# Face-to-Face Meetings in Urban Labor Markets

Eleonora Patacchini* Qi Wu†

**Abstract**

We examine the impact of face-to-face interactions with pre-existing social connections in the workplace – commonly known as referrals – on worker mobility. Using detailed cellphone geolocation data and sociodemographic information for 3.3 million workers in a large urban labor market over two years, we find that in-person meetings can double the positive effects of referrals on job movement. These interactions yield higher returns than other communication methods, such as phone calls and messaging apps. The benefits are particularly pronounced for women, migrants, full-time employees, high-income earners, private sector workers, and individuals transitioning from the IT, finance, and manufacturing industries. Our analysis reveals that a higher frequency of meetings drives the premium observed among women, high-income job switchers, and private-sector employees. In contrast, the composition of these meetings is crucial to explaining the premium for migrants, full-time workers, and employees in specific industries.

**Keywords** Agglomeration Economies, Employment, Communication Technology
**JEL Classification: R23, J60, L15**

*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.

# 1 Introduction

Job mobility is crucial for a dynamic labor market and economic growth. A well-documented finding in the literature is that pre-existing social ties play a significant role in recruitment and hiring decisions within firms (Bayer, Ross and Topa, 2008; Burks et al., 2015; Friebel et al., 2023). Recent studies highlight that communication with referrals – friends already employed at a workplace – plays a crucial role in driving worker mobility (Barwick et al., 2023). However, little attention has been given to how the mode of communication affects these mobility patterns.

This paper explores whether face-to-face interactions among friends shape the magnitude of referral effects in the labor market. It further investigates which types of workers and jobs are most influenced by referrals and whether the frequency or composition of these interactions explains variations in referral effects.

Addressing these questions is critical not only for understanding the mechanisms driving labor market dynamics but also for unpacking broader phenomena such as the determinants of agglomeration, the role of cities as engines of economic growth, and why globalization has not diminished the importance of location and place. It may also provide insights into the impact of the COVID-19 pandemic and help evaluate the relative importance of in-person versus virtual meetings.

Despite its significance, measuring the returns to face-to-face interactions in the labor market (and other markets) has been challenging due to the lack of data, as such meetings typically go unrecorded. This paper advances the field by leveraging highly granular cell-phone geolocation data to construct a digital map of interactions and worker movements across workplaces over more than two years in a major urban labor market: Beijing, China. We enhance this dataset with information on alternative communication methods, including phone calls and mobile chat apps. Additionally, we incorporate highly detailed thematic maps to identify first, job types (private or public sector) and industry classifications.

We leverage the richness of this extraordinary dataset to address long-standing challenges in the referral literature that examines observational data. These challenges include the difficulty of identifying actual and meaningful interactions, distinguishing referral effects from homophily and sorting, and ruling out the possibility that social contacts are formed after job change decisions are made.

First, we use consistent patterns of in-person interactions we observe months before job changes to identify pre-existing social ties. Because of the continuous-time observation of interactions, we can identify friends more precisely using interactions specifically during weekends and holidays. Second, while most studies in the referral literature use the mere

presence of a friend in the workplace as an indicator of referrals – making it challenging to distinguish referrals from unobserved characteristics shared by movers and their friends – we focus on the importance of the meeting technology. By conditioning on the presence of a friend, we mitigate concerns related to homophily. Moreover, the frequency of phone calls, which we observe and include in our model, likely captures the effects of similarities in preferences between friends (homophily). Finally, we control for geospatial attributes that are correlated with job flows (the presence of commercial centers, industrial clusters, etc.) and location choices by the interaction of the origin and destination block fixed effects. In our empirical design, we compare individuals whose old and new jobs are located in the same origin-destination block pair and with friends already working in different workplaces within the destination neighborhood but with different intensities of face-to-face interactions with these friends.

Our results show that face-to-face interactions with friends increase an individual's propensity to move to a workplace by about 1.25 times relative to the probability of choosing other job locations nearby. Such an effect doubles the positive effect of referrals.

We conduct a comprehensive set of robustness analyses to explore alternative explanations for the referral effect. These include homophily – where the intensity of face-to-face interactions reflects similarities in observed or unobserved characteristics – location attributes, such as working population and number of friends, and switchers' characteristics, such as their overall level of face-to-face interactions. The returns to face-to-face meetings persist under these specifications and surpass the effects of phone calls and other communication channels in magnitude.

While identifying the precise mechanism through which face-to-face interactions influence labor mobility is highly challenging with observational data, we make progress by leveraging the richness of our dataset. Specifically, we compare the effects of face-to-face interactions with friends currently working at the new job location to those with friends who have recently moved away or will relocate there in the future. This approach provides insight into whether face-to-face interactions facilitate labor mobility through an information-sharing channel. Our results indicate that face-to-face meetings with friends currently working at a location have a significant impact on the probability of switching jobs, an effect that is both larger and statistically distinct from meetings with the other two types of friends. Since friends currently working at a location are more likely to be informed about job opportunities or provide better access to them than those who are not, this evidence suggests that direct, in-person connections facilitate access to valuable labor market information, ultimately influencing individuals' career decisions.

In terms of effect heterogeneity, the benefits of in-person interactions are more pronounced

for women, migrants, full-time employees, and high-income earners. In terms of job types, the benefits accrue to workers in the private sector and from IT, finance, and manufacturing industries. Meeting quantity and quality explains the differential effects. The higher frequency of meetings explains the premium observed for women, high-income job switchers, and private-sector employees. On the other hand, the specific composition of the meetings – who these workers are meeting with – is the key factor in explaining the premium for migrants, full-time workers, and employees in specific industries.

Our paper closely relates to two key areas of research: labor economics, specifically the literature on referrals (Topa, 2011; Schmutte, 2015), and economic geography, particularly the work on agglomeration economies (Duranton and Puga, 2004, 2020). Both strands of literature share a common limitation: they primarily infer interactions based on observed characteristics, such as former colleagues, family ties, shared immigrant or ethnic backgrounds, or residential proximity. Additionally, little is known about actual interactions or the role of interaction technologies.

Two notable exceptions closely related to our paper also leverage cellphone data, as we do. Barwick et al. (2023) explores the relationship between call patterns and job referrals. It provides direct evidence that referrers pass job information to job seekers by identifying friends as social contacts for which actual interaction via mobile phone call is observed and by comparing call patterns between referrals and non-referrals friends. Differently from this study we use repeated in-person interaction well before a job change is observed to define friends and focus on the importance of interaction technology in shaping the referral effects. Atkin, Chen and Popov (2022) demonstrates that serendipitous face-to-face meetings between workers at different establishments in Silicon Valley increase inter-firm citations. Like this study study, we measure face-to-face interactions using cellphone co-location data derived from location pings. However, unlike their data, which comes from SafeGraph and relies on pings from smartphones with open (or background) apps, our data provider records the location of all cellphones every 15 minutes without requiring apps to be open or calls to be made. This approach significantly improves the precision of our digital map. Additionally, the extended time horizon of our data allows us to identify pre-existing friendships. At the same time, the greater depth of our dataset enables us to examine heterogeneity in effects based on individual and job characteristics and the quantity and characteristics of the meetings.

Our work contributes to the emerging literature on how new data sources – such as online social networks and phone calls – enhance our understanding of economic activity by revealing how people interact and share information (Bailey et al., 2018; Chetty et al., 2022*a*,*b*; Buchel et al., 2019). Our paper demonstrates that large-scale, longitudinal data at

a granular geographical level can help explain agglomeration forces in cities (Glaeser et al., 2015). Specifically, while face-to-face interactions are often cited as key drivers of positive externalities in influential theories of urban agglomeration (Jacobs, 1960), their empirical significance has never been conclusively proven[1]. By providing the first evidence of their role in worker mobility, we take a crucial step in advancing this line of research.

The paper proceeds as follows. Section 2 discusses our data. Section 3 introduces key definitions and descriptive evidence. Section 4 presents baseline regressions on how face-to-face interactions affect the referral effect, conducts a set of robustness analyses and placebo tests to rule out alternative explanations, and compares face-to-face interactions with other communication methods. Section 5 examines effect heterogeneity across worker and job types. Section 6 concludes.

# 2 Data

## 2.1 Cellphone Location Data and Worker Flow

Our primary data source is cellphone location data from one major Chinese telecommunications company. The data set covers timestamped geolocation tracking of the universe of over 3.3 million users in Beijing, China from January 2018 to December 2019. For each user, we also observe their demographic characteristics, including their gender, age, birth county and a wealth index ranging from 1 to 8 as a proxy for income level where 1 indicates the lowest and 8 indicates the highest income group.[2] We consider individuals in groups 1 to 5 as low-income earners, while those in groups 6 to 8 as high-income earners.

We identify workers and job switchers as follows. First, we divide the city into uniform size cells, each measuring a 250-by-250-meters area, in total 262,832 cells. Second, we pool a map of road-delineated city blocks. In total there are 10,807 blocks. Third, we overlap the first grid a thematic map of the city, which shows polygons of a diverse set of locations, including firms, schools, green areas, hospitals, apartment buildings, shopping centers etc. and focus on cells entirely covered by establishments and thus firms or offices of public offices

[1] Storper and Venables (2004) outline key theoretical features of face-to-face interaction that enhance productivity and knowledge spillovers. First, multidimensional communication technology improves the effectiveness and quality of transmitted information. Second, co-presence fosters emotional closeness, which can reduce incentive and coordination problems. Third, investing time in face-to-face interactions signals commitment, helping to build trust. Finally, co-presence serves as a motivator by stimulating desire, imitation, and competition.

[2] The wealth index is constructed based on multiple sources of information, including phone price, subscription plan fee, residential location house price, mode of travel transportation (airplain, highspeed railway), and whether visit certain high-expenditure places (e.g. golf clubs) etc.

of civil servants. Our final sample contains 19,190 cells in 1,860 blocks.[3] We define a user is working in cell $i$ if cell $i$ is the most frequent location between 9am and 5pm on weekdays for at least 2 consecutive months. In our sample, all most frequent locations are identified by pings in at least 20 different in cell $i$ in a given month. In our analysis, we refer to this most frequent location as a workplace. We further classify workers as full-time and part-time workers based on the days they show up constantly in a given workplace. If a worker shows up in a workplace for more than 4 weeks or at least 2 consecutive months with at least 15 days each month, he or she is a full-time worker; while if he or she shows up for at least 2 consecutive months but with less than 15 days each month, this worker is considered as a part-time worker. We identify job changes based on the work location history of individuals. To be classified as a job switcher, an individual must have worked in at least 2 distinct work cells in the past 12 months, and the distance between each old and new workplace must be at least 1 kilometer (km). [4] Applying these definitions, over a 12-month period, we identify 272,937 job changes from a total working population of 3,295,737. This implies a rate of about 8 percent job switchers, in line with a job switching rate of about 7 to 8 percent reported in studies using different data sources (Nie and Sousa-Poza, August 2017; Barwick et al., 2023).

Table 1 shows the number of job changes and job switching rates across the 12 months from January to December 2019. Column 1 shows the total number of job changes that occurred in each month and Column 2 reports the corresponding percentage among all job changes for the entire period. Most job changes occur in September and March, in line with the regular hiring season and job change season after the Chinese New Year. Column 3 shows the rate at which workers are switching jobs each month, calculated as the number of job changes divided by the total work population in each month. On average, the monthly job change rate is about 7.3 percent across the 12-month period.

## 2.2 Cellphone Location Data and Face-to-face Meetings

We construct face-to-face meetings between workers as follows. First, we define a *visit* to a cell. A worker visits a cell $c$ in half-hour-day-month-year $h$ if he or she spends at least 15 minutes in $c$ during time-period $h$. A *meeting* is defined as two workers who work in different cells both being in the same cell $c$ at the same time $h$. Then the total number of meetings between workers in cells $i$ and $j$ is multiplying the number of workers of cell $i$ who visit cell

[3] Almost all these cells are located in the six inner city districts. The inner city districts are the economic and population center of the city, accounting for around 75-80 percent of Beijing's total GDP and 60-65 percent of the work population.

[4] Given that work locations are geocoded with a precision radius of 100-200 meters, the 1 km cutoff ensures individuals are not classified as switchers if they change office within a firm.

$c$ at time $h$, by the number of workers of cell $j$ who visit the same at the same time $(c, h)$, and sum over all $(c, h)$ combinations:

$$Meeting_{ij} = \sum_{hc} e_{ich} e_{jch}$$

where $e_{ich}$ is the number of workers in cell $i$ who visit cell $c$ at time $h$. To mitigate the measurement error of chance meetings, we exclude meeting cells that correspond to large population density, such as shopping centers, hotels, hospitals, train stations, etc. We also exclude meetings associated with delivery workers and meetings between co-workers, that is between workers who work in the same cell.[5]Among all meetings, 2.8% happen in the switcher's cell, 2.9% in friends' cell, and 94.3% in other places.[6]

### 2.3 Face-to-Face meetings and Other Communication Methods Data

We merge the person meetings data with information on other communication methods. We gather the monthly WeChat app data volume for all users and their monthly total data usage from July 2018 to December 2019. We also obtain the universe of phone calls for all users from January to June 2019.[7]

## 3 Definition of Variables and Preliminary Evidence

In this section, we discuss how we use the richness of our data to tackle the measurement challenge of measuring face-to-face interactions with pre-existing referral friends. We then compare our sample of switchers with administrative data to gain confidence into our data. Finally, we describe meeting patterns of job switchers with referral and non-rederral friends, highlightng features that support our identification strategy.

