Rural-to-Urban Migration, Perceived Inequality and Subjective Social Status in China

Lai Wei

Affiliation:
Department of Sociology, Princeton University
Direct correspondence to Wallace Hall 125, Princeton University, Princeton NJ 08540.
E-mail: laiw@princeton.edu

Abstract:
Using three waves of longitudinal data from China Family Panel Studies, this research investigates the impact of migration, both leaving and returning, on subjective social status and perceived inequality among rural-to-urban migrant workers in China. I find that leaving migration has a positive effect on perceived inequality and a negative effect on subjective social status. However, no evidence shows that returning migration has significant effect on either perceived inequality or subjective social status. Employing a novel operationalization of reference group, this research shows the linkages between migration, perceived inequality, reference group and subjective social status. Leaving migration has a negative causal effect on subjective social status, because triggered by increasing perceived inequality, migrant workers now have higher standards for evaluating social status. Returning migration cannot reverse inequality perception due to the persistence of knowledge, so the opposite process cannot be triggered by returning, and subjective social status does not recover even after migrant workers return.
1. Introduction

According to the sixth Chinese census in 2010, the size of migrant population in China reached 221 million, making it the largest migration in human history (Liang, 2016). Similar to international migration, rural-to-urban migration in China is driven by a dual labor market (Piore, 1979), which is maintained by a series of strict institutional arrangements, including the household registration (hukou) system. Migration opportunity clearly improves migrants’ objective well-being, as measured by income, but it is unclear how it influences subjective well-being. Income increase might lead to improving subjective well-being, but systematic discrimination and lack of emotional support in cities for migrant workers could offset such positive effect.

This research focuses on one specific type of subjective well-being, subjective social status. Migrant workers change their values and attitudes upon migration to adapt to the changing sociocultural environments. One set of values that are likely to be shaped by migration experience is about social comparison. Entering into cities expose migrant workers to new cultural norms and material life, but more importantly, to the striking urban-rural inequality. Past researches have found that rural residents have higher subjective well-being than urban residents in China, despite their relatively disadvantaged material condition, and they attribute it to the low inequality in rural areas (Knight & Gunatilaka, 2010a). For migrant workers, does migration reverse this rural advantage in equality? If migrant workers actively compare themselves with urban residents, and evaluate their social status based on their urban experience, then we could observe a decline in subjective social status when they enter cities, despite their income increase. On the other hand, if migrant workers hold no intention of staying in cities, and their entire goal of migration is just to improve family’s life in their hometown, as new economics of migration predicts (Stark & Taylor, 1989; Stark & Taylor, 1991), then we expect to observe a positive effect of migration on migrant worker’s subjective social status.

The circular nature of rural-to-urban migration in China further complicates the issue. Due to the difficulty of gaining permanent resident status, the majority of migrant workers would eventually return to hometown, and even when they are still active workers in cities, they regularly return home at least annually during the Chinese New Year. It is necessary to take returning into account when we consider the impact of migration. For instance, the effect of migration on subjective social status could be negative in the short term, when migrants are still in cities, but positive in the long term, when migrants return home. Another possible scenario is that the negative impact persists even after they return.

Using a longitudinal survey data, this research studies the impact of both out-migration and return-migration on rural-to-urban migrant worker’s subjective social status in China. I find that migration has a negative effect on subjective social status, which persists even after migrant workers return. I argue that the negative effect of migration could be explained by its positive effect on inequality perception. Increased inequality perception leads to higher reference standards in evaluating social status, which is gauged by anchoring vignette method, and higher reference standards means lower self-evaluations. The persistent negative effect of migration is a result of persistent inequality perception, which is knowledge-based and cannot be reversed by return migration.

In the next section, I am going to briefly review three strands of literature that are relevant to this article. First, I am going to review literature on immigration and subjective social status. Empirically, this article is among the first to identify the causal effect of migration on subjective well-being. Theoretically, this article complements classic researches on migration’s influence on subjective well-being by pointing out the role of persistent inequality perception. The second is on the measurement of reference group.
Employing a novel method, this research operationalizes reference group as a continuous “reference standard”, which is better equipped to answer questions that the prior operationalization of reference group cannot tackle. Third, I review literature on relative deprivation in China. This article points out one important mechanism causing China’s prevalent relative deprivation, that is domestic migration.

2. Literature Review

2.1 Immigration, reference group and subjective social status

Social comparison has received plenty of attention in immigration researches. It appeared first in the researches arguing against neoclassical economics which sees migration as a process of income maximization. They find that utility is not solely determined by absolute income, but also relative income, so people may refuse to migrate even that increases income. (Stark & Taylor, 1989; Stark & Taylor, 1991). Reference group plays a key role in the decision of migration, and whether migrants change their reference group is indeed important for the subjective consequence of migration.

Empirical findings concerning this question are mixed. New economics of labor migration sees returning as the natural outcome of migration (Stark & Bloom, 1985; Stark & Taylor, 1989; Stark & Taylor, 1991), and depicts returners as successful migrants who bring remittances to families and the home region (Demurger & Xu, 2011; Dustmann & Mestres, 2010; Marchetta, 2012; Yang, 2011). This strand of literature has largely focused on the economic consequences of return migration, while subjective well-being receives no attention. By emphasizing the economic gain of migration, they have implicitly assumed improvement of subjective well-being.

