduces the average working hours and the likelihood of overwork.

Specifically, the hourly wage increases by 3.25 *yuan*, or 24% of its mean, as migrants socially assimilate to the local place. On the other hand, local assimilation also significantly reduces the average work time per day by 1.17 hours, which eliminates the likelihood of working longer than eight hours per day, the standard working hours.

Further analyses on the mechanisms suggest that social assimilation significantly raises the propensity of interacting with locals and living in local neighborhoods. It also significantly increases the likelihood of finding a job with help from locals, likely through the extended access to local networks. In line with these results, we find a negative effect of social assimilation on working hours and propensity of overwork reduced by controlling for job characteristics, including types of occupations, industries, and work units. However, the benefits of assimilation on labor market outcomes among migrants are still significant after accounting for these job characteristics, indicating there might be differences in the quality of jobs even within the same occupation, industry, and type of work unit.

References

Abramitzky, Ran, Boustan, Leah Platt, Eriksson, Katherine. 2012. "Europe's Tired, Poor, Huddled Masses: Self-selection and Economic Outcomes in the Age of Mass Migration," American Economic Review 102(5): 1832-1856.

Alesina, Alberto, Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance," Journal of Economic Literature 43(3): 762-800.

Akerlof, George A., Rachel E. Kranton. 2000. "Economics and Identity," Quarterly Journal of Economics 115: 715-753.

Banerjee, A. V., Duflo, E. 2005. "Growth Theory through the Lens of Development Economics," In Philippe Aghion, Steven Durlauf (eds.) Handbook of Economic Growth, 1: 473-552.

Battu, Harminder, McDonald Mwale, Yves Zenou. 2007. "Oppositional Identities and the Labor Market," Journal of Population Economics 20(3): 643-667.

Battu, Harminder, Yves Zenou. 2010. "Oppositional Identities and Employment for Ethnic Minorities: Evidence from England," The Economic Journal 120: F52-F71.

Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, Yves Zenou, Andrea Ichino, Etienne Wasmer. 2011. "Ethnic Identity and Labour Market Outcomes of Immigrants in Europe," Economic Policy 26(65): 59-92.

Borjas, George. 1985. "Assimilation, Changes in Cohort Quality, and Earnings of Immigrants," Journal of Labor Economics 3(4): 463-489.

Bryan, Gharad, Melanie Morten. 2019. "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia" Journal of Political Economy 127(5): 2229-2268.

Casey, Teresa, Christian Dustmann. 2010. "Immigrants' Identity, Economic Outcomes and the Transmission of Identity across Generations," The Economic Journal 120: F31-F51.

Chiswick, Barry R. 1978. "The Effect of Americanization on the Earnings of Foreign-born Men," Journal of Political Economy 86(5): 897-922.

Constant, Amelie F., Klaus F. Zimmermann. 2008. "Measuring Ethnic Identity and Its Impact on Economic Behavior," Journal of European Economic Association 6: 424-433.

Constant, Amelie F., Klaus F. Zimmermann. 2009. "Work and Money: Payoffs by Ethnic Identity and Gender," Research in Labor Economics 29: 3-30.

Constant, Amelie F., Klaus F. Zimmermann. 2011. “Migration Ethnicity and Economic Integration,” In: Jovanovic M. N. (eds.) International Handbook of Economic Integration. Edward Elgar Publishing, Cheltenham, pp 145-168.

Constant, Amelie F., Liliya Gataullina, Klaus F. Zimmermann. 2009. “Ethnosizing Immigrants,” Journal of Economic Behavior & Organization 69: 274-287.

Delaporte, Isaure. 2019. “Ethnic Identity and the Employment Outcomes of Immigrants: Evidence from France,” CESifo Working Paper No. 7651. https://ssrn.com/abstract=3422232

Edin, Per-Anders, Peter Fredriksson, Olof Åslund. 2003. “Ethnic Enclaves and the Economic Success of Immigrants — Evidence from a Natural Experiment,” The Quarterly Journal of Economics 118(1): 329-357.

Falck, Oliver, Stephan Heblich, Alfred Lameli, Jens Sudekum. 2012. “Dialects, Cultural Identity, and Economic Exchange,” Journal of Urban Economics 72:225-239.

