Public Opinion Quarterly, Vol. 85, No. 1, 2021, pp. 123–146

MIGRANT INFLOWS AND ONLINE EXPRESSIONS OF REGIONAL PREJUDICE IN CHINA

XI WANG
YANG ZHANG*
SHANGGUAN WANG
KANG ZHAO

Abstract China started market-oriented reforms in the late 1970s, and a fundamental social change in post-reform China has been massive migration. The primary purpose of this research is to examine the effect of migrant inflows on online expressions of regional prejudice, which is defined as expressions of antipathy toward migrants. In particular, we examine the dynamics of expressions of regional prejudice on Weibo, a popular Twitter-like social media site in China. Furthermore, we explain online expressions of regional prejudice through the lens of the group-threat and contact theories, and seek potential conditions that facilitate the development of a positive relationship between local residents and migrants. Our findings lend strong support to the group-contact argument that migrant inflows reduce online expressions of regional prejudice in China. The contact effect is larger when migrant inflows are measured in terms of migrant labor and interprovince migrants, who represent China's typical migrants. Furthermore, the contact effect is only observed in relatively small cities, while a small threat effect is observed in mega and super-mega cities where the population-control system is strictly maintained.

Xi Wang is an assistant professor in the School of Information at the Central University of Finance and Economics, Beijing, China. Yang Zhang is an assistant professor of political science at Southwest Jiaotong University, Chengdu, China. Shangguan Wang is an undergraduate student in the Department of Business Analytics at the University of Iowa, Iowa City, IA, USA. Kang Zhao is an associate professor in the Department of Business Analytics at the University of Iowa, Iowa City, IA, USA. The authors thank Jiaqi Lei, Zhiya Zuo, Tingyao Zhu, and the undergraduate volunteers from the Central University of Finance and Economics for data annotation, and Shanghui Qi and Jingyao Chen for putting together the yearbook data. This work was supported by the MOE (Ministry of Education in China) Project of Humanities and Social Sciences [19XJC810001 to Y.Z.]; and a grant from the Program for Innovation Research at the Central University of Finance and Economics to X.W. *Address correspondence to Yang Zhang, Southwest Jiaotong University, School of Public Affairs and Law, 111 North Yiduan, Erhuan Road, Chengdu, China; email: yzhang_polisci@163.com.

doi:10.1093/poq/nfab004 Advance Access publication June 21, 2021

During a cold winter night in November 2017, a fire killed 19 migrant renters and injured eight more in suburban Beijing, China. The accident attracted widespread media attention, and special safety inspections were immediately carried out by the local government. Because these inspections determined their homes to be unsafe, many migrant residents were evicted from their rented apartments on short notice and became homeless overnight. On Weibo, a Twitter-like social media site in China, this issue received heated discussion. While some local residents in Beijing expressed sympathy toward the evicted migrant renters, many others supported the government's action and believed that the large migrant population is a major cause for a host of social problems in the city.

Understanding the recent history of migration within China is important in order to understand the public's mixed reactions toward the plight of migrants. In the late 1970s, China loosened state control over its economy and embraced market-oriented reforms. The sociopolitical conditions brought on by this change brought about an enormous inflow of rural residents into cities, a fundamental social change in post-reform China (Zhou and Xie 2019). Many local residents in cities blame migrants for the depletion of urban resources, such as public housing, education, and healthcare services. The distrust and prejudice between locals and migrants has imposed a significant challenge to social harmony and the stability of Chinese society (Tse 2016).

What contributes to the prejudice toward migrants that is expressed on the Internet? In addition, what conditions are likely to reduce the expression of prejudice? We explore the origins of online expressions of regional prejudice in China from the perspectives of group-threat theory (Blalock 1967) and group-contact theory (Allport, Clark, and Pettigrew 1954). Both theories explain how the size of a migrant population affects intergroup relations but yield contradictory predictions. Group-threat theory predicts that larger migrant inflows will increase regional prejudice, while group-contact theory predicts that larger migrant inflows will decrease regional prejudice.

We tested these two competing theories with an analyses of a large corpus of Weibo messages about the issue of migration. Over the past decade, as internet services covered wider areas in China and the population of smartphone users grew rapidly, social media have become an important part of people's lives and have facilitated communication between different social groups. In the context of China, Weibo is one of the most popular social media sites, with nearly half of all Chinese internet users (376 million by September 2017) on Weibo each month (Weibo Data Center 2017; China Internet Network Information Center 2018). Similar to Twitter, Weibo makes personal opinions public. Previous research has relied primarily on survey and experimental data to measure prejudice (Pettigrew and Tropp 2006), but the Weibo data have some advantages over these approaches. First, Weibo

allows one to post a message while hiding one's true personal identity, thus reducing the effect of social desirability bias. Second, Weibo posts reflect voluntarily expressed bias, rather than responses elicited in an unrealistic context by researchers. Third, the large volume of data from Weibo allow an estimate of prejudice to be constructed for each region, making it possible to uncover geographic patterns of expressed prejudice across diverse regions.

Our research approach combines data-science techniques with econometric methods that are commonly applied in social science inquiry. In particular, webscraping is used to collect the textual data from Weibo, and machine-learning classifiers are used to create the measures of online expressions of regional prejudice. Linear regression analysis was used to test the predictors of online regional prejudice. Our research findings lend support to the group-contact theory argument that migrant inflows reduce expressions of prejudice toward migrants. Further analyses show that the effect of migrant inflows is stronger when certain contact conditions are satisfied. The contact effect grows when migrant inflows are measured in terms of interprovince migrants and migrant labor, who represent typical migrants. Support for group-contact theory only exists in relatively small cities where a relaxed population policy eases the way for migrant integration. However, in mega and super-mega cities where the local government is required to maintain strict population control, a small group-threat effect was observed.

Migrant Inflows: Threat or Contact?

The current paper examines online expressions of prejudice toward migrants in China. There is substantial internal migration in post-reform China (Zhou and Xie 2019). In 2018, the Chinese migrant population was estimated to be approximately 240 million (National Bureau of Statistics 2019).

Much of the research examining Chinese migrants has focused on institutional discrimination rather than individual expressions of prejudice. A major focus of these studies is the institutional discrimination maintained by the household registration (*hukou*) system in which households are registered as being in a specific region (Cheng and Selden 1994; Knight, Song, and Huaibin 1999; Wu and Treiman 2007; Chan and Buckingham 2008; Pakrashi and Frijters 2017). Although the *hukou* system has provided the government with an important administrative means to cope with demographic pressures since 1955 (Wu and Treiman 2007), it results in systematic discrimination against those who migrate from the region where their household is registered because by leaving their home region, they may lose their eligibility to receive a wide range of welfare benefits offered by the government, such as public housing, education, healthcare, and a pension. The *hukou* system is more strictly enforced in large cities in order to

avoid potential negative consequences of migration such as slums or unhealthy conditions.

