Education Level, Urban Development Level and Mobility Household Aggregation——Empirical analysis based on a sample of 2000 in 2018

Xinyi Huang*

HNU-ASU Joint International Tourism College, Hainan University, Haikou, Hainan, 570228, China

*Corresponding author: huangxinyi0309@163.com

Keywords: education level, urban development level, mobile population, household aggregation, Probit model

Abstract: The issue of migrant families is one of the most significant problems to be solved in urbanization. Separation results in a sharp decline in the happiness of numerous families. Thus, to enhance the happiness of citizens, it is essential to investigate the influence of regional education level and urban development level on the aggregation of migrant families. This paper uses a quantitative research method as well as Probit models to delve into the mutual effects of these variances and then obtains the following results in the heterogeneity analysis: (1) Mobile population in high education level areas tend to be separated from their children for a long time. (2) Mobile populations in high economy-level areas have higher probabilities of living separately from their own children for a long period. (3) Areas with large proportions of industrial structure offer more employment opportunities and can tackle family separation issues.

1. Introduction

By 2020, the floating population in China has reached 376 million, an increase of 70% compared with 2010. The level of urbanization continues to intensify. Such a large-scale mobile population brings new urban development directions and requirements. Regional public resource utilization, public service supply, rationalization of industrial structure, the development needs of the mobile population, etc., are all agendas that require continuous attention. Therefore, finding how migration characteristics affect the mobile population promotes the stable development of population mobility, the policy formulation of household registration, and education.

The degree of family aggregation is a critical indicator of the stability and the future development of the migrant population during the migration process. The established literature pointed out that the education returns associated with education level and the difference between resources and consumption, represented by urban development level, directly affect the degree of family aggregation of the migrant population. However, when the urban development level of different cities is used as a variable to explore the influence on the aggregation degree of migrant families, few studies have considered the diversity of education levels across regions. Based on this, this paper primarily focused on two levels of migrant household aggregation, education level, and urban development level. Moreover, we applied the latest 2018 CMDS, released by the Migrant Population Service Centre of the National Health and Family Planning Commission, as the research data. OLS model and logit model were used for robustness testing. In the end, results were obtained by heterogeneity analysis of the data.

2. Literature Review

2.1 Educational standards and family aggregation of mobile population

According to human capital theory, individuals acquire certain knowledge, labor and management skills, etc., through education and training, which is an investment that can help them produce greater value in social life. Based on this, Cai (2021) suggests that access to education is particularly essential for the mobile population. Also, it is one of the significant factors that influence the mobility

characteristics of the mobile population. Based on this, parents are more likely to invest in areas with higher educational standards to get a higher return on education. Huang and Zhu (2021) proved that the returns to education, especially higher education, are higher for the mobile population than local residents, being a positive cycle. In other words, those capable among the mobile population self-select to metropolitans with higher levels of education to receive larger returns to education. Therefore, parents who have jobs in areas with high educational standards are more likely to have their children move with them.

Sun and Wang (2019) suggest an indirect positive relationship between the stage of schooling of the children and the mobile family reunion. Areas with a developed education industry cover a longer period of quality instruction. In order to give their children more access to better educational resources as possible, parents tend to increase the amount of time living with their children. In addition, high educational standards are an essential factor for the mobile population to move in. Since they already wanted to have their children move with them before they move to such cities, they are supposed to act quickly. Hence, areas that can access high levels of education, such as college, are conducive to families moving with their children and living together for long periods.

Accordingly, the following hypothesis is proposed.

Hypothesis 1-1: The higher the level of education, the lower the likelihood that the mobile population will be separated from their children.

Hypothesis 1-2: The higher the level of education, the shorter the time of separation of the mobile population from their children.

