

Title: My neighbor next floor: the built environment and social preferences

Date: November 21, 2025

Authors: Marco Castillo (Department of Economics, Texas A&M University; CESifo Research Network; IZA; marco.castillo@tamu.edu, Ragan Petrie (Department of Economics, Texas A&M University; CESifo Research Network; rpetrie@tamu.edu) and Rong Rong (Department of Resource Economics, University of Massachusetts Amherst, rrong@umass.edu.)

Acknowledgements: We thank seminar participants at University of Amsterdam (CREED), University of Queensland (CUBES), Middlebury College (VEEP), University of Sydney, Chapman University, University of Southern California, Texas A&M University, Virginia Commonwealth University, Ohio State University, University of Virginia, Advances in Field Experiments Conference (AFE2023) and the 2nd Workshop on Field Experiments in Economics and Business for many helpful comments. The work was not possible without the support of Xiangdong Qin (Shanghai Jiao Tong University), many expert research assistants in Shanghai and a George Mason University Seed Grant. Special thanks goes to Patricia Petrie who introduced us to the work of Jane Jacobs (the inspiration for this research). The study has human subjects approval from George Mason University (IRB2010-6845). Castillo, Petrie and Rong declare no conflicts of interest.

Abstract: We assess the effect of the built environment on low-cost helping behavior toward neighbors. The setting is households in Shanghai, China that, due to rapid development, were involuntarily, i.e. as good as randomly, relocated to different building structures. Using a natural field experiment of misdelivered mail, we test the causal effect of spatial proximity and the built environment on helping behavior. Living one floor apart reduces the willingness to help a neighbor by 16 percentage points, similar to adding one more apartment per floor. Small spatial barriers can profoundly shape social interactions, and helping behavior, in urban settings.

JEL codes: D64, R23, C93

1 Introduction

The built environment can be designed to facilitate ease of movement, healthy living, and social interactions among its residents. It is found to be correlated with physical health (Coussens et al., 2005; Zhong et al., 2022), mental health (Evans, 2003), obesity (Papas et al., 2007), travel (Ewing and Cervero, 2010), crime (Jacobs, 1961; Newman, 1972) and social connections among neighbors (Festinger et al., 1950; Glaeser and Sacerdote, 2000). An important motivation for urban design is to attract diverse populations and foster positive social interactions (Hillier, 2007). Dense, mixed-use development and walkable streets are argued to make an area safe and prosperous (Jacobs, 1961), and sprawl isolates individuals and reduces social cohesion (Putnam, 2000). It is important to understand how the built environment affects social interactions and the willingness to help neighbors, as this affects the nature of civil society and how cities and buildings are designed. However, inference is hampered by self selection into geographical location, neighborhoods and building structures (Krzizek, 2003). This paper addresses these issues with a field experiment and involuntary, i.e. as good as random, relocation into different residential structures to examine how the built environment and spatial proximity affect low-cost helping behavior among neighbors.

We use a behavioral measure of kindness towards a neighbor. Our field experiment examines whether individuals are willing to help a neighbor living in the same residential building by returning a misdelivered piece of mail. This is a natural field experiment (Harrison and List, 2004), in that those who receive the misdelivered envelope and those who should have gotten the envelope do not know they are participating in a study. The cost to return the envelope is very low, as all mailboxes are centrally located at the main entrance to the apartment building. The neighbor merely needs to move the misdelivered mail from their mailbox into the recipient's. This simple act of kindness speaks to how neighbors treat each other and serves as a measure of social cohesion.

Our setting is Shanghai, China, where roughly 12.4% of residential housing units were resettled into buildings of various sizes from 2000-09, during a period of rapid redevelopment

in the city.¹ Most households had only one or two housing options available at the time of relocation, and they ended up in tall or short buildings with few or many apartments per floor. We provide evidence that housing assignments were as good as random, both across tall and short buildings and location within a building.

