Face-to-Face Meetings, Worker Mobility, and Referrals

Eleonora Patacchini* Qi Wu†

Abstract

We study how face-to-face interactions shape worker mobility through social networks. Using granular cellphone geolocation and sociodemographic data on 3.3 million workers in a large urban labor market, we exploit settings in which multiple friends of the same worker relocate to jobs within the same destination area. Within the same worker and destination, face-to-face interaction with a friend increases the probability of moving to that friend's workplace by 36-43 percent relative to remote communication, with effects substantially larger than those of phone calls or digital messaging. Consistent with referral models, the effect emerges only after the friend joins the destination firm and disappears in pre-move placebo periods. A data-driven heterogeneity analysis using regularized machine learning reveals a pronounced targeting gradient: effects in the highest predicted-return settings are more than twice as large as the population average. We interpret these findings through a simple referral model in which interaction technology shapes the effectiveness of information transmission, showing that mobility depends not only on the presence of social ties but on how information flows within them.

Keywords Employment, Communication Technology, Agglomeration Economies, Machine Learning

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.

We thank Costas Arkolakis, Benjamin Goldman, Gordon Hanson, Yanos Zylberberg, and all participants at the CESifo Conference on Big Data in Urban Economics in June 2025 for their valuable comments and suggestions.

1 Introduction

Worker mobility is essential for a dynamic labor market and sustained economic growth (Davis and Haltiwanger, 1992). A well-established finding in the literature is that employee referrals — where current workers help recruit friends or acquaintances — significantly influence mobility patterns (Bayer et al., 2008; Burks et al., 2015; Friebel et al., 2023; Hoffman and Stanton, 2024). Although referral-based communication is a key channel through which workers transition between jobs (Barwick et al., 2023; Dustmann et al., 2016; Topa, 2011; Schmutte, 2016), little is known about how the mode of communication, whether face-to-face or remote, shapes these results.

Canonical referral models (e.g., Montgomery, 1991; Topa, 2001; Calvó-Armengol and Jackson, 2004; Ioannides and Loury, 2004) typically abstract from the nature of interaction within a link, implicitly assuming that information flows costlessly and symmetrically across social ties. In this literature, the mere existence of a connection is sufficient for referrals, job information, or access to jobs.

This paper examines this assumption by asking whether differences in interaction mode within social ties shape worker mobility. Influential agglomeration theories highlight face-to-face contact as central to information exchange and coordination that sustain labor market matching (Jacobs, 1960; Storper and Venables, 2004; Rosenthal and Strange, 2004), yet it remains unclear whether differences in interaction mode shape these processes.[1] Addressing this question has broader implications: it informs theories of agglomeration, the persistence of cities in an era of globalization, and the labor-market consequences of remote work following COVID-19. If co-presence enhances the effectiveness of information transmission, location remains economically valuable not only for production but also for matching workers to jobs.

Providing direct evidence on the role of face-to-face interaction, however, has proven difficult because such encounters are rarely observed. We overcome this constraint using high-frequency cellphone geolocation data covering more than two years in a major urban labor market, Beijing, China. The data capture physical co-location every 15 minutes, independent of app usage, and are complemented by phone-call records, data from mobile chat apps, detailed firm maps, and job postings. This richness allows us to observe pre-existing social ties, track workplace mobility at high frequency, and directly measure interaction mode.

[1] Storper and Venables (2004) outline key theoretical features of face-to-face interaction that may improve labor market matching. First, face-to-face communication provides rich, multidimensional information through visual, auditory, and non-verbal cues, enhancing the effectiveness of communication. 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.

We leverage this rich data environment to construct an empirical design that isolates the role of interaction mode within social ties. Our most demanding specification, which we term the friend-move design, focuses on cases where multiple friends of a worker relocate to jobs within the same destination neighborhood. Holding worker, destination, and friendship fixed, we compare whether workers are more likely to follow friends with whom they interact face-to-face than those with whom they interact only remotely. We find that face-to-face interaction increases the probability of following a friend to a workplace by 36-43 percent relative to remote-only interaction. Because friendships and interaction modes are measured before mobility events, the design mitigates concerns about homophily, sorting, and reverse causality. This pattern is consistent with referral-based information transmission rather than alternative explanations. In particular, the mobility effect of face-to-face interaction emerges only after a friend joins a new workplace, not before. Estimated effects collapse in pre-move placebo periods and are not explained by differential vacancy postings, helping rule out anticipatory search or firm-level hiring shocks. Specifications with friend fixed effects confirm that the estimates are driven by within-friend variation in interaction mode rather than differences across contacts.

The primary advantage of the friend-move design is that it speaks directly to the mechanism. It allows us not only to estimate the effect of face-to-face interactions but also to identify how social contacts shape job matching conditional on mobility choices. To guide interpretation, we develop a parsimonious theoretical model of referral-based mobility that maps naturally into the friend-move design and delivers testable predictions. Our simple model highlights features shared with canonical referral models but rarely tested empirically: referrals operate at the level of specific jobs rather than broad locations, take effect only *after* a friend is employed at a firm, and improve job matching rather than simply shifting overall mobility. The model yields four predictions: (i) face-to-face interaction affects mobility only after a friend moves; (ii) the effect operates at the workplace level rather than broadly across neighborhoods; (iii) conditional on friendship and destination, interaction mode matters; and (iv) friends shape destination choice, while face-to-face interaction shapes workplace choice within the destination. Our empirical results closely match these predictions, supporting the mechanism that in-person interactions facilitate job information transmission and access rather than reflecting general social or locational preferences.

While the friend-move design provides a clean causal benchmark, it imposes demanding data requirements. For heterogeneity analysis and additional evidence, we rely on a broader baseline specification that includes a larger set of job switchers. This specification compares switchers who have friends already working in different workplaces within the destination neighborhood, without requiring that the friend move before the focal worker. The estimates

from this broader design are quantitatively similar, indicating that the role of face-to-face interaction reflects a general mechanism within destination areas rather than a feature of the restricted design. Face-to-face effects remain sizable under extensive robustness checks and dominate other communication modes: when estimated jointly with phone calls and WeChat, in-person meetings have nearly twice the impact of WeChat, while calls are statistically insignificant.

The effects of face-to-face interaction are not uniform across workers: the benefits of in-person contact are more pronounced for migrants, full-time employees, and high-income earners. To examine additional dimensions of heterogeneity in the returns to face-to-face interaction beyond pre-specified categories, we complement our theory-guided analysis with a data-driven approach. Using a regularized machine-learning approach (LASSO), we identify moderators of treatment heterogeneity in a training sample and predict out-of-sample effects, ranking settings by expected responsiveness to face-to-face interaction.

This exercise serves several purposes. Most importantly, it evaluates whether the observed heterogeneity is systematic rather than sample-specific by testing whether predicted returns materialize in holdout data. It also quantifies the gains from targeting, asking how much larger mobility responses are in high predicted-return environments relative to the population average. We find a pronounced targeting gradient: treatment effects in the highest predicted-return settings are more than twice as large as in the overall population.

Importantly, the characteristics that predict large returns are predominantly network-based — such as meeting intensity and the connectedness and income of contacts — rather than features specific to particular firms or locations. Moreover, the targeting gradient remains stable across alternative training-testing partitions defined along economically meaningful dimensions. Together, these results suggest that the heterogeneity uncovered by the model reflects structural features of social interaction and job matching rather than idiosyncratic features of particular institutional or geographic contexts, with direct implications for where in-person interaction is most valuable.

Our paper closely relates to two key areas of research: labor economics, specifically the literature on referrals (Topa, 2011; Schmutte, 2015; Hoffman and Stanton, 2024), 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.[2]

[2]A notable exception is Barwick et al. (2023), who estimate referral effects by identifying social connections through actual interaction patterns, specifically using data on cellphone calls. Their work represents an

Closely related is Atkin et al. (2022), which shows that chance face-to-face encounters between workers at different firms in Silicon Valley increase inter-firm patent citations. Similar to their approach, we measure face-to-face interactions using cellphone co-location data based on location pings. However, our dataset differs in a key way: whereas Atkin et al. (2022) relies on smartphone pings triggered by open or background apps, our data provider captures the location of all cellphones every 15 minutes – independent of app usage or phone activity. This leads to a much more precise and comprehensive digital map of physical proximity. Moreover, our dataset covers a longer period, allowing us to identify pre-existing social ties. It also offers greater depth, enabling us to explore heterogeneity in the effects of in-person meetings by individual and job characteristics, as well as by the frequency and nature of face-to-face meetings. Our findings complement those of Atkin et al. (2022) in meaningful ways. While their study shows that face-to-face interactions facilitate knowledge spillovers – measured through increased inter-firm patent citations – we find that such interactions also play a critical role in shaping worker mobility. Together, these results highlight two distinct yet interconnected channels through which face-to-face contact contributes to agglomeration externalities. On one hand, in-person interactions enhance the flow of ideas and innovation; on the other, they help workers access job opportunities through social connections.

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., 2022a,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; Storper and Venables, 2004; Rosenthal and Strange, 2004), their empirical significance remains an underexplored area. By providing the first evidence of their role in worker mobility, we take a crucial step in advancing this line of research.

The remainder of the paper proceeds as follows. Section 2 describes the data. Section 3 presents definitions and descriptive evidence. Section 4 introduces the empirical strategy and theoretical framework. Section 5 provides robustness checks. Section 6 presents additional evidence including comparing face-to-face interactions with other communication methods, effect heterogeneity across worker types, and Chernuzhukov LASSO-based approach to perform a targeting exercise and investigate external validity. Section 7 concludes.

important step forward by moving beyond inferred or self-reported relationships. However, this paper does not provide evidence that the mode of communication, such as mobile phone calls, in-person meetings, or online interactions, plays a meaningful role in shaping referral effects.

2 Data

2.1 Cellphone Location Data and Worker Flow

Our primary data source is cellphone location data from a major Chinese telecommunications provider. The dataset covers time-stamped geolocation tracking for 3.3 million users in Beijing from January 2018 to December 2019. For each user, we observe demographic characteristics, including gender, age, birth county, and a wealth index.[3]

To precisely identify workplaces, we partition the city into 250-by-250-meter grid cells and overlay this grid with administrative block boundaries and a detailed land-use map. This procedure allows us to isolate cells that are fully covered by establishments — specifically firms and civil service offices — thereby mapping physical space to economically meaningful workplaces.[4] Our final sample consists of 19,190 workplace cells located across 1,860 city blocks.[5]

User locations are recorded either when individuals actively engage with the network (e.g., calls, messages, internet use) or, when inactive, through passive pings at 15-minute intervals. We define a user's monthly primary workplace as the cell in which they spend the most time between 9 a.m. and 5 p.m. on weekdays. Individuals present in the same workplace cell for more than four weeks, or at least two consecutive months with at least 15 days of presence per month, are classified as full-time workers; those present for at least two consecutive months but fewer than 15 days per month are classified as part-time workers. 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 cells in the past 12 months, and the distance between each old and new workplace must be at least 1 km.[6] In addition, we require the job switchers to be full-time workers in the new workplace.

[3]The wealth index is constructed by the provider using information on handset price, subscription fees, residential property values, travel mode (e.g., air travel, high-speed rail), and visits to high-expenditure locations. It ranges from 1 (lowest) to 8 (highest). We classify groups 1-5 as low-income and 6-8 as high-income.

[4]These establishments relate to 19 industries based on industrial classification for national economic activities from National Bureau of Statistics of China 2019 including Agriculture, forestry, animal husbandry and fishery; Mining; Manufacturing; Utilities; Construction; Wholesale and retail trades; Transportation, storage and post; Accommodation and restaurants; Finance; Real estate; Information transmission, software and information technology services; Leasing and business services; Scientific research and technical services; Water conservancy, environment and public facilities management; Residential services and repairs; Education; Health and social work; Culture, sports and Entertainment; and Public administration, social security and social organizations.

[5]Nearly all of these cells lie in Beijing's six inner-city districts, which account for approximately 75-80 percent of the city's GDP and 60-65 percent of its working population.

[6]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. None of the firms in our sample occupies more than 1 square km.

Applying these definitions, over 12 months, we identify 272,937 job switchers from a total working population of 3,295,737. Table 1 presents the number of monthly job switchers and switching rates for January to December 2019. Column 1 reports the number of job switchers per month, Column 2 shows each month's share of total annual job switchers, and Column 3 shows the monthly switching rate, calculated as the number of switchers divided by the total working population. As expected, most job changes occur in March and September, aligning with typical post-Chinese New Year transitions and hiring cycles. The annual job switching rate is approximately 7.3 percent, consistent with the 7-8 percent range reported in other studies using different data from China (Nie and Sousa-Poza, 2017; Barwick et al., 2023).

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 they spend at least 15 minutes in c during time-period h. A *meeting* is defined as two workers who work in different cells both being present in the same cell c at the same time h. We calculate the total number of meetings between workers in cells i and j by multiplying the number of workers from cell i who visit cell c at time h by the number of workers from cell j who are present in the same cell at the same time. We then sum this product over all combinations of locations c and times h:

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

where e_{ich} is the number of workers in cell i who visit cell c at time h.[7] To mitigate the measurement error of chance meetings, we implement several data restrictions. First, we exclude meetings in cells that correspond to large population densities, such as shopping centers, hotels, hospitals, train stations, etc. Second, we remove meetings involving delivery workers[8]. The resulting meeting distribution shows that 2.8% occur in the switcher's work cell, 2.9% in friends' work cell, and 94.3% in other places.

[7]This procedure follows closely Atkin et al. (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.

[8]Delivery workers are identified by the data provider based on usage patterns of apps related to express delivery and food delivery riders, daily location trajectories, and call patterns.

2.3 Other Communication Channels

We merge the face-to-face meeting data with information on other communication methods. The telecommunication company provided complete records of phone calls between all users from January to June 2019. Additionally, we obtained monthly data on WeChat app usage and total data consumption for each user, covering the period from July 2018 to December 2019. WeChat is the dominant networking application in China with 1.08-1.16 billion monthly active users in 2018 and 2019 – approximately 80 percent of China's total population.[9] In compliance with privacy regulations, our WeChat data are limited to aggregate app-level usage volumes – we cannot collect activity within apps (e.g., message content or dyadic interactions) without additional user and platform consent.

2.4 Vacancy Data

To gauge the dynamics of local labor market conditions, we collect job postings from Zhilian (zhaopin.com), a leading online job posting and recruitment platform in China.[10] Each posting reports the job title, job description, job category, required years of experience, required education level, salary range, number of vacant positions, firm name, and address. We obtained 1,270,955 job postings in Beijing from January 2019 to December 2019 and merged them with our main dataset based on locations.