2.2 Urban development level and family reunion of mobile population

According to Zhu (2021), Hu and Kang (2021), based on the principle of Family Economic Rationality, the reunion effect and child development effect make it easier for the mobile population in the all-family migration mode to avoid risks and accumulate family capital, which directly affects the mobile population's willingness to reside. Urban development factors such as elementary education, social welfare, employment, industrial structure, and so on are all factors that impact the mobile population's choice on whether to settle in the inflow area (Chen et al., 2021). On the one hand, areas with high levels of economic development have fewer resources available to individuals, especially mobile populations. It causes a high consumption and dilution of community resources, with a decline in the quality of transportation, services, etc., further leading to a decline in the per capita standard of living. On the other hand, developed areas are not the most livable places due to the fast pace of life. Specifically, a variety of pollution such as noise, air pollution, water pollution, which is normal in such cities, is harmful to people's health. Therefore, Yang et al. (2020) pointed out that the new generation of the mobile population is more willing to migrate to medium or lower development level areas with family members instead of high-level.

Li (2020) and Ai et al. (2020) deem that the mobile population living in areas with high levels of urban development often faces high housing prices, living costs, and education costs. Moreover, the financial conditions gap between them and the local people lead to low consumption of the mobile population. Consequently, the longer the children live with their parents, the higher the various costs required. It is difficult for mobile people to afford the cost of long-term co-residence for their children when they live unstably, being one of the main obstacles for mobile families to aggregate. In addition, such urbanization is "individual urbanization" rather than "family urbanization". The long-term co-habitation of children requires parents to build complex social networks and access to human capital to meet the needs of their children's education, training, and so on. In developed regions, the labor market is often saturated. It is difficult for mobile people to take care of their children while their own needs cannot be met. Therefore, high-level development cities hinder the children living with their parents on a long-term basis in mobile families. (Chen et al., 2019).

Based on the above research, we hope to test the following hypotheses:

Hypothesis 2-1: The higher level of urban development, the higher the probability that the mobile population will be separated from their children.

Hypothesis 2-2: The higher level of urban development, the longer the mobile population is

separated from their children.

3. Data and Model Setup

3.1 Research Data Source

We hope to delve into the influence of education level and urban development level on the household aggregation of the mobile population. We chose household aggregation of the mobile population as the explained variable, education level and urban development level as the explanatory variables. Meanwhile, we found that the 2018 CMDS meets the requirements of this paper. There are 2000 of these typical samples from 10 cities in each regional province, including Beijing, Shenyang, Hohhot, Nanjing, etc.

3.2 Introduction and Selection of Variables

3.2.1 Main Variables

The degree of family aggregation of the mobile population is evaluated by whether the mobile population is separated from their children and the separated time. Core explanatory variables are education standards and urban development. The former reflects selections of the college entrance examination papers and whether children of mobile population are allowed to take the college entrance examination in the cities. The latter is represented by population size, GDP per capita, industry structure, etc. Moreover, the representation of these dummy variables was based on Liu (2020). The industry structure is specified as the ratio of the output value of the tertiary industry to the secondary industry, while house price spatial difference is the ratio of affordable housing prices to annual per capita disposable income.

Those data all refer to cities' education and residency policy as well as the 2019 Statistical Yearbook released by the government.

3.2.2 Control Variables

Due to the differences among individuals in the sample, the following variables were selected as control variables for analysis in this paper: gender, age, ethnicity, education level, mobility range, etc.

Table 1. Symbols and definitions of variables

Variable name	Description
Separation (y_1)	0= Separated;1= Not separated
Separation time (y_2)	0= y_1<1(year); 1= 1<y_1<5; 2= 6<y_1<10; 3= y_1>10
Selections of the college entrance examination papers (x_1)	1= NationalI; 2= NationalII;3= NationalIII;4= Beijing; 5= Tianjin;6= Shanghai; 7= Jiangsu
Take the college entrance examination out of the domicile place (x_2)	0= Allowed; 1= Not allowed
Population size (x_3)	The number of permanent residents (Unit: million)
GDP per capita (x_4)	1=x_4 <2.8 (million yuan); 2=2.8-8.0; 3= x_4 >8.0)
Industry Structure (x_5)	1= x_5 <1; 2= 1< x_5<3; 3= x_5>3
House spatial differences (x_6)	1= x_6<0.8; 2= 0.8< x_6<2; 3= 2< x_6<3.5; 4= x_6>3.5
Gender (x_7)	0= Male; 1= Female
Age (x_8)	1= x_8<35;2= 35< x_8<45; 3= x_8>45
Ethnicity (x_9)	1= Han, 2= Hui, 3=Man, 4= Meng, 5= Tujia, 6= Zhuang, 7=Yi, 8= Tibet, 9= Miao, 10=Dong, 11= Korea, 12= Yao, 13= Other
Household registration type (x_{10})	1= Other, 2= Agricultural, 3= Agricultural to resident, 4= Resident, 5=Non-agricultural, 6= Non-agricultural to resident
Education level (x_{11})	0= Never received, 1= Elementary, 2= Junior high, 3= Senior high polytechnic, 4=Junior College, 5= Undergraduate, 6= Graduate