However, some researches do find evidence that migrant workers change reference group upon migration, and that has negative impact on their subjective well-being. Knight & Gunatilaka, (2010b) noticed that migrant workers in China are less happy compared with rural residents, and they attribute the finding to increasing expectation. Similarly, Wang (2017) found that migrant workers increase their reference standard when they enter cities, and their low subjective social status compared with rural residents can be explained by the shift in reference standard. Instead of seeing migrant worker’s reference group as exclusively in the sending region or in the receiving region, some other researches take a “transnationalism” perspective and argue that their reference group is actually a composite of the two, with evidence presented in the case of rural-to-urban migration in China (Wu & Wang, 2017) and Mexico-US migration (Gelatt, 2013).

Out-migration has an impact on migrant worker’s reference group and consequently their subjective well-being, but what if we take into account return migration? Four opposing hypotheses can be derived based on the abovementioned literature. If we view returning as an inevitable outcome as argued in new economics of labor migration, then the reference group stays in the home region across all periods, including after returning. If we see reference group as a contextual phenomenon, then we should expect migrant workers switch into urban framework upon migration but switch back when return. If we understand reference group in terms of expectation and aspiration, then it will not switch back after returning, because expectation is based on migrant worker’s experience and knowledge of urban life, which are irreversible. Finally, if we take a “transnationalism” perspective, we expect to see that reference group switches back but not entirely when return, because migrant workers are still influenced by the urban framework. No research has been conducted yet on how return migration influences reference group and subjective well-being, mostly due to data limitation.
Empirically, this article is among the first to identify the causal effect of migration on subjective well-being. Past researches have generally failed to make causal inference due to data limitation. Collection of longitudinal data, which social scientist rely on for valid causal identification, of migrant workers is difficult because they are hard to track, especially in the case of international migration. In view of the difficulty, most existing researches on the relationship between migration and subjective well-being employ cross-sectional data of internal migration, especially in the context of China (Knight & Gunatilaka, 2010b; Akay & Zimmerman, 2012), with Stillman et al. (2015) and Lönnqvist et al. (2015) as two exceptions. Taking advantage of the migration lottery policy from Tonga to New Zealand, Stillman et al. (2015) managed to remove self-selection bias of migration through a natural experiment design. They found that those who won the lottery and successfully migrated to New Zealand reported lower happiness than those who stayed in the home region. Lönnqvist et al. (2015) collected pre-migration samples of Ingrian-Finnish migrants from an immigration training program and successful followed them after their migration, finding an increase in life satisfaction. However, this article did not have information about the staying population, so a valid counterfactual is not established.

Students of return migration face similar dilemma of data collection. Researches about return migration in China is plentiful, but all of them are based on surveys of households in rural areas in the form of cross-section data (Demurger & Xu, 2011; Wang & Fan, 2006; Zhao, 2002). As a result, this strand of studies focused on the comparison between return migrants and non-migrants to evaluate the influence of migration experience. For instance, Zhao (2002) explored the difference in economic activities return migrants and non-migrants participate in. To the author’s knowledge, this study is the first one that uses large-scale longitudinal survey data to study the causal effect of migration on subjective well-being.

2.2 Operationalizing reference group

Originated in the middle of the twentieth century, reference group and relative deprivation have been among the most well-known and insightful sociological concepts (Stouffer et al., 1949; Merton and Kitt, 1950; Festinger, 1954). The theory suggests that people base their evaluations of situations at least partially on comparing themselves with others. It inspired many important substantive researches. In a now classic book of social movement studies and political sociology, Gurr (1970) established his whole argument upon the concept “relative deprivation”. Reference group theory is also the foundation of the explanation of an influential empirical puzzle, the Easterlin Paradox, which suggests that people’s happiness will not increase if the income of all members of the society increase altogether because happiness is based on the relative position of one’s income in a reference group (Easterlin, 1995).

The influence of reference group theory makes the immaturity of its operationalization all too surprising. Enormous numbers of researches have been conducted employing this concept, with most of them trying to use relative income to explain happiness as responses to the Easterlin paradox (for a thorough review, see Clark (2008)). The prevalent way existing researches operationalize reference group is to assume a priori a group respondent makes comparison with, like neighborhood (Luttmer, 2005), company (Brown et al., 2008), province (Wu & Wang, 2017) or even country (Helliwell, 2003), and then include both the respondents’ absolute income and the average income of the group in a regression to compare their effects. Some researches include more than one group average income to examine different comparison effects (Firebaugh & Schroeder, 2009; Gelatte, 2013). In this approach, researchers subjectively impose a reference group to respondents instead of identifying one.
The problem with the mainstream approach is that it runs the risk of capturing effects other than comparison. For instance, average income of a region also reflects public-good consumption (Clark et al., 2008). With only an indirect measure of reference group, it is never certain whether the finding is a far-stretched interpretation. One modification of the existing approach is to directly ask respondents who they usually compare themselves with (Knight et al., 2009; Alderson & Katz-Gerro, 2016). Researches in both US and China show that happiness is negatively associated with one’s tendency to compare themselves with others.