In contrast to this work focused on institutional discrimination, our research examines prejudice expressed by individuals, where prejudice can be defined as "an antipathy based on a faulty and inflexible generalization" (Allport, Clark, and Pettigrew 1954, p. 9), which is characterized by irrationality and emotional evaluation (Hainmueller and Hangartner 2013). Specifically, our research analyzes individual social media posts to assess and explain regional variation in prejudice toward migrants.

One important factor that varies across regions and may affect expressions of prejudice is the amount of migrant inflow or the migrant group size. Theoretical predictions regarding the effect of migrant group size on prejudice are conflicting. Group-threat theory (Blalock 1967) predicts a negative effect of migrant inflows on regional prejudice, but the group-contact theory (Allport, Clark, and Pettigrew 1954) predicts a positive effect. Empirical evidence regarding this relationship is also mixed. In Pottie-Sherman and Wilkes's meta-analysis, 25 percent of the effects are positive, 15 percent are negative, and 60 percent are not significant (Pottie-Sherman and Wilkes 2017). Among three explanations for these differences they explore, only differences in the measurement of immigrant group size account for the different results across studies.[1] We examine support for the group-threat and group-contact theories by testing the relationship between migrant group size and expressions of prejudice, and we examine how this relationship is affected by the measurement of migrant group size (i.e., using all migrants or those considered "typical") and the strictness of the *hukou* system (i.e., one possible indicator of authority support or opposition to migration).

Group-Threat Theory

First developed from an observation of racial relations in the United States (Blalock 1967), group-threat theory suggests that an increase in the size of the minority population is associated with greater prejudice and discrimination (Alba, Rumbaut, and Marotz 2005), because larger minority group size heightens the competition for scarce resources and the potential for political mobilization (Blalock 1967). The threat from the minority group may be realistic (Blumer 1958; Bobo 1983), but it can also be symbolic (Kinder and Sears 1981) because the dominant group may see a large minority group as a threat to their morals, values, and beliefs (Hainmueller and Hangartner 2013). Group-threat theory therefore leads us to the following hypothesis.

1. The other two explanations are differences in model specification and differences in the geographic unit of analysis (Pottie-Sherman and Wilkes 2017).

Threat hypothesis 1: Migrant group size will be positively associated with increased online expressions of regional prejudice (threat effect).

Group-threat theory suggests prejudice toward migrants in China may result from realistic conflict over resources. China's labor market today is best described as a split labor market. In a split labor market, there is a large price differential of labor across different social groups for the same occupation, which produces a three-way conflict among business and the two labor groups. The price differential between local and migrant workers results from differences in their economic, information, and political resources as well as motives. In China, these group differences are sustained largely by the *hukou* system (Ling 2015; Wang, Guo, and Cheng 2015; Zhang and Wu 2017). The management scheme of public goods is localized so the local government is only responsive to the demand of local residents (Zhou and Wang 2016). The strict *hukou* system imposes high costs on settlement, and many migrants only intend to work temporarily in destination cities. Owing to weak bargaining power and weak willingness to locate to a settlement, the price of migrant labor is much lower than that of local labor. Consequently, businesses have an incentive to displace higher-paid local labor with cheap migrant labor, creating tension between the two labor groups. Consistent with this argument, Zhu (2018) found that among people living in Chinese cities, exposure to information about the negative bearing of migrant inflows on the local labor market and public finance reduces support for migrants. This leads us to hypothesize that threat will be greater when migrant group size is assessed using migrant laborers.

Threat hypothesis 2: The threat effect will be stronger when migrant group size is measured in terms of migrant laborers.

Migrants in China may also pose a symbolic threat to residents. Symbolic threat involves group differences in customs, values, norms, and beliefs (Kinder and Sears 1981), and it is most salient between groups whose cultures are highly dissimilar (Stephan et al. 1998). Previous studies have shown that anti-immigrant sentiment is origin-based and that native residents hold stronger prejudice toward immigrants who are culturally distant (Hainmueller and Hangartner 2013). There exists great cultural variation across different regions in China (e.g., language, accents), and cultural distance is correlated with geographic distance. We therefore expect that migrants coming from distant regions are more likely to invoke a symbolic threat, leading to the following hypothesis.

Threat hypothesis 3: The threat effect will be stronger when migrant group size is measured in terms of migrants who are coming from more distant regions.

Moreover, the *hukou* system protects local residents' economic interests by limiting the possibilities for the migrants. Therefore, in the regions where the

hukou regulations are strictly implemented, migrant inflows are less likely to be perceived as a threat by local residents and trigger regional prejudice, leading to the following hypothesis.

Threat hypothesis 4: The threat effect will be weaker in regions where the hukou *system is highly enforced.*

Group-Contact Theory

In contrast to group-threat theory, the main thesis of the contact theory is that contact with the members of a group reduces intergroup prejudice (Allport, Clark, and Pettigrew 1954), and research examining group-threat theory has demonstrated a routine and robust contact effect across a wide range of social settings (Pettigrew and Tropp 2006). For instance, in one of their meta-analyses, Pettigrew and colleagues (2011) found that 94 percent of the prejudice studies report a negative relationship between contact and prejudice. Social contact is effective in improving the relationship between racial and ethnic groups, as well as other—often stigmatized—groups, such as homosexuals, the mentally ill, and the disabled.

Furthermore, survey and experimental studies have provided evidence supporting the contact thesis concerning Chinese migrants. One survey of 885 adult residents in Jiangsu Province showed that intergroup friendship led to more positive attitudes toward rural-urban migrants (Nielsen et al. 2006). Data from a nationally representative survey showed that interaction—either intimate or not—with an individual migrant increased urban residents' willingness to engage in contact with migrants in general (Xue 2018). Furthermore, an experimental study provided evidence that contact between urban and rural middle school students (via a shared puzzle-solving task) resulted in more positive intergroup attitudes (in comparison to a control group; Gu et al. 2016). This leads to the following hypothesis about contact.

Contact hypothesis 1: Migrant group size will be negatively associated with online expressions of regional prejudice (contact effect).

If migrant inflows are able to induce more contact between locals and migrants in China, we might expect that migrant group size would be negatively associated with prejudice. Previous research on prejudice has identified a much larger contact effect for majority groups than for minority groups (Tropp and Pettigrew 2005). Therefore, migrant inflows should mainly affect the online expressions of regional prejudice of local residents more than those of migrants.