We classify the occupation of each posting according to the job title, job description, and job category into one of 18 occupation categories based on the U.S. Census 2020 occupational classification system.[11] Appendix Section A shows detailed occupation classifications. The top occupations demanded in Beijing 2019 include management (37.51%), business and finance operations (20.81%), computer and mathematical (8.58%), architecture and engineering (7.06%) and service (5.88%).[12] The occupation of a given location is the most common occupation among all postings.[13]

[9] WeChat Year in Review 2018 January 14, 2019 By CIW Team: https://www.chinainternetwatch.com

[10] Zhilian is one of the top three online job posting and hiring platform in China, with a market share of about 20 percent in 2019. Source: https://www.analysys.cn/article/detail/20019506

[11] https://www.bls.gov/oes/current/oes_stru.htm

[12] The remaining occupations include: sales and related (5.73%), arts, design, entertainment, sports, and media (4.49%), educational instruction and library (3.27%), office and administrative support (2.42%), life, physical, and social science (2.23%), healthcare practitioners and technical (0.95%), legal (0.39%), transportation and material moving (0.37%), construction and extraction (0.16%), production (0.05%), installation, maintenance, and repair (0.04%), community and social service (0.03%), farming, fishing, and forestry (0.02%).

[13] It is coded as missing if the most common occupation accounts for less than half of all postings at the same location. On average, the most common occupations account for roughly 80% of all postings.

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 friends. We then compare our sample of switchers with administrative data to gain confidence in our data. Finally, we describe meeting patterns of job switchers with referrer and nonreferrer friends, highlighting 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 encounters. To address this issue, we leverage the depth of our longitudinal data, which tracks individuals continuously over an extended period. We define "friends" based on pre-existing relationships before a job change, as well as on frequent and consistent patterns of interaction 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, before 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.

To minimize the risk of misclassifying other types of relationships as friendships, we carefully exclude meetings likely to involve family members and coworkers, or family visits. Specifically, we exclude: (1) meetings occurring between 12:00 a.m. and 8:00 a.m.; (2) meetings between individuals who reside or work in the same cell for more than ten days in a given month;[14] (3) meetings held at an individual's residential location on more than twenty days in a given month; and (4) high-frequency weekend visits (top 1 percent of meetings) between individuals with an age difference greater than twenty years. We define two individuals as friends if they meet face-to-face on more than three separate days and at least sixteen times over six months. These thresholds correspond to the 90th percentile

[14]By excluding coworkers from this definition, we ensure that our results aren't simply capturing cases where an entire firm relocates – bringing its employees along – rather than genuine face-to-face interactions influencing individual job changes.

of the respective distributions. In other words, we identify friends as those who interact significantly more frequently than average – specifically, individuals in the top 10 percent in terms of both the number of days and the total time spent in face-to-face interactions. 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. We validate our friend definition using information from phone calls and official statistics. First, we examine the number of phone calls between friends and non-friends, as defined above. If our definition captures friendship ties well, 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 B1 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 per month on weekends and weekdays, 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. The differences are statistically significant. These findings provide evidence that our proposed friend definition captures meaningful social interactions rather than incidental encounters. Next, we compare our data with information from the 2020 China Family Panel Studies (CFPS). According to our data, 35% of job switchers move to a friend's workplace. This aligns closely with the CFPS 2020 data, where 39.44% of the respondents indicated they found jobs through referrals by relatives, friends, or acquaintances.[15]

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 3 (and Figure B1), this aggregation does not materially affect our result. In over 91 percent of cell pairs with a positive number of switchers, there is only one switcher, and in more than 73 percent of cases with positive friend counts, each switcher has only one friend in the destination block's cells. As a result, the difference between individual-level data and aggregated data is immaterial.

[15] Authors' calculation based on 2020 CFPS question KG13A Employment pathway "Which pathways of employment have you tried to gain this main job?"

Job Switchers. Figure 1 presents a comparison between the composition of workers in our sample and the working population in Beijing as reported by official statistics. We find that the worker characteristics in our sample, represented by the blue bars, closely match the overall composition of the workforce shown by the green bars. On average, 34 percent of workers are female, 65 percent are migrants, 70 percent belong to low-income groups, and 71 percent are employed full-time.[16] The job switchers in our sample, shown by the orange bars, have a broadly similar profile, with one notable exception: they are significantly less likely to be full-time workers. This is consistent with expectations, as part-time and temporary workers typically change jobs more frequently than full-time employees.[17]

Face-to-Face Meetings. Figure 2 displays monthly averages for job switchers and non-switchers. Panel A shows that switchers consistently have more friends than non-switchers each month. Panel B reveals a similar pattern for the number of face-to-face meetings, with switchers meeting friends more often on average. Both measures exhibit seasonal variation, peaking in warmer months and declining from November onward.

To better understand how in-person meetings with referrals relate to job switching, we next examine changes in in-person meetings with referrer and nonreferrer friends in the months leading up to a switch using the following model:

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

where Meetings_{ijt} is the number of in-person meetings between switchers in cell i and their friends in cell j in month t. Referrer_j takes a value of one if the switcher has a referrer friend in cell j. A referrer friend refers to a friend employed in the cell where the switcher ultimately moves to. t_i is the month of job switch. s denotes the number of months before the job switch, and it takes a value from -15 to -1. For example, if the job switch happens in January 2019, then meetings in December 2018 are considered as 1 month before the

[16] We use data from the 2020 CFPS nationally representative survey. To define full-time employees in CFPS, we first identify those ever worked 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 besides 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 the 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 break 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.

[17] Empirical studies consistently link non-standard employment arrangements (e.g., part-time, temporary, or gig work) to higher turnover rates, driven by factors such as lower job satisfaction, weaker organizational attachment, and greater labor market flexibility (Kalleberg, 2000; Hyatt and Spletzer, 2013).

job change. Months more than 15 months before the job switch are binned with the tail $s = -15$. For comparison, we define the reference group as the number of in-person meetings with nonreferrer friends two months before the job change, since t=-1 marks the final period of meeting observation before the transition. The month fixed effects μ_t are included to capture the seasonality in face-to-face meetings (see Figure 2). Origin cell fixed effects λ_{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 nonreferrer friends, a trend that holds across all months before 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 explore the communication patterns 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 is 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 2 presents the results. Columns 1 to 4 consider all ij cell pairs where switchers in cell i have at least one friend in j. The dependent variable is the number of face-to-face meetings between cells i and j from 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 cell-level attributes, including the size of the working population as a proxy for population density and the number of friends in the destination cell. 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. In the restricted sample (Columns 5 to 8), the correlations are higher in magnitude. The coefficient on calls is approximately 2.8, indicating that an additional call is associated

with 2.8 more meetings between switchers and their friends. The coefficient on WeChat volume is about 1.5, suggesting that a one-gigabyte (GB) increase in WeChat data usage by switchers in cell i is associated with an increase of 1.5 meetings between switchers in cell i and their friends in cell j. These patterns suggest that different modes of communication are complementary rather than substitutes.

4 Empirical Strategy and Results

4.1 Face-to-face Meetings and Workplace Choices

We begin our exploration of patterns in the data by employing the following linear probability model:

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

where the dependent variables M_{ij} are defined as follows: (1) a dummy variable equals 1 if switchers from cell i change jobs to cell j, or (2) the number of job switchers transitioning from cell i to cell j. 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. b_i and b_j denote the road-delineated city blocks of cell i and j. 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 are better able to isolate the causal effect of communication technology on job location choices, while mitigating the influence of related, unmeasured factors, such as industry composition, labor market demand, commuting distance, local amenities, and intrahousehold dynamics. We include the origin block fixed effects λ_{b_i} to account for unobserved switchers' characteristics, such as job preferences, and destination block fixed effects λ_{b_j} to control for unobserved location attributes such as the industrial composition and the number of available jobs in the destination area. Identification therefore comes from comparisons across workplaces within the same destination neighborhood. The coefficient α is identified by whether a worker has a friend employed at a given workplace, while β, the primary parameter of interest, is identified by differences in interaction method (face-to-face versus remote) holding the destination neighborhood and the presence of a friend constant.

Summary Statistics. Table 3 reports the summary statistics for the regression sample. On average, the probability of switching from cell i to cell j in one year is 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. Appendix Figure B1, which shows that among area pairs with positive switcher counts, 91.75% have exactly 1 switcher. Similarly, among pairs with at least 1 friend, 73.45% have only 1 friend.

In our sample, 9.2 percent of the constructed ij pairs have switchers in cell i with at least one friend working in cell j. Notably, 2.8 percent of these pairs involve a switcher who had face-to-face meetings with friends in cell j during the 6 months, with an average of 3 meetings per such pair. The average annual WeChat data usage by the switchers is 0.33 GB. On average, there are 0.0002 calls between ij pairs during the period from January to June 2019.

Turning to meeting composition, conditional on ij pairs with nonzero meetings, the data reveals that 39 percent of the meetings involve female friends, while 45 percent 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 two groups – small (less than 50 unique call contacts per month) and large (50 or more unique contacts per month). 52 percent of the meetings are with friends from small social circles, and 48 percent are with friends who have large social circles.

Results. Table 4 reports the estimation results for Model 6. In all columns, the standard errors are two-way clustered by origin and destination blocks. 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 6, we observe the coefficient on "Friend" decreases by approximately 70 percent, from 0.128 in Column 1 to 0.031 in Column 2. This significant reduction underscores the value of using granular data, as we do, by demonstrating how geographical fixed effects, defined at a fine spatial scale, 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 0.025, indicating that, on average, having a friend increases the probability of moving to a specific cell by 2.5 percentage points. The coefficient on the interaction term is also precisely estimated at 0.024 and is statistically significant at the 1 percent level. This finding implies that face-to-face meetings with friends double the referral effect. Combined, these effects raise the average propensity to choose a cell within a block from 5 percent to 9.9 percent – almost doubling the likelihood of moving relative to the baseline. This highlights how powerful face-to-face interaction can be in amplifying the effects of social ties in job mobility.

In Column 4, we explore a possible nonlinear effect of face-to-face meetings on 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 not greater than 34, the first tertile. The "medium" group covers cases where the number of meetings is greater than 34 but no more than 108 (the second tertile), while the "high" group corresponds to cases where the number of meetings exceeds 108. In Column 4, the coefficients for the dummies of "low," "medium," and "high" meeting intensity are 0.019, 0.021, and 0.033, 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. We thus maintain the specification in Column 3 as our baseline.

Because the analysis focuses on movers, the resulting sample may differ systematically if job mobility is related to social interactions. To evaluate this channel, we extend the analysis to include nonmovers and explicitly model the choice between staying in the current workplace and moving to alternative destinations. We construct a choice set that includes a "stay" option alongside potential destination cells in blocks where switchers from the same origin cell actually move to and re-estimate the model on the full sample. We find that the estimated coefficients on both the presence of a friend in a destination and the intensity of face-to-face interactions are similar in magnitude to those obtained in the movers-only specification, with differences driven by the change in the mean of the dependent variable. These results support the interpretation that the estimated effects reflect genuine differences in mobility and job-matching behavior rather than artifacts of sample selection.

Appendix Table B2 shows the results when the number of job switchers transitioning from cell i to cell j rather than the probability of job changing from cell i to cell j is used as the dependent variable. The results are similar. Our preferred specification (Column 3) suggests that having face-to-face meetings with friends would increase the number of job switchers moving to a given work cell by 0.081 (0.033+0.048), which increases the mean (0.053) by about 1.5 times. The fact that the results are similar to those in Table 4 is not surprising, since in most cells there is only one switcher.

4.2 The Friend-Move Design

While our baseline design already addresses many of the standard concerns in the referral literature, by conditioning on origin and destination neighborhoods, measuring friendships

prior to mobility, and comparing workplaces within narrowly defined locations, it cannot fully isolate the mechanism through which social ties affect job choice. To sharpen identification and directly isolate the role of interaction technology holding the friendship fixed, we leverage the richness of our data and introduce a more demanding friend-move design, which exploits situations in which multiple friends of the same worker relocate to jobs within the same destination neighborhood.

Let i index the individual worker and $j \in J_i$ denote the set of friends who move to a given neighborhood. We observe the presence of a social link and variation in the mode of interaction – face-to-face versus remote-only – between i and each j measured several months before any observed job move. By conditioning on the destination neighborhood and the existence of the friendship network, we compare, for the same focal worker and neighborhood, the likelihood of following a friend who is met in person versus one who is only contacted remotely. Formally, we estimate:

$$F_{ij} = \alpha_i + \gamma_{b_j} + \tau \text{EverMeet}_{ij} + \nu_{ij} \tag{3}$$

where F_{ij} indicates whether worker i follows friend j into a given workplace, α_i are focal-worker fixed effects, γ_{b_j} are destination neighborhood fixed effects, and EverMeet_{ij} measures the method of interaction. By exploiting pre-determined friendships and the timing of interactions, the friend-move design sharply mitigates confounding from homophily, sorting, reverse causality, and correlated search or mobility shocks, allowing us to identify the role of interaction technology (τ) in job matching within destinations. This approach thus makes the mechanism interpretation unusually clean.

Additionally, we also estimate a specification that includes friend j fixed effects. In this setup, the coefficient on face-to-face interactions is identified from within-friend variation, asking whether, for the same friend, focal individuals are more likely to follow that friend into a workplace when they interact face-to-face rather than only remotely. Formally, the regression can be written as

$$F_{ij} = \alpha_i + \phi_j + \tau \text{EverMeet}_{ij} + \nu_{ij} \tag{4}$$

where ϕ_j captures time-invariant attributes of friend j, including overall quality, prominence, reputation, occupation, firm type, and career stage. By comparing movements tied to the same friend, this specification effectively nets out any confounding due to friend-specific characteristics that might otherwise influence the likelihood of following them.

4.3 Conceptual framework

To clarify the mechanism underlying the friend-move design and to guide the interpretation of the estimates, we next present a simple search-and-referral model that highlights the role of interaction technology in job matching. We build on the network-search framework of Calvó-Armengol and Jackson (2004).

Setup We consider a single search period where worker i is choosing a job j within the set of all available jobs $\mathcal{J}$. The worker can learn about job opportunities through two distinct channels: formal search and social referrals. Each job j has a baseline probability $\pi_j > 0$ of being discovered and accepted through formal channels independently of social contacts such as advertisements or direct applications. Each worker i has a pre-existing set of friends F_i. A friend $f \in F_i$ who is employed at job $j \in \mathcal{J}$ can act as a referral source. The effectiveness of this referral depends critically on the communication technology with the social tie.