Fouka, Vasiliki. 2020. “Backlash: The Unintended Effects of Language Prohibition in US Schools after World War I,” The Review of Economic Studies 87(1): 204-239.

Ginsburgh, Victor, and Shlomo Weber. 2020. “The Economics of Language,” Journal of Economic Literature 58(2): 348-404.

Gorinas, Cedric. 2014. “Ethnic Identity, Majority Norms, and the Native–Immigrant Employment Gap,” Journal of Population Economics 27:225-250.

Herrmann-Pillath, Carsten, Alexander Libman, Xiaofan Yu. 2014. “Economic Integration in China: Politics and Culture,” Journal of Comparative Economics 42(2): 470-492.

Islam, Asadul, Paul Raschky. 2015. “Genetic Distance, Immigrants’ Identity, and Labor Market Outcomes,” Journal of Population Economics 28(3): 845-868.

Konya, Istvan. 2007. “Optimal Immigration and Cultural Assimilation,” Journal of Labor Economics 25(2): 367-391.

Liu, Yuyun, Xianxiang Xu, and Zekai Xiao. 2015. “The Pattern of Labor Cross-dialects Migration” Economic Research Journal (in Chinese) 50(10): 134-146.

Lubotsky, Darren. 2007. “Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings,” Journal of Political Economy 115(5): 820-867.

Nekby, Lena, Magnus Rödin. 2010. “Acculturation Identity and Employment among Second and Middle Generation Immigrants,” Journal of Economic Psychology 31(1): 35-50.

Pendakur, Krishna, Ravi Pendakur. 2005. “Ethnic Identity and the Labour Market,” working paper. http://www.sfu.ca/~pendakur/pendakur_and_pendakur_ethnic_identity.pdf

Sequeira, Sandra, Nathan Nunn, and Nancy Qian. 2020. “Immigrants and the Making of America,” Review of Economic Studies 87(1): 382-419.

Spolaore, Enrico, and Romain Wacziarg. 2009. "The Diffusion of Development," Quarterly Journal of Economics 124(2): 469-529.

Suedekum, Jens. 2018. “Economic Effects of Differences in Dialect,” IZA World of Labor 414 doi: 10.15185/izawol.414

Talhelm, T., X. Zhang, S. Oishi, C. Shimin, D. Duan, X. Lan, S. Kitayama. 2014. “Large-scale Psychological Differences within China Explained by Rice versus Wheat Agriculture,” Science 344 (6184): 603-608.

Tombe, Trevor, Xiaodong Zhu. 2019. “Trade, Migration, and Productivity: A Quantitative Analysis of China,” American Economic Review 109(5): 1843-1872.

Verdier, Thierry, Yves Zenou. 2017. “The Role of Social Networks in Cultural Assimilation,” Journal of Urban Economics 97: 15-39.

Xu, Baohua, Miyata Ichiro. 1999. Chinese Dialect Dictionary. Zhonghua shuju, Beijing.