There are four explanations of why contact with an individual outgroup member reduces prejudice toward the outgroup in general (Pettigrew 1998). First, if stereotype-based prejudice is primarily promoted by ignorance about outgroup members (Stephan and Stephan 1984), and holding stereotypes

about outgroup members causes prejudice (Stephan and Stephan 2000), contact should reduce stereotypes and therefore prejudice. Second, anxiety about an outgroup may cause prejudice (Stephan and Stephan 2000) and repeated contact generates affective ties, reduces anxiety, and helps develop intergroup empathy (Reich and Purbhoo 1975) and friendship (Pettigrew 1997). Third, persons with strong prejudice often try to avoid interaction with the targets of prejudice, so contact itself can create uncomfortable dissonance (Aronson 1978). To resolve the dissonance, they may readjust their attitudes toward the outgroup by lowering their level of prejudice. Finally, intergroup contact may lead to ingroup reappraisal (Pettigrew 1998), whereby intergroup contact reshapes people's view of their ingroup norms and customs, leading "to a less provincial view of outgroups in general," a process called "deprovincialization" (Pettigrew 1998).

Intergroup contact is more likely to lead to reduced prejudice under some conditions than under others. Four classic optimal contact conditions include equal group status, common goals, intergroup cooperation, and authority support (Allport, Clark, and Pettigrew 1954). Empirical evidence demonstrates the effectiveness of these four conditions (Pettigrew and Tropp 2006). Researchers have also discovered the importance of salient outgroup membership (Hewstone 2000; Pettigrew and Tropp 2006). When outgroup saliency is low, the contact effect is interpersonal and fails to generalize to the group level (Pettigrew 1998). In contrast, contact with typical outgroup members is likely to reduce intergroup prejudice to a greater degree.

Migrant workers and long-distance migrants are considered typical migrants in China. Rural-urban migrant workers made up the first massive wave of migration in post-reform China, and they still constitute the largest proportion of the migrant population today (National Bureau of Statistics 2011). Most local residents' first face-to-face encounters with migrants involve migrant workers, who have also received substantial media coverage. Short-distance migrants travel to nearby cities, and local residents are apt to regard them as ingroup members instead of outgroup members. Long-distance migrants, however, can be easily recognized because of their distinct lifestyles, dialects, and even appearance, and they are often categorized as outgroup members by locals. Therefore, we expect that migrant workers and long-distance migrants, who are perceived as the typical representatives of the migrant group, are more likely to evoke regional prejudice in their destination cities, leading to the following hypotheses.

Contact hypothesis 2: The contact effect will be stronger when migrant group size is measured in terms of migrant laborers.

Contact hypothesis 3: The contact effect will be stronger when migrant group size is measured in terms of migrants who are coming from more distant regions.

One optimal contact condition, authority support, should have particular interest to policymakers who are concerned with intergroup tensions (Allport, Clark, and Pettigrew 1954). In Pettigrew's words, "authority support establishes norms of acceptance" (Pettigrew 1998, p. 67). He points out that explicit authority sanction makes intergroup interaction readily accepted and facilitates the contact effect in prejudice reduction (Pettigrew 1998). The authority support condition sheds light on the role of the Chinese *hukou* system in shaping the intergroup attitudes between locals and migrants. As mentioned earlier, the *hukou* system is an important administrative means of population control for the Chinese government. Public resources are allocated in a way that is largely based on *hukou* status (Solinger 1999; Wallace 2014). *Hukou* status is a central component of the local and migrant identities (Afridi, Li, and Ren 2015). A strict *hukou* system sets many restrictions on the conversion of *hukou* status, and maintains a dual system of education (Ling 2015) and labor and housing markets (Wang and Zuo 1999) in cities, which imposes great barriers to intergroup contact between local and migrant residents. Such a *hukou* system is expected to significantly dampen the contact effect in reducing prejudice. This logic leads to the following hypothesis about contact:

Contact hypothesis 4: The contact effect will be weaker in regions where the hukou system is highly exclusive.

Methods

We combine social media data with census and yearbook data and employ a mixed-method approach for data analysis. In particular, we utilize supervised learning models to classify a large corpus of Weibo messages and create a measure of online expressions of regional prejudice. For hypothesis testing, we rely on linear regression models to examine the relationship between migrant inflows, *hukou* policy, and regional prejudice expressed on social media.

WEIBO DATA

Twitter messages have been used to measure prejudice against Muslims (Awan 2014), immigrants (Sanguinetti et al. 2018), and minority groups in general (Waseem and Hovy 2016). Weibo is a Twitter-like social media site in China with an enormous user group. Between December 14, 2014, and April 15, 2015, we collected over 4 million online messages from Weibo. Each message includes at least one of the 13 filtering words regarding China's internal migrants: "本地" (local), "外地" (non-local), "外来" (outsider), "外省" (other province), "农民工" (peasant worker), "常住人口" (permanent population), "流动人口" (floating population), "暂住人口" (transient population), "户籍" (census register), "户口" (*hukou*), "落户"

(new *hukou* registration), "居住证" (residential permit), and "暂住证" (temporal residential permit). These filters are the most commonly used Chinese words that concern either migrant identity or migration policies. There is little regional variation in the usage of these words. The filtering words themselves are not sensitive and are unlikely to be subject to government censorship. Out of the raw data, we excluded the messages that are duplicated, leaving only 1.4 million messages for subsequent analysis.

We further preprocessed the Weibo textual data for the measurement of online expressions of regional prejudice. There are two forms of Weibo messages: initial posts and reposts.[2] Weibo reposts follow the format "comment//@ initial post." In a repost, one can, but is not required to, add a comment of up to 140 Chinese characters in length along with the initial post. We kept all of the initial posts, and for the reposts retained only the comment (i.e., the text before//@). This allows reposters' personal opinions to be better represented and reduces the measurement error of online expressions of regional prejudice caused by the redundant text. Out of the 1.4 million messages, we filtered out those with empty comments (10.2 percent) and those that are primarily news articles (65.9 percent), advertisements (2.7 percent), scientific findings (0.1 percent), and anything else irrelevant (7.7 percent). We also filtered out the 8.4 percent Weibo posters whose profile pages hide their locations of residence. As a result, about 186,000 messages remained for further analysis; a supervised learning model indicated that 81.7 percent of these messages were composed by local Weibo users, so the prejudice contained in the messages in a large part reflects local users' prejudice toward migrants, not vice versa.

ONLINE EXPRESSIONS OF REGIONAL PREJUDICE

City serves as the unit of analysis. The cities include both prefecture-, subprovincial-, and provincial-level cities. In building the measures of online expressions of regional prejudice, we started with classifying individual Weibo messages in terms of sentiment, then obtained the city-level measures via aggregation.