Referral Channels When a friend f is employed at job j, they can facilitate a match through two distinct channels: network presence or quality of information. If worker i has a friend $f \in F_i$ at job j, the worker can access information or get a referral. This occurs with an effective probability that increases the match chance by σ. Face-to-face interaction facilitates the transmission of rich, tacit information, such as details about team culture, the hiring manager's preferences, or unadvertised needs. This high-quality referral provides an additional increase in the match probability of ρ.

We define information quality with an indicator variable $q_{ij} \in \{0, 1\}$. It equals 1 if worker i had face-to-face contact with a friend at job j in the period prior to the search, and 0 if the contact was remote-only. A key assumption is that $q_{ij} = 1$ implies a friend is present at j; face-to-face contact is a specific high-bandwidth subset of the general network connection.

Match Probability We model the total probability P_{ij} that worker i matches with job j as the sum of arrival rates from independent information channels. This additive structure is standard in search-and-matching theory, where the likelihood of matching with a specific job is small and well-approximated by the sum of probabilities from distinct sources (Pissarides, 1979; Petrongolo and Pissarides, 2001).

The match probability is given by:

$$P_{ij} = \underbrace{\pi_j}_{\text{Formal Search}} + \underbrace{\sigma \cdot \mathbf{1}\{\text{friend at } j\}}_{\text{Network Presence}} + \underbrace{\rho \cdot q_{ij}}_{\text{Bandwidth Effect}} \tag{5}$$

where σ captures the presence-of-friend effect, and ρ is the additional marginal effect of

having a high-bandwidth, face-to-face contact. ρ is our primary parameter of theoretical interest, as it identifies the mechanism through which face-to-face interactions transfer rich, non-codifiable information and referrals.

4.3.1 Testable predictions

The key features of this model, shared with canonical referral models but rarely tested empirically, are that referrals operate at the level of specific firms or jobs rather than broad locations; take effect only after a vacancy exists and a network contact is employed at the firm; and improve job matching rather than simply shifting overall mobility propensities. From our simple model, we derive four clear, testable predictions:

I. **Timing of the effect:** $\rho > 0$ only after the friend has moved. Prior to the friend's relocation, face-to-face interaction cannot convey job-specific information.

II. **Workplace-level effect:** $\rho > 0$ at the level of specific jobs within a destination, rather than at the neighborhood level. If face-to-face interaction only affected broad mobility choices, the effect within the neighborhood would be zero.

III. **Interaction technology matters:** Among friends in the same destination job or area, face-to-face interactions increase the probability of following the friend relative to remote-only interactions ($\rho > 0$), holding the friendship and destination fixed.

IV. **Presence versus interaction technology:** The model distinguishes two margins of social referrals. The presence effect, $\sigma > 0$, captures whether a worker relocates to a destination area where they have any friends. The technology effect, $\rho > 0$, captures whether, conditional on moving to that destination, the worker matches to the job of a face-to-face friend, reflecting more precise job matching.

In addition, the effects of face-to-face interactions should not be explained by friend characteristics alone.

4.4 Mapping to the Friend-Move Design

The friend-move design directly translates the key primitives of the model, allowing us to test its central predictions. By focusing exclusively on cases in which a friend has already moved to a new workplace, the friend-move design conditions on the existence of a potential referral node and permits investigation on whether the effect of face-to-face interactions appears only after the friend's move (Prediction I). This sharply contrasts with designs that

merely observe the presence of a friend, which may confound network effects with sorting or homophily.

Prediction II and III relate to the role of interaction technology. By comparing individuals with pre-existing ties at a workplace but differing in whether they interact face-to-face or only remotely, the friend-move design isolates the effect of interaction technology. Prediction II is that the effect of such interactions is observed at the level of individual workplaces rather than broadly across destinations, while Prediction III is that conditional on the friendship and destination, face-to-face interactions increase the likelihood of following a friend relative to remote-only interactions.

Prediction IV distinguishes two margins of influence: the presence of a friend versus the intensity of interaction. The friend-move design allows us to separately identify the effect of having a friend in a destination neighborhood, captured empirically by σ, from the effect of face-to-face interactions with that friend, captured by ρ. The presence effect σ measures how social networks shape broader mobility decisions, whereas the interaction effect ρ identifies how in-person interactions influence which specific workplace within the destination a worker joins. Together, these features of the friend-move design allow us to disentangle the role of social networks in mobility choice from the causal impact of in-person interactions on job matching, providing a clean test of the theoretical predictions while addressing concerns of reverse causality, homophily, and correlated shocks.

The friend move design (Model 3) restricts the sample to cases where multiple friends of worker i have recently moved to the same destination block b_j. In this setting, worker fixed effects (α_i) absorb time-invariant heterogeneity in search effectiveness, and destination block fixed effects (γ_{b_j}) absorb the baseline appeal π_j of the area. Critically, the presence-of-a-friend effect σ is also held constant, as the comparison is among friends who have all moved to the same block. The remaining variation used to identify the coefficient τ is the variation in information quality q_{ij} between the focal worker i and different friends in that same destination area b_j. Therefore, τ precisely estimates ρ in the theoretical model: the causal effect of high-quality, face-to-face–enabled information access on the final job-matching decision, netting out the general "presence" effect of merely having a friend nearby.

Results Table 5 presents the results of the friend-move design. This exercise is empirically demanding as identification effectively relies on switchers who had at least two friends that changed jobs earlier and to the same block, but differs in their mode of interaction (face-to-face versus remote-only). From this group, we pair each with friends who changed jobs at least three months before. Column 1 and 2 show the estimation results for Models 3 and 4, respectively (Predictions II, III, and IV). The average probability that a focal switcher follows

a friend to his or her new workplace is 0.00003. Columns 1 and 2 show that face-to-face interaction increases this probability by about 0.00011-0.00013.

Columns 3 and 4 exploit the timing of friends' job transitions to test Prediction I. Specifically, we pair focal switchers with friends who change jobs after the focal worker and estimate whether face-to-face interactions predict "following" the friend's move. Given limited mobility events at fine temporal resolution, we pool pre- and post-move periods rather than estimating fully dynamic effects. The estimates show no differential likelihood of following face-to-face versus remote friends prior to the friend's job transition. In contrast, a positive and significant effect emerges only after the friend has relocated to the destination workplace. This temporal pattern is consistent with a mechanism in which in-person interactions facilitate the transmission of job-relevant information or access once the opportunity becomes available, rather than reflecting pre-existing unobserved similarity or correlated mobility shocks.

The estimated returns to face-to-face interactions imply a 36 to 43 percent increase in the probability of following a friend relative to remote-only interactions, closely matching the baseline magnitudes reported in Column 3 of Table 4. Given this similarity, and the substantially more restrictive data requirements of the friend-move design, we adopt the specification in Column 3 of Table 4 as our baseline for the subsequent analysis.

5 Sensitivity Analysis

Drawing on the longitudinal dimension and richness of our data, we begin the sensitivity analysis by examining vacancy posting dynamics during the face-to-face period. This check directly addresses the possibility that local labor demand shocks drive both face-to-face interactions and mobility decisions and therefore plays a central role in validating our identification strategy. We then present a broader set of robustness checks to assess the sensitivity of our results to alternative specifications and remaining sources of unobserved heterogeneity.

Unobserved Factors that Change Over Time. One might worry that a sudden hiring surge in a given area could increase the likelihood that friends in that location interact face-to-face with a prospective switcher. The switcher might then decide to change jobs there, not because of the face-to-face contact per se, but due to the underlying labor market shock. This scenario seems unlikely: our results hold when controlling for phone calls between friends, and there is no clear reason why a hiring surge would encourage in-person interactions (outside the workplace) rather than remote communication. Nevertheless, we investigate this possibility directly. We compare the number of vacancies across destination areas where

switchers engaged in face-to-face interactions with friends before their job change and those where they did not. If hiring surges were driving face-to-face interactions, we would expect to see systematic differences in vacancy postings around the months of interaction. Figure 4 tests this prediction. It plots, using July 2018 as the reference period, the time trend in the vacancy gap between destination areas with and without face-to-face interactions. The gap remains stable over time, providing no evidence that local hiring surges explain our findings.[18] We further conduct a leave-one-out sensitivity analysis by sequentially removing individuals who change jobs or engage in face-to-face interactions in specific months, one month at a time. Appendix Figure B3 compares the coefficient estimates on "Friend" and "Friend × EverMeet" from these robustness checks with the baseline estimate. In Panel A, we remove face-to-face interactions in different months while retaining the original regression sample. In Panel B, we remove both switchers and their face-to-face interactions, reconstructing the regression sample each time. In both cases, the estimates remain statistically indistinguishable from the baseline. These results indicate that our findings are not driven by time-varying differences between individuals or areas with and without face-to-face interactions, and that the observed averages are not artifacts of unobserved shocks or anomalies in particular months.

Overall In-Person Socializing of the Switchers. In Column 1 of Table 6, we control for the total number of face-to-face meetings of switchers from cell i. This specification isolates the relative pull of the destination cell from the overall propensity to interact socially in person. The results show the coefficients on "Friend" and its interaction with "EverMeet" remain similar. This stability implies that the in-person meetings drive mobility because of their location-specific content not because they are correlated with more overall in-person socializing of the movers. No evidence of relative reallocation effects shows that absolute (not relative) exposure to face-to-face interactions drives the moving decision.

Strong Ties. To ensure that our parameter captures the effect of meeting technology – rather than merely reflecting the influence of stronger ties, who may be more likely to meet in person – we additionally control, in Column 2 of Table 6, for the demeaned initial intensity of face-to-face interactions between switchers and their friends, as measured during the friend definition stage (i.e., weekends and holidays from January to June 2018). The meeting frequency varies considerably across the different initial intensity levels, as shown in Figure B2. The results show that the coefficient on "Friend" decreases from 0.025 to 0.01,

[18] Areas with face-to-face interactions have slightly more vacancies than areas without face-to-face interactions in most months. Baseline results remain stable after controlling for the number of vacancies in cell j during July-December 2018 as shown in Appendix Table B6.

while the coefficient on the interaction term remains relatively stable at around 0.02. This suggests that tie strength influences how much 'having a friend' in the workplace affects job location choices, with weaker ties contributing less to relocation decisions than stronger ones.[19] However, while tie strength explains part of the effect, it does not fully account for the impact of face-to-face interactions. In other words, seeing a friend in person adds a layer of influence on the decision to relocate, even after accounting for the strength of the relationship.

Job Charactersitics. To examine whether unobserved job characteristics may affect our estimates, we restrict the analysis to a subsample of workers whose new workplaces had comparable job opportunities in other locations with friends within the same block. These alternatives are defined by vacancy postings in the same occupation, salary range, and with the same experience and education requirements (see Section 2 and Appendix Section A). Specifically, a location is considered to offer a similar opportunity to the new workplace if it contains at least one job posting in the same occupation and salary range, requiring the same level of experience and education as those of the new workplace.[20] Column 3 of Table 6 reports the results. The coefficient on "Friend*EverMeet" is 0.025, statistically and economically similar to the baseline estimate of 0.024. These findings suggest that our results are not merely driven by job availability, reinforcing the conclusion that face-to-face meeting opportunities play a central role in shaping location choices.

Homophily. To examine homophily in further detail, in Table 6 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.

In Column 4 of Table 6, 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, or in the same income group as their friends working in cell j. Second, in addition to using parametric functions of observables, we use a popular unsupervised machine learning algorithm – k-means clustering – to nonparametrically profile

[19]This evidence is in line with recent large-scale studies finding that strong ties provide significant help for job finding (see, e.g. Gee et al. (2017)).

[20]If a location has multiple job postings, the required education, experience, and salary are calculated as the weighted average of those values across the postings, with weights based on the number of vacancies for each posting. The salary range is defined as the midpoint of the offered salary range with a margin of ± 5,000 rmb (the interquartile range of the monthly salary is 5,250 to 11,468 rmb). Alternative thresholds, such as ± 0.5 standard deviations of the distribution of salary midpoints, do not alter the results.

switchers and friends. 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 median of their distribution across cells.[21] 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 that indicates whether the origin cell i of the switchers is in the same cluster as the destination cell j of their friends. The results in Column 4 of Table 6 show that the coefficient on the interaction between "Friend" and "EverMeet" is approximately 0.02, 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.

Agglomeration Effects. To the extent that face-to-face meetings are more common in areas with larger populations simply because more people are present, the estimated effect may reflect population size rather than the meetings themselves. To explore this possibility, we control for working population size in each alternative destination cell, using it as a proxy for population size.[22] Column 5 of Table 6 presents the results. While the coefficient on having a friend in the workplace declines sharply from 0.025 to 0.007, the effect of face-to-face meetings is only partially absorbed by the control for population size, falling from 0.024 to 0.015. This reduction suggests that part of the effect of in-person interactions is indeed driven by differences in population size across locations. However, the remaining significant effect indicates that face-to-face interaction still plays an independent and meaningful role in shaping mobility decisions.

Preferences for Friend Closeness. Areas with more friends may be more attractive, and also more likely to contain at least one face-to-face friend. To examine this possibility, we control for the number of friends in destination neighborhoods. Column 6 of Table 6 shows the results. The coefficient on the interaction between "Friend" and "EverMeet" is 0.024, closely aligned with the estimate in Column 3 of Table 4, and remains statistically significant at the 1 percent level. This finding suggests that even after accounting for the size of the social network in an area, face-to-face meetings convey additional value that increases the likelihood of moving in a specific workplace.

[21] For example, an origin cell i is female dominant if the share of female switchers is above its median in the female share distribution across cells.

[22] The working population in a cell is measured as the number of individuals working in that cell, based on cellphone location data.

Local Amenities. In Column 7 of Table 6, we control for the amenities at the level of each alternative location within the destination area. Specifically, Column 7 introduces a rich set of controls, including the number of parking lots, restaurants, hotels, schools, marketplaces, sports and entertainment sites at the cell level. The effect of face-to-face meetings remains robust and is estimated as 0.022.

Unobserved Heterogeneity at the Origin-Destination Level. In Column 8 of Table 6, we add the interactions of origin and destination block fixed effects. As a result, the coefficient α is identified by exploiting variation in location choices among individuals moving within the same origin–destination block pair but differing in their friend network. The coefficient β, the primary parameter of interest, is 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. The coefficient on the interaction between "Friend" and "EverMeet" is 0.031, in line with the estimate in Column 3 of Table 4.

6 Additional Evidence

6.1 Face-to-Face Meetings and Other Communication Channels

To assess the relative importance of in-person meetings, phone conversations, and mobile messaging applications in job location choices, we leverage supplementary data on phone calls and WeChat activity. We focus on a period with comprehensive coverage of all communication technologies –January to June 2019–and examine job transitions occurring in the subsequent months, 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 in Column 3 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 0.0016, an effect almost twice that of a one-standard-deviation increase in WeChat volume and also larger than that of calls, which remains statistically insignificant.