### 3.1 Friend Definition

A key challenge in measuring face-to-face interactions through individuals' co-appearance in the same location is distinguishing meaningful social interactions from random encoun-

[5]Delivery workers are identified by the data provider based on APP usage pattern, daily location trajectories and call pattern.

[6]This procedure follows closely Atkin, Chen and Popov (2022) who construct face-to-face meetings among workers at establishments in Silicon Valley to measure their effect on patent citations among firms these workers belong to.

[7]The earliest month we can track phone call is January 2019.

ters. To address this issue, we leverage the time dimension and define "friends" based on pre-existing relationships prior to a job change, as well as frequent and consistent meeting patterns over time. Specifically, we adopt the following timeline:

1. Friendship Identification: Friends are identified based on regular meetings that occur during weekends and holidays between January 2018 and June 2018.

2. Pre-Job-Change Interactions: Face-to-face interactions with job switchers are observed between July 2018 and December 2018, prior to the job change.

3. Job Changes: Job changes take place between January 2019 and December 2019.

This framework allows us to ensure that the relationships we classify as "friends" reflect meaningful and sustained social connections, rather than incidental encounters.

After establishing the time period for friend identification, we examine the distribution of face-to-face interactions among all workers on weekends and holidays between January 2018 and June 2018. To minimize the risk of misclassifying other relationships as friendships, we exclude certain cases from the face-to-face meeting counts: (1) Meetings during 12 am - 8 am; (2) Meetings between individuals who reside or work in the same cell for more than 10 days in a given month; (3) Meetings taking place at an individual's residential location for more than 20 days in a given month; and (4) Extremely frequent visits (top 1 percent of meetings) on weekends, combined with an age difference of more than 22 years, which likely represent family visits. We define individuals $m$ and $k$ are friends if they meet face-to-face on more than 3 days and at least 16 times over the six-month period. These thresholds correspond to the 90th percentile of the respective distributions. Using this definition, we identify 2.76 million workers with at least one friend, representing approximately 85 percent of the working population.

**Validation of friend definition** Next, we validate our friend definition by examining the number of phone calls between friends and non-friends, as defined above. If phone calls are positively correlated with face-to-face meetings, we would expect the number of calls between friends to exceed that between non-friends. To test this hypothesis, we randomly sample 12,000 user pairs: 6,000 classified as friends under our proposed definition and 6,000 classified as non-friends.

Table A1 presents the validation results. The dependent variable in Column 1 is the average number of calls per month between each user pair. Columns 2 and 3 further break this down into the average number of calls on weekends and on weekdays per month, respectively. The results confirm that friends make significantly more calls than non-friends, and this difference is statistically significant. Interestingly, friends make slightly more calls on weekends

than on weekdays, whereas the reverse pattern holds for non-friends. These findings provide evidence that our proposed friend definition captures meaningful social interactions rather than incidental encounters.

## 3.2 Descriptive Evidence

For confidentiality reasons, we are not allowed to disclose information at the individual level. As a result, we aggregate the number of job switchers and meetings at the cell pair level. However, as we will show in Table 2 (and Figure A2), in the large majority of the cases there is just one switcher (in more than 95 percent of the cell pairs) and one friend (in more than 91 percent of the cell pairs). As a result, the difference is immaterial.

**Job Switchers** Figure 1 reports the summary statistics for urban employees in Beijing based on the China Family Panel Studies (CFPS) national representative survey in 2020, the overall work population and job switchers at the cell level in our sample. Among all workers in urban Beijing, on average, 45 percent of them are female, 69 percent are migrants, 83 percent are full-time employed and 57 percent belong to the low-income group[8].

Turning to the overall work population in our sample, the worker characteristics exhibit some similar patterns as the national representative sample. Across all workers in a given cell, on average 34 percent are female, 65 percent are migrants and 70 percent are in low-income groups. One exception is that for all workers, 71 percent of them are full-time workers. The cell-level means for the job switchers are fairly similar. Among the switchers in a given cell, 33 percent are female and 67 percent are migrants. Additionally, 57 percent of the switchers belong to the low-income groups (income groups 1 to 5), while 25 percent of them are full-time workers. This difference between stable workers and all workers is mechanical, due to the way we define our job switchers – the job switching event may be more common among part-time or temporary workers, leading to their overrepresentation in the switcher sample. In addition, our job switcher sample has a slightly lower share of female compared to the broader urban workforce in Beijing.

[8]To define full-time employees in CFPS, we first identify those ever work in part-time jobs if either of the following three conditions is satisfied: (a) they have not work full-time, (b) they have worked as part-time, and (c) they have other jobs beside their main job. Then those who have not worked in part-time jobs are potential full-time employees. Next, we further exclude the following two groups of individuals from the potential full-time employees: (a) work in agricultural sector, (b) weekly working hours is less than 30 hours (as compared to a mean of 47 and median of 40 among urban employees). To define low-income vs. high-income in CFPS, we breakurban employed workers in Beijing into 8 quintiles based on the annual wage income. Groups 1 to 5 are low-income, 6-8 are high-income.

**Face-to-Face Meetings** Figure 2 presents the raw data. We present monthly averages separately for job switchers and nonswitchers. Panel A of Figure 2 shows the number of friends a worker meets with each month. When comparing switchers and nonswitchers, we find that job switchers consistently have more friends than nonswitchers, a pattern that persists across all observed months. Panel B of Figure 2 illustrates the number of meetings for switchers and nonswitchers. Similar to the number of friends, job switchers on average engage in significantly more face-to-face meetings per month compared to nonswitchers. The decreasing trends in both panels reflect seasonal patterns: Workers tend to meet more friends during warmer months, followed by a gradual decline starting in November as winter approaches.

We begin our formal investigation by comparing the face-to-face meetings among job switchers with referrer and non-referrer friends during the months leading up to the job switch in the following event study analysis:

$$\text{Meetings}_{ijt} = \sum_{s=-15}^{-1} \beta_s 1\{t-t_i = s\}\text{Referrer}_j + \sum_{s=-15, s\neq -2}^{-1} \gamma_s 1\{t-t_i = s\}\text{Nonreferrer}_j + \mu_t + \lambda_{c_i} + \nu_{ijt} \tag{1}$$

where $\text{Meetings}_{ijt}$ is the number of meetings between switchers in cell $i$ and their friends in cell $j$ in month $t$. $t_i$ is the month of job switch. $s$ denotes the number of months prior to the job switch and it takes value from -15 to -1. For example, if the job switch happens in January 2019, then meetings in December 2018 is considered as 1 month before the job change. Months that are more than 15 months before the job switch are binned with the tail $s = -15$. For comparison, the reference group is the number of meetings with non-referrer friends two months before the job change. Month fixed effects $\mu_t$ are included to capture the seasonality in face-to-face meetings. Origin cell fixed effects $\lambda_{c_i}$ are included to control for switchers' unobserved time-invariant characteristics.

Figure 3 illustrates the results. Two key patterns emerge from the analysis. First, job switchers consistently meet more with referrer friends than non-referrer friends, a trend that holds across all months prior to the job change. Second, examining the pattern of meetings with referrer friends over time, we find that the frequency remains stable, with no noticeable spikes around the job change. This stability reassures us that the pre-job-change interactions we observe with friends, particularly referrers, are not endogenously driven by the decision to switch jobs but rather reflect existing social relationships and regular interactions.

**Other Communication Technologies** Besides face-to-face meetings, we have information on the use of other communication technologies including phone calls and WeChat –

the dominant networking application in China.

To capture the communication between switcher-friend pairs, we examine the correlation between (1) face-to-face meetings between switchers in cell $i$ and friends in cell $j$, (2) calls between switchers in cell $i$ and friends in cell $j$, and (3) WeChat data volume associated with switchers in cell $i$ who have friends in cell $j$, during the pre-job-change period. As call data are available starting in January 2019, we focus on individuals who changed jobs between July and December 2019, and examine their communication patterns from January to June 2019.

Table 3 presents the results. Columns 1 to 4 limit to $ij$ cell pairs where switchers in cell $i$ have at least one friend in $j$. The dependent variable is the number of meetings between cells $i$ and $j$ during January to June 2019. The variable "Calls" is the total number of calls between switchers from $i$ and friends in $j$, while "WeChat volume" is the data usage of switchers in cell $i$ who have friends in cell $j$. We find a positive correlation between face-to-face meetings, calls, and WeChat volume across all columns. This positive correlation persists after controlling for block fixed effects and cell-level population density measures, including the number of working individuals and the number of friends. Next, we narrow our focus to $ij$ pairs where switchers from $i$ eventually move to $j$. This restricted sample reflects the communication patterns between switchers and their referrer friends, providing suggestive insights into the role of communication in job transitions. In the restricted sample (Columns 5 to 8), the correlations are higher in magnitude. The coefficient on calls is approximately 2.4, indicating that an additional call is associated with 2.4 more meetings between switchers and their friends. Similarly, the coefficient on WeChat volume is about 1.4, suggesting that a one-gigabyte increase in WeChat data usage by switchers in cell $i$ is associated with an increase of 1.4 meetings between switchers in cell $i$ and their friends in cell $j$.

# 4 Empirical Model and Results

## 4.1 Face-to-face Meetings and Location Choices

To quantify the impact of pre-job-change face-to-face interactions with referrers on job switchers' location choices, we employ the following linear probability model:

$$M_{ij} = \alpha \text{Friend}_{ij} + \beta \text{Friend}_{ij} \times \text{EverMeet}_{ij} + \lambda_{b_i} + \lambda_{b_j} + \lambda_{b_i b_j} + \varepsilon_{ij} \tag{2}$$

where the dependent variables $M_{ij}$ are defined as follows: (1) a dummy variable equals 1 if switchers from cell $i$ changing jobs to cell $j$, and (2) the number of job switchers transitioning from cell $i$ to cell $j$. $\text{Friend}_{ij}$ is a dummy variable equal to 1 if switchers from cell $i$ have at

least one friend working in cell $j$. EverMeet$_{ij}$ is a dummy variable equal to 1 if there is at least one face-to-face meeting between switchers in cell $i$ and their friends in cell $j$. For this analysis, we restrict the potential destination choices of cell $i$ switchers to cells located *within* the blocks containing their new workplace cell. By focusing on cells within the same block, we significantly reduce the influence of unobserved factors that may shape job location decisions, such as industry composition, labor market demand, commuting distance, local amenities, and intrahousehold dynamics. This restriction allows us to better isolate the causal effect of communication technology on job location choices while mitigating the impact of related, unmeasured factors. We include the origin block fixed effects $\lambda_{b_i}$ to account for unobserved switchers' characteristics, such as job preferences, and destination block fixed effects $\lambda_{b_j}$ to control for unobserved location attributes such as the industrial composition and the number of available jobs in the destination area. Additionally, the interaction of origin and destination block fixed effects are included to account for potential correlations stemming from worker flows associated with massive firm reallocations. As a result, the coefficient $\alpha$ is identified by exploiting variation in location choices among individuals moving within the same origin–destination block pair but differing in their friend relationships. The coefficient $\beta$, the primary parameter of interest, identified by exploiting variation in location choices among individuals moving within the same origin–destination block pair and having the same friend network in the destination block, but with different face-to-face interactions with their friends. It captures the effect of face-to-face meetings with friends on individuals' work location decisions.

**Summary statistics** Table 2 reports the summary statistics for the regression sample. On average, the probability of switching from cell $i$ to cell $j$ in one year is about 0.047. The average number of switchers transitioning from $i$ to $j$ is 0.053, with a minimum of 0, a maximum of 4 and the 99th percentile being 1. On average, 9.2 percent of the constructed $ij$ pairs have switchers in cell $i$ with at least one friend working in cell $j$. Additionally, 2.8 percent of $ij$ pairs have a switcher in cell $i$ who had face-to-face meetings with friends in cell $j$ over the 6 months. The average face-to-face meetings between switchers in cell $i$ and their friends in cell $j$ is 3. The average annual WeChat data usage by the switchers is 0.33 GB. On average, there are 0.0002 calls between i-j pairs in the restricted sample (Jan-June 2019).

Turning to meeting composition, conditional on $ij$ pairs with non-zero meetings, the data reveals that 39 percent are meeting involve female friends, while 45 are with low-income friends, and 63 percent are with migrant friends. We further consider the size of the friends' social circles, measured by the number of unique call contacts per month. We categorize

it into three groups – small (less than 50 unique call contact per month), middle (50-100 unique call contact per month) and large (more than 100 unique call contact). 52 percent of the meetings are with friends from small social circles, 30 percent are with friends from middle-sized social circles, and 18 percent are with friends that have large social circles.