The modification is convincing in demonstrating comparison effects, but it relied even more on a stringent assumption underlying the traditional approach, which is that people only compare themselves with one single and concrete group when they form their opinions. This makes little sense in real life, because people are embedded in multiple social groups, and the reference point they compare themselves with is more likely to be an abstract aggregation of people around them, instead of one concrete group, like family members or colleagues. When the reference group imposed is too narrow, researchers miss comparison effects caused by other possible reference groups, and when the reference group imposed is too broad, like all people in the same province, it is hard to capture the comparison effect, because respondents are unlikely to compare themselves with all others in the same geographical region.

Assuming a single and concrete reference group is an over-simplification of the complicated process of social comparison. This article proposes an alternative to study social comparison phenomenon. Instead of reference group, this article sees the foundation of social comparison as “reference standard”, a continuous variable that people use when determining their subjective social position. Evans et al. (1992) and Kelley & Evans (1995) are the first ones to recognize the relationship between reference group and the subjective social status, and Wang (2017) first extracted reference standard from the response to subjective social status question. Following Wang (2017), reference standard is operationalized as the latent cut-points respondents employ when asked to rate their subjective social status, which is identifiable with anchoring vignette data. The ordinal response to the question can be decomposed into two parts, latent value and cut-points. An intuitive way to understand the difference of the two is that latent value is about one’s understanding of him/herself and cut-points are about his/her understanding of others, or of the society. After controlling for income or other SES variables, people with higher reference standards perceive lower relative status, and vice versa. Compared with the traditional approach, it is more informative in descriptively demonstrating people’s reference points and more flexible when used as dependent or independent variable in multivariate analysis. By definition, people’s reference groups switch with the change of their social position, but due to the difficulty of identifying reference group, most studies treat reference group as static, with almost no attention paid to the dynamics of comparison. Operationalizing reference standard as a continuous variable makes the study of comparison dynamics possible with longitudinal data.

2.3 Relative deprivation in China

Based on World Value Survey, Chen and Fan (2013) conducted an international comparison of the distribution of subjective social status. They found that in the 44 countries surveyed, China has the fourth lowest average subjective social status. The countries above China include some of the most underdeveloped countries, like Yemen, Rwanda and Zimbabwe. This research confirmed their finding that subjective social status is concentrated on the lower end in China, instead of centering on the middle as past research would expect (Evans & Kelley, 2004). Researches in subjective social status have found
that country’s average subjective social status is positively associated with its level of development (Evans & Kelley, 2004; Anderson & Curtis, 2012), and according to this theory, China is an outlier.

Theories have been proposed to explain the severe relative deprivation in China. Two most notable ones are what I refer to as transitioning thesis and inequality thesis. Gao (2013) argued that market transformation has brought fundamental changes to people’s class identification in China. The foundation of class identification has shifted from danwei, all-encompassing employment organization in socialist society, to market performances, like income. Compared with concrete organizations, reference groups based on income do not have concrete boundaries, and make upward-comparison more likely, leading to prevalent relative deprivation. Some other researches do not think that the foundation of class identification has shifted to market factors but believe that it is totally fragmented by market transformation (Li, 2005). High social mobility, low association between income, occupation and education, as well as cultural diversity, which are characteristics of a transitioning society, all lead to the absence of solid reference group for people to evaluate their social status.

Inequality thesis sees social inequality, instead of social transformation, as the main mechanism leading to relative deprivation. Chen and Fan (2013) found that local inequality level has significant negative effect on people’s subjective social status, which is larger than the positive effect of economic development. Hence, despite the rapid developments in China in the recent decades, subjective social status decreases due to the skyrocketing economic inequality.

Both social inequality and social transformation play roles in relative deprivation in China, but neither is sufficient condition for relative deprivation. Relative deprivation, in its simplest terms, is that people compare themselves with those who are better off than them. In other words, upward comparison is essential to relative deprivation. Transitioning thesis is meaningful in pointing out the uncertainty of reference groups, but that is not necessarily negative for subjective social status. We can easily imagine a transitioning society, like those who are experiencing a socialism transformation, where most people identify themselves as high in social status. Inequality thesis points out that some people are richer than others, but it can only influence people’s subjective social status if they are actively comparing themselves with the richer ones. We can imagine a highly unequal society where people do not feel relatively deprived, simply because they do not compare themselves with, or have no knowledge of the richer people.

This research tries to complement both theses by specifying one mechanism leading to relative deprivation in China, that is rural-to-urban migration. Like the transitioning thesis, I view reference group
as a dynamic phenomenon, and believe that the switch of reference group can lead to relative deprivation. Migration is one possible scenario of reference group switch and given the huge inequality between rural and urban areas in China, such switch is highly likely to cause upward comparison and lead to relative deprivation. This article also acknowledges that inequality leads to relative deprivation, but the problem is how? I argue that inequality cannot influence subjective evaluations until it is actually perceived by individuals, and for rural residents, migration into cities could be a significant exposure to social inequality.

Theoretical framework of this research is shown in Figure 1. Out-migration has a negative effect on subjective social status, because triggered by increasing perceived inequality, migrant workers now have higher standards for evaluating social status. Return-migration cannot reverse inequality perception due to the persistence of knowledge, so the opposite process cannot be triggered, and subjective social status does not recover even after returning.