Appendix

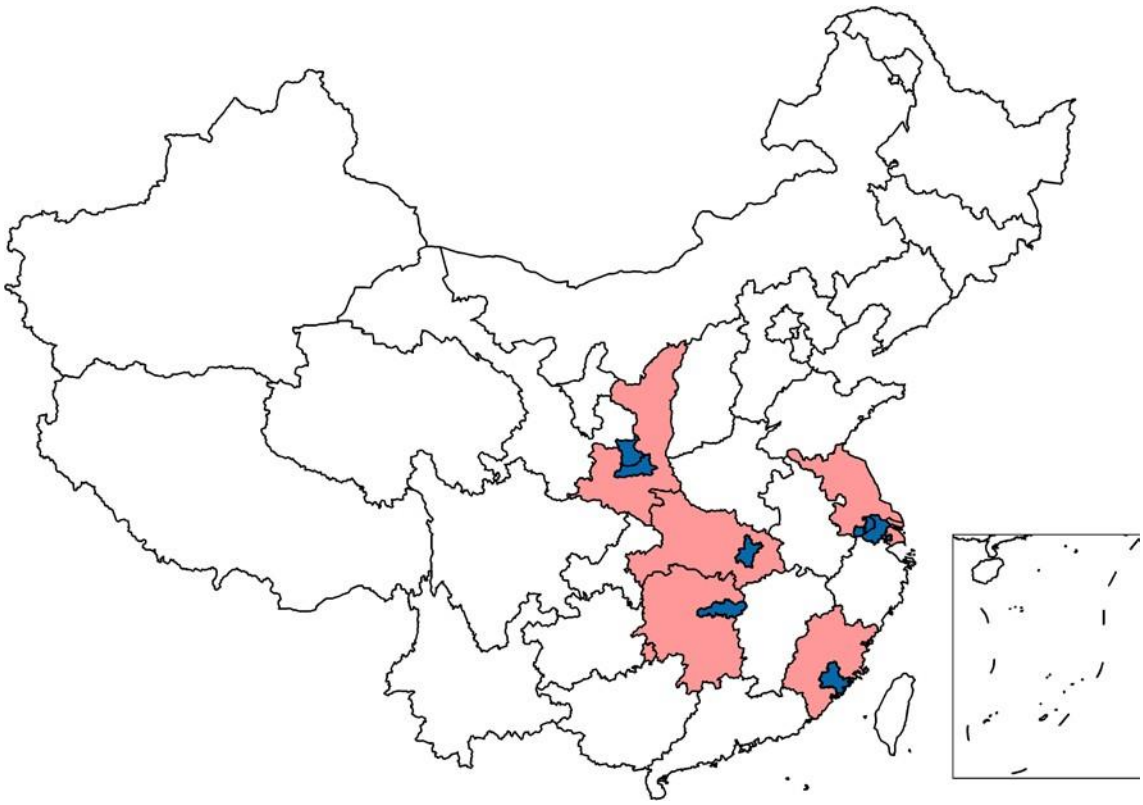

Figure 1: Sampled prefectures

Notes: The figure illustrates location of provinces (in pink) and prefectures (in blue) in the sample. These include Songjiang district in Shanghai, Suzhou and Wuxi prefectures in Jiangsu Province, Quanzhou prefecture in Fujian Province, Wuhan prefecture in Hubei Province, Changsha prefecture in Hunan Province, and Xi'an and Xianyang prefectures from Shan'xi Province.

Source: Own construction.

Table 1: Sample characteristics by social assimilation

	Feel belongs to the local citizen?		Difference	
	Yes	No	(1)-(2)	*p*- value
	(1)	(2)	(3)	(4)
Labor market outcomes				
Monthly income	3195.45	3016.73	178.72	0.000
Hourly wage	14.68	13.09	1.59	0.000
Average days worked per week	5.98	6.08	-0.09	0.000
Average hours worked per day	9.00	9.34	-0.34	0.000
Average hours worked per week	54.19	57.09	-2.91	0.000
Overwork (days per week > 5)	0.74	0.80	-0.06	0.000
Overwork (hours per day > 8)	0.42	0.53	-0.11	0.000
Overwork (hours per week > 40)	0.77	0.84	-0.07	0.000
Demographic characteristics				
Age	33.19	32.63	0.56	0.002
Male (yes=1)	0.55	0.55	-0.01	0.612
Never married (yes=1)	0.24	0.32	-0.07	0.000
Married first time (yes=1)	0.74	0.66	0.08	0.000
Married second or more times (yes=1)	0.01	0.01	0.00	0.613
Divorced (yes=1)	0.01	0.01	0.00	0.016
Widowed (yes=1)	0.00	0.00	0.00	0.612
Education level below middle school	0.09	0.13	-0.04	0.000
Education level of middle school	0.55	0.61	-0.06	0.000
Education level of high school	0.19	0.16	0.02	0.003
Education level of college	0.17	0.10	0.07	0.000
Education level above college	0.00	0.00	0.00	0.040

Notes: The number of observations is 9,790, and 45% of the respondents feel they belong to the local citizens. Column (1) reports the sample mean of migrants that feel they belong to the local citizens, whereas Column (2) describes the sample mean of migrants that do not feel they belong to the local citizens. Column (3) reports the difference in means between the two groups. The last column reports the p-value on testing the hypothesis that the difference is equal to 0.