**Results** Table 4 reports the results. In Column 1, we only include "Friend" as the regressor, without any block fixed effects. The coefficient is precisely estimated at 0.128. After incorporating the block fixed effects as specified in Model 2, we observe the coefficient on "Friend" decreases by approximately 70 percent, from 0.128 in Column 1 to 0.039 in Column 2. This significant reduction underscores the value of using granular data, as we do, by demonstrating how block fixed effects effectively capture correlated unobservables. In Column 3, we include the interaction term between "Friend" and "EverMeet" to examine the impact of face-to-face interactions, which is the primary focus of our analysis. The coefficient on "Friend" is precisely estimated at 0.031, indicating that, on average, having a friend increases the probability of moving to a specific cell by 0.031. Relative to the mean propensity to choose a cell within a block (0.05), this represents an increase of approximately 0.62 times. Importantly, the coefficient on the interaction term is also precisely estimated at 0.031 and is statistically significant at the 1 percent level. This finding implies that face-to-face meeting with friends further increases the moving propensity by 0.031. Combining these two effects, having a face-to-face meeting with friends will increase the mean propensity to move to a cell by about 1.25 times. In columns 4, we further control for the total number of face-to-face meetings of switchers from cell $i$. The results show the coefficients on "Friend" and its interaction with "EverMeet" remain similar, which suggests the effect of face-to-face meetings on job location choices is consistent across switchers regardless of their overall interaction frequency with friends. Column 3 is our preferred specification.

To ensure that our parameter captures the effect of meeting technology rather than the influence of strong ties, we further control for the initial intensity of face-to-face interactions between switchers and their friends during the friend definition stage (i.e., on weekends and holidays from January to June 2018) in Column 5. The results show that the coefficient on "Friend" decreases from 0.03 to 0.01, while the coefficient on the interaction term remains relatively stable at around 0.03. This suggests that tie strength plays a role in shaping job location choices, but it does not fully explain the impact of meeting technology with friends. In other words, this evidence supports the idea that face-to-face interactions with friends reflect more than just tie strength – they also capture the causal influence of meeting technology.

Next, we explore a possible nonlinear effect of face-to-face meetings on the job location

choices. Specifically, we generate three dummy variables –low, medium, and high – indicating different levels of meetings between switchers in cell $i$ and their friends in cell $j$. These dummy variables represent different meeting intensities, with thresholds determined by the tertiles of the meeting distribution. The "low" group includes cases where the number of meetings between cells $i$ and $j$ is greater than zero but no more than 34. The "medium" group covers cases where the number of meetings is above 34 but no more than 108, while the "high" group corresponds to cases where the number of meetings exceeds 108.

Column 6 of Table 4 reports the nonlinear effects. In Column 6, the coefficients for the "low," "medium," and "high" meeting intensity dummies are 0.023, 0.026, and 0.039, respectively. This pattern indicates that as the frequency of meetings between switchers and friends increases, so does the likelihood of relocating to a particular workplace. However, the upward trend is not particularly pronounced, with a noticeable increase occurring only at the upper end of the distribution.

Appendix Table A2 show the results when the number of job switchers transitioning from cell $i$ to cell $j$ is used as dependent variable. The results are similar. Column 3 with our preferred specification suggests that having face-to-face meetings with friends would increase the number of job switchers moving to a given work cell by 0.083 (0.046+0.037), which increases the mean (0.053) by about 1.57 times. The results are similar to that in Table 4 given that in most of the cells there is just one switcher.

**Robustness and Placebo Tests** Table 4 provides evidence against homophily by controlling for the initial intensity of face-to-face calls between switchers and their friends. To examine homophily in further detail, in Table 5 we exploit variations in observed characteristics among switchers and friends and directly control for homophily in the model. Further evidence against homophily will be found in table 7, where we control for the intensity of the relationship between switchers and friends as measured by phone call frequency, and in table 6 where we compare the effects for different types of friends.

In Table 5, we construct two sets of controls for similarity in observed characteristics. First, we add three dummies indicating whether the switchers in cell $i$ are of the same gender, migration status and in the same income group as their friends working in cell $j$. Second, instead of using parametric functions of observables as in column 1 we use a popular unsupervised machine learning algorithm – k-means clustering – to nonparametrically profile switchers and friends. that groups switchers and their friends into clusters based on similarity. Specifically, we define a cell as dominated by switchers or friends with certain characteristics if the share of switchers or friends with that characteristic is above the 75th percentile of their

distribution across cells. [9] We consider a vector of 12 variables: whether a cell is dominant in female, male, migrant, local workers/friends, and in each of the 8 income groups (see section 2.1). We then use the clustering algorithm to group cells into 10 clusters and control for a dummy variable indicating whether the origin cell $i$ of switchers is in the same cluster of the destination cell $j$ of their friends. The results in Columns 1 of Table reftab:robustness show that the coefficients on the interaction between "Friend" and "EverMeet" is approximately 0.03, similar to the baseline estimate in Table 4, indicating that our granular fixed effects are adequate in capturing sorting and that our results are not driven by homophily.

A potential explanation for the estimated effect is that more face-to-face meetings may merely reflect the effect of higher population density. To investigate this pathway, we control for the working population[10] in each alternative destination cells as a proxy for population density. Columns 2 and 5 show the result. While the effects of having a friend in the workplace drop dramatically (from 0.031 to 0.002), the effect of the meeting technology is only marginally absorbed by the effect of population density.

To further differentiate the impact of face-to-face meetings from simply having friends, Column 3 examines whether the effect of in-person interactions is driven solely by the number of friends. To do this, we control for the number of friends that switchers in cell $i$ have in cell $j$. The coefficient on the interaction between "Friend" and "EverMeet" is 0.024 in Column 3, and statistically significant at the 1 percent level. It suggests that conditional on the same number of friends, face-to-face meeting with friends still matter. Appendix Table A3 shows these robustness checks when the number of switchers changing jobs between areas is used as alternate dependent variable. The results remain similar to the baseline estimates in Appendix Table A2 in all columns.

We conclude this section by comparing the effects of face-to-face interactions with friends who are currently working at the new work location to those with friends who have recently moved away or will relocate there in the future. This analysis provides insight into whether the role of face-to-face interactions in labor mobility is driven by an information-sharing channel. Since friends currently working at a location are likely to be more informed about job opportunities or offer better access to them than those who are not, we expect a stronger association between face-to-face interactions with informed friends and job location choices compared to meetings with friends who are elsewhere. At the same time, this serves as a further test against homophily. If the intensity of face-to-face interactions merely reflects similarities in observed and unobserved characteristics between friends, we should see no

[9] For example, an origin cell $i$ is female dominant if the share of female switchers is above its 75th percentile in the female share distribution across cells.

[10] Working population in a cell is the number of workers primarily working in a cell from the cellphone location data.

difference in the estimated effects across different types of friends. To implement our exercise, we focus on job switchers who change jobs during the middle three months of the year – July to September 2019 – and consider face-to-face meetings in the preceding three months – April to June 2019. We distingush three types of friends: (1) friends who currently work in cell $j$ throughout 2019; (2) friends who previously worked in cell $j$ (in April-June 2018) but left before the switchers' move (by April 2019); (3) friends who were not in cell $j$ since July 2018 but joined after September 2019. We construct three dummy variables for face-to-face meetings before the job change: "EverMeet", "EverMeet (left)" and "EverMeet (will join later)", corresponding to meetings with the three types of friends, respectively. Table 6 reports the results. Columns 1-3 include each meeting dummy separately, while Column 4 adds all three in one regression. The results indicate that face-to-face meetings with friends who worked in cell $j$ throughout the year have a significant effect on the probability of switching (coefficient = 0.21), which is both larger and statistically distinct from the effects of meetings with the other two types of friends who are less likely to possess relevant job information (coefficients = 0.13 and 0.12, respectively). A similar pattern is observed in Appendix Table A4, which analyze the number of switchers moving from $i$ to $j$.

## 4.2 Other Communication Channels

Our findings so far highlight the importance of face-to-face interactions in driving referral effects. A natural question that follows is how these effects compare to those facilitated by different communication technologies. To explore this, we leverage supplementary data on phone calls and WeChat activity to examine the relative impacts of in-person meetings, phone conversations, and mobile messaging apps on job location choices, controlling for the presence of a friend.

Specifically, we focus on a common observation window for communication methods – January to June 2019 – and examine job changes occurring from July to December 2019. Table 7 shows the result. To make data from different technologies comparable, we first standardize the measures for face-to-face meetings, calls and WeChat volume by subtracting their means and dividing by their standard deviations. Columns 1-3 of Table 7 include each standardized measure separately, while Column 4 includes all three measures in a single regression. Results suggest that a one standard deviation increase in face-to-face meetings with friends would increase the probability of switching to the friends' workplace by 0.001, which is higher than the effects of one standard deviation increase in calls (0.0002) or WeChat volume (0.0005).

Looking at the extensive margin, we consider the following dummies for the three tech-

nologies: "EverMeet" for switchers from cell $i$ having at least one face-to-face meeting with friends in cell $j$, "EverCall" for having at least one call, and "EverWechat" for switchers who have friends in $j$ having nonzero total WeChat data volume. Column 5 reports the result. The coefficient on the interaction term between "Friend" and "EverMeet" is precisely estimated at around 0.03, almost 4 times the coefficients on the interaction with "EverCall" and "EverWeChat". This suggests that having face-to-face interactions with friends significantly increase the likelihood of switching to a friend's workplace, while the effect of calls and WeChat activity remain modest.

# 5 Heterogeneity

Face-to-face interactions may hold greater value for certain workers and job types compared to others. In this section, we examine the varying impacts of in-person interactions on job location choices across different worker and job types. Additionally, we explore the potential mechanisms driving these differences.

## 5.1 Heterogeneity By Worker Characteristics

To investigate heterogeneity patterns in the returns to face-to-face interactions, we re-estimate our baseline model (Equation 2) for subsets of switchers with different characteristics. As the worker flow and pre-job-change interactions are aggregated at cell pair level, we restrict the analysis to cell pairs where the majority of job switchers in the origin cell $i$ share a specific characteristic. For instance, if the proportion of female switchers in a given cell exceeds the 75th percentile of its distribution, we classify this cell as "female-dominant" and include it in the analysis for female workers. We examine heterogeneity along four dimensions: (1) gender (female vs. male); (2) migration status (migrant vs. local); (3) employment type (full-time vs. part-time); and (4) income level (low-income vs. high-income). Since, as previously documented, the vast majority of cell pairs contain only one switcher (see Table 2) , we can reasonably interpret the results at the individual switcher level.

Panel A of Table 8 reports the effect heterogeneity employing the specification in model 2. Each coefficient is estimated from a separately regression on the respective subsample. Columns 1 and 2 compare the effect of face-to-face meeting with referrers for female and male workers. To ensure the validity of this comparison, we exclude cases where the proportion of female or male switchers in the origin cell falls below the 75th percentile, as these cells do not exhibit a clear majority gender. Conditional on having a friend, among female workers, having face-to-face meetings with friends increase the probability of relocating to a given

cell by 0.037, compared to 0.0199 for male workers. A t-test confirms that this difference is statistically different from zero (p-value = 0.042).

Columns 3 and 4 contrast the effects of face-to-face interaction with friends between migrant and local workers. For migrant workers, such interactions increase the probability of moving to a given cell by about 0.038, whereas the effect for local workers is substantially smaller at 0.0165. The t-test also indicates a statistically significant difference between these two groups.

Comparing the coefficients in Columns 5 and 6, we observe a pronounced gap in the effects of face-to-face meetings with friends between full-time workers and part-time workers. Full-time workers experience a larger benefit with their moving probability increasing by about 0.04, almost 4 times the effect observed for local workers (0.013). Finally, Columns 7 and 8 compare the effects for high-income and low-income workers. Interestingly, face-to-face meetings with friends yield a stronger effect for high-income workers (0.035) than for low-income workers (0.023).

In summary, our findings suggest that conditional on having a friend, face-to-face meetings with the friend are particularly valuable for female workers, migrants, full-time workers and high-income workers.

## 5.2 Heterogeneity By Job Characteristics

Next, we analyze the heterogeneous effects of face-to-face interaction with friends on location choices between private and public sectors and across industrial sectors. Similar to Section 5.1, we restrict the analysis to cell pairs for which the origin cell $i$ belongs to a specific type [11] Our first analysis contrasts switchers moving from areas dominated by the public versus private sectors[12]. We then extend the analysis to explore heterogeneity across switchers originating from areas dominated by specific leading industries. According to employment statistics from Beijing Municipal Bureau of Statistics in 2019[13] and China Statistical Yearbook 2020[14], the top industrial sectors that contribute to over 60 percent of GDP in Bejing are IT and software, finance, leasing and business services, wholesale trade and retail, and manufacturing.

[11]Specifically, we classify a cell as dominated by a particular sector if most of the land (more than 75%) is occupied by firms in that sector.

[12]With the 16-industry classification, "Public" sector includes industries "Public Management, Social Security and Social Organizations", "Health and Social Work", "School/Science Sites" under "Education", "Water Conservancy, Environment and Public Facilities Management", "Scientific Research and Technical Services", "Culture Sites", and "Scenic area"; "Private" sector includes firms and enterprises in all other industries not classified as public.

[13]https://nj.tjj.beijing.gov.cn/nj/main/2019-tjnj/zk/indexeh.htm

[14]https://www.stats.gov.cn/sj/ndsj/2020/indexeh.htm

Table 9 reports the heterogeneous results by sector which employs the specification as in model 2 with different controls. Panel A uses the preferred specification in Table 4 and includes no additional controls. Columns 1 and 2 compare the effect of face-to-face interaction with friends on location choice for switchers originating from areas dominated by public sector jobs versus those from areas dominated by private sector jobs. The coefficient in Column 1 is 0.0176, suggesting that face-to-face interactions with friends increase the probability of switching to a given area by approximately 0.018 for switchers from public sectors. In contrast, the coefficient in Column 2 is 0.034, indicating that the probability of switching to a given area rises by about 0.034 for switchers from private sectors who interact face-to-face with friends working there. These results highlight that switchers from private sectors are more likely to benefit from face-to-face interactions with friends than their counterparts from public sectors, with the difference being statistically significant at the 1 percent level.