3. Data and Method

3.1 Data

This article is based on China Family Panel Studies (CFPS), a comprehensive longitudinal survey of Chinese households. There are currently four waves of available data, and this article mainly uses the 2012, 2014 and 2016 waves.

In China, traditional general social surveys have difficulty sampling migrant population, because a large proportion of them live in temporary worker dormitories instead of normal households. Those who are sampled are usually non-representative of all migrant workers working in urban area. There are specialized surveys targeting migrant workers, but they lack information of the general population. The unique strength of CFPS in collecting information of migrant population is that it does not only sample from the urban areas, but also the rural areas. In interviews in rural households, interviewees are required to provide basic information about all family members, which usually include ones who are working in cities. The aim of the survey is to collect information of the whole household, so they will try to track the migrant workers mentioned by family members and conduct interviews with them.

The definition of migrant workers in this research includes both samples from urban areas and samples from rural areas. We define migrants as either who are reported by family members that are working away from home or who is rural hukou and lives in an area out of his home prefecture. Because the majority of the article is comparing migrant workers with rural population, respondents with urban hukou are excluded from the sample. As for the main dependent variables, “subjective social status” refers to the survey response to the question “what do you think is your social status in your local area” on a 1-5 scale, and “inequality perception” or “perceived inequality” refers to the response to the question “according to your opinion, how serious is the gap between the rich and the poor in today’s society” on a 1-10 scale.

Two parts of analyses are conducted below. The first is propensity score matching difference-in-difference analysis, which aims at showing the pattern of the changes of subjective social status across the migration cycle. This part uses all three waves of data. Two kinds of respondents are kept in the sample. The first are those who remain at rural home in the first wave, leave to work in cities in the second wave and return home in the third wave, from whom we can observe the complete migration cycle. This group is the treatment group of the study. The second is those who remain at home across all three waves, from
whom the control group is constructed through propensity score matching based on the attributes in the first wave. The difference between the two group gives us the causal effect of migration on subjective social status. The second part is hierarchical ordered probit modelling, which only includes data of the 2012 wave, and all respondents of rural hukou are kept in the sample.

3.2 Causal effect of migration on subjective social status and inequality perception

My aim of the research is to study the impact of migration on subjective social status and inequality perception with three waves of longitudinal data. One possible solution is using fixed effect, which simply observes individual variation over time with the change of migration status. However, this setting cannot take time trend into account. Actually, the main dependent variables fluctuate significantly across waves, so it will be hard to disentangle the treatment effect and the general time trend. Difference-in-difference is another common tool to make causal inference using longitudinal data. After identifying the treatment and control groups, the control group can serve as a counterfactual for the treatment group to estimate treatment’s causal effect. The key assumption of difference-in-difference design is common time trend. We can only use the control group as counterfactual when there is reason to believe that the treatment group would experience the same trend if it had not received treatment. In this case, it is a too strong assumption to assume common trend between migrant workers and the entire rural population, because migrant worker is a very special subgroup, so I use propensity score matching to construct the proper control group.

Specifically, I first identified the rural residents who stayed at their rural hometown in the first wave (2012), migrated into cities and worked as migrant workers in the second wave (2014), and returned home in the third wave (2016) as the treatment group, from whom I can observe the whole migration cycle. For the control group, they stay as rural residents across the three waves. As can be seen from table 1, there are far more samples in the potential control group than in the treatment group, which gives adequate space for matching methods because the treatment units are unlikely to be off support.

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Second Wave (2014)</th>
<th>Third Wave (2016)</th>
<th>Total Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>216</td>
<td>0</td>
<td>216</td>
</tr>
</tbody>
</table>

| Control Group   | First Wave (2012) | 0                  | Control Group               | 0                  | 19719                          |

Note: Total observations refers to the observations used in the construction of each group, not the observations used in the final analysis. In the control group case, only the information of samples who are matched to the treatment group is used in the DID analysis.

To construct the control group using propensity score matching, I first run a probit regression, with individual attributes in the first wave of data as predictors, and binary migration status in the second wave of data as outcome. This step uses respondents’ information in 2012 to predict their probability of becoming migrant workers in 2014. The predictors are all collected in the pre-treatment period and are hence comparable between the treatment and control group. In other words, there is no risk of post-treatment conditioning, which is harmful for the estimation of causal effects (Montgomery et al., 2018).
Then the control group is constructed using kernel matching algorithm, which employs information from the control group most effectively. The propensity score estimation results and balance check results are shown in Table 2. 70.8% of migrant workers identified in this research are male, and 27.8% are married. Their average annual family income is around 33,590 yuan. The majority of them only received primary or middle school education (68%), and they are overwhelmingly young, with 52.8% younger than 30, and another 25% between 30 and 40. According to the balance check results, all covariates are balanced except for age, marital status and gender. Compared with the control group, the treatment group includes more male, married, and young people. Subjective social status and perceived inequality are properly balanced between treatment and control group in the first wave of data. Difference between the two groups in the second and the third wave of data gives us causal estimates of migration. Essentially, two causal effects are estimated. The difference in the second wave of data, or the comparison between those who are migrating and those who remain in rural areas gives us the “migration effect”. The difference in the third wave of data, or the comparison between those who return to hometown and those who have never left gives us the “lingering effect” of migration. Both effects are quantities of interest to this research, since we do not only care about the impact of migration as migrant workers are in the cities, but also the persisting effects after they return.

