

Difference and Integration: Population Mobility in Ethnic Minority Areas of China

Runren Zhou

Abstract

Using the 20% subsample from 2005 National 1% Population Sampling Survey, I study the factors affecting the population movement of ethnic minorities in China with binary probit model and instrument variable method. The results show that living in compact communities of ethnic minorities would reduce the willingness of individuals in ethnic groups to move to other prefectures. Possible explanations are: 1.the living pattern increases the closedness of ethnic groups; 2.the living pattern reflects the incomplete integration of ethnic regions by Central Plain Dynasty in history, an effect that still exists today. My study also finds that lower proficiency in Mandarin leads to less population movement, and ethnic compact communities would influence local language in education and religious forces, thus affecting people's choice of migration.

Keyword: ethnic minority · population movement · ethnic amalgamation

1 Introduction

Large-scale and frequent population movement has become one of the basic facts of China's economy and society. According to the seventh census data in 2021, China's current floating population has reached 376 million, accounting for more than 40% of the urban population, which means that the traditional foundations of Chinese society, known as "from the soil", has been transformed into "seasonal migration" (Duan, 2021). In the context of the growing aging of population and the decline in natural growth rate, the role of population mobility in the spatial population distribution pattern has become more and more important. However, although both scales of the floating population of ethnic minorities and *Han* people keep increasing, there has always been a noticeable difference between their population mobility. As shown in Figures 1 and 2, the mobility of ethnic minorities remains well below that of *Han* people; the percentage of ethnic minorities in the floating population enjoys sustained growth overall, but it is still lower than the percentage of ethnic minorities in the total national population in 2015.

[Figure 1: Percentage of floating population in the total *Han* people and ethnic minorities]

[Figure 2: Scale and percentage of ethnic minority in floating population]

The *Han* population accounts for more than 90% of China's population. Their predecessor, the *Huaxia* nationality, mainly settled in the middle and lower reaches of the Yellow River and mainly lived on farming. Today the *Han* Population are widely distributed in the Northeast China Plain, the North China Plain, the Middle and Lower Yangtze Plain, the Sichuan Basin, and the Lingnan region. Since Qin Dynasty, about two thousand years ago, these areas have been ruled by Central Plains Dynasty of all generations, and have become the political and economic basics of the unified dynasties. Smallholder farming, long-term political integration, and similar historical and cultural backgrounds have resulted in the formation of relatively close connections and fierce sense of identity within the traditional *Han* Chinese areas, which provide favorable conditions for the current population movement. The definition of ethnic minorities can be traced back to the identification of ethnic minorities after the founding of the People's Republic of China. Their main settlements are located in eight provinces (Xinjiang, Inner Mongolia, Qinghai, Ningxia, Tibet, Guizhou, Yunnan, and Guangxi). These regions have also been under the rule of Central Plains Dynasty in history, but most of the time, especially for the first five ethnic autonomous regions, they were entrusted to local elites so that the duration and degree of their integration with the Central Plain region was limited in terms of language, customs, and level of education. The current relatively low population mobility of ethnic minorities can be partially explained by those considerable differences. For individuals, going out as a migrant worker is an important way to increase income and improve their lives. Therefore, studying the status quo and factors of ethnic minority population mobility can be conducive to narrowing the gap of income and quality of life between ethnic groups and ensuring that people of all ethnic groups across the country can share the fruits of development jointly.

This paper studies the factors that affect the population movement of ethnic minorities from the perspective of outflow areas, in order to better understand their decision-making behaviors. The study uses a 20% sub-sample of micro data from the 2005 National 1% Population Sampling Survey and adopts the binary probit model and instrumental variable methods. The results show that the life pattern of ethnic minorities that living in compact communities will reduce the willingness of them to leave their hometowns for other places to work or live for the long term. Their migration will also be hindered without the proficiency in Mandarin. And the ethnic compact communities will interact with the local language education and religious power, thereby having an impact on people's choices.

The remaining parts of this paper are arranged as follows: the second part is a review of the relevant literature; the third part introduces the data used in the research, constructs an empirical model, analyzes the results and conducts a robustness check; the fourth part explores possible influence mechanisms and conducts heterogeneity analysis; the fifth part concludes the paper.

2 Literature Review

The term of population movement in this paper refers to the phenomenon of “separation of people's real residence and their registered permanent residence (known as *Hukou*)” under China's household registration system. In fact, such separation can cover two kinds

of population movement: short-term, repetitive or periodic population movement between regions; or the migration from the former place of residence to a new one, which is always a long-term or even permanent process. However, due to the restrictions of the household registration system, even if many people have changed their place of residence for a long time or permanently, their *Hukou* remains in the places where they come from, which makes it difficult to determine whether to define the change of their residence as “long-term population movement” or “population migration”. Therefore, this paper directly adopts the definition of “separation of people’s real residence and their registered permanent residence (*Hukou*)”. Since many phenomena and underlying causes of population movement and population migration are highly similar, such a simplified definition will not have a significant impact on the results of this paper.

The research on the theory of population migration originated in the paper “The Laws of Migration”. That paper for the first time points out that distance, gender, destination transportation convenience, migration motivation, etc. are all factors that affect population migration, and the migration population presents a progressive characteristic (EG Ravenstein, 1885). The following classic push-pull theory of population migration believes that: on the one hand, there are insufficient employment opportunities, poor infrastructure, low average income and other factors in the outflow area to promote population moving out; better infrastructure and a more developed economy can stimulate population inflow (R. Herberle, 1938). E. S. Lee made a more systematic summary of this theory in his article “Migration Theory” (E. S. Lee, 1966). The theory of population migration for developing countries has been enriched by researches such as the dual structure model (W. A. Lewis, 1954) and the Todaro model (Michacl P. Todaro, 1970).

The empirical research on population movement mainly focuses on the factors that affect population mobility, and the spatial pattern and trend of the distribution of floating population. Personal factors such as gender, years of education, marital status, occupational characteristics, etc. (Barbieri, 2005; Y. Wu, 2013; X. Wang, 2015), political factors such as household registration system and taxation system (Alonsoc, 1980; Cohen, 2004), economic factors such as income level, employment opportunities, and industrial structure (Zhao, 1998; L. Li, 2001; Q. Li, 2003; A. Liu, 2017; Y. Wang et al., 2017) and social factors such as social public services and population size (Zhao, 1998; Gedik, 2010) will jointly constitute a comprehensive factor affecting population mobility (C. Duan et al., 2019). From the perspective of spatial distribution pattern and flow trends, the Beijing-Tianjin-Hebei region, Yangtze River Delta and Pearl River Delta are still the main gathering places of migrants in China (T. Yu, 2012). Also, there are some other trends like inland-toward movement (G. Cao et al., 2011) and the migrant population spreading from center cities to surrounding cities in Yangtze River Delta and Pearl River Delta (T. Liu et al., 2015).

There are relatively few studies on the migration of ethnic minority populations in China. Most existing studies concentrate on the perspective of sociology, with case analysis and descriptive statistics as the main methods. Relevant literature points out that ethnic minorities generally tend to move to places where they can join a larger community of the same ethnicity in order to enjoy the convenience brought by cultural belonging and language customs (R. Ma, 2004); ethnic minorities leave the place of birth less frequently, the inter-provincial migration rate is much lower than that of the *Han* people, and the mi-

gration rate of ethnic minorities living in cities is higher than that in the countryside. At the same time, population movement of ethnic minorities is more affected by social factors than economic factors (S. Zhang et al., 2005). The characteristics of the floating population of ethnic minorities summarized in the existing literature include: most of them are from ethnic autonomous areas and rural areas, more women than men, years of education and labor contract signing rate are lower than the national average (C. Duan, 2011); their jobs often have obvious ethnic features and intermarriage between ethnic minorities and *Han* people conduces to the increase of their eastward migration (X. Zheng et al., 2002).

Some scholars have explored specific ethnic groups and ethnic ghettos, and find that the rate of population migration in Inner Mongolia is lower than that of the whole country, and the migration rate of *Han* population in Inner Mongolia is higher than that of Mongolians especially those ones who are peasants or shepherds (C. Duan et al., 2017); research on population mobility in Kashgar, Xinjiang shows that distance, language, customs, education level, and cultural adaptation restrict Uyghur labor from going out to work (R. Ma, 2007).

In summary, the existing literature on the population movement of ethnic minority has a large supplementary space in terms of data utilization, measurement methods, and mechanism exploration. In this paper, individual micro data, macroeconomic data and historical data are used comprehensively, and the econometric model is used for causality inference and mechanism testing, which fills the gaps in the relevant literature.

3 Empirical Strategies

3.1 Data and Descriptive Statistics

The micro individual-level data used in this paper is a 20% sub-sample of the 2005 national 1% population sampling survey data, with a total of about 2.58 million observations. There are about 1.55 million observations after keeping labor force between 16 to 60 years old and not in school. This sub-sample is a good representative of the overall distribution of various ethnic groups in China in 2005. As shown in Table 1, among the eighteen ethnic groups which make up more than 99% of the population, the difference between the real population percentage and the population percentage of the sub-sample is greater than 0.1% only for *Han* and *Zhuang* people and less than 0.1% for others.

[Table 1: Percentage of the population of all ethnic groups in the total population]

The three core variables in the model are *migration*, the dependent variable, *minority* and “percentage of ethnic minorities in the total population of a region” which is denoted as *minprop*. The data sources and construction methods of them are introduced in turn.

The definition of “migration” is further refined on the basis of “separation of people’s real residence and registered permanent residence (*Hukou*)”. Specifically, if the observation meets the following two requirements, the dependent variable ”migration” variable is 1, otherwise it is 0:

1. The registered permanent residence of the individual is different from the place of residence at the time of the survey at the prefecture-level city level.