Similarly, Columns 3 to 8 of Table 9 compare the effects of face-to-face interaction with friends for switchers originating from specific industries versus those from other industries. We find that switchers from IT and software, finance, and manufacturing industries benefit significantly more from face-to-face interaction with friends compared to their counterparts from other industries. In contrast, Appendix table A5 shows no significant differences in the impact of face-to-face interactions between workers in business services or wholesale/retail trade industries compared to workers not in these specific industries.

## 5.3 Mechanisms Behind Differential Effects

What could explain the differential effects of face-to-face interactions on workers and jobs of different types? We consider two potential mechanisms. First, the *quantity channel*: the number of face-to-face meetings may play a crucial role. If certain types of workers tend to engage in frequent face-to-face meetings with friends, they may derive greater benefits from these interactions. Second, the *quality channel*: even if the number of meetings is similar across worker types, differences in the composition or quality of these interactions may explain the observed variations. Specifically, we examine whether having the majority of face-to-face interactions with friends of a particular gender, migration status, income level, or social circle size shapes the returns to face-to-face interactions. We will look at the effects of meetings with a diverse type of friends in Section 5.4.

To explore these mechanisms, we begin by providing some descriptive evidence. Table 10 presents the frequency of face-to-face interactions across worker types. We limit our sample to observations where the origin cell has dominant characteristics in gender, migration status,

employment type, or income level as in Table 8. In Columns 1 through 4, we include dummy variables indicating whether the origin cell is female-dominant, migrant-dominant, full-time dominant, or high-income dominant, respectively. The positive coefficients on these four dummies indicate systematic differences in meeting frequencies across groups. Specifically, female workers meet more often with friends than male workers, migrant workers engage in more face-to-face interactions than local workers, full-time workers participate in more meetings than part-time workers, and high-income workers meet more frequently than low-income workers. These findings provide initial evidence that the quantity channel may contribute to the heterogeneity in the effects of face-to-face interactions.

Table 11 reports heterogeneity in meeting composition across worker types. The dependent variables are dummy variables indicating whether the majority of meetings of switchers (in origin cell $i$) are with friends (in cell $j$) who are female, migrant, low-income friends and friends of different social circles[15]. We measure social circle as the average number of unique call contacts between January and March 2019. Each coefficient is estimated from a separate regression. As in Table 10, we include dummy variables indicating whether the origin cell is female-dominant, migrant-dominant, full-time dominant, or high-income dominant, and limit the sample to observations where the origin cell has dominant characteristics in gender, migration status, employment type, or income level. The positive (or negative) coefficients on the four dummies indicate systematic differences in meeting compositions across groups. Noticeably, we find that full-time and migrant workers meet more connected individuals, whereas female and high-income workers do not.

To understand the extent to which the quantity and quality of face-to-face meetings are important channels for the differential impact of meeting on the referral effect for different types of workers, we employ the sample-splitting analysis with the baseline model (Equation 2) with variations in controls.

First, we add in the number of face-to-face meetings between areas as an additional control. If meeting quantity plays an important role in explaining the differential effects across groups, we would expect that the coefficients on the interaction between "Friend" and "EverMeet" become similar after adding such a control. Panel B of Table 8 reports the results. The gap in effects between female and male workers decreases from -0.011 to -0.008 and becomes statistically insignificant (p-value = 0.208) when controlling for the number of meetings. Similarly, the gap closes between high-income and low-income workers and becomes no longer statistically different from zero (p-value = 0.676). Therefore, the

[15] "Majority" here is defined as cases where the share of meetings with a certain type of friend (e.g., female friends) exceeds the mean of the nonzero distribution of meeting shares. Specifically, the meeting share is calculated as the number of meetings with that group divided by the total number of meetings associated with a switcher.

findings suggest that meeting quantity helps explain the premium for female and high-income workers. However, although the gap narrows as compared to that in Panel A, the difference still remains statistically significant between migrant and local workers and between full-time and part-time workers.

We then additionally control for meeting composition, including a set of dummies indicating whether the majority of meetings of switchers in a cell $i$ are with friends in cell $j$ who are female, migrant, low-income, or with different sizes of social circles. These dummies are the same as the dependent variables used in Table 11. As reported in Panel C, the differences in effects between full-time and part-time workers, as well as between migrants and locals, reduce substantially and become statistically insignificant with these composition controls. This suggests that meeting composition is the key channel driving the premium for full-time and migrant workers. Corroborating the findings from the descriptive statistics, meeting with more connected individuals appears to be an important mechanism underlying the face-to-face interaction effects for full-time and migrant workers.

Turning to the heterogeneous effects across switchers originating from areas dominated by different industries, we also find some descriptive evidence on differences in meeting frequency and meeting composition. Table 12 suggests that switchers from private sectors tend to have more meetings than those from public sectors, though this difference is not statistically different from zero. In addition, switchers from top industries – such as IT and software, finance, and manufacturing engage in more frequent meetings with friends compared to their counterparts. Table 13 further indicates that switchers originating from the public sectors are less likely to meet with connected individuals, whereas those from specific top industries are more likely to do so. As a formal analysis, we incorporate meeting quantity and composition as controls in the baseline specification, following the approach in Panels B and C of Table 8. The corresponding results are presented in Panels B and C of Table 9. Adding in the number of meetings as a control in Panel B, the gap between public and private sectors narrows and becomes statistically insignificant (p-value = 0.462), suggesting that meeting quantity helps explain the premium for workers from private sectors compared to those from public sectors. Further controlling for meeting composition in Panel C, the difference between IT versus non-IT, finance versus non-finance, and manufacturing versus non-manufacturing workers also diminishes and becomes statistically insignificant. These findings suggest that meeting with more connected individuals plays a key role in explaining the differential benefits of face-to-face interaction effects for workers in top industries and their counterparts.

### 5.4 Diverse vs. Homophilous Meetings

As workers of different types enjoy differential benefits from meeting with different friends, we further zoom into the compositional effects of meeting with similar or dissimilar friends. We construct two meeting composition dummies: (1) "Majority same type" indicating whether the majority meetings are with friends of the same type as the job switchers, and (2) "Majority different type" indicating whether the majority meetings are with friends of a different type from the job switchers. We then use the sample-splitting analysis as in Panel C of Table 8, and replace the set of meetings composition dummies with these same-type and different-type dummies in two separate regressions. [16]

Figure 4 represents the effects of meeting with same- vs. different-type friends across worker types. Each dot is the point estimate for the coefficient on "Majority same type" or "Majority different type" separately on a subsample, surrounded by bars indicating the 95% confidence interval. We find that for female-dominant origin cells, the effect of meeting with same-gender friends is precisely estimated at 0.009, suggesting that meeting with female friends additionally increase female workers' propensity to move to a given cell by about 0.01, about 20 percent increase from the mean propensity. However, the effect of meeting with different-gender friends is -0.0087, suggesting meeting with male friends reduces the moving propensity of female workers. Similarly, for high-income dominant origin cells, the effect of meeting with same-type friends is precisely estimated at 0.0132, while the effect of meeting with different-type friends is -0.0128. Although we do not see a significant difference for the other types of workers, the coefficients generally show a similar pattern – positive for same-type and negative for different-type meetings (except for local workers). These findings suggest that homophilous face-to-face interactions are more valuable than diverse interactions in shaping job location choices.

Turning to the meeting with same- versus different-type friends across sectors, we consider a worker has the majority of meetings of the same type if the worker and friends work in the same sector. Using the sample-splitting analysis, Figure 5 reports the results. The estimates show that the effects of homophilous meetings are more valuable than diverse meetings.

## 6 Conclusion

Our analysis highlights the crucial role of face-to-face interactions, which are central to influential theories on the advantages of urban density (Jacobs, 1960; Storper and Venables,

[16] For example, for female-dominant cells where a majority of the switchers are female, we re-estimate the baseline model with "Majority same type" as an additional control first, then with "Majority different type" as an alternative additional control again.

2004). Our investigation is important for at least two reasons. First, despite the importance of social network in job transitions (Topa, 2011; Schmutte, 2016), accurately measuring these networks remains a challenge. Our analysis illustrates how the availability of longitudinal and geocoded location data is an important step towards overcoming the measurement challenge. Second, our analysis indicates that face-to-face interactions are an important pathway for understanding labor market dynamics, thus uncovering a mechanism of agglomeration economies. Our results thus provide insights for urban planner and policy makers aiming at designing cities maximizing the advantages of agglomeration forces.

# References


**Atkin, David, M. Keith Chen, and Anton Popov.** 2022. "The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley." National Bureau of Economic Research Working Paper 30147.

**Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel.** 2018. "The Economic Effects of Social Networks: Evidence from the Housing Market." *Journal of Political Economy*, 126(6): 2224–2276.

**Barwick, Panle Jia, Yanyan Liu, Eleonora Patacchini, and Qi Wu.** 2023. "Information, Mobile Communication, and Referral Effects." *American Economic Review*, 113(5): 1170–1207.

**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.

**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.

**Chetty, Raj, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, Drew Johnston, Martin Koenen, Eduardo Laguna-Muggenburg, Florian Mudekereza, Tom Rutter, Nicolaj Thor, Wilbur**

**Townsend, Ruby Zhang, Mike Bailey, Pablo Barberá, Monica Bhole, and Nils Wernerfelt.** 2022*a*. "Social capital I: measurement and associations with economic mobility." *Nature*, 608(7921): 108–121.

**Chetty, Raj, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, Drew Johnston, Martin Koenen, Eduardo Laguna-Muggenburg, Florian Mudekereza, Tom Rutter, Nicolaj Thor, Wilbur Townsend, Ruby Zhang, Mike Bailey, Pablo Barberá, Monica Bhole, and Nils Wernerfelt.** 2022*b*. "Social capital II: determinants of economic connectedness." *Nature*, 608(7921): 122–134.

**Duranton, Gilles, and Diego Puga.** 2004. "Chapter 48 - Micro-Foundations of Urban Agglomeration Economies." In *Cities and Geography*. Vol. 4 of *Handbook of Regional and Urban Economics*, , ed. J. Vernon Henderson and Jacques-FranÃ§ois Thisse, 2063–2117. Elsevier.

**Duranton, Gilles, and Diego Puga.** 2020. "The Economics of Urban Density." *Journal of Economic Perspectives*, 34(3): 3â26.

**Friebel, Guido, Matthias Heinz, Mitchell Hoffman, and Nick Zubanov.** 2023. "What Do Employee Referral Programs Do? Measuring the Direct and Overall Effects of a Management Practice." *Journal of Political Economy*, 131(3): 633–686.

**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.

**Jacobs, Jane.** 1960. *Dark age ahead: Author of the death and life of great American cities.* New York: Random House.

**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.

**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*.

**Storper, Michael, and Anthony J. Venables.** 2004. "Buzz: face-to-face contact and the urban economy." *Journal of Economic Geography*, 4(4): 351–370.

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

# Figures and Tables

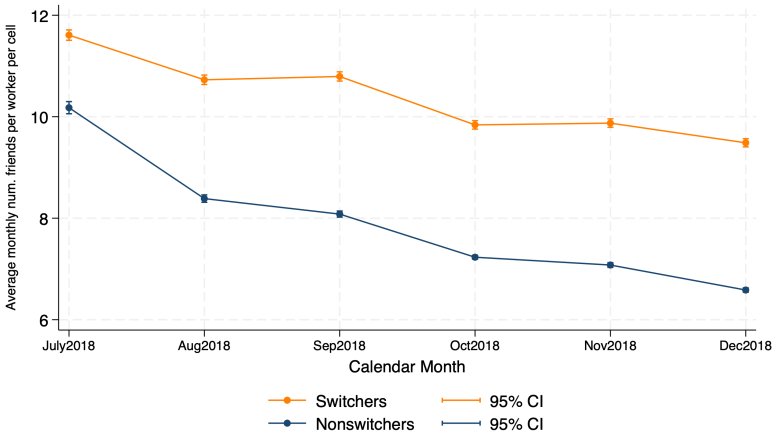
**Figure 1:** Comparison with official statistics



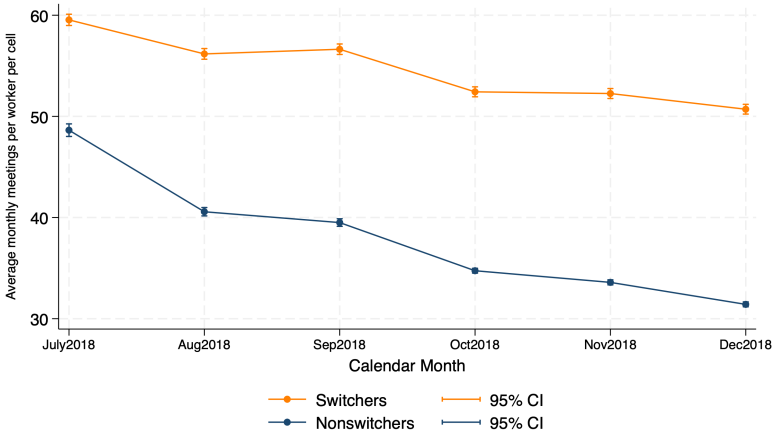

*Notes*: Figure 1 shows the average fraction of workers that are female, migrant, full-time and low-income in three samples. The first bar (in green) examines all urban employees in Beijing in CFPS 2020. The second bar (in blue) uses all workers in a cell in our cellphone data, and the third bar (in orange) limits to switchers in a cell. Number of cells = 13585. To define "full-time" employees in CFPS, we first identify those ever work in part-time jobs if either of the following three conditions is satisfied: (a) they have not worked full-time, (b) they have worked as part-time, and (c) they have other jobs beside their main job. Then those who have not worked in part-time jobs are potential full-time employees. Next, we further exclude the following two groups of individuals from the potential full-time employees: (a) work in agricultural sector, (b) weekly working hours is less than 30 hours. To define "low-income" in CFPS, we break urban employees into 8 quintiles based on annual wage income. Groups 1 to 5 are low-income, 6-8 are high-income.