The classification of individual messages involved supervised learning (figure 1). First, we randomly sampled 7,000 messages as the annotated set from the 186,000 preprocessed messages. Each message in the annotated set was coded by three college students who are active Weibo users.[3] We provided standard instructions and examples to guide the coding process (see Appendix 1). A message was coded "negative" if it expressed a negative attitude toward migrants and "not negative" if it did not. Expressing a

2. In the data set examined, 84.0 percent of the Weibo messages are initial posts, and 16.0 percent are reposts.

3. The intercoder reliability score is 70.1 percent in terms of Fleiss's Kappa (Fleiss 1971).

Figure 1. Workflow of the classification of the Weibo messages.

negative attitude was considered as a proxy of regional prejudice expressed on social media. When the three coders disagreed, the message was labeled in accordance with the majority decision.

The annotated messages were then utilized to train sentiment classifiers. The two categories of the messages are not balanced in proportion: the "not negative" messages constitute 89 percent, but the "negative" messages only constitute 11 percent. Hence, we oversampled the "negative" messages during the training process. We included some extra features in the classifiers: one set of the features comprises the word vectors generated by the word2-vec word-embedding technique (Mikolov et al. 2013; Goldberg and Levy 2014), and the other set of features comprises a topic identifier based on a 10-topic latent Dirichlet allocation model.[4]

We built three classifiers using different machine-learning algorithms (ensemble model, extra tree, and gradient boost), and validated their performance using 10-fold cross-validation (Appendix 2). The ensemble model has the best performance in terms of precision, recall, and F-measure, and was chosen to categorize the sentiment of the 179,000 unlabeled messages.

A Weibo user may send multiple messages about migration and convey conflicting sentiment toward migrants. In measuring an individual poster's regional prejudice, we took all of the relevant messages into account. In particular, a user is labeled "prejudiced" if the "negative" messages outnumber the "not negative" messages. For the dependent variable, we measure online expressions of regional prejudice at the city level by calculating the percentage of posters that are labeled "prejudiced." Figure 2 maps online expressions of regional prejudice across the cities of China. Some geographic patterns appear. It seems that online expressions of regional prejudice are concentrated in central China and dissipate in the peripheral areas, which cover the coastal regions in the east and autonomous regions in the west. Marketization was initiated in the coastal regions, which witnessed the first wave of migrant inflows in post-reform China. In addition, autonomous regions are home to China's largest minority peoples, such as the Hui, Tibetan, and Uyghur, and have a long history of inter-ethnic interactions. The peripheral areas are where the contact between distinct social groups takes place most frequently.

MIGRANT GROUP SIZE

Migrant group size is measured as the percentage residents of the city who are migrants in the year 2015. The data source is the 2015 1% Chinese Population Survey. This survey is also called the 2015 Minute Census,

4. The LDA models with different numbers of topics were compared. The 10-topic model has the best performance in terms of partition and interpretability. LDA is a baseline topic model and has some limitations. There are some alternative topic models (e.g., structural topic models) that can incorporate covariates and outperform LDA on short texts (Roberts et al. 2014).

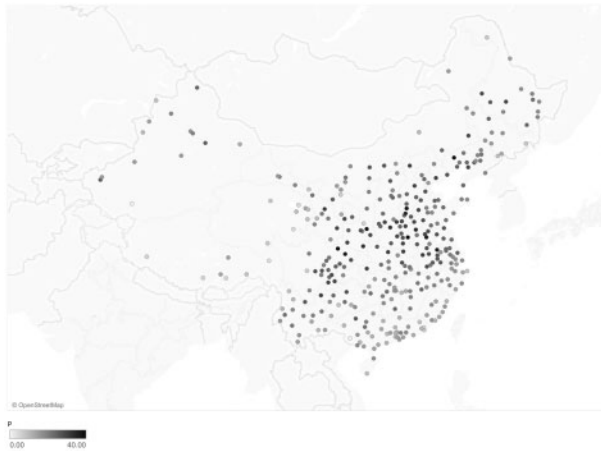

Figure 2. Distribution of online expressions of regional prejudice across Chinese cities. Each dot represents a Chinese city. The red dots indicate the cities with a relatively high level of online expressions of regional prejudice, and the blue dots indicate the opposite.

because it includes all of the statistical items in the decennial census but involves collecting data from a sample of only 1 percent of the population. To acquire the dataset of the 2015 Minute Census for each province, we searched through a number of online databases and libraries, and contacted many provincial bureaus of statistics. However, we were not able to obtain access to the dataset for Shandong Province and three autonomous regions—Tibet, Xinjiang, and Inner Mongolia. The 49 cities within these four regions are therefore excluded for our analysis.

In addition to total migrant group size, it is possible to assess the size of intraprovince and interprovince migrant groups. Intraprovince migrants move to a new city that is in their home province. Interprovince migrants move out of their home province. Interprovince migrants are further divided into adjacent- and disjoint-province migrants. As the names suggest, adjacent-province migrants flow into a neighboring province and disjoint-province migrants travel beyond the neighboring provinces.[5] It is also possible to assess migrant group size separately for migrants who migrated for different

5. Although the Chinese government has taken efforts to facilitate regional integration, there is no city straddling the border between two provinces so far, which makes it relatively easy to distinguish between migrants from adjacent and disjoint provinces. For the cities in Guangxi province, we were only able to acquire the data about intra- and interprovince migrants in general rather than the data disaggregated by specific origin provinces. Therefore, the variables concerning adjacent and disjoint provinces are missing for these cities.

Table 1. Descriptive statistics of variables

	N	Min	Max	Mean	Standard deviation
Prejudice	336	0	40	19.20	7.50
Origin of migration					
Intra-province	286	0	40.98	7.01	6.18
Interprovince	286	0	57.09	4.56	8.03
Adjacent-province	272	0	30.26	1.91	3.57
Disjoint-province	272	0	31.09	2.78	5.05
Purpose of migration					
Job	286	0.08	48.43	3.22	6.42
Family	286	0.05	11.85	1.12	1.67
School	286	0	9.96	0.53	1.05
(Super-)mega city	336	0	1	0.04	0.20
Socioeconomic environment					
GDP per capita	336	0.88	28.66	4.82	3.24
College graduate	286	3.39	42.34	11.25	6.45
Ethnic minority	286	0.03	99.43	12.81	23.22

NOTE.—Full sample of cities. (Super-)mega city is a binary variable. All the other variables except GDP per capita are measured in percentage.

reasons. The 2015 Minute Census shows that the most common purpose of migration is jobs, followed by family and education (table 1). Migrants were considered laborers if the purpose of their migration was for jobs.

