Information, Mobile Communication, and Referral Effects

Panle Jia Barwick* Yanyan Liu† Eleonora Patacchini‡ Qi Wu§

Abstract

This paper uses the universe of cellphone records from a Chinese telecommunication provider for a northern Chinese city to examine the role of information exchange in urban labor markets. We provide the first direct evidence of increased communication among referral pairs around job changes. Information provided by social contacts mitigates information asymmetry and improves labor market performance.

Keywords Information, Mobile Communication, Urban Labor Market, Social Networks

JEL Classification: R23, J60, L15

*Department of Economics, Cornell University and NBER. Email: panle.barwick@cornell.edu

†International Food Policy Research Institute. Email: Y.Liu@cgiar.org

‡Department of Economics, Cornell University and CEPR, EIEF, IZA. Email: ep454@cornell.edu

§Guanghua School of Management, Peking University. Email: qiwu@gsm.pku.edu.cn. We thank the editor, three anonymous referees, Susan Athey, Patrick Bayer, Giacomo De Giorgi, Jessie Handbury, Tatsiramos Konstantinos, Mike Lovenheim, Michele Pellizzari, Steve Ross, and various seminar participants for helpful comments. Barwick acknowledges generous support from the National University of Singapore during her sabbatical visit.

1 Introduction

A society's ability to disseminate and exchange job-related information among firms and job seekers crucially determines its labor market outcomes. The existing literature has documented the importance of social connections for job information dissemination and referrals (Ioannides and Loury, 2004). However, the extent and nature of information exchange between referrers and referees are not well understood, mainly because they are difficult to measure.

The wide prevalence of location-aware and Global Positioning System (GPS) technologies in mobile phone devices offers researchers a novel way to quantify the extent of information flow among individuals while also tracking their movements in physical space. Datasets derived from these geocoded phone communication records present three unique advantages over traditional data. First, such records trace a more accurate profile of individuals' social networks than do surveys commonly used in the literature. Second, the frequency and intensity of call records provide a direct measure of information exchange. Third, the panel character of these datasets makes it feasible to follow individuals over time and space and control for individual unobserved attributes.

In this paper, we analyze the impact of information exchange on labor market outcomes. Our analysis is made feasible by a unique dataset that contains the universe of de-identified and geocoded cellphone records from a major Chinese telecommunication service provider in a northern Chinese city over the course of 12 months.[1] These detailed records enable us to construct individuals' social networks, measures of information flow among individuals, and variables on their employment status, history of work locations, home locations, and demographic attributes. We supplement these phone records with administrative data on firm attributes (industry classification and payroll) and auxiliary datasets on residential housing prices and job postings to obtain additional socioeconomic measures.

We first present descriptive evidence that information flow, measured by the frequency of phone calls, exhibits spatial patterns similar to and correlates strongly with worker flows. Our core analysis examines the role of job-related information shared by social contacts (friends) on job change. When an individual moves to a preexisting friend's workplace, we refer to this friend as a referrer. According to our definition, at least one in every four jobs is based on a referral. The referral effect in this analysis is defined as the effect of having social contacts in a given workplace on individuals' work location choices.[2] This effect is quantified by the difference in a job seeker's propensity to choose a friend's workplace versus

[1] We are not allowed to disclose the name of the service provider and the city studied due to the data sharing agreement.

[2] We use *social contacts* and *friends* interchangeably in this paper.

his/her propensity to choose a work location in the same neighborhood but where none of his/her friends works. We control for geospatial attributes that are correlated with job flows (the presence of commercial centers, industrial clusters, etc.) by the interaction of the origin and destination neighborhood fixed effects. In other words, we compare individuals whose old and new jobs are located in the same origin–destination neighborhood pair but who have different social networks, and we examine their choices of workplace locations with and without friends. Having a referrer in a location increases an individual's likelihood of moving there by nearly threefold relative to his/her propensity to choose other work locations—a pattern consistent with those found in previous studies on other countries (Topa, 2001; Burks et al., 2015).

To provide direct evidence that referrers pass job information to job seekers, we use event studies to examine the intensity of information flow between workers and their referrers, proxied by the number of phone calls and call duration. The call volume between job switchers and their referrers exhibits an inverted-U shape that peaks at the time of the job change. In contrast, the information flow between workers and nonreferrer friends remains stable throughout the sample period, with no noticeable differences during the months that precede job switches. As far as we know, this is the first empirical analysis documenting the increasing communication intensity around job changes between referrers and referees.

In terms of effect heterogeneity, referrals are particularly important for young workers, people switching jobs from rural to urban areas, and those who change sectors. These results are in line with the observation that information asymmetries are more severe in these settings and hence that referrals are more valuable. We also provide evidence that stronger social ties are associated with a larger referral effect, corroborating results in the literature on weak versus strong ties. In addition, conditional on the number of phone calls exchanged and the share of strong ties, the diversity of individuals' social contacts also matters. Individuals whose social networks exhibit greater socioeconomic diversity are more likely to use referrals when changing jobs.

We compare our definition of referrers with two commonly used measures of social contacts in the literature, namely residential neighbors (Bayer, Ross and Topa 2008) and individuals belonging to the same ethnic group or immigrant community (Edin, Fredriksson and Aslund 2003), the latter of which is analogous to individuals sharing the same birth county in our setting. We are able to validate these measures using call records. Residential neighbors and people born in the same county are more likely to communicate with each other. Nonetheless, while the effect of referrals identified under these two alternative definitions is positive and significant, it is much lower than our baseline estimate. As a result, the reported estimates in the literature that are derived from these social network proxies are likely to be a lower bound of the true referral effect.

A long-standing challenge in the referral literature that examines observational data is the difficulty in distinguishing a referral effect from homophily and sorting. We conduct an extensive set of robustness analyses to examine alternative explanations for the referral effect, including homophily, unobserved location attributes, local labor market demand, and preference for working with friends. We also examine the extent of reverse causality and repeat the analysis using alternative definitions of friends. Our results hold in these robustness analyses.

Job information passed on via referrals is valuable for workers. Specifically, referral jobs are associated with higher wages and nonwage benefits, shorter commutes, and a greater likelihood of transitioning from part time to full time and from regular jobs to premium ones. Information transmitted through referral networks is also beneficial to firms. Firms receiving referred workers are more likely to have successful recruits, achieve higher matching rates, and experience faster growth. These results provide suggestive evidence that referrals improve labor market efficiency by facilitating better matches between workers and job vacancies.

We conclude our analysis with extensions that shed light on the external validity of our findings. A key premise of our analysis is that mobile phone communications serve as a proxy for the amount of information exchanged among individuals. Although our data lack information on communication via other channels, such as text messages, mobile apps (such as WeChat), and web-based media (such as emails), we present evidence that different information channels are complements: people who frequently communicate via phone conversations are also more likely to use other channels. Next, we replicate our analysis for individuals who experienced unemployment before successfully finding a job. This complements the bulk of our analysis described above that examines individuals who switch from one job to another with minimal job disruption. Event studies reveal remarkably similar patterns in terms of information flow between unemployed individuals and their referrers. The number of phone calls between these referral pairs also exhibits an inverted-U shape that peaks at the time of reemployment. In addition, the estimated referral effect is very close to the baseline estimate derived from individuals without an employment gap. While this analysis is constrained to individuals who find a job within a short time window post-unemployment, these findings provide suggestive evidence that the communication patterns and referral effect that we document are potentially applicable to all job seekers, whether employed or unemployed.

Our work contributes to the emerging literature (Bailey et al., 2018; Glaeser et al., 2015; Donaldson and Storeygard, 2016) that demonstrates how the increasingly available information on individual digital footprints provides a finer-grained picture of social activities (mobility, urban development, etc.). A pioneering study by Henderson, Storeygard and Weil (2012) exploits satellite data to conduct an analysis on urban economic activities at a finer level of spatial disaggregation than is typical of traditional studies. Using predicted travel time from Google Maps, Akbar et al. (2018) construct city-level vehicular mobility indices for 154 Indian cities and propose new methodologies to improve our understanding of urban development. Existing research has documented that mobile phone usage predicts human mobility (Gonzalez, Hidalgo and Barabasi, 2008), migration (Blumenstock, Chi and Tan, 2019), poverty and wealth (Blumenstock, Cadamuro and On, 2015), credit repayment (Bjorkegren and Grissen, 2018), restaurant choices (Athey et al., 2018), and residential location choices (Buchel et al., 2019). Our work contributes to this literature by combining mobile phone records with traditional socioeconomic data to shed light on urban labor market mobility at fine geographical and temporal scales.

Another relevant strand of literature examines the role of social networks in job searches (Topa, 2011; Schmutte, 2016). To identify referred workers, this literature uses surveys or assumes interactions and exchange of job information between certain social ties, such as former colleagues (Cingano and Rosolia 2012; Glitz 2017; Saygin, Weber and Weynandt 2018), family ties (Kramarz and Skans 2014), individuals who belong to the same immigrant community or ethnic group (Edin, Fredriksson and Aslund 2003; Munshi and Rosenzweig 2013; Beaman 2012; Dustmann et al. 2016; Aslund, Hensvik and Skans 2014), residential neighbors (Bayer, Ross and Topa 2008; Hellerstein, McInerney and Neumark 2011; Hellerstein, McInerney and Neumark 2014; Schmutte 2015), and Facebook friends (Gee, Jones and Burke 2017). The paper closest to ours is Bayer, Ross and Topa (2008), who also study the importance of referral effects in an urban market. Using census data on residential and employment locations, they document that individuals who reside in the same city block are more likely to work together than those who live in nearby blocks, and the authors interpret these findings as evidence of social interactions. We contribute to this literature by providing a more refined measure of social networks and presenting direct evidence of information exchange among referral pairs. We also introduce complementary data on vacancies and firm attributes to cover a diverse set of economic outcomes.

Our study is also related to the literature on weak versus strong ties. The seminal study by Granovetter (1973) argues that weak ties could be more important because of their access to a diverse set of information. [3] Considering Facebook users from 50 countries, Gee et al. (2017) document that strong ties are more important than weak ties in job search at the

[3] This study has spurred a large literature on whether weak ties are more effective for information transmission. Aral and Alstyne (2011) show that the importance of weak ties and strong ties could be context dependent. Kramarz and Skans (2014) find that strong social ties are an important determinant of where young workers find their first job.

margin though collectively weak ties are more important because they are numerous. Our results corroborate the findings in the existing literature that the marginal referral effect is more pronounced among strong social ties.

Finally, our work is related to the empirical literature on information economics. Recent studies have shown that increasing information transparency boosts consumers' perceptions of product attributes (e.g., Smith and Johnson 1988), drives up average product quality (e.g., Jin and Leslie 2003; Bai 2018), improves consumer choices (e.g., Hastings and Weinstein 2008; Barahona et al. 2021), and enables households to better respond to environmental disamenities (e.g., Barwick et al. 2020; Cutter and Neidell 2009). Our analysis contributes to this strand of literature by quantifying the importance of information exchange through referrals in facilitating urban labor market mobility.

The paper proceeds as follows. Section 2 discusses our data, the institutional background, and descriptive evidence. Section 3 presents baseline regressions on the referral effect and event studies on the information flow among referral pairs, examines effect heterogeneity, and compares our referral measure with proxies for social interactions that are commonly used in the literature. Section 4 reports an extensive set of robustness analyses and rules out alternative explanations. Section 5 analyzes the benefits of referrals to both workers and firms. Section 6 provides evidence on external validity. Section 7 concludes.

2 Data and Descriptive Evidence

2.1 Data

We compiled a large number of datasets for our analysis. Besides data on geocoded phone records, we assembled administrative data on firm attributes and auxiliary data on neighborhood attributes, residential housing prices, and vacancies (job postings). The online appendix provides more details on data construction and the auxiliary data sets used in the analysis.

Geographical Units At the highest level, the city that we study is divided into 23 administrative districts and counties.[4] These districts and counties are further broken into 1,406 neighborhoods delineated by major roads. A neighborhood is similar to but smaller in size than a census block in the United States. There are 917 neighborhoods in the urban center of

[4]The city consists of an urban core (divided into 8 districts) and 15 surrounding suburban and rural counties. These 8 districts and 15 counties are all equal parts of the city and are under its administrative authority.

the city and 489 neighborhoods in the surrounding suburbs.[5] The lowest-level geographical unit is a *location*, a geographic position returned by a cellular tower station, which represents a building complex or an establishment within a neighborhood. In total, there are close to 18,000 locations, with on average 13 distinct locations per neighborhood.[6]

We overlay two GIS shape files (maps) to obtain the spatial attributes for each location and neighborhood. The first shape file delineates administrative divisions, roads, highways, railways, and parks, as well as points of interests, such as hospitals, schools, shopping malls, parking lots, and restaurants. The second shape file depicts neighborhood boundaries.

Call Data China's cellphone penetration rate is high. According to the China Family Panel Studies (CFPS), a nationally representative longitudinal survey of individuals' social and economic status since 2010, 85 percent of respondents sixteen years and older report possessing a cellphone.

We obtain anonymized and geocoded call data from a major Chinese telecommunications company (hereafter Company A). The data set contains the universe of phone records for all of Company A's mobile phone subscribers (excluding commercial entities) in a northern city, covering the period of November 2016 to October 2017. Company A serves between 30 and 65 percent of all mobile phone users in the city that we study.[7]

Call records include individual identifiers (IDs), location at the time of usage, and the time and duration of usage.[8] The data to which we have access are aggregated to the weekly level and contain the encrypted IDs of the calling party and receiving party, call frequency and duration in seconds, and details on whether a user is Company A's customer. We also observe demographic information about each customer, such as age, gender, and place of birth. The birth county enables us to distinguish migrants from local residents. The existing literature has shown that migrants are much more likely to refer and work with other migrants from their birth city and province (Dai et al., 2018).