**Figure 2:** Number of friends and meetings over time

**(a)** Number of friends ever meet per worker

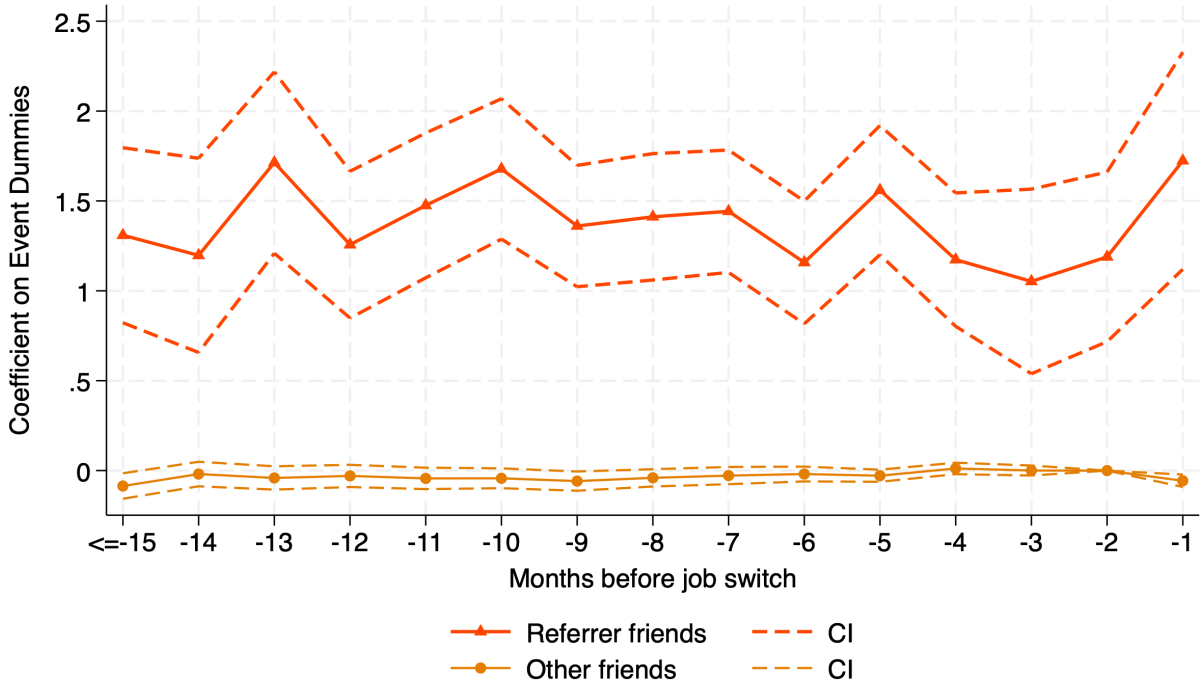



**(b)** Number of meetings per worker



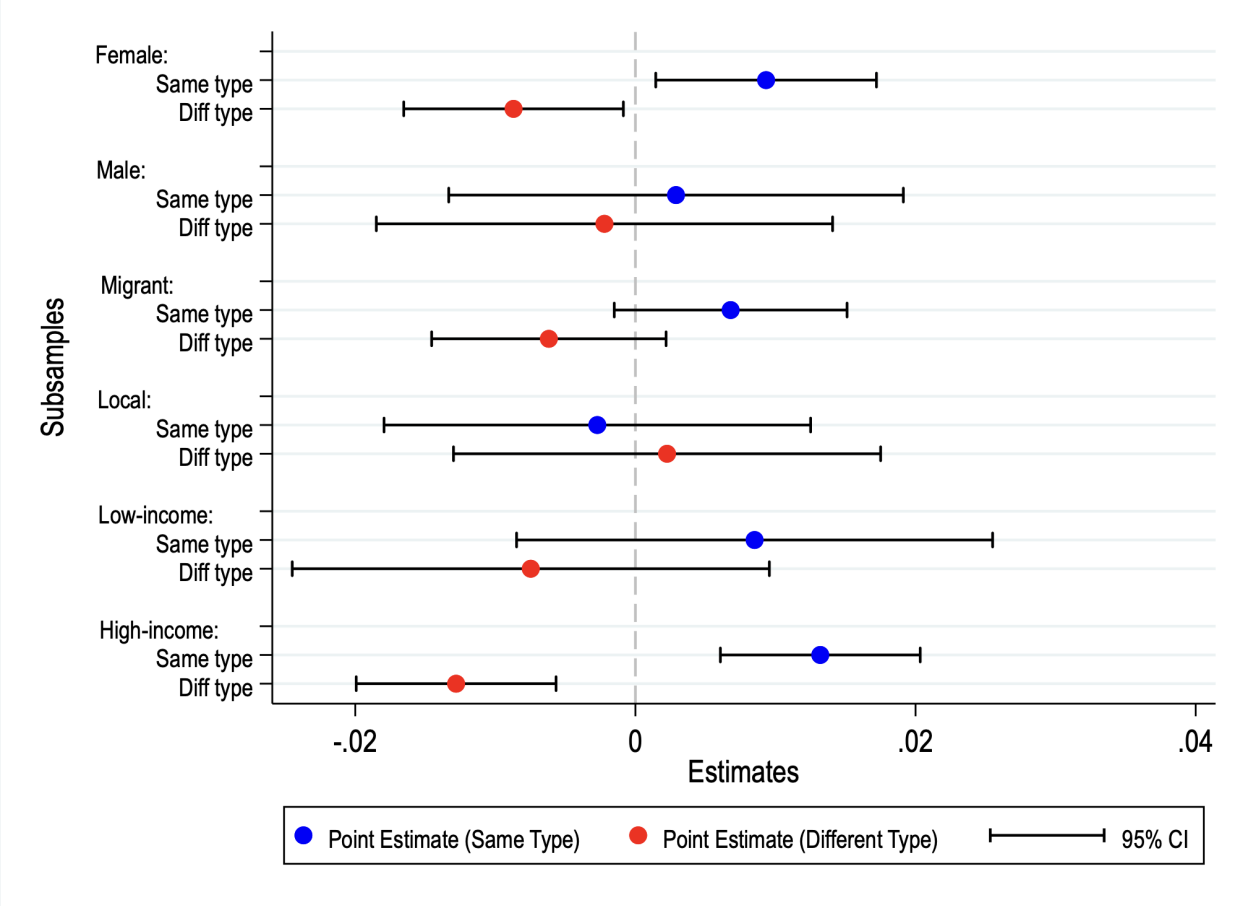

*Notes*: Figures 2a and 2b plot the average number of friends ever meet and the number of meetings per worker have in a given month and their 95 percent confidence intervals, respectively. In both figures, the orange line (with triangles) denotes average monthly friends and meetings related to job switchers across cells. The blue line (with dots) denotes that related to all workers across cells. The point estimates and confidence intervals are obtained from regressing number of friends/meetings on calendar month dummies on a sample where the unit of observation is a cell-month pair.

**Figure 3:** Face-to-face meetings with Referrer vs. Nonreferrer friends



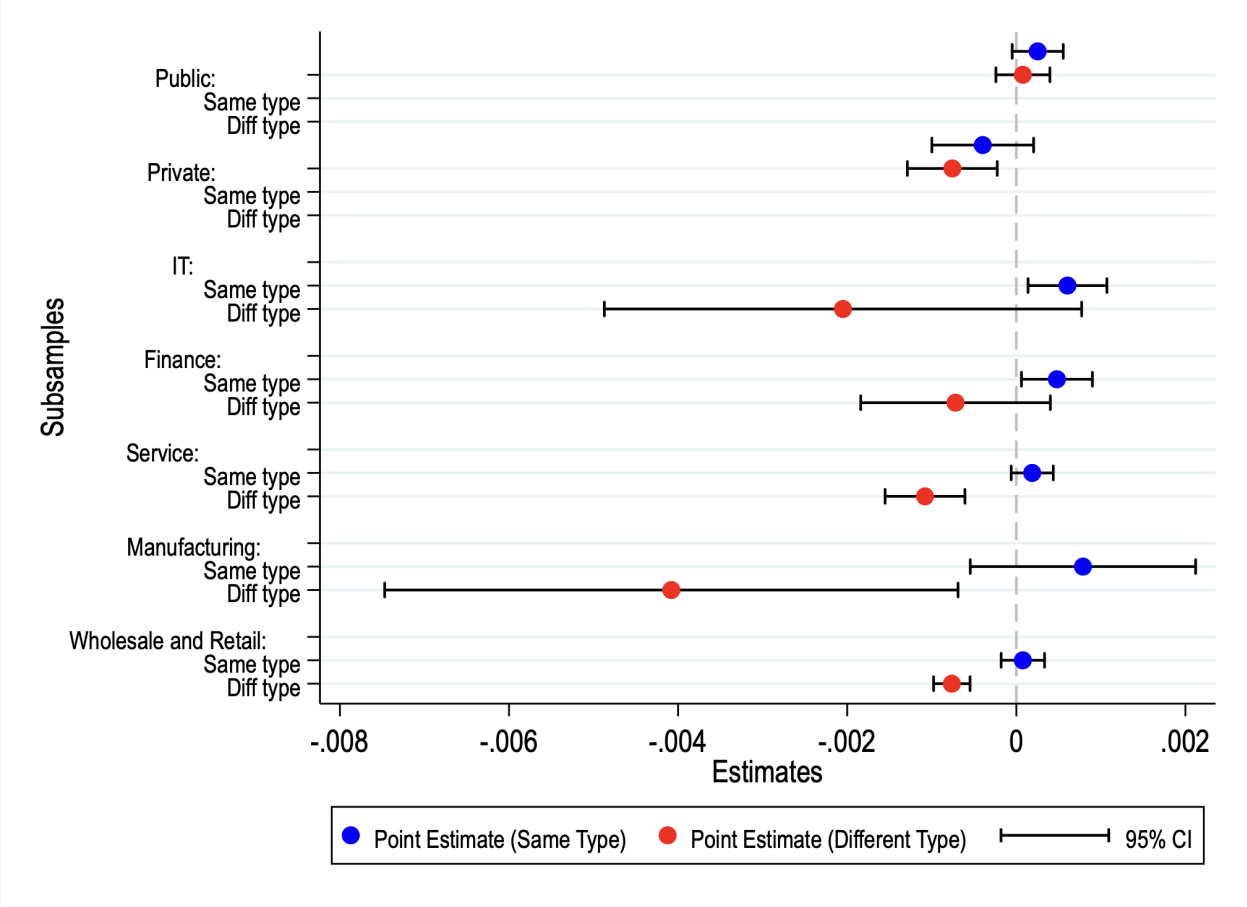

*Notes*: Figure 3 plot the event-study coefficients and their 95 percent confidence intervals for the number of face-to-face meetings between switchers in the origin cell and referrer friends vs. nonreferrer friends in a given cell before and the job switch, respectively. The red line (with trangles) denotes meetings between switchers and referrer friends (149,701 obs). The blue line (with dots) denotes meetings between switchers and their nonreferrer friends (62,266,236 obs). The reference group is the number of meetings between switchers and nonreferrer friends two months prior to the job switch. The period prior to the 15th month before the job switch is binned with the 15th month. The point estimates and confidence intervals are obtained from regressing number of meetings on event dummies on a sample where the unit of observation is at $ij$ cell pair-month level. Calendar month fixed effects and the origin cell fixed effects are included in the regression.

**Figure 4:** Effect of meeting with friends of same-type vs. different-type: by worker characteristics




*Notes*: Figure 4 represents the effects of meeting with same- vs. different-type friends across worker types. Each coefficient is estimated on a separate subsample where a majority of the switchers are female, male, migrant, local, low-income and high-income workers, respectively. In each subsample, we re-estimate the baseline model with "Majority same type" or "Majority different type" as an additional control. Each dot is the point estimate for the coefficient on "Majority same type" or "Majority different type" separately on a subsample, surrounded the 95% confidence interval.

**Figure 5:** Effect of meeting with friends of same-type vs. different-type: by sector




*Notes*: Figure 5 represents the effects of meeting with same- vs. different-type friends across sectors. Each coefficient is estimated on a separate subsample where a majority of the switchers are originating from areas dominated by public sectors, private sectors, or specific top industries (including IT, finance, leasing and business service, and manufacturing), respectively. In each subsample, we re-estimate the baseline model with "Majority same type" or "Majority different type" as an additional control. Each dot is the point estimate for the coefficient on "Majority same type" or "Majority different type" separately on a subsample, surrounded the 95% confidence interval. We define friend's industry as the dominant industry of the friend's primary work area. "Same type" and "different type" refers to whether the majority of meetings are with friends working in the same sector (public/private, specific industries) or in a different sector.