Mobility range (x_{12})	1= Interprovincial, 2= Intercity, 3= Intercounty				
Variables	Sample size	Average value	Std.	Minimum	Maximum
Mobility Selection	2,000	.790	.408	0	1
Separation time	2,000	1.717	.846	0	3
Examination paper selection	2,000	2.785	1.870	1	7
Exam out of the domicile place	2,000	.203	.402	0	1
Population size	2,000	1017.491	560.688	312.6	2154.2
GDP per capita	2,000	2.698	.459	2	3
Industry Structure	2,000	1.993	.450	1	3
Spatial differences in houses	2,000	1.098	.297	1	2
Gender	2,000	.504	.500	0	1
Age Ethnicity	2,000	1.941	.806	1	3

3.3 Descriptive statistical analysis of the data

This paper uses the average data of 2018, with the cross-sectional dimension instead of temporal. On the one hand, cross-sectional data highlight spatial differences. Because the goal is to find how to improve the aggregation of mobile households under the new social trend, the timeliness of the early data is yet to be considered. Moreover, compared with panel data, it can observe more indicators. However, the effect if time change still cannot be excluded, being the limitation of this paper.

The proportion of not separated from their children is 78.95%, indicating that most of the migrant population chooses their children to live together. That means they receive local education in the migrated cities. The vast majority of all samples have a separation time of 1-10 years, accounting for 74.4% of the total. It can be seen that mobile people who live out of their hometown within 1-10 years have a higher probability of being separated from their children.

Table 2. Descriptive statistical analysis of the data

Household registration type	2,000	1.344	1.403	1	3
Education level	2,000	2.651	1.174	2	6
Mobility range	2,000	2.470	1.357	0	6

The education standard indicators include the selections of the college entrance examination papers and whether allowed to take the examination out of the domicile place. The cities with the largest proportion of adapting National I and National II papers. Nearly 80% of the cities is considered being allowed to take the examination out of the domicile place. The indicators of urban development level mainly include urban population size, GDP per capita, and so on.

Among samples, the gap between extremums of population size is significant. The spatial differences in house prices are all distributed between 0-2, with most of them less than 1. The indicators of personal characteristics mainly include gender, age, and household registration types, as shown in Table 2.

4. Empirical Analysis

Based on the research data in the previous section, we conducted a quantitative analysis, using Probit, Logit, and OLS models to explore the effects of two aspects respectively on the degree of household aggregation. The expression of the resulting probability distribution is

$$P_i = f\ (\alpha + \beta x_i) = f\ (\mathcal{E}_i) \tag{1}$$

Where α is a constant and β is the coefficient to be estimated, denoting the relationship between *P* and *x*. P represents the probability that the explained variable is affected by one of the explanatory variables.

4.1 Effects of education level on family cohesion of mobile population

As shown in Table 3, the marginal effect of selections of papers on whether to separate is 0.003, indicating that the choice of paper has a weak positive impact on separation. Therefore, parents who immigrate to non-national paper areas, such as Beijing, will be more likely to face separation from their children, thus, having a higher probability of increased separation time.

According to the results, the marginal effect of whether allowed to take the examination out of the domicile place and separation is -0.021, which means that the situation of not allowing the exam outside the place of residence will lead to more separation of the floating population and their children. Hence, the relationship between the two variables is negatively correlated. However, whether allowed to take the examination out of the domicile place is positively related to the time of separation, implying that not being allowed will shorten the time of separation.