6.2 Heterogeneity by Worker Characteristics

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

To investigate heterogeneity in the returns to face-to-face interactions, we extend our baseline model (Model 6) by interacting the variables "Friend" and "Friend × EverMeet" with indicators for the dominant characteristics of switchers. As the worker flow and pre-job-change interactions are aggregated at the cell pair level, we restrict our analysis to cell pairs where a majority of job switchers in the origin cell i share a specific characteristic. For example, if the share of female switchers in a cell exceeds the median across all cells, we classify it as "female-dominant" and include it in the corresponding analysis. We examine heterogeneity by worker characteristics 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 3), we can reasonably interpret the results at the individual level.

Panel A of Table 8 reports estimates of the differential effect of face-to-face meetings ("Friend × EverMeet") across these characteristics, each from a separate regression. To ensure meaningful comparisons, we exclude cells without a clear majority, e.g., for gender, we drop cells where the female share falls below the median. We find no statistically significant differential return by gender in Column 1. Column 2 suggests that, conditional on having a friend, face-to-face meetings increase the probability of relocating to a given cell by an additional 0.5 percentage points for migrant workers relative to locals. Column 3 shows a pronounced gap between full-time and part-time workers. The benefit of face-to-face interaction is 1.2 percentage points larger for full-time employees than part-time workers. Similarly, Column 4 shows that high-income workers experience a 0.7 percentage points stronger effect than low-income workers. The full regression results, including all controls, are provided in Appendix Tables B7 and B8.

The next question we ask is whether the channels underlying these benefits differ across worker groups. To explore this, we examine how variation in both meeting frequency and meeting composition relates to heterogeneous returns. Panels B and C of Tables 8 progressively incorporate controls for interaction quantity and composition. Specifically, we examine the composition of these meetings by interacting "Friend*EverMeet" with indicators for specific meeting types, namely, whether the majority of meetings were with female friends, migrant friends, high-income friends, or friends with different social circle sizes. Con-

trolling for meeting frequency accounts for most of the premium observed among migrant and high-income workers, whose interaction premia become statistically insignificant once exposure is held constant. In contrast, the gap for full-time workers persists until we account for compositional features of their meetings.

Appendix Tables B7 and B8 provide additional insight into these mechanisms. While both the total number of meetings and meetings with high-income friends are significant predictors across groups, a significant interaction with large-circle friends emerges only for full-time workers. Specifically, the main effect of meeting large-circle friends is not statistically significant, but its interaction with full-time status is positive and significant (0.008, $p < 0.1$). This pattern suggests that, unlike part-time workers, full-time workers derive a distinct advantage from interacting with highly connected contacts. Interactions with well-connected individuals therefore appear to constitute a key mechanism through which face-to-face exposure translates into mobility outcomes for this group.

These findings indicate that both the intensity and specific composition of meetings shape heterogeneous returns to interaction. In the next section, we extend this analysis by examining a broader dimension of compositional exposure, namely, the diversity of meeting partners across key socioeconomic characteristics.

6.3 Diverse vs. Homophilous Meetings

We extend our analysis by contributing to the growing literature on how diversity in social contacts relates to economic decision-making (Page, 2007; Eagle et al., 2010; Chetty et al., 2016). Recognizing that workers may experience different outcomes depending on whom they meet in person, we examine the compositional features of face-to-face interactions by distinguishing between meetings with similar and dissimilar contacts. We construct two indicators of meeting composition:(1) "Majority same type," equal to one when most meetings occur with friends sharing the same characteristics as the job switcher, and (2) "Majority different type," equal to one when most meetings occur with friends of a different type. We then partition the sample according to dominant characteristics and estimate two specifications separately, alternately including the same-type and different-type indicators as additional regressors.[23]

Figure 5 summarizes the estimated associations between meeting composition and mobility across worker groups. Each point reports the coefficient on either "Majority same type" or "Majority different type" estimated within a given subsample, with bars indicating 95% confidence intervals. For female-dominant origin cells, meetings that are predominantly with

[23]For instance, in female-dominant origin cells, where most switchers are female, we re-estimate the baseline specification twice: once including "Majority same type" and once including "Majority different type."

same-gender friends are associated with an approximately 0.006 higher probability of job mobility for female workers (about a 12 percent difference relative to the mean propensity), while meetings that are predominantly with male friends are associated with a reduction of similar magnitude. These differences may reflect variation in the content, framing, or interpretation of information exchanged across gendered social ties, rather than directional behavioral responses. Among migrants, the pattern is similar but smaller: meetings concentrated among migrant friends are associated with higher mobility (about 0.004), whereas meetings concentrated among local friends are associated with a comparable decrease, consistent with migrants relying more heavily on same-type contacts when navigating destination labor markets. A strong income gradient also emerges. High-income workers display larger mobility differences when meetings are concentrated among other high-income contacts (0.007) and smaller mobility differences when meetings are concentrated among lower-income contacts (-0.007). In contrast, low-income workers exhibit stronger mobility associations with cross-income interactions, consistent with the idea that exposure to higher-income contacts is linked to access to information and opportunities that may be less available within more homogeneous networks.

Taken together, these results indicate that diversity in meeting composition along key dimensions — gender, migration status, and income — is systematically associated with distinct mobility patterns. Combined with the evidence in Section 6.2, the findings highlight that not only the frequency of interactions but also the composition of meetings constitutes an empirically salient dimension of social exposure for economic decision-making. In other words, who workers meet appears to matter alongside how often they meet, pointing to the importance of interaction content and social diversity in shaping labor market outcomes.

6.4 Data-Driven Heterogeneity and Targeting Performance

While Table 8 documents heterogeneity in the return to face-to-face interaction along theory-motivated dimensions, we next examine whether the data reveal structure in treatment heterogeneity beyond pre-specified categories. To do so, we complement the preceding analysis with a data-driven framework that uses machine learning to identify and rank the settings in which face-to-face interactions generate the largest mobility responses.

This exercise serves several purposes. First, it provides an independent assessment of treatment heterogeneity that is not restricted to a small set of ex-ante partitions, allowing the data to uncover combinations of characteristics associated with higher returns. Second, by evaluating predicted treatment effects out of sample, it validates whether heterogeneity is systematic and economically meaningful rather than sample-specific. Third, the resulting

ranking enables a transparent assessment of targeting gains by quantifying how much larger mobility responses are in high-return settings relative to the population average. Finally, by re-estimating the prediction model across alternative training-testing partitions, we assess the stability of the identified patterns across subsamples, thereby strengthening the case for the external validity of our findings.

We implement a Machine Learning procedure — LASSO regularization — to estimate predicted treatment returns using a high-dimensional set of worker, workplace, and meeting characteristics. Following Chernozhukov et al. (2025), we train the model on an initial subsample of job switchers moving within the first nine months and evaluate it out of sample on a held-out testing partition covering switchers moving in the last three months. Specifically, we augment the baseline model using LASSO with interactions between "Friend × EverMeet" and the full set of characteristics to capture rich heterogeneous effects. The model is estimated using variables residualized against origin and destination block fixed effects in the training sample, with penalty parameter selected via 10-fold cross-validation. In the out-of-sample prediction, to ensure the consistency with the learned parameters, we first construct "benchmarked residuals" for the testing sample, by subtracting the block-level means calculated from the training sample from the raw testing data. We then apply the frozen LASSO coefficients to the benchmarked residuals of the interaction between"Friend × EverMeet" and characteristics to obtain predicted returns and construct targeting quartiles. This approach allows us to assess whether face-to-face interactions generate systematically larger mobility responses in settings where predicted returns are highest. Details can be found in Appendix Section B.

Table 9 reports the results. Panel A presents the out-of-sample targeting exercise. We partition the testing sample into high- and low-predicted-return groups based on the LASSO predictions and re-estimate the baseline face-to-face effect within each bin using the same fixed-effects specification as Model 6 in the main analysis. The results show a clear gradient. The estimated face-to-face effect in the top quartile is substantially larger than in the full sample and markedly exceeds the effect in the low predicted-return group. Expressed relative to baseline mobility rates (Column 4), the impact of face-to-face interaction is more than three times as large in the high predicted-return group as in the overall population. These findings indicate that the model successfully ranks settings by expected responsiveness and that targeting in-person interaction toward high-return environments yields sizable mobility gains.

Panel B reports the most important moderators selected by the LASSO procedure. The top predictors of treatment heterogeneity include the number of in-person meetings, full-time employment status, whether the majority of meetings occur with low-income contacts,

whether meetings disproportionately involve highly connected ("large-circle") contacts, and employment in Information Transmission, Software, and IT Services. These predictors align closely with mechanisms emphasized in referral and job-search models, where both the intensity of interaction and the informational resources embedded in social ties shape matching outcomes.

We further replicate the targeting exercise within the friend-move design in Table 10. Despite the more restrictive data requirements of this specification, the same qualitative pattern emerges: treatment effects are again larger in the high predicted-return group. Moreover, predicted returns collapse to zero in pre-move placebo periods, reinforcing the interpretation that the heterogeneity identified by the model reflects variation in referral-based information transmission rather than spurious correlation or sorting. This consistency across empirical designs strengthens the mechanism interpretation of face-to-face interaction as facilitating job-specific information and access.

Importantly, the moderators selected by the data-driven procedure are predominantly network-based measures – such as meeting intensity, the connectedness and income of contacts, and interaction composition – rather than features specific to particular firms, industries, or local labor market structures. This pattern suggests that the mechanisms emphasized in the paper operate through general properties of social interaction and information transmission rather than idiosyncratic aspects of particular institutional settings. As a result, the machine-learning exercise not only strengthens the internal validity of our heterogeneity findings but also provides indirect evidence of their external relevance.

To further probe external validity, we re-estimate the prediction model using alternative trainingâtesting splits defined along economically meaningful dimensions: working population size, proximity to subway stations, sector (private vs. public), and industry composition. Specifically, we (i) train on destination blocks in the top 70 percent of the working population distribution and test on the bottom 30 percent; (ii) train on blocks within 1 km of a subway station and test on those farther away; (iii) train on cells dominated by private-sector employment and test on public-sector cells; and (iv) train on areas dominated by the two largest industries (wholesale/retail trade and water/environment/public facilities management) and test on all other industries.[24] Table 11 reports the results. Across these splits, the targeting gradient remains robust: predicted high-return settings consistently exhibit larger realized treatment effects. This stability indicates that the heterogeneity uncovered by the model

[24]In terms of 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 establishments in all other industries not classified as public.

reflects structural features of social interaction and job matching rather than idiosyncratic characteristics of a particular institutional or geographic context.

7 Conclusion

Most empirical studies of referrals rely on reduced-form correlations between having a friend at a firm and employment outcomes, correlations that are consistent with multiple mechanisms, including homophily, shared preferences, and correlated search behavior. We introduce a friend-move design that isolates a sharper margin: settings in which socially connected friends are employed in the same destination area but differ in whether they interact face-to-face. This variation allows us to identify the role of interaction technology in facilitating information transmission and job access within social networks.

Using high-frequency geolocation data, we show that face-to-face interaction substantially increases the probability that workers follow their contacts into specific workplaces, with effects emerging only after the contact joins the destination firm. These findings provide direct evidence on a central, yet empirically under-validated, assumption in referral and agglomeration models: that the economic value of social networks depends not only on the existence of ties but on the opportunity for in-person interaction. Urban density, in this sense, facilitates matching not merely by connecting workers, but by enabling repeated face-to-face contact through which job-specific information flows.

Our heterogeneity and machine-learning analyses further show that meeting intensity, and network composition, especially the connectedness of contacts strongly predict mobility responses, and a targeting exercise reveals that effects in the highest predicted-return environments are more than twice as large as the population average. These results highlight substantial gains from directing in-person interaction — through workplace design, hybrid-work policies, or networking infrastructure — toward settings where referral-based information transmission is most valuable.

Taken together, the evidence underscores that interaction technology is a central determinant of labor market mobility. As remote work reshapes where and how workers connect, understanding not only who is connected but how they interact becomes critical for firms, cities, and policymakers seeking to sustain efficient job matching and economic dynamism.

References

Athey, S. and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360.

Atkin, D., Chen, M. K., and Popov, A. (2022). The returns to face-to-face interactions: Knowledge spillovers in silicon valley. Working Paper 30147, National Bureau of Economic Research.

Bailey, M., Cao, R., Kuchler, T., and Stroebel, J. (2018). The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy*, 126(6):2224–2276.

Barwick, P. J., Liu, Y., Patacchini, E., and Wu, Q. (2023). Information, mobile communication, and referral effects. *American Economic Review*, 113(5):1170–1207.

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

Buchel, K., Ehrlich, M. V., Puga, D., and Viladecans-Marsal, E. (2019). Calling from the outside: The role of networks in residential mobility. Working Paper wp2019-1909, CEMFI.

Burks, S. V., Cowgill, B., Hoffman, M., and Housman, M. (2015). The value of hiring through employee referrals. *Quarterly Journal of Economics*, pages 805–839.

Calvó-Armengol, A. and Jackson, M. O. (2004). The effects of social networks on employment and inequality. *American Economic Review*, 94(3):426â454.

Chernozhukov, V., Demirer, M., Duflo, E., and FernÃ¡ndezâVal, I. (2025). Fisherâschultz lecture: Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in india. *Econometrica*, 93(4):1121–1164.

Chetty, R., Hendren, N., and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855â902.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barberá, P., Bhole, M., and Wernerfelt, N. (2022a). Social capital I: measurement and associations with economic mobility. *Nature*, 608(7921):108–121.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg,

E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barberá, P., Bhole, M., and Wernerfelt, N. (2022b). Social capital II: determinants of economic connectedness. *Nature*, 608(7921):122–134.

Davis, S. J. and Haltiwanger, J. (1992). Gross job creation, gross job destruction, and employment reallocation*. *The Quarterly Journal of Economics*, 107(3):819–863.

Duranton, G. and Puga, D. (2004). Chapter 48 - micro-foundations of urban agglomeration economies. In Henderson, J. V. and Thisse, J.-F., editors, *Cities and Geography*, volume 4 of *Handbook of Regional and Urban Economics*, pages 2063–2117. Elsevier.

Duranton, G. and Puga, D. (2020). The economics of urban density. *Journal of Economic Perspectives*, 34(3):3â26.

Dustmann, C., Glitz, A., Schonberg, U., and Brucker, H. (2016). Referral-based Job Search Networks. *The Review of Economic Studies*, 83(2):514–546.