**Table 1:** Job Changes by Month

| Month | Number of job changes | Percent among total job changes | Switching rate |
|---|---|---|---|
| Jan | 25,059 | 9.18% | 0.72% |
| Feb | 26,024 | 9.53% | 0.75% |
| Mar | 27,208 | 9.97% | 0.95% |
| Apr | 23,788 | 8.72% | 0.70% |
| May | 23,563 | 8.63% | 0.69% |
| Jun | 20,666 | 7.57% | 0.62% |
| Jul | 23,037 | 8.44% | 0.69% |
| Aug | 21,078 | 7.72% | 0.62% |
| Sep | 27,362 | 10.03% | 0.84% |
| Oct | 18,757 | 6.87% | 0.57% |
| Nov | 19,539 | 7.16% | 0.60% |
| Dec | 16,856 | 6.18% | 0.51% |
| Total | 272,937 | 100% | |

*Notes*: This table shows the number and percent of job changes, and switching rate in each month in 2019. Switching rate is the number of job changes over the number of working population in each month in 2019.

**Table 2:** Summary Statistics

| Variable | Mean | SD | Min | Max |
|---|---|---|---|---|
| *Panel A. Worker flow and F2F communication* | | | | |
| Number of switchers from $i$ to $j$ | 0.053 | 0.26 | 0 | 4 |
| Probability of switching from $i$ to $j$ | 0.047 | 0.212 | 0 | 1 |
| Number of friends in $j$ of switchers from $i$ | 0.132 | 0.50 | 0 | 6 |
| Probability of having a friend in $j$ of switchers from $i$ | 0.092 | 0.29 | 0 | 1 |
| Number of face-to-face meetings between $i$ and $j$ | 3.038 | 26.887 | 0 | 444 |
| Probability to face-to-face meetings between $i$ and $j$ | 0.028 | 0.165 | 0 | 1 |
| Number of face-to-face meetings between $i$ and $j$ (conditional on having meetings) | 3.038 | 26.887 | 0 | 444 |
| *Panel B. Share of meeting with xx friends, conditional on having meetings* | | | | |
| Female | 0.39 | 0.46 | 0 | 1 |
| Low-income | 0.45 | 0.47 | 0 | 1 |
| Migrant | 0.63 | 0.46 | 0 | 1 |
| network size (50,100] | 0.30 | 0.43 | 0 | 1 |
| network size (100,.) | 0.18 | 0.37 | 0 | 1 |
| *Panel C. Other communication methods* | | | | |
| Wechat volume in origin cell (GB) | 0.334 | 2.216 | 0 | 8.885 |
| Calls between $i$ and $j$ (restricted sample) | 0.0002 | 0.067 | 0 | 55 |

*Notes*: This table shows summary statistics in the regression sample. In Panel A, the number of area pairs is 4,810,573. In Panel B, we conditional on area pairs with at least one meeting in the pre-job-change period. The number of area pairs is 134,633. In Panel C, the number of area pairs is 4,810,573; for calls is 2,169,538 in the restricted sample.

**Table 3:** Complementarity of F2F meetings, Phone Calls and WeChat Volume of Switchers Before Job Change

| Dependent variable: | Switchers in $i$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| F2F between cell $i$ and $j$ | with at least one friend in $j$ | | | | move to $j$ July-Dec 2019 | | | |
| Jan-June 2019 | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Calls | 1.096** | 1.094** | 1.095** | 1.094** | 2.422* | 2.452** | 2.396* | 2.426** |
| | (0.442) | (0.441) | (0.441) | (0.441) | (1.249) | (1.247) | (1.233) | (1.232) |
| Wechat volume (GB) | 0.946*** | 0.917*** | 0.940*** | 0.916*** | 1.585*** | 1.501*** | 1.473*** | 1.431*** |
| | (0.019) | (0.019) | (0.019) | (0.019) | (0.092) | (0.093) | (0.096) | (0.096) |
| Work population in $i$ | | 0.002*** | | 0.002*** | | 0.002*** | | 0.001** |
| | | (0.0001) | | (0.0001) | | (0.0003) | | (0.0003) |
| Work population in $j$ | | 0.002*** | | 0.002*** | | 0.002*** | | 0.002*** |
| | | (0.0001) | | (0.0001) | | (0.0003) | | (0.0003) |
| Number of friends | | | 0.001*** | 0.0002*** | | | 0.002*** | 0.002*** |
| | | | (0.00005) | (0.0001) | | | (0.0003) | (0.0003) |
| Observations | 2,700,372 | 2,700,372 | 2,700,372 | 2,700,372 | 125,387 | 125,387 | 125,387 | 125,387 |
| R-squared | 0.342 | 0.342 | 0.342 | 0.342 | 0.475 | 0.477 | 0.478 | 0.479 |
| Block i FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Block j FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Block i*j FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes*: This table shows complementarity across F2F, calls and WeChat associated with switchers and friends before the job change. The unit of observation is a $ij$ cell pair. Columns 1 to 4 consider $ij$ pairs where switchers in cell $i$ have at least one friend in $j$. Columns 5 to 8 limit to $ij$ pairs where switchers in cell $i$ land in jobs in cell $j$ July-December 2019. The dependent variable is the total number of face-to-face meetings between switchers in cell $i$ and friends in cell $j$ in January-June 2019. "Calls" is the total number of calls between switchers and friends in January-June 2019. "WeChat volume (GB)" is the total WeChat data volume of switchers in cell $i$ who have friends in $j$ in January-June 2019. Columns 2, 4, 6 and 8 controls for the number of work population in cells $i$ and $j$. Columns 3, 4, 7 and 8 controls for the number of friends in cell $j$ of switchers from cell $i$. All columns include origin block fixed effects, destination block fixed effects and origin-by-destination block-pair fixed effects. Standard errors are clustered by block pair and reported in parentheses.

**Table 4:** Effect of F2F on Job Location Choices

| Dependent variable: Probability of switching from $i$ to $j$ | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Friend | 0.128*** | 0.039*** | 0.031*** | 0.032*** | 0.013*** | 0.013*** |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Friend*EverMeet | | | 0.031*** | 0.031*** | 0.029*** | |
| | | | (0.001) | (0.001) | (0.001) | |
| F2f meetings: Low | | | | | | 0.023*** |
| | | | | | | (0.002) |
| F2f meetings: Medium | | | | | | 0.026*** |
| | | | | | | (0.002) |
| F2f meetings: High | | | | | | 0.039*** |
| | | | | | | (0.002) |
| Cell $i$ total F2F (x1000) | | | | -0.0001*** | -0.0001*** | -0.0001*** |
| | | | | (0.00001) | (0.00001) | (0.00001) |
| Observations | 4,805,218 | 4,787,639 | 4,787,639 | 4,787,639 | 4,787,639 | 4,787,639 |
| R-squared | 0.030 | 0.168 | 0.169 | 0.169 | 0.169 | 0.169 |
| Initial intensity control | No | No | No | No | Yes | Yes |
| Block i FE | No | Yes | Yes | Yes | Yes | Yes |
| Block j FE | No | Yes | Yes | Yes | Yes | Yes |
| Block i x block j FE | No | Yes | Yes | Yes | Yes | Yes |

*Notes*: This table examines the effect of face-to-face meetings on job changers' location choices. The unit of observation is a $ij$ cell pair. The dependent variable is the probability that switchers change jobs from cell $i$ to cell $j$. "Friend" is a dummy variable that equals one if switchers in cell $i$ have at least one friend working in cell $j$. "EverMeet" is a dummy variable that equals one if switchers from cell $i$ ever face-to-face meeting with their friends working in cell $j$. "F2F meetings: Low/Medium/High" are dummies for the number of face-to-face meetings between cells $i$ and $j$ being above zero and in each tertile, with cutoffs at 34 and 108 meetings. Columns 2-6 include origin block fixed effects, destination block fixed effects and origin-by-destination block-pair fixed effects. Columns 4-6 controls for the total number of face-to-face meetings of switchers from cell $i$. Columns 5 and 6 further control for the initial meeting intensity (meetings and days of meetings) between switchers and friends on weekends and holidays from January to June 2018. Standard errors are clustered by block pair and reported in parentheses.

**Table 5:** Robustness

| Dependent variable: Probability of switching from $i$ to $j$ | (1) | (2) | (3) |
|---|---|---|---|
| Friend | 0.031*** | 0.002*** | 0.020*** |
| | (0.001) | (0.001) | (0.001) |
| Friend*EverMeet | 0.026*** | 0.015*** | 0.024*** |
| | (0.001) | (0.001) | (0.001) |
| Number of workers at j | | 0.0003*** | |
| | | (3.21e-06) | |
| Number of friends | | | 0.010*** |
| | | | (0.001) |
| Observations | 4,787,639 | 4,787,639 | 4,787,639 |
| R-squared | 0.169 | 0.180 | 0.169 |
| Whether friend has same gender, migrant status and income group | Yes | No | No |
| Whether cells in same cluster | Yes | No | No |
| Block i FE | Yes | Yes | Yes |
| Block j FE | Yes | Yes | Yes |
| Block i*Block j FE | Yes | Yes | Yes |

*Notes*: This table conducts robustness checks of the baseline result in Table 4. The unit of observation is a $ij$ cell pair. Columns 1 includes for homophily controls: (1) three dummy variables indicating whether the dominant gender, migrant status, income group of switchers in cell $i$ are the same with that among friends in cell $j$; (2) a dummy variable indicating whether switchers in cell $i$ are in the same group as their friends in cell $j$ using k-means clustering algorithm. K-means clustering algrithm is performed based on a vector of characteristics, including whether the dominant gender is female, male, migrant, local, and in each of the eight income group bins. Columns 2 controls for the number of work population in cell $j$. Columns 3 controls for the number of friends of switchers in cell $j$. All columns include origin block fixed effects, destination block fixed effects and origin-by-destination block-pair fixed effects. Standard errors are clustered by block pair and reported in parentheses.

**Table 6:** Placebo Tests

| Dependent variable: Probability of switching from $i$ to $j$ | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Friend | 0.043*** | 0.043*** | 0.043*** | 0.043*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Friend*EverMeet | 0.216*** | | | 0.209*** |
| | (0.045) | | | (0.045) |
| Friend*EverMeet (left) | | 0.139 | | 0.130 |
| | | (0.119) | | (0.122) |
| Friend*EverMeet (will join later) | | | 0.148** | 0.117* |
| | | | (0.058) | (0.061) |
| Observations | 1,253,932 | 1,253,932 | 1,253,932 | 1,253,932 |
| R-squared | 0.159 | 0.159 | 0.159 | 0.159 |
| Block i FE | Yes | Yes | Yes | Yes |
| Block j FE | Yes | Yes | Yes | Yes |
| Block i*Block j FE | Yes | Yes | Yes | Yes |

*Notes*: This table performs placebo tests of the baseline result in Table 4. The unit of observation is a $ij$ cell pair. The dependent variable is the probability that switchers change jobs from cell $i$ to cell $j$ during July-September 2019. "EverMeet", "EverMeet(left)" and "EverMeet(will join later)" are dummies indicating whether there were face-to-face meetings during April-June 2019 with different sets of friends: (1) "EverMeet" considers meetings with friends who did not change jobs in 2019; (2) "EverMeet (left)" considers meetings with friends who previously worked in cell $j$ in April-June 2018 but left before April 2019; (3) "EverMeet (will join later)" considers meetings with friends who previously worked elsewhere since July 2018 but switched to cell $j$ after September 2019. All columns include origin and destination block fixed effects, and origin-by-destination block-pair fixed effects. Standard errors are clustered by block pair and reported in parentheses.

**Table 7:** F2F and Other Communication Methods

| Dependent variable: Probability of switching from $i$ to $j$ | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Friend | 0.049*** | 0.051*** | 0.039*** | 0.038*** | 0.040*** |
| | (0.001) | (0.001) | (0.002) | (0.002) | (0.006) |
| Friend*std_F2F meetings (x10) | 0.019*** | | | 0.010*** | |
| | (0.003) | | | (0.003) | |
| Friend*std_calls (x10) | | 0.003* | | 0.002 | |
| | | (0.002) | | (0.002) | |
| Friend*std_WeChat volume (x10) | | | 0.005*** | 0.005*** | |
| | | | (0.0004) | (0.0004) | |
| Friend*EverMeet | | | | | 0.032*** |
| | | | | | (0.004) |
| Friend*EverCall | | | | | 0.008 |
| | | | | | (0.054) |
| Friend*EverWeChat | | | | | 0.006 |
| | | | | | (0.006) |
| Observations | 2,169,538 | 2,169,538 | 2,169,538 | 2,169,538 | 2,169,538 |
| R-squared | 0.165 | 0.165 | 0.165 | 0.165 | 0.165 |
| Block i FE | Yes | Yes | Yes | Yes | Yes |
| Block j FE | Yes | Yes | Yes | Yes | Yes |
| Block i*Block j FE | Yes | Yes | Yes | Yes | Yes |

*Notes*: This table compares the effect of face-to-face meetings, calls and WeChat volume on job changers' location choices. The unit of observation is a $ij$ cell pair. The dependent variable is the probability that switchers change jobs from cell $i$ to cell $j$ during July-December 2019. Face-to-face meetings, calls and WeChat volume are from January to June 2019. Columns 1-4, we interact "Friend" with standardized measures, i.e. minus the mean and divide by the standard deviation, and scale up by 10. In Column 5, we interact "Friend" with nonzero dummies: "EverMeet", "EverCall" are for switchers from cell $i$ ever meeting face-to-face or making calls with their friends in cell $j$, respectively; "EverWechat" for nonzero total WeChat data volume of switchers from cell $i$ who have friends in $j$. All columns include origin block fixed effects, destination block fixed effects and origin-by-destination block-pair fixed effects. Standard errors are clustered by block pair and reported in parentheses.