Table 3. Empirical analysis-Probit model

	Coefficient		Std.		Marginal effect	
	Separation	Time	Separation	Time	Separation	Time
x_1	-.010$^{(.834)}$	.028$^{(1.722)}$	.017	.029	-.003	.002
X_2	-.071$^{(.819)}$	.153$^{(1.771)}$	.078	.141	-.021	.012
X_3	-2.14*10$^{-4(1.028)}$***	9.61*10$^{-5(1.704)}$	5.54*10^{-5}	9.80*10^{-5}	-6.13*10^{-5}	7.59e^{-06}
X_4	-.277$^{(.879)}$	.149$^{(1.401)}$	.069	.111	-.008	.012
X_5	-.257$^{(1.322)}$***	.169$^{(1.468)}$	.072	.117	-.074	.013
X_6	-.185$^{(1.001)}$	.164$^{(1.621)}$	.102	.197	-.053	.013
X_8	-.280$^{(1.367)}$***	.206$^{(1.424)}$**	.039	.068	-.079	.016
X_9	.092$^{(.686)}$***	-.047$^{(1.866)}$	.029	.030	.027	-.004

(*The numbers in parentheses are constant terms)

Among the control variables, age and ethnicity had a more significant effect on whether to separate from children, with a 28% difference between age categories and 9.2% difference between ethnic categories.

4.2 Effects of urban development level on the household aggregation of mobile population

As shown in Table 3, all four indicators of the level of urban development have negative effects on whether to separate. Specifically, each 1% increase in population size causes 0.002% of the mobile population to separate from their children. In areas with higher per capita GDP, higher output ratios of tertiary industry and secondary industry, and large spatial differences of house prices, the probability of separation will increase. Thus, the higher level of urban development increases the likelihood of living separately from their children.

For separation time, all four urban variables contribute positively to it, though population size is weak. Each 1% increase in GDP per capita, the ratio of the output, and spatial variation of house prices can lead to an increase in separation time of about 1.2%. It can be seen that the more developed the city is, the longer the separation time of the migrant population from their children.

Table 4. Parameters related to OLS and Logit models

	OLS Coefficient		Logit OR value	
	Separation	Time	Separation	Time
x_1	-.003$^{(.798)}$	.038$^{(1.612)}$***	.983$^{(3.939)}$	1.070$^{(22.314)}$
X_2	-.021$^{(.794)}$	.492$^{(1.617)}$***	.884$^{(3.848)}$	1.424$^{(25.148)}$
X_3	-3.70*10$^{-4(.854)}$***	2.15*10$^{-4(1.413)}$***	1.0002$^{(5.532)}$***	2.00*10$^{-4(21.660)}$
X_4	-.008$^{(.811)}$	.091$^{(1.471)}$*	1.401$^{(4.270)}$)	.338$^{(10.901)}$
X_5	-.072$^{(.932)}$***	.301$^{(1.117)}$***	1.470$^{(9.056)}$***	.385$^{(12.607)}$
X_6	-.0567$^{(.852)}$	.416$^{(1.259)}$***	1.465$^{(5.327)}$	.382$^{(17.708)}$
X_8	-.084$^{(.952)}$***	.255$^{(1.221)}$***	1.625$^{(10.638)}$***	.485$^{(11.188)}$**
X_9	.020$^{(.762)}$	-.014$^{(1.735)}$	.901$^{(3.001)}$**	-.105$^{(31.196)}$
X_{11}	-.005$^{(.802)}$	-.078$^{(1.910)}$***	.998$^{(4.038)}$	-.002$^{(26.879)}$
X_{12}	.007$^{(.777)}$	-.121$^{(1.914)}$***	1.030$^{(3.486)}$	.030$^{(25.508)}$

(*The numbers in parentheses are constant terms)

4.3 Robustness tests

Table 4 shows that the models' high significance levels are consistent with the core model. Comparing the marginal effects, the OLS coefficient, and the OR value, their features are similar. Consequently, we can conclude that the core model has robustness.

4.4 Heterogeneity Analysis

Zheng (2021) and Chen et al. (2020) point out that urban population size synergizes with economic development and industrial agglomeration. While Lin et al. (2019) deems that the return to education of the mobile population is significantly influenced by urban population size. Therefore, when it is used as an explanatory variable, it is necessary to explore the heterogeneity of this variable when it is involved as an explanatory variable. Thus, this subsection applies the Probit model to analyze the heterogeneity of household aggregation of the sample at different city sizes.