Table 2: OLS estimates of association between assimilation and labor market outcomes

			work time			overwork		
	monthly income	hourly wage	aver. days worked per week	aver. hours worked per day	aver. hours worked per week	overwork (days per week>5)	overwork (hours per day>8)	overwork (hours per week>40)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Feel local member (yes=1)	59.20	0.66***	-0.06***	-0.16***	-1.48***	-0.04***	-0.06***	-0.04***
	(40.80)	(0.23)	(0.02)	(0.05)	(0.39)	(0.01)	(0.01)	(0.01)
Mean of outcome	3097.18	13.8	6.04	9.19	55.78	0.77	0.48	0.81
Observations	9,761	9,761	9,790	9,790	9,790	9,790	9,790	9,790
Control VARs	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Original province FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: The table reports the results from OLS regressions as specified in equation (1) in the text. The other control variables include age, age squared, dummy of male, marital status (including dummies of married first time, married second or more times, divorced, widowed), education categories (including dummies of education level of middle school, education level of high school, education level of college, and education level above college). The standard errors in parentheses are clustered by community. *** p<0.01, ** p<0.05, * p<0.1.

Table 3. IV estimation on the impact of assimilation on labor market outcomes

	feel local member (yes=1)	Monthly income	hourly wage	work time			overwork		
				aver. days worked per week	aver. hours worked per day	aver. hours worked per week	overwork (days per week>5)	overwork (hours per day>8)	overwork (hours per week>40)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Feel local member (yes=1)		39.58	3.25	-0.23	-1.17***	-9.21**	-0.22**	-0.44***	-0.26**
		(400.30)	(2.09)	(0.21)	(0.44)	(3.76)	(0.11)	(0.15)	(0.11)
Linguistic distance (0-1-2-3)	-0.10***								
	(0.01)								
KP *F* -statistic	52.31								
Control VARs	YES	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Original province FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	9,780	9,751	9,751	9,780	9,780	9,780	9,780	9,780	9,780
Panel B									
Feel local member (yes=1)		-31.37	4.29	-0.29	-1.54**	-12.01*	-0.31*	-0.58**	-0.37**
		(618.83)	(3.31)	(0.34)	(0.74)	(6.28)	(0.18)	(0.25)	(0.18)
Can speak local language (yes=1)	0.15***	28.33	-0.55	0.04	0.22	1.77	0.04	0.07	0.06*
	(0.02)	(106.75)	(0.58)	(0.06)	(0.14)	(1.14)	(0.03)	(0.05)	(0.03)
Can understand local language (yes=1)	0.11***	23.40	-0.12	-0.02	0.01	-0.17	0.02	0.02	0.01
	(0.02)	(81.64)	(0.46)	(0.05)	(0.11)	(0.93)	(0.03)	(0.04)	(0.03)
Linguistic distance (0-1-2-3)	-0.06***								
	(0.01)								
KP *F* -statistic	21.08								
Control VARs	YES	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Original province FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	9,780	9,751	9,751	9,780	9,780	9,780	9,780	9,780	9,780
Mean of outcome	0.45	3097.18	13.8	6.04	9.19	55.78	0.77	0.48	0.81

Notes: The table reports the results from IV regressions as specified in equations (1) and (2) in text. The other control variables are the same as those in Table 2. Panel B further controls for dummies indicating whether the migrant can speak the local language, and can understand the local language. "KP F-statistic" denotes the cluster-robust Kleibergen-Paap (KP) F-statistic on testing weak instruments. The standard errors in parentheses are clustered by community. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: IV estimation on the influence of assimilation on networks and neighborhood choice

	Networks			Neighborhood		
Outcome variables	interact with ethnic people (yes=1)	interact with local people (yes=1)	member of ethnic organization (yes=1)	Neighbors are mostly local citizens (yes=1)	Neighbors are mostly non-local citizens (yes=1)	The number of local and non-local neighbors is similar (yes=1)
	(1)	(2)	(3)	(4)	(5)	(6)
Feel local member (yes=1)	-0.12*	0.52***	-0.08	0.21*	-0.37***	0.16
	(0.06)	(0.14)	(0.12)	(0.12)	(0.13)	(0.13)
Mean of outcome	0.93	0.42	0.12	0.23	0.49	0.28
Observations	9,780	9,780	9,780	9,405	9,405	9,405
Control VARs	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES
Original Province FE	YES	YES	YES	YES	YES	YES