**Table 8:** Heterogeneous Effects By Worker Characteristics

| Dependent variable: | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Probability of switching from $i$ to $j$ | Female | Male | Migrant | Local | Full-time | Part-time | High-income | Low-income |
| *Panel A: No control* | | | | | | | | |
| Friend*EverMeet | 0.0307*** | 0.0199*** | 0.0380*** | 0.0165*** | 0.0396*** | 0.0127** | 0.0346*** | 0.0229*** |
| | (0.00247) | (0.00473) | (0.00250) | (0.00433) | (0.00218) | (0.00520) | (0.00228) | (0.00475) |
| Diff | | -0.0108 | | -0.0215 | | -0.0269 | | -0.0117 |
| t-stats | | -2.031 | | -4.293 | | -4.776 | | -2.222 |
| p-value | | 0.042 | | 0.000 | | 0.000 | | 0.026 |
| *Panel B: Control for number of meetings* | | | | | | | | |
| Friend*EverMeet | 0.0243*** | 0.0161*** | 0.0260*** | 0.0126** | 0.0253*** | 0.0136** | 0.0225*** | 0.0197*** |
| | (0.00293) | (0.00577) | (0.00298) | (0.00549) | (0.00258) | (0.00641) | (0.00272) | (0.00594) |
| Diff | | -0.00816 | | -0.0134 | | -0.0117 | | -0.00273 |
| t-stats | | -1.260 | | -2.148 | | -1.697 | | -0.417 |
| p-value | | 0.208 | | 0.032 | | 0.090 | | 0.676 |
| *Panel C: Control for number of meetings AND meeting composition* | | | | | | | | |
| Friend*EverMeet | 0.0177*** | 0.0156 | 0.0199*** | 0.0120 | 0.0119*** | 0.00440 | 0.0186*** | 0.000661 |
| | (0.00500) | (0.0101) | (0.00502) | (0.00955) | (0.00441) | (0.0113) | (0.00445) | (0.0112) |
| Diff | | -0.00208 | | -0.00788 | | -0.00753 | | -0.0180 |
| t-stats | | -0.184 | | -0.730 | | -0.620 | | -1.493 |
| p-value | | 0.854 | | 0.465 | | 0.535 | | 0.136 |
| Observations | 1,193,872 | 1,196,716 | 1,107,900 | 1,208,641 | 1,138,143 | 1,194,109 | 1,179,683 | 1,195,088 |

*Notes*: The unit of observation is a $ij$ cell pair. Each column restricts to the subsample where the majority of switchers in the origin cell share the same characteristics, with the proportion of switchers in that characteristics exceeds the 75th percentile of the distribution. Panel A employs the preferred specification as in Column 3 of Table 4. Panel B adds number of meetings as an additional control. Panel C further controls for meeting composition, including dummies indicating whether the majority of meetings of switchers in a cell $i$ are with friends in cell $j$ who are female, migrant, low-income friends and friends of different social circles. In all panels, we report the difference, t-test statistics and p-value between the coefficients in the even-numbered column and the odd-numbered column before it. All columns include origin block fixed effects, destination block fixed effects and origin-by-destination block-pair fixed effects. Standard errors are clustered by block pair and reported in parentheses.

**Table 9:** Heterogeneous Effects By Sector

| Dependent variable: Probability of switching from $i$ to $j$ | (1) Public | (2) Private | (3) IT and Software | (4) Not IT and Software | (5) Finance | (6) Not Finance | (7) Manufacturing | (8) Not Manufacturing |
|---|---|---|---|---|---|---|---|---|
| *Panel A: No control* | | | | | | | | |
| Friend*EverMeet | 0.0176** | 0.0342*** | 0.0397*** | 0.0253*** | 0.0344*** | 0.0259*** | 0.0420*** | 0.0285*** |
| | (0.00880) | (0.00292) | (0.00198) | (0.00168) | (0.00159) | (0.00221) | (0.00274) | (0.00144) |
| Diff | | 0.0167 | | -0.0144 | | -0.00850 | | -0.0135 |
| t-stats | | 1.802 | | -5.554 | | -3.127 | | -4.364 |
| p-value | | 0.0716 | | 2.79e-08 | | 0.00177 | | 1.28e-05 |
| *Panel B: Control for number of meetings* | | | | | | | | |
| Friend*EverMeet | 0.0148 | 0.0231*** | 0.0295*** | 0.0174*** | 0.0245*** | 0.0190*** | 0.0341*** | 0.0192*** |
| | (0.0107) | (0.00343) | (0.00230) | (0.00207) | (0.00186) | (0.00274) | (0.00328) | (0.00173) |
| Diff | | 0.00829 | | -0.0121 | | -0.00551 | | -0.0149 |
| t-stats | | 0.736 | | -3.918 | | -1.663 | | -4.004 |
| p-value | | 0.462 | | 8.92e-05 | | 0.0964 | | 6.22e-05 |
| *Panel C: Control for number of meetings AND meeting composition* | | | | | | | | |
| Friend*EverMeet | 0.0131 | 0.0165*** | 0.0209*** | 0.0165*** | 0.0181*** | 0.0187*** | 0.0216*** | 0.0172*** |
| | (0.0169) | (0.00581) | (0.00387) | (0.00357) | (0.00316) | (0.00478) | (0.00553) | (0.00297) |
| Diff | | 0.00346 | | -0.00439 | | 0.000548 | | -0.00438 |
| t-stats | | 0.193 | | -0.834 | | 0.0956 | | -0.697 |
| p-value | | 0.847 | | 0.404 | | 0.924 | | 0.486 |
| Observations | 251,125 | 685,639 | 936,464 | 3,283,597 | 1,445,527 | 2,774,545 | 484,105 | 3,736,356 |

*Notes*: The unit of observation is a $ij$ cell pair. Each column restricts to the subsample where the majority of work locations in the origin cell belongs to the same sector, with the proportion of work locations in that sector exceeds the 75th percentile of the distribution. "Public" sector includes industries "Public Management, Social Security and Social Organizations", "Health and Social Work", "School/Science Sites" under "Education", "Water Conservancy, Environment and Public Facilities Management", "Scientific Research and Technical Services", "Culture Sites", and "Scenic area"; "Private" sector includes firms and enterprises in all other industries not classified as public sector. The three panels use the same specification as in Table 8. Standard errors are clustered by block pair and reported in parentheses.

**Table 10:** Heterogeneity By Worker Characteristics: Frequency of F2F meetings

| Dependent variable: Number of meetings between cells $i$ and $j$ | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| *Whether the origin cell is* | | | | |
| Female | 0.149*** | | | |
| | (0.051) | | | |
| Migrant | | 0.429*** | | |
| | | (0.098) | | |
| Full-time | | | 0.867*** | |
| | | | (0.100) | |
| High-income | | | | 0.312*** |
| | | | | (0.071) |
| Constant | 1.817*** | 2.009*** | 1.994*** | 2.596*** |
| | (0.026) | (0.047) | (0.049) | (0.035) |
| Observations | 2,390,981 | 2,316,806 | 2,332,537 | 2,374,962 |
| R-squared | 0.256 | 0.259 | 0.272 | 0.298 |
| block i | Yes | Yes | Yes | Yes |
| block j | Yes | Yes | Yes | Yes |
| block i*block j | Yes | Yes | Yes | Yes |

*Notes*: The unit of observation is a $ij$ cell pair. Columns 1-4 limit to the subsample where the majority of switchers in the origin cell are of certain characteristics. "Female" is a dummy variable equals one if the fraction of female switchers origin cell is above the 75th percentile of its distribution, and zero if the share of male switchers is above the 75th percentile of the distribution. Column 1 excludes observations where the gender characteristics of the origin cell is in the middle range. Similarly, "Migrant" dummy indicates the majority of the switchers in the origin cell are migrants, as opposed to local workers; "Full-time" dummy indicates the majority of the switchers in the origin cell are full-time, as opposed to part-time workers; "High-income" dummy indicates the majority of the switchers in the origin cell are high-income, as opposed to low-income workers. All columns include origin block fixed effects, destination block fixed effects and origin-by-destination block-pair fixed effects. Standard errors are clustered by block pair and reported in parentheses.

**Table 11:** Heterogeneity By Worker Characteristics: Meeting Composition

| Dependent variable: | | | | Individuals of social circle | |
|---|---|---|---|---|---|
| Majority of the meetings with | Female | Migrant | Low-income | 50-100 | above 100 |
| *Whether the origin cell is* | | | | | |
| Female | 0.0006*** | 0.0007*** | 0.0004*** | 0.0003*** | 0.0001* |
| | (0.0001) | (0.0002) | (0.0002) | (0.0001) | (0.0001) |
| Full-time | 0.0021*** | 0.0035*** | 0.0021*** | 0.0017*** | 0.0012*** |
| | (0.0003) | (0.0004) | (0.0003) | (0.0002) | (0.0002) |
| Migrant | 0.0009*** | 0.0015*** | 0.0013*** | 0.0008*** | 0.0005*** |
| | (0.0002) | (0.0003) | (0.0002) | (0.0002) | (0.0001) |
| High-income | 0.0003** | 0.0006*** | 0.0003 | 0.0004*** | -2.38e-06 |
| | (0.0002) | (0.0002) | (0.0002) | (0.0001) | (0.0001) |
| block i | Yes | Yes | Yes | Yes | Yes |
| block j | Yes | Yes | Yes | Yes | Yes |
| block i*block j | Yes | Yes | Yes | Yes | Yes |

*Notes*: The unit of observation is a $ij$ cell pair. The dependent variables are dummy variables indicating whether the majority of meetings between switchers in cell $i$ and friends in cell $j$ are with female, migrant, low-income friends and friends of middle and large social circle, respectively. Social circle are measured by monthly average number of call contacts: middle if on average have 50-100 unique call contacts, large if above 100 call contacts each month. Each coefficient is obtained from a separate regression. The regressor in each row exclude observations where the characteristics of the majority of switchers in the origin cell are not clear. "Female" is a dummy variable equals one if the fraction of female switchers origin cell is above the 75th percentile of its distribution, and zero if the share of male switchers is above the 75th percentile of the distribution. Column 1 exclude observations where the gender characteristics of the origin cell is in the middle range. Similarly, "Migrant" dummy indicates the majority of the switchers in the origin cell are migrants, as opposed to local workers; "Full-time" dummy indicates the majority of the switchers in the origin cell are full-time, as opposed to part-time workers; "High-income" dummy indicates the majority of the switchers in the origin cell are high-income, as opposed to low-income workers. All columns include origin block fixed effects, destination block fixed effects and origin-by-destination block-pair fixed effects. Standard errors are clustered by block pair and reported in parentheses.

**Table 12:** Heterogeneity By Sector: Frequency of F2F meetings

| Dependent variable: Number of meetings between cells $i$ and $j$ | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *Whether the origin cell is* | | | | | |
| Public | -0.267 | | | | |
| | (0.227) | | | | |
| IT and Software | | 1.917*** | | | |
| | | (0.190) | | | |
| Finance | | | 1.892*** | | |
| | | | (0.146) | | |
| Manufacturing | | | | | 1.439*** |
| | | | | | (0.259) |
| Constant | 3.200*** | 2.785*** | 2.562*** | 2.673*** | 3.045*** |
| | (0.0608) | (0.0421) | (0.0499) | (0.0522) | (0.0297) |
| Observations | 936,810 | 4,221,358 | 4,221,358 | 4,221,358 | 4,221,358 |
| R-squared | 0.257 | 0.247 | 0.248 | 0.247 | 0.247 |
| Block i | Yes | Yes | Yes | Yes | Yes |
| Block j | Yes | Yes | Yes | Yes | Yes |
| Block i*Block j | Yes | Yes | Yes | Yes | Yes |

*Notes*: The unit of observation is a $ij$ cell pair. Each limits to the subsample where the majority of work locations in the origin cell are of certain sector. "Public" is a dummy variable equals one if the fraction of work locations in the origin cell is above the 75th percentile of its distribution, and zero if the share of private work locations is above the 75th percentile of the distribution. That is, Column 1 excludes observations where the share of public or private work locations is in the middle range. Similarly, "IT and Software" dummy indicates the majority of the work locations in the origin cell belongs to IT and software industry, and it takes value zero when the majority of work locations belongs to other industry. "Finance", and "Manufacturing" dummies indicate the majority of the work locations in the origin cell belongs to finance, and manufacturing industry, respectively. All columns include origin block fixed effects, destination block fixed effects and origin-by-destination block-pair fixed effects. Standard errors are clustered by block pair and reported in parentheses.

