

Analysis of Settlement Intention of Floating Population in New Urban Areas Based on Decision Tree Model

Haixia Lai

Fuzhou New Area Management Committee, Fuzhou, Fujian, China

Abstract: In the initial stage of new urban area construction, it is urgent to introduce a certain amount of population to activate idle resources and promote the benign cyclical development of the economy. By using the CHAID decision tree method, this paper analyzes the characteristics of the population willing to settle in Fuzhou New Area in the early stage of construction, where urban construction is relatively advanced but population introduction is slow. The aim is to provide decision-making references for increasing the permanent population of the new urban area. The study found that the overall settlement intention of the floating population is not strong. Factors such as household registration migration, house purchasing intention, and children's education, which are based on family considerations, have a significant impact on the settlement intention of the floating population. Among them, the group willing to migrate with the whole family and whose children have reached school age has the highest settlement intention.

Keywords: New Urban Area; Decision Tree; Population Introduction; Settlement Intention

1. Introduction

A Survey conducted by the Urban and Small Town Reform and Development Center of the National Development and Reform Commission showed that, as of May 2016, there were over 3,500 new urban areas at the county level and above across China. According to a previous survey released by the center, the planned population of these new urban areas reached 3.4 billion. With the rapid urbanization in China, new urban areas have become important spatial carriers for implementing the new urbanization strategy in various regions. Most of these new urban areas,

including national-level new areas, have evolved from the demolition of townships and development zones. However, the speed of population aggregation has lagged behind the construction speed of new urban areas, leading to a shortage of population input and difficulties in sustaining commercial facilities in the early stages of new urban development. How to retain talent and achieve deep integration between industry, city, and people has become a key concern for every new urban area in its initial stage of construction. In recent years, some new urban areas in China have experienced high vacancy rates, affecting the sustainable economic and social development of these areas. Therefore, experiences in introducing population to new urban areas deserve more attention.

According to the Seventh National Population Census conducted in 2020, the proportion of China's urban population reached 63.89%, showing a year-on-year growth trend. Urbanization is considered a significant process of population redistribution, primarily driven by population migration [1]. After the initial rural-to-urban migration, a large number of China's migrant populations have been in a state of non-permanent circular migration between rural and urban areas for an extended period, forming a three-dimensional differentiation pattern of settling in the destination, returning to their hometown, and continuing to migrate, as well as a dual or even multiple livelihood strategy for migrant family members [2]. Previous studies have found that the factors influencing the diverse migration directions of the floating population include demographic characteristics such as age, marital status, gender, and education level [3], as well as the monthly income of the migrant population and their families, employment and occupational characteristics, social insurance, and the characteristics of the labor market in the destination (especially the status of labor-intensive industries in the economic

structure of the destination) [4], hukou (household registration) factors, social networks, social identity, social integration [5], equalization of public goods and services [6], life satisfaction, and regional differences [7,8]. Existing research has primarily focused on regions with relatively long development histories or conducted nationwide empirical studies [9-12], lacking generalizability for new urban areas with short construction periods, small local populations, and an urgent need to introduce population to stimulate urban vitality.

To promote the deep integration of industry, city, and people in new urban areas, further accelerate population introduction, and promote the comprehensive development of regional economy and society, this study takes the floating population in the built-up area of Fuzhou New Area as the research object. Based on the actual situation of the initial construction of new urban areas, an analytical framework is established from both individual and family perspectives, adhering to the concept of “people-oriented”. This study deeply analyzes the main influencing factors affecting the settlement of the floating population in new urban areas, and ranks the importance of these influencing factors based on a decision tree model. By identifying the key points for population introduction, this study provides reference significance for the population introduction and the promotion of deep integration of industry, city, and people in the initial stage of construction of other new urban areas across the country.

2. Data Sources and Research Methods

2.1 Data Sources

This study was based on a survey of the floating population who have worked in the built-up area of Fuzhou New Area for more than half a year. The survey was conducted online and randomly selected respondents. The questionnaire was well-structured and logically sound, with a Cronbach’s alpha coefficient of 0.927 and a KMO value of 0.943, indicating good reliability and validity. The survey content included: gender, age, marital status, workplace, education level, children’s

schooling situation, plans for household registration migration, willingness to settle in the new area, and ratings of the livability of the new area.