| Table 2 Propensity Score Estimation and Balance Check Results of Propensity Score Matching |
|-----------------------------------------------|---------------------------------|---------------------------------|----------------|
| Propensity Score Estimation                  | Mean Control Group | Mean Treatment Group | Difference |
| Gender                                        | 0.379*** (3.74)     | 0.669               | 0.708       | 0.039*** (3.45) |
| Marital Status                                | 0.016 (0.13)        | 0.249               | 0.278       | 0.028** (2.62)  |
| Party Membership                              | -0.428 (1.16)       | 0.016               | 0.014       | -0.002 (0.81)   |
| Logged Family Income                          | 0.138** (2.73)      | 10.395              | 10.422      | 0.027 (1.15)    |
| Logged Asset                                  | -0.110* (2.06)      | 11.808              | 11.808      | 0.000 (0.01)    |
| Education: Illiterate                         | 0.080 (0.22)        | 0.166               | 0.153       | -0.013 (1.47)   |
| Education: Primary School                     | 0.169 (0.50)        | 0.335               | 0.333       | -0.001 (0.10)   |
| Education: Middle School                      | 0.062 (0.18)        | 0.340               | 0.347       | 0.007 (0.62)    |
| Education: High School                        | 0.123 (0.61)        | 0.139               | 0.139       | 0.000 (0.62)    |
| Age: 16-30                                    | 4.531*** (5.54)     | 0.468               | 0.528       | 0.060*** (4.91) |
| Age: 30-40                                    | 4.257*** (5.22)     | 0.248               | 0.250       | 0.002 (0.17)    |
| Age: 40-50                                    | 3.883*** (4.71)     | 0.206               | 0.181       | -0.025** (2.58) |
**Note:** All variables are from the baseline wave (2012). “Propensity Score Estimation” are results from Probit regression predicting migrant status in the second wave of data. Coefficients are reported and the absolute value of z scores are in parentheses. “Balance Check” reports the difference of each covariate between treated and constructed control group (Treatment - Control). The control groups are constructed through kernel matching algorithm with gaussian kernel and bandwidth of 0.005. Absolute values of t scores are in parentheses. The regression coefficient and group mean difference for region indicators are omitted. $p<0.05 \ast; p<0.01 \ast\ast; p<0.001 \ast\ast\ast$

### 3.3 Explaining the pattern of change

As shown in Figure 1, I argue that the negative association between migration and subjective social status is mediated by inequality perception and reference standard. Difference-in-difference analysis illustrates that migration has negative effect on subjective social status and inequality perception, and what remains to be shown is the process of “inequality perception-reference standard-subjective social status”.

In an ordinal response question, one can see the answer to the question contains two parts of information: the latent value and the thresholds. This research operationalizes reference standard as the latent threshold respondents use when asked to answer the question “on a scale of 1-5, which one do you think is your local status”. Usually, the thresholds are unidentifiable, but with anchoring vignette data, we can model latent value and thresholds simultaneously and hence separate the determinants of the two (King et.al, 2004). By threshold modelling, we can see how inequality perception influences reference standard, and how that affects responses of subjective social status.

The two anchoring vignettes employed in the survey are “After graduating from primary school, Mr./Ms. Chen makes his/her living as a street vendor, earning 1,000 yuan each month. In your opinion, what is the relative social status of Mr./Ms.Chen in your local area?” and “After graduating from a medical school, Mr./Ms. Zhou has become a doctor, earing 5,000 yuan each month. In your opinion, what is the relative social status of Mr./Ms. Zhou in your local area?” They provide information we need for estimating threshold heterogeneity. The distribution of self-rating and vignette-ratings are shown in Figure 2. The distribution of subjective social status is clearly right skewed with more respondents concentrating in the lower end than in the higher end. As for the two vignettes, the rating for vignette 1 is concentrated on the lower end, whereas the rating for vignette 2 is concentrated in the higher end, which is consistent with our expectation.

| Age: 50-60 | 3.395*** | 0.039 | 0.028 | -0.011* |
| Age: 60-70 | 3.611*** | 0.040 | 0.014 | -0.026*** |
| Perceived Inequality | 0.024 | 7.022 | 7.111 | 0.090 |
| Subjective Social Status | 0.019 | 2.667 | 2.708 | 0.042 |
| Life Satisfaction | -0.079 | 3.121 | 3.097 | -0.024 |
| Region Effects | YES | YES | YES | YES |
The techniques used to analyze anchoring vignette data is hierarchical ordered probit model. With the assumptions that respondents rate themselves and the vignettes based on the same set of thresholds (King et.al, 2004), we can model the latent value of self-rating following traditional probit model as:

$$Y_i^* = \beta_0 + \beta X_{it} + \eta_i, \quad \eta_i \sim N(0, \sigma^2)$$

with a matching scheme between latent and observed value as:

$$Y_i = k \text{ if } \tau_{ik-1} \leq Y_i^* \leq \tau_{ik}$$

(2)