Cellphone usage records are automatically collected when individuals send text messages,

[5]These neighborhoods are constructed by our data provider for billing purposes. The average size of an administrative district/county is 712.00 km^2. The average sizes of a neighborhood in the urban core and a neighborhood in the suburb are 0.45 km^2 and 25.03 km^2, respectively. See Online Appendix Figure S1 for a section of the city map.

[6]The distribution of locations by neighborhood is skewed, with some neighborhoods having many more locations (and switchers) than the average. We control for neighborhood fixed effects in our analysis below.

[7]China has three major telecommunications companies. To keep the data vendor anonymous, we report its market share in a range.

[8]Individuals are identified based on their anonymized IDs, which is made possible by the "real-name authentication system" implemented in 2011. Since January 2017, mobile phones that do not meet the real-name authentication requirements cannot operate in China. For individuals with multiple phones, we observe usage on the most commonly used phone. If they subscribe to services from multiple carriers (which is uncommon), we observe their activities only within Company A.

make calls, use apps, or browse the Internet. An important advantage of our data is that the geocoded location is also recorded every 15 minutes whenever the mobile device is turned on, even if it is idle. The serving cellular tower station closest to the device records a geographic position in longitude and latitude that is accurate up to a 100–200 meter radius—roughly the size of a large building complex. For each individual and week, we observe the location with the most frequent phone usage (calls, texts, apps, etc.) between 9 am and 6 pm during the weekdays (which we call a work location), and the location with the most frequent usage between 10 pm and 7 am in the same week (which we call a residential location).[9] In contrast to traditional datasets in social science studies that lack fine-grained geographical information about human interactions, these geocoded location data trace out individuals' spatial trajectories over time and allow us to construct a diverse set of social ties (including friends, neighbors, past and present coworkers, and friends' coworkers).

Constructing individuals' workplace history using recorded geocodes is the most crucial step of our analysis. Since we do not directly observe the employment status or place of work, we take a conservative approach to mitigate measurement errors in our work-related variables. The raw data correspond to 1.8 million individuals. We focus on working-age individuals above 16 with valid work locations for at least 45 weeks—a period long enough to precisely identify workplaces. Locations that are visited during working hours on a daily basis for weeks in a row are likely to be a workplace rather than shopping centers or recreational facilities. Doing so gives us 560,000 individuals.[10] After we further restrict our sample to individuals who have at most two working locations throughout the sample period (which excludes salespersons and individuals who engage in out-of-town business travel or make family visits) and for whom we have complete demographic information, our final sample is reduced to 456,000 users. We carry out the core empirical analysis using this sample and conduct robustness checks in Section 4 using less stringent sample selection criteria.

Auxiliary Data We merge the call data with administrative firm-level records and data on job vacancies and housing prices. These data sources allow us to construct a large number of attributes for each location and neighborhood, including industry composition, average wage, number of employees and vacancies, most common occupations among job postings, and housing prices. For individuals in our final sample, we observe their work and residential location, friends, and neighbors, as well as the workplaces and home locations of both friends

[9] Phone usage during 7 am –9 am and 6 pm –10 pm is excluded because people are likely on the move during these time intervals.

[10] Several factors contribute to sample attrition. China's cellphone market is dynamic, with a high fraction of customers switching carriers during each month, especially among people on prepaid plans. In addition, the work location information is missing for periods when individuals travel outside the city or experience frequent location changes (common for unemployed or part-time workers, salespersons, etc.).

and neighbors. The online appendix provides more details on these additional data sets.

Job Switchers We define individual i as a *job switcher* if the following criteria are satisfied. First, individual i must have worked in two locations, must be observed for at least four weeks in either location, and must have switched locations only once. Second, the distance between these two locations must be at least one kilometer (km). We choose the cutoff of 1 km to avoid erroneously identifying someone as a switcher because individuals' work locations are geocoded up to a radius of 100–200 meters and the average distance between neighborhood centroids is 1.4 km. Among the 456,000 users in our final sample, 8 percent (38,102) are identified as job switchers. Though constructed with different data sources, this on-the-job switching rate is similar to that reported in the literature on China's labor market, which is around 7 percent (Nie and Sousa-Poza, August 2017). Job-to-job mobility is lower in China than in Western countries (for example, it is 15–18 percent in the European Union as documented in Recchi 2009), partly because of the hukou system that imposes significant restrictions on individual migration across provinces or from rural to urban areas (Whalley and Zhang, 2007; Ngai, Pissarides and Wang, 2017). Our switchers found jobs in a total of 5,800 unique work locations spread across 1,100 neighborhoods. Two-thirds of these locations are in the urban core, with the remainder in surrounding rural areas.

2.2 China's Labor Market and Referrals

China's labor market has several noticeable features. In particular, its hukou system, established in the 1950s, categorizes individuals as agricultural or nonagricultural on the basis of their birth place, partly to anchor peasants to the countryside. According to Zhang and Wu (2018), China's labor market has a two-tier system: urban cities and rural areas. The large divide that separates these two tiers in terms of job opportunities, social benefits, and amenities (education, health care, etc.) has created a large number of migrant workers in urban cities who take jobs at low wages and with long working hours and who are often denied social benefits.

As in the United States and European countries, referrals are common among Chinese workers. Figure 1 compares the popularity of different job search methods among Chinese and U.S. workers using data from the 2014 CFPS (red dotted bars) and the 2014 U.S. Current Population Survey (blue solid bars), respectively. Workers in China are more likely to rely on informal search methods (38 percent of workers in China find jobs through friends, in comparison to 30 percent in the United States), while formal search methods such as searching vacancy ads, registering with job agencies, or directly contacting employers are more prevalent among workers in the United States. In addition, referrals are more important for young workers in China, with a higher fraction of young respondents citing referrals as their main channel for finding a job.[11]

The top panel of Table 1 presents descriptive statistics of all individuals in our sample. About one-third (36 percent) of users are women and 89 percent of users are younger than 60, reflecting the higher mobile-phone penetration among men and the younger population. Three-quarters of our sample users were born in the local province; the rest migrated from other provinces. Thirty-nine percent of users were born in the city that we study. The last column presents the national average of the 2014 CFPS survey among individuals who use a cellphone.[12] Our sample exhibits demographics similar to the national average, except that it contains a smaller fraction of individuals under age 25 and fewer women, partly because we focus on individuals with stable jobs.

Social Ties and Referrals The bulk of our analysis focuses on job switchers and their social ties. As the bottom panel of Table 1 illustrates, job switchers and nonswitchers have similar demographics, except for age. Job switchers are more likely to be in their thirties and on average two years younger than nonswitchers. They are more likely to be migrants and have a higher fraction of friends who use Company A's mobile service, although the magnitude of these differences is modest.

Switcher i's social contacts include everyone who makes a phone call to or receives a phone call from individual i at least once during our sample period.[13] The call data provide rich information on users' social network but contain information on work locations only for Company A's customers. On average, 50 percent of a user's friends are customers of Company A. One might be concerned about potential sample selection bias if Company A's customer network overrepresents certain demographic groups. This is unlikely to be a major issue. First, Company A's user network is geographically spread out and covers all street-blocks of the city. Second, pricing and plan offerings are similar across mobile service providers. Nonetheless, to examine the robustness of our results, Section 4 separates individuals into two groups based on whether friend coverage is above the median and documents similar findings for both groups.

Figure 2 reports the average weekly number of social contacts with whom switchers com-

[11] According to the 2014 CFPS, 64 percent of Chinese job seekers under age 20 and 44 percent between ages 20 and 24 have received referrals. These figures are considerably higher than their counterparts among U.S. job seekers of similar ages, at 19 percent and 24.7 percent, respectively.

[12] The CFPS sample is restricted to adults with phone-related expenses that exceed RMB 30 per month to ensure proper phone usage.

[13] These are "one-way" contacts. An alternative definition requires a contact to both make a phone call to *and* receive a phone call from individual i at least once during the sample period. The use of these two definitions leads to very similar results (Section 4).

municated over our sample period. It varies between 23 and 25 for most weeks. Importantly, no spikes in the number of social contacts exist during the weeks leading to the job switch. Instead, the number of friends decreases modestly prior to the job switch and becomes slightly higher after the switch.[14] These patterns suggest that social links established prior to the job change are likely exogenous; otherwise, we should expect a spike in the number of contacts during weeks approaching the job change. Nonetheless, to mitigate concerns over endogenous links formed surrounding the job change, we use social contacts established *three months* prior to the job switch throughout the empirical analysis.[15] Section 4 documents the robustness of the results to consideration of alternative cutoffs (excluding social contacts formed within 1, 2, 3, 4, or 5 months of the job switch).

When a switcher moves to a preexisting social contact's workplace, we define this contact as the switcher's referrer. Among the 38,102 job switchers, 4,703 (12 percent) have missing information for friend locations (Panel A of Online Appendix Table S1). Among the switchers with nonmissing locations for at least one friend, 25 percent find a job through a referral. Note that this should be interpreted as a lower bound of the frequency of referred job changes, as referrers with fewer than 45 weeks of nonmissing work locations are not counted. In Panel B of Online Appendix Table S1, we include all social contacts with at least four weeks of nonmissing work locations. About two-fifths (43 percent) of switchers move to work with a referrer friend. In light of this difference, Section 3 presents estimates with our preferred friend definition, while Section 4 repeats these analyses using all friends with at least four weeks of work information.

2.3 Motivating Evidence

Before we delve into a formal regression analysis, we present descriptive evidence on the spatial patterns of job changes and phone communications.

Job Changes Most job changes occur between districts with comparable levels of socioeconomic development (Column 1 of Table S2). Urban-to-urban job changes are most frequent (49.6 percent), followed by rural-to-rural job changes (35 percent) and switches between urban and rural districts (15.4 percent). Wealthy districts (districts with above-median housing prices) account for a disproportionately high fraction of job changes, with close to three-quarters (73.5 percent) of worker flows occurring between wealthy district

[14] Note that the phone numbers of commercial entities—such as those related to moving services—are excluded from our data (Section 2.1).

[15] There are no official statistics on the duration of job search for on-the-job transitions in China. However, according to the 2014 CFPS, the median time needed to transition from an old job to a new one for people with employment gaps is 3 months. As a result, we use the 3-months threshold in the main analysis.

pairs. This reflects differences in the overall economic activities between more developed and less developed areas. About a fifth (18.9 percent) of job changes occur between districts with below-median housing prices, and only 7.6 percent are between wealthy and nonwealthy districts. There is an equally strong correlation in terms of amenities between the old and new job districts, where amenities are measured by the total number of restaurants, schools, major roads, and parking lots. Three-quarters (75.4 percent) of the job changes take place between high- (above-median) amenity districts, 17.6 percent between low- (below-median) amenity districts, and only 7 percent between high- and low-amenity districts.

These patterns are perhaps to be expected since most job changes in our sample are local: three-quarters of the new jobs are within 9.7 km of the old job and within 8.8 km of the job switcher's home. Such localized job moves are partially driven by China's hukou system, which limits rural-to-urban migration.

Information Flow The spatial patterns of phone communication resemble those of job changes, with the majority of phone calls also occurring between districts with comparable levels of socioeconomic development (Column 2 of Table S2). Urban district pairs account for a much higher fraction of the total number of phone calls (63.7 percent) than other district pairs. Rural district pairs contribute 13.1 percent of the total call volume, with the remaining 23.2 percent occurring between urban and rural districts.[16] Wealthy district pairs and nonwealthy district pairs account for 80.9 percent and 13.1 percent of the communication flow, respectively, with the remaining 6 percent of phone calls taking place between wealthy and nonwealthy district pairs. The patterns are similar in terms of district amenities, with high-amenity district pairs and low-amenity district pairs accounting for 82.1 percent and 12.3 percent of the total call volume, respectively.

Information Flow and Worker Flows There is a strong correlation between the information flow and worker flows. The raw correlation between these two series at the district level exceeds 0.94. Some correlation arises naturally from heterogeneous spatial and economic attributes, such as the high communication volume and job turnover among economic centers and urban cores. To better examine the relationship between the two series, we regress the number of job changes on the call volume and include origin and destination fixed effects. The patterns are similar whether we use more aggregate geographical areas (such as district pairs) or less aggregated areas (such as neighborhood or location pairs), though the correlations are stronger at the more aggregated level (Online Appendix Table S3).

[16] Using the total call volume in minutes delivers similar patterns.

The top panel of Table 2 reports the correlation patterns when the unit of observation is a neighborhood pair. Column 1 examines standard economic indicators that predict the flow of job changes, such as the geographic distance and the absolute differences in housing prices and amenities between the two areas. As expected, areas that are closer in distance and more similar in economic development and amenities experience more job flows. Nonetheless, explaining worker flows among fine spatial areas like neighborhood pairs is a demanding exercise, as shown by the R squared of 0.03. Column 2 evaluates how these economic indicators correlates with information flow across areas. The coefficients are intuitive given the patterns documented above: areas that are closer and more similar in economic development levels share higher call volumes. Column 3 adds the communication flow as an additional regressor for worker flows. The R squared jumps from 0.03 in Column 1 to 0.17 in Column 3, a fourfold increase, suggesting that information flow has a much stronger correlation with worker flows than the standard socioeconomic measures. The effect size is both statistically and economically significant. An inverse hyperbolic sine specification (which accommodates zero values) suggests that a doubling of the call volume is associated with a 16 percent increase in worker flows.

Existing studies have shown that mobile phone usage can predict economic activities (Kreindler and Miyauchi, 2019). The lower panel of Table 2 shows the results of a simple prediction exercise where we use the regression coefficients estimated from the first half of the sample to predict worker flows for the second half of the sample. Including the call volume as a regressor significantly improves the prediction accuracy. For example, the root mean squared error (RMSE) declines from 0.253 job changes to 0.181, and the mean absolute percentage error (MAPE) shrinks from 1.162 percent to 1.037 percent.