HUKOU POLICY

The restrictiveness of the *hukou* policy is largely dependent on the size of the city. In November 2014 the State Council issued the "Notice on Adjusting the Category of City Size." This official document sets out the criteria of categorizing cities in terms of urban resident population: small city (< 0.5 million), middle-sized city ($>= 0.5$ million and < 1 million), type I big city ($>= 1$ million and < 3 million), type II big city ($>= 3$ million and < 5 million), mega city ($>= 5$ million and < 10 million), and super-mega city (≥ 10 million). An important official guide for *hukou* policy reforms is the "National New Type Urbanization Plan (2014–2020)" issued by the National Development and Reform Commission in March 2014 (Wang et al. 2015). It proposes varying objectives of *hukou* policy reforms for each category of cities. Although the collection of Weibo data started before the two official documents, they reflect the earlier practice of *hukou* administration. We refer to the 2015 Yearbook of Urban and Rural Construction for the size

of urban resident population,[6] and rely on the State Council's categorization scheme to classify the 336 cities in our dataset. As a result, there are 81 small cities, 100 middle-sized cites, 126 type I big cities, 15 type II big cities, 10 mega cities, and 4 super-mega cities. The major gap in *hukou* policy reforms exists between (super-)mega and other types of cities. The national government designed very different goals for *hukou* administration with regard to these two groups of cities: small, middle-sized, and big cities are encouraged to relax and even eliminate *hukou*-based restrictions; however, mega and super-mega cities are required to maintain stricter population control policies than other cities. Therefore, we create a binary variable as a proxy for how restrictive the *hukou* system operates. It is coded zero for small, middle-sized, and big cities, and one for mega and super-mega cities.

SOCIOECONOMIC ENVIRONMENT

Previous studies have shown the crucial role of socioeconomic environment in intergroup and interpersonal relations (Branton and Jones 2005). In our analysis on online expressions of regional prejudice, we also consider a set of socioeconomic indicators at the city level as control variables. In particular, we measure the level of economic development as GDP per capita, the level of education as the percentage of college graduates, and ethnic composition as the percentage of ethnic minority residents. The first measure is based on the 2015 Statistical Yearbook of each province, and the last two measures are based on the 2015 Minute Census. The descriptive statistics of the variables are presented in table 1.

Models and Results

In this section, we employ linear regression models to test the threat and contact hypotheses. The unit of analysis is city. Some small-sized cities have a limited number of observations, and their prejudice measure is likely to be driven just by a few Weibo users. In the regression analysis, therefore, we exclude 119 cities where fewer than 100 Weibo users revealed an opinion toward migrants. Our final analysis is based on a total of 191 cities (15 small cities, 58 middle-sized cities, 92 type I big cities, 12 type II big cities, 10 mega cities, and 4 super-mega cities). The dependent variable is Online Expression of Regional Prejudice. The key independent variable is Migrant Group Size, which is also decomposed by the origin and purpose of migration. Among the relevant variables, the proportions of interprovince, adjacent province, disjoint-province, and job-seeking migrants are highly correlated

6. The 2015 Yearbook of Urban and Rural Construction was accessed from http:// www. mohurd. gov. cn/ xytj/ tjzljsxytjgb/ jstjnj/ .

with a correlation coefficient larger than 0.8 (Appendix 3). It seems that these variables are conceptually indistinguishable and are just different measures for the concept of typical migrants, the migrants who travel over a long distance for finding a job. Concerned with the problem of multicollinearity, we choose to include the highly correlated variables in separate regression models. (Super-)mega City is a moderating variable that interacts with Migrant Group Size. Three variables—GDP per Capita, College Graduate, and Ethnic Minority—control the impact of socioeconomic environment.

The results from the linear regression models generally support the contact hypotheses (see table 2). Model 1 tests the overall effect of migrant group size on online expressions of regional prejudice. We find that a 1 percent increase in the migrant population reduces online expressions of regional prejudice by 0.16 percent, controlling the socioeconomic variables.[7] This finding is consistent with the first contact hypothesis and rejects the corresponding threat hypothesis. Model 2 is designed to evaluate the second threat and contact hypotheses concerning the purpose of migration. This model finds evidence supporting the contact theory. It shows that there is a negative association between the number of migrant laborers and regional prejudice expressed on social media. Nonetheless, we find no such an association with regard to migrant students and family members.

The third threat and contact hypotheses posit that the effect of migrant group size on online expressions of regional prejudice grows when it is measured in terms of long-distance migrants, but the two hypotheses do not agree on the direction of the effect. Model 3 lends support to the contact interpretation that the effect is negative, such that the contact effect is more pronounced for inter- (rather than intra-) province migrants. As Model 3 shows, intra-province migrants have no effect on online expressions of regional prejudice but interprovince migrants have a negative effect (-0.20) that is significant at the 99 percent confidence level and is higher than the effect of migrants in general (-0.16). Model 4 decomposes interprovince migrants into adjacent- and disjoint-province migrants, neither of which, however, is of statistical significance. The Vuong test, which is designed for comparing a pair of non-nested models (Vuong 1989), further shows that Model 3 and Model 4 are in fact indistinguishable, with p-value close to 1. These findings suggest that local residents are apt to perceive long-distance migrants and migrant workers as typical migrants and that more contact with

7. Given that the labor market and education resources in China are intensely competitive, regional prejudice is likely to be tolerated by Chinese migrants and just a minor consideration of them for choosing the destination city. The positive correlation here is thus more likely to be driven by the causal effect of migrant flows on online expressions of regional prejudice, rather than the other way around.

Table 2. Linear regression analysis of online expressions of regional prejudice

Independent variables	Model 1	Model 2	Model 3	Model 4	Model 5
Migrant group size	−0.16**				−0.20**
	(0.06)				(0.06)
Origin of migration					
Intra-province		−0.10	−0.08	−0.09	
		(0.11)	(0.11)	(0.11)	
Interprovince			−0.20**		
			(0.07)		
Adjacent-province				−0.07	
				(0.22)	
Disjoint-province				−0.27	
				(0.15)	
Purpose of migration					
Job		−0.22*			
		(0.10)			
Family		0.08			
		(0.44)			
School		0.16			
		(0.72)			
(Super-)mega city					−8.94*
					(4.28)
Migrant group size					0.23*
× (Super-)mega city					(0.11)
Socioeconomic environment					
GDP per capita	−0.42	−0.38	−0.37	−0.44	−0.40
	(0.27)	(0.30)	(0.28)	(0.28)	(0.27)
College graduate	0.12	0.04	0.07	0.08	0.15
	(0.10)	(0.14)	(0.11)	(0.12)	(0.10)
Ethnic minority	−0.14**	−0.15**	−0.14**	−0.16**	−0.14**
	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)
Constant	24.34**	24.51**	24.26**	24.92**	24.39**
	(1.20)	(1.24)	(1.20)	(1.21)	(1.25)
R^2	0.16	0.16	0.17	0.19	0.19
N	191	191	191	184	191

NOTE.—The dependent variable of all the five models is the percentage of the Weibo users in a city who expressed prejudice toward migrants. The cities with less than 100 Weibo users revealing opinion toward migrants are excluded. Standard errors in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$.

typical migrants is most likely to lead to reduced expressions of regional prejudice.