Table 5. Heterogeneity analysis of separation

	Overall	Size= (0,1000) (n=1,185)	Size= (1000,1500) (n=210)	Size= (1500, ∞) n= (605)
x_1	-.010$^{(0.542)}$ (.017)	.004$^{(0.838)}$ (.020)	- -	.023$^{(0.001)}$*** (.067)
X_2	-.071$^{(0.359)}$ (.078)	- -	- -	.349$^{(0.002)}$*** (.114)
X_4	-.227$^{(0.689)}$ (.069)	.140$^{(0.114)}$ (.089)	- -	- -
X_5	-.257$^{(0.000)}$***(.072)	-.335$^{(0.006)}$** (.121)	- -	.021$^{(0.860)}$ (.177)
X_6	-.185$^{(0.070)}$ (.102)	- -	- -	.021$^{(0.860)}$ (.177)
X_8	.206$^{(0.001)}$*** (.068)	-.189$^{(0.000)}$*** (.056)	-.190$^{(0.127)}$ (.125)	-.428$^{(0.000)}$*** (.071)
X_9	.092$^{(0.001)}$*** (.029)	.100$^{(0.003)}$** (.038)	-.427$^{(0.389)}$ (.496)	.031$^{(0.606)}$ (.061)
X_{11}	-.017$^{(0.467)}$ (.023)	-.061$^{(0.043)}$* (.033)	-.095$^{(0.209)}$ (.076)	.062$^{(0.114)}$ (.040)
X_{12}	.026$^{(0.604)}$ (.050)	.075$^{(0.291)}$ (.067)	-.144$^{(0.278)}$ (.132)	-.264$^{(0.013)}$* (.106)

Table 6. Heterogeneity analysis of separated time

	Overall	Size= (0,1000) (n=1,185)	Size= (1000,1500) (n=210)	Size= (1500, ∞) n= (605)
x_1	.028$^{(0.324)}$ (.029)	.001$^{(0.984)}$ (.033)	- -	-.074$^{(0.578)}$ (.134)
X_2	.153$^{(0.279)}$ (.141)	- -	- -	-.130$^{(0.592)}$ (.242)
X_4	.149$^{(0.177)}$ (.111)	-$^{(1.000)}$ (.143)	- -	- -
X_5	.169$^{(0.148)}$ (.117)	.191$^{(0.239)}$ (.163)	- -	-.021$^{(0.926)}$ (.232)
X_6	.164$^{(0.406)}$ (.197)	- -	- -	-.021$^{(0.926)}$ (.232)
X_8	.206$^{(0.002)}$** (.068)	.206$^{(0.019)}$* (.088)	.119$^{(0.522)}$ (.186)	.199$^{(0.140)}$ (.135)
X_9	-.047$^{(0.118)}$ (.030)	-.013$^{(0.718)}$ (.037)	- -	-.164$^{(0.006)}$** (.059)
X_{10}	-.071$^{(0.087)}$ (.041)	-.104$^{(0.041)}$* (.051)	.046$^{(0.783)}$ (.166)	-.046$^{(0.588)}$ (.086)

Table 5 and Table 6 show the grouped regression results of different models, with a total sample

size of overall, less than 10 million, 10 to 15 million, and more than 15 million. In Table 5, among different models, the coefficients of paper selection, exam out of the domicile place, industrial structure, age, ethnicity, education level, and mobility range are significantly different. In Table 6, coefficients of ethnicity and household registration type are significantly different. In addition, age has a positive effect on time at population less than 15 million, whereas the magnitude of the effect is not significantly different. Therefore, population size plays a moderating role in the relationship not only between the effects of paper selection, exam out of the domicile, etc., on separation but also between the effects of household registration type and ethnicity on separation time.

To sum up, heterogeneity existed in the process that examination paper selection, industrial structure, ethnicity, and so on influence mobile population separate from children and the time.