**Table 13:** Heterogeneity By Sector: Meeting Composition

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dependent variable: | | | | Individuals of social circle | |
| Majority of the meetings with | Female | Migrant | Low-income | 50-100 | above 100 |
| *Whether the origin cell is* | | | | | |
| Public | -0.0005* | -0.0013** | -0.0008* | -0.0009** | -0.0001 |
| | (0.0003) | (0.0006) | (0.0004) | (0.0004) | (0.0003) |
| IT and Software | 0.0039*** | 0.0049*** | 0.0033*** | 0.00304*** | 0.0021*** |
| | (0.0006) | (0.0007) | (0.0006) | (0.0005) | (0.0004) |
| Finance | 0.0039*** | 0.0056*** | 0.0041*** | 0.0035*** | 0.0023*** |
| | (0.0004) | (0.0005) | (0.0004) | (0.0003) | (0.0003) |
| Manufacturing | 0.0023*** | 0.0038*** | 0.0033*** | 0.00187*** | 0.0023*** |
| | (0.0008) | (0.0008) | (0.0008) | (0.0007) | (0.0005) |
| Block i | Yes | Yes | Yes | Yes | Yes |
| Block j | Yes | Yes | Yes | Yes | Yes |
| Block i*block j | Yes | Yes | Yes | Yes | Yes |

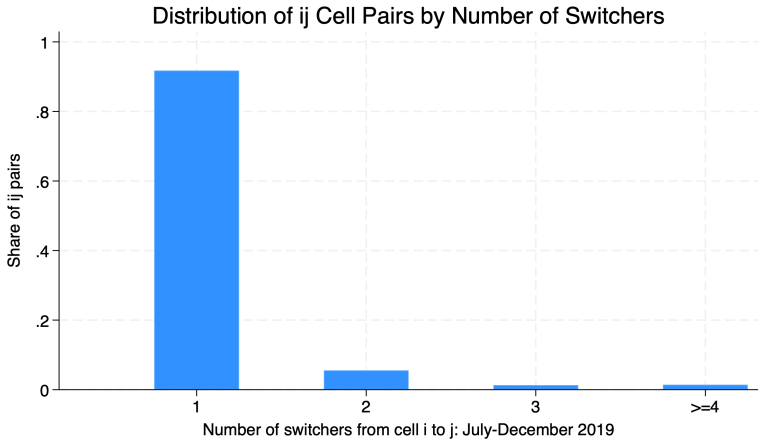
*Notes*: The unit of observation is a $ij$ cell pair. The dependent variables are dummy variables indicating whether the majority of meetings between switchers in cell $i$ and friends in cell $j$ are with female, migrant, low-income friends and friends of middle and large social circle, respectively. Social circle are measured by monthly average number of call contacts: middle if on average have 50-100 unique call contacts, large if above 100 call contacts each month. Each coefficient is obtained from a separate regression. The regressor in each row exclude observations where the sectors of the majority of work locations in the origin cell are not clear. "Public" is a dummy variable equals one if the fraction of work locations in the origin cell is above the 75th percentile of its distribution, and zero if the share of private work locations is above the 75th percentile of the distribution. "IT and Software" dummy indicates the majority of the work locations in the origin cell belongs to IT and software industry, and it takes value zero when the majority of work locations belongs to other industry. "Finance", and "Manufacturing" dummies indicate the majority of the work locations in the origin cell belongs to finance, and manufacturing industry, respectively. All columns include origin block fixed effects, destination block fixed effects and origin-by-destination block-pair fixed effects. Standard errors are clustered by block pair and reported in parentheses.

# Online Appendices

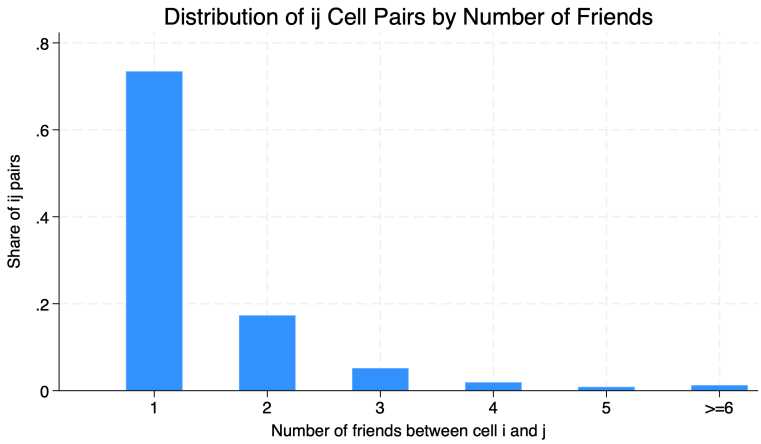
## A Supplementary Figures and Tables

**Figure A1:** Distribution of area pairs

**(a)** By Number of Switchers from $i$ to $j$

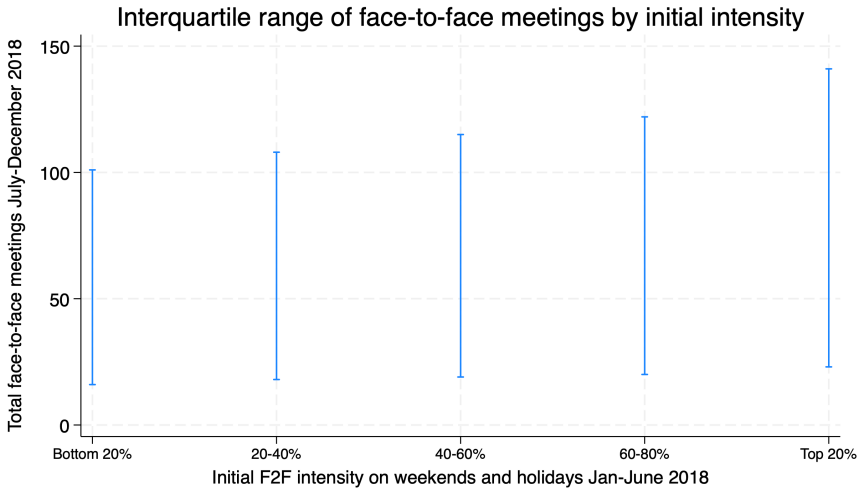



**(b)** By Number of Friends between $i$ and $j$




*Notes*: This figure plots the distribution of $ij$ area pairs by number of switchers from $i$ to $j$ and number of friends between $i$ and $j$, respectively. 4,582,775 out of 4,810,573 $ij$ pairs in the regression sample have zero switchers (95%), 227,798 have positive number of switchers (5%). Figures A1a plots the distribution of $ij$ area pairs among 227,798 pairs with positive number of switchers, 91.75% of them have only 1 switchers, 5.5% have 2 switchers, 1.3% have 3 switchers and 1.42% have 4 or more switchers. 4,366,320 out of 4,810,573 $ij$ pairs in the regression sample have zero friends (91%), 444,253 have positive number of friends (9%). Figures A1b plots the distribution of $ij$ area pairs among these 444,253 pairs, 73.45% of them have only 1 friend, 17.33% have 2 friends, 5.19% have 3 friends and 4% have 4 or more friends.

**Figure A2:** Interquartile range of face-to-face meetings by initial intensity




*Notes*: This figure shows the interquartile range of total face-to-face meetings between switchers from $i$ and their friends in $j$ from July to December 2018, grouped by the initial number of meetings that occurred during weekends and holidays from January to June 2018. The sample is restricted to $ij$ cell pairs with at least one face-to-face meeting during July to December 2018. The x-axis divides $ij$ pairs into five equal-sized bins based on the initial number of meetings. The interquartile range is represented by the distance between the 25th and 75th percentiles of the distribution of total face-to-face meetings within each bin.

**Table A1:** Phone calls with friends and non-friends

| Dep. Var. | (1) Average number of calls per month | (2) Average number of calls per month: weekends | (3) Average number of calls per month: weekdays |
|---|---|---|---|
| Friend | 4.995*** | 2.729*** | 2.265*** |
| | (0.057) | (0.057) | (0.046) |
| Non-friend | 2.921*** | 1.374*** | 1.547*** |
| | (0.044) | (0.041) | (0.032) |
| Observations | 12,000 | 12,000 | 12,000 |
| R-squared | 0.515 | 0.241 | 0.288 |

*Notes*: One unit of observation is a pair of two individuals. The regression sample consists of random samples of 6000 pairs for friend and non-friend relationship, respectively. The dependent variable is the average monthly calls between each individual pair during Jan-June 2019 in Column 1. It is the average monthly calls during weekends and weekdays in Columns 2 and 3.

**Table A2:** Effect of F2F on Job Location Choices: An Alternative Dependent Variable

| Dependent variable: Number of switchers from $i$ to $j$ | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Friend | 0.156*** | 0.049*** | 0.037*** | 0.037*** | 0.012*** | 0.013*** |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Friend*EverMeet | | | 0.047*** | 0.046*** | 0.043*** | |
| | | | (0.002) | (0.002) | (0.002) | |
| F2f meetings: Low | | | | | | 0.033*** |
| | | | | | | (0.002) |
| F2f meetings: Medium | | | | | | 0.036*** |
| | | | | | | (0.003) |
| F2f meetings: High | | | | | | 0.064*** |
| | | | | | | (0.003) |
| Cell $i$ total F2F (x1000) | | | | 0.000002 | -0.0001*** | -0.0001*** |
| | | | | (0.00001) | (0.00001) | (0.00001) |
| Observations | 4,805,218 | 4,787,639 | 4,787,639 | 4,787,639 | 4,787,639 | 4,787,639 |
| R-squared | 0.030 | 0.176 | 0.177 | 0.177 | 0.177 | 0.178 |
| Initial intensity control | No | No | No | No | Yes | Yes |
| Block i FE | No | Yes | Yes | Yes | Yes | Yes |
| Block j FE | No | Yes | Yes | Yes | Yes | Yes |
| Block i x block j FE | No | Yes | Yes | Yes | Yes | Yes |

*Notes*: This table examines the effect of face-to-face meetings on job changers' location choices as in Table 4 but with an alternative dependent variable. The unit of observation is a $ij$ cell pair. The dependent variable is the number of switchers changing jobs from cell $i$ to cell $j$.

**Table A3:** Robustness: An Alternative Dependent Variable

| Dependent variable: Number of switchers from $i$ to $j$ | (1) | (2) | (3) |
|---|---|---|---|
| Friend | 0.037*** | 0.003*** | 0.016*** |
| | (0.001) | (0.001) | (0.003) |
| Friend*EverMeet | 0.037*** | 0.027*** | 0.033*** |
| | (0.002) | (0.002) | (0.002) |
| Number of workers at j | | 0.0004*** | |
| | | (4.02e-06) | |
| Number of friends | | | 0.019*** |
| | | | (0.002) |
| Observations | 4,787,639 | 4,787,639 | 4,787,639 |
| R-squared | 0.177 | 0.188 | 0.178 |
| Whether friend has same gender, migrant status and income group | Yes | No | No |
| Whether cells in same cluster | Yes | No | No |
| Block i FE | Yes | Yes | Yes |
| Block j FE | Yes | Yes | Yes |
| Block i*Block j FE | Yes | Yes | Yes |

*Notes*: This table conducts robustness checks of the baseline result in Table 4 and uses the same specifications as in Table 5 but with alternative dependent variable. The dependent variable is the number of switchers changing jobs from cell $i$ to cell $j$.

**Table A4:** Placebo Tests: An Alternative Dependent Variable

| Dependent variable: Number of switchers from $i$ to $j$ | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Friend | 0.052*** | 0.052*** | 0.052*** | 0.051*** |
| | (0.003) | (0.003) | (0.003) | (0.002) |
| Friend*EverMeet | 0.331*** | | | 0.319*** |
| | (0.069) | | | (0.0690) |
| Friend*EverMeet (left) | | 0.190 | | 0.176 |
| | | (0.160) | | (0.166) |
| Friend*EverMeet (will join later) | | | 0.225** | 0.179* |
| | | | (0.092) | (0.091) |
| Observations | 1,253,932 | 1,253,932 | 1,253,932 | 1,253,932 |
| R-squared | 0.170 | 0.170 | 0.170 | 0.170 |
| Block i FE | Yes | Yes | Yes | Yes |
| Block j FE | Yes | Yes | Yes | Yes |
| Block i*Block j FE | Yes | Yes | Yes | Yes |

*Notes*: This table performs placebo tests of the baseline result in Table 4 but with an alternative dependent variable. The dependent variable is the number of switchers moving from $i$ to $j$ during July-September 2019. Regressors "EverMeet", "EverMeet(left)" and "EverMeet(will join later)" and the specifications are the same as in Table 6.

**Table A5:** Heterogeneous Effects By Specific Industries

| Dependent variable: | (5) | (6) | (9) | (10) |
|---|---|---|---|---|
| | | Not | | Not |
| Probability of switching from $i$ to $j$ | Service | Service | Wholesale and retail | Wholesale and retail |
| Friend | 0.033*** | 0.031*** | 0.035*** | 0.030*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| EverMeet | 0.032*** | 0.031*** | 0.030*** | 0.032*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Observations | 2,117,700 | 2,102,288 | 1,463,394 | 2,756,522 |
| R-squared | 0.166 | 0.160 | 0.161 | 0.165 |
| Block i FE | Yes | Yes | Yes | Yes |
| Block j FE | Yes | Yes | Yes | Yes |
| Block i*Block j FE | Yes | Yes | Yes | Yes |
| Diff | | -0.0003 | | 0.002 |
| t-stats | | -0.113 | | 0.733 |
| p-value | | 0.910 | | 0.463 |

*Notes*: This table uses the same specification as Columns 3-8 in Table 9 but reports the heterogeneous effects by the other 2 industries contributing to Beijing's GDP in 2019: (1) leasing and business service, and (2) wholesale and retail trade.