Model 5 tests the last threat and contact hypotheses, which concern how the *hukou* system moderates the effect of migrant group size on online

expressions of regional prejudice. We use mega city status as a proxy for the restrictiveness of *hukou* policy. The independent effect of this variable on online expressions of regional prejudice is significantly negative. However, its moderating effect tells a different story. We include the interaction term between migrant group size and mega city status in Model 5. The positive coefficient of the interaction term suggests that strict *hukou* policy dampens the contact effect and even creates a small threat effect. To be specific, a 1 percent increase in the migrant population raises online expressions of regional prejudice by 0.03 percent in mega or super-mega cities, whereas it reduces online expressions of regional prejudice by 0.20 percent in other types of cities. Because of the frequent interaction between diverse groups, mega and super-mega cities have a natural tendency to develop a culture of intergroup tolerance. This notwithstanding, the highly restrictive *hukou* system sustained in these cities generates a negative dynamic of migrant inflows that breeds online expressions of regional prejudice. In sum, our regression analysis provides strong support for the contact theory that migrant inflows reduce online expressions of regional prejudice. In general, migrant group size decreases online expressions of regional prejudice. The contact effect is stronger when migrant group size is measured in terms of migrant labor and interprovince migrants, which are likely to be seen as typical migrants by local residents. Our research also supports one of Allport's optimal contact conditions—the support of authorities, law, or custom (Allport, Clark, and Pettigrew 1954). The contact effect is only observed in small, middle-sized, and big cities in which the *hukou* system is less strict. The contact effect disappears and turns into a small threat effect in mega and super-mega cities where the *hukou* system is strictly maintained and shows no sign of relaxation in the near future. We replicated the regression analysis using the full sample of cities. The results from this analysis are shown in Appendix 4 and are generally consistent with the findings with the sample of larger cities, except that they show no strong evidence for the moderating effect of the *hukou* system.

Discussion and Conclusions

Two questions led our research: (1) Do migrant inflows increase or decrease online expressions of regional prejudice? (2) If migrant inflows indeed have an effect on online expressions of regional prejudice, what conditions can facilitate this effect to a greater extent? To answer these two questions, we used a mixed-method strategy (using machine-learning tools to construct city-level measures of prejudice toward migrants from social media posts) to test hypotheses generated by both group-threat (Blalock 1967) and group-contact (Allport, Clark, and Pettigrew 1954; Pettigrew and Tropp 2006) theories. Consistent with group-contact theory, migrant group size was negatively

associated with regional prejudice expressed on social media. The contact effect was found to be larger with regard to migrant workers and interprovince migrants, who are often perceived as typical migrants in the eyes of local residents. Moreover, we found that the contact effect only exists in China's relatively small cities, where the *hukou* conversion policy is relaxed. The contact effect, however, disappears and becomes a small threat effect in mega and super-mega cities, such as Beijing and Shanghai, where the local government is required to maintain strong population control.

These findings provide new insight into the seemingly competing arguments of the group-threat and contact theories. Our research implies that the direction of the effect of minority group size on intergroup prejudice depends on whether a correspondence exists between minority group size and the likelihood of interaction between majority and minority group members. In this sense, the intergroup threat and contact effects are context-based. As the context is being reshaped, the threat effect may wither away and even turn into the contact effect, and vice versa. Our research findings also support one of the proposed optimal contact conditions—the support of authorities, law, and custom (Allport, Clark, and Pettigrew 1954). This condition suggests that authorities play an important role in creating a social context that facilitates intergroup contact. The Chinese government, because of its concern with regional imbalance in economic development, allows relatively small cities to loosen their *hukou* systems (Wang et al. 2015). Because of this political decision, new migrants face lower barriers to being integrated into the destination society and are more likely to have intimate contact with local residents. Owing to the *hukou* restrictions, China's internal migration resembles international immigration (Wang and Zuo 1999). The findings from our research should have particular interest to policymakers in China and other countries who are concerned with the increasing tensions between native and immigrant groups.

Our research also shows the promise of measuring attitudes via social media data. Traditional social science research mainly relies on social surveys and experiments to measure opinions, attitudes, and psychological traits. Instead, we employed supervised learning models and built an attitudinal measure based on a large collection of social media messages. Such a measure has some advantages over survey- and experiment-based measures, particularly for measuring sensitive subjective constructs such as prejudice. First, social media users may tend to express their opinions more truthfully in the absence of researchers and in a more anonymous environment. Second, social media messages reflect the attitudes that people feel and express in real life, rather than those elicited in an unrealistic setting by researchers. Finally, the size of social media data is huge, which makes it possible to create fine-grained aggregate-level measures of attitudes. For example, the Weibo data allowed us to measure online expressions of regional prejudice for each Chinese city.

Our research is subject to some limitations. First, heavy internet censorship in China causes the problem of missing data, a common problem in the study of Chinese social media. This problem may lead to a measurement error concerning online expressions of regional prejudice. A possible but difficult solution is to continuously track updates on social media (King, Pan, and Roberts 2013). Another source of measurement error is the complex nature of textual data. Weibo messages involve high uncertainty in their meanings because of sarcasm, metaphor, simile, satire, and rapidly changing jargon and "insider" language, which only specific online communities can understand. No machine-learning method, however, has been invented to achieve the accuracy at the level of a sophisticated human coder in processing complex text data. Third, Weibo users, especially active Weibo users, are not representative of the entire Chinese population (Zhang 2018). They tend to be younger, better educated, and have more online and face-to-face contact with different social groups. They are likely to be less prejudiced toward migrants than those who rarely use social media and who choose to keep silent on social media. The strength of social survey research is that it can obtain a representative sample of the target population through a well-designed sampling process and measures of personal characteristics such as age, educational attainment, and social ties. Therefore, our future research plan is to test whether the findings from this study can be replicated using nationally representative survey data.