2. The reason for such difference is not study, training or business trip.

The definition of “minority” is based on question R5 of the survey. If the individual is a ethnic minority, and her/his *Hukou* is in the ethnic enclave of that ethnicity, then the binary variable that defines the minority identity of “minority” is 1, otherwise it is 0. This is because of the diversity of ethnic minorities in China, and their distribution locations, cultural customs, etc., are quite different, and cannot be generalized. For example, when a Mongolian citizen lives in the Inner Mongolia Autonomous Region, it is more likely to be affected by the ethnic compact communities. At this time, this variable is set to 1. On the contrary, the ethnic identity of a *Dai* citizen living in Inner Mongolia obviously cannot match the ethnic identity, history and culture of the local ethnic minority groups, and the variable is 0 at this time. Whether a province is regarded as a minority ethnic enclave is determined by the distribution of ethnic minorities in “China’s Ethnic Statistical Yearbook”. Under this definition, the percentage of floating population in the minority population is 5.83%, which is lower than the national average rate, 8.40%.

The total population and the number of ethnic minorities at the prefecture-level city level come from the “China Population Statistics Yearbook” and “Chinese Ethnic Population Data”. In regressions, this paper uses the percentage of ethnic minorities in each city in 1990 instead of 2005. This is because the scale of population flow in 2005 has been quite large, and the percentage of ethnic minorities in the region may have changed due to population movement, forming a reverse cause and effect. According to Figure 3 and 4, except for some prefecture-level cities in southern Tibet, western Yunnan, and northeastern China, the percentage of ethnic minority populations in most regions did not change much from 1990 to 2005; if the percentage of ethnic minority population of a prefecture-level city is high, the population mobility of ethnic minority in the area will be relatively low.

[Figure 3: Percentage of ethnic minorities in the total population by prefecture-level regions, 1990 and 2005]

[Figure 4: Percentage of floating ethnic minority population in the total ethnic minority population by prefecture-level regions, 2005]

The economic data as the control variables comes from the CEIC China Economic Database and the statistical yearbooks of cities, and the elevation data comes from ASTER Global Digital Elevation Model V003.

3.2 Econometric Model

In order to show that ethnic minority individuals living in ethnic minority areas will be less inclined to migrate, this section constructs the following binary probit model:

$$y_{ijk} = \beta_0 + \beta_1 minority_i \times minprop_j + \beta_2 minority_i + \beta_3 minprop_j + \mathbf{x}'_{1i} \gamma_1 + \mathbf{x}'_{2j} \gamma_2 + \lambda_k + u_{ijk}$$

[Table 2: Meaning of notations in econometric model]

[Table 3: Description statistics of variables]

3.3 Instrumental Variable

The benchmark model may face potential endogenous problems. The main variable concerned in this paper is the interaction item of minority_i and minprop_j . The minority identity represented by minority_i is random to the individual, and a series of regional control variables that may be related to ethnic identity are added to the regression equation, so this variable is likely to be exogenous. In order to alleviate the endogeneity of minprop_j , which is the percentage of the minority population in the region, this paper uses the time of duration of the establishment of prefectures and counties in the history of region j as an instrumental variable. The time span is from the end of the Warring States Period to the fall of the Qing Dynasty in 1911. Historically, the easternmost regions where the Central Plains Dynasty implemented effective county system management was Liaodong Peninsula, the northmost region was the Great Wall and the northwesternmost regions were Qilian Mountains. The integration of the Yunnan-Guizhou region in the southwest was not basically completed until the complete of “Bureaucratization of Native Officers” in mid-Qing Dynasty. In the Qing Dynasty, Xinjiang and Inner Mongolia, which had a relatively high percentage of ethnic minority population, were governed by Ambans, generals, or League & Banners. It can be seen that the time of duration of the establishment of prefectures and counties is obviously related to the percentage of ethnic minority population.

The specific method of constructing instrumental variables is as follows: using the historical coordinates of prefectures and counties given by the Chinese Historical Geographic Information System (CHGIS) v6, lay these points out on a map of China divided by prefecture-level regions. When a certain point falls within the administrative scope of a prefecture-level area, it is considered that the prefecture-level area has been affected by the historical counties represented by this point. For example, the administrative centre of Zhongshan Prefecture established between 174 and 231 AD was located in the present-day Shijiazhuang city, then it is believed that Shijiazhuang City established prefectures and counties during this time. If there are two or more administrative centres of prefectures and counties in a current prefecture-level area, and their setting time overlaps, the time of duration is calculated according to the maximum interval composed of their setting time. For example, there are two administrative centres of Langmeng County established in 23-264 AD and Yangqu County established in 215-528 AD located in Taiyuan City, then the time of duration provided by these two counties for setting up counties is 23-528 AD. Since the endogenous variables concerned in this paper are in the form of interaction terms, the regression needs to use $\text{minority}_i \times \text{time of duration of setting up counties and prefectures}_j$ as the instrumental variable of $\text{minority}_i \times \text{minprop}_j$.

3.4 Results of Regression

The results of the baseline regression and the results with instrumental variables are presented in Table 4. Column 1 includes only the dummy variable of whether the individual is a minority; column 2 includes the regional minority percentage and its interaction term with the minority dummy variable; column 3 is the linear probability model regression with instrumental variables, and column 4 is the probit regression with instrumental variables.

[Table 4: Baseline regression with instrumental variables regression]

The results in column 1 show that minority status significantly decreases individuals' willingness to leave their hometowns. This result is not stable in columns 2, 3, and 4, but this paper is mainly concerned with the coefficients of the interaction terms, and whether or not one is a minority is used as a control variable for individuals. The coefficients of the interaction terms in columns 2, 3, and 4 are all significantly negative, indicating that for an individual who is a ethnic minority in his or her own ethnic enclave, the higher the percentage of the local minority population, the lower the probability of leaving his or her domicile. The results of the linear probability model in column 3 illustrate that if the percentage of local minority population increases by 10%, it will decrease the probability of minority's population movement in the area by 0.93% relative to *Han* Chinese.

Since columns 2 and 4 use a nonlinear model, the coefficients of the interaction terms do not directly reflect the interaction effects, and the marginal effects cannot be found directly on the interaction terms using existing statistical software. In order to show the interaction effect and its significance more intuitively, this paper uses the stata package developed by Ai & Norton (2003) to find the second-order mixed partial derivatives about minority_i and minprop_j for each observation to calculate the interaction effect, and selects six provinces with dense concentrations of ethnic minorities, namely Xinjiang, Inner Mongolia, Tibet, Ningxia, Yunnan, and Guangxi. The interaction effect is showed by the vertical axis and the z-statistic by the horizontal axis to make a scatter plot. If the horizontal coordinate of the point in the plot is to the left of 0 and the vertical coordinate is below -1.96, the interaction effect of the observation point is significantly negative. Taking an individual minority in Xinjiang with an interaction effect of -0.2 as an example, a 1% increase in the minority population percentage of the city where the individual is located will reduce its probability of making population movement by 0.2% compared to the *Han* Chinese. As shown in Figure 5, except for the positive but insignificant interaction effect in the Tibet Autonomous Region, the majority of observations in the remaining five provincial-level administrative regions are located in areas with significant negative effects, again demonstrating that minority individuals living in minority ethnic areas are less likely to engage in population movement.

[Figure 5: Scatter-plot of interaction effect-z statistic for six provincial administrative regions with dense concentrations of ethnic minorities]

This paper offers two possible explanations for the empirical finding that ethnic minorities living in their own ethnic areas are less likely to engage in population movements.

First, China's ethnic identification scheme and ethnic policy after the foundation of PRC in 1949 have greatly imitated the relevant practices of the former Soviet Union, i.e., the identification was based on Stalin's ethnic theory and each ethnic group was assigned a territory (R. Ma, 2011). Such a policy protects the rights and interests of minority ethnic people. In this context, the geographic clustering of the same ethnic group has two effects: first, it strengthens the sense of community within the ethnic group and the belief in the "specific ethnicity - specific region" correspondence, and reduces the willingness of population movement from the point of view of outflow. Second, it reduces the exchange and integration between different ethnic groups and regions, weakening the ability of the nation,

the largest community, to create a collective identity, deepening stereotypes and even barriers between ethnic groups, and increasing the cost of making ethnic minority people get accustomed to the culture, customs, and life of the population inflow areas.

Second, the indicator of regional minority percentage also reflects the degree of political and economic integration of the region by the historical central government, the effects of which last until now. In ethnic minority regions - often also frontier regions - there is often greater political discretion and less economic ties to the Central Plains, so for local ethnic minority people, entering the traditional *Han* Chinese areas for work lacks historical tradition and implies more unknowns and risks.

3.5 Robustness Check

In this paper, the baseline regression makes use of all samples with household locations in the central and western provinces, but since the urban and rural populations do not conduct population movement in the same way and for the same purpose, which may confound the results. Therefore, samples from villages and towns in the central and western provinces and samples from villages in the central and western provinces are selected separately in this section for testing. As shown in Tables 5 and 6, the absolute values of the coefficients of the interaction terms of the minority dummy variables and the regional minority population percentages are slightly reduced, but both are still negatively significant, proving that the regression results are still robust in the subsamples.

[Table 5: Robustness Check 1: Village and Town Sample in Central and Western Provinces]

[Table 6: Robustness Check 2: Village Sample in Central and Western Provinces]

In addition, population movement is defined in the benchmark regression as the “separation of people’s real residence and *Hukou*” and the reason for leaving the former location is not business trip or training and study. Since people who leave their hometowns for work are the focus of this paper, this section defines population movement only as “separation of people’s real residence and *Hukou*” and the reason for leaving is working. The coefficient of the interaction term in Table 7 is still robustly negative and significant, which verifies the applicability of the above findings for migrant workers.