2.2 Research Methods

A decision tree CHAID model was established to link the respondents’ willingness to settle in the new area with potential related factors, screening for the significant characteristics of those willing to settle. The decision tree method divides data into training and prediction sets. The data is first classified based on the training set. If the independent and dependent variables added have no significant correlation, they will be excluded from the decision tree model. This process is repeated until the final decision set is formed. A decision tree consists of decision nodes, branches, and leaves. The first node is the root node, each branch is a new decision node or a leaf of the tree. In the process of traversing the decision tree from top to bottom, the influence of the independent variables at each level of nodes on the dependent variable decreases. If there is no statistically significant division at a certain node, the node stops branching and becomes a leaf node.

The CHAID decision tree method is a decision tree algorithm that analyzes and compares numerous classification rules based on the dependent variable set in the model to be constructed, finding the optimal classification variables and results, and thus determining which type of group is more likely to have this outcome [13]. The dependent variable can be nominal or numerical. The criteria for selecting classification variables for the two types of variables are based on the p-value of the χ^2 and F distribution, respectively. The stopping rules in the construction of the decision tree are all based on the statistical significance of the χ^2 or F distribution [14].

2.2.1 Variable setting

A total of 234 valid questionnaires were collected, among which 50 people were willing to settle in the new area, accounting for 21.4%. From both individual and family perspectives, variables were set as follows (Table 1):

Table 1. Variable Definition

Variable Type	Variable Label	Variable Definition
	Dependent variable	

Settlement intention	Whether you are willing to settle in the new area	1=Yes,0=No
	Independent variable	
Individual perspectives	Age	1=Under 18 years old,2=Between 18 and 25 years old,3=Between 26 and 35 years old,4=Between 36 and 45 years old,5=Between 46 and 55 years old,6=Between 56 and 65 years old,7=Over 65 years old
	Education	1=High school diploma or below,2=Bachelor,3=Master and above
	Marital status	1=Unmarried,2=Married
	Workplace	1=Government,2=Institutions,3=State-owned enterprise,4=Private enterprise,5=Joint venture,7=Foreign-funded enterprise
	Whether the salary can meet the needs of life	1=Yes,0=No
	Whether there is a residence in the new area	1=Yes,0=No
Family perspectives	Whether you are willing to buy a house	1=Yes,0=No
	Children's schooling	1=Not of school age,2=Have reached school age but are not attending school,3=Have reached school age and are enrolled in primary or secondary school in the new area,4=Have reached school age and are attending primary or secondary school in another area,5=Pursuing a university degree or higher,6=Have completed their studies,7=No child
	Household registration migration plans	1=Migrate the whole family,2= Only migrate themselves,3=Migrate part of their family members,4=Unwilling to migrate,5=No idea

2.2.2 Establishment of the decision tree model
The CHAID algorithm was used to analyze the relevant demographic and sociological characteristics of the floating population in the built-up area of the new area. The classification rules were as follows: (1) Tree growth: the significance level of branch splitting $\alpha = 0.05$; (2) Tree pruning: the growth level of the decision tree was set to 3 layers, and the minimum sample size of the parent node and child node was 10. If the sample size at a node does not meet the requirements, no further division will be performed.

3. Analysis of the Willingness of the Floating Population to Settlement Intention and Its Influencing Factors

3.1 Current Situation of the Willingness of the Floating Population to Settlement Intention

The survey results showed that the willingness of the floating population in the new area to settlement intention was relatively weak. Only 21.4% of the respondents were willing to settle in the new area, while 15.8% chose to live in other cities after working in the new area for a

period of time, 16.7% chose to return to their hometown after working in the new area for a period of time, and 46.1% chose to work in the new area but live in other areas. Compared with the willingness of over 50.0% of the floating population in large cities [15], the willingness of the floating population in the new area is very low.

Referring to previous studies, respondents who chose "willing to settle in the new area" in the questionnaire were classified as having the intention to settlement intention, while those who chose "to move to other cities after working in the new area for a period of time," "to return to their hometown after working in the new area for a period of time," and "to work in the new area but live in other areas" were classified as having no intention to settlement intention. The comparison found that individual factors and family factors had different degrees of influence on the willingness of the existing population in the new area to settlement intention. Among them, family factors had a significant statistical impact on the willingness to settlement intention (Table 2).