Appendix 1. Annotation examples of Weibo messages

Weibo message (translated)	Label
What are people thinking on the Internet! This tragedy could have been avoided by just doing one thing, that is, have the migrants go back to wherever they came from and give Shanghai back to the Shanghaiese... Everything will be perfect...	Negative
We are living in the same city. Please be tolerant of one another. Maybe you come here from far away. Maybe you were born here. Maybe you are a Shenzhen native. Maybe you are a newcomer to this city.	Not negative
Crowded is only an excuse, the ladder to the death is wide enough. The event of Shanghai Stampede[1] exactly reflects the lacking ability of local police in public control. The news should not only pay attention to the cold numbers, but also should include humanism in reporting. The behavior of money drop must be the reason of the stampede, but the one who was eulogized might be the creator of the accident. Finding out the person directly responsible for the accident is more important than the compensation.	Not negative

Appendix 2. Performance scores of the Weibo classifiers

Classifier	Precision	Recall	F measure
Ensemble model	90.4%	86.6%	88.0%
Extra tree	79.6%	85.4%	81.7%
Gradient boost	83.0%	84.0%	83.5%

NOTE.—The performance scores of the prejudice classifier are based on the 7,000 labeled messages rather than the full Weibo or our entire dataset. They are also the averaged results from 10-fold cross-validation.

Appendix 3. Correlation between migrant-relevant variables

	Migrant group size	Intra-province	Inter province	Adjacent-province	Disjoint-province	Job	Family	School
Migrant group size	1.00							
Intra-province	0.78	1.00						
Inter-province	0.90	0.43	1.00					
Adjacent-province	0.84	0.40	0.93	1.00				
Disjoint-province	0.87	0.41	0.97	0.80	1.00			
Job	0.86	0.37	0.98	0.94	0.93	1.00		
Family	0.60	0.37	0.61	0.54	0.60	0.66	1.00	
School	0.51	0.60	0.32	0.25	0.34	0.29	0.54	1.00

NOTE.—The cities with less than 100 Weibo users revealing opinion toward migrants are excluded.

Appendix 4. Linear regression analysis of online expressions of regional prejudice (full sample of cities)

	Model 1	Model 2	Model 3	Model 4	Model 5
Migrant group size	-0.12*				−0.15**
	(0.05)				(0.05)
Origin of migration					
Intra-province		−0.02	−0.00	−0.02	
		(0.10)	(0.09)	(0.09)	
Interprovince			-0.18**		
			(0.07)		
Adjacent-province				−0.10	
				(0.20)	
Disjoint-province				−0.22	
				(0.14)	
Purpose of migration					
Job		−0.20*			
		(0.09)			
Family		0.09			
		(0.35)			
School		0.04			
		(0.62)			
(Super-)mega city					−7.28
					(4.16)
Migrant group size					0.19
× (Super-)mega City					(0.11)
Socioeconomic environment					
GDP per capita	−0.42	−0.38	−0.36	−0.42	−0.40
	(0.23)	(0.26)	(0.23)	(0.23)	(0.23)
College graduate	0.09	0.01	0.03	0.03	0.11
	(0.09)	(0.12)	(0.10)	(0.10)	(0.09)
Ethnic minority	−0.14**	−0.14**	−0.14**	−0.15**	−0.14**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	23.49**	23.57**	23.39**	23.99**	23.49**
	(0.99)	(1.03)	(0.99)	(1.00)	(1.04)
R^2	0.21	0.21	0.22	0.24	0.22
N	286	286	286	272	286

NOTE.—Full sample of cities. Standard errors in parentheses.

*$p < 0.05$

**$p < 0.01$

Data Availability Statement

REPLICATION DATA AND DOCUMENTATION will be available within 12 months of publication at https://doi.org/10.7910/DVN/HL1LH8.

References

Afridi, Farzana, Sherry Xin Li, and Yufei Ren. 2015. "Social Identity and Inequality: The Impact of China's Hukou System." *Journal of Public Economics* 123:17–29.

Alba, Richard, Rubén G. Rumbaut, and Karen Marotz. 2005. "A Distorted Nation: Perceptions of Racial/Ethnic Group Sizes and Attitudes toward Immigrants and Other Minorities." *Social Forces* 84:901–19.

Allport, Gordon W., Kenneth Clark, and Thomas Pettigrew. 1954. *The Nature of Prejudice*. Cambridge: Addison-Wesley Publishing Company.

Aronson, Elliot. 1978. *The Jigsaw Classroom*. Oxford: Sage.

Awan, Imran. 2014. "Islamophobia and Twitter: A Typology of Online Hate against Muslims on Social Media." *Policy & Internet* 6:133–50.

Blalock, Hubert M. 1967. *Toward a Theory of Minority-Group Relations*. New York: Wiley.

Blumer, Herbert G. 1958. "Race Prejudice as Sense of Group Position." *Pacific Sociological Review* 1:3–7.

Bobo, Lawrence. 1983. "Whites' Opposition to Busing: Symbolic Racism or Realistic Group Conflict?" *Journal of Personality & Social Psychology* 45:1196–1210.

Branton, Regina P., and Bradford S. Jones. 2005. "Reexamining Racial Attitudes: The Conditional Relationship between Diversity and Socioeconomic Environment." *American Journal of Political Science* 49:359–72.

Chan, Kam Wing, and Will Buckingham. 2008. "Is China Abolishing the Hukou System?" *China Quarterly* 195:582–606.

Cheng, Tiejun, and Mark Selden. 1994. "The Origins and Social Consequences of China's Hukou System." *China Quarterly* 139:644–68.

China Internet Network Information Center. 2018. *The 2018 Statistical Report on China's Internet Development*. http://cnnic.cn/gywm/xwzx/rdxw/201801/t20180131_70188.htm.

Fleiss, Joseph L. 1971. "Measuring Nominal Scale Agreement among Many Raters." *Psychological Bulletin* 76:378–82.

Goldberg, Yoav, and Omer Levy. 2014. "Word2vec Explained: Deriving Mikolov et al.'s Negative-Sampling Word-Embedding Method." *ArXiv:1402.3722 [Cs, Stat]*. http://arxiv.org/abs/1402.3722.

Gu, Jun, Ingrid Nielsen, Jason Shachat, Russell Smyth, and Yujia Peng. 2016. "An Experimental Study of the Effect of Intergroup Contact on Attitudes in Urban China." *Urban Studies* 53:2991–3006.

Hainmueller, Jens, and Dominik Hangartner. 2013. "Who Gets a Swiss Passport? A Natural Experiment in Immigrant Discrimination." *American Political Science Review* 107:159–87.

Hewstone, Miles. 2000. "Contact and Categorization: Social Psychological Interventions to Change Intergroup Relations." In *Stereotypes and Prejudice: Essential Readings*, edited by Charles Stangor, 394–418. New York: Psychology Press.

Kinder, Donald R., and David O. Sears. 1981. "Prejudice and Politics: Symbolic Racism versus Racial Threats to the Good Life." *Journal of Personality & Social Psychology* 40:414–31.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107:326–43.