[Table 7: Robustness Test 3: Migrant Workers]

4 Cultural Factors: Language and Religion

4.1 Use of ethnic languages

China’s ethnic minorities have a rich variety of languages. Except for the *Hui* and *Manchu* who generally use Chinese, all other 53 ethnic minorities have their own languages. The flourishing ethnic language culture is a reflection of China’s efforts to protect the rights

and interests of minority citizens and support cultural diversity, but it has also created communication barriers between people who speak different languages. Data show that there are significant differences in the learning and use of Mandarin and ethnic languages among different ethnic minorities, and such differences will directly affect ethnic minority people's feasibility for population movement.

This section draws the following images based on the Survey on the Use of Chinese Language and Character published by Steering Group Office for Survey of Language Situation in China in 2006. Figure 6 shows that in the western provinces, where ethnic minorities are widely distributed, the percentage of people who can talk to people in Mandarin is low, even below 20% in Tibet. Figure 7 shows that the percentage of people who have difficulty in learning Mandarin is high in Xinjiang, Tibet, Yunnan, and Inner Mongolia, with Xinjiang exceeding 35%. As shown in Figure 8, among the types of difficulties encountered in learning to speak Mandarin, "no one around me speaks it and there are few opportunities to speak it" is a common difficulty. *Manchu* and *Hui* are mostly "influenced by the Chinese dialect and have difficulty in changing their accent", while other ethnic minorities are mostly "influenced by their own language and have difficulty in changing their accent". From Figure 9, it can be seen that a significant percentage of ethnic minorities want to use their own language as the language of instruction in primary and secondary schools. All these findings indicate that the task of popularizing Mandarin in China is still arduous, and ethnic minorities still face a considerable degree of language barriers.

[Figure 6: Percentage of respondents who can speak in Mandarin by provincial-level administrative regions]

[Figure 7: Percentage of respondents who encountered difficulties in learning to speak Mandarin "due to the influence of their own language and the difficulty in changing their accent" by provincial-level administrative regions]

[Figure 8 Percentage of major problems encountered by ethnic minorities in learning to speak Mandarin]

[Figure 9: Percentage of respondents who want local primary and secondary schools to use a particular language as the language of instruction]

To evaluate the impact of language use on the mobility of minority populations, this section constructs an indicator of linguistic ethnicization for each minority, as detailed in Table 8. This is a variable between 0 and 1, with closer to 1 representing a higher degree of linguistic ethnicization and a lower degree of Sinicization. It is calculated as follows

Linguistic ethnicization = $(1 + \text{the percentage of the minority that can communicate in the ethnic language} - \text{the percentage of the minority that can communicate in Mandarin})/2$

[Table 8: Calculation of linguistic ethnicity index]

The baseline regressions of minority_i and $\text{minority}_i \times \text{minprop}_j$ were replaced with $\text{minority}_i \times \text{linguistic ethnicization}$ and $\text{minority}_i \times \text{linguistic ethnicization} \times \text{minprop}_j$, and the results are presented in Table 9. The results in row 1 of the table show that minorities with higher linguistic ethnicization are significantly less likely to migrate, and the results in row 2 indicate that both a higher percentage of minority population in the *Hukou* region and a lower average level of Mandarin use discourage minority population mobility. Similar to the baseline regression, this regression remains robust and significant under the subsample of mid-western villages and towns and mid-western villages.

[Table 9: Effects of Ethnic Language]

In fact, the degree of linguistic ethnicization is affected by ethnic clustering in the long run. This is partly because ethnic compact communities make the use of only the ethnic language sufficient for daily life, and partly because the communities often have schools that teach in the ethnic language in order to preserve ethnic culture. Since 1978, the government has legislated for bilingual education, and the 1982 Constitution emphasized the importance of ethnic minority languages. However, in the actual development process, minority primary and secondary schools and *Han* primary and secondary schools are often separated and closed to each other; independently established minority schools have serious problems of mono-cultural and mono-lingual environment (Wan & Liu, 2012), which create quite a few obstacles for minority people to learn the national common language.

4.2 The role of religion

Religious beliefs are an integral part of spiritual life, and religious organizations have an important impact on real social life. Participation in religious organizations and involvement in religious activities promotes the construction of social networks, reduces the cost of participation in group events (Ruan et al., 2014), and to a considerable extent influences the ways individuals behave. The percentage of formal believers in China's population is not high, but the percentage of those who do not believe in or worship God at all is quite low (Lu et al., 2014), suggesting that religion has become relevant to the lives of most of people of China. In China's ethnic minority regions, religious forces have developed to a considerable scale during the history, with Tibetan Buddhism popular in Tibet, Inner Mongolia, and the western and southern parts of Qinghai, Islam influencing Xinjiang, Ningxia, Gansu, and the eastern and northern parts of Qinghai, and Theravada Buddhism prevalent in Yunnan. There are even some ethnic groups where religion played a decisive role in their formation and development, such as Islam, which has an important influence on the formation of the *Hui* dietary customs and bleeds into their culture. These religious forces play a positive role in enriching people's spiritual life, but they also organize and fix people with religious creeds, groups and rituals, increasing the psychological cost of population mobility for believers. For example, Lanzhou has become a major inflow of Uyghur migrants because its large Muslim community and more mosques have shortened the psychological distance between Uyghurs and Lanzhou city, easing the obstacles to their daily religious life (Zhou, 2015); some scholar also points out that the reason why people move to places with a higher concentration of their ethnic groups is to seek a sense of cultural belonging and language

and custom convenience (R. Ma, 2004).

In order to verify that religion is the channel through which the percentage of regional minority population influences the mobility of minority population, this section uses the number of religious places per capita in each prefecture-level cities as a proxy variable for the strength of local religion, replacing all minprop_j in the baseline regression with *the number of religious places per capita_j*, and the results are presented in Table 10. Column 1 of Table 10 shows that the percentage of minority population has a significant positive effect on the number of religious sites per capita, and every 10% increase in this percentage will increase the number of religious sites per capita by 0.52. Columns 2 and 4 regress the *migration* dummy variable on the number of religious places per capita and the percentage of ethnic minorities in the region, respectively, and find that both have a significant negative effect on the movement of ethnic minorities. Column 3 still uses the time of duration of setting up prefectures and counties as an instrumental variable for the number of religious sites per capita, and obtains similar results as column 2. Column 5 adds both of them to the regression, and the results show that the probability of population movement of ethnic minorities is lower in regions with stronger religious power. The absolute value of the coefficient of the interaction term between the minority dummy variable and the regional minority percentage decreases slightly compared to column 4, but is still negatively significant, indicating that religion is not the only channel through which the regional minority percentage has an impact.

[Table 10: Influence of Religion]

5 Conclusion

The relatively lower mobility of ethnic minorities reflects the fact that China's ability to govern frontier areas is still inadequate, and the degree of integration of the national labor market still needs to be improved. From the perspective of the individual, the main destination of population movement is to go out to work and do business (as shown in Figure 10), which provides opportunities for households with surplus labor, especially those in rural areas where income channels are originally narrower, to obtain higher income. The higher income gained by working outside improves the living standard of these households and gives them more incentive and financial resources to invest in their children's education, enhancing the mobility of social classes. Thus, the stronger adhesion of minority groups to their domicile implies lower average income levels and social status, and this trend is perpetuated and reinforced through inter-generational transmission. From the perspective of the labor market, the eastern and central regions of China with high labor demand have experienced a labor supply shortage, while the labor force in the vast minority regions has still not been fully released. This historical and culturally induced inability of the national labor market to reach natural equilibrium prevents the advantages of China's large overall population and economic volume from being fully utilized and reduces the efficiency of domestic labor allocation.

[Figure 10: Reasons for Population Movement in the 2005 Sample]

To summarize the full paper, this study explored the factors affecting the mobility of ethnic minorities using a 20% subsample of the 2005 National 1% Population Sampling Survey micro-data, using a binary probit model and an instrumental variable approach. It is found that living in compact communities of ethnic minorities reduces the willingness of individuals among them to engage in population movement. Possible explanations are that ethnic compact communities enhance the closedness of ethnic minority groups and reflect the incomplete integration of ethnic regions by Central Plains Dynasty in history, an effect that lasts until today. Through heterogeneity analysis and mechanism testing, this study also found that lower proficiency in Mandarin would be detrimental to population mobility, and that ethnic compact communities would affect local language education and religious power, thus influencing the behavioral choices of population movement.

References

- [1] Barbieri, A. F., Carr. D. L. Gender-specific out-migration, deforestation and urbanization in the Ecuadorian Amazon. *Global & Planetary Change*, 2005, 47(2/4):99-110.
- [2] Cao, G., Liu, T. Rising Role of Inland Regions in China's Urbanization in the 21st Century: The New Trend and Its Explanation. *Acta Geographica Sinica*, 2011, 66(12):1631-1643.
- [3] Cao. G., Bian, X. Comprehensive evaluation of the urbanization level in Yangtze River Delta Region: An index framework based on the population, economic structure and land use. *Acta Geographica Sinica*, 2011, 30(12):2139-2149.
- [4] Cohen, J. H. Remittance outcomes and migration: Theoretical contests, real opportunities. *Studies in Comparative International Development*, 2005, 40(1):88-112.
- [5] Duan, C., Chi, S. Study on Floating Minorities of China. *Population Journal*, 2011, (03):21-28.
- [6] Duan, C., Feng, L., Qin, M. Situation and Comparison of Characteristics of the Floating Population in Typical Minority Nationality Areas——A Study of the Population in Inner Mongolia Autonomous Region. *Journal of Inner Mongolia University (Philosophy and Social Sciences)*, 2017, 49(06):5-10.
- [7] Duan, C., Liu, T., Lyu, Li. Research on the current situation of population mobility in China and its impact. *Shandong Social Science*, 2017, (09):63-69.
- [8] Duan, C., Lyu, L., Wang, H., Xie, D. From Rural China to Migrating China: Rethinking Migration Transition in China. *Population Research*, 2020, 44(01):19-25.
- [9] Duan, C., Xie, D., Lyu, L. Migration Transition in China. *Population Research*, 2019, 43(02):12-20.
- [10] Duan, C., Yang, G. Trends in Destination Distribution of Floating Population in China. *Population Research*, 2009, 33(06):1-12.