Having illustrated the high correlation between information exchange and job flows, we now turn to the main component of our empirical analysis, which focuses on a specific channel whereby information is at work: information on job openings shared among social contacts. The existing literature has documented that 30 to 60 percent of all jobs are typically found through informal contacts rather than formal search methods, a pattern that holds across countries and over time, regardless of the occupation or industry (Topa, 2001; Burks et al., 2015). We next use call data to depict individuals' social network and quantify the magnitude of referral effects and the benefits of referrals for workers and firms.

3 Communication and Job Changes

Our formal analysis begins with a regression analysis that quantifies the referral effect. Then, we conduct event studies to illustrate the time series patterns of the information flow between job seekers and their referrers surrounding job changes. The stark contrast with the patterns for job seekers and nonreferrer pairs presents concrete evidence that referrers provide job-related information to job switchers. We then explore effect heterogeneity and document that referrals are more effective in settings in which job information is more valuable. The last subsection examines social interaction proxies that are commonly used in the literature.

3.1 Regression Analysis

Our regression sample includes all job switchers as defined in Section 2.1.[17] To quantify the referral effect that shapes job seekers' location choices, we compare the propensity for switcher i to find a job at a friend's workplace with that of finding a job at a nearby location:

$$M_{il} = \beta \text{Friend}_{il} + \boldsymbol{X_i Z_l \gamma} + \varepsilon_{il}$$

where M_{il} is 1 if i moves to location l. We restrict individual i's choices to locations *within* the neighborhood that contains his/her new workplace. This is done purposefully. Job location choices are influenced by many factors, including industry composition and labor demand, commuting distance and local amenities, and intrahousehold bargaining, many of which are unobserved in our setting. Limiting individuals' choices to locations within the neighborhood of new workplace greatly reduces the extent of heterogeneity across choices and allows us to better isolate the effect of referrals from competing explanations for job changes. One implication of this regression design is that demographic variables (or individual fixed effects) do not help explain location choices because they are invariant across locations.

The key regressor is $Friend_{il}$, a dummy variable for having at least one friend working in location l. We include a rich set of interactions between demographic attributes and location amenities. The demographic controls ($\boldsymbol{X_i}$) consist of a constant, gender, migration status, and age group categories (ages 25–34, 35–44, 45–59, and 60 and above). We also include i's total number of social contacts (irrespective of cellphone carrier) to capture differences in personality and social outreach. Amenities at each work location ($\boldsymbol{Z_l}$) are captured by the number of restaurants, number of roads and parking lots (which measures a location's accessibility), and number of schools within a 500-meter radius. To allow for differential preferences toward local amenities, we interact gender with schools and parking lots, age group dummies with restaurants, the migrant dummy with the number of roads, and the number of i's social contacts with all location attributes. The results are similar if we control for the full interactions between all demographic characteristics and location attributes. To

[17] Section 6 reports the results for individuals who experienced unemployment gaps and then subsequently found a new job during our sample period.

facilitate interpretation of the coefficient magnitudes, Online Appendix Table S4 tabulates summary statistics for key variables referenced in various regression samples.

The referral effect is measured by β. There are three main threats to a causal interpretation. The first is driven by unobserved spatial confounders, where a positive correlation can arise with exogenous worker flows. We address this problem by adding the *interaction* of the origin and destination neighborhood fixed effects:

$$M_{il} = \beta \text{Friend}_{il} + \boldsymbol{X_i Z_l \gamma} + \lambda_{\tilde{c},c} + \varepsilon_{il} \tag{1}$$

where $\lambda_{\tilde{c},c}$ is a dummy for the pair of neighborhoods ($\tilde{c}$,c) that contains individual i's previous and current workplace. There are a total of 20,811 neighborhood-pair fixed effects. This is a demanding specification wherein the key coefficient β is estimated from the within-origin–destination variation. In other words, β is identified from location choices among individuals who move between the same origin–destination neighborhood pair but have different friend networks.

The second long-standing challenge in the literature using observational data is the difficulty in distinguishing a referral effect from the effects of homophily and sorting. If individuals share similar preferences and skills with their friends, then a positive β could be driven by sorting rather than referrals. In addition, not all locations have desirable openings. An individual might move to location l not because of a referral but because other locations lack appropriate job opportunities. In other words, the friend dummy might simply proxy for locations specializing in jobs that require similar skills shared by individuals and their friends.

Leveraging the richness and structure of our data, we propose the following battery of tests. First, we limit our analysis to workers for whom there is at least one other location within the same neighborhood that has vacancy listings in the same occupation and offering the same salary range as the job taken.[18] This mitigates the concern that individuals sort into friends' locations that provide the only employment opportunity in the area.

Second, we distinguish between friends who are currently working in location l and friends who used to work there but moved away prior to the job switch. Given that sorting on unobserved preferences or skills should happen regardless of a friend's *current* location, we would expect to find similar β estimates for both types of friends if our finding is driven by sorting.[19] Third, we distinguish between friends who work versus friends who live at location

[18]The occupation of location l is the most common occupation among all postings. It is coded as missing if the most common occupation accounts for less than a third of all postings at the same location. The results are robust if we replace the salary range with the expected job compensation as measured by the average payroll (see Section 5.1).

[19]This exercise assumes that homophily is time invariant—an assumption that is likely to hold for our

l. Larger estimates for friends who work in location l would be consistent with the referral effect: affiliation with the workplace enables friends who work there to have an information advantage on jobs openings. Fourth, we borrow the "second degree of influence" concept from the network literature (Evtushenko and Kleinberg, 2021) and examine the importance of friends of friends. These second-degree connections are similar to switchers but do not directly communicate with switchers by construction. We expect their coefficient to be smaller than the referral coefficient if they do not carry job information.

The third threat is related to differences in local labor market demand and location attributes (such as employment size) that may be correlated with the referrer dummy. We address this threat by controlling for an extensive set of location-level characteristics. They include the number of vacancies in each occupation and their interactions with i's demographics; the number of employees and firm attributes in each location, such as firm age, revenue, real capital stock, and average payroll per employee, as well as the interaction of these variables with switcher i's demographics; the total number of calls made or received by individuals (excluding switcher i) working in location l to proxy for location-level economic activities, and the total number of calls by individuals (excluding switcher i) living in that location; and turnover at the new work location prior to the job change.[20]

Results Table 3 reports the coefficient estimates for model (1). All columns include fixed effects for old-by-new neighborhood pairs and interactions of demographic and location attributes.[21] The standard errors are clustered at the neighborhood-pair level, though the statistical significance of key parameter estimates is robust to the choice of clusters. The mean propensity to choose a location within a neighborhood is 0.09. The referral coefficient in Column 1 is 0.34. This effect is economically large and precisely estimated: the probability of an individual moving to location l increases nearly threefold when he/she has a friend working there.

Column 2 limits the sample to the subset of switchers who have at least one alternative work location within the same neighborhood that has openings in the same occupation and with the same salary range as the job that they take. This exercise speaks to the concern that having a friend in a workplace proxies for job openings that require similar skills shared by individual i and his/her friends. If this were true, the presence of a friend in a workplace should not matter as much for individuals facing multiple similar job opportunities. Instead,

sample period of twelve months.

[20]Turnover is measured by the ratio of the average monthly number of people leaving location l during the three months before the job switch over the average monthly number of people leaving location l for the rest of the sample period. We also include a dummy for locations without turnovers.

[21]Online Appendix Table S5 presents supplementary results for Column 1 that begin with no controls and incrementally add more regressors and fixed effects.

in Column 2, we continue to see a positive and large effect of having a friend in a work location. Note that this robustness analysis is made possible by the vacancy information that we collected.

As it is important to control for the availability of job openings, all regressions hereafter use this subset of switchers who face multiple similar job opportunities. This only moderately reduces the number of observations from 1,120,797 to 915,251. In subsequent discussions, we refer to the estimate of the referral effect in Column 2 (0.35) as the baseline estimate.

Column 3 contrasts the effect of friends currently in the new work location with that of friends who recently moved away, while Column 4 compares the effects of friends working versus friends living in the new workplace. In both cases, friends currently working in the new location have a much larger impact on the choice probability: they are five times as influential as friends who recently moved away and twice as effective as friends who live but do not work in the same location. The differences in these parameter estimates are statistically significant at the one percent level.[22]

Column 5 compares referrer friends with second-degree links—friends of friends—who also work in location l. These individuals reflect homophily because as friends of friends, they are similar to individual i. On the other hand, they are not friends with switcher i and hence are unlikely to communicate job information to switcher i. The coefficient for friends of friends is only 0.14, much smaller than the coefficient for referrer friends (0.35), which remains stable across columns.[23] In Column 6, the referral estimate is 0.33, similar to the baseline estimate in Column 2, despite the inclusion of an extensive set of location-level controls that capture differences in employment size and labor market demand.

3.2 Event Study

To provide direct evidence on how referrals work, we now turn to the detailed call records and examine communication patterns between job switchers i and their referrer friends over time. Specifically, we examine the call frequency dynamics during the event window from

[22] The smaller coefficient for 'Friends who moved away' is not driven by undesirable and time-varying work environment that is associated with friends' departure and reduces switchers' probability of moving to that location. It remains robust when we control for turnovers before job switch and an extensive set of location attributes.

[23] As shown in Online Appendix Table S6, the coefficient for friends of friends is partially driven by job clusters among different occupations in addition to homophily, especially finance and professional services. Nonetheless, the referral effect is robust to controlling for occupation clusters.

11 months before to 9 months after the job switch with a rich set of fixed effects:[24]

$$\text{Freq}_{ijt} = \sum_{s=-11}^{9} \gamma_s \text{Referral}_{ij} \cdot 1[t = s] + \sum_{s=-11, s\neq -1}^{9} b_s \text{Nonreferral}_{ij} \cdot 1[t = s] + \lambda_i + \tau_t + \epsilon_{ijt}$$

where Freq_{ijt} is the number of calls between switcher i and his/her friend j in month t. Referral_{ij} takes value one if switcher i moves to friend j's workplace during the sample period and zero otherwise. Nonreferral_{ij} takes value one for all other friends. Note that friend types do not vary over time by construction. The key coefficients $\{\gamma_s, b_s\}$ vary by event month s ($s = 0$ for the month of the job change). The call frequency between a switcher and nonreferrer friends before the job change at time $s = -1$ is the reference category. All regressions include individual fixed effects λ_i, which control for personality traits such as outgoingness or introversion, and month fixed effects τ_t, which control for holiday and seasonality. With the individual fixed effects, the event study coefficients γ_s and b_s do not reflect the level of call frequency between friend pairs. Instead, they capture dynamic changes in the call frequency *relative to* an individual's baseline frequency of talking to nonreferrer friends prior to the job switch.

To facilitate identification of the event study coefficients and to increase the underlying sample size for the tail months (Schmidheiny and Siegloch, 2020), we bin the period more than ninth months before the job switch with the ninth month and the period more than eighth months post the job switch with the eighth month. The standard errors are clustered at the individual level, though the results are robust if we cluster by the work neighborhood instead. Figure 3 presents the results. The confidence intervals for nonreferrer friends are much tighter than those for referrer friends. This is because switcher–nonreferrer pairs are more common: there are 4.76 million switcher–nonreferrer–month observations but just 238 thousand switcher—referrer–month observations.

The communication patterns between referrer pairs and nonreferrer pairs are distinct, even after we include a rich set of fixed effects. First, switchers have more frequent calls with referrer friends than with other friends. This pattern corroborates findings from the literature (Gee et al., 2017) indicating that referrer friends are closer social contacts. Second, the intensity of information flow between referrer pairs exhibits an inverted-U shape that peaks just before the job change.[25] In contrast, the information flow between switchers and other friends remains stable throughout the sample period, with no noticeable change in the

[24]Since we consider a three-month window prior to the job change to define switchers' social network, the event window after the job change can be a maximum of nine months. We use a monthly instead of a weekly event window to average out noise in the time trends.

[25]The inverted-U shape in calls is not driven by changes in the number of social contacts, which remains stable as shown in Figure 2.

months prior to the job switch. Lastly, the communication intensity between referrers and referees remains elevated post–job switch. Information flow appears to increase with the dimensions of social interaction, as referrers and referees are friends before the job switch and become friends and colleagues afterward.

One might be concerned that individuals sometimes share news about a job offer with friends, which would also lead to intensified communication before they move to the new workplace. First, note that the increased communication starts well before the job change. Second, if this were driven by switchers informing friends about their employment change, we should expect to observe a spike in the communication volume with *both* referrer *and* nonreferrer friends. The fact that we do not see such an increase in communication with nonreferrer friends indicates that this concern is likely unfounded. Finally, some phone calls between the referrer pairs could be inquiries about workplace amenities (instead of job openings per se). We regard all such calls as communication with referrers that facilitates a job change.

Falsification Event Study To provide direct evidence for the falsification tests reported in Table 3, we repeat the event study and examine the patterns of communication with different types of friends. We consider two separate event study regressions to ease the readability of the figure, though results are similar when we pool all friend types in one regression.

$$\begin{aligned}\text{Freq}_{ijt} = \sum_{s=-11}^{9} \gamma_s \text{Referral}_{ij} \cdot 1[t=s] + \sum_{s=-11}^{9} \alpha_s \text{MovedAway}_{ij} \cdot 1[t=s] + \\ \sum_{s=-11, s\neq -1}^{9} b_s \text{OtherFriends}_{ij} \cdot 1[t=s] + \lambda_i + \tau_t + \epsilon_{ijt}\end{aligned}$$

and

$$\begin{aligned}\text{Freq}_{ijt} = \sum_{s=-11}^{9} \gamma_s \text{Referral}_{ij} \cdot 1[t=s] + \sum_{s=-11}^{9} \alpha_s \text{LiveAtNewPlace}_{ij} \cdot 1[t=s] + \\ \sum_{s=-11, s\neq -1}^{9} b_s \text{OtherFriends}_{ij} \cdot 1[t=s] + \lambda_i + \tau_t + \epsilon_{ijt}\end{aligned}$$

where MovedAway_{ij} flags friends who moved away before switcher i joined the new work location and $\text{LiveAtNewPlace}_{ij}$ takes value one if friend j lives in the neighborhood that contains switcher i's new work location but does not work at the new job location. These two event studies are shown in Figures 4a and 4b. Echoing findings in Figure 3, the communication patterns between the referrer pairs that peak immediately before the job switch are in sharp contrast to the patterns of communication between other types of friends. Specifically, the frequency of calls between other types of friends is relatively flat during the months surrounding the job switch, indicating that information exchanged between these friend pairs is unlikely to be specific to the job change. The pattern of communication with friends who moved away is noisy due to the limited number of observations, and there is no evidence that the communication peaks close to the job switch. These pictures provide direct evidence that the increased intensity of communication with switchers is specific to friends who work in the new workplace. While friends living in the new workplace's neighborhood might have relevant information on local amenities and friends who used to work in the new workplace might share similar skills and preferences, there is no systematic evidence that they provide job-specific information, in contrast to referrer friends.