Eagle, N., Macy, M., and Claxton, R. (2010). Network diversity and economic development. *Science*, 328(5981):1029–1031.

Friebel, G., Heinz, M., Hoffman, M., and Zubanov, N. (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.

Gee, L. K., Jones, J., and Burke, M. (2017). Social networks and labor markets: How strong ties relate to job finding on facebookâs social network. *Journal of Labor Economics*, 35(2):485–518.

Glaeser, E. L., Kominers, S. D., Luca, M., and Naik, N. (2015). Big data and big cities: The promises and limitations of improved measures of urban life. Working Paper 21778, National Bureau of Economic Research.

Hoffman, M. and Stanton, C. T. (2024). People, practices, and productivity: A review of new advances in personnel economics. Working Paper 32849, National Bureau of Economic Research.

Hyatt, H. R. and Spletzer, J. R. (2013). The recent decline in employment dynamics. *IZA Journal of Labor Economics*, 2(1):5.

Ioannides, Y. M. and Loury, L. D. (2004). Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, 42:1056–1093.

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

Kalleberg, A. L. (2000). Nonstandard employment relations: Part-time, temporary and contract work. *Annual Review of Sociology*, 26(Volume 26, 2000):341–365.

Montgomery, J. D. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *The American Economic Review*, 81(5):1408–1418.

Mullainathan, S. and Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2):87â106.

Nie, P. and Sousa-Poza, A. (August 2017). What chinese worker value: An analysis of job satisfaction, job expectations and labor turnover in china. *IZA Discussion Papers No. 10963*, 108.

Page, S. E. (2007). *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies (New Edition)*. Princeton University Press.

Petrongolo, B. and Pissarides, C. A. (2001). Looking into the black box: A survey of the matching function. *Journal of Economic Literature*, 39(2):390â431.

Pissarides, C. A. (1979). Job matchings with state employment agencies and random search. *The Economic Journal*, 89(356):818–833.

Rosenthal, S. S. and Strange, W. C. (2004). Chapter 49 - evidence on the nature and sources of agglomeration economies. In Henderson, J. V. and Thisse, J.-F., editors, *Cities and Geography*, volume 4 of *Handbook of Regional and Urban Economics*, pages 2119–2171. Elsevier.

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

Schmutte, I. (2016). How do social networks affect labor markets? *IZA World of Labor*.

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

Topa, G. (2001). Social interactions, local spillovers and unemployment. *The Review of Economic Studies*, 68(2):261–295.

Topa, G. (2011). Labor markets and referrals. *Handbook of Social Economics*, 1B:1193–1221.

Figures and Tables

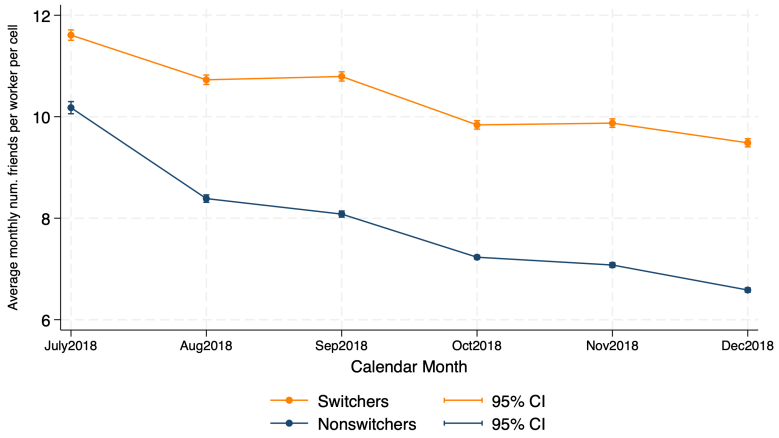
Figure 1: Comparison with Official Statistics



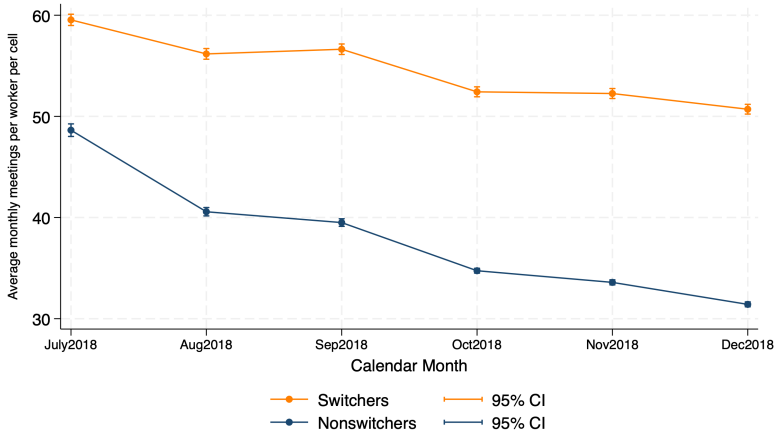
Notes: Figure 1 reports the average fraction of workers who are female, migrants, full-time, and low-income across three samples. The green bars correspond to all urban employees in Beijing in CFPS 2020. The blue bars include all workers observed in a cell in our cellphone dataset. The orange bars restrict the sample to job switchers within a cell. Number of cells = 13,585. In CFPS, "full-time" employees are defined in two steps. First, we identify individuals who have ever worked part-time (i.e., those who (a) have not worked full-time, (b) report part-time work, or (c) report holding additional jobs besides their main job). Individuals not meeting any of these conditions are classified as potential full-time workers. Second, we exclude from this group those employed in the agricultural sector or reporting fewer than 30 weekly working hours. "Low-income" employees are defined by ranking urban workers into eight groups based on annual wage income; groups 1â5 are classified as low-income and groups 6â8 as high-income.

Figure 2: Number of Friends and In-Person Meetings over Time

(a) Number of Friends Ever Met per Worker

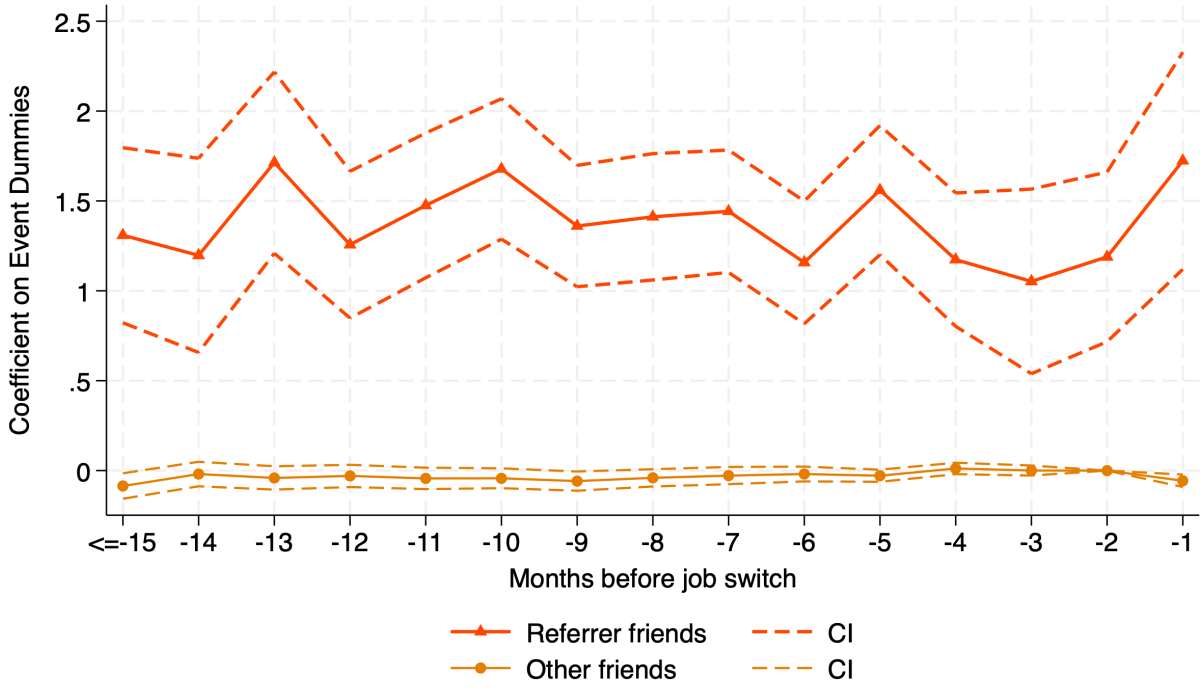


(b) Number of Meetings per Worker



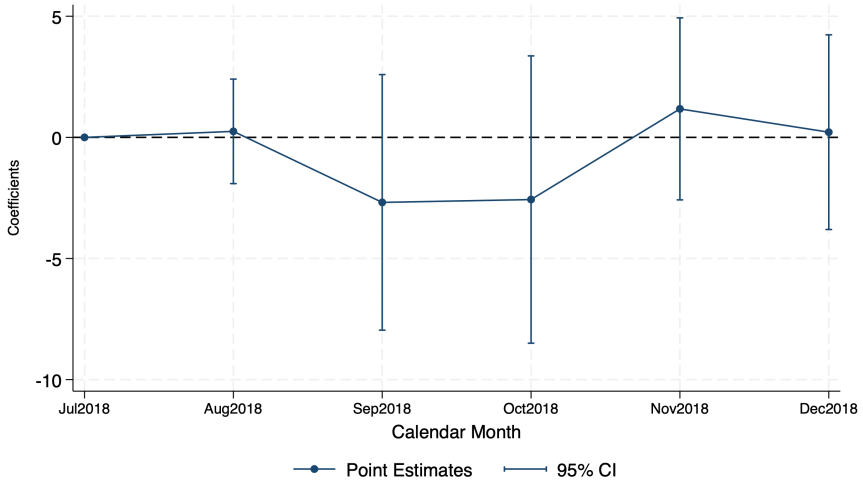
Notes: Figures 2a and 2b plot the average number of friends ever met and the average number of in-person meetings per worker in a given month, along with their 95 percent confidence intervals. In both figures, the orange line (triangles) represents job switchers across cells, while the blue line (dots) represents all workers across cells. Point estimates and confidence intervals are obtained from regressions of the number of friends or meetings on calendar-month dummies, using a sample where the unit of observation is a cellâmonth pair.

Figure 3: Face-to-Face Meetings with Referrer Friends vs. Nonreferrer Friends



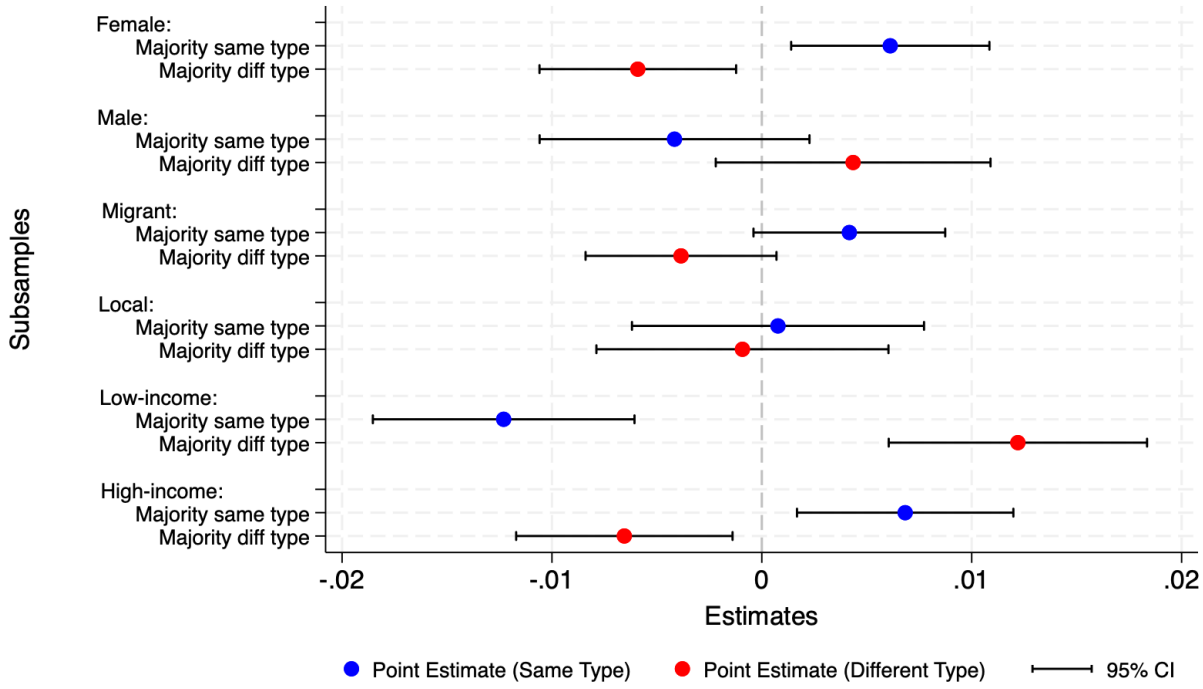
Notes: Figure 3 reports the estimated coefficients and corresponding 95 percent confidence intervals from Model 1 for the number of face-to-face meetings between job switchers in the origin cell and their referrer friends versus non-referrer friends prior to the job switch. The dark orange line (triangles) represents meetings between switchers and referrer friends (149,701 observations). The light orange line (dots) represents meetings between switchers and non-referrer friends (62,266,236 observations). The omitted category is the number of meetings between switchers and non-referrer friends two months before the job switch. Months earlier than 15 months prior to the switch are binned with the 15th month. Standard errors are clustered at the origin-cell level.

Figure 4: Job Vacancies in Destination Areas by Face-to-Face Meetings with Friends



Notes: Figure 4 plots the difference in the number of vacancies between destination cells with and without face-to-face interactions, relative to July 2018.

Figure 5: Effect of In-Person Meetings with Same-Type vs. Different-Type Friends, by Worker Characteristics



Notes: Figure 5 reports the estimated effects of meeting with same-type versus different-type friends across worker types. Each coefficient is estimated on a separate subsample of cell pairs in which the majority of switchers in the origin cell are female, male, migrant, local, low-income, or high-income workers, respectively. "Majority same type" is an indicator equal to one if the majority of meetings are with friends of the same type as the job switchers, while "Majority different type" equals one if the majority of meetings are with friends of a different type. In each subsample, we re-estimate the baseline model including either "Majority same type" or "Majority different type" as an additional regressor and report the corresponding coefficient along with its 95 percent confidence interval.

Table 1: Job Switch By Month

Month	Number of job switchers	Percent among total job switchers	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 reports the number and percentage of job switchers, as well as the monthly switching rate for each month in 2019. The switching rate is defined as the number of job switchers divided by the working population in the corresponding month.