- [11] Duan, C., Zhao, C., Lyu, L. Changes of Destination Distribution of Floating Population in China (2000-2015). *Population Economics*, 2020, (01):89-99.
- [12] Gedik, A. The effects of in- and out-migration on urban growth in Turkey (1965-85) and a comparison with the developed countries. *Papers in Regional Science*, 1992, 71(4):405-419.
- [13] Harris, J. R., Todaro, M. P. Migration, Unemployment Development: A Two-Sector Analysis. *American Economic Review*, 1970, 60(1):126-142.
- [14] Heberle, R. The Causes of Rural-Urban Migration a Survey of German Theories. *American Journal of Sociology*, 1938, 43(6):932-950.
- [15] Lee, E. S. Quasi-linearization, non-linear boundary value problems and optimization. *Chemical Engineering Science*, 1966, 21(2):183-194.
- [16] Lewis, W. A. Report on Industrialization and the Gold Coast. 1954.
- [17] Li, L. Internal population migration in China since the economic reforms: A review. *Acta Geographica Sinica*, 2001, (04):453-462.
- [18] Li, Q. An Analysis of Push and Pull Factors in the Migration of Rural Workers in China. *Social Sciences in China*, 2003, (1):125-136+207
- [19] Liu, A. Analysis on the Spatial Agglomeration and Its Influence Factors of the Floating Population in Beijing-Tianjin-Hebei. *Population Economics*, 2017, (06):71-78
- [20] Liu, T., Qi, Y., Cao, G. China's floating population in the 21st century: Uneven landscape, influencing factors, and effects on urbanization. *Acta Geographica Sinica*, 2015, 70(04):567-581.
- [21] Ma, R. China's Ethnic Question under the Trend of Nationalism. *People's Tribune*, 2011, (01):24-25.
- [22] Ma, R. Factors Affecting the Changes in Ethnic Relations. *Northwest Minorities Research*, 2003, (04):5-29.
- [23] Ma, R. Rural Migrant Workers from South Xinjiang Uygur to Coastal Cities——A Case Study on Shufu County, Xinjiang Uygur Autonomous Region. *Chinese Journal of Population Science*, 2007, (05):23-35+95.
- [24] Ravenstein, E. G. The Laws of Migration, Part I. *Journal of the Statistical Society of London*, 1885, 48(2):167-235.
- [25] Wang, X., Tian, B. Research on Characteristics and Influence Factors of Yanbian Korean Population Flow. *Population Journal*, 2015, 37(03):78-87.
- [26] Wang, Y., Tong, Y., Liu, A. The Formation of Spatial Distribution Pattern of Population in China's Capital Circle“Game” of Agglomeration and Dispersion. *Population Journal*, 2017, 39(04):5-16.

- [27] Wu, Y. Current Situation and Influence of Border Area Population Outflow[D]. Jilin University, 2013
- [28] Yu, T. Spatial-temporal Features and Influential Factors of the China Urban Floating Population Growth. *Chinese Journal of Population Science*, 2012, (04):47-58+111-112.
- [29] Zhang, S., Zeng, M. The Distributional Change and the Migration Situation of the Ethnic Population: An Analysis of the Data from China's Fifth Census of 2000. *Ethno-national Studies*, 2005, (01):17-25+107-108.
- [30] Zhao, Y., Wen, J. Chinese rural social security and land holding[M]. 1998.
- [31] Zheng, X., Zhou, J. A Research into the Ethnic Population Floating and Urban Ethnic Relations. *Journal of South-Central University for Nationalities (Humanities and Social Sciences)*. 2002, (04):52-59.

Appendix. Figures and Tables

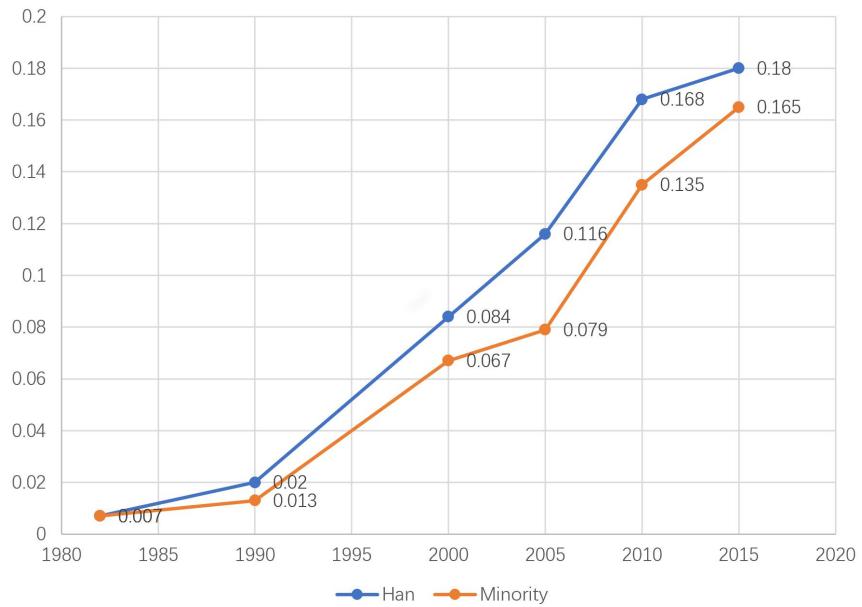


Figure 1: Percentage of floating population in the total *Han* people and ethnic minorities