Similarly, the latent value of vignette rating is modeled as:

$$Z_{ij}^* = \theta_j + \nu_{ij}, \quad \nu_{ij} \sim N(0,1)$$

with a matching scheme between latent and observed value as:

$$Z_{ij} = k \text{ if } \tau_{ijk-1} \leq Z_{ij}^* \leq \tau_{ijk}$$

(4)

Because we have assumed same thresholds in self- and vignette-rating, thresholds can be modelled as a function of a set of individual-level covariates $V_i$:

$$\tau_{ik} = \gamma^k_0 + \gamma^k V_i, \quad k = 1,2,3,4$$

(5)

The set of equations can be estimated through maximum likelihood estimation. Following Xu and Xie (2015), we also impose the constraint that $\gamma^k_0 = \gamma_0$ in some model specifications to extract one instead of four values as a more parsimonious measure of reference standard.
There is a total of five outcomes in the model, which are the latent value and four cut-points. If parallel change is assumed, then the outcomes are reduced to only the latent value and one cut-point. Latent value can be understood as an objective measure of social status. If respondents employ different sets of cut-points, they might give different responses of subjective social status even if their latent values are the same. According to the research hypothesis, those who have higher perceived inequality would employ higher reference standard in the evaluation of social status. Figure 3 is a visual illustration of this hypothesis. Assuming a constant latent social status $Y^*$, one would rate himself as lower middle in social status if his perception of inequality is low but would change the rate to low if his perception of inequality is high.

Figure 3 Illustration of Hierarchical Ordered Probit Models

Anchoring vignette data is only available in the 2012 and 2014 waves of China Family Panel Studies, so I am not able to conduct this part of analysis using the same data in the PSMDID analysis. I only use data from 2012 for the estimation of hierarchical ordered probit models, with inequality perception as main independent variable in the estimation cut-points. Control variables are also included in the regression analysis, which are the same as the controls in the PSMDID analysis.

4. Results
4.2 Propensity score matching difference-in-difference results

Figure 3 and Table 3 show the result of propensity score matching difference-in-difference analysis. The subjective social status does not differ significantly between treatment and control group in the baseline wave. In the second wave, when the treatment group migrated into cities whereas the control group stayed in home region, the subjective social status of treatment group is significantly lower than that of the control group, showing the negative causal effect. In the third wave, when the treatment group returned, the gap between the two groups narrows from 0.175 to 0.132, but is still significant, indicating the persisting negative effect of migration on subjective social status even after migrant workers return home.

![Graph showing causal effect of migration on subjective social status and inequality perception from propensity score matching difference-in-difference analysis.](image)

The right column of Figure 4 and Table 3 shows the impact of migration on perceived inequality. In the baseline wave, the control and treatment group do not differ in perceived inequality. In the second wave when the treatment group migrates into cities, they show significantly higher inequality perception compared with the control group who stay at the home region. After the migrant workers return, the gap...
between the two groups does not narrow but gets larger. In short, migration has negative effect on perceived inequality, which persists after the migrant workers return.

Table 3 Causal Effect of Migration on Subjective Social Status and Inequality Perception from Propensity Score Matching Difference-in-Difference Analysis

<table>
<thead>
<tr>
<th></th>
<th>Subjective Social Status</th>
<th>Perceived Inequality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before Migration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>2.614</td>
<td>7.074</td>
</tr>
<tr>
<td>Treated</td>
<td>2.647</td>
<td>7.096</td>
</tr>
<tr>
<td>Difference</td>
<td>0.033</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>During Migration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>2.873</td>
<td>7.354</td>
</tr>
<tr>
<td>Treated</td>
<td>2.699</td>
<td>7.625</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.175***</td>
<td>0.271***</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>After Migration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>2.779</td>
<td>6.969</td>
</tr>
<tr>
<td>Treated</td>
<td>2.647</td>
<td>7.309</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.132***</td>
<td>0.340***</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.054)</td>
</tr>
</tbody>
</table>

Note: Standard error in parentheses. Before migration refers to the 2012 wave. During migration refers to the 2014 wave. After migration refers to the 2016 wave. 

The difference-in-difference result in the case of perceived inequality is primarily driven by the treatment group. Compared with the first wave, in the second wave both treatment group and control group experienced an increase in perceived inequality, and the increase is larger for the treatment group, so it is fair to make the argument that migration leads to increasing perceived inequality. In the case of subjective social status, however, the result is driven by the control group. Compared with the first wave, control group experienced significant increase in subjective social status, but for the treatment group, the change in subjective social status is minimal. Indeed, across the three waves, subjective social status of treatment group fluctuates little. If the gap between treatment and control group occurs because of the changing conditions of control instead of treatment group, it does not mean the violation of any assumptions of difference-in-difference method, and the causal estimate is still valid, but it makes the interpretation of the result dubious, because after all the treatment group does not seem to be affected by the treatment.

The concern can be resolved if we can show that there is strong time trend across waves. If that is the case, then the stability of subjective social status for the treatment group does not mean the absence of treatment effect, because the effect could be offset by the general time trend. Table 4 shows the general time trend of subjective social status. For the complete sample, subjective social status increases from 2012 to 2014, and decreases from 2014 to 2016 but is still higher than the 2012 level. This trend is significant and consistent for urban residents, rural residents as well as migrant workers. Comparing Table 4 with Table 3, we notice that the control group in the propensity score matching difference-in-difference analysis also follows the general pattern. With such a strong general time trend, we have
confidence to state that the treatment group is an outlier in the temporal change of subjective social status, and we conclude that this unusual pattern is caused by the migration process.