The referral effect is economically large, precisely estimated, and stable across all columns in Table 3. In addition, the communications between referrer pairs exhibit remarkably different dynamic patterns from those between other types of friends. These results cannot be reconciled with sorting and indicate that referrers carry useful information that facilitates matching between workers and job openings.

3.3 Effect Heterogeneity

Referrers could facilitate the match between job seekers and vacancies in different ways. For example, current employees might share job opportunities with their social contacts (providing information to workers). Alternatively, employees might inform their employers of their friends' work attitude and labor market prospects (providing information to firms). Although we cannot disentangle these different mechanisms, we test their common implication that referrals mitigate information frictions in the hiring process. We thus examine whether referrals are more important when information asymmetry is more severe.

Individuals who live far from the new work location, who have limited work experience, or who change industrial sectors are likely to be disadvantaged in obtaining information about new job openings. Similarly, employers are less likely to be knowledgeable about such workers. In Table 4, we interact Friend$_{il}$ with the distance between the old and new workplaces, the distance between home and the new work location, a dummy for young workers (between 25 and 34), a dummy for moving from rural to urban locations, and a dummy for changing sectors. Referrals facilitate job transitions in *all* these situations, especially for rural workers migrating to urban areas and for people changing industrial sectors. For these two groups of individuals, the point estimates of the referral effect are

0.66 and 0.53, respectively—a significant boost above the base estimate of 0.35.

The evidence in Table 4 helps rule out several alternative explanations. One is that our results are simply driven by preferences—that individuals enjoy the company of friends and hence prefer to work in the same place with them. However, this theory cannot explain the stronger referral effect when information asymmetry is more severe (or the communication patterns documented in Section 3.2). Another explanation is that our estimates are driven by nepotism: friends and family being hired instead of the best available candidates (Hoffman, 2017). This is probably not a first-order concern since the presence of nepotism would not imply a stronger effect when information asymmetry is more severe. Moreover, as shown in Online Appendix Section S4 and Table S7, referrals exhibit assortative patterns. In particular, referrals are more common among people in the same age range, whereas nepotism often involves individuals from different age groups (e.g., children of relatives) (Wang, 2013; Foley, 2014). We examine alternative explanations in more detail in Section 4.

3.4 Comparison with the Existing Literature

Proxies for Social Ties How does our referral measure compare to those in the existing literature? There are two common approaches to inferring social networks in observational studies. The first, pioneered by Bayer, Ross and Topa (2008), uses residential neighbors as a proxy. Exploiting data from the Boston metropolitan area, the authors treat as friends individuals who live in the same census block and find that friends are more likely to work in the same census block than individuals living in the same census block group (or 10 nearest blocks) but not in the same block. The second approach assumes that social interactions are stronger within an ethnic group and defines friends as coworkers who are members of the same minority group (Bandiera, Barankay and Rasul, 2009; Dustmann et al., 2016). We re-estimate model (1) using these alternative definitions of friendship in Table 5. "Residential neighbor" is a dummy variable that takes value one if workplace l contains at least one individual who shares the same residential location as i. Ethnicity, which is inapplicable in China's context, is replaced with birth county, as the literature documents strong social ties among individuals from the same birth region (Zhao, 2003). "Same birth county" takes value one if individual i has a coworker in location l who was born in the same county. Columns 1 and 2 include only these alternative friend definitions. Column 3 contrasts neighbors with referrer friends who are not neighbors, while Column 4 compares coworkers who share the same birth county with referrer friends who work in the same location but have different birth counties.[26]

[26] "Residential neighbors" include individuals who are both neighbors and referrer friends, and "Same birth county" includes individuals who share the same birth county and are referrer friends. These measures stack

The results in Table 5 confirm the findings in the literature that both neighbors and coworkers from the same birth counties are good proxies for referrers. The coefficients on neighbors and the same birth county are 0.21 and 0.10, respectively, when they are the only measure of an individual's social network. Given the average moving probability of 0.09, having a social tie of either type significantly increases the probability of switching to location l. On the other hand, the effect of our friend measure based on actual communication dominates the effects of both types of social ties by a large margin. The difference in coefficient magnitude is both statistically significant and economically sizable, and in the case of "Same birth county", the effect of our friend measure is four times as large (Column 4).

To examine the extent of social interactions among these referral proxies, we use the disaggregated call records between neighbors, coworkers, and individuals sharing the same birth county. In Column 1 of Online Appendix Table S8, we randomly draw one percent of individuals (including both switchers and nonswitchers) and examine the average monthly calls between pairs of individuals living in the same neighborhood and pairs of individuals living in the same location within a neighborhood. This mimics the empirical setting in Bayer, Ross and Topa (2008), which contrasts individuals living in the same census block with individuals living in the same census block-group but not in the same block. Neighborhood fixed effects are included to control for observed and unobserved characteristics. Conditional on living in the same residential neighborhood, neighbors living in the same location make 4.5 times as many calls as two random individuals residing in the same neighborhood. In Column 2, we compare the frequency of calls between coworkers with that of calls between pairs of individuals working in the same neighborhood. Similarly to what we find for neighbors, we find that coworkers on average make four times as many phone calls as two random individuals working in the same neighborhood. The results in Column 3 are based on the same sample as in Column 2 but the specification examines the importance of sharing the same birth county for individuals working in the same neighborhood. People with the same birth county communicate more frequently than individuals born in different counties, though the difference is much smaller than the differences for the other two types of social ties.

In sum, our results support the findings from analyses of social interactions using neighbors, coworkers, and coethnic individuals as proxies for social ties but suggest that the estimated referral effects based on these proxies are likely to be a lower bound.

Weak versus Strong Ties The literature on weak versus strong ties in general finds that the marginal effect of a strong tie is stronger than that of a weak tie in job search (Gee

the odds against us in terms of detecting a significant difference in these estimated referral effects.

et al., 2017; Bian, 1997). We measure tie strength by call intensity and revisit this question in Online Appendix Table S9. We follow the baseline specification (Column 2 of Table 3) and include interactions between the "Friend" dummy and measures of call intensity. In Column 1, "Call intensity" is the (de-meaned) number of calls between switcher i and his/her referrer friend at location l prior to the job switch. In Column 2, "$Call_{il}/Call_i$" is a (de-meaned) ratio of the number of calls between switcher i and location l over all calls made by i prior to the job switch. The call-frequency ratio takes into consideration differences across individuals (some people are more outgoing than others) and is a better measure of tie strength. The coefficient of the interaction term is positive and significant in both columns, suggesting that the referral effect strengthens with tie strength. For example, a one-standard-deviation increase in $Call_{il}/Call_i$ is associated with a 6 percent increase in the referral effect.[27] These patterns corroborate the evidence from the existing literature, despite the differences in tie strength measures and data contexts.

Social Contact Diversity In addition to tie strength, another important factor that has been highlighted in the sociology and economics literature is social contact diversity (Ottaviano and Peri, 2006; Ashraf and Galor, 2011; Alesina, Harnoss and Rapoport, 2016). A high volume of information exchange that is limited to the same social group might not be as beneficial as information from a more diverse setting that taps into different social entities. Following Eagle, Macy and Claxton (2010), we define two diversity measures using the normalized Shannon entropy: social entropy and income entropy.[28] These entropy measures reflect the complexity of an individual's network in terms of socioeconomic status. To examine the importance of diversity, we regress the probability of changing jobs using referrals on entropy measures, the total number of calls and the fraction of strong ties as well as demographics in Online Appendix Table S10. Social entropy and income entropy, which reflect the socioeconomic diversity of individuals' information sources, have a sizable and significant impact on successful referral. A one-standard-deviation increase in social and income entropy is associated with a 6 percent and 3 percent increase in the probability of using referrals, respectively. Higher entropy measures reflect a more diverse source of

[27]The s.d. of $Call_{il}/Call_i$ is 0.06. Note that while the referral effect increases with tie strength, it is important even among weak social ties, as reflected by the large coefficient on "Friend". Online Appendix Figure S2 presents the event study controlling for tie strength. The communication pattern between referrer pairs remains hump-shaped.

[28]Loosely speaking, social entropy measures whether an individual is equally likely to converse with any contact or concentrates his/her communication with few social contacts. Online Appendix S3 presents precise definitions of these measures. Further details on information theory and entropy measures can be found in Cover and Thomas (2006).

information on job opportunities and increase the occurrence and success of referrals.[29]

4 Robustness Analysis of the Referral Effect

This section conducts additional robustness checks on the referral effect estimated above. We first evaluate the importance of homophily and preferences for working with friends before we examine reverse causality and the robustness of our estimates to different definitions of friends.

Homophily and Preferences for Working with Friends Table 3 provides evidence against homophily by comparing different types of friends (such as friends who recently moved away and friends living but not working in location l). To examine homophily in further detail, Column 1 of Table 6 directly controls for observable attributes that reveal similarity between switchers and referrers. Our test is inspired by the literature on social network formation (Fafchamps and Gubert, 2007) showing that homophily (common preferences, tastes and attitudes) is the main driver of friendship formation. It follows that individuals sharing a larger number of common friends and friends with common features should have higher levels of homophily than individuals sharing a lower overlap. Hence, including individuals' characteristics and their social ties' attributes should be a direct and effective way to control for homophily. These regressors include dummies for whether the referrer in location l has the same gender as switcher i, is in the same age group and from the same birth county, and has similar wealth (proxied by the housing price);[30] the share of mutual friends and the share of common work neighborhoods covered by i's and l's social network;[31] and differences in age, gender, migration status, birth county, and housing prices between switchers' social ties and referrers' social ties. If switcher i and referrer l mingle with similar friends, then they are likely to be similar as well.

In a similar vein, Column 2 uses a popular unsupervised machine learning algorithm, the K-means clustering algorithm, to nonparametrically profile switchers and their social ties (including referrer friends) instead of using parametric functions of observables as in

[29] Higher entropy measures, especially income entropy, are also associated with a higher probability of changing jobs. It is worth noting that these results echo the findings in Eagle, Macy and Claxton (2010), where a strong correlation exists between information diversity and socioeconomic development across UK communities.

[30] The housing price is "similar" if the difference is within one standard deviation on the housing price distribution. In locations with more than one friend, friend attributes are constructed with the average.

[31] The share of mutual friends and the share of common work neighborhoods are calculated with the Jaccard index: $J(A, B) = \frac{A \cap B}{A \cup B}$. For example, the share of mutual friends between i and l is equal to the number of mutual friends divided by the total number of unique friends among i's and l's social ties.

Column 1.[32] The K-means clustering is performed based on individuals' own attributes (gender, age, whether born in the local city or local province, housing price), and their social ties' attributes (number of neighborhoods covered by their friends, share of friends who are female, share of migrants, share in each age bin, and their average house price). We group switchers and their social ties into 100 clusters and control for a dummy variable indicating whether switcher i and the referrer friend at location l are in the same cluster. Most of these homophily controls have intuitive signs. Nonetheless, the referral effect remains similar to the baseline (0.33 in Column 1 and 0.34 in Column 2), indicating that the baseline controls in model (1)—a rich set of demographic and location-attribute interactions and fixed effects for old-by-new neighborhood pairs—are adequate at capturing sorting and that our results are not driven by homophily.

Another potential explanation for the estimated referral effect is preference for working with friends. For example, workers might prefer to colocate with their friends, even if there is no job-related information shared among them.[33] While such a preference is a plausible and potentially relevant explanation, individual preferences are rarely explicitly measured and difficult to examine with observational datasets. To examine the importance of a preference for working with friends, we leverage the spatial variation in switchers' social networks.[34] If people have strong preferences to be with friends, then we should expect that switchers are more likely to move to places with more friends, all else equal. Column 3 of Table 6 follows the baseline specification but also controls for a dummy that takes value one if the new workplace contains more friends than switcher i's old workplace and dummies for whether the number of similar friends at the new workplace is higher than that at the old workplace.

Indeed, people prefer to mingle with friends. However, having more friends in the new than in the old workplace increases the switching probability by 0.002—much smaller than the estimated referral effect, which increases the switching probability by 0.35. These results suggest that a preference for working with friends is not a major threat in our setting.

Reverse Causality Our analysis defines switcher i's social network as the one formed three months prior to his/her job switch. As discussed in Section 2.2, the three-month cutoff is chosen to reflect a reasonable job search duration. It also mitigates concerns over reverse

[32] K-means clustering uses iterative procedures to partition the data into k non-overlapping groups or clusters. The procedure begins with k randomly picked initial group centers. Each individual is assigned to the group with the closest center to minimize the within-group Minkowski distance metric with argument 2 (i.e., the L2 Euclidean distance). The mean of the observations assigned to each of the groups is computed, and the process is repeated until all observations remain in the same group from the previous iteration.

[33] Park (2019) uses a field experiment and shows that people are willing to forgo 6 percent of their wage to work with friends.

