



ARTICLE



<https://doi.org/10.1057/s41599-023-02345-y>

OPEN

The impact of social exclusion and identity on migrant workers' willingness to return to their hometown: micro-empirical evidence from rural China

Haiping Xu¹, Wenjia Wu^{1✉}, Chuqiao Zhang^{1✉}, Yibo Xie^{1✉}, Jinge Lv¹, Shahzad Ahmad¹ & Zengrui Cui¹

The return of rural migrant workers to employment is an endogenous force that drives the development of the rural economy and plays a vital role in rural revitalization. Using data from the China Migrants Dynamic Survey (CMDS) conducted in 2017, this study comprehensively examines the influence of social exclusion and the role of identity on migrant workers' willingness to return home (MWRH). The findings indicate a significant impact of social exclusion on MWRH, and this relationship remains robust even after employing instrumental variables to address endogeneity concerns. Identity emerges as a crucial mediating factor through which social exclusion affects MWRH, demonstrating a partial mediating effect. Notably, the impact of social exclusion on MWRH is particularly prevalent among rural residents in the western and northeastern regions, the older generation, those with lower incomes, and those with limited intra-provincial mobility. Promoting a positive image of migrant workers and encouraging their pursuit of an improved quality of life can help alleviate social exclusion, strengthen their sense of identity, and foster deeper roots in cities. Furthermore, providing supportive platforms for returning migrant workers can enhance their active willingness to return home.

¹International Business School of Hainan University, Haikou, China. ✉email: 20095138210012@hainanu.edu.cn; 21111202000021@hainanu.edu.cn; 994325@hainanu.edu.cn

Introduction

The rapid advancement of industrialization and urbanization, improved infrastructure, access to public services, and effective allocation of factor resources in urban areas attract a large number of rural surplus labor. The rural-urban migration from the agricultural sector has provided essential labor and human capital for the urban non-agricultural industry, thus promoting the effective allocation of resources (Wang, Li, 2021). By the end of 2021, the number of migrant workers reached 292.51 million, an increase of 2.4% over the previous year. Among them, 171.72 million were migrant workers, an increase of 1.2% over the previous year¹. However, the vast majority of young and middle-aged farmers in poor rural villages move to urban areas in search of urban economic opportunities, resulting in old and young women and children becoming left-behind people. President Xi Jinping proposed the policy of implementing the rural revitalization strategy in the reports of the 19th National Congress of the Communist Party of China, which means that China gives priority to the acceleration of the modernization of agriculture and rural areas. The sustainable development of the countryside is the people's prosperity, and it is vital to introduce and retain the talents of rural households, as they are the main force of rural revitalization. The rural labor force is the key to revitalizing the countryside. The No. 1 central document in 2020 emphasizes that "we will further implement actions to cultivate innovative and entrepreneurial leaders in rural areas, bring into and vigorously encourage and support migrant workers to return to their hometowns to start their businesses". The No. 1 central document in 2022 also points out that "promoting the construction of entrepreneurial parks for people returning to their hometowns and implementing various support policies"².

However, most migrant workers in China currently confront difficulties at the cultural and psychological levels in urban integration (Ai et al., 2021). With the social transformation in urban areas and the complex results of economic and social relations, the issues of social exclusion with new transitional characteristics have arisen in many places. Many company workers and citizens were prejudiced against migrant workers who move to the city and discriminate against them in attitude and behavior (Li, 2011), which will undoubtedly impact the willingness of migrant workers to return home. In 2020, the CPC Central Committee and the State Council stated that the social mobility channels for labor and talent should be further opened up, urban and rural labor should be fully guaranteed equal employment rights, and reduce the phenomenon of identity discrimination (Wang, Li, 2021). In reality, migrant workers move from rural to urban to earn a livelihood, and their working and living places have changed.

Therefore, based on the perspective of identity, this paper studies the influence of social discrimination on migrant workers' willingness to return home (MWRH), which is of theoretical reference value and practical significance for understanding social discrimination among migrant workers and identifying their urban identity, as well as for a more comprehensive understanding and promotion of integrated urban-rural development and comprehensive rural revitalization. Following Sayce (1998) and Kaber (2000), we use the terms "exclusion", "discrimination" interchangeably. However, it is acknowledged that each of these terms can be defined in different ways and may contain a variety of components. The rest of the article is organized as follows. Section "Literature review and research hypotheses" outlines the related literature and research hypotheses; section "Data sources, variable selection, and econometric model construction" provides the data sources, variable selection, and econometric model construction; section "Results" presents the

empirical findings; section "Robustness tests" examines the robustness tests; section "General discussion" offers concluding remarks.

Literature review and research hypotheses

Influencing factors of MWRH. Many studies regarding the factors influencing migrant workers returning home can be roughly divided into two aspects: Firstly, from the "pull" of rural areas, research has highlighted the environment of hometown, favorable policies, natural environment, comfortable housing, the development degree of non-agricultural industries in rural areas and cost of living in hometown. For example, Xu (2014) takes nearly 400 young generation migrant workers in Shaanxi Province and finds that the income level of family farming, policy support, and social climate in their hometown significantly positively affect their willingness to return home. Secondly, from the perspective of urban "push", study the threshold of urban household registration (Zhang et al., 2020), the burden of children's education (Jiang, 2009), the pressure to adapt to the city (Cai, Wang, 2007), the lack of non-agricultural experience (Vera Junge et al., 2015), the institutional arrangements related to social security for migrant workers (Qin & Zhou, 2014), social network exclusion (Pan & He, 2017), urban identity (Wang, Li, 2021), and income (Zhu et al., 2010) on MWRH. Li and Long (2009) of the survey data of migrant workers conclude that education is the most significant factor affecting the willingness of migrant workers to stay or return to rural areas. Meanwhile, working hours and marital status also impact the willingness.

Social exclusion and MWRH. The megacities' rich resources attract a large group of rural migrant workers, and the migration of the agricultural population to cities is the primary trend of population movement in China. The migration of poor agricultural households significantly affects raising their income and escaping rural poverty (Wang, Li, 2021). Some scholars consider the economic process of migrant workers in the city to be a social integration process. The primary challenge encountered by the migrant population is how to integrate into the social life of the destination. Their social identity will be differentiated based on individual characteristics and structural factors at the relocation site. In particular, the outcome of the differentiation is determined mainly by the interaction between migrants and the local society (Alba, Nee, 1997). The social integration and social identity of migrant workers are essential to building social harmony in the city. However, integrating into the city is not a one-way process of migrant workers but a two-way interaction between the workers and citizens. This requires the adaptation of the workers to the city and their transformation, as well as the acceptance of them by the city (Zhang, 2007).

From the conceptual point of view, deviant behavior against them because they belong to a particular class or group can lead to social discrimination (Wang, 2013). Social discrimination is often associated with conflicts of interest (Mayda, 2005), cultural perceptions, and social rules (McDaniel et al., 2011; Sides and Citrin, 2007).

There have been previous studies on social discrimination. The Social Exclusion theory of Frank Parkin, a representative of New Weberian, claims that factors such as family background, language, and religious beliefs can be the reason for excluding one group from another (Parkin, 1983). In addition, Allport et al. (1954) discover that one group's negative views (disgust, boredom) and concepts towards another group may cause discrimination. Starting from the 1980s, Europe began to conduct comprehensive and systematic studies on social discrimination,

proposing the formulation of related social policies to alleviate the problems of discrimination. Chinese research on social discrimination has mainly focused on exploring the causes, effects, and solutions. According to some scholars, the cause for the social discrimination of migrant workers includes institutional factors and economic status disparity (Li, 2011; Li, 2004). In terms of the effects, Ai et al. (2021) argue that social discrimination affects migrant workers' willingness to settle in cities. Some scholars believe that social discrimination significantly impacts rural cultural identity (Zhang, 2018).

In recent years, after migrant workers flow into cities, social discrimination has become increasingly prominent. Many enterprises, workers, and residents in cities have discrimination and prejudice against migrant workers in their behaviors and attitudes (Li, 2011). On the one hand, migrant workers exposed the dilemma of discrimination and social integration in urban-rural integration and new urbanization. According to the theory of Migrant Assimilation, social discrimination against migrant workers is likely to weaken their willingness to settle down in the city, leading to their return home (Ai et al., 2021). On the other hand, migrant workers leave the countryside for a long time and no longer engage in agricultural activities. They neither have many rights in the countryside nor the same treatment as urban dwellers (Chen, 2005) and become "rootless non-citizens". If the effects of social discrimination cannot be overcome, migrant workers will desire to return to their hometowns.

Hypothesis 1: Social discrimination has a significant positive effect on MWRH.

Identity and MWRH. The word "identity" originates from Philosophy and is a deep-seated questioning of the individual's existence and the meaning of life (Zhang et al., 2012). In addition, from the Etymology, identity has two meanings, one is sameness, and the other is otherness. Meanwhile, previous research has also documented the definition of identity. For instance, Deaux (1993) indicates that identity is an individual's perception of the group. Similarly, Stets et al. (2003) find that the essence of identity is providing a set of behavior standards. In the study of migrant population, Li et al. (2012) define the identity of the floating population as "the migrant's perception of their identity (urban people, rural people, foreigner, migrant worker), which is manifested by a sense of belonging and thinking about who they are. Hou et al. (2016) claim that identity is a subjective psychological willingness to identify, which is the identification of local citizens and their belonging to the local urban space.

However, identity is essential for a person. Because a specific identity corresponds to a specific philosophy, culture, and norms, it can effectively guide individual behavior and decision-making, and provide a reference standard for social activities (Blunt et al., 2003). Moreover, identity is an important dimension reflecting the social integration of the population (Hou et al., 2016). The floating population can truly integrate into the local society only after completing the psychological transformation of identity (Li et al., 2012; Cui (2012)). The individual's identity in the group and sense of social integration determines their ability to engage in productive activities (Basu, 2013). Identity is a crucial process of migrant social integration (Lin et al., 2022). With the accelerated urbanization in China, the migration of the agricultural population to cities is the primary trend of population movement (Wang, Li, 2021). In recent years, a significant number of scholars have begun to focus on the identity of migrant workers and how to realize real urban-rural integration, which depends on whether migrant workers become city builders and new citizens (Chen et al., 2017). A stronger sense of urban identity among migrant workers pushes the

governments of migrant cities to make necessary adjustments in employment policies, fundamental rights and interests, and public services. Hence, those migrant workers enjoy the same rights as citizens (Yang et al., 2016). Therefore, they are more likely to reside in a city, and their willingness to return home will subsequently decrease.

Hypothesis 2: There is a significant negative effect of identity on MWRH.

Social exclusion and identity. Under the new urbanization, many people move from rural to urban areas, which can easily lead to conflicts between natives and newcomers (Li, 2017). The relationship of mutual antagonism and exclusion can reinforce the group's culture and identity (Zhang, 2018). Social discrimination is a form of exclusivity. Social discrimination is not only reflected in the "natives-newcomers" conflict but also in the institutions and markets. As a result of the dual labor market, many immigrants are excluded from the mainstream market and work in informal occupations (Portes, 1994; Portes & Zhou, 1996). Bertrand and Duflo (2017) claim that discrimination is the unfair treatment of a minority of members who do not share a characteristic compared to the majority of members sharing the characteristic in similar situations.

Moreover, inflow migrants interact with the residents in the regions, and the effect of interactions would affect the sense of social identity and then affect their choice between returning home and settling in the city. Akerlof and Kranton (2000) systematically apply Social Identity theory to gender discrimination in the labor market.

In China, the existing literature on the relationship between social discrimination and identity mainly focuses on the perspective of belonging. Firstly, migrant workers can only be employed in the secondary labor market without labor security, which is below their psychological expectations. Thus, they face difficulty developing a sense of belonging to the city. In addition, there is discrimination against migrant workers in the city, sometimes city residents do not accept the workers, and social interaction is characterized by "involution". However, migrant workers rarely communicate with city residents and cannot integrate into city life because they do not think they belong there (Shi, 2011). Wang (2013) finds that local discrimination significantly affects migrant workers' sense of belonging. Similarly, Chen et al. (2017) conclude that the more pleasant the neighborhood, the less discrimination there is, and the stronger the sense of belonging of migrant workers.

Hypothesis 3: Social discrimination has a significant positive effect on MWRH through identity.

The pathway map see Fig. 1.

Data sources, variable selection, and econometric model construction

Data sources. This study is based on the database of China Migrants Dynamic Survey (CMDS) by the China National Health and Family Planning Commission in 2017, which covers the basic information about migrant worker families and individuals, mobility and employment characteristics, social integration, family planning, and public health. The sample covers 31 provinces nationwide among the floating population. The probability proportional sampling (PPS) method was used. The floating population is aged 15–59 who lived in the local area one month before the survey was selected as the investigation object through random sampling.

According to the selected variables, urban samples, the "do not know" and "inapplicable" samples are deleted. A total of 47,011 samples were collected in the study.

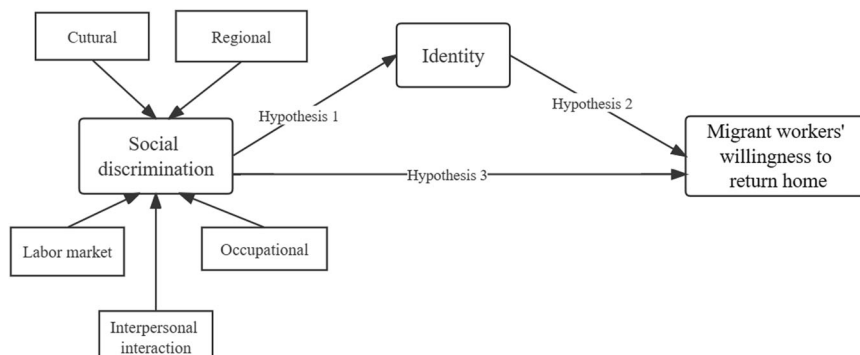


Fig. 1 Pathway map based on the hypothesis. The relationship between the three hypotheses is shown in Fig. 1. Hypothesis 1 explores the relationship between social exclusion and MWRH. Hypothesis 2 represents the connection between identity and MWRH. Hypothesis 3 represents the role of identity in social exclusion and MWRH.

Variables

Explained variable. The main explained variable is MWRH. We formulated it based on relevant questions in the CMDS database. The first is a question about whether migrant workers intend to stay in urban area? The second one is about staying here or returning hometown or elsewhere when they do not plan to stay here? If respondents choose “Yes” in the first question, it is considered that they have the intention to settle down. If they select “No”, proceed to the second question. If they chose “Return home” in the second question, they are considered to be willing to return home. “Have they will to return home” is assigned a value of “1”; the others are “0”.

Explanatory variables. The explanatory variable is social discrimination. Zhang (2018) finds that social discrimination mainly includes the discrimination of migrants by residents (based on interests or culture), labor market discrimination, and institutional discrimination. Similarly, Ai et al. (2021) measured social discrimination from four dimensions: interpersonal interaction, social security, labor market, and cultural discrimination. Combined with the summary of existing studies, this study measured social discrimination by five dimensions: interpersonal interaction discrimination, labor market discrimination, job discrimination, cultural discrimination, and regional discrimination.

(1) Interpersonal interaction discrimination: Migrant workers’ interactions with urban residents impact their social identity (Zhang, 2007). This paper measures whether migrant workers have formed an excellent social network in the city by “Who do you hang out with the most in your spare time?”. Of these, “Countrymen” and “Other local people” as residents and assigned a value of “0”, “Countrymen,” “Other foreigner,” and “Rarely interacts with people” are assigned a value of “1”.

(2) Labor market discrimination: The institutional roots of social discrimination against migrant workers are the urban access and employment systems (Li, 2011). The high rural-urban disparity in the labor market can force many migrant workers into informal occupations leading to low-income levels and an unemployment crisis. This study measures whether migrant workers are excluded from the urban labor market by “what kind of labor contract do you sign with the employer?” “Failure to sign labor contracts” and “completion of one-time work tasks” are regarded as discrimination and assigned a value of “1”; the others are “0”.

(3) Occupational discrimination: As the main content of objective stigmatization, discrimination is manifested as language discrimination, behavior discrimination, and occupation discrimination (Yu, 2006). Occupations are roughly divided into two types: standard and non-standard. The former is characterized by

high wage levels, excellent working conditions, stable employment, and perfect social welfare, the latter is characterized by low wages, poor working conditions, unstable employment, and inadequate welfare security (Li et al., 2012). Migrant workers are born in the countryside and are artificially separated by occupation, so they can only engage in those dirty, tired, complex, and poor jobs (Cheng, Yin, 2004). Therefore, the current occupational choices of migrant workers may also reflect whether they are discriminated against. Referring to the article by Ai et al. (2021), this study uses “Your current occupation” to describe whether they have experienced occupational discrimination. 1 for if migrant workers engage in informal occupations (courier, cleaner, builder, housekeeper, etc.), and 0 otherwise.

(4) Cultural discrimination: The significant rural-urban disparity leads to less communication between urban and rural residents. Rural identity makes migrant workers have obstacles in urban integration, while urban cultural identity plays a promoting role in this process. Therefore, in this study, “Do you agree with the statement that it is important for me to follow the customs of my hometown?” and “Do you agree with the statement that there is a big difference between my health habits and local citizens?” are measured by migrant workers’ acceptance of mainstream local values. “Agree” is assigned a value of 1, otherwise 0.

(5) Regional discrimination: As an ancient civilization deeply influenced by thousands of years of Confucian culture, “family Culture” has always been one of the core cultures of China, with regional characteristics. The “acquaintance society” is formed by “family Culture,” a tradition of humane interaction, leading to a preference for the local worker (Zhang, 2018). Thus, migrant workers have become “strangers” and are discriminated against by the locals. Regional discrimination makes the wage rate of the stranger much lower than that of the local population. This paper measures whether migrant workers are subject to regional discrimination by “Do you agree with the statement that local people look down on strangers?” “Agree” is assigned a value of 1, the others are 0.

Referring to Shi (2013), the scores of the above five dimensions are summed by equal-weighted summation. Thus, a comprehensive social discrimination index is constructed as the independent variable of this study. The range of values is 0–5, and higher values indicate the severity of discrimination.

Mediating variable. Mediating variables refer to the mechanisms by which the explanatory variables affect the explained variables. This paper takes “identity” as the mediating variable to further explore whether social discrimination affects MWRH by influencing their local urban identity. The first question is, “I feel I am a local”. The second question is, “I think the locals would like to

Table 1 Variable definitions.

Variables	Description and evaluation
MWRH	0 = Not returning home; 1 = Returning home
Social discrimination	0 = No discrimination; 1 = One discrimination; 2 = Two discriminations; 3 = Three discriminations; 4 = Four discriminations; 5 = Five discriminations
Identity	1 = Low sense of identity; 2 = Average identification; 3 = High level of identification; 4 = Highest sense of identity
Control variables: individual level	Gender 0 = Female; 1 = Male
	Ethnicity 0 = Han; 1 = otherwise
	Political affiliation 1 = Chinese Communist Party member; 0 = otherwise
	Marital status 1 = have spouse; 0 = otherwise
	Education 1 = Primary level and below; 2 = Junior High School; 3 = Senior High School; 4 = University and above
	Age Actual age
	Nature of housing 0 = No owner-occupied housing; 1 = Have purchased your own home
Control variables: family level	Income Take the logarithm of the average monthly household income
	Num. of family members living together Actual number of members
	Availability of contracted land 1 = Yes; 0 = No
	Availability of homestead 1 = Yes; 0 = No
	Are there difficulties at home 1 = Yes; 0 = No
Control variables: mobility level	Num. of cities on the move Number of cities actually mobile
	Flow range 1 = Cross-provincial; 2 = Cross-City; 3 = Cross County

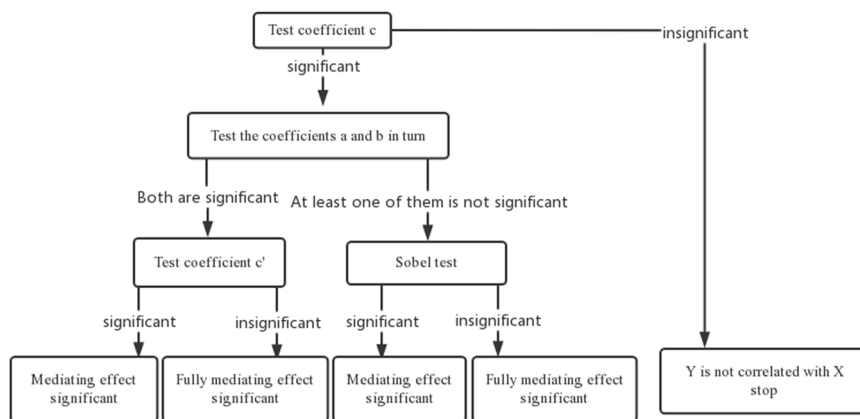


Fig. 2 Mediating effect test program. The independent variable is X, the dependent variable is Y, and the mediating variable is M. The specific mediating effect test steps are shown in Fig. 2. It should be noted that when $c' < c$, and the coefficients a, b, and c' are significant, it means that the intermediary variable plays a partial mediating role, while when $c' < c$, but c' is not significant, it plays a fully mediating role.

accept me as one of them”. The third question is, “I like the city where I live now”. The fourth question is, “I would like to integrate with and be one of the locals”. All items were presented on a 4-point scale (1 = “ Strongly agree” to 4 = “ Totally agree”). Higher values indicate a higher level of identity.

Control variables. Basic demographic and mobility variables were included as control variables, including age, gender, ethnicity, political affiliation, marital status, education level, monthly household income, household size, whether there is contracted land in the village, whether there is a home base in the village, whether there are difficulties in the household, the number of mobile cities, and the range of flow. These variables are commonly used in previous studies focusing on the psychological status of migrants in China (Wang, Li, 2021; Ai et al., 2021). A summary of the variables is described in Table 1.

Econometric model construction. With reference to the literature on the mediation effect test, this paper adopts the stepwise causality of Baron and Kenny (1986) and Wen, Ye (2014).

In addition, this paper validates the mediating effect by the Sobel test, which is re-validated using the Bootstrap test. The error rates of type I and type II in the mediation effect test of this method are relatively low, which tests both partial and complete mediating effects. The independent variable is X, the dependent variable is Y, and the mediating variable is M. The specific steps of the mediation effect test are shown in Fig. 2. Note that when $c' < c$ and the coefficients a, b, and c' are significant, it indicates that the mediating variable is partially mediating, while when $c' < c$ but c' is not substantial, the mediating variable is fully mediating.

$$Y = cX + e_1 \tag{1}$$

$$M = aX + e_2 \tag{2}$$

$$Y = c'X + bM + e_3 \tag{3}$$

Based on this, the model is constructed as follows:

$$Willingness = \alpha_1 + \beta_1 Discrimination + \beta_2 Controls + \varepsilon_1 \tag{4}$$

$$Identity = \alpha_2 + \gamma_1 Discrimination + \gamma_2 Controls + \varepsilon_2 \quad (5)$$

$$Willingness = \alpha_3 + \delta_1 Discrimination + \delta_2 Identity + \delta_3 Controls + \varepsilon_3 \quad (6)$$

Where α reference to the constant term, β , γ , and δ are the coefficient of variation. ε represents an error term. Equation (4) tests the effect of social discrimination on MWRH. Equation (5) tests the impact of social discrimination on identity. The results of Eqs. (4), (5), and (6) are used to test the effect of identity in social discrimination on MWRH.

The following is the method's test procedure: The first step is to test whether the coefficient β_1 is significant; if it is significant, the aggregate effect is present. The second step is to test the coefficients γ_1 and δ_2 (Baron & Kenny, 1986). If the coefficients γ_1 and δ_2 are both significant, it indicates that identity plays a role in the impacts of social discrimination on MWRH. The third step tests the coefficient δ_1 . If coefficient δ_1 is significant, it shows a partially mediated effect; otherwise, it denotes a fully mediated effect.

Results

Descriptive statistics. The descriptive statistics are shown in Table 2. A total of 47,011 individuals aged 15 to 80 were interviewed, and the average age was approximately 36. The highest percentage of participants was Han and married, with 90% and 84%, respectively. The respondents were generally less educated, with an average score of 2.26. The score is between middle school and high school. In addition, rural residents with homestead and contracted land accounted for 74% and 60% of the total sample. The most likely reason is that about 26 % of farmers have moved entirely to cities. Migrant workers reported higher social discrimination scores, with 2.48.

The impact of social exclusion on MWRH. Table 3 shows the effect of social discrimination on MWRH based on the Probit model. The variants of our empirical model are grouped into four phases to progressively build more robust results. All estimation results were performed with robust standard errors. Model (1) is the regression result with only social discrimination and MWRH. Models (2)-(4) are gradually adding three levels of control variables. The results provide strong evidence to support the claim that social discrimination has affected MWRH. As the model predicts, MWRH as reflected by the social discrimination is positive in Models (1)–(4) with a significance level of 1%, indicating that MWRH is increased if they suffer social discrimination in the city³.

Meanwhile, the marginal effect of social discrimination level is 0.96%, which means that the probability of MWRH increased by 0.83% for each unit increase in the case of social discrimination. Therefore, Hypothesis 1 is supported. In other words, the more social discrimination is experienced, the higher MWRH. Migrant workers' thoughts of returning to their hometowns are largely because they have been unfairly treated or discriminated against in the cities, which makes them feel physically and mentally uncomfortable. To verify the model fit, this paper conducted the goodness-of-fit, Wald, and Hosmer-Lemeshow tests; the results are shown in Table 4. As seen in Table 4, the model fits well, further confirming the reliability of the results.

As expected, the results for the control variables are also consistent with expectations. The control variables of age and political affiliation passed the significant tests at 1%, and the regression coefficient is positive. Compared to the other variables, the marginal effect of political affiliation on MWRH is more

Table 2 Descriptive text statistics of variables.

Variables	Samples	Mean	SD	Min.	Max.
MWRH	47,011	0.03	0.17	0	1
Social discrimination	47,011	2.48	1.06	0	5
Identity	47,011	3.25	0.51	1	4
Gender	47,011	0.59	0.49	0	1
Ethnicity	47,011	0.90	0.29	0	1
Political affiliation	47,011	0.04	0.19	0	1
Marital status	47,011	0.84	0.40	0	1
Education	47,011	2.26	0.89	1	4
Age	47,011	36.39	9.87	15	80
Nature of housing	47,011	0.24	0.43	0	1
Income	47,011	8.70	0.57	3.91	12.21
Num. of family members living together	47,011	3.23	1.19	1	10
Availability of contracted land	47,011	0.60	0.49	0	1
Availability of homestead	47,011	0.74	0.44	0	1
Are there difficulties at home	47,011	0.58	0.49	0	1
Num. of cities on the move	47,011	2.10	2.02	1	80
Mobility range	47,011	1.66	0.75	1	3

Table 3 The influence of social exclusion on willingness is estimated by the Probit model.

	Model 1	Model 2	Model 3	Model 4
Social discrimination	0.187*** (0.011)	0.150*** (0.013)	0.151*** (0.013)	0.142*** (0.013)
Gender		0.000 (0.024)	-0.009 (0.024)	-0.035 (0.025)
Ethnicity		0.063 (0.040)	0.062 (0.040)	0.050 (0.040)
Marital status		-0.261*** (0.029)	-0.206*** (0.039)	-0.203*** (0.039)
Age		0.006*** (0.001)	0.005*** (0.001)	0.005*** (0.001)
Education		-0.066*** (0.014)	-0.055*** (0.015)	-0.047*** (0.015)
Political affiliation		0.222*** (0.060)	0.227*** (0.060)	0.228*** (0.060)
Nature of housing		-0.478** (0.041)	-0.457** (0.041)	-0.426** (0.042)
Num. of family members living together			0.010 (0.013)	0.010 (0.013)
Income			-0.134*** (0.027)	-0.170*** (0.028)
Availability of homestead			0.085*** (0.031)	0.070** (0.031)
Availability of contracted land			0.051** (0.026)	0.052** (0.026)
Are there difficulties at home			-0.105*** (0.024)	-0.102*** (0.024)
Mobility range				-0.073*** (0.017)
Num. of cities on the move				0.030*** (0.004)
Observations	47,011	47,011	47,011	47,011

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 4 Model fit test.

Test	Statistic
goodness-of-fit test	96.84%
Wald test	594.88***
Hosmer-Lemeshow test	$P = 0.052$

Note. $Ns P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

robust. Party members are 1.51% more likely to return home than non-party members. Compared to non-Party members, Party members may be better advocates of national policies, so they are more likely to want to return to their hometowns to participate in rural revitalization. Perhaps a more surprising result is that migrant workers with relatives in their hometowns who are sick are less willing to return home. They may have to earn money in the city to work to solve the difficulties in their hometowns.

Analysis of the mediating effect. Table 5 shows the results of the mediating effect test of identity in the effect of social discrimination on MWRH. In model 1, the impact of social discrimination on MWRH is positive when the control variables are added, indicating that social discrimination significantly contributes to the MWRH. The more social discrimination migrant workers experience, the more they feel that the city does not accept them and thus want to return home. Moreover, the variable of social discrimination is negatively significant at the 1% level in Model 2, showing that social discrimination negatively impacts on identity. Social discrimination, identity, and control variables are put in model 4. Hypothesis 2 is supported; the more serious the social discrimination of migrant workers, the lower their identity. Combined analysis of models (1), (2), and (3) reveals that the effect of social discrimination on MWRH remains significantly positive after adding mediating variables, but the coefficient is reduced. Meanwhile, the variable of identity significant at the 1% level indicates that a part of MWRH is influenced by identity, and identity plays a mediating role in the effect of social discrimination on MWRH. The mediating effect accounted for 50.77%. In addition, the Sobel test and the Bootstrap method confirmed the intermediary role of identity. The Z-value of the Sobel test is 23.63 and passes the 1% significance level test. The Bootstrap test is carried out after 1000 iterations, and the result shows that the confidence interval is (0.0032,0.0067), excluding 0. Hypothesis 3 is supported.

Endogeneity test. The variable of social discrimination may cause endogeneity issues. The MWRH has no direct impact on whether they suffer from social discrimination. Therefore, the endogenous problem in this paper does not exist endogeneity caused by reverse causation, but may be caused by sample selection bias and omitted variable bias. For example, the psychological state of the respondents may affect their willingness to return home. People with a good mindset may perceive that they are subject to less social discrimination, and their willingness to return home decreases⁴.

To solve the endogenous problem, the most common-used way is to apply the instrumental variable method. Theoretically, the instrumental variable must satisfy the conditions of correlation and exogeneity. Thus, “whether to apply for a residence permit” is an instrumental variable. In China, people can apply for a residence permit if they leave their household registration address, live in other cities for more than six months, and have a legal and stable job, residence, or continuous study in the city where they live. After applying for a residence permit, people can enjoy basic employment, public cultural services, public culture, and sports, and legal assistance in local areas, which could reduce

Table 5 Mediating effect of identity.

	(1) MWRH	(2) Identity	(3) MWRH
Social discrimination	0.142*** (0.013)	-0.206*** (0.005)	0.081*** (0.013)
Identity			-0.350*** (0.014)
Gender	-0.035 (0.025)	-0.007 (0.010)	-0.042* (0.026)
Ethnicity	0.050 (0.040)	-0.038** (0.017)	0.058 (0.042)
Marital status	-0.203*** (0.039)	0.015 (0.015)	-0.192*** (0.041)
Age	0.005*** (0.001)	0.010*** (0.001)	0.009*** (0.001)
Education	-0.047*** (0.017)	0.017*** (0.006)	-0.033** (0.015)
Num. of family members living together	0.010 (0.013)	0.013** (0.005)	0.007 (0.013)
Income	-0.170*** (0.060)	-0.068*** (0.024)	-0.193*** (0.062)
Political affiliation	0.228*** (0.060)	0.008 (0.024)	0.234*** (0.062)
Nature of housing	-0.426*** (0.042)	0.412*** (0.013)	-0.316*** (0.044)
Flow range	-0.073*** (0.017)	0.169*** (0.007)	-0.007 (0.018)
Number of cities on the move	0.030*** (0.004)	-0.027*** (0.003)	0.025*** (0.005)
Availability of homestead	0.070** (0.031)	-0.162*** (0.012)	0.031 (0.032)
Availability of contracted land	0.052** (0.026)	0.007 (0.011)	0.059** (0.027)
Are there difficulties at home	-0.102*** (0.024)	-0.033*** (0.010)	-0.109*** (0.025)
Proportion of mediating effect	0.228***	0.008	0.234***
Observations	47011	50.77%	47011

Note. $Ns P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 6 Estimation of instrumental variables.

	First stage Explained variable: Social discrimination	Second stage Explained variable: MWRH
Whether to apply for a residence permit	0.039*** (0.009)	-
Social discrimination	-	7.005*** (3.813)
Control variables	Yes	Yes
Observations	46,170 $F = 786.11$, $R^2 = 0.2035$	46,170 $Chi(2) = 107.74$ ***

Note. $Ns P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

social discrimination to some degree. In addition, whether to apply for a residence permit is not directly related to MWRH. Therefore, it can be considered that “whether to apply for a residence permit” is an appropriate instrumental variable for social discrimination. This study uses the IV Probit model for the estimation. Table 6 shows the regression results, in which relevant

variables are controlled to increase the robustness of the results. According to the first-stage regression result of Table 6, the *F*-value is 786.11 ($P = 0.000$), which means the instrumental variables are significant to the endogenous variable at the 1% level and are consistent with correlation of the instrumental variable and endogenous variable. In addition, the validity of instrumental variables needs to be further tested. The results are shown in Table 7. The significance levels of AR and Wald are both 1% significance level, indicating that the instrumental variable is not weak. Therefore, “whether to apply for a residence permit” is a suitable instrumental variable for social discrimination. The conclusions of this paper are verified once again.

Further analysis. Cities people benefit from improved facilities, a developed economy, higher incomes, and more job opportunities (Wang, Li, 2021; Zhu, 2021), and all these facilities attract many migrant workers. In this paper, the range of mobility also significantly affects MWRH. Therefore, this study further analyzes the differences in the impact of social discrimination on MWRH in different regions, ages, flow ranges, and incomes.

Regional differences. According to the China National Bureau of Statistics, this paper divides cities into the eastern, middle, western, and northeastern regions⁵. The empirical results are shown in Table 8. Specifically, the coefficients of the eastern, middle, western, and northeastern regions are 0.141, 0.125, 0.146, and 0.260, respectively, indicating that migrant workers in the western and northeastern regions have a higher willingness to return to their hometowns. In other words, migrant workers in less economically developed regions are more willing to return to their hometowns than economically developed regions. Economically developed regions are relatively rich in infrastructure and resources, leading to a significant increase in employment opportunities and a much higher quality of human capital. Therefore, migrant workers are less likely to be discriminated against and choose to return to their hometowns (Table 9).

Age differences. According to the age differences of migrant workers Zhu, Leng (2018), migrant workers born before 1980 are divided into a new generation of migrant workers and an old generation of migrant workers in this paper. The regression results are shown in Table 10. The samples of the new generation of migrant workers and the old generation of migrant workers pass the significance level test of 1%, and the coefficients are 0.139 and 0.149, respectively. The new generation of migrant workers is less willing to return home than the old generation. The possible reason is that the new generation of migrant workers has a higher education level, health status, and anti-pressure ability. In contrast, the older generation of migrant workers have a more vital idea of “fallen leaves return to their roots” with the increase of age, so they are more willing to return to their hometown (Table 11).

Table 7 Estimation of Cragg-Donald Wald.

Test	Statistic
AR	116.36***
Wald	14.54***

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 8 Regional differences in the impact of social exclusion on the MWRH.

	(1) Eastern	(2) Middle	(3) Western	(4) Northeastern
Social discrimination	0.141*** (0.018)	0.125*** (0.034)	0.146*** (0.024)	0.260*** (0.073)
Control variables	Yes	Yes	Yes	Yes
Observations	25438	7006	11384	3183

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 10 Age differences in the impact of social exclusion on the MWRH.

	(1) New generation of migrant workers	(2) Older generation of migrant workers
Social discrimination	0.139*** (0.020)	0.149*** (0.017)
Control variables	Yes	Yes
Observations	17,311	29,700

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 9 Mediating effects of regional differences.

	Eastern		Middle		Western		Northeastern	
	MWRH	Identity	MWRH	MWRH	MWRH	MWRH	Identity	MWRH
Social discrimination	0.133*** (0.018)	-0.210*** (0.007)	0.073*** (0.019)	0.156*** (0.019)	-0.192*** (0.008)	0.096*** (0.019)		
Identity			-0.356*** (0.020)			-0.352*** (0.021)		
Control variables	Yes	Yes	Yes	Yes	Yes	Yes		
Observations	27,121	27,121	27,121	19,890	19,890	19,890		
Social discrimination	0.146*** (0.024)	-0.191*** (0.010)	0.087*** (0.024)	0.260*** (0.019)	-0.202*** (0.008)	0.206*** (0.019)		
Identity			0.332*** (0.027)			-0.267*** (0.060)		
Control variables	Yes	Yes	Yes	Yes	Yes	Yes		
Observations	11,384	11,384	11,384	3183	3183	3183		

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 11 Mediating effects of age differences.

	New generation		Older generation			
	MWRH	Identity	MWRH	MWRH	Identity	MWRH
Social discrimination	0.139*** (0.020)	-0.189*** (0.008)	0.079*** (0.021)	0.149*** (0.017)	-0.217*** (0.006)	0.088*** (0.018)
Identity			-0.358*** (0.021)			-0.342*** (0.020)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17,311	17,311	17,311	29,700	29,700	29,700

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 12 Flow range differences in the impact of social exclusion on the MWRH.

	(1)	(2)
	Inter-provincial flow	Intra-provincial flow
Social discrimination	0.135*** (0.017)	0.151*** (0.020)
Control variables	Yes	Yes
Observations	24,539	22,472

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Flow range. The distance from home to the workplace is also an essential factor influencing the willingness to return home. In this paper, the group of migrant workers is divided into inter-provincial and intra-provincial flow according to the range of mobility. The results of the regression are shown in Table 12. The coefficients of the inter-provincial and intra-provincial flow samples are 0.135 and 0.151 with a significance level of 1%. It can be seen that the migrant workers within the province are more willing to return home than migrant workers across the province. Inter-provincial flow means that farther away from their hometowns and relatives; migrant workers need more time and energy to work outside. Therefore, to reduce the time costs and transport costs, these migrant workers may continue to work in the city with the idea that they will make a big difference in their career in the city and then return home (Table 13).

Income differences. Referring to Wang, Li (2021), this paper divides the monthly household income of migrant workers into four groups: below 3000 yuan, 3000 yuan to 5000 yuan, 5000 yuan to 8000 yuan, and more than 8000 yuan. The results are shown in Table 14. When the family monthly income of migrant workers is grouped into regressions, the coefficients are 0.206, 0.167, 0.107, and 0.106, respectively, and all of them pass the 1% significance level test. It can be seen that compared with other income levels, migrant workers with income below 3000 yuan are more willing to return home. In addition, with the increase in income, the coefficient decreases, which means that the desire to return home gradually decreases. The possible reason is that low-paid jobs are characterized by high intensity and low education levels of practitioners, which makes them unable to integrate into the city. In addition, based on the Push-Pull Theory, the high income of the (urban area) is a kind of urban “pull”. The higher the income level in the city, the stronger the “pull” on migrant workers, thus attracting them to the city (Table 15).

Robustness tests

In this section, we conduct two groups of sensitivity analysis to examine the robustness of our results. First, this study adopts the

Logit model to replace the Probit model. The regression results are shown in Table 16, which passed the significant tests at 1%. The direction of the explanatory variables’ influence on the explained variables is unchanged. Social discrimination has a significant positive influence on the willingness to return home, which is consistent with the results of the previous regression.

In addition, to further test the robustness of the results, the entropy evaluation method (EEM)⁶ is adopted to calculate the social discrimination in this paper. The regression results are shown in Table 17. It can be seen that after changing the calculation method of explanatory variables, the impact of social discrimination on MWRH still passes the significant tests at the 1%, and the mediating effect still exists.

Finally, as demonstrated in the previous regression results, age significantly affects MWRH; the older they are, the more inclined they may be to settle in a fixed place and be reluctant to move due to the loss of labor force, frailty, and other reasons. On the other hand, young people aged 20 and below may also stay fixedly in one place for a long time because they have to go to school. Therefore, this paper adopts the method of reducing the sample size by excluding the samples of age 20 and below, age 60 and above, and the samples of age 20 and below and age 60 and above for the robustness test, respectively. The regression results are shown in Tables 18 and 19. The results show that after excluding samples aged 20 and below, 60 and above, and both 20 and below and 60 and above, the regression results are still significant at the 1% significance level, and the mediation effect still holds.

In conclusion, the research findings of this paper are robust, social discrimination has a positive impact on MWRH, and it affects MWRH through identity.

General discussion

Conclusion and policy implications. As the main group of the floating population, migrant workers play a crucial role in integrated urban-rural development, and the study of MWRH is also significant for rural revitalization. Therefore, this study explores the impact of social discrimination on MWRH from the perspective of identity and draws the following four conclusions: (1) Social discrimination has a significant effect on MWRH. In other words, migrant workers with higher discrimination are more inclined to return home. The results held even after the instrumental variables mitigated the potential endogeneity problem. (2) Identity has a significant inhibitory effect on MWRH; the lower the identity, the intermediate mechanism by which social discrimination affects MWRH. Social discrimination leads to the low identification of migrant workers’ social identity, and they then choose to return home. (3) There is heterogeneity in the impact of social discrimination on MWRH. The willingness to return home is more significant in the less developed areas, the older generation of migrant

Table 13 Mediating effects of flow range differences.

	Inter-provincial		Intra-provincial			
	MWRH	Identity	MWRH	MWRH	Identity	MWRH
Social discrimination	0.135*** (0.017)	-0.221*** (0.007)	0.070*** (0.018)	0.151*** (0.020)	-0.184*** (0.007)	0.098*** (0.020)
Identity			-0.354*** (0.019)			-0.329*** (0.022)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24,539	24,539	24,539	22,472	22,472	22,472

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 14 Income differences in the impact of social exclusion on the MWRH.

	(1) Below 3000	(2) 3000-5000	(3) 5000-8000	(4) Above 8000
Social discrimination	0.206*** (0.039)	0.167*** (0.020)	0.107*** (0.023)	0.106*** (0.033)
Control variables	Yes	Yes	Yes	Yes
Observations	3595	17,606	16,355	9475

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

workers, low income, and the migrant workers who move within the province. Accordingly, the likely policy implications are as follows:

Firstly, promoting the positive image of migrant workers and eliminating urban residents' inherent cognition and prejudice against migrant workers. The government should urge the media to actively shoulder their social responsibilities and form a publicity mechanism with the city's publicity department as the leader, the functional departments cooperating, and all kinds of media making concerted efforts. In urban publicity, pay attention to the positive image of migrant workers, report on their practical difficulties, and guide the whole society to form a consensus that "migrant workers are the principal part of industry development in China, an important force in promoting the modernization of the country, and have made great contributions to economic and social development", and create an intense atmosphere in the whole society to learn from and help migrant workers.

Secondly, encouraging migrant workers to improve their quality and integrate into urban development naturally. The most crucial difference between the countryside and the city is that fewer laws and regulations make the countryside more "casual" than the city. This "casualness" makes migrant workers discriminated against and treated unfairly when they enter the city. Therefore, it is essential to improve the quality of migrant workers in all aspects. Carry out legal literacy and general education among migrant workers to guide them to enhance their concept of legal issues. In addition, carrying out social ethics education and guiding migrant workers to abide by the code of professional conduct.

Lastly, finding a balance between urban expansion and rural contraction. The vital factor of urban expansion and rural contraction is the flowing people, especially the unilateral flow from the countryside to the city. Therefore, the government should speed up the improvement of the labor flow system, reduce and eliminate the obstacles and difficulties of labor flow, comprehensively promote the equalization of public services, and

narrow the gap between urban and rural areas. Meanwhile, it provides a good platform for migrant workers to return to their hometowns to make outstanding achievements in the vast countryside.

Theoretical contributions. This paper systematically analyzes the impact of social discrimination on the willingness of migrant workers to return to their hometowns and provides new insights from theoretical and practical perspectives. It is of great significance to study the migrant population in integrated urban-rural development. The existing literature has mainly studied the effects of policies (Cai, Wang, 2007), institutions (Zhang et al., 2020; Qin & Zhou, 2014), the general social environment (Pan & He, 2017; Li & Long, 2009), and the good development of the countryside (Xu, 2014) on the willingness of rural migrant workers to return to their hometowns, and there has been insufficient analysis of social discrimination, identity and other psychological factors affecting the willingness to return home are not sufficiently analyzed. Meanwhile, there are fewer studies related to identity as a mediating mechanism between social discrimination and willingness to return home. Based on analyzing the composition of the indicators, this study empirically analyzes identity as a mediating variable, sounding the internal transmission mechanism of social discrimination on the willingness of rural migrant workers to return to their hometowns and enriching the theoretical research on the consequences of social discrimination and the factors that affect the willingness of rural migrant workers to return to their hometowns. Finally, the paper tests the heterogeneous role of social discrimination on the willingness of migrant workers to return home in different contexts. The impact of social discrimination on MWRH is particularly prevalent among rural residents in the western and northeastern regions, the older generation, those with lower incomes, and those with limited intra-provincial mobility. This can inform the development of targeted policy measures at the national level.

Practical implications. Providing a policy basis for social discrimination affects the labor force's willingness to return to the countryside. "Rural revitalization" is a key part of China's development strategy of "strengthening up", which is significant. Many rural migrant workers are attracted by the rich resources of cities to work in cities. In the process, many rural migrant workers suffer discrimination and rejection from various aspects of the city. This study provides policy references to alleviate the discrimination of rural migrant workers, improve their identity, and eliminate the obstacles to their development in cities by examining the impact of social discrimination on the willingness of laborers to return to their hometowns.

Table 15 Mediating effects of income differences.

	Below 3000 Identity		3000-5000 Identity		5000-8000 Identity		Above 8000 Identity	
	MWRH	MWRH	MWRH	MWRH	MWRH	MWRH	MWRH	MWRH
Social discrimination	0.168*** (0.020)	-0.194*** (0.008)	0.112*** (0.021)	0.210*** (0.039)	-0.198*** (0.019)	0.142*** (0.040)		
Identity			-0.343*** (0.038)			-0.401*** (0.042)		
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17,606	17,606	17,606	3595	3595	3595	3595	3595
Social discrimination	0.106*** (0.033)	-0.216*** (0.011)	0.044 (0.034)	0.108*** (0.023)	-0.210*** (0.009)	0.049** (0.024)		
Identity			-0.334*** (0.038)			-0.332*** (0.025)		
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9475	9475	9475	16,355	16,355	16,355	16,355	16,355

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 16 Social exclusion, identity and MWRH estimated by Logit model.

	(1) MWRH	(2) MWRH	(3) MWRH
Social discrimination	0.327*** (0.030)		0.187*** (0.030)
Identity		-1.505*** (0.042)	-1.531*** (0.058)
Control variables	Yes	Yes	Yes
Observations	47,011	84,450	47,011

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 17 The entropy evaluation method (EEM) to calculate social exclusion.

	(1) MWRH	(2) Identity	(3) MWRH
Social discrimination	0.475*** (0.035)	-0.836*** (0.016)	0.243*** (0.037)
Identity			-0.333*** (0.010)
Control variables	Yes	Yes	Yes
Observations	84,450	84,450	84,450

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 18 The effect of social exclusion on MWRH after downsizing the sample.

	(1) MWRH Delete sample 60 years and over	(2) MWRH Delete 20 years and under	(3) MWRH Delete 20 years and under and 60 years and over
Social discrimination	0.140*** (0.013)	0.141*** (0.020)	0.139*** (0.014)
Identity		-1.505*** (0.042)	-1.531*** (0.058)
Control variables	Yes	Yes	Yes
Observations	46,369	45,190	44,548

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 19 Mediating effects of downsizing the sample.

	Delete samples 60 years and over		Delete 20 years and under		Delete 20 years and under and 60 years and over	
	MWRH	Identity	MWRH	MWRH	Identity	MWRH
Social discrimination	0.140*** (0.013)	-0.204*** (0.005)	0.081*** (0.014)	0.140*** (0.013)	-0.206*** (0.005)	0.080*** (0.041)
Identity			-0.348*** (0.015)			0.139*** (0.014)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	45,190	45,190	45,190	46,369	46,369	44,548

Note. Ns $P \geq 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Data availability

Raw data collected and analyzed in the current study are available in the China Migrants Dynamic Survey: <https://chinaldrk.org.cn/wjw/#/home>.

Received: 9 March 2023; Accepted: 6 November 2023;

Published online: 07 December 2023

Notes

- 1 The data comes from China's National Bureau of Statistics.
- 2 http://www.gov.cn/zhengce/2022-02/22/content_5675035.htm
- 3 All coefficients are processed in standardization (similarly hereinafter). Standardization formula: $x' = \frac{x-u}{s}$.
- 4 From Interim Provisions on Residence Permit in China.
- 5 Eastern region: Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Hainan; Middle region: Shaixi, Anhui, Jiangxi, Henan, Hubei, Hunan; Western region: Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang; Northeastern region: Liaoning, Jilin, Heilongjiang.
- 6 The entropy evaluation method (EEM) determines the objective weight according to the index variability, avoids the deviation of human factors, and has strong credibility.

References

- Ai XQ, Cheng X, Li GZ (2021) A study on the influence mechanism of social discrimination on migrant workers' willingness to settle in cities: a mediating effect based on identity. *Popul Dev* 27(1):73–85
- Akerlof GA, Kranton RE (2000) Economics and identity[J]. *Q J Econ* 115(3): 715–753
- Alba R, Nee V (1997) Rethinking assimilation theory for a new era of immigration. *Int Migr Rev* 31(4):826–874
- Allport GW, Clark K, Pettigrew T (1954) *The nature of prejudice*, Boston: Beaton Press
- Baron RM, Kenny DA (1986) The Moderator-mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *J Pers Soc Psychol* 51(6):1173–1182
- Basu K (2013) Group identity, productivity and well-being policy implications for promoting development. *J Hum Dev Capab* 14(3):323–340
- Bertrand M, Duflo E (2017) Field experiments on discrimination. *Handb Econ Field Exp* 1:309–393
- Blunt A, Gruffud P, May J (2003) *Cultural Geography in Practice*. London: Edward Arnold p 71–73
- Cai H, Wang J (2007) "Migrant workers" permanent migration intend to study. *J Soc Sci Res* 6(86-113):243
- Chen KR, Long MQ, Li GC (2017) Sense of belonging of immigrant in urban villages of megacity: a case study of Shangbu village in Shenzhen. *Reg Res Dev* 36(5):64–74
- Chen YF (2005) "Peasant-labor": system and identity. *Soc Res* 3:119–132,244
- Cheng H, Yin NB (2004) Political economy analysis of migrant workers' employment discrimination. *Rural Econ* 2:20–23
- Cui YA (2012) Study on migrant' psychological integration and self-identity. *sociological. Research* 27(5):141–244
- David P (1999) *Sociology* (Trans. Li, Q et al.) China Renmin University Press, Beijing
- Deaux K, Dane FC, Wrightsman LS (1993) *Social psychology in the 80s* (6th ed.). Pacific Grove: Brooks/Cole Publishing Company
- Hou YJ, Yao H (2016) Patterns of migrant identity: a latent class analysis. *population. Research* 40(2):38–49
- Jiang Q (2009) Analysis on influencing factors of migrant workers' return intention in Sichuan Province. Sichuan Agricultural University
- Junge V, Diez JR, Schätzl L (2015) Determinants and consequences of internal return migration in Thailand and Vietnam. *World Dev* 71:94–106
- Kabeer N (2000) Social exclusion, poverty and discrimination towards an analytical framework. *IDS Bulle* 31(4):83–97
- Li H, Ni SG, Huang LY (2012) Self-identity of floating population and Its Influential Factors. *J Northwest Norm Univ (Social Science edition)* 49(4):68–74
- Li L (2004) Social discrimination: The invisible barrier of migrant workers' children's education. *Reference of Basic Education* (05):37–38
- Li Q, Long WJ (2009) Analysis on influencing factors of migrant workers' willingness to stay in City and return home. *China Rural Econ* 2(46-54):66
- Li XG (2011) The problem of social discrimination to peasant workers in the process of urbanization in China. *J Hainan Univ (Humanities and Social Science edition)* 29(4):68–73
- Li Y (2017) Interest threat, cultural Exclusion and frustration resentment: Immigrant exclusion under the new "native guest" relationship, (02)

- Mayda AM (2005) International migration: A panel data analysis of economic and non-economic determinant. Available at SSRN 725441
- Mcdaniel EL, Irfan N, Allyson FS (2011) Divine Boundaries: How Religion Shapes Citizens' Attitudes Toward Immigrants. *Am. Politics Res.* 39(1):205–233
- Pan ZQ, He Q (2017) Living space, Social interaction and subjective status cognition: A study on migrant workers' identity. *Hunan Soc Sci* (01):80–87
- Parkin F (1983) Marxism and class theory: A bourgeois critique
- Portes A (1994) The Informal Economy and Its Para-doxes, In Smelser NJ, Swedberg R (eds) *The Handbook of Economic Sociology*, Chapter 17, Princeton University Press (9)
- Portes A., Zhou M (1996) Self-employment and the Earnings of Immigrant, A. S. R. Albany: Apr, vol 61. p 219
- Qin XZ, Zhou JB, Xin Y, et al. (2014) The effect of urban-rural dual medical insurance structure on migrant workers' willingness to return home: A case study of migrant workers in Beijing. *China Rural Econ* (02)
- Sayce L (1998) Stigma, discrimination and social exclusion: What's in a word? *J Ment Heal* 7(4):331–343
- Shi XY (2011) Analysis on the reasons of migrant workers' lack of sense of belonging in cities: from the perspective of social mobility. *J Shandong Inst Adm* 05:4–7
- Shi ZL (2013) The development of migrant' capability and urban integration. *China Population. Resour Environ* 23(01):89–96
- Sides J, Citrin J (2007) European opinion about immigration: The role of identities, interests and information. *Br J Polit Sci* 37(3):477–504
- Stets JE, Biga CF (2003) Bringing Identity Theory into Environmental Sociology. *Social Theory* (21) : 398–423
- Wang YJ (2013) Settlement intention of rural migrants in Chinese cities: findings from a twelve-city migrant survey. *Popul Res* 37(4):19–32
- Wang ZT, Li HB (2021) Identity, personal skills and migrant workers' willingness to return their home-an empirical study based on the data of CMDS. *Econ Fabric* 20(3):34–43
- Wen ZL, Ye BJ (2014) Analysis of mediating effects: the development of methods and models. *Adv Psychol Sci* 22(5):731–745
- Xu JP (2014) Analysis of factors affecting new generation of migrant workers willingness to returning home and farming-Based on the survey in Shanxi Province. *Guangdong Agri Sci* 9(22):205–211
- Yang JH, Zhang JJ, Wu M (2016) My home is where my heart is: regional disparity and self-identity of migrants in China. *Popul Econ* 4:21–33
- Yu GL (2006) *Social Psychology*. Beijing: Beijing Normal University Press
- Zhang JP, Huang K, Wang JH, Huang M (2020) Return migration and Hukou registration reform in Chinese cities. *Econ Res J* 55(7):175–190
- Zhang SH, Li HY, Liu F (2012) A review of research on identity. *Psychol Res* 5(01)
- Zhang XJ (2018) "Group exclusion and partial acceptance"—an empirical analysis of the inter-group relationship between citizens and migrant workers. *Guangxi Soc Sci* 5:175–178
- Zhang ZX (2007) From urban-rural segmentation to regional segmentation: a new perspective of urban migrant population research. *Popul Res* 6:16–24
- Zhu HG, Kang LY, Weng ZL (2010) An empirical analysis of the influencing factors of migrant workers' willingness to return home to start their own business: based on the survey data of 1145 migrant workers returning home in Jiangxi Province. *China Rural Observation* (05)
- Zhu ZK (2021) How public health services affect migrant workers' willingness to stay in cities: An analysis based on the dynamic monitoring survey of China's floating population. *China Rural Econ* (10):125–144
- Zhu ZK, Leng CX (2018) Housing status, social status and urban identity of migrant workers: an empirical analysis based on social integration survey data. *China Rural Observ* 1:96–110

Acknowledgements

This research is funded by the National Natural Science Foundation of China (No.72063006), the Natural Science Foundation of Hainan Province (No.720MS027) and the Cooperation project of the Chinese Academy of Engineering and Local Governments as well as Enterprises Academy-Locality (No. ZGGCY-YDH-201811-03).

Author contributions

HX and WW wrote the manuscript. HX initiated the idea and edited the manuscript. All authors contributed to the article and approved the submitted version. All contributors meet authorship criteria and have contributed to the manuscript equally. All authors contributed to the paper and approved the submitted version.

Competing interests

The authors declare no competing interests.

Ethics approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Correspondence and requests for materials should be addressed to Wenjia Wu, Chuqiao Zhang or Yibo Xie.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023