Table 2: Complementarity of Face-to-Face Meetings, Phone Calls, and WeChat Volume of Switchers Before Job Change

Dependent variable:	Switchers in i							
Face-to-face meetings between	with at least one friend in j				move to j July-Dec 2019			
cell i and j Jan-June 2019	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Calls	1.091***	1.090***	1.091***	1.090***	2.920**	2.925**	2.819*	2.827*
	(0.422)	(0.421)	(0.422)	(0.421)	(1.426)	(1.426)	(1.467)	(1.465)
Wechat volume (GB)	0.838***	0.827***	0.836***	0.826***	1.614***	1.575***	1.474***	1.458***
	(0.0412)	(0.0405)	(0.0411)	(0.0405)	(0.0873)	(0.0866)	(0.0959)	(0.0948)
Number of workers at i		0.002***		0.001***		0.001***		0.0004***
		(0.0002)		(0.0002)		(0.0001)		(0.0001)
Number of workers at j		0.002***		0.002***		0.001***		0.001***
		(0.0003)		(0.0003)		(0.0002)		(0.0002)
Number of friends			0.0005***	0.0002***			0.003***	0.002***
			(0.0001)	(0.0001)			(0.0003)	(0.0003)
Observations	4,063,119	4,063,119	4,063,119	4,063,119	467,431	467,431	467,431	467,431
R-squared	0.021	0.021	0.021	0.021	0.145	0.146	0.150	0.151
Block i FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Block j FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table examines complementarity across face-to-face meetings, phone calls, and WeChat interactions between job switchers and their friends prior to the job change. The unit of observation is an ij cell pair. Columns 1-4 include all pairs in which switchers in cell i have at least one friend in cell j. Columns 5-8 restrict the sample to pairs in which switchers from cell i obtain jobs in cell j between July and December 2019. The dependent variable is the total number of face-to-face meetings between switchers and their friends in January-June 2019. "Calls" denotes the total number of phone calls between switchers and friends during the same period. "WeChat volume (GB)" measures the total WeChat data usage (in gigabytes) of switchers in cell i who have friends in cell j over January-June 2019. Even-numbered columns control for the working population size in cells i and j. Columns 3, 4, 7, and 8 additionally control for the number of friends in cell j for switchers from cell i. All specifications include origin and destination block fixed effects. Standard errors are two-way clustered by origin and destination block and reported in parentheses. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table 3: Summary Statistics

Variable	Mean	SD	Min	Max
Panel A. Worker flow and face-to-face 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)	108.53	119.89	1	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 ($\geq$50)	0.48	0.47	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 (January-June 2019)	0.0002	0.067	0	55

Notes: This table reports summary statistics for the regression sample. In Panel A, the sample includes 4,810,573 cell pairs. In Panel B, the sample is restricted to cell pairs with at least one face-to-face meeting in the pre-job-change period, yielding 134,633 cell pairs. In Panel C, the number of cell pairs is 4,810,573 for face-to-face meetings and 2,169,538 for calls.

Table 4: Effect of Face-to-Face Meetings on Job Location Choices

Dependent variable: Probability of switching from i to j	(1)	(2)	(3)	(4)	(5) Add Nonmovers
Friend	0.128***	0.031***	0.025***	0.025***	0.01***
	(0.013)	(0.002)	(0.002)	(0.002)	(0.001)
Friend*EverMeet			0.024***		0.005***
			(0.002)		(0.001)
F2F meetings: Low				0.019***	
				(0.002)	
F2F meetings: Medium				0.021***	
				(0.002)	
F2F meetings: High				0.033***	
				(0.003)	
Observations	4,805,218	4,805,200	4,805,200	4,805,200	13,931,711
R-squared	0.030	0.229	0.230	0.230	0.253
Dep. var. mean	0.053	0.053	0.053	0.053	0.016
Block i FE	No	Yes	Yes	Yes	Yes
Block j FE	No	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 an ij cell pair. The dependent variable is an indicator equal to one if switchers move from cell i to cell j. "Friend" is a dummy equal to one if switchers in cell i have at least one friend working in cell j. "EverMeet" is a dummy equal to one if switchers from cell i have ever met face-to-face with friends working in cell j. "F2F meetings: Low/Medium/High" are indicators for the number of face-to-face meetings between cells i and j being positive and falling into each tertile of the distribution (cutoffs at 34 and 108 meetings). Column 5 expands the sample to include individuals in the origin cell who did not switch jobs (non-switchers). Each non-switcher cell is matched to its own cell (stay) and to cells within the same block that received switchers from the same origin cell. Columns 2-5 include origin and destination block fixed effects. Standard errors are two-way clustered by origin and destination block and reported in parentheses. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table 5: Effect of Face-to-Face Meetings on Job Location Choices: Friend-Move Design

Dependent variable:	(1)	(2)	(3)	(4)
Probability of i following friend j	i moves after j		i moves before j	
EverMeet	0.00011**	0.00013**	-0.00001	-0.00002
	(0.00001)	(0.00001)	(0.00004)	(0.00004)
Observations	1,259,905	1,253,664	3,934,460	3,921,194
R-squared	0.003	0.083	0.002	0.038
Focal i FE	Yes	Yes	Yes	Yes
Block j FE	Yes	No	Yes	No
Friend j FE	No	Yes	No	Yes

Notes: This table examines the effect of face-to-face meetings on job changers' location choices using a friend-move design. The unit of observation is an ij focal-friend pair, where i is "focal" job switchers and j is a friend who changed jobs. Columns 1 and 2 present the main results, where friend j changed jobs at least three months before i; Columns 3 and 4 present a test of theoretical prediction (I), where friend j changed jobs *after* focal i. The dependent variable is an indicator equal to one if focal i moves to friend j's new workplace. "EverMeet" is a dummy equal to one if focal i has ever met face-to-face with friend j. Columns 1 and 3 include focal i and destination block fixed effects. Columns 2 and 4 include focal i and friend j fixed effects. Standard errors are clustered by the focal individual's origin cell and the month of the job switch and reported in parentheses. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table 6: Robustness

Dependent Variable: Probability of switching from i to j	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Friend	0.025***	0.010***	0.025***	0.025***	0.007***	0.025***	0.020***	0.031***
	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Friend*EverMeet	0.025***	0.022***	0.025***	0.020***	0.015***	0.024***	0.022***	0.031***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Cell i total F2F (x1000)	-0.0001***							
	(0.00002)							
Number of meetings: initial stage		0.0002***						
		(0.0001)						
Days of meetings: initial stage		0.002***						
		(0.0003)						
Number of workers at j					0.0001***			
					(0.00002)			
Number of i's friends in block j						-0.00002***		
						(0.00001)		
Observations	4,805,200	4,805,200	1,498,297	4,805,200	4,805,200	4,805,200	4,805,200	4,787,639
R-squared	0.230	0.230	0.153	0.230	0.238	0.230	0.232	0.169
Similar job opportunities nearby	No	No	Yes	No	No	No	No	No
Whether friend has same gender, migrant status and income group	No	No	No	Yes	No	No	No	No
Whether cells in same cluster	No	No	No	Yes	No	No	No	No
Amenities	No	No	No	No	No	No	Yes	No
Block i FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Block j FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Block i x j FE	No	No	No	No	No	No	No	Yes

Notes: This table reports robustness checks of the baseline results in Table 4. The unit of observation is an ij cell pair. Columns 1-2 control for the total number of face-to-face meetings of switchers from cell i; Columns 2 additionally controls for demeaned initial meeting intensity (number of meetings and distinct meeting days) on weekends and holidays, January to June 2018. Column 3 restricts the sample to switchers whose destination cell has at least one comparable job opportunity in another cell within the same block (same occupation, salary range, and experience/education requirements). Column 4 adds homophily controls: indicators for whether dominant gender, migrant status, and income group match between switchers in cell i and friends in cell j, and an indicator for belonging to the same group k-means cluster (based on dominant gender, migrant status, and eight income-group bins). Column 5 controls for the working population in cell j. Column 6 controls for the number of friends of switchers in the block containing j. Column 7 adds local amenity controls in cell j (parking lots, restaurants, hotels, schools, marketplaces, sports and entertainment venues). All columns include origin and destination block fixed effects. Column 8 further includes origin-by-destination block-pair fixed effects. Standard errors are two-way clustered by origin and destination block, and reported in parentheses. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table 7: Face-to-Face Meetings and Other Communication Methods

Dependent Variable: Probability of switching from i to j	(1)	(2)	(3)	(4)
Friend	0.039***	0.041***	0.040***	0.037***
	(0.003)	(0.003)	(0.003)	(0.003)
Friend*std_F2F meetings (x10)	0.014***			0.016***
	(0.003)			(0.003)
Friend*std_calls (x10)		0.003*		0.002
		(0.001)		(0.002)
Friend*std_WeChat volume (x10)			0.006***	0.009***
			(0.002)	(0.002)
Observations	2,179,087	2,179,087	2,179,087	2,179,087
R-squared	0.235	0.235	0.235	0.235
Block i FE	Yes	Yes	Yes	Yes
Block j FE	Yes	Yes	Yes	Yes

Notes: This table compares the effects of face-to-face meetings, phone calls, and WeChat volume on job changers' location choices. The unit of observation is an ij cell pair. The dependent variable is an indicator equal to one if switchers move from cell i to cell j between July and December 2019. Measures of face-to-face meetings, calls, and WeChat volume are computed over January-June 2019. Columns 1-4 interact "Friend" with standardized measures of each communication variable (demeaned and divided by the standard deviation), scaled by a factor of 10. All specifications include origin and destination block fixed effects. Standard errors are two-way clustered by origin and destination block and reported in parentheses. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table 8: Heterogeneous Effects By Worker Characteristics

Dependent Variable: Probability of switching from i to j	(1)	(2)	(3)	(4)
Panel A: No control				
Friend*EverMeet	0.025***	0.021***	0.015***	0.020***
	(0.003)	(0.002)	(0.002)	(0.002)
Friend*EverMeet*Female	-0.001			
	(0.003)			
Friend*EverMeet*Migrant		0.005*		
		(0.003)		
Friend*EverMeet*Full-time			0.012***	
			(0.002)	
Friend*EverMeet*High-income				0.007***
				(0.002)
R-squared	0.230	0.230	0.230	0.230
Panel B: Control for number of meetings				
Friend*EverMeet	0.017***	0.018***	0.011***	0.015***
	(0.002)	(0.003)	(0.003)	(0.003)
Friend*EverMeet*Female	0.001			
	(0.002)			
Friend*EverMeet*Migrant		-0.0003		
		(0.003)		
Friend*EverMeet*Full-time			0.009***	
			(0.003)	
Friend*EverMeet*High-income				0.004
				(0.003)
R-squared	0.230	0.230	0.230	0.230
Panel C: Control for number of meetings AND meeting composition				
Friend*EverMeet	0.013***	0.020***	0.012***	0.014***
	(0.004)	(0.004)	(0.004)	(0.004)
Friend*EverMeet*Female	0.005			
	(0.004)			
Friend*EverMeet*Migrant		-0.006		
		(0.004)		
Friend*EverMeet*Full-time			0.005	
			(0.005)	
Friend*EverMeet*High-income				0.004
				(0.005)
R-squared	0.230	0.230	0.230	0.230
Observations	4,795,743	4,785,658	4,791,002	4,800,901
Block i FE	Yes	Yes	Yes	Yes
Block j FE	Yes	Yes	Yes	Yes

Notes: This table reports the heterogeneous effects of face-to-face interactions by worker characteristics. The unit of observation is an ij cell pair. Each column reports the coefficient on the interaction between "Friend × EverMeet" and switcher characteristics. A characteristic is defined as dominant if it is shared by the majority of switchers in the origin cell and the proportion of switchers with that characteristic exceeds the median of its distribution. Panel A uses the preferred specification from Column 3 of Table 4. Panel B additionally controls for "Friend" interacted with the number of meetings. Panel C further controls for "Friend" interacted with meeting composition, including indicators for whether the majority of meetings of switchers in cell i with friends in cell j involve female, migrant, high-income friends, or friends from different-sized social circles. All specifications include origin and destination block fixed effects. Standard errors are two-way clustered by origin and destination block and reported in parentheses. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table 9: Effect Heterogeneity and Out-of-Sample Targeting using ML

Panel A: Out-of-Sample Targeting

	(1)	(2)	(3)	(4)
	N. obs	Coef.	Dep Var. Mean	Effect (% Mean)
Full test sample	932,529	0.015***	0.05	28.46
		(0.004)		
Top quartile of predicted returns	232,783	0.031***	0.03	94.51
		(0.010)		
Bottom quartile of predicted returns	232,807	0.016***	0.12	13.79
		(0.006)		
Lift ratio (Top/Full)		2.09		3.32

Panel B: Effect Heterogeneity

Top 5 Selected Moderators	Rank	Standardized Coef.
Number of meetings	1	0.00131
Full-time worker	2	0.00064
Majority meetings are with low-income friends	3	-0.00062
Majority meetings are with large-circle friends	4	0.00040
IT, Software and IT service industry	5	0.00036

Notes: This table reports heterogeneous effect of face-to-face interactions and assesses the out-of-sample performance using machine learning methods. We augment the baseline model (Model 6) using LASSO with interactions between "Friend × EverMeet" and the full set of characteristics to capture rich heterogeneous effects. We train the model on an initial subsample of job switchers moving within the first nine months and evaluate it out of sample on a held-out testing partition of switchers moving in the last three months. Panel A reports the baseline estimates for subsamples partitioned by the LASSO-predicted returns on the held-out test sample. Column 2 reports the coefficient estimates and standard errors on Friend×EverMeet. Standard errors are two-way clustered by origin and destination block and reported in parentheses. Predicted-return quartiles are defined using out-of-sample predictions from the LASSO model estimated on the training sample. Column 4 reports the treatment coefficient relative to the mean of the dependent variable within each predicted-return bin. The lift ratio is defined as the ratio of the coefficient (or economic magnitude) in the top-quartile subsample relative to the full test sample. Panel B displays the top five moderators (interacted with Friend×EverMeet) selected by the LASSO procedure. These are presented as standardized coefficients and sorted in descending order by their absolute values. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table 10: Cross-Sample Prediction: Friend-Move Design

	(1) N. obs	(2) Coef.	(3) Dep Var. Mean	(4) Effect (% Mean)
Panel A: Friend move before focal switchers				
Full test sample	1,259,905	0.0001**	0.0003	32.02
		(0.00005)		
Top quartile of predicted returns	313,306	0.0002**	0.0004	46.51
		(0.00009)		
Bottom quartile of predicted returns	315,637	0.00001	0.0003	3.96
		(0.00009)		
Lift ratio (Top/Full)		1.69		1.45
Panel B: Friend move after focal switchers				
Full test sample	3,934,460	-0.00001	0.0004	—
		(0.00003)		
Top quartile of predicted returns	977,344	-0.00007	0.0004	—
		(0.00005)		
Bottom quartile of predicted returns	988,033	-0.00001	0.0004	—
		(0.00006)		