[34] Among all location–switcher pairs that contain at least one friend, 69 percent of locations have exactly one friend, 16 percent have two friends, and 15 percent have three or more friends.

causality whereby some social ties in location l are established after switcher i has found a job there. Online Appendix Table S11 uses the baseline specification but with increasingly stringent cutoffs to define switchers' social network—from one month (Column 1) to five months (Column 5) prior to the job change. The referral effect is robust under different cutoffs, consistent with the fact that few links are formed immediately before the job switch.

Alternative Friend Definition We conduct a few additional robustness checks using different definitions of a friend. The baseline analysis limits friends to those who have at least 45 weeks of nonmissing work locations. This mitigates measurement errors in friends' job locations but omits a large fraction of friends for whom we observe fewer than 45 weeks of location information. The first column of Online Appendix Table S12 replicates the baseline analysis (Column 2 of Table 3) with all friends who have at least four weeks of nonmissing work locations. This enlarges the number of individual–friend pairs from 401,437 to 979,595. The estimated referral effect remains robust: having a friend in a location increases an individual's probability of moving there by 36 percentage points.

Social ties are one-way contacts in the baseline analysis. Column 2 of Online Appendix Table S12 defines individual i's friends as social contacts for whom we observe two-way communications: people who both make phone calls to and receive phone calls from individual i. In addition, all friends with at least four weeks of nonmissing work locations are included in the analysis. The estimate of the referral effect (0.38) is slightly larger than but comparable to the one under our base specification (0.35). As work locations are missing for friends outside Company A's customer network, one might be concerned about potential sample selection biases. Columns 3 and 4 split the switcher sample based on whether the friend coverage is above or below the median (the cutoff is 48 percent). The difference in the estimates of the referral effect between these two subsamples is modest and not significant.

5 Referral Benefits for Workers and Firms

5.1 Referral Benefits for Workers

Having established the robustness of the referral estimate, we turn to examining whether referrals improve referees' labor market outcomes, conditional on finding a job. Our framework for analyzing the benefit of referrals is conceptually similar to model (1):

$$Y_{ilr} = \beta \text{Friend}_{ilr} + \boldsymbol{X_i Z_l \gamma} + \lambda_c + \alpha_r + \varepsilon_{ilr} \tag{2}$$

where Y_{ilr} denotes the labor market outcome of worker i who switches to work location l in neighborhood c and lives in residential neighborhood r. We control for the same set of demographic variables such as gender, age group dummies, migration status, and log number of social contacts. Because we do not observe individuals' socioeconomic background and status, such as education and wealth, we include in all regressions the residential neighborhood fixed effect (α_r), which captures luxurious complexes versus low-income neighborhoods, as a proxy.

We construct five different measures of job quality. The first is the expected wage at the new job, measured by the average annual payroll (in thousand RMB) among firms in the same location weighted by their number of employees. Wage dispersion is often driven by across-firm rather than within-firm variations (Card et al., 2018). As individuals housing value is correlated with their labor income, we use *coworkers'* housing price as a second measure for job compensation. Specifically, we construct the difference between the average housing price of coworkers at the new workplace and that of coworkers at the previous job. Large positive differences are more likely to be associated with increases in wages and other pecuniary benefits.

The other three measures of job amenities include whether the move is from a part-time to a full-time job, whether there is a reduction in commuting distance, and whether the move is from a non-state-owned enterprise (SOE) firm to an SOE, because openings at SOEs are sought after for their job security and pension benefits (Zhu, 2013).[35] Although none of these measures of job outcomes is perfect, collectively, they speak to both the financial and nonfinancial aspects of job quality.

Results Since our labor market outcomes are constructed from different data sources, the number of observations across the specifications in Table 7 varies from 15,881 to 29,117 and reflects the varying degrees of missing information. Referral jobs pay higher expected wages than nonreferral jobs. The point estimate of the wage premium is RMB 620, or about 2 percent of the average wage reported in our sample.[36] Turning to differences in the home values of coworkers in the new versus old workplace, referral jobs are associated with a 0.5 percent higher housing price per square meter (the average housing price in the city is RMB

[35]SOEs account for a small fraction of the total number of firms but more than 30 percent of China's GDP and 20 percent of total employment (State Assets Supervision and Administration Commission 2017). Many SOEs appear in the Fortune Global 500 list and are among the largest conglomerates in the world. Private and foreign companies trail behind SOEs in terms of firm size and revenue. Employment opportunities at SOEs are coveted for their job security, generous benefits, and sometimes higher wages than those in nonstate sectors. A workplace is classified as a SOE if the majority of workers at that location are employed by SOE firms.

[36]The annual wage is measured in thousand RMB, and the mean is 31.

$13,000/m^2$, or \$2,000/m^2.)

Having at least one friend at the new workplace increases an individual's probability of moving from a part-time job to a full-time one by 1.4 percentage points, which is a 2 percent increase in the likelihood of working full time.[37] About a third (31 percent) of job changes involve a shorter commute. Referred jobs are associated with a 30 percent increase in the likelihood of working closer to home. Finally, having a referrer friend raises the probability of moving to an SOE firm by 1.2 percentage points, an 11 percent increase from the mean (0.11). Higher wages/compensation are an indication of enhanced worker productivity, and shorter commutes and full-time positions reflect better job amenities.

Our results are robust to alternative specifications. Online Appendix Table S13 repeats the analysis using all social contacts with at least four weeks of nonmissing work locations. Referral jobs are associated with a 1.3 percent increase in the wage premium, a 0.6 percent increase in job benefits (as proxied by coworkers' housing prices), and a 3 percent increase in the likelihood of working full time; these outcomes are similar to the findings in the baseline specification. The effects on the likelihood of having a shorter commute and transitioning to an SOE firm are also similar.[38]

5.2 Referral Benefits for Firms

With a few exceptions, most empirical studies on job referrals abstract from firm outcomes, because comprehensive data on the performance of both employees and employers are hard to obtain.[39] We merge the call data with administrative firm-level data based on locations and examine variation across a large number of firms in different industries.

Although our analysis in this section is descriptive because we lack suitable instruments, we use a variety of strategies to establish the robustness of our findings. We show that the estimates are robust to a rich set of firm and worker controls, which raises our confidence that these estimates are not simply picking up unobserved firm and employee quality. Instead, firms are likely to benefit from employee-provided referrals, consistent with the fact that referral-based hiring programs are common (Burks et al., 2015).

We successfully merge between 5,000 and 10,000 firms, 67 percent of which are manufac-

[37]Hours worked is derived from phone usage during workdays at the workplace and is conservative by the nature of such records. Part time (full time) is defined as 30 hours or less (more than 30 hours). On average, 57 percent of the switchers work full time before the job change, reflecting the conservativeness of this measure.

[38]This evidence is also robust when adding other measures of social contacts (as in Table 3) that may capture homophily and other factors. See Online Appendix Table S14.

[39]A notable exception is Burks et al. (2015), who use data from nine large firms in three industries (call centers, trucking, and high-tech) to analyze whether firms benefit from referrals.

turing firms that require production facilities.[40] Our main specification focuses on locations matched to large firms with more than 100 employees, which constitute about 20 percent of firms in our sample. The average employment for these firms is 150; thus, they are likely to occupy an entire location. While limiting the sample to large firms significantly reduces the sample size, it reduces the likelihood of erroneously linking workers to unrelated firms.[41] Online Appendix Table S15 reports the results from replicating the analysis with all firms. The results are similar both statistically and economically, which is reassuring. In the rest of this section, we use "location" and "firm" interchangeably.

We compare the performance of firms that hire through referrals to firms that hire through other channels via the following model:

$$Y_i = \gamma \text{Referral}_i + \boldsymbol{Z_i}\boldsymbol{\beta} + \lambda_c + \varepsilon_i \tag{3}$$

where i denotes a firm. We examine three measures of firm performance Y_i: (1) net inflow of workers, or the number of hires minus separations; (2) the match rate, measured by net inflow over vacancies; and (3) the firm growth rate, measured by net inflow over total number of employees.[42] We limit our analysis to locations with at least one hire; otherwise, the estimate of γ would be artificially inflated since the number of hires is at least one for locations with referrals by construction.

Referral_i is a dummy variable that takes value one if at least one worker who switches to firm i has a friend working there, while λ_c denotes neighborhood fixed effects—the same as in model (2). $\boldsymbol{Z_i}$ denotes firm attributes and employee characteristics. Firm attributes include age, dummies for 18 different industries, a dummy for SOEs, the average number of employees (firm size) and the average inflation-adjusted capital stock from 2010 to 2015. To capture preexisting trends, we also control for the average employment growth rate from 2010 to 2015. In addition, we include a firm's referral network size, defined as the number of unique social contacts owned by employees who work in firm i prior to the arrival of new hires. Worker attributes include the shares of female workers and migrants, the average employee age, and the average housing price of preexisting employees.

Results The parameter estimate γ captures the effect of using referrals on firms' performance. The dependent variables in Table 8 are in logs, and hence, γ directly reflects semi-elasticities: the percentage change in the outcome variable when firms hire through re-

[40] The exact number of successful merges is withheld to keep the city anonymous.

[41] For the same reason, we repeat the referral analysis while limiting the sample to locations matched with large firms (those with more than 100 employees). The results are very similar to the baseline estimates.

[42] There are two measures of worker inflows: gross inflows and net inflows. The results reported below use net inflows (defined as total inflows minus outflows), though they are similar to those based on gross inflows.

ferrals. To the extent that firms that grow quickly are more likely to hire through employee referrals, our estimate could be biased upward. To address this problem, we estimate model (3) with an increasingly rich set of controls for firm growth and employee quality.

The Referral_i coefficient estimates are remarkably similar across the specifications with different sets of controls for firm and employee attributes. Firms that recruit through referrals are associated with more hires, better matching rates, and higher growth rates. According to the most saturated specification (Column 4), successful referrals increase a firm's net labor inflow by 63 percent, enhance the job matching rate by 84 percent (the average matching rate for large firms is 1.53), and raise the firm growth rate by 45 percent (the median growth rate is 4 percent for large firms).[43] We replicate our analysis with other selection criteria (for example, using all friends with at least four weeks of nonmissing work locations as in Online Appendix Table S16 and friends with at least three months or six months of work locations) and obtain robust findings.

6 External Validity

We conclude our analysis with extensions that shed light on the external validity of our findings. We first examine how call volumes relate with other communication channels such as text messages and mobile apps (WeChat). Then, we repeat the analysis for individuals who experience unemployment spells but subsequently find a new job within our sample period.

Phone Calls and Other Communication Channels A key premise of our analysis is that call volume serves as a proxy for the amount of information exchanged among individuals. In practice, there are alternative communication channels, such as text messages, emails, and apps (such as WeChat). One potential concern with our analysis is that people might use text messages or WeChat in lieu of phone calls to communicate with friends. If there is a negative correlation between phone calls and alternative information channels, then observing a high phone call volume does not imply more information exchange between individual i and his/her social contacts, as the increased call volume could be offset by reduced numbers of text messages and app usage.

We do not observe WeChat or app usage for individuals in our sample and thus consider

[43] The matching rate is defined as inflow over vacancies. It can exceed one as not all openings are posted and the data coverage on posted vacancies is incomplete. Davis, Faberman and Haltiwanger (2013) analyze the *Job Openings and Labor Turnover Survey* and report that the vacancy yield, defined as the flow of hires during the month expressed as a percentage of the vacancy stock on the last day of the previous month, averaged 1.3 from 2001 and 2006 and varied from 0.7 to 3.1 across major industries.

three alternative measures: whether an individual's phone is compatible with the 4G network as a proxy,[44] the Internet data allowance on an individual's cellphone plan, and individuals' internet browsing behavior (duration in thousand minutes) for one week in May 2017. There is a positive and significant correlation between monthly calls and each of these three measures. The pattern holds for both switchers and nonswitchers and is robust across different specifications (Online Appendix Table S17). We then collect information on the individual usage of text messages and WeChat from 20,000 randomly selected cellphone users in a comparable city in China in November 2020. Individuals who talk more (measured by either the number of calls or call duration) also send more text messages and use WeChat more frequently (Online Appendix Table S18). The positive correlations hold under individual fixed effects.

These patterns suggest that different communication channels are complements: individuals who make more phone calls also send more text messages, use WeChat more intensively, and browse the Internet more. As a result, while our call records do not cover other information channels, they serve as a good proxy for the amount of information exchanged between an individual and his/her social ties.[45]

Switchers with Employment Gaps Our data also provide an opportunity to examine the referral effect for individuals with employment gaps. However, this analysis is considerably more challenging since it is difficult to distinguish unemployment from other factors that also lead to intermittent work location patterns (travel, sick leave, part-time jobs, etc.). In addition, the referral analysis for switchers with employment gaps is limited to workers who experience unemployment and manage to find another job within our sample period (and hence experience a relatively short unemployment duration). In the end, we identify a total of 3,638 individuals with one employment gap and valid friend work location information, of whom 1,677 find jobs through referrals (see Online Appendix Section S5 for more details). The modest sample size reflects the challenges in measuring unemployment; as a result, we interpret the results below using reemployed individuals as suggestive. Nonetheless, this analysis could be informative about the generalizability of our main results.

We begin with an event study to examine the patterns of communication between reemployed individuals and their referrers at different stages (pre-unemployment, during unemployment, and post-reemployment) and contrast them with the patterns of communication

[44] WeChat is the dominant social media app in China with 1.2 billion users as of 2020 (Tecent, 2020). The performance of WeChat and other communication apps is significantly better with the 4G network.