Notes: The table reports the IV estimates on migrants' network and neighborhood choice, using linguistic distance between original and destination provinces as an instrumental variable for the sense of belonging to the local citizens. The other control variables include age, age squared, dummy of male, marital status (including dummies of married first time, married second or more times, divorced, widowed), education categories (including dummies of education level of middle school, education level of high school, education level of college, and education level above college). The standard errors in parentheses are clustered by community. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: IV estimation on the influence of assimilation on job searching

Outcome variables	find the job through family/relatives, or friends/classmates (yes=1)	find the job on their own, or start a business on their own (yes=1)	find the job through local people (yes=1)	find the job through government, social agency, internet, job fair, and others (yes=1)
	(1)	(2)	(3)	(4)
Feel local member (yes=1)	-0.44***	0.21*	0.17**	0.06
	(0.15)	(0.12)	(0.07)	(0.10)
Mean of outcome	0.47	0.31	0.06	0.16
Observations	9,774	9,774	9,774	9,774
Control VARs	YES	YES	YES	YES
County FE	YES	YES	YES	YES
Original Province FE	YES	YES	YES	YES

Notes: The table reports the IV estimates on channels of migrants' job searching, using linguistic distance between original and destination provinces as an instrumental variable for the sense of belonging to the local citizens. The other control variables include age, age squared, dummy of male, marital status (including dummies of married first time, married second or more times, divorced, widowed), education categories (including dummies of education level of middle school, education level of high school, education level of college, and education level above college). The standard errors in parentheses are clustered by community. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: IV estimation on the impact of assimilation on labor market outcomes conditional on job characteristics

	feel local member (yes=1)	monthly income	hourly wage	work time: aver. days worked per week	work time: aver. hours worked per day	work time: aver. hours worked per week	overwork: overwork (days per week>5)	overwork: overwork (hours per day>8)	overwork: overwork (hours per week>40)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Feel local member (yes=1)		311.04	4.07**	-0.19	-0.98**	-7.74**	-0.19*	-0.36**	-0.22**
		(381.43)	(2.04)	(0.21)	(0.44)	(3.80)	(0.11)	(0.15)	(0.11)
Linguistic distance (0-1-2-3)	-0.10***								
	(0.01)								
KP *F* -statistic	50.26								
Mean of outcome	0.45	3097.18	13.8	6.04	9.19	55.78	0.77	0.48	0.81
Observations	9,780	9,751	9,751	9,780	9,780	9,780	9,780	9,780	9,780
Control VARs	YES	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Original province FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Occupation, Industry, Unit type	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: The table reports the results from IV regressions as specified in equations (1) and (2) in the text. The other control variables include age, age squared, dummy of male, marital status (including dummies of married first time, married second or more times, divorced, widowed), education categories (including dummies of education level of middle school, education level of high school, education level of college, and education level above college), and dummies indicating types of occupation, industry, and work unit. "KP F-statistic" denotes the cluster-robust Kleibergen-Paap (KP) F-statistic on testing weak instruments. The standard errors in parentheses are clustered by community. *** p<0.01, ** p<0.05, * p<0.1.

Table 7. IV estimation on the influence of assimilation on migrant's attitude and perception

Outcome variable (1=absolutely disagree, 2=disagree, 3=agree, 4=totally agree)	willingness to be neighbors of local citizens	willingness to be friends of local citizens	willingness to be involved in community/units and become a member of them	feeling that themselves belongs to the city currently lived	willingness for themselves or their relatives to marry local citizens	feeling that local citizens willing to accept themselves to be one of them	feeling that local citizens willing to be neighbors of themselves	feeling that local citizens like and respect non-local citizens
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Feel local member (yes=1)	0.39**	0.21	0.23	0.67***	0.68***	0.46***	-0.05	-0.11
	(0.15)	(0.15)	(0.15)	(0.19)	(0.21)	(0.18)	(0.20)	(0.20)
Mean of outcome	3.52	3.54	3.49	3.27	3.24	3.25	2.97	2.96
Observations	9,780	9,780	9,780	9,780	9,780	9,780	9,780	9,780
Control VARs	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: The table reports the IV estimates on migrants' attitudes towards local citizens and their perceptions of local citizens' attitude, using linguistic distance between original and destination provinces as an instrumental variable for the sense of belonging to the local citizens. The other control variables include age, age squared, dummy of male, marital status (including dummies of married first time, married second or more times, divorced, widowed), education categories (including dummies of education level of middle school, education level of high school, education level of college, and education level above college). The standard errors in parentheses are clustered by community. *** p<0.01, ** p<0.05, * p<0.1.