Notes: This table reports estimates from the friend-move design (Model 3) across out-of-sample targeting bins. The out-of-sample partition is based on the predicted returns generated from the LASSO model estimated in the training sample (see Table 9). Panel A presents results for cases in which the friend moves before the switcher. Panel B presents a placebo test in which the friend moves after the focal switcher. Standard errors are clustered by the focal individual's origin cell and the month of job switch and reported in parentheses. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table 11: External Validity and Out-of-Sample Targeting using ML

	(1) N. obs	(2) Coef.	(3) Dep Var. Mean	(4) Effect (% Mean)
Panel A: Train in high employment density areas, test in low employment density areas				
Full test sample	2,883,709	0.025***	0.06	42.69
		(0.003)		
Top quartile of predicted returns	644,102	0.053***	0.06	88.39
		(0.005)		
Bottom quartile of predicted returns	645,152	0.020***	0.13	15.57
		(0.003)		
Lift ratio (Top/Full)		2.10		2.07
Panel B: Train in subway-adjacent areas (≤1km), test in non-adjacent areas				
Full test sample	2,835,558	0.015***	0.02	63.20
		(0.003)		
Top quartile of predicted returns	708,833	0.029***	0.02	118.60
		(0.004)		
Bottom quartile of predicted returns	710,633	0.014***	0.04	31.95
		(0.005)		
Lift ratio (Top/Full)		1.97		1.88
Panel C: Train on private-sector workers, test on public-sector workers				
Full test sample	764,435	0.017***	0.05	36.50
		(0.004)		
Top quartile of predicted returns	190,925	0.033***	0.04	89.89
		(0.005)		
Bottom quartile of predicted returns	191,079	0.016*	0.10	15.69
		(0.008)		
Lift ratio (Top/Full)		1.99		2.46
Panel D: Train on workers from top 2 industries, test on others				
Full test sample	1,074,277	0.018***	0.06	30.08
		(0.003)		
Top quartile of predicted returns	268,285	0.032***	0.06	53.23
		(0.005)		
Bottom quartile of predicted returns	268,853	0.013*	0.13	10.23
		(0.007)		
Lift ratio (Top/Full)		1.79		1.77

Notes: This table reports baseline estimates (Model 6) across out-of-sample targeting bins constructed from alternative training-testing partitions defined along economically meaningful dimensions: working population size (Panel A), proximity to transit infrastructure (Panel B), sector of employment (Panel C), and industry type (Panel D). In Panel A, the training sample consists of destination blocks in the top 70 percent of the working population distribution, and the remaining 30 percent form the test sample. In Panel B, the training sample consists of destination blocks located within 1 km of a subway station, and the test sample comprises those located more than 1 km away. Panel C trains on workers in the private sector and tests on workers in the public sector. Panel D trains on workers in areas dominated by the two largest industries in the sample (wholesale/retail trade and water/environment/public facilities management) and tests on workers in all other industries. Standard errors are two-way clustered by origin and destination block and reported in parentheses. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Online Appendix

A Job Posting Occupations

We classify the occupation of each posting according to the job title, job description, and job category into one of 18 occupation categories based on the U.S. Census 2020 occupational classification system[1]. Below are detailed occupations we use:

1. Management Occupations – includes customer service manager, warehouse manager, production manager, hospital manager, human resource manager, CEO, retail shop manager and vice manager, sales manager, education administrator, etc.

2. Business and Financial Operations Occupations – includes accountants and auditors, financial and investment analysts, personal financial advisors, insurance underwriters, financial examiners, loan counselors and officers, tax examiners and revenue agents, tax preparers, budget analysts, credit analysts, property appraisers and assessors, compliance officers, claims adjusters and investigators, management analysts, human resource workers, meeting and event planners, and market research analysts and marketing specialists.

3. Computer and Mathematical Occupations – includes software developers, computer programmers, web developers, digital interface designers, computer systems analysts, information security analysts, database administrators, network architects, IT support specialists, data scientists, and computer research scientists.

4. Architecture and Engineering Occupations – includes civil engineers, mechanical engineers, electrical engineers, quality control technicians, surveyors, drafters, industrial engineers, and materials testing technicians.

5. Life, Physical, and Social Science Occupations – includes research scientists (biological, agricultural, and environmental), laboratory technicians, food scientists, nutrition specialists, environmental analysts, and biotechnology researchers.

6. Community and Social Service Occupations – includes social workers (child/family, healthcare, and mental health), community service specialists, counseling professionals (mental health, rehabilitation, and behavioral), and social/human service assistants.

[1] https://www.bls.gov/oes/current/oes_stru.htm

7. Legal Occupations –includes lawyers, judicial law clerks, paralegals/legal assistants, and other legal support workers.

8. Educational Instruction and Library Occupations – includes school teachers, university lecturers, vocational trainers, special educators, tutors, teacher assistants, curriculum developers, early childhood educators, and library professionals.

9. Arts, Design, Entertainment, Sports, and Media Occupations – includes graphic designers, video editors, photographers, web content editors, new media operators, directors, producers, actors, translators, interpreters, animators, sound technicians, lighting technicians, fashion designers, makeup artists, art directors, broadcast hosts, sports coaches, and post-production specialists.

10. Healthcare Practitioners and Technical Occupations – includes physicians, surgeons, nurses, nurse practitioners, physician assistants, pharmacists, physical therapists, occupational therapists, respiratory therapists, optometrists, opticians, clinical laboratory technicians, radiologic technologists, dental hygienists, dietitians, nutritionists, acupuncturists, emergency medical technicians, paramedics, and other specialized healthcare professionals.

11. Service Occupations – includes cashier, customer service, front desk, fire fighter, nail polisher, cleaner, massage, flight attendants, food server, cooks, laundry workers, counter attendants, security guards, and surveillance control workers.

12. Sales and Related Occupations – includes retail salespersons, wholesale representatives, insurance agents, real estate brokers, advertising sales agents, financial services sales agents, cashiers, product promoters, telemarketers, travel agents, models, marketing specialists, medical sales representatives, and other sales professionals across various industries.

13. Office and Administrative Support Occupations – includes administrative assistants, customer service representatives, data entry clerks, office clerks, receptionists, secretaries, human resources assistants, billing specialists, shipping/receiving clerks, inventory control staff, loan processors, and other general office support workers.

14. Farming, Fishing, and Forestry Occupations – includes agricultural workers, livestock breeders, crop growers, forestry technicians, fishing workers, agricultural inspectors, and other farming and cultivation specialists.

15. Construction and Extraction Occupations – includes construction laborers, electricians, plumbers, structural welders, construction equipment operators, construction and building inspectors, and other extraction workers.

16. Installation, Maintenance, and Repair Occupations – includes automotive technicians, heavy vehicle and mobile equipment service technicians and mechanics, industrial machinery repairers, electrical equipment installers, appliance repair technicians, general maintenance workers, and other installation, maintenance, and repair workers.

17. Production Occupations – includes machine operators, assemblers, welding, soldering, and brazing workers, production line workers, quality control inspectors, computer numerically controlled tool programmers and operators, woodworkers, metalworkers, food processing workers, packaging operators, and other manufacturing and industrial production staff.

18. Transportation and Material Moving Occupations – includes truck drivers, delivery drivers, warehouse workers, freight movers, stock handlers, shipping/receiving clerks, cargo agents, logistics coordinators, and other transportation and material moving workers.

B Data-Driven Heterogeneity via LASSO and Out-of-Sample Predictions

This section provides additional details on how we estimate data-driven heterogeneity in the returns to face-to-face interaction using LASSO regularization. Following the predictive framework of Mullainathan and Spiess (2017), we leverage the out-of-sample predictive power of LASSO to construct "predicted return" scores. This approach allows us to partition the testing sample into targeting quartiles and verify the existence of systematic heterogeneity, consistent with the "Generic Machine Learning" inference protocol proposed by Chernozhukov et al. (2025).

Specifically, we construct out-of-sample predictions of the marginal return of face-to-face interaction by isolating the treatment-interaction components of the fitted model. We then evaluate targeting performance in held-out samples both for our baseline workplace-choice design and the friend-move design.

B.1 Baseline workplace-choice design

Empirical setup. As described in Section 4, our baseline specification is:

$$M_{ij} = \alpha \text{Friend}_{ij} + \beta \text{Friend}_{ij} \times \text{EverMeet}_{ij} + \lambda_{b_i} + \lambda_{b_j} + \varepsilon_{ij} \tag{6}$$

where λ_{b_i} and λ_{b_j} denote the origin and destination block fixed effects. The coefficient of interest, β, captures the incremental return to face-to-face interaction conditional on the existence of a friendship tie.

Allowing for heterogeneous returns. Let X_{ij} denote a high-dimensional vector of moderators, including worker characteristics (gender, migration status, employment type, income level), job characteristics (public vs. private sector and 17 specific industries), and meeting characteristics (meeting frequency and composition variables as described in Section 6.2). We allow the return to face-to-face interaction to vary with these observables. We define the heterogeneous treatment effect as the marginal return:

$$\tau(X_{ij}) = \beta + \sum_k \theta_k X_{ijk} \tag{7}$$

Our objective is to identify the sparse set of moderators $\{X_k\}$ that characterize the environments where face-to-face interaction is most pivotal for job mobility.

Estimation via Orthogonalized LASSO. Adhering to the "honest" estimation framework (Athey and Imbens, 2016), we train the model on a partition of switchers moving within the first nine months of the sample. To identify the predictors of treatment heterogeneity while maintaining consistency with our baseline fixed-effects design, we implement the following procedure within the training partition:

1. **Partialling Out Fixed Effects:** We first residualize all variables – including the outcome (M_{ij}), the presence of friend indicator Friend_{ij}, the interaction treatment ($\text{Friend}_{ij} \times \text{EverMeet}_{ij}$), and the vector of characteristics (X_{ij}) – with respect to the origin and destination block fixed effects. This ensures that the subsequent selection is driven by within-block variation, consistent with the identification strategy in Equation 6.

2. **LASSO Selection:** We regress the residualized outcome on the residualized friend presence indicator, residualized treatment interaction, its interactions with X_{ij}, and

the residualized characteristics X_{ij}. We use LASSO regularization with 10-fold cross-validation to select the optimal penalty parameter. We include the residualized X_{ij} variables in the selection stage to "absorb" baseline differences in mobility propensities, ensuring that the interaction coefficients (θ_k) isolate marginal treatment returns rather than level effects.

Out-of-sample predicted returns and evaluation. We construct "predicted return" scores for a held-out testing sample (the final three months) using the coefficients estimated from the training sample:

$$\widehat{\tau}(X_{ij}) = \hat{\beta} + \sum_k \hat{\theta}_k res_X_{ijk} \tag{8}$$

Note that $\widehat{\tau}(X_{ij})$ represents the potential marginal gain from face-to-face interaction for a given pair. We use this index to rank observations independently of their actual treatment status. Given our workplace-choice empirical structure, we aggregate these scores to the origin-cell level by averaging across alternative workplaces in the destination block. We then rank observations in the testing partition by this average predicted return and define targeting quartiles. Finally, we re-estimate the baseline specification (6) within the top and bottom quartiles to observe the realized face-to-face effects.

B.2 Friend-move design

Cross-sample predicted returns and mechanism validation. We apply the estimated LASSO model to the friend-move design (an alternative testing sample) and compute out-of-sample predicted returns in this testing sample as:

$$\widehat{\tau}^{FM}(X_{ij}) = \hat{\beta} + \sum_k \hat{\theta}_k res_X^{FM}_{ijk} \tag{9}$$

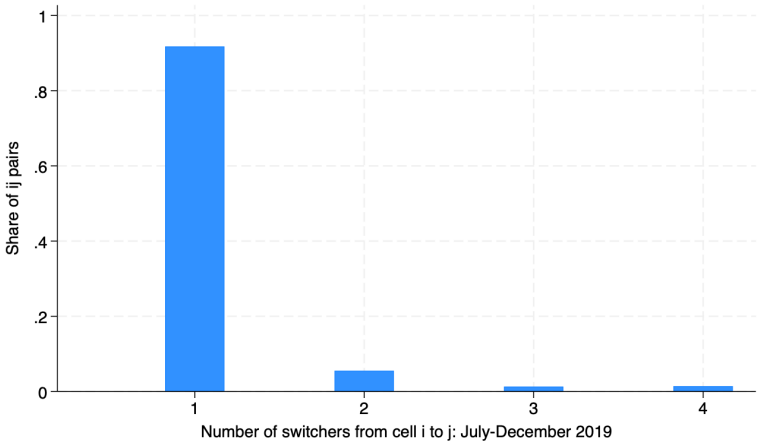
We then partition the friend-move sample into predicted-return quartiles and re-estimate Model (3) in Section 4.2 within each group.

We additionally evaluate predictions in pre-move placebo periods, where friends move after job switchers. Because referral-based information is not yet available in these periods, the model's predicted returns should not translate into differential realized effects. In our empirical results, estimated effects are concentrated in post-move periods and collapse toward zero in the pre-move placebo periods. This divergence supports the interpretation of the friend-move effects as operating through referral-related information transmission or access.