[45] We repeat our baseline referral analysis separately for switchers whose phones are compatible or incompatible with the 4G network in Online Appendix Table S19. The referral effect estimates are similar across both groups. Thus, referrals are important regardless of whether individuals can use WeChat.

between these reemployed individuals and their nonreferrer friends:

$$\begin{aligned}\text{Freq}_{ijt} = &\sum_{s=-5}^{3} \gamma_s \text{Referral}_{ij} \cdot 1[t = s, r < 0] + \sum_{r=-3}^{4} \gamma_r \text{Referral}_{ij} \cdot 1[t = r, s \geq 0] \\ &+ \sum_{s=-5, s\neq -1}^{3} b_s \text{nonreferral}_{ij} \cdot 1[t = s, r < 0] + \sum_{r=-3}^{4} b_r \text{nonreferral}_{ij} \cdot 1[t = r, s \geq 0] + \lambda_i + \tau_t + \epsilon_{ijt}\end{aligned}$$

As there are two events (the unemployment and the reemployment), we use s to denote the event window index for unemployment and r to denote the index for reemployment. Similarly to the setup in the analysis using on-the-job switchers, the reference category is the frequency of calls between reemployed individuals and their nonreferrer friends during the month immediately before the unemployment, $s = -1$.[46]

The event study is presented in Figure 5, which exhibits patterns similar to those in Figure 3. Most notably, the communication pattern with referrals has an inverted-U shape during unemployment (the period of job search) and peaks prior to reemployment.[47] Similarly to on-the-job switchers, people with employment gaps experience more intense communication with their referrers before finding a new job. We repeat the event study using different unemployment definitions. The qualitative pattern of more pronounced communication between job seekers and their referrers during the search period (the unemployment spell) is present in all event studies that we conduct (Figure S3).

Using the number of monthly calls as a proxy for job search intensity, we find suggestive evidence that individuals who switch jobs after unemployment spells search more actively for job opportunities than on-the-job switchers (Online Appendix Table S20).[48] Similarly to on-the-job switchers, people who experience employment gaps also benefit from referrals (Online Appendix Table S21). The referral effect for this population varies from 0.31 to 0.33, in line with our findings in Section 3.

[46] As in the event study for on-the-job switchers, we bin the period prior to the fifth month before unemployment with $s = -5$ and the period after the fourth month post-reemployment with $r = 4$. The coefficients to the left of the first vertical line (s=0) are estimated using observations prior to unemployment, and the coefficients to the right of the second vertical line (r=0) are estimated using observations post-reemployment. In between these two vertical lines, the coefficient for s=1 is estimated using the first month of unemployment for people with one or multiple months of an employment gap, and the coefficients for r=-1 is estimated using the last month of unemployment for people with one month or multiple months of an employment gap. Similarly for $s = 2$ and $r = -2$. Finally, the coefficient at $s \geq 3, r \leq -3$ is estimated using observations when the employment gap is at least three months.

[47] The uptick in calling intensity one month prior to unemployment likely reflects China's Labor Contract Law, which requires a 30-day notice before the termination of an employment contract.

[48] We define the search period as the unemployment spell for individuals with employment gaps and the three months prior to the job change for on-the-job switchers.

7 Conclusion

This paper provides the first direct evidence of increased communication around job changes among referral pairs. We use geocoded mobile phone records matched to administrative firm-level data and auxiliary data on housing prices and vacancies to illustrate that information provided by social contacts mitigates information asymmetry and improves labor market performance. Future studies on the mechanisms that govern how information exchange through referrals increases labor market efficiency would be extremely valuable.

References

Akbar, Prottoy A, Victor Couture, Gilles Duranton, and Adam Storeygard. 2018. "Mobility and Congestion in Urban India." National Bureau of Economic Research Working Paper 25218.

Alesina, Alberto, Johann Harnoss, and Hillel Rapoport. 2016. "Birthplace diversity and economic prosperity." *Journal of Economic Growth*, 21(2): 101–138.

Aral, Sinan, and Marshall Van Alstyne. 2011. "The Diversity-Bandwidth Trade-off." *American Journal of Sociology*, 117(1): 90–171.

Ashraf, Quamrul, and Oded Galor. 2011. "Cultural Diversity, Geographical Isolation, and the Origin of the Wealth of Nations." NBER Working Paper 17640.

Aslund, Olof, Lena Hensvik, and Oskar Skans. 2014. "Seeking similarity: How immigrants and natives manage in the labor market." *Journal of Labor Economics*, 32(3): 405–41.

Athey, Susan, David Blei, Robert Donnelly, Francisco Ruiz, and Tobias Schmidt. 2018. "Estimating Heterogeneous Consumer Preferences for Restaurants and Travel Time Using Mobile Location Data." *AEA Papers and Proceedings*, 108: 64–67.

Bai, Jie. 2018. "Melons as Lemons: Asymmetric Information, Consumer Learning and Quality Provision." Working Paper.

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2018. "Social Connectedness: Measurement, Determinants, and Effects." *Journal of Economic Perspectives*, 32(3): 259–80.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2009. "Social Connections and Incentives in the Workplace: Evidence from Personnel Data." *Econometrica*, 77(4): 1047–1094.

Barahona, Nano, Cristobal Otero, Sebastian Otero, and Joshua Kim. 2021. "Equilibrium Effects of Food Labeling Policies." *Stanford Working Paper.*

Barwick, Panle Jia, Shanjun Li, Liguo Lin, and Eric Zou. 2020. "From Fog to Smog: the Value of Pollution Information." *NBER Working Paper No. 26541.*

Bayer, Patrick, L. Stephen Ross, and Giorgio Topa. 2008. "Place of work and place of residence: Informal hiring networks and labor market outcomes." *Journal of Political Economy*, 116(6): 1150–96.

Beaman, Lori. 2012. "Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the U.S." *Review of Economic Studies*, 79(1): 128–61.

Bian, Yanjie. 1997. "Bringing Strong Ties Back in: Indirect Ties, Network Bridges, and Job Searches in China." *American Sociological Review*, 62(3): 366–385.

Bjorkegren, Daniel, and Darrell Grissen. 2018. "Behavior Revealed in Mobile Phone Usage Predicts Credit Repayment." *World Bank Economic Review.* Accepted.

Blumenstock, Joshua E., Gabriel Cadamuro, and Robert On. 2015. "Predicting poverty and wealth from mobile phone metadata." *Science*, 350: 1073–1076.

Blumenstock, Joshua, Guanghua Chi, and Xu Tan. 2019. "Migration and the Value of Social Networks." C.E.P.R. Discussion Papers CEPR Discussion Papers 13611.

Buchel, Konstantin, Maximilian V. Ehrlich, Diego Puga, and Elisabet Viladecans-Marsal. 2019. "Calling from the outside: The role of networks in residential mobility." CEMFI Working Paper wp2019-1909.

Burks, Stephen V., Bo Cowgill, Mitchell Hoffman, and Michael Housman. 2015. "The Value of Hiring Through Employee Referrals." *Quarterly Journal of Economics*, 805–839.

Cai, Hongbin, Hanming Fang, and Lixin Collin Xu. 2011. "Eat, Drink, Firms, Government: An Investigation of Corruption from the Entertainment and Travel Costs of Chinese Firms." *The Journal of Law and Economics*, 54(1): 55–78.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics*, 36(S1): S13–S70.

Cingano, Federico, and Alfonso Rosolia. 2012. "People I know: Job Search and Social Networks." *Journal of Labor Economics*, 30(2): 291–323.

Cover, Thomas M., and Joy A. Thomas. 2006. *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing).* Wiley–Interscience.

Cutter, W. Bowman, and Matthew Neidell. 2009. "Voluntary information programs and environmental regulation: Evidence from 'Spare the Air'." *Journal of Environmental Economics and Management*, 58(3): 253 – 265.

Dai, Ruochen, Dilip Mookherjee, Kaivan Munshi, and Xiaobo Zhang. 2018. "Community networks and the growth of private enterprise in China." *VoxDev.*

Davis, Steven J., R. Jason Faberman, and John C. Haltiwanger. 2013. " The Establishment-Level Behavior of Vacancies and Hiring *." *The Quarterly Journal of Economics*, 128(2): 581–622.

Donaldson, Dave, and Adam Storeygard. 2016. "The View from Above: Applications of Satellite Data in Economics." *Journal of Economic Perspectives*, 30(4): 171–98.

Dustmann, Christian, Albrecht Glitz, Uta Schonberg, and Herbert Brucker. 2016. "Referral-based Job Search Networks." *The Review of Economic Studies*, 83(2): 514–546.

Eagle, Nathan, Michael Macy, and Rob Claxton. 2010. "Network Diversity and Economic Development." *Science*, 328(5981): 1029–1031.

Edin, Per-Anders, Peter Fredriksson, and Olof Aslund. 2003. "Ethnic Enclaves and the Economic Success of Immigrants: Evidence from a Natural Experiment." *The Quarterly Journal of Economics*, 118(1): 329–357.

Evtushenko, Anna, and Jon Kleinberg. 2021. "The paradox of second-order homophily in networks." *Scientific Reports*, 11(1): 13360.

Fafchamps, Marcel, and Flore Gubert. 2007. "Risk Sharing and Network Formation." *American Economic Review*, 97(2): 75–79.

Foley, Meghan. 2014. "JPMorgan Gives Regulators Evidence of Nepotism in China." Retrieved March 13, 2014, from https://www.cheatsheet.com/money-career.

Gee, Laura, Jason Jones, and Moira Burke. 2017. "Social Networks and Labor Markets: How Strong Ties Relate to Job Finding on Facebook's Social Network." *Journal of Labor Economics*, 35(2): 485–518.

Gee, Laura K., Jason J. Jones, Christopher J. Fariss, Moira Burke, and James H. Fowler. 2017. "The paradox of weak ties in 55 countries." *Journal of Economic Behavior & Organization*, 133: 362–372.

Glaeser, Edward L, Scott Duke Kominers, Michael Luca, and Nikhil Naik. 2015. "Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life." National Bureau of Economic Research Working Paper 21778.

Glitz, Albrecht. 2017. "Coworker networks in the labour market." *Labour Economics*, 44: 218–230.

Gonzalez, Marta C, Cesar A Hidalgo, and Albert-Laszlo Barabasi. 2008. "Understanding individual human mobility patterns." *Nature*, 453: 779.

Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology*, 78(6): 1360–1380.

Hastings, Justine S., and Jeffrey M. Weinstein. 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *The Quarterly Journal of Economics*, 123(4): 1373–1414.

Hellerstein, Judith, Melissa McInerney, and David Neumark. 2011. "Neighbors and coworkers: The importance of residential labor market networks." *Journal of Labor Economics*, 29(4): 659–95.

Hellerstein, Judith, Melissa McInerney, and David Neumark. 2014. "Do labor market networks have an important spatial dimension?" *Journal of Urban Economics*, 79(4): 39–58.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review*, 102(2): 994–1028.

Hoffman, Mitchell. 2017. "The value of hiring through employee referrals in developed countries." *IZA World of Labor*.

Ioannides, Yannis M., and Linda Datcher Loury. 2004. "Job information networks, neighborhood effects, and inequality." *Journal of Economic Literature*, 42: 1056–1093.

Jin, Ginger Zhe, and Phillip Leslie. 2003. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." *The Quarterly Journal of Economics*, 118(2): 409–451.

Kramarz, Francis, and Oskar Nordström Skans. 2014. "When Strong Ties are Strong: Networks and Youth Labour Market Entry." *The Review of Economic Studies*, 81(3): 1164–1200.

Kreindler, Gabriel E., and Yuhei Miyauchi. 2019. "Measuring Commuting and Economic Activity inside Cities with Cell Phone Records." Working Paper.

Munshi, Kaivan, and Mark Rosenzweig. 2013. "Networks, Commitment, and Competence: Caste in Indian Local Politics." National Bureau of Economic Research Working Paper 19197.

National Bureau of Statistics of China. 2014. "The Third National Economic Census Report." Retrieved December, 2014, from http://www.stats.gov.cn.

Ngai, L Rachel, Christopher A Pissarides, and Jin Wang. 2017. "China's mobility barriers and employment allocations." *Journal of European Economic Association*. Forthcoming.

Nie, Peng, and Alfonoso Sousa-Poza. August 2017. "What Chinese Worker Value: An Analysis of Job Satisfaction, Job Expectations and Labor Turnover in China." *IZA Discussion Papers No. 10963*, 108.

Ottaviano, Gianmarco I.P., and Giovanni Peri. 2006. "The economic value of cultural diversity: evidence from US cities." *Journal of Economic Geography*, 6(1): 9–44.

Park, Sangyoon. 2019. "Socializing at Work: Evidence from a Field Experiment with Manufacturing Workers." *American Economic Journal: Applied Economics*, 11(3): 424–55.

Recchi, Ettore. 2009. "Chapter 4: The Social Mobility of Mobile Europeans." In . Vol. 1 of *Pioneers of European integration citizenship and mobility in the EU: Citizenship and mobility in the EU*, , ed. Adrian Favell and Ettore Recchi, 72 – 97. Edward Elgar Publishing.

Saygin, Perihan Ozge, Andrea Weber, and Michele A. Weynandt. 2018. "Coworkers, Networks and Job Search Outcomes." *ILR Review*. Forthcoming.

Schmidheiny, Kurt, and Sebastian Siegloch. 2020. "On Event Studies and Distributed-Lags in Two-Way Fixed Effects Models: Identification, Equivalence, and Generalization." *ZEW - Centre for European Economic Research Discussion Paper No. 20-017.*

Schmutte, Ian. 2015. "Job referral networks and the determination of earnings in local labor markets." *Journal of Labor Economics*, 33(1): 1–33.

Schmutte, Ian. 2016. "How do social networks affect labor markets?" *IZA World of Labor*.

Smith, V. Kerry, and F. Reed Johnson. 1988. "How do Risk Perceptions Respond to Information? The Case of Radon." *The Review of Economics and Statistics*, 70(1): 1–8.

Tecent. 2020. "Tecent Announces 2020 Third Quarter Results." Retrieved November, 2020, from https://www.tencent.com/en-us/investors/financial-news.html.

Topa, Giorgio. 2001. "Social Interactions, Local Spillovers and Unemployment." *The Review of Economic Studies*, 68(2): 261–295.