Our experimental design randomizes the spatial proximity of the target neighbor whose mail is misdelivered and the subject neighbor who gets the mail. These neighbor pairs are either next door to one another or located one floor apart. All envelopes, except one treatment, misplaced in the subject’s mailbox have the target’s name and address. In three treatments, we experimentally vary the urgency of the mail through the use of postal and courier services. In a fourth treatment, we raise the cost to return the envelope by using the target’s name and the subject’s address. In this case, the subject would need to know where the target actually lives to return the envelope. These four treatments allow to stress test the willingness to help as costs and benefits of returning the envelope change.

The mail was privately misdelivered by us to insure subjects and targets reacted naturally and reports of receipt, or non-receipt, of envelopes was due to the actions of the neighbor. The envelope contained a letter from Shanghai Jiao Tong University outlining a research study on mail delivery services that the research team would come to collect the following week for payment. Only target households were surveyed for receipt of the letter. We did not contact subjects. Care was taken in the experimental design to detect false negative and false positive reports of envelope receipt. There are no false positive reports, as expected, and we account for false negatives in our analysis.

Evidence that housing assignments are as good as random allows for causal interpretation along two dimensions. First, one concern might be that previous neighbors are placed in apartments together or next to one another. We find no evidence for this. Our experimental design can then test the causal effect of spatial distance of living quarters on helping behav-

¹From 2000-09, 630,624 residential units were resettled in Shanghai, and there were 5.1 million households registered in 2009 (SSB, 2010).

ior.² Second, given that we find no evidence for assignment to buildings based on personal characteristics or previous residence, we can then use building assignment to test the causal effect of the built environment on helping behavior.

Treatment assignment in our study avoids interference across target-subject pairs by design, a concern in many field experiments (Rubin, 1978; Athey et al., 2018). No target household could be a subject household or vice versa, and no target or subject were assigned to more than one partner household. We also minimized the number of pairs within a building (average 1.95 pairs per building). While interference is avoided, it is at a cost, as it can prevent probabilistic assignment, a necessary condition for unbiased estimation of treatment effects (Imbens and Rubin, 2015). Once a target household is designated, it cannot be a subject household, and a subject household cannot be a target. Thus, our design impedes the use of properties of random assignment on potential outcomes for inference. To address this, we develop a randomization test that accounts for bias that emerges when interference is avoided by design. We embed this test in Athey et al. (2018)’s framework.

Our findings show that the built environment and spatial proximity of residences significantly affect willingness to help neighbors. When a neighbor lives one floor above, instead of next door, 0.21 fewer envelopes are returned (on average, 1.1 envelopes, out of two, are reported returned). Recall that the cost to return an envelope to a neighbor on any floor is the same. All mailboxes are centrally located, and they are ordered so that the mailbox of the neighbor next door and next floor are both adjacent to the subject’s. When the cost to return an envelope increases, in the fourth treatment, because the recipient’s name and address are mismatched, 0.47 fewer envelopes are returned. We further examine the effects

²Spatial distance can affect social interactions and relationships. Indeed, the “proximity principle” in social psychology posits that people have a tendency to form relationships with those who are close by (Newcomb, 1960; Ebbesen et al., 1976; Festinger et al., 1950). Spatial proximity has also been found to affect donations, i.e. increasing as real geographical distance decreases (Touré-Tillery and Fishbach, 2017).

of the built environment on helping behavior, including a rich set of controls for household characteristics, relocation status, current and previous residence characteristics and community. Having one more apartment per floor in a building is equivalent to living next door, but in the opposite direction. An additional apartment on the floor reduces the likelihood of returning an envelope by 18 percentage points, while being the neighbor next door increases the likelihood by 16 percentage points.

The effect of spatial proximity manifests in reported social interactions as well. Respondents are more likely to nod and smile at their neighbor and less likely to consider the neighbor to be a stranger when the neighbor lives next door. We conclude that the spatial layout of residential structures and the location of neighbors with respect to one another has a profound effect on the willingness to engage in small acts of kindness towards one's neighbors. Using mediation analysis, when neighbors live next door to one another, they engage in more social interactions, thus increasing familiarity, and this drives helping. The built environment, on the other hand, has a direct effect on helping a neighbor and is not mediated by social interactions.

Robustness checks that account for sample selection confirm our main results that one more apartment per floor reduces the likelihood of helping a neighbor by 12-26 percentage points, depending on the specification. We further explore the effects of the built environment on social network formation. Buildings with fewer apartments per floor lead to networks that are more dense with familiar links. These types of networks can support informal cooperation ([Jackson et al., 2012](#)) and support our finding of a higher willingness to help a neighbor in a built environment with fewer apartments per floor.

The large city setting for our study and our findings speak to the urban studies literature on the layout of cities, residential housing and social interactions. While it is well established that relationships are influenced by social proximity ([Bochner et al., 1976](#); [Wellman, 1979](#)), spatial proximity is also influential. In city settings with high-rise buildings, there is a larger pool of people to draw from for friendships ([Churchman, 1999](#)), and the spatial layout of

buildings and cities are correlated with social relationships and casual encounters ([Jacobs, 1961](#); [Skjaeveland and Garling, 1997](#); [Gehl, 2013](#); [Mazumdar et al., 2018](#)), social isolation ([Gifford, 2007](#); [Kearns et al., 2012](#)), perceptions of mutual trust and willingness to help others ([Cohen et al., 2008](#)), professional collaborations ([Wineman et al., 2009](#)) and crime ([Newman, 1973](#)). Our findings can confidently speak to the causal effect of spatial proximity and the built environment on behavior by reducing the confounds of self selection inherent in correlational research. We show the influence of spatial proximity within city residential buildings. Closeness, created by constraining the pool of neighbors on a given floor (i.e. fewer apartments) and living next door, increases the willingness to help a neighbor.

Our paper relates to findings that spatial proximity generated by office or university dorm assignment correlates with work-related outcomes, such as the gender-pay gap ([Cullen and Perez-Truglia, 2023](#)) and scientific collaborations ([Salazar Miranda and Claudel, 2021](#); [Boudreau et al., 2017](#)), as well as the formation of friendships ([Festinger et al., 1950](#); [Marinos and Sacerdote, 2006](#)). Merely crossing paths with others increases familiarity ([Milgram, 1977](#)), as does spatial proximity in residential buildings ([Felder et al., 2023](#)), and repeated casual contact among strangers on the street promotes support for more inclusive social policies ([Bollen, 2023](#)) and may reduce crime ([Zahnow et al., 2021](#)). Our findings establish the separate direct effect of the built environment and indirect effect of spatial proximity through enhanced social interactions on helping a neighbor.

As established social connections enhance trust ([Glaeser et al., 2000](#)) and altruism ([Leider et al., 2009](#)), neighbors who engage in social interactions may be more inclined to help one another as well. Learnings from psychology and biology provide insights for this pathway. At a fundamental level, altruism is motivated by the desire to help one's progeny to ensure its continued existence ([Hamilton, 1964](#); [Eisenberg et al., 2013](#); [Batson and Powell, 2003](#)), and people are more likely to feel empathy and help those they perceive to be similar to themselves ([Krebs, 1975](#)). Prosocial motivation, however, is not confined to family, as it also emerges among non-kin as direct reciprocity ([Trivers, 1971](#); [Nowak, 2006](#)). In our setting, increased

social interactions among neighbors within a building, brought on by spatial proximity and the built environment, may produce empathic feelings that then lead to a greater willingness to help a neighbor.³

Finally, our paper makes a methodological contribution with the development of a new re-randomization test that addresses an issue of treatment assignment. There is a robust literature that has developed methods to address inference when interference cannot be avoided (Rubin, 1978; Athey et al., 2018). By avoiding interference with our randomization procedure, inference based on large-sample properties is biased, and traditional p-values are incorrect in our setting. The methodology developed in the paper addresses this and could be used in other settings to obtain exact p-values.

The paper is organized as follows. Section 2 describes the redevelopment and relocation of households in Shanghai, China. Section 3 describes the field experiment. Section 4 derives our estimation approach when avoiding interference. Section 5 describes the sample and evidence for as good as random assignment to the built environment. Section 6 presents results, robustness checks, mediation and network structure effects. Section 7 concludes.

2 Shanghai development and relocation

Shanghai is one of the most populous cities in the world, and the most populated in China, with over 24 million inhabitants in the municipality in 2020 (NBSC, 2022). While today Shanghai is considered a world-renowned financial center that showcases China’s booming economy, its urban landscape looked different thirty years ago. In the late 1970’s and 80’s, a series of sweeping economic reforms were introduced in China by Deng Xiaoping which initiated China’s move away from a planned economy to a market-oriented one. These policy reforms decentralized decision-making, giving more power to local governments and state owned enterprises, and changed land use rights systems. These reforms precipitated

³The extent to which one is altruistic towards in- and out-groups, or “moral universalism,” is not universally constant and varies by sex and age (Enke et al., 2022).

several policies by the Shanghai Municipal Government aimed at urban renewal of the city in the 1990's. For instance, one program had the goal of redeveloping 365 hectares of urban area, primarily consisting of old houses.⁴

As a result of these programs, the city experienced rapid and intense development in the city center and outskirts. Buildings in the inner city were removed, and many residents were relocated to new developments on the outskirts of the city (Wu, 2004b,a).⁵ From 1993-2000, 17.7% of the total building stock of Shanghai from 1990 was demolished (Zhu, 2004). To put the rate and magnitude of the redevelopment in Shanghai in perspective, between 1991 to 1995, 300,000 household were relocated, a rate 2.4 times that of the previous decade (SCG, 1996; Shih, 2010). And that number would continue to grow, to 660,000 by 1999 (Yao and Jiang, 2002; Wu, 2004b,a) and over one million households by 2006 (SSB, 2007; Shih, 2010).

From 1992-1999, 48% of intra-urban housing moves were due to relocation from demolished structures (Li and Song, 2009). The relocations were often involuntary. Surveys of intra-urban Shanghai households conducted in 2000 and 2001 found that 58.3% of relocations were involuntary, with 59.2% resulting directly from real estate or infrastructure development (Wu, 2004b,a). Many of these households had only one available alternative for new housing.

Households were also asked to complete the relocation in a short period of time. It was not uncommon for residents to learn of a relocation plan after the land had already been leased to developers (Shih, 2010). Modes of compensation for relocated households varied, but typically households were offered alternative living accommodations (i.e. an owned apartment), monetary compensation or a combination. The welfare effects of these relocations are somewhat mixed. The quality of housing and amenities tend to be better (Li and Song, 2009), but residents located further from the city center face longer commutes and

⁴The urban extent of Shanghai in 1991 was 195,581 hectares (AUE, 2023).

⁵From 1980-90, 96% of land development was outside the city center (Sun and Deng, 1997; Zhu, 2004).

less job accessibility (Day and Certero, 2010; He, 2010). There is no evidence that housing demolition had adverse political effects during the period 2010-2018 (Sha and Zou, 2023).

The nature of these relocations placed individual and groups of households in a variety of different buildings, in terms of building height and number of apartments per floor, with limited or no options to choose from. We use this variation in built environment to examine its effect on willingness to help a neighbor. Crucial to the strength of our identification is that the built environment in which the relocated resident was placed is uncorrelated with other factors that might explain helping. A small percent (8.5%) of the neighbor pairs in our sample shared the same previous address, however, there is no evidence that this is correlated with the built environment or our experimental design. In Section 5.2, we present comprehensive evidence that relocation in our sample into the built environment is as good as random. Before we do so, we describe the field experiment and estimation approach.

3 Field experiment

The objective of the field experiment is to see how the built environment affects low-cost helping behavior among neighbors. The guiding premise for the design is that establishing social relationships carry a cost. If the built environment adds to that cost, i.e. by creating spatial distance between neighbors or increasing the number of neighbors on an apartment floor, the built environment will impact relationship building and thus the willingness to help a neighbor.⁶

To that end, we conducted the experiment in 2010 in residential communities with recently, and majority, relocated households, who were thus unlikely to have selected the built environment in which they live. Experimental variation is introduced by randomizing the spatial distance within the apartment building between the person who could be helped (the target) and the person who could help (the subject). Another source of experimental variation is four treatments that vary the characteristics and urgency of the misdelivered mail

⁶We discuss a formal model and predictions in Appendix C.

the subject could return to the target.

3.1 Sample selection of communities

To identify housing communities with relocated households, we first conducted a phone survey with a large sample of the population of community offices in Shanghai.⁷ Each housing community typically has a central office that is responsible for community upkeep and tracking residents (e.g. household size, job status). We identified 3,917 community offices in Shanghai from a 2005 list that covered 18 districts and 168 sub-districts. A stratified random sample by district was drawn to obtain a subsample consisting of 864 community offices to be surveyed. We called each of these offices by phone and made several attempts to call back if there was no answer. This resulted in survey responses from 546 community offices (63% response rate). The survey determined the buildings in the community, the community population and the relocation status of its residents.

An ideal experiment to uncover causal effects of the built environment on behavior would randomly assign residents of a community to various building structures, they would not move and there would be variation in the architectural style of buildings, i.e. building height and apartments per floor. In practice, this experiment is difficult to run because households typically have a say in where they live. Housing location would then be correlated with unobservable characteristics of the household. In Shanghai, however, due to rapid re-development, many residents were involuntarily relocated to new apartments in buildings of various sizes and had one or two options available. We use involuntary relocation to strengthen our identification of the built environment on behavior (see Section 5.2). Candidate communities for our study are majority relocated, and there is a mix of short and tall buildings.⁸ These criteria give the best chance to test the effect of variation in the built environment on behavior.

⁷The community phone survey, field experiment and household survey were conducted June-September 2010.

⁸Another criteria was that sample communities are within 30km from the city center.

From the over 200 communities that met our criteria, we then drew a random sample to contact in order to obtain seven communities for the study.⁹ The community officer of the candidate community is contacted by phone and provided a brief description of the study, i.e. a research study on urbanization and community conducted by Shanghai Jiao Tong University. In exchange for participation, the officer receives a personal gift from us upon agreeing to participate, and a lump sum payment of 300 RMB is made to the community office account upon completion of the study. The officer also receives 530 RMB as compensation for providing access to apartment mailboxes for the field experiment and assistance with the household surveys.¹⁰ If the community officer shows interest in participating in our study, we arrange an appointment to meet with the officer in person to explain the study details.

At the in-person meeting with the community officer, we present an introductory letter from Shanghai Jiao Tong University. If the officer consents to the study, we complete a survey of the community and building characteristics and obtain a map of the community layout. The building survey collects information about residential buildings in the community (e.g., number of buildings, building height, year built), and we take photos of the buildings to document the built environment. The community officer agrees to assist the research team and keep project information confidential.

3.2 Design

The field experiment provides a behavioral measure of willingness to help a neighbor in a natural setting. An envelope with the target recipient's name and address is misdelivered,

This was done to ease fieldwork logistics and minimized study costs.

⁹A sample of 25 was drawn and put in random order to contact. If the community did not ultimately fit our sampling requirements or the community officer declined participation, we moved on to the next community until we secured our target of seven communities. Figure [A.1](#) shows a map of our sample. Budgetary constraints prevented us from running the field experiment in more than seven communities.

¹⁰At the time of the field work in 2010, 1RMB=US\$0.1474.

by us, to the subject’s mailbox, where both the target and subject live in the same building. This is a natural field experiment (Harrison and List, 2004), and neither the target nor the subject know they are in a study. Mailboxes are centrally located at or near the main entrance of the building with slots to insert mail, and they are configured such that mailboxes of the apartment next door and floors above and below are adjacent, independent of the apartment building size.¹¹ Thus, the cost to return the letter to any neighbor in the building is very low, and the cost to return a letter to a proximate neighbor, i.e. next door or next floor, is the same in small and large buildings. The outcome of interest is whether the subject returns the letter to the target.¹²

Two aspects are varied experimentally: the spatial distance between the residences of the target and subject and the characteristics of the misdelivered envelope. Randomization of spatial distance is done at the apartment level within each building. Apartments are randomly assigned the role of target. A subject then is randomly chosen to be within one apartment of the target, either from the apartment on the same floor to the right of the target or on the floor above in the apartment directly above the target or to the right.¹³ If a randomly chosen subject has already been assigned to be a target, the subject is discarded and another subject is randomly chosen among the remaining. On average, there were 1.80 target-subject pairs (s.d. 1.91, min 1, max 12) in each building, and the average number of apartments per building was 56 (s.d. 60.2). An average of 8.0% (s.d. 4.1%) of apartments were treated in any given building, and treatment saturation within a building is minimized to reduce spillovers across pairs within a building. This procedure results in paired sets

¹¹Photos of typical building mailboxes are in Figure A.2.

¹²As in most countries, it is illegal in China to conceal or open another person’s mail. However, there are no records of misdelivered mail, so this is difficult to enforce. Given the scant risk of being discovered, we interpret returning the letter to the neighbor as a prosocial act.

¹³In the case the target is the end apartment, the apartment on the same floor to the left and above to the left are used.

of apartments that are either a target or a subject, but not both, and insures there is no interference in treatment assignment.

Once the pairs are assigned within a building, we randomly assign pairs to one of the four letter treatments, such that the four treatments are equally represented within and across buildings within a community. We include additional households in a placebo treatment to test for false positives in reporting. In the placebo treatment, no envelope is misdelivered, or delivered, however we survey the placebo households, at the same time we survey the target households, and ask if they received a letter. No household in the placebo treatment claimed to receive a letter.

The four envelope treatments vary the outside of the envelope and the service used to deliver the letter, i.e. regular postal or courier service. Each envelope includes the return address of Shanghai Jiao Tong University, and the recipient’s name and address are hand written on the front of the envelope. Inside the envelope is a letter on university letterhead outlining a research project to test mail delivery services. The recipient is asked to keep the letter, and the research team will come to retrieve it in ten days for a payment of 10 RMB.¹⁴ For each treatment, we also place an identical envelope directly in the target’s mailbox. This allows for identification of false negatives in reporting.

All envelopes are deposited in mailboxes by the research team. This insures that both subjects and targets do receive the envelopes in their mailboxes, and envelope reports reflect actions of subjects not mail services. Delivery is done discretely in the afternoon when few residents are around. Ten days later, the envelope survey team returns to the community and knocks on the target household’s door to ask if they received the envelopes. Subject households are not contacted for the envelope survey. Target households should report either receiving one or two envelopes, with the latter reported if the subject returns the misdelivered

¹⁴The text of the letter is in Appendix B.1. The envelope is a standard, white, opaque envelope. Each letter contained a unique code that we checked against our records, thus eliminating the possibility of manufacturing a letter for payment.

mail. If the target indicates that they received an envelope(s), we ask to see the letter(s) as confirmation and pay the target 10 RMB cash per letter shown. If the target has not checked their mailbox, we ask to accompany them to check it. During the same survey, placebo households are also asked if they received an envelope. The envelope survey team is blind to treatment assignment.

[Table 1 about here]

The four envelope treatments, plus the placebo, are summarized in Table 1. The first treatment (T1) is an envelope sent via regular China Post. The second (T2) is sent via express courier. Courier services charge a higher price than China Post and guarantee a fast and secure deliver. The third (T3) is identical to T2 but with an “urgent” stamp. The fourth (T4) is sent via regular China Post but with the target’s name and the subject’s address, thus the subject would need to know where the target lives to return the envelope.¹⁵ The envelope treatments were designed to stress test the willingness to help a neighbor by varying costs and benefits. T2 and T3 signal importance of the letter because the sender of the envelope paid more to get the letter to the target. We expect this would increase willingness to help compared to T1, even in the absence of a relationship between the target and subject. T4 raises the cost for the subject to return the envelope. We expect this would decrease the willingness to help compared to T1, except if the subject knew the target.

In total, 338 target households were treated across the four treatments with two envelopes each, one deposited in the target mailbox to test for false negatives and one deposited in the subject mailbox to test for helping behavior. In addition, to test for false positives,

¹⁵Photos of treatment envelopes are in Figure A.3. All envelopes but T4 included the target’s name and address as the recipient. The T1 envelope included a China Post postage stamp and a date processed stamp. The T2 envelope included a courier receipt common to courier services in China. The T3 envelope was the same as T2 and included an “urgent” stamp on the front of the envelope. The T4 envelope was similar to T1 but had the target’s name and the subject’s address.

107 households were in the placebo treatment that received no envelope.¹⁶ In terms of spatial distance between the subject and target, 38% of targets were next door, 35% in the apartment directly above and 27% in the apartment not directly above. Treatment assignment is balanced on covariates (Table A.2).

3.3 Household survey

In addition to the field experiment of misdelivered mail, we conduct a household survey with target-subject pairs. The household survey is carried out by a completely different survey team after the envelope survey team asked target households if they received envelopes. The household survey is administered separately to target households and subject households. The purpose of the household survey is to understand urban community and takes approximately ten minutes to complete (see Appendix B). The survey asks the head of household basic demographic information, relocation status and options, and previous housing characteristics. In addition, we collect information on social interactions with four randomly chosen neighbors located either in the apartment next door, in the same apartment number or the apartment next door number to the respondent on floors above or below, and we make sure to include the target-subject pair.¹⁷ We ask if the respondent considers a particular neighbor a stranger and smiles or nods at the neighbor.

¹⁶More households were ex ante assigned to the four treatments (~ 500) than were ex post treated. This is because we could only verify the apartment was occupied and the mailbox was secure, and not broken, after we arrived to deliver the mail.

¹⁷Specifically, we sample from those who share the same apartment number x as the respondent (i.e. $20x$) and are located on a floor above (i.e. $30x$) or below (i.e. $10x$), as well as from those who are in the apartment to the right of the respondent, i.e. $(x + 1)$ and are located on a floor above, same floor and floor below. If a respondent is located in a corner apartment such that there is no apartment to the right, we sample from the apartment to the left ($x - 1$).

3.4 Identification of biased reporting

Two features introduced in our experimental design help to address concerns of false positives and false negatives in reporting receipt of the envelope. First, we randomly selected placebo households, who were not target-subject pairs, and asked the household if they received any mail. Since none of the placebo households reported receiving the envelope, we conclude that we do not have a problem with false positives. Second, to address the possibility of false negatives, we delivered a second identical treatment envelope to each target household. This allows us to measure bias introduced by non-response since each target household should report receiving at least one envelope. In Section 6, we provide treatment effect estimates of spatial distance that account for false negatives. In particular, we estimate treatment effects in the presence of non-response by assuming that non-response is independent of treatment assignment.¹⁸

4 Estimation approach when avoiding interference

Any experiment is an assignment problem. Our experiment avoids interference by design, i.e. no target can be a subject, no subject can be a target and one target is assigned to one subject. However, probabilistic assignment is not guaranteed to hold, a necessary condition for unbiased estimation of treatment effects (Imbens and Rubin, 2015). We derive the correct randomization tests for the sharp null hypothesis of no built environment effects. These tests account for our experimental design and are used in Section 6 to examine treatment effects from the field experiment. The full derivation and procedure of the tests are in Appendix D.

We illustrate the assignment problem. Let N be the set of apartments in a building. For target apartment $i \in N$, we denote the neighbors of i by $g(i)$. In our experimental setting, neighbors are apartments that have the same number on the floor above apartment i , or are

¹⁸Non-response in binary dependent variables produces non-classical errors on the left-hand side of a regression (see Hausman, 2001). We discuss our estimation strategy to deal with non-response bias in Section 6.

the next apartment to i on the same floor or the floor above that apartment. Figure 1(a) represents neighbors of target apartment 101 according to this definition.

[Figure 1 about here]

Based on this information, we now construct a bipartite graph (N, h) . Let T be a set of targets such that $T \subset N$. Let $S = N \setminus T$ be the set of subjects. We let $(i, j) \in h$ if $i \in T$, $j \in S$, and $j \in g(i)$. In this framework, no interference is equivalent to the existence of an assignment of each target in T to one subject in S . This is so because no target is assigned to each other, and no target shares the same subject. Figure 1(b) illustrates a possible assignment based on our assignment protocol that does not violate no interference.¹⁹ Figure 1(c) illustrates a possible assignment that violates no interference. A necessary condition for an experimental design to satisfy no interference is for the number of targets to be less than the number of subjects for each possible set of targets.²⁰

We construct a test of the sharp null hypothesis of no treatment effects. We abstract from the possibility of stratification and use of covariates for simplicity. Start with (N, h) , a bipartite graph constructed as above such that the target group T has at least one complete assignment. This graph is constructed by randomly choosing a set of targets and defining those not in T as subjects. These assignments cannot be found for all subsets of N since targets cannot be subjects.

We are interested in the effect of a treatment assignment, $W \in \{0, 1, 2\}^{|T|}$, on outcome y . Outcome y equals one if the envelope is returned and zero if it is not returned. Let $y_i(W)$ denote the i -th unit's, $i = 1, \dots, |T|$, potential outcome under treatment assignment $W = (W_i, W_{-i})$, where W_i is the treatment received by i and W_{-i} are the treatments received by other targets. In our study, we have three possible treatments. Let,

¹⁹This assignment would not be used in practice because we want target-subject pairs located on the same floor as well as different floors.

²⁰This is Hall (1935)'s theorem for the existence of assignments in bipartite graphs.

$$W_i = \begin{cases} 0, & \text{if same floor} \\ 1, & \text{apartment exactly above} \\ 2, & \text{apartment not exactly above} \end{cases}$$

Let $Y_{obs(W)}$ denote the vector of observed outcome values for a given treatment assignment W :

$$Y_{obs(W),i} = y_i(0, W_{-i}) \times \mathbf{1}[W_i = 0, W_{-i}] + y_i(1, W_{-i}) \times \mathbf{1}[W_i = 1, W_{-i}] + y_i(2, W_{-i}) \times \mathbf{1}[W_i = 2, W_{-i}] \quad (1)$$

A sharp null hypothesis of no treatment effect assumes that $Y(W) = Y_{obs(W)}$ for any feasible treatment assignments W . The interference problem outlined above suggests that this sharp null hypothesis is violated. That is, potential outcomes are not constant across the treatment assignment of neighbors or their neighbors. The following remark shows conditions for the sharp null hypothesis to hold and be testable using a permutation test.

Remark 1. *A necessary and sufficient condition for the sharp null hypothesis $Y(W) = Y_{obs(W)}$ for all W to hold is for no two targets in T to share a neighbor or be neighbors.*

Proof: See Appendix D.

Remark 1 implies that if an experiment respects no interference but not all permutations are assignments, we cannot conclude that a randomization test is a valid test of the sharp null hypothesis $Y(W) = Y_{obs(W)}$ since the distribution of counterfactuals might not be invariant under these transformations. We build on [Athey et al. \(2018\)](#), who propose creating artificial experiments where a sharp null hypothesis can be tested. Intuitively, in our experiment, the sharp null hypothesis $Y(W) = Y_{obs(W)}$ can be tested using a randomization test on a subset of the original experiment where Remark 1 is valid. The most powerful test would be conducted in the largest subset of the original experiment in which Remark 1 holds. We

use this randomization test in Table 4 to test the sharp null of no treatment effects. The derivation and procedures of our randomization test are found in Appendix D.

5 Description of sample

5.1 Buildings and households

Within our sample of seven communities, there are 429 buildings. Table 2 reports characteristics. Buildings are an average height of 7.74 floors, with a minimum of 5 and a maximum of 24 floors. There are 2.39 apartments per floor, with a minimum of 2 and a maximum of 8. There is usually only one entrance per building, and 21% of buildings have an elevator. Buildings were constructed between 1988 and 2009. The average percent of relocated households in each building is 76%, with some buildings containing no relocated households and some containing all. The high proportion of relocated households reflects our sampling strategy of choosing communities with a majority of relocated households.

[Table 2 about here.]

Household characteristics are reported in