Table 4 Average Value of Subjective Social Status by Wave and Group

<table>
<thead>
<tr>
<th></th>
<th>Urban Residents</th>
<th>Rural Residents</th>
<th>Migrant Workers</th>
<th>Complete Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>2.629</td>
<td>2.726</td>
<td>2.389</td>
<td>2.703</td>
</tr>
<tr>
<td>2014</td>
<td>2.827</td>
<td>2.995</td>
<td>2.614</td>
<td>2.971</td>
</tr>
<tr>
<td>2016</td>
<td>2.730</td>
<td>2.847</td>
<td>2.554</td>
<td>2.830</td>
</tr>
<tr>
<td>F-value</td>
<td>88.66***</td>
<td>283.72***</td>
<td>15.52***</td>
<td>297.84***</td>
</tr>
</tbody>
</table>

Note: p<0.05 *; p<0.01 **; p<0.001***

Migration is an important factor causing relative deprivation. Figure 5 shows the observed and counterfactual distribution of subjective social status of migrant workers. The counterfactual distribution is simply the distribution of the control group in the propensity score matching analysis. We can see that the distribution of migrant worker’s subjective social status is clearly right skewed, with more respondents concentrating on the lower end instead of the higher end. But the counterfactual distribution, which is a proxy for migrant worker’s distribution of subjective social status had they not migrated, is almost entirely centered in the middle, indicating a low level of relative deprivation. At least for migrant workers surveyed in this research, migration is a fundamental cause of their lower concentration of subjective social status.

Figure 5 Observed and Counterfactual Distribution of Subjective Social Status of Migrant Workers

Note: On the left shows the distribution of subjective social status of the treatment group in 2014, as defined in the PSMDID analysis section. On the right shows the distribution of subjective social status of control group in 2014, as constructed by pairwise propensity score matching.

4.2 Hierarchical Ordered Probit Model
In the last section we have demonstrated how migration influences subjective social status and perceived inequality. In this section, we show how the change of these two subjective variables are associated. This association is crucial to understanding why out-migration influences people’s subjective social status but return-migration does not.

To explore the relationship between subjective social status and perceived inequality, I estimated hierarchical ordered probit model (HOPIT) using the 2012 wave of CFPS, which has anchoring vignette data. As discussed in the method section, the advantage of HOPIT model is that it decomposes the total variation into latent-value heterogeneity and cut-point heterogeneity, the latter of which we conceptualize as “reference standard” and see as an important quantity of interest. Latent value and cut-points are modelled simultaneously. Table 5 presents the results of HOPIT and OPROBIT models.

Only respondents who are from rural origin are included in the regression. From the OPROBIT model, we can see that compared with non-migrants, migrant workers have significantly lower subjective social status. After controlling for a series of important individual attributes, the negative association becomes smaller, but still remains at a high level, with a z value of -7.53. However, the OPROBIT model does not allow cut-point heterogeneity, so we cannot tell whether the negative association comes from migrant workers’ lower latent value or higher reference standard. In other words, we do not know whether they have low self-evaluation because they see themselves too low or see others too high. HOPIT model helps us probe reference standard heterogeneity. The same set of predictors are included in the estimation of cut-points as in the estimation of latent value, and we focus on our main independent variable, which is perceived inequality. The result of parallel shift model gives us the estimate with the restriction that the effect of perceived inequality on all four cut-points are the same. The coefficient is positive and significant, meaning that increasing inequality perception leads to increase in reference standard. Non-parallel shift model allows inequality perception to have different effects on the four cut-points. We can see that the positive effect is mostly from low-lower middle and lower middle-middle cut-point, whereas the coefficient for the other two higher cut-points are negative and non-significant. As for the latent-value estimation, after accounting for cut-point heterogeneity, migrant workers do not display significantly different subjective social status compared with non-migrants, so the difference between the two groups is caused by cut-point heterogeneity.

### Table 5 Ordered Probit and Hierarchical Ordered Probit Models of Subjective Social Status

<table>
<thead>
<tr>
<th>Outcome: Subjective Social Status</th>
<th>Ordered Probit</th>
<th>Ordered Probit</th>
<th>HOPIT Parallel</th>
<th>HOPIT Non-parallel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Migration</td>
<td>-0.352***</td>
<td>-0.265***</td>
<td>0.0558</td>
<td>0.0517</td>
</tr>
<tr>
<td></td>
<td>(-11.07)</td>
<td>(-7.53)</td>
<td>(1.22)</td>
<td>(1.10)</td>
</tr>
<tr>
<td>Controls</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

### Outcome: Cut-points

<table>
<thead>
<tr>
<th>Low-Lower Middle Perceived Inequality</th>
<th>0.00438**</th>
<th>0.0141***</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(2.68)</td>
<td>(5.74)</td>
</tr>
<tr>
<td>Controls</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Lower Middle-Middle Perceived Inequality</td>
<td>0.0087***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>YES</td>
<td>Middle-Upper Middle</td>
</tr>
<tr>
<td>---------</td>
<td>-----</td>
<td>---------------------</td>
</tr>
<tr>
<td>Controls</td>
<td>YES</td>
<td>Upper Middle-High</td>
</tr>
<tr>
<td>N</td>
<td>18,349</td>
<td>18,349</td>
</tr>
</tbody>
</table>

Note: Controls for latent-value estimation include gender, marital status, party membership, logged family income, logged asset, education level dummies, age group dummies, regional dummies. Controls for cut-point estimation include migration status, gender, marital status, party membership, logged family income, logged asset, education level dummies, age group dummies, regional dummies. The coefficients of HOPIT model is adjusted to be comparable to the coefficient of ordered probit model. Z values in parentheses.