Table 2. Population Profile with and without Settlement Intentions

Variables		The percentage of people without settlement intention(%)	The percentage of people with settlement intention(%)	P Value
Sex	Male	47.3	68.0	0.009
	Female	52.7	32.0	
Age	Under 18 years old	0	0	0.033

	Between 18 and 25 years old	8.2	8.0	
	Between 26 and 35 years old	45.7	36.0	
	Between 36 and 45 years old	28.3	50.0	
	Between 46 and 55 years old	14.1	6.0	
	Between 56 and 65 years old	3.8	0	
	Over 65 years old	0	0	
Education	High school diploma or below	18.5	28.0	0.322
	Bachelor	63.0	54.0	
	Master and above	18.5	18.0	
Marital status	Unmarried	29.9	22.0	0.272
	Married	70.1	78.0	
Workplace	Institutions	47.8	30.0	0.002
	State-owned enterprise	28.3	22.0	
	Private enterprise	23.4	46.0	
	Joint venture	0.5	0	
	Foreign-funded enterprise	0	2.0	
Whether there is a residence in the new area	Yes	18.5	34.0	0.018
	No	81.5	66.0	
Children's schooling	Not of school age	13.6	14.0	0.027
	Have reached school age but are not attending school	2.2	2.0	
	Have reached school age and are enrolled in primary or secondary school in the new area	1.1	8.0	
	Have reached school age and are attending primary or secondary school in another area	33.2	46.0	
	Pursuing a university degree or higher	7.1	2.0	
	Have completed their studies	4.3	0.0	
	No child	38.6	28.0	
Whether you are willing to buy a house	Yes	14.1	62.0	0.000
	No	85.9	38.0	
Household registration migration plans	Migrate the whole family	1.6	44.0	0.000
	Only migrate themselves	10.9	14.0	
	Migrate part of their family members	13.0	28.0	
	Unwilling to migrate	50.5	8.0	
	No idea	23.9	6.0	

3.2 Analysis of Influencing Factors on the Willingness to Settlement Intention

Among the demographic variables, gender, age, highest education level, marital status, monthly net income after tax, whether the salary can meet the needs of life, and whether there is a residence in the new area were all automatically excluded from the model, while family household registration migration plans, willingness to purchase a house, children's schooling situation, and workplace were the most significant factors affecting residents' willingness to settlement intention and were retained in the model. The decision tree has 3

layers (Figure 1), and the model has a prediction accuracy of 89.3% with a standard error of 0.02, indicating that the model has good classification ability.

The analysis results show that the influence of individual factors such as gender, age, highest education level, marital status, and salary level of the floating population in the new city on residential choices is not as significant as that of family factors such as household registration migration, children's schooling, and willingness to purchase a house, indicating that people consider more from a family perspective when making residential choices.

Among them, the correlation between household registration migration plans and the willingness to settlement intention is the highest. Among those who are willing to migrate the whole family, the proportion of people who choose to settle in the new area is 88%, while the proportion of those who only migrate themselves or part of their family members and are unwilling to migrate is only 32.3% and 4.9%, respectively. People who are willing to migrate the whole family show a stronger willingness to settlement intention, which is consistent with the research results of Lin Sainan and others [15], and families who are willing to migrate the whole family and whose children have reached school age have the highest willingness to settlement intention, reaching 100%. The willingness of those who choose to migrate only their own household registration or only part of their family members' household registration is most affected by the workplace. The willingness of employees of private enterprises is higher than that of employees of state-owned enterprises, institutions, and joint-stock enterprises. This is

mainly because the latter often come to work in the new area due to the relocation of their units or the establishment of branches, but their families and social interactions are still in their original place of work, and therefore they choose to separate their work and residence. This is somewhat different from the view of scholars that the stability of employment is an important consideration for the floating population to choose whether to settle in a city, and good employment prospects and career development can promote their urban settlement, indicating that it is necessary to further consider the employment form of employees and the nature of the workplace in the relocation destination. Among those who are unwilling to migrate their household registration, the willingness to purchase a house has a significant impact on whether they are willing to settle in the new area. The proportion of people with the intention to purchase a house who choose to settle in the new area reaches 21.4%, which is significantly higher than that of those without the intention to purchase a house.