Table 8. IV estimation on the heterogeneous impact of assimilation on labor market outcomes

	monthly income	hourly wage	aver. days worked per week	worktime: aver. hours worked per day	aver. hours worked per week	overwork (days per week>5)	overwork: overwork (hours per day>8)	overwork (hours per week>40)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Feel local member (α_1)	-551.79	0.11	-0.17	-1.57***	-11.39**	-0.08	-0.58***	-0.17
	(444.60)	(2.29)	(0.23)	(0.56)	(4.44)	(0.12)	(0.18)	(0.12)
Male×Feel local member (α_2)	1,029.07**	5.46*	-0.12	0.69	3.81	-0.25*	0.25	-0.16
	(516.79)	(2.83)	(0.25)	(0.60)	(4.95)	(0.14)	(0.20)	(0.13)
$\alpha_1+\alpha_2$	477.28	5.58**	-0.28	-0.88*	-7.58*	-0.33**	-0.33*	-0.33**
p -value of $\alpha_1+\alpha_2=0$	0.340	0.039	0.269	0.086	0.094	0.013	0.063	0.013
KP *F* -statistic	26.27	26.27	25.96	25.96	25.96	25.96	25.96	25.96
Panel B								
Feel local member (α_1)	83.19	3.61*	-0.24	-1.20***	-9.55***	-0.22**	-0.48***	-0.26**
	(382.05)	(1.98)	(0.20)	(0.42)	(3.57)	(0.10)	(0.14)	(0.10)
Young×Feel local member (α_2)	-78.79	-0.65	0.02	0.04	0.62	0.00	0.08	-0.01
	(208.25)	(1.22)	(0.09)	(0.22)	(1.77)	(0.05)	(0.07)	(0.05)
$\alpha_1+\alpha_2$	4.401	2.964	-0.225	-1.15**	-8.93**	-0.22*	-0.41**	-0.27**
p -value of $\alpha_1+\alpha_2=0$	0.992	0.198	0.329	0.015	0.029	0.053	0.012	0.022
KP *F* -statistic	26.73	26.73	26.43	26.43	26.43	26.43	26.43	26.43
Panel C								
Feel local member (α_1)	1,386.13**	12.10***	-0.31	-1.81***	-13.64***	-0.40***	-0.56***	-0.44***
	(617.63)	(3.74)	(0.23)	(0.54)	(4.50)	(0.13)	(0.18)	(0.13)
High-Edu×Feel local member (α_2)	-5,996.01***	-39.40***	0.35	2.83***	19.77**	0.79**	0.54*	0.78**
	(1,881.61)	(11.43)	(0.38)	(0.98)	(7.73)	(0.31)	(0.29)	(0.32)
$\alpha_1+\alpha_2$	-4609.88***	-27.30***	0.04	1.02	6.13	0.39	-0.02	0.34
p -value of $\alpha_1+\alpha_2=0$	0.005	0.006	0.918	0.251	0.375	0.156	0.941	0.229
KP *F* -statistic	21.64	21.64	21.52	21.52	21.52	21.52	21.52	21.52

Notes: The table reports the results from IV regressions as specified in equation (3) in text. The variable male is dummy coded and equal to 1 if migrants are male and 0 otherwise. Similarly, the migrants are defined as young if they are 30 years old or younger than 30 years old, and are highly educated if they have completed college education or above. All the regressions also include the same control variables as those in Table 2, as well as destination-county fixed effects, and original-province fixed effects. "KP F-statistic" denotes the cluster-robust Kleibergen-Paap (KP) F-statistic on testing weak instruments. The standard errors in parentheses are clustered by community. *** p<0.01, ** p<0.05, * p<0.1.