Knight, John, Lina Song, and Jia Huaibin. 1999. "Chinese Rural Migrants in Urban Enterprises: Three Perspectives." *Journal of Development Studies* 35:73–104.

Ling, Minhua. 2015. "'Bad Students Go to Vocational Schools!' Education, Social Reproduction and Migrant Youth in Urban China." *China Journal* 73:108–31.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Efficient Estimation of Word Representations in Vector Space." *ArXiv:1301.3781 [Cs]*. http://arxiv.org/abs/1301.3781.

National Bureau of Statistics. 2011. *The Database of 2010 Chinese Population Census*. Beijing: Chinese Statistics Press.

National Bureau of Statistics. 2019. *The 2018 Statistical Report of Chinese National Economic and Social Development.* http://www.stats.gov.cn/tjsj/zxfb/201902/t20190228_1651265.html.

Nielsen, Ingrid, Chris Nyland, Russell Smyth, Mingqiong Zhang, and Cherrie Jiuhua Zhu. 2006. "Effects of Intergroup Contact on Attitudes of Chinese Urban Residents to Migrant Workers." *Urban Studies* 43:475–90.

Pakrashi, Debayan, and Paul Frijters. 2017. "Migration and Discrimination in Urban China: A Decomposition Approach." *Review of Income and Wealth* 63:821–40.

Pettigrew, Thomas F. 1997. "Generalized Intergroup Contact Effects on Prejudice." *Personality and Social Psychology Bulletin* 23:173–85.

———. 1998. "Intergroup Contact Theory." *Annual Review of Psychology* 49:65–85.

Pettigrew, Thomas F., and Linda R. Tropp. 2006. "A Meta-Analytic Test of Intergroup Contact Theory." *Journal of Personality and Social Psychology* 90:751–83.

Pettigrew, Thomas F., Linda R. Tropp, Ulrich Wagner, and Oliver Christ. 2011. "Recent Advances in Intergroup Contact Theory." *International Journal of Intercultural Relations* 35:271–80.

Pottie-Sherman, Yolande, and Rima Wilkes. 2017. "Does Size Really Matter? On the Relationship between Immigrant Group Size and Anti-Immigrant Prejudice." *International Migration Review* 51:218–50.

Reich, Carol, and Mary Purbhoo. 1975. "The Effect of Cross-Cultural Contact." *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement* 7:313–27.

Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58:1064–82.

Sanguinetti, Manuela, Poletto Fabio, Bosco Cristina, Patti Viviana, and Stranisci. Marco 2018. "An Italian Twitter Corpus of Hate Speech against Immigrants." Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 1768–75. Miyazaki, Japan: European Language Resources Association (ELRA). Available at https://www.aclweb.org/anthology/L18-1279.pdf.

Solinger, Dorothy J. 1999. *Contesting Citizenship in Urban China: Peasant Migrants, the State, and the Logic of the Market*. Berkeley: University of California Press.

Stephan, Walter G., and Cookie W. Stephan. 1984. "The Role of Ignorance in Intergroup Relations." In *Groups in Contact: The Psychology of Desegregation*, edited by Norman S. Miller and Marilynn B. Brewer, 229–55. Orlando, FL: Academic Press.

———. 2000. "An Integrated Threat Theory of Prejudice." In *Reducing Prejudice and Discrimination*, edited by Stuart Oskamp, 23–45. Hove, East Sussex: Psychology Press.

Stephan, Walter G., Oscar Ybarra, Carmen Martnez Martnez, Joseph Schwarzwald, and Michal Tur-Kaspa. 1998. "Prejudice toward Immigrants to Spain and Israel: An Integrated Threat Theory Analysis." *Journal of Cross-Cultural Psychology* 29:559–76.

Tropp, Linda R., and Thomas F. Pettigrew. 2005. "Relationships between Intergroup Contact and Prejudice among Minority and Majority Status Groups." *Psychological Science* 16:951–57.

Tse, Chun Wing. 2016. "Urban Residents' Prejudice and Integration of Rural Migrants into Urban China." *Journal of Contemporary China* 25:579–95.

Vuong, Quang H. 1989. "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses." *Econometrica* 57:307–33.

Wallace, Jeremy. 2014. *Cities and Stability: Urbanization, Redistribution, and Regime Survival in China*. New York: Oxford University Press.

Wang, Feng, and Xuejin Zuo. 1999. "Inside China's Cities: Institutional Barriers and Opportunities for Urban Migrants." *American Economic Review* 89:276–80.

Wang, Haining, Fei Guo, and Zhiming Cheng. 2015. "A Distributional Analysis of Wage Discrimination against Migrant Workers in China's Urban Labour Market." *Urban Studies* 52:2383–2403.

Wang, Xin-Rui, Eddie Chi-Man Hui, Charles Choguill, and Sheng-Hua Jia. 2015. "The New Urbanization Policy in China: Which Way Forward?" *Habitat International* 47:279–84.

Waseem, Zeerak, and Dirk Hovy. 2016. "Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter." Proceedings of the NAACL Student Research Workshop, 88–93. San Diego, CA: Association for Computational Linguistics.

Weibo Data Center. 2017. "The 2017 Report of Weibo Users' Development.” Available at http://data.weibo.com/report/reportDetail?id=404&sudaref=www.baidu.com&display=0&retcode=6102.

Wu, Xiaogang, and Donald J. Treiman. 2007. "Inequality and Equality under Chinese Socialism: The Hukou System and Intergenerational Occupational Mobility." *American Journal of Sociology* 113:415–45.

Xue, Sen. 2018. "Does Contact Improve Attitudes towards Migrants in China?" *Economics of Transition and Institutional Change* 26:149–200.

Zhang, Yang. 2018. "The Internet in China: New Methods and Opportunities." *Journal of Chinese Political Science* 23:121–33.

Zhang, Zhuoni, and Xiaogang Wu. 2017. "Occupational Segregation and Earnings Inequality: Rural Migrants and Local Workers in Urban China." *Social Science Research* 61:57–74.

Zhou, Xiang, and Yu Xie. 2019. "Market Transition, Industrialization, and Social Mobility Trends in Postrevolution China." *American Journal of Sociology* 124:1810–47.

Zhou, Yisu, and Dan Wang. 2016. "Understanding the Constraints on the Supply of Public Education to the Migrant Population in China: Evidence from Shanghai." *Journal of Contemporary China* 25:563–78.

Zhu, Boliang. 2018. "Economic Considerations and Attitudes toward Migrant Workers." SSRN Scholarly Paper ID 2819056. Rochester: Social Science Research Network. https://papers.ssrn.com/abstract=2819056.