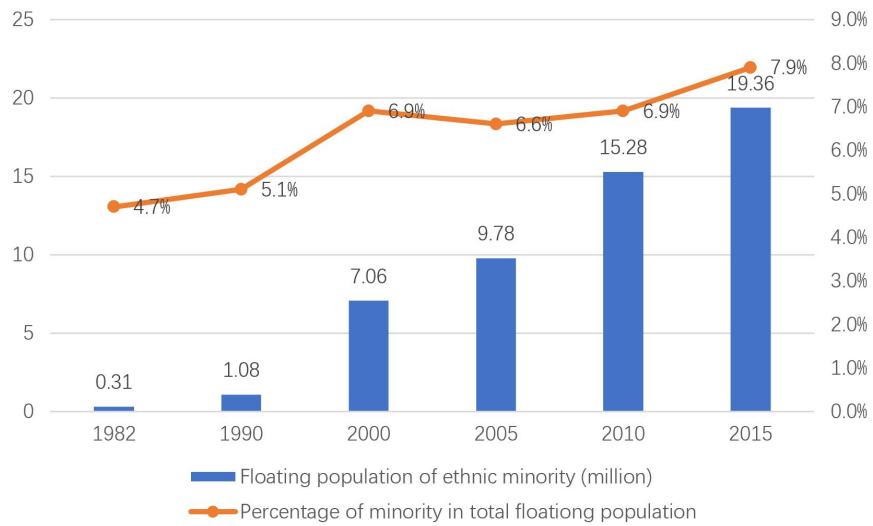


Figure 2: Scale and Percentage of minority in total floating population

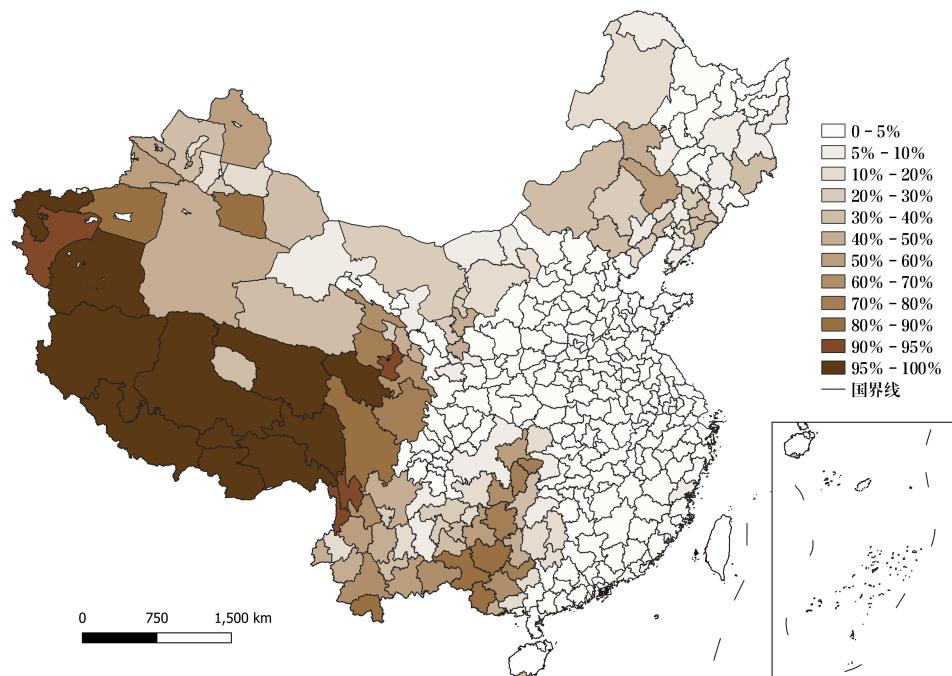
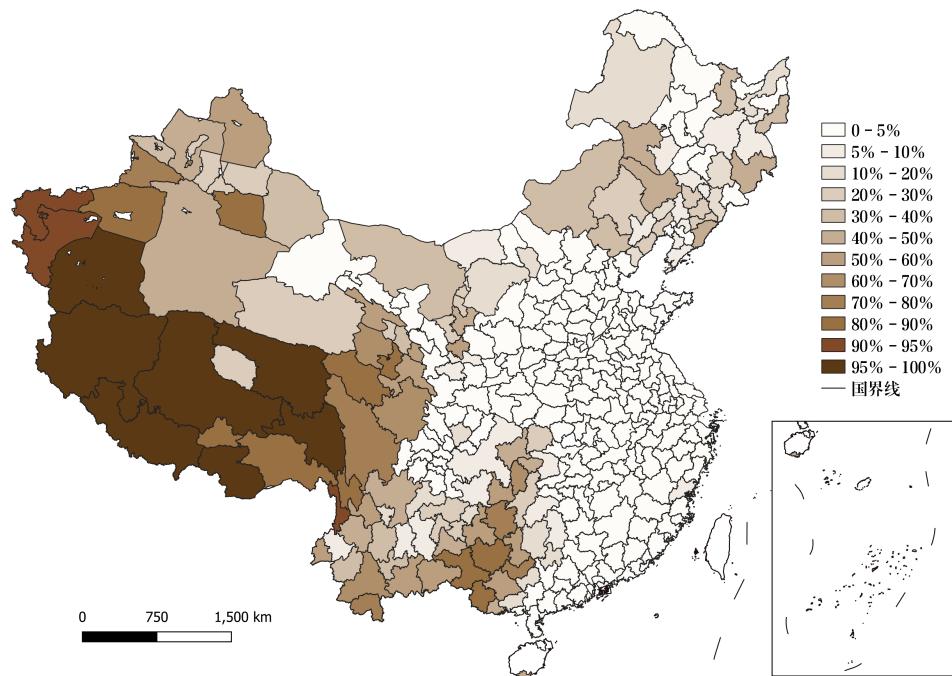


Figure 3: Percentage of ethnic minorities in the total population by prefecture-level regions, 1990 (upper) and 2005 (lower)

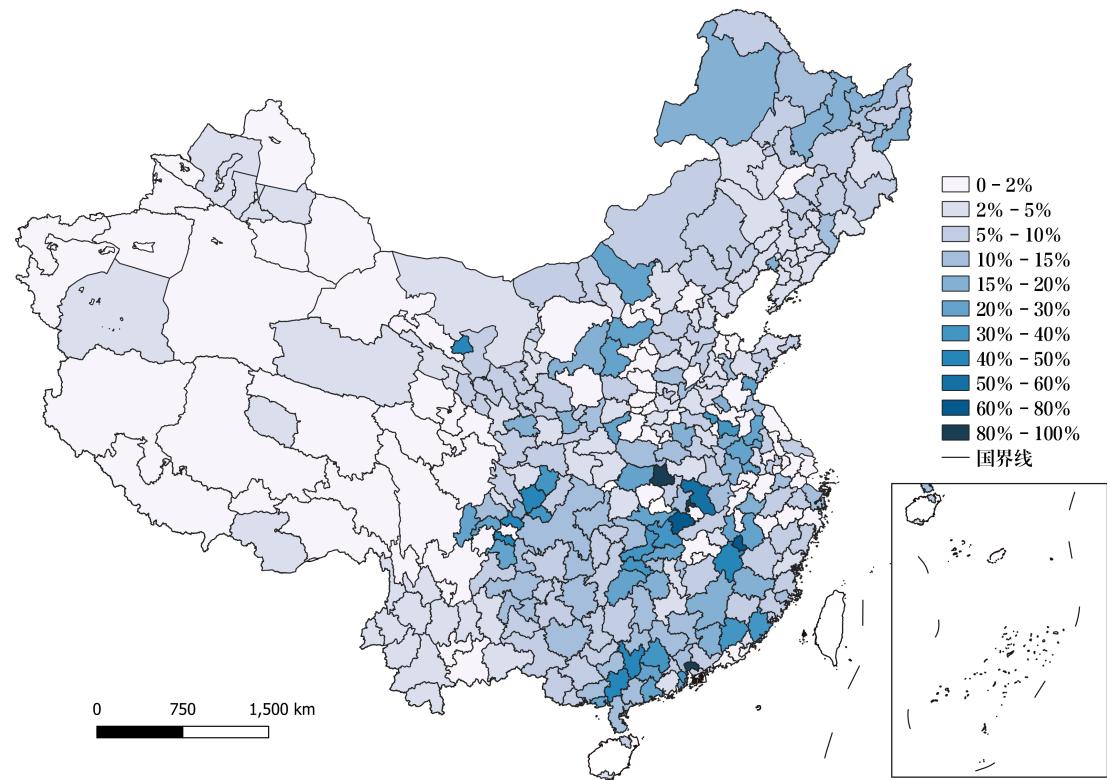
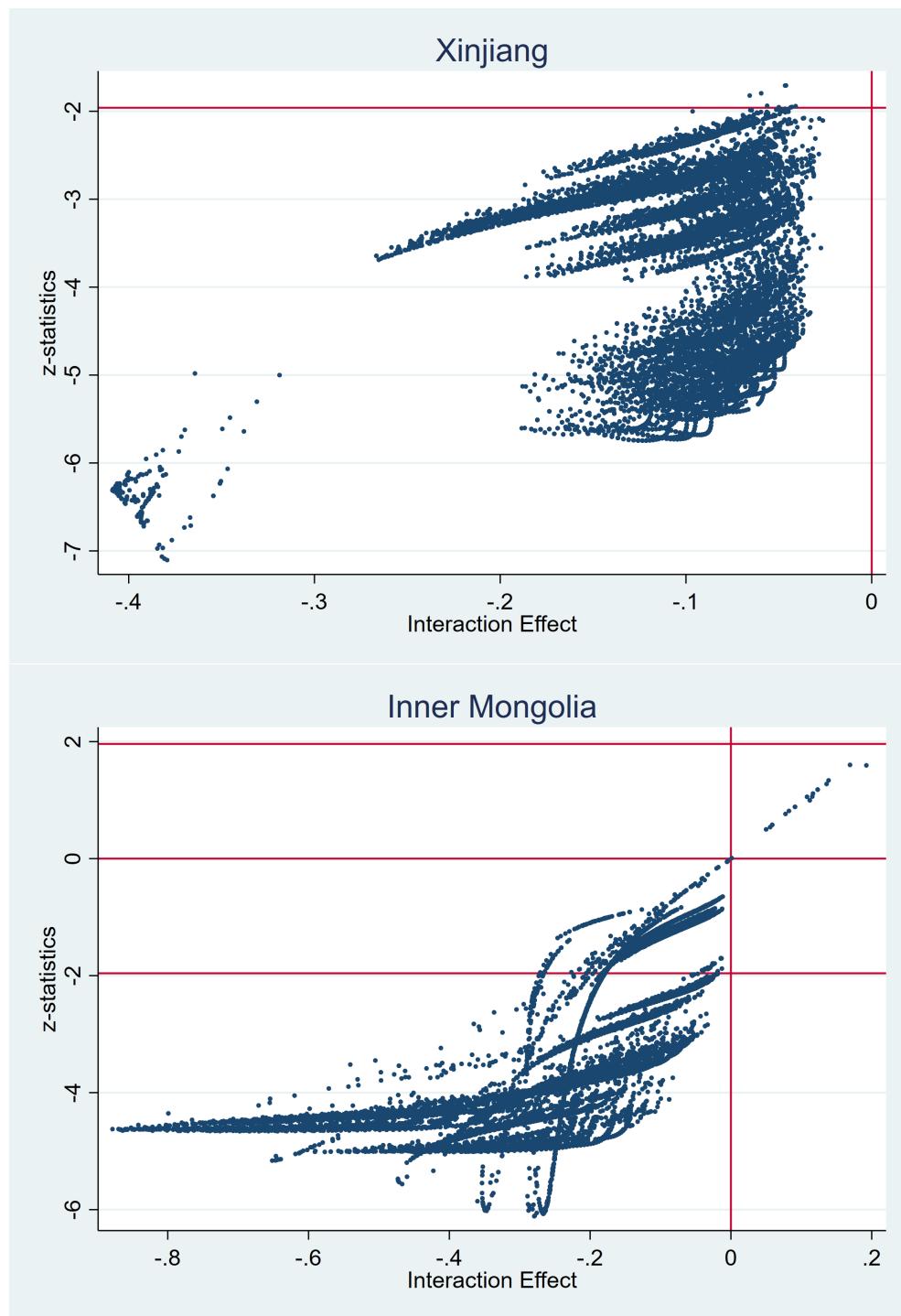
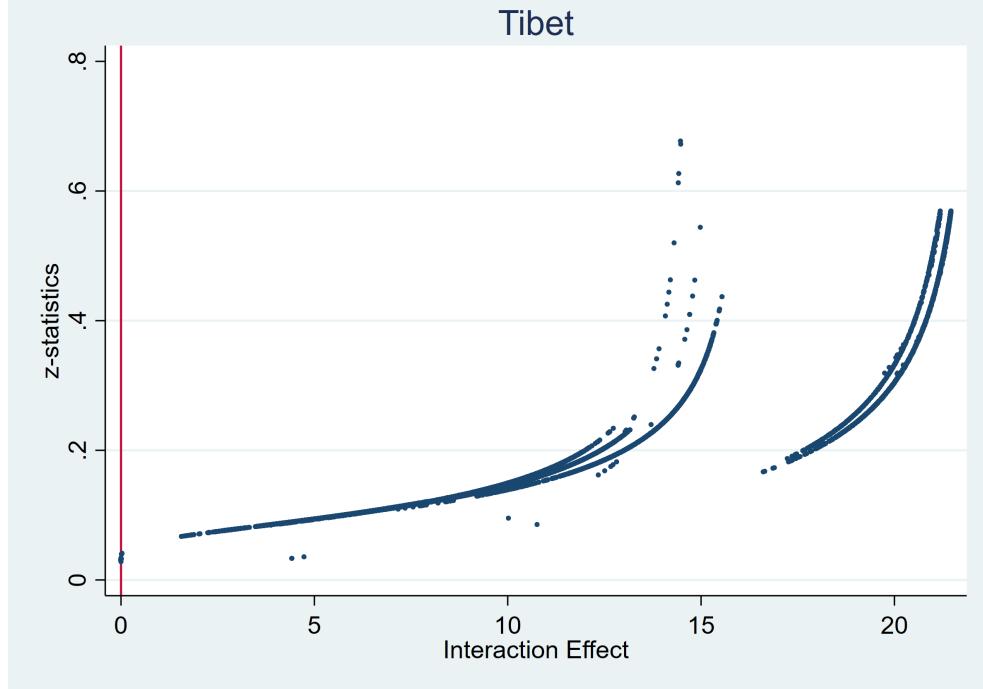
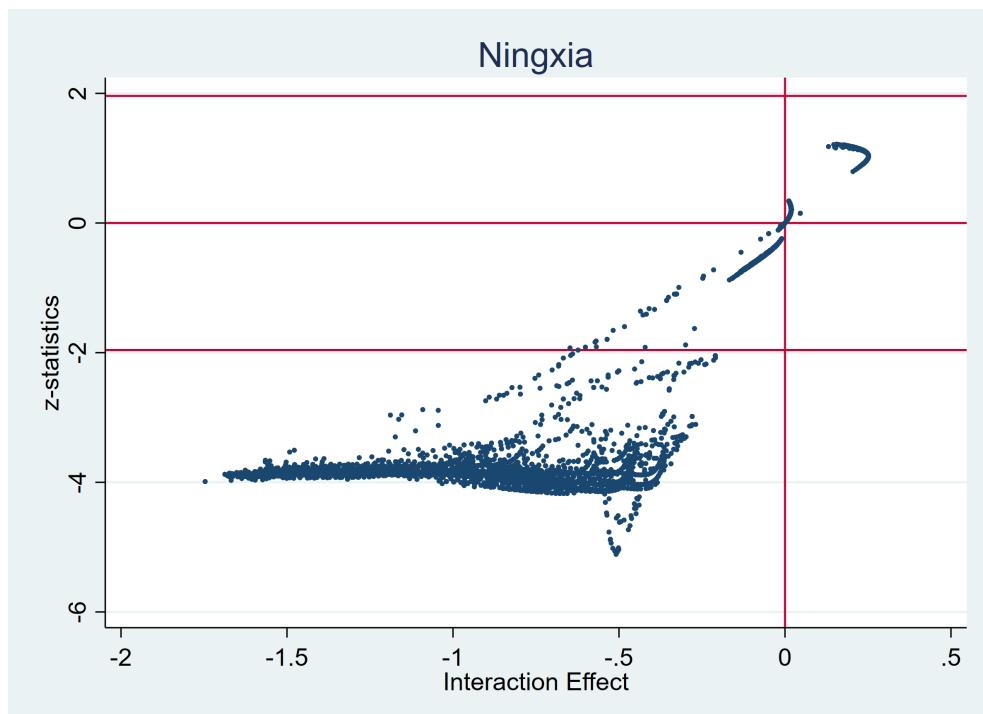