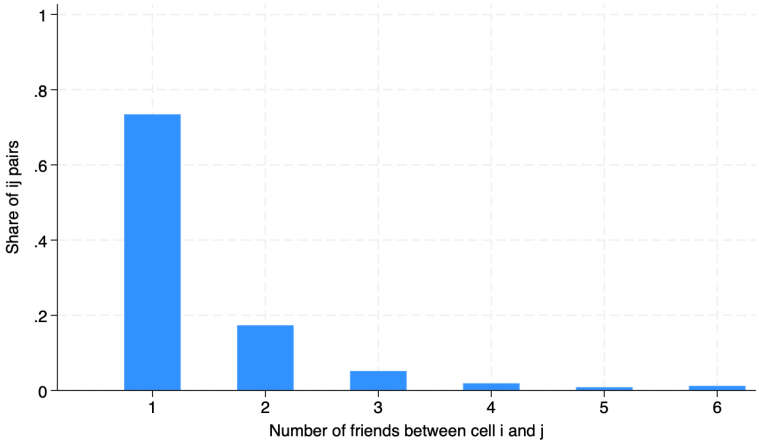
C Supplementary Figures and Tables

Figure B1: Distribution of Cell Pairs

(a) By Number of Switchers from i to j

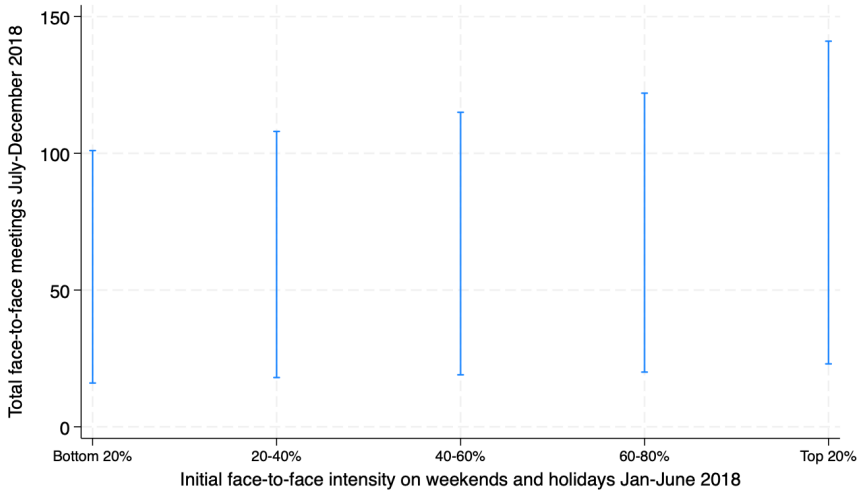


(b) By Number of Friends Between i and j



Notes: This figure plots the distribution of ij cell pairs by the number of switchers moving from cell i to cell j and by the number of friends between cells i and j in the baseline regression sample. In the regression sample, 4,582,775 out of 4,810,573 cell pairs (95 percent) have zero switchers, while 227,798 pairs (5 percent) have at least one switcher. Figure B1a plots the distribution among the 227,798 pairs with positive switchers: 91.75 percent have only one switcher, 5.5 percent have two, 1.3 percent have three, and 1.42 percent have four or more. Similarly, 4,366,320 out of 4,810,573 cell pairs (91 percent) have zero friends, while 444,253 (9 percent) have at least one friend. Figure B1b plots the distribution among these 444,253 pairs: 73.45 percent of them have one friend, 17.33 percent have two, 5.19 percent have three, and 4 percent have four or more.

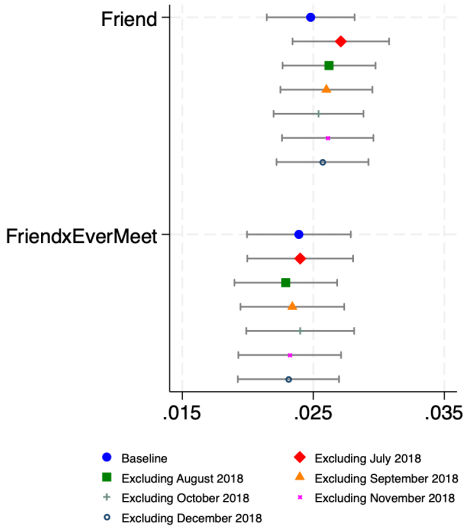
Figure B2: Interquartile Range of Face-to-Face Meetings by Initial Intensity



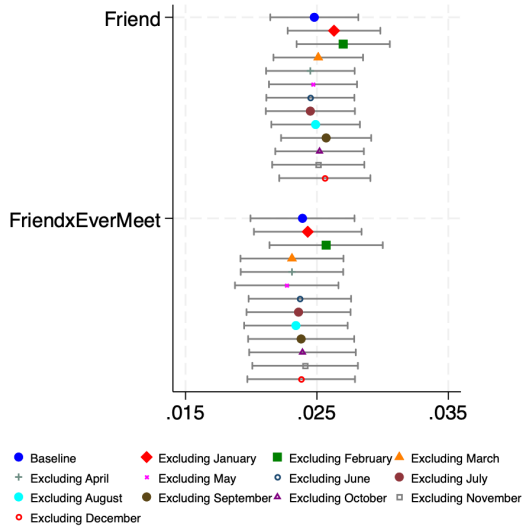
Notes: This figure shows the interquartile range of face-to-face meetings between switchers from i and their friends in j from July to December 2018, grouped by the initial number of meetings that occurred on weekends and holidays from January to June 2018. The sample is restricted to ij cell pairs with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

Figure B3: Leave-One-Out Sensitivity Analysis

(a) Removing Face-to-Face Interactions in Different Months



(b) Removing Switchers in Different Months



Notes: This figure presents robustness checks obtained by sequentially removing face-to-face communications or switchers in a given month. The figure plots the estimated coefficients on "Friend" and "Friend x EverMeet," along with their 95% confidence intervals. Panel B3a excludes face-to-face interactions occurring in specific months and reconstructs the key independent variables "Friend" and "Friend x EverMeet". Panel B3b excludes switchers in each month and their associated communications, reconstructing both the dependent and independent variables accordingly.

Table B1: Phone Calls with Friends and Non-Friends

Dependent variable:	(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: The unit of observation is a pair of individuals. The regression sample consists of random samples of 6,000 friend pairs and 6,000 non-friend pairs. The dependent variable in Column 1 is the average monthly number of calls between each pair during January-June 2019. Columns 2 and 3 report the average monthly number of calls during weekends and weekdays, respectively. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table B2: Effect of Face-to-Face Meetings on Job Location Choices: An Alternative Dependent Variable

Dependent Variable: Number of switchers from i to j	(1)	(2)	(3)	(4)	(5) Add Nonmovers
Friend	0.156***	0.0454***	0.0325***	0.0327***	0.0134***
	(0.0154)	(0.00304)	(0.00231)	(0.00232)	(0.00175)
Friend*EverMeet			0.0479***		0.0145***
			(0.00378)		(0.00237)
F2F meetings: Low				0.0343***	
				(0.00324)	
F2F meetings: Medium				0.0368***	
				(0.00333)	
F2F meetings: High				0.0737***	
				(0.00693)	
Observations	4,805,218	4,805,200	4,805,200	4,805,200	13,931,711
R-squared	0.030	0.205	0.206	0.206	0.046
Block i FE	No	Yes	Yes	Yes	Yes
Block j FE	No	Yes	Yes	Yes	Yes

Notes: This table examines the effect of face-to-face meetings on job changers' location choices, following the specification in Table 4 but with an alternative dependent variable. The unit of observation is an ij cell pair. The dependent variable is the number of switchers changing jobs from cell i to cell j. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table B3: Robustness: An Alternative Dependent Variable

Dependent Variable: Number of switchers from i to j	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Friend	0.0325***	0.00787***	0.0322***	0.0319***	0.00544***	0.0329***	0.0274***	0.0373***
	(0.00230)	(0.00216)	(0.00347)	(0.00230)	(0.00145)	(0.00225)	(0.00212)	(0.000877)
Friend*EverMeet	0.0479***	0.0439***	0.0429***	0.0382***	0.0342***	0.0483***	0.0452***	0.0465***
	(0.00375)	(0.00344)	(0.00494)	(0.00336)	(0.00365)	(0.00371)	(0.00360)	(0.00170)
Cell i total F2F (x1000)	1.16e-06							
	(9.03e-06)							
Number of meetings: initial stage		0.000426***						
		(7.78e-05)						
Days of meetings: initial stage		0.00329***						
		(0.000558)						
Number of workers at j					0.000349***			
					(4.42e-05)			
Number of friends in block j						-1.95e-05**		
						(7.69e-06)		
Observations	4,805,200	4,805,200	1,498,297	4,805,200	4,805,200	4,805,200	4,805,200	4,787,639
R-squared	0.206	0.207	0.136	0.206	0.216	0.206	0.208	0.177
Similar job opportunities nearby	No	No	Yes	No	No	No	No	No
Whether friend has same gender, migrant status and income group	No	No	No	Yes	No	No	No	No
Whether cells in same cluster	No	No	No	Yes	No	No	No	No
Amenities	No	No	No	No	No	No	Yes	No
Block i FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Block j FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Block i x j FE	No	No	No	No	No	No	No	Yes

Notes: This table reports robustness checks of the baseline results in Table 4, using the same specifications as in Table 6 but with an alternative dependent variable. The dependent variable is the number of switchers changing jobs from cell i to cell j. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table B4: Effect of Face-to-Face Meetings on Job Location Choices: Frequency and Composition of Meetings

Dep. Var. Probability of switching from i to j	(1)	(2)	(3)
Friend	0.025***	0.025***	0.025***
	(0.002)	(0.002)	(0.002)
Friend*EverMeet	0.024***	0.017***	0.0134***
	(0.002)	(0.002)	(0.003)
Observations	4,805,200	4,805,200	4,805,200
R-squared	0.23	0.23	0.23
Friend*Number of F2F meetings	No	Yes	Yes
Friend*F2F Meeting composition dummies	No	No	Yes
Block i FE	Yes	Yes	Yes
Block j FE	Yes	Yes	Yes

Notes: This table examines the effect of face-to-face meeting frequency and composition on job location choices, conditional on having at least one face-to-face meeting with friends. The unit of observation is an ij cell pair. Column 1 uses the preferred specification from Column 3 of Table 4. Column 2 additionally controls for "Friend" interacted with the number of meetings. Column 3 further controls for "Friend" interacted with meeting composition, including indicators for whether the majority of meetings of switchers in cell i and friends in cell j involve female, migrant, low-income friends or friends from different-sized social circles. All specifications include origin and destination block fixed effects. Standard errors are two-way clustered by origin and destination block and reported in parentheses. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table B5: Effect of Frequency and Composition of Meetings: An Alternative Dependent Variable

Dep. Var. Number of switchers from i to j	(1)	(2)	(3)
Friend	0.0325***	0.0327***	0.0328***
	(0.00231)	(0.00232)	(0.00232)
Friend*EverMeet	0.0479***	0.0275***	0.0176***
	(0.00378)	(0.00295)	(0.00447)
Observations	4,805,200	4,805,200	4,805,200
R-squared	0.206	0.206	0.206
Friend*Number of F2F meetings	No	Yes	Yes
Friend*F2F Meeting composition dummies	No	No	Yes
Block i FE	Yes	Yes	Yes
Block j FE	Yes	Yes	Yes
Block i x block j FE	Yes	Yes	Yes

Notes: This table reports results using the same specifications as in Table B4 but with an alternative dependent variable. The dependent variable is the number of switchers changing jobs from cell i to cell j. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table B6: Effect of Face-to-Face Meetings on Job Location Choices: Controlling for Number of Vacancies

Dependent variable: Probability of switching from i to j	(1)	(2)
Friend	0.025***	0.025***
	(0.002)	(0.002)
Friend*EverMeet	0.024***	0.024***
	(0.002)	(0.002)
Observations	4,805,200	4,805,200
R-squared	0.230	0.230
Number of vacancies at j	No	Yes
Block i FE	Yes	Yes
Block j FE	Yes	Yes

Notes: This table examines the effect of face-to-face meetings on job changers' location choices, following the specification in Table 4 and additionally controlling for the number of vacancies in cell j during July to December 2018. The unit of observation is an ij cell pair. *, **, and *** represent significance levels of 0.10, 0.05, and 0.01, respectively.

Table B7: Heterogeneous Effects by Gender and Migrant Status, Full Results

Dependent variable:	**Female**			**Migrant**		
	(1)	(2)	(3)	(4)	(5)	(6)
Probability of switching from i to j	No control	+Qty	+Comp	No control	+Qty	+Comp
Heterogeneity:						
Friend*EverMeet*X	-0.001	0.001	0.005	0.005*	-0.0003	-0.006
	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)
Meeting Quantity:						
Number of meetings		0.0001***	0.0001***		0.00003**	0.00003*
		(0.00002)	(0.00002)		(0.00002)	(0.00002)
Number of meetings*X		-0.00002	-0.00002		0.00004**	0.00004**
		(0.00002)	(0.00002)		(0.00002)	(0.00002)
Meeting Composition: Majority meetings with						
female friends			0.004			0.001
			(0.003)			(0.003)
migrant friends			0.007**			-0.001
			(0.003)			(0.004)
high-income friends			0.008***			0.008***
			(0.002)			(0.002)
large-circle friends			0.005			0.003
			(0.004)			(0.004)
female friends*X			0.001			0.006
			(0.004)			(0.004)
migrant friends*X			-0.007*			0.005
			(0.004)			(0.004)
high-income friends*X			-0.0002**			-0.0004***
			(0.0001)			(0.0001)
large-circle friends*X			-0.002			0.000083
			(0.004)			(0.005)
Observations	4,795,743	4,795,743	4,795,743	4,785,658	4,785,658	4,785,658
R-squared	0.230	0.230	0.230	0.230	0.230	0.230
Block i FE	Yes	Yes	Yes	Yes	Yes	Yes
Block j FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the full set of control variables for the heterogeneous effects by gender and migrant status presented in Table 8.

Table B8: Heterogeneous Effects by Employment Type and Income, Full Results

Dependent variable:	**Full-time**			**High Income**		
	(1)	(2)	(3)	(4)	(5)	(6)
Probability of switching from i to j	No control	+Qty	+Comp	No control	+Qty	+Comp
Heterogeneity:						
Friend*EverMeet*X	0.012***	0.009***	0.005	0.007***	0.004	0.004
	(0.002)	(0.003)	(0.005)	(0.002)	(0.003)	(0.005)
Meeting Quantity:						
Number of meetings		0.00004***	0.00004***		0.00005***	0.00005***
		(0.00001)	(0.00001)		(0.00002)	(0.00002)
Number of meetings*X		0.00003**	0.00003**		0.00002	0.00002
		(0.00002)	(0.00002)		(0.00002)	(0.00002)
Meeting Composition: Majority meetings with						
female friends			0.004			0.006*
			(0.004)			(0.004)
migrant friends			0.004			0.0001
			(0.004)			(0.003)
high-income friends			0.008***			0.008***
			(0.002)			(0.002)
large-circle friends			-0.003			0.004
			(0.004)			(0.004)
female friends*X			0.002			-0.002
			(0.004)			(0.004)
migrant friends*X			-0.003			0.003
			(0.004)			(0.004)
high-income friends*X			-0.001***			-0.0001
			(0.0001)			(0.0001)
large-circle friends*X			0.008*			-0.001
			(0.004)			(0.004)
Observations	4,791,002	4,791,002	4,791,002	4,800,901	4,800,901	4,800,901
R-squared	0.230	0.230	0.230	0.230	0.230	0.230
Block i FE	Yes	Yes	Yes	Yes	Yes	Yes
Block j FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the full set of control variables for the heterogeneous results by employment type and income presented in Table 8. The results indicate that only full-time workers exhibit a statistically significant differential effect for meetings with large-circle friends.