5. Results and Discussion

5.1 Results

This paper aims to delve into the effects of education level and urban development on the degree of household aggregation of the mobile population. The hypotheses are based on two aspects. Then the interrelationships between the variables are obtained through Probit models, followed by robustness tests to further confirm the hypotheses. Finally, the following results were obtained:

From the perspective of education level, the difficulty of college entrance examination papers varies from city to city, reflecting the differences in educational standards and competition. The policy of not allowing taking college entrance examination out of domicile place has separated mobile families, and the separation time tends to increase.

From the urban development level, cities with large populations are economically developed. What is more, the diverse constructures of industrial structures show large spatial differences in housing prices. It means that the mobile population is incapable of owning property and living stably in the working cities, discouraging families from immigrating entirely.

5.2 Discussion

Policies such as reducing the restrictions of the household registration system and the social welfare system, improving the policies related to the college entrance exams in different areas, etc., on the one hand, allow migrant children to receive education, have entrance examinations, and acquire higher education. Thus, barriers between educational systems and resources in diverse areas are broken down, promoting educational equity on a larger scale. On the other hand, they can perfect the social welfare protection system to release the living pressure of the mobile population and facilitate their family aggregation.

Furthermore, vigorously carrying out employment-related training for the mobile population to improve the current situation of the mobile population's low education degree and insufficient accumulation of human capital. Likewise, expanding the scales of employment platforms is also considered effective. Besides, implementing policies to promote social integration between the mobile population and local residents can grant equal social benefits to both the local and mobile population. It is also vital for building a harmonious and advanced society.

References

[1] Huang, J. &Zhu, M. (2021) A study on the difference of higher education returns between migrant population and local labor force. *China Population Science, 05*: 77 - 87+127.

[2] Sun Y. & Wang X. (2019). A study on the association between family reunion, ICT use, and urban integration of non-domestic population. *Journal of Harbin University of Commerce, 06*: 62 - 77.

[3] Hu, X. & Kang, Y. (2021). Analysis of family-oriented migration and urban settlement intention of the migrant population. *Statistics and Decision Making, 37* (19): 76 - 79. DOI: 10.13546/j.cnki.tjyjc.2021.19. 017.

[4] Zhu, P. (2021). Family aggregation-a comparative study of family urban-rural attributes and familial migration. *Social Development Research, 8* (02): 72 - 88+243.

[5] Cai, S. (2021). Knowledge mapping analysis of the study of family-based migration of Chinese population. *World Geography Research*, 1-14. Retrieved from http: //kns.cnki.net/kcms/detail/31.1626.p.20210915.1828.006.html.

[6] Yang, C., Bo, G. & Han, H. (2020). Urban settlement intentions of migrant population and influencing factors--An example from Anhui Province. *World Geography Research, 29* (06): 1136 - 1147.

[7] Chen, S., Wang, P., Liu, Y. & Li, G. (2019). Household perspective of urban settlement intention of migrant population and its influencing mechanism. T*ropical Geography, 39*(01):58-68. DOI: 10.13284/ j. cnki. rddl.003096.

[8] Ai, X., Chang, J. & Li, G. (2020). A study on the willingness to settle, income difference and household consumption level of migrant population. *East China Economic Management, 34*(01):94-100. DOI: 10.19629/j.cnki.34-1014/f.190710001.

[9] Li, C., Wang, X. & Che, Y. (2020). Research on social integration, willingness to settle in town and differences in household consumption level of mobile population. *Journal of Beijing Union University, 18* (03): 106 - 115. DOI: 10.16255/j.cnki.11-5117c.2020.0043. J.

[10] Chen, M., Wu, Y., Liu, G. & Wang, X. (2020). City economic development, housing availability, and migrants' settlement intentions. *Growth and Change* (3), DOI: 10.1111/grow.12416.

[11] Zheng, J., Wang, T. & Chen, H. (2021). Industrial Technology Economics, 40 (10): 52 - 61.

[12] Chen, D., Yan, Z. & Wang, W. (2020). Industrial Agglomeration and Urban Innovation. *China Population Science, 05*: 27 - 40+126.

[13] Lin, Y. & Geng, C. (2019). Effects of the Urban Population Scale of Floating Population on Education Returns. *Urban Problems, 02*:89-95. DOI: 10.13239/j.bjsshkxy.cswt.190212.