Figure 4: Percentage of floating ethnic minority population in the total ethnic minority population by prefecture-level regions, 2005





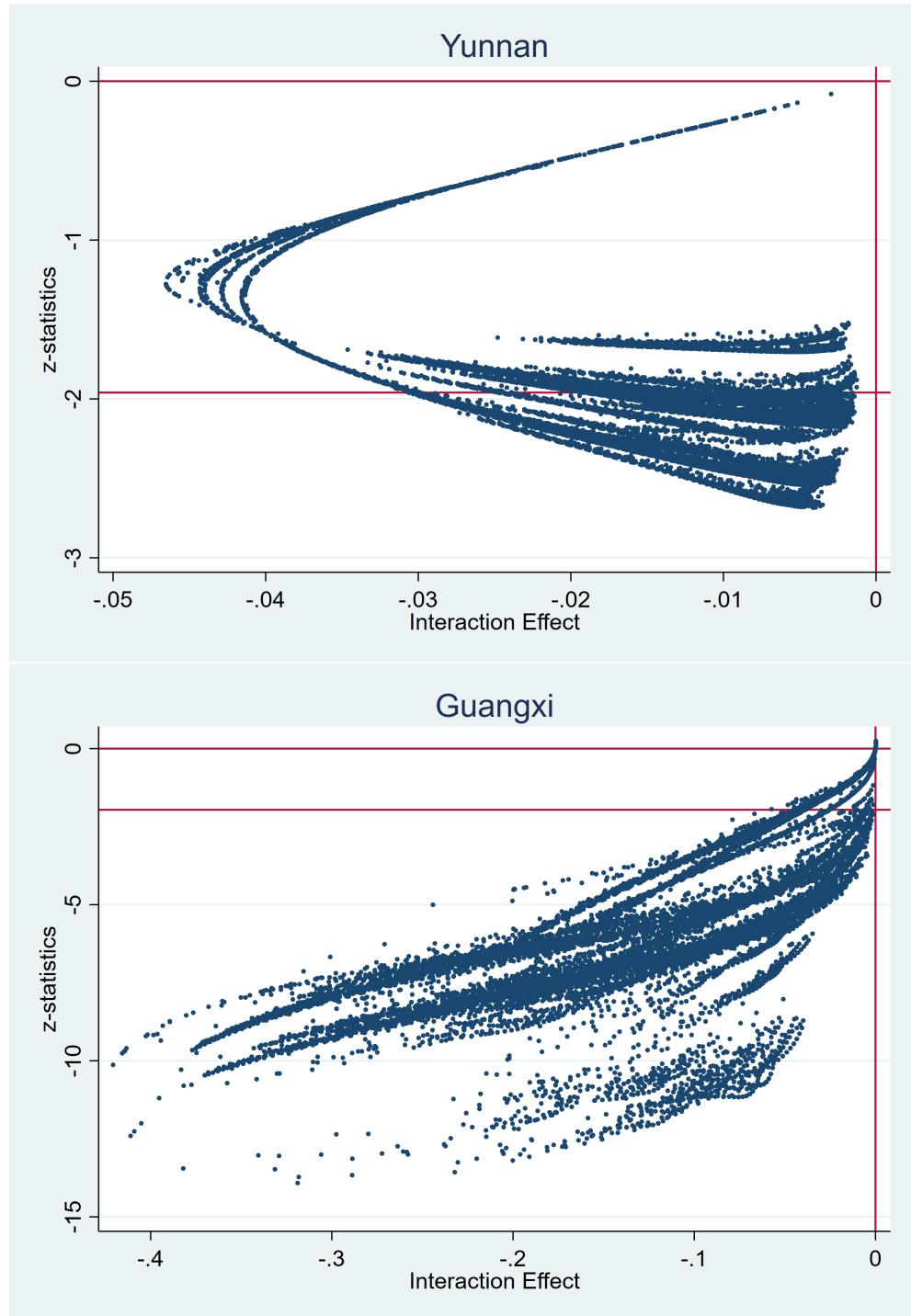


Figure 5: Scatter plot of interaction effect-z statistic for six provincial administrative regions with dense concentrations of ethnic minorities

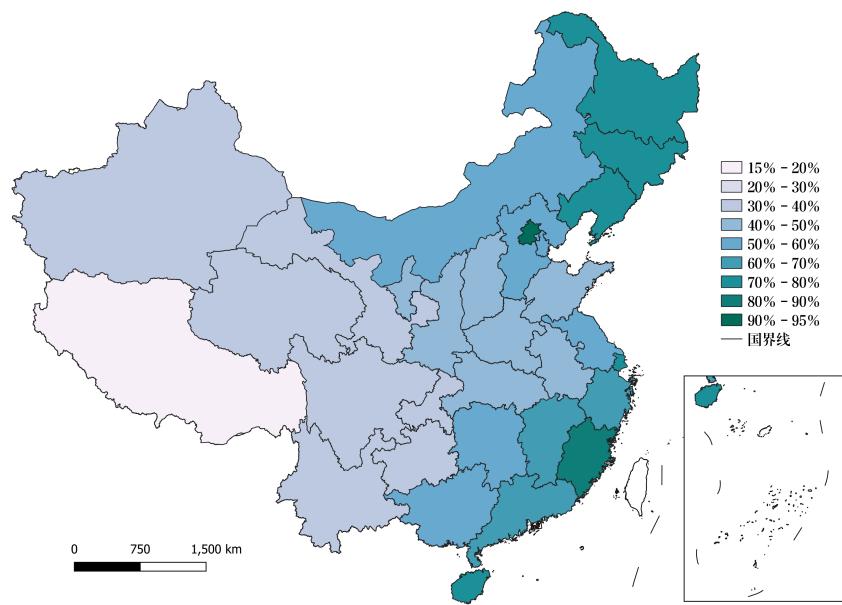


Figure 6: Percentage of respondents who can speak in Mandarin by provincial-level administrative regions

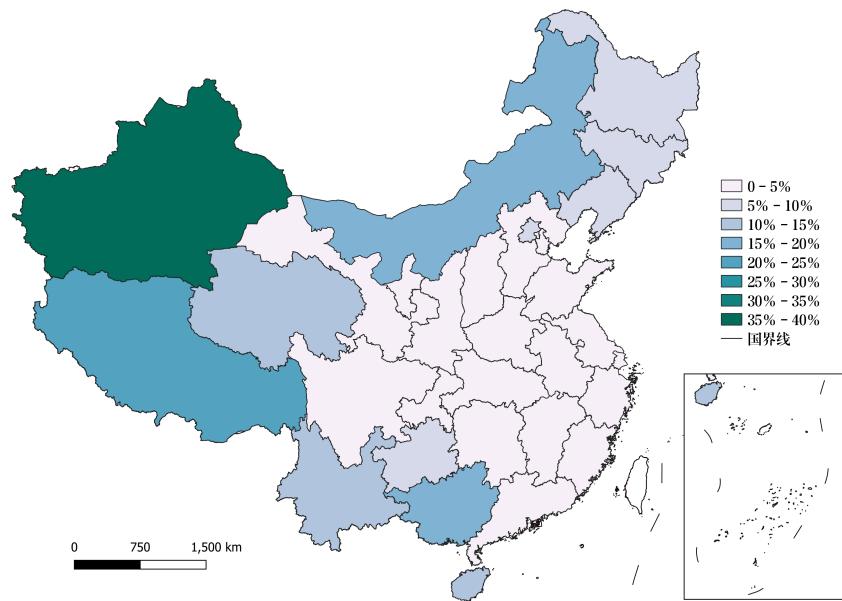


Figure 7: Percentage of respondents who encountered difficulties in learning to speak Mandarin “due to the influence of their own language and the difficulty in changing their accent” by provincial-level administrative regions

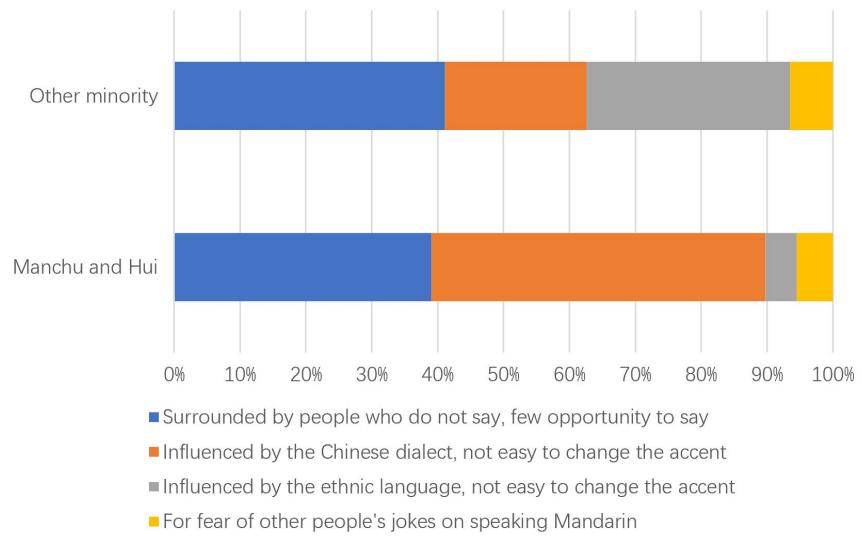


Figure 8: Percentage of major problems encountered by ethnic minorities in learning to speak Mandarin

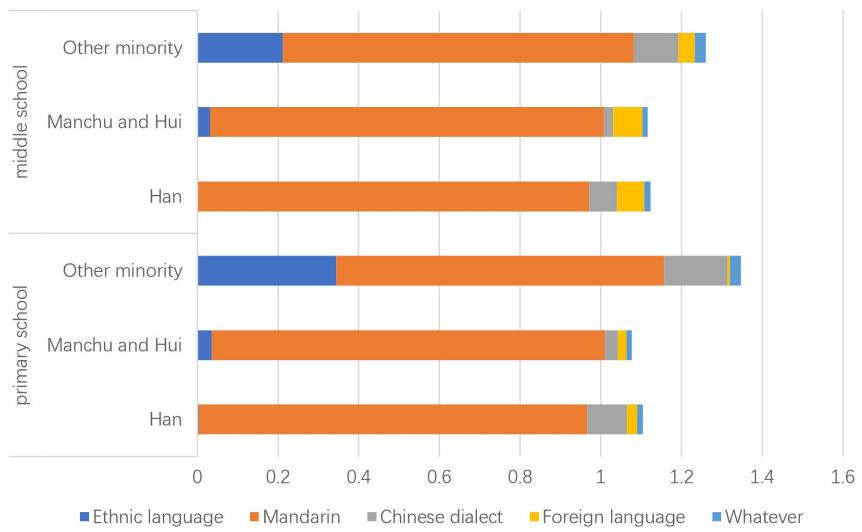


Figure 9: Percentage of respondents who want local primary and secondary schools to use a particular language as the language of instruction

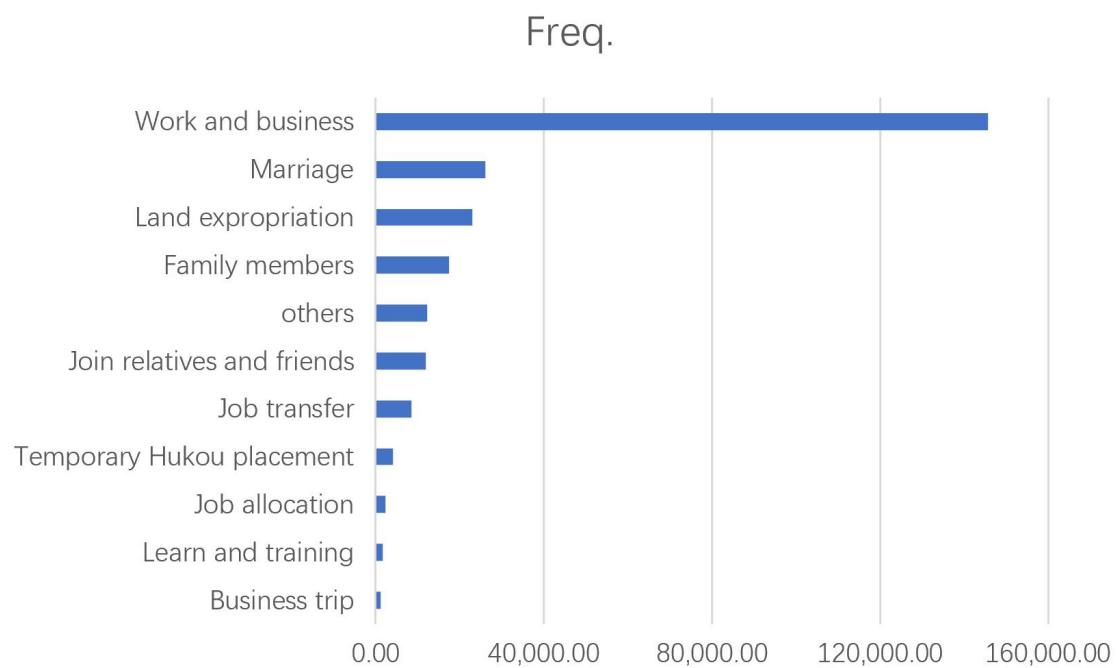


Figure 10: Reasons for Population Movement in the 2005 Sample

Ethnicity	National Overall	Sub-sample	Difference
Han	90.95%	91.08%	0.13%
Mongolian	0.49%	0.49%	0%
Hui	0.77%	0.75%	-0.02%
Tibetan	0.57%	0.62%	0.05%
Uighur	0.74%	0.77%	0.03%
Miao	0.75%	0.74%	-0.01%
Yi	0.79%	0.76%	-0.03%
Zhuang	1.37%	1.21%	-0.16%
Buyei	0.26%	0.23%	-0.03%
Korean	0.14%	0.14%	0%
Manchu	0.82%	0.82%	0%
Dong	0.28%	0.27%	-0.01%
Yao	0.24%	0.24%	0%
Bai	0.14%	0.15%	0.01%
Tujia	0.65%	0.67%	0.02%
Hani	0.12%	0.12%	0%
Kazakh	0.09%	0.10%	0.01%
Dai	0.08%	0.09%	0.01%
Li	0.11%	0.11%	0%

Table 1: Percentage of the population of all ethnic groups in the total population

Variables	Meaning and definition
y_{ijk}	Whether the individual i whose <i>Hukou</i> is in province k prefecture-level city j belongs to floating population
$minority_i$	Whether the individual i is an ethnic minority who lives in the ethnic enclave of that ethnicity
min_prop_j	Percentage of ethnic minority in the total population of prefecture-level city j in 1990
$minority_i \times min_prop_j$	Interaction term of $minority_i$ and min_prop_j
x'_{1i}	Control variables of individual i , including gender, <i>Hukou</i> type, education, health, marriage, age, squared age
x'_{2j}	Control variables of city j , including GDP、GDP per capita、fiscal revenue、length of road、average altitude、percentage of employees in the primary industry、population density、registered unemployment rate、fixed-line and cell phone ownership rates, etc.
λ_k and u_{ijk}	Dummy variable indicating province k and error term

Table 2: Meaning of notations in econometric model

Variables	Unit	Obs	Mean	Sd	Min	Max
Individual						
gender	0/1	1550876	.494	.5	0	1
whether collective household	0/1	1550876	.961	.194	0	1
education	1~7	1550876	2.934	1.047	1	7
<i>Hukou</i> type	0/1	1550876	.706	.455	0	1
health	0~3	1550876	1.042	.242	1	3
marriage	0/1	1550876	.812	.391	0	1
age	year	1550876	38.214	11.465	16	60
squared age	\	1550876	1591.797	889.608	256	3600
<i>migration</i>	0/1	1550876	.099	.298	0	1
<i>minority</i>	0/1	1550876	.093	.29	0	1
Region						
Han population in 1990	person	325	3122961.7	2654856.8	1432	27375841
total population in 1990	person	325	3393207.6	2621527.9	61639	28865870
percentage of ethnic minority population in 1990	\	325	0.139	0.241	0	0.977
fiscal revenue	million RMB	325	3844.459	10585.575	17	143390
total population in 2005	10 000 people	325	383.684	289.764	14.9	2798
land area	sq. km	325	25921.926	48543.174	1113	470954
average altitude	meter	325	735.801	933.482	7.337	5044.376
population density	people/sq. km	325	5.335	1.373	-0.243	8
number of employed	1000 people	325	817.332	1424.978	42.9	14563
number of unemployed	1000 people	325	21.08	24.319	0.252	275
unemployment	\	325	0.051	0.037	0.001	0.166
number of employees in the primary industry	1000 people	325	207.117	527.708	0.1	6783.2
percentage of employment in the primary sector	\	325	0.161	0.227	0	0.82
number of landline subscribers	1000 households	325	1039.181	1204.102	9	9966.544
number of mobile phone subscriber	1000 households	325	1262.369	1918.932	18	14597.845
landline usage rate	\	325	0.289	0.271	0.027	2.556
mobile phone usage rate	\	325	0.358	0.568	0.015	6.659
length of road	kilometer	325	86040.366	6422.491	483.3	98218
GDP per capita	RMB	325	14606.543	11539.933	2394	88562
GDP	a billion RMB	325	61.281	93.696	0.79	919.713

Table 3: Description statistics of variables

Dependent Variable:	<i>migration</i> (whether belongs to floating population)			
	(1) Probit	(2) Probit	(3) IV	(4) Probit IV
<i>minority</i>	-0.238*** (-29.45)	-0.0976*** (-9.70)	-0.00501 (-1.41)	0.0814** (2.08)
<i>minority</i> × <i>minprop</i>		-0.610*** (-20.08)	-0.0928*** (-5.56)	-1.486*** (-7.91)
City control variables	Y	Y	Y	Y
Province Dummy	Y	Y	Y	Y
N	1023046	1023046	1023046	1023046

Note: t value in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Baseline regression with instrumental variables regression

Dependent Variable:	<i>migration</i> (whether belongs to floating population)			
	(1) Probit	(2) Probit	(3) IV	(4) Probit IV
<i>minority</i>	-0.215*** (-20.92)	-0.122*** (-9.54)	-0.00932*** (-3.16)	0.0274 (0.55)
<i>minority</i> × <i>minprop</i>		-0.427*** (-10.60)	-0.0324** (-2.27)	-1.213*** (-4.75)
City control variables	Y	Y	Y	Y
Province Dummy	Y	Y	Y	Y
N	751657	751657	751657	751657

Table 5: Robustness Check 1: Village and Town Sample in Central and Western Provinces

Dependent Variable: <i>migration</i> (whether belongs to floating population)		(1)	(2)	(3)	(4)
		Probit	Probit	IV	Probit IV
<i>minority</i>		-0.205*** (-16.43)	-0.130*** (-8.48)	-0.00334 (-1.21)	0.0336 (0.62)
<i>minority</i> × <i>minprop</i>			-0.396*** (-8.10)	-0.0507*** (-3.61)	-1.324*** (-4.41)
City control variables	Y	Y	Y	Y	Y
Province Dummy	Y	Y	Y	Y	Y
N	573377	573377	573377	573377	573377

Note: t value in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Robustness Check 2: Village Sample in Central and Western Provinces

Dependent Variable: <i>migration</i> (whether belongs to floating population)		(1)	(2)	(3)	(4)
		Probit	Probit	IV	Probit IV
<i>minority</i>		-0.228*** (-28.74)	-0.133*** (-13.10)	-0.00914*** (-2.64)	0.0889** (2.23)
<i>minority</i> × <i>minprop</i>			-0.431 *** (-14.37)	-0.0794 *** (-4.96)	-1.523 *** (-7.90)
City control variables	Y	Y	Y	Y	Y
Province Dummy	Y	Y	Y	Y	Y
N	1023046	1023046	1023046	1023046	1023046

Table 7: Robustness Test 3: Migrant Workers

Ethnicity	Percentage of people who can talk to people in Chinese	Percentage of people who can talk to people in ethnic language	linguistic ethnicization
Achang	0.9875	0.8615	0.437
Bai	0.8354	0.9137	0.53915
Baoan	0.9911	0.4925	0.2507
Burang	0.7764	0.9693	0.59645
Buyi	0.9455	0.5033	0.2789
Korean	0.8411	0.9399	0.5494
Daur	0.925	0.8713	0.47315
Dai	0.8858	0.9828	0.5485
De'ang	0.8965	0.999	0.55125
Dongxiang	0.857	0.7105	0.42675
Dong	0.8904	0.5727	0.34115
Dulong	0.4812	0.9579	0.73835
Russian	0.9927	0.4951	0.2512
Oroqen	0.9858	0.5972	0.3057
Ewenke	0.7867	0.9423	0.5778
Gelao	0.9987	0.0146	0.00795
Hani	0.6828	0.9461	0.63165
Kazakh	0.4237	0.9908	0.78355
Han	1	0	0
Hezhen	0.9999	0.0267	0.0134
Hui	0.9966	0.046	0.0247
Jinuo	0.8195	0.9686	0.57455
Kinh	0.9772	0.9337	0.47825
Jingpo	0.9209	0.975	0.52705
Kirgiz	0.1221	0.9828	0.93035
Lahu	0.8146	0.9485	0.56695
Li	0.9551	0.8918	0.46835
Lisi	0.7118	0.9872	0.6377
Lhoba	0.3543	0.8974	0.77155
Manchu	0.9999	0	0
Maonan	0.8803	0.4753	0.2975
Memba	0.1322	0.9752	0.9215
Mongolian	0.7138	0.7552	0.5207
Miao	0.9212	0.597	0.3379
Mulao	0.8626	0.9123	0.52485
Naxi	0.8045	0.9834	0.58945
Nu	0.6027	0.9819	0.6896

Table 8: Calculation of linguistic ethnicity index

Pumi	0.8681	0.9777	0.5548
Qiang	0.9861	0.1466	0.08025
Salar	0.7395	0.6241	0.4423
She	0.9999	0.002	0.00105
Shui	0.5867	0.9077	0.6605
Tajik	0.066	0.9975	0.96575
Tatar	0.4325	0.9896	0.77855
Tujia	0.9939	0.0663	0.0362
Tujia	0.5958	0.8402	0.6222
Wa	0.8427	0.9926	0.57495
Uyghur	0.1988	0.9974	0.8993
Uzbeks	0.1609	0.9812	0.91015
Xibo	0.8936	0.9387	0.52255
Yao	0.8947	0.749	0.42715
Yi	0.8143	0.6907	0.4382
Yugur	0.9814	0.6426	0.3306
Tibetan	0.5187	0.904	0.69265
Zhuang	0.7999	0.8616	0.53085

Note: The Alpine ethnic group is mainly distributed in the mountainous areas of Pingtung and Taitung counties in southern Taiwan, for which data are missing.

Table 8: Calculation of linguistic ethnicity index (continued)

Dependent Variable:	<i>migration</i> (whether belongs to floating population)			
	(1) Probit	(2) Probit	(3) IV	(4) Probit IV
<i>minority</i> × <i>linguistic ethnicization</i>	-0.621*** (-32.81)	-0.255*** (-8.74)	-0.0304*** (-6.57)	-0.270*** (-4.76)
<i>minority</i> × <i>linguistic ethnicization</i> × <i>minprop</i>		-0.968*** (-15.31)	-0.0547*** (-4.80)	-0.928*** (-6.23)
City control variables	Y	Y	Y	Y
Province Dummy	Y	Y	Y	Y
N	1023026	1023026	1023026	1023026

Note: t value in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Effects of Ethnic Language

Dependent Variable:	religious sites per capita		<i>migration</i> (whether belongs to floating population)		
	(1) OLS	(2) Probit	(3) Probit IV	(4) Probit	(5) Probit
<i>minprop</i>	5.206*** (2.99)			0.119*** (5.32)	0.0997*** (4.33)
<i>minority</i> × religious sites per capita		-0.0194*** (-8.74)	-0.0487*** (-5.71)		-0.00823*** (-3.54)
<i>minority</i>			-0.228*** (-28.71)	-0.218*** (-26.55)	-0.0976*** (-9.70)
religious sites per capita		0.00741 *** (3.49)	0.0324 *** (4.41)		0.00439 ** (2.04)
<i>minority</i> × <i>minprop</i>				-0.610*** (-20.08)	-0.566*** (-17.38)
City control variables	Y	Y	Y	Y	Y
Province Dummy	Y	Y	Y	Y	Y
Ns	248	1023046	1023046	1023046	1023046

Note: t value in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Influence of Religion