Topa, Giorgio. 2011. "Labor Markets and Referrals." *Handbook of Social Economics*, 1B: 1193–1221.

Wang, Shing Yi. 2013. "Marriage Networks, Nepotism, and Labor Market Outcomes in China." *American Economic Journal: Applied Economics*, 5(3): 91–112.

Whalley, John, and Shunming Zhang. 2007. "A numerical simulation analysis of (Hukou) labour mobility restrictions in China." 83(2): 392–410.

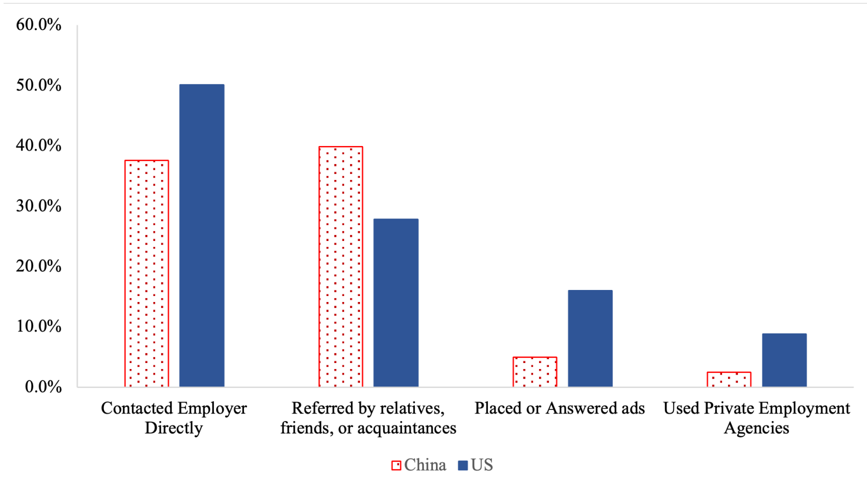
Zhang, Junsen, and Jia Wu. 2018. "The Chinese labor market, 2000–2016." *IZA World of Labor*, , (437).

Zhao, Yaohui. 2003. "The Role of Migrant Networks in Labor Migration: The Case of China." *Contemporary Economic Policy*, 21(4): 500–511.

Zhu, Julie. 2013. "Chinese SOEs and "invisible benefits"." Retrieved May 13, 2013, from https://www.ft.com.

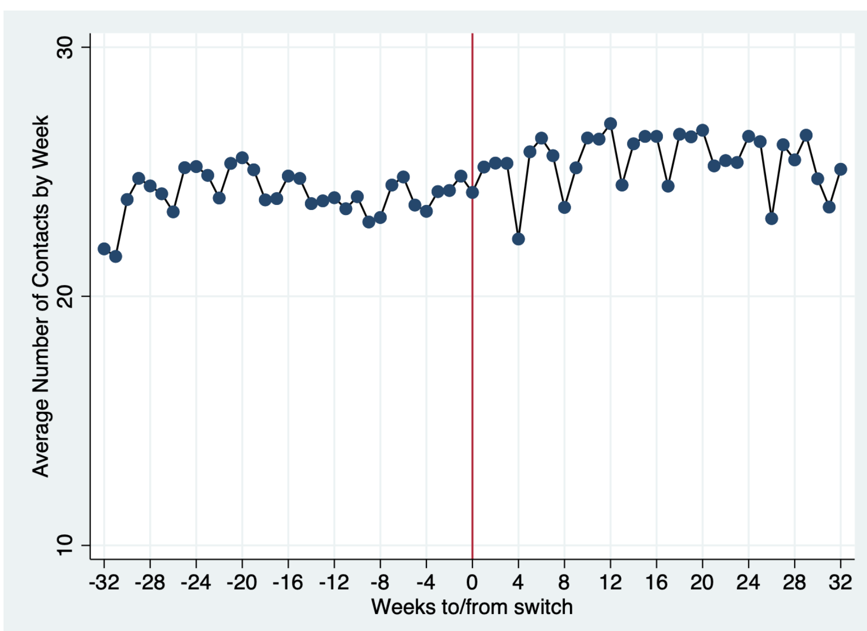
8 Figures and Tables

Figure 1: Job Search Methods in China vs. United States (2014)



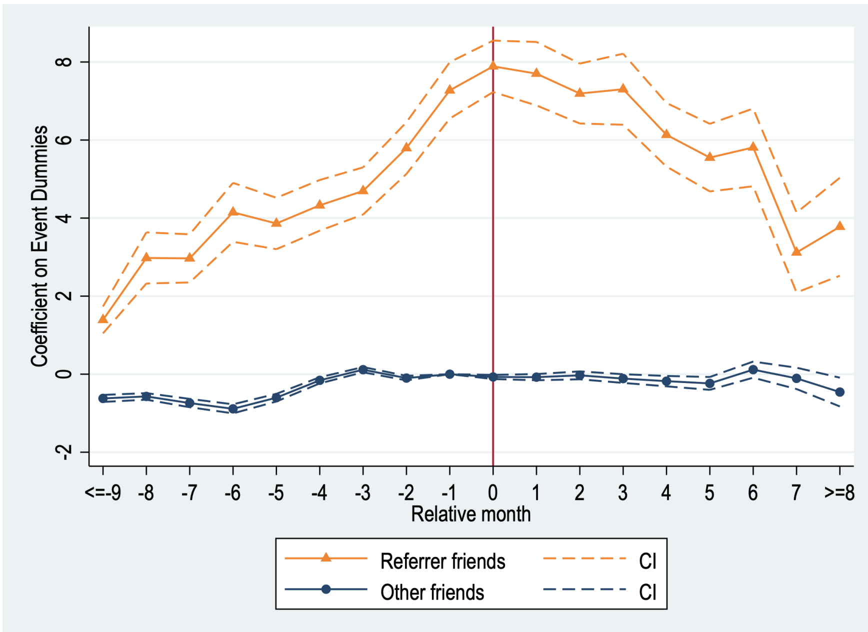
Notes: the data sources are 2014 China Family Panel Studies (CFPS) and 2014 U.S. Current Population Survey (CPS). This figure shows the frequencies of job search methods that are common among the Chinese and US job seekers. Red dotted (blue solid) bars represent China (United States). The search methods are not mutually exclusive (some job seekers used more than one method).

Figure 2: Job Switchers' Social Contacts Before and After Job Switches



Notes: This figure plots the average number of social contacts who communicate with a job switcher in each week before and after the job switch. The vertical line indicates the week of the job change.

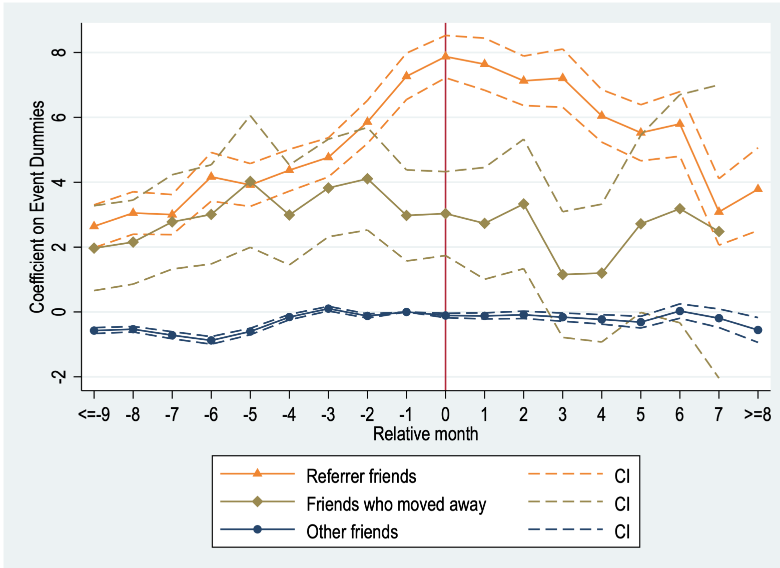
Figure 3: Calls to Referrer vs. Nonreferrer Friends



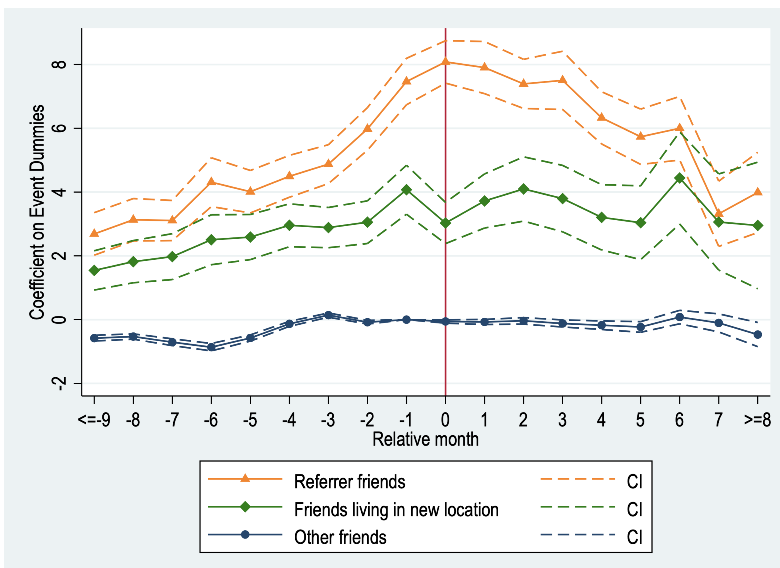
Notes: This figure plots the coefficient estimates and their 95 percent confidence intervals for an event study that examines the number of calls between a job switcher and his/her referrer and nonreferrer friends. The vertical line indicates the month of the job change. The orange line (with triangles) represents calls between switchers and their referrers (238,092 obs). The blue line (with dots) represents calls between switchers and their nonreferrer friends (4,759,176 obs). The vertical line indicates the month of the job switch. The reference group is the frequency of calls between switchers and nonreferrer friends one month prior to the job switch. The period prior to the ninth month before the job switch is binned with the ninth month, and the period after the eighth month after the job switch is binned with the eighth month. Switcher fixed effects and calendar month fixed effects are included in the regression.

Figure 4: Calls to Referrer Friends vs. Other Types of Friends

(a) Calls to Referrer Friends and Friends Who Move Away

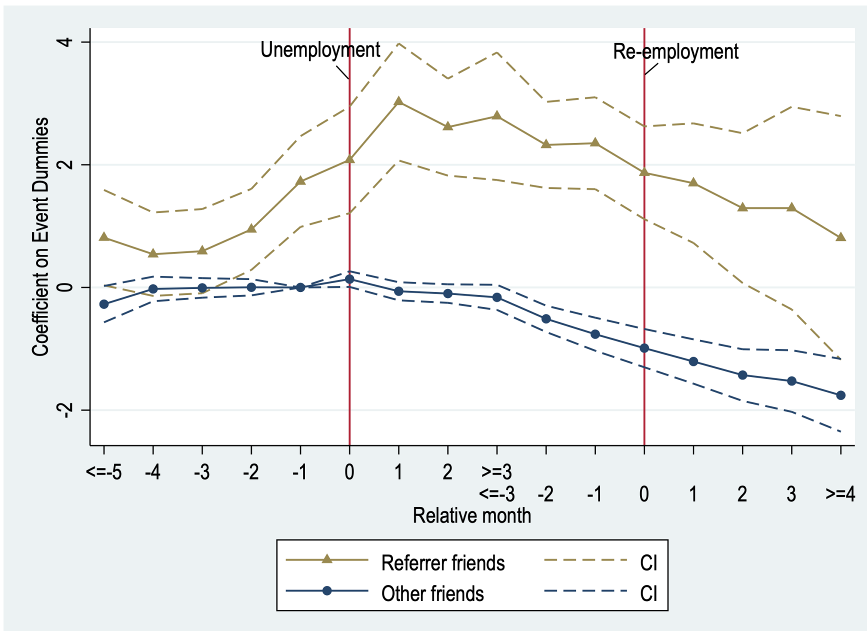


(b) Calls to Referrer Friends and Friends Living in the New Workplace



Notes: Figures 4a and 4b plot the event-study coefficients and their 95 percent confidence intervals for the number of calls between a switcher and four types of friends before and after the job switch, respectively. The vertical line indicates the month of the job change. In both figures, the orange line (with triangles) denotes calls between switchers and their referrer friends (238,092 obs). In Figure 4a, the brown line (with diamonds) denotes calls between switchers and nonreferrer friends who move away before the job switch occurs (29,148 obs). The blue line (with dots) denotes calls between switchers and the rest of their nonreferrer friends (4,730,028 obs). In Figure 4b, the green line (with diamonds) denotes calls between switchers and nonreferrer friends who live in the same neighborhood as the new job location (62,472 obs). The blue line (with dots) denotes calls between switchers and the rest of their nonreferrer friends (4,696,704 obs). In both figures, the reference group is the frequency of calls between switchers and nonreferrer friends one month prior to the job switch. The period prior to the ninth month before the job switch is binned with the ninth month, and the period after the eighth month after the job switch is binned with the eighth month. Calendar month fixed effects and individual fixed effects are included in the regression.

Figure 5: Calls to Referrer and Nonreferrer Friends by Individuals with Unemployment Spells



Notes: This figure plots the coefficient estimates and their 95 percent confidence intervals for an event study that examines the number of calls between individuals with unemployment spells and their referrer and nonreferrer friends. The analysis is limited to 3,638 individuals who become unemployed for at least eight weeks and then reemployed with valid work locations during our sample period. The brown line (with triangles) denotes calls between a reemployed individual and his/her referrer friend (53,244 obs). The blue line (with dots) denotes calls between the reemployed individual and his/her nonreferrer friends (747,804 obs). The first vertical line denotes the month immediately before unemployment. The second vertical line denotes the month of reemployment. The x-axis labels denote months relative to the two events. Calendar month fixed effects and individual fixed effects are included in the regression.

Table 1: Summary Statistics

(a) All users

	Mean	2014 CFPS national survey Mean
Female	0.36	0.45
Age 25–34	0.29	0.23
Age 35–44	0.26	0.24
Age 45–59	0.27	0.27
Age 60 and above	0.11	0.09
Age (midpoint)	40.18	39.28
Born in local province	0.75	0.76
Born in local city	0.39	-
Fraction of social contacts in Firm A	0.50	-
Job switcher	0.08	0.07

(b) Switchers vs. nonswitchers

	Nonswitchers Mean	Switchers Mean	Diff.	t-stat
Female	0.36	0.36	-0.001	-0.45
Age (midpoint)	40.36	38.23	2.13	32.49
Born in local province	0.75	0.74	0.01	3.62
Born in local city	0.39	0.38	0.002	0.70
Fraction of social contacts in Firm A	0.50	0.51	-0.004	-0.53

Notes: The sample is restricted to 455,572 individuals with valid work information for at least 45 weeks during the sample period. Panel (a) reports summary statistics for the sample individuals (the first column) and the national average of the 2014 CFPS survey (the second column). We limit the CFPS sample to individuals with phone-related expenses that exceed RMB 30 per month, weighted by the sampling weights. Panel (b) reports the summary statistics for switchers and nonswitchers separately and t-test statistics for the differences in sample means. "Fraction of social contacts in Firm A" is the fraction of an individual's social contacts who are Company A's customers. "Job switcher" is a dummy for individuals who change jobs, based on the criteria described in the text. The CFPS sample's job switching rate is the "on-job switch rate" calculated by Nie and Sousa-Poza (August 2017). "Age (midpoint)" uses the midpoint of each age range.

Table 2: Information Flow and Worker Flows

	(1)	(2)	(3)
Dependent variable:	Worker flows between l and k	Total calls between l and k (thousand)	Worker flows between l and k
Panel A: Regression analysis			
Distance between l and k	-0.003	-0.006	-0.003
	(0.0003)	(0.001)	(0.0002)
Difference in housing price (thousand)	-0.001	-0.005	-0.001
	(0.0004)	(0.001)	(0.0004)
Difference in amenities	-0.001	-0.002	-0.0004
	(0.0001)	(0.0002)	(0.0001)
Total calls between l and k (thousand)			0.101
			(0.016)
Observations	987,713	987,713	987,713
R-squared	0.033	0.043	0.170
Area l + Area k	Yes	Yes	Yes
Panel B: Prediction exercise			
RMSE	0.253		0.181
MAPE	1.162%		1.037%

Notes: This table examines the relationship between information flow and worker flows. The unit of observation is a neighborhood pair. The dependent variable in Columns 1 and 3, "Worker flows between l and k", is the total number of workers moving between area l and area k. The dependent variable in Column 2 is the information flow between l and k. In Panel A, "Distance between l and k" is the geographical distance in kilometers between the centroids of areas l and k. "Difference in housing price" is the (absolute) difference in the average housing prices in areas l and k. "Difference in amenities" is the (absolute) difference in the total numbers of amenities (restaurants, roads, parking lots, and schools) in areas l and k. Standard errors are reported in parentheses and two-way clustered by area l and area k. Panel B reports the results of a prediction exercise whereby we run regressions using the first half of the sample and predict worker flows for the second half of the sample, following the same specification as in the top panel. Then, we compare the observed and predicted worker flows and report the root mean squared error (RMSE) and mean absolute percentage error (MAPE) of the prediction exercise.

Table 3: Effect of Friends on Job Location Choices

Dependent variable: Probability i switches to location l	(1)	Individuals with similar job opportunities nearby (2)	(3)	(4)	(5)	(6)
Friend	0.34	0.35	0.34	0.34	0.34	0.33
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Friend moved before the switch			0.07			
			(0.03)			
Friend living but not working in i's new workplace				0.16		
				(0.02)		
Friend of i's nonreferrer friends					0.14	
					(0.01)	
Observations	1,120,797	915,251	915,251	915,251	915,251	915,251
R-squared	0.14	0.12	0.12	0.13	0.13	0.14
Controls for local amenities	Yes	Yes	Yes	Yes	Yes	Yes
Controls for local labor market characteristics	No	No	No	No	No	Yes
Old x new work neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of neighborhood-pair FE	20,811	16,468	16,468	16,468	16,468	16,468

Notes: This table examines the effect of referrals on job switchers' location choices. The unit of observation is a switcher-location pair. "Friend" is a dummy variable that equals one if switcher i has at least one social contact working at location l. Columns 2 to 6 limit to job switchers for whom there is at least one other location within the same neighborhood that has vacancy listings in the same occupation and salary range as the one taken. In Column 3, "Friend moved before the switch" is a dummy variable that equals one if switcher i has at least one social contact at location l who used to work there but left before i switched to location l. In Column 4, "Friend living but not working in i's new workplace" is a dummy variable that equals one if switcher i has at least one social contact who lives but does not work at location l. In Column 5, "Friend of i's nonreferrer friends" is a dummy variable that equals one if switcher i has at least one second-degree social contact who also work at location l (those who are friends of switcher i's nonreferrer friends but do not directly communicate with switcher i). Column 6 controls for location-level vacancies, employment, firm attributes, turnovers, and total calls by the working and living population, as described in the text. All columns control for the old-by-new work neighborhood-pair fixed effects and location amenities (the number of restaurants, roads, parking lots, and schools within a 500-meter radius), gender interacted with the number of schools and parking lots, age group dummies interacted with the number of restaurants, migrant dummy interacted with the number of roads, and individual i's number of social contacts interacted with all location attributes. Standard errors are clustered by neighborhood pair and reported in parentheses.

Table 4: Referral Effect and Information Asymmetry

Dependent variable: Probability i switches to location l	(1)	(2)	(3)	(4)	(5)	(6)
Friend	0.35	0.33	0.33	0.33	0.34	0.32
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Friend×Distance(job1-job2)		0.002				
		(0.0004)				
Friend×Distance(home-job2)			0.002			
			(0.0003)			
Friend×Young				0.04		
				(0.01)		
Friend×Rural to urban					0.32	
					(0.05)	
Friend×Changing sector						0.21
						(0.02)
Observations	915,251	915,251	915,251	915,251	915,251	915,251
R-squared	0.12	0.13	0.13	0.12	0.13	0.13
Old x new work neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of neighborhood-pair FE	16,468	16,468	16,468	16,468	16,468	16,468

Notes: This table uses the same specification as that in Column 2 of Table 3 and interacts the "Friend" dummy with measures that reflect the extent of information asymmetry. "Young" refers to switchers between 25 and 34 years old. "Rural to urban" flags switchers who move from the rural to urban part of the city. "Changing sector" is 1 if the switcher changes job sector. Columns 2–6 also control for the baseline level of the interacted variable. Standard errors are clustered by neighborhood pair and reported in parentheses.

Table 5: Comparison with Referral Proxies in the Literature

Dependent variable: Probability i switches to location l	(1)	(2)	(3)	(4)
Friend definition				
Residential neighbor	0.21		0.18	
	(0.01)		(0.01)	
Same birth county		0.10		0.09
		(0.01)		(0.01)
Friend, not neighbor			0.25	
			(0.01)	
Friend, not same birth county				0.36
				(0.02)
Observations	915,251	915,251	915,251	915,251
R-squared	0.15	0.10	0.19	0.15
Controls	Yes	Yes	Yes	Yes
Old x new work neighborhood FE	Yes	Yes	Yes	Yes
Number of work neighborhood-pair FE	16,468	16,468	16,468	16,468
Residential neighborhood FE	Yes	No	Yes	No
Number of residential neighborhood FE	1,067	NA	1,067	NA
Birth county FE	No	Yes	No	Yes
Number of birth county FE	NA	17	NA	17

Notes: This table uses the same specification as that in Column 2 of Table 3 and contrasts the estimates based on our referral measure with those based on proxies from previous literature. "Residential neighbor" is a dummy that equals one if switcher i has at least one residential neighbor who works in the new work location. "Same birth county" is a dummy that equals one if there is at least one individual who works in switcher i's new work location and shares the same birth county as switcher i. Old-by-new work neighborhood-pair fixed effects are included in all columns. Columns 1 and 3 also include residential neighborhood fixed effects. Columns 2 and 4 include birth county fixed effects. Standard errors are reported in parentheses and clustered by old-by-new work neighborhood pair and residential neighborhood in Columns 1 and 3 and by old-by-new work neighborhood pair and birth county in Columns 2 and 4.

Table 6: Referral Effect: Robustness Checks

Dependent variable: Probability i switches to location l	(1)	(2)	(3)
Friend	0.33	0.34	0.35
	(0.03)	(0.01)	(0.01)
Dummy for Number of friends at $l \geq$ Number of friends at old place			0.002
			(0.001)
Homophily controls			
Whether friend l has same gender, age group, birth county and housing price as i	Yes	No	No
Share of mutual friends and neighborhoods between i and friend l	Yes	No	No
Difference between demographics of i's friends and demographics of l's friends	Yes	No	No
Dummy for "i and friend l in the same cluster" (n. clusters = 100)	No	Yes	No
Controls for preference for working with friends			
Whether more similar friends at l than at old workplace	No	No	Yes
Observations	915,251	915,251	915,251
R squared	0.12	0.12	0.12
Old x new work neighborhood FE	Yes	Yes	Yes
Number of neighborhood-pair FE	16,468	16,468	16,468

Notes: This table examines the robustness of the estimates to inclusion of homophily and preference for working with friends. It replicates Column 2 of Table 3 with additional controls. Column 1 controls for whether the friend at location l has the same gender, age group, birth county, and similar housing price as switcher i; the share of mutual friends and share of overlapping work neighborhoods between i's social ties and friend l's social ties (defined by the Jaccard index); and the demographic differences between i's and the referrer friend's social contacts. For locations with more than one friend, friend attributes are constructed with the average. In Column 2, we group switchers and their friends into 100 clusters using the k-means clustering algorithm based on each individual's own characteristics and their social ties' characteristics. Column 3 includes a dummy that takes value one if the number of friends at location l is equal to or greater than that at the old workplace and dummies indicating whether the number of friends with the same gender, age group, birth county, and similar housing price at location l is equal to or greater than that at the old workplace. Standard errors are clustered by neighborhood pair and reported in parentheses.

Table 7: Referral Benefits for Workers

	Income Effect		Job Quality		
	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Wage at new job	ΔCoworker HP	PT to FT	Shorter commute	Non-SOE to SOE
Friend	0.62	0.07	0.014	0.09	0.012
	(0.31)	(0.04)	(0.007)	(0.01)	(0.005)
Observations	17,615	23,323	19,431	29,117	15,881
R-squared	0.79	0.53	0.11	0.12	0.56
Residential neighborhood FE	Yes	Yes	Yes	Yes	Yes
New work neighborhood FE	Yes	Yes	Yes	Yes	Yes

Notes: This table examines whether jobs secured through referrals provide higher (financial and nonfinancial) benefits to workers. The sample size varies due to missing observations. "Wage at new job" is the annual payroll per worker in thousand RMB, weighted by employee sizes among firms in the new work location. "ΔCoworker HP" is the difference between the average house price (thousand RMB per m^2) of coworkers in the new workplace and that of coworkers in the old workplace. "PT to FT" is a dummy that equals one if the switcher works part time (30 hours or less per week) before the job change and full time (more than 30 hours) afterward. "Shorter commute" equals one if the commuting distance at the new workplace is shorter than that at the previous workplace. "Non-SOE to SOE" is a dummy that equals one if the new workplace is in an SOE-dominant location (with the majority of employees working in SOE firms) while the previous location is not. All regressions include gender, age group dummies, migration status, and log number of social contacts as controls. Standard errors are two-way clustered by residential neighborhood and new work neighborhood and reported in parentheses.

Table 8: Referral Benefits for Firms

Dependent variable: Panel A: Log of net inflow	(1)	(2)	(3)	(4)
Referral	0.60	0.62	0.62	0.63
	(0.13)	(0.14)	(0.14)	(0.14)
Observations	[600,1000]	[600,1000]	[600,1000]	[600,1000]
R-squared	0.64	0.65	0.65	0.66
Panel B: Log of matching rate	(5)	(6)	(7)	(8)
Referral	0.91	0.86	0.86	0.84
	(0.24)	(0.26)	(0.26)	(0.27)
Observations	[600,1000]	[600,1000]	[600,1000]	[600,1000]
R-squared	0.85	0.86	0.86	0.87
Panel C: Log of firm growth rate	(9)	(10)	(11)	(12)
Referral	0.49	0.45	0.44	0.45
	(0.11)	(0.10)	(0.10)	(0.11)
Observations	[600,1000]	[600,1000]	[600,1000]	[600,1000]
R-squared	0.76	0.83	0.83	0.83
Controls				
Firm attributes	No	Yes	Yes	Yes
Previous growth rate	No	No	Yes	Yes
Employee attributes	No	No	No	Yes
Neighborhood FE	Yes	Yes	Yes	Yes

Notes: This table examines whether firms benefit from successful referrals. A unit of observation is a location that has at least one firm with more than 100 employees and positive hiring. We report a range of the observation number to keep Company A anonymous. The number of neighborhood fixed effects is 225. "Referral" takes value 1 if at least one switcher joining the firm has a friend working in the firm. "Net inflow" is the number of hires minus separations. "Matching rate" is defined as the net inflow over the number of vacancies. "Firm growth rate" is measured as the net inflow over the employee size. Firm attributes include age, industry, SOE dummy, average employee size, real capital, and employee growth rate from 2010 to 2015. Employee attributes include the shares of female and migrant employees and the average age of preexisting employees. Firm network size, measured by the number of distinct contacts of the firm's preexisting employees, and the number of Company A's users at each location are controlled for in all columns. Standard errors are clustered by neighborhood and reported in parentheses.