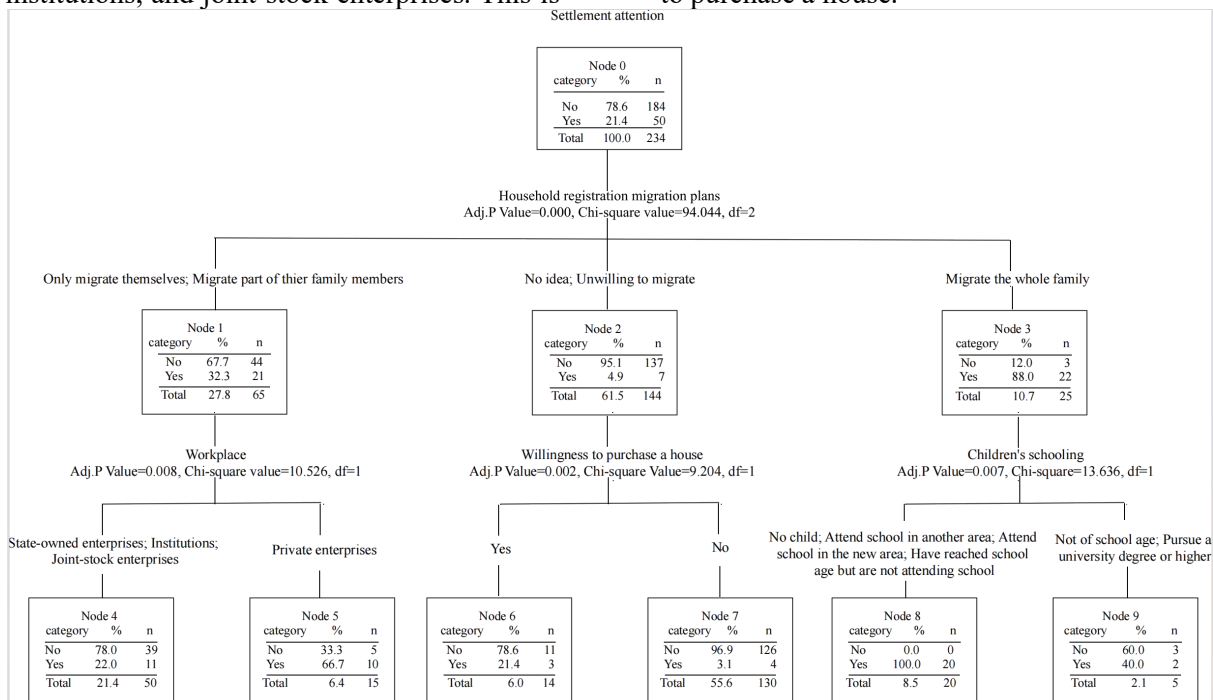


Figure 1. The CHAID Decision Tree Model

4. Conclusion and Discussion

In the early stages of new urban area development, the pace of industrial development, population introduction, and urban construction is inconsistent, mainly manifested in the advanced urban construction,

but the relatively lagging industrial development and population introduction. The problem of insufficient population introduction has affected regional economic development, and urban development planning urgently needs to shift from pursuing the integration of industry and city to the integration of industry,

city, and people. Studies have found that the willingness of the floating population to settlement intention is generally weak, and rational choices based on the family have a profound impact on the willingness of the floating population to settlement intention. The main conclusions and discussions of this article are as follows:

(1) Improving the household registration policy is a powerful tool to attract the floating population to settlement intention. The survey shows that the population willing to migrate the whole family and whose children have reached school age has the highest willingness to settlement intention. Currently, various places in China are launching a “competition for talent” by lowering the threshold for household registration, including restarting the purchase of housing for household registration, exploring rental housing for household registration, increasing the points for buying a house, and expanding the scope of dependent household registration. However, simple household registration system reforms, granting local household registration to the floating population, cannot effectively promote the settlement and integration of the floating population into the city. Against the background of the gradual relaxation of the household registration policy, taking attracting family household registration as the starting point, improving the mechanism for the household registration of accompanying family members, and perfecting supporting services such as childcare, leisure and entertainment, and employment consultation for accompanying family members will help the floating population integrate into the city. It is suggested that pilot projects be carried out in areas with concentrated populations in the early stage to drive the entire region’s public service resources and commercial resources while introducing population.

(2) Leveraging high-quality education resources to play a leading role in attracting population. Attracting the children of migrant workers to attend school can achieve a rapid increase in population in the short term. To fully leverage the leading role of education, it is first necessary to increase the number of high-quality schools at different education stages, as well as the number of new students and transfer students at the compulsory education stage, to create quality conditions

for the children of migrant workers to study in the new area. Secondly, it is necessary to introduce policies such as housing security for the families of students to solve the housing problems for families accompanying their children to study.

(3) Strengthening housing security helps attract the floating population to settlement intention. The survey found that those who are unwilling to register but have the intention to buy a house have a significantly higher willingness to move in than those who do not have the intention to buy a house. The decision to buy a house reflects the floating population’s recognition of the local living and working conditions and community public service facilities, and thus has a higher willingness to settlement intention. Therefore, improving the supply and guarantee system of various types of housing such as public rental housing, talent housing, and commercial housing, increasing the types of housing suitable for families, and improving community supporting services so that more people can afford to buy houses and live comfortably are another core measure to attract the floating population in the new urban area to settlement intention.

The shift from industry-city integration to industry-city-people integration is an effective way to promote high-quality development of new urban areas. This survey found that the coordinated efforts of new urban areas in housing, education, family household registration migration, and life support will help attract the floating population to truly settle in new urban areas, form a certain population base, and improve the utilization rate of urban commercial facilities and public resources. The shortcomings of this study are that due to the early stage of new urban area construction, there are few employees of enterprises entering, and the survey objects are not diversified. Factors such as education level, income, and educational attainment, which have been shown in previous studies to have a significant impact on the willingness of the floating population to settlement intention, have been excluded from the decision tree model, and only a few factors with significant impacts have been retained. As the construction of new urban areas continues to advance, further analysis and comparison of the factors that attract the floating population

to settlement intention in different periods will be carried out, including a more systematic analysis of individual and family factors.

References

- [1] Zhu Yu, Ding Jinhong. Population Geography in China over the Past 40 Years - Progress in an Interdisciplinary Research Field. *Progress in Geography*, 2017,36(4):466-482.
- [2] Lin Liyue, Zhu Yu. The Spatial Type Selection of Initial Migration of Floating Population and Its Influencing Factors: Based on a Survey in Fujian Province. *Geographical Science*, 2014,34(5): 539-546.
- [3] Zhu Yu, Chen Wenzhe. The settlement intention of China's floating population in the cities: Recent changes and multi-faceted individual-level determinants. *Population, Space and Place*, 2010,16(4): 253-267.
- [4] Wang Chunlan, Ding Jinhong. Analysis of the Influencing Factors of Urban Residence Intentions of Floating Population. *Southern Population*, 2007, 22(1): 22-29.
- [5] Zou Jing, Chen Jie, Wang Hongwei. How social integration affects the residential choices of the floating population - Based on the 2014 national floating population monitoring data research, *Journal of Shanghai University of Finance and Economics*, 2017,19(5): 64-79.
- [6] Zeng Tonggang, He Xionglang. Settlement intention of westward migrants in China and its influencing factors. *Progress in Geography*, 2024, 43(11): 2213-2225.
- [7] Liu Yuqi, Liu Ye, Li Zhigang. The Settlement Intention of New Immigrants in Chinese Cities and Its Influencing Mechanism. *Scientia Geographica Sinica*, 2014, 34(7): 780-787.
- [8] Lin Liyue, Zhu Yu, Xu Lifang. Floating Population's Environmental Perception of Inflow Places and Its Influence on Settlement Intention: A Survey Based on Fuzhou City. *Human Geography*, 2016, 31(1): 65-72.
- [9] Yang Xue, Wei Hongying. New Characteristics and Influencing Mechanism of Long-term Residence Intention of Floating Population. *Population Research*, 2017, 41(5): 63-73.
- [10] Wu Kaize. Housing and Employment: A Study on the Influence Factors of Settlement Desire of Young Talents from Outside the City in Large Cities. *Journal of East China Normal University*, 2023, (4): 162-175.
- [11] Chen Yanru, Gu Yue, Song Weixuan, et al. Differences in Residence Intention of Floating Population in Yangtze River Delta and Its Influencing Factors. *Economic Geography*, 2022,42(9):59-67.
- [12] Zhan Dongsheng, Zhang Wenzhong, Dang Yunxiao, et al. Perception of Urban Livability of Chinese Floating Population and Its Influence on Settlement Intention. *Progress in Geography*, 2017, 36(10): 1250-1259.
- [13] Zhang Yuejiao, Xu Xin, Dai Li, et al. Application of CHAID model in the screening of influencing factors of neonatal low birth weight in Gongyi City. *Modern Preventive Medicine*, 2012, 39(1): 18-20.
- [14] Xu Lei, He Jia, Meng Hong, et al. Decision tree technology and its application in medicine. *Magazine Magazine*, 2004, 17(2): 161-164.
- [15] Linkannan, Liang Qi, Li Zhigang, et al. The impact of "family migration" on the willingness of the settlement of the migrant population in small and medium-sized cities-take Wenzhou as an example. *Geographic research*, 2019, 38(7): 1640-1650.