Figure 5 provides an illustration of the working mechanism of cut-points. On the left shows the reporting pattern if we assume cut-point homogeneity. The response to the survey question is totally determined on respondents’ self-evaluation of social status, but not on their perception of the social hierarchy. We can imagine this as in the scenario that all respondents are given a perfectly detailed guideline of answering this question, so any incomparability due to interpretation uncertainty is removed. On the right shows the actual estimated cut-point heterogeneity on perceived inequality. The coefficients for the two lower cut-points are positive, meaning that as inequality perception increases, respondents’ standard of what counts as “lower middle” vs “low” social status, and what counts as “middle” vs “lower-middle” social status increases. For people with low socioeconomic status, if they have higher perceived inequality, they are more likely to identity as low instead of lower middle, and as lower middle instead of middle in social status. But the pattern does not hold for the high SES end. The slope of the two higher lines are actually negative, meaning that for high SES respondents, if they perceive high social inequality, they are more likely to overrate rather than underrate their social status. This is understandable given that perceived inequality is measure with the question “how serious do you think the rich-poor gap is”. For the rich side of the gap, the bigger they perceive the gap is, the higher will they place themselves.
Figure 6 Cut-Point Heterogeneity of Subjective Social Status on Perceived Inequality

Note: Left graph is an illustration of reporting pattern without cut-point heterogeneity and with equal distribution of the five responses. Right graph is an illustration of the estimates from the HOPIT Non-parallel model from Table 5. The slopes of the four lines are the estimates of the effect of perceived inequality on the four cut-points. The intercepts are the constant terms from the four cut-point estimations.

In summary, migration has negative effect on both subjective social status and perceived inequality, which persist even after returning. The decreasing subjective social status is partly a result of increasing perceived inequality, because increasing perceived inequality leads to higher reference standard of evaluating social status. The persistence of migration’s negative effect on subjective social status is due to the persistence of knowledge about social inequality. Knowledge is irreversible, so even for migrants who have returned home, migration still has negative effect on their self-evaluation of social status.

5. Discussion

Previous researches have long depicted return migrants as heroes who, after arduous but fruitful work away from home, gains respect and prestige from families and local peers upon returning. This research, however, finds that return migrants actually have lower subjective social status compared with the control group who stay at hometown. By emphasizing the economic gain of migration, many studies implicitly assume migration has positive effect on subjective well-being as well. For migrant workers who are bound to return, like seasonal international workers or rural-to-urban migrant workers in China, even if there is subjective suffering due to discrimination or lack of social support in the receiving region, it should no longer exist when they get back to home region. Underlying such an argument is the
assumption that factors leading to migrant worker’s subjective suffering are totally contextual, but I find that some of those factors are persistent. Inequal perception is one of them. Migrant workers might regain social support, live in more comfortable housing conditions, or reassimilate into local culture once they return, but inequality perception cannot be reversed due to the persistence of memory and knowledge. Therefore, their disadvantage in subjective social status will continue even after they return.

Reference group is central to many subjective well-being indicators, not only subjective social status, but also life satisfaction and happiness. Past researches have difficulty employing reference group both as independent and dependent variables, mostly due to the underdevelopment of measurement. This study tries to operationalize reference group as “reference standard”, which enables me to conduct more granular analysis compared with past studies employing the concept. Migration is only one possible cause of reference group switch, it remains to be explored how other important life events, like marriage and employment, influence people’s evaluation of themselves and evaluation of the society.

How does social inequality in China influences people’s subjective well-being? This article claims that inequality can only influence subjective evaluation when they are actually perceived by individuals, and migration is an important mechanism in this perception process. This study only investigates the case of rural-to-urban migration and finds that migration plays a big role in migrant worker’s relative deprivation. However, rural-to-urban migration is only one type of migration in today’s China. College admission, job market entrance, among others, are all possible events leading to changes in geographical location. Moreover, migration is not the only channel of inequality perception. Rising media exposure and more frequent social contact could also increase inequality perception, but little research has been done about how they influence subjective well-being, especially relative deprivation.

This research only studies the short-term effect of migration in a four-year window, which is one major limitation. It is unclear if migrant workers stay in cities, or return home for longer periods of time, their reference group and subjective well-being would have further changes. Another related problem is the representativeness of the treatment group in this study for the entire migrant worker population. The treatment group changed their geographical location twice in four years, which is quite frequent even for rural-to-urban migrants in China who are highly circular. More analyses need to be done to address whether this group is only a special subgroup of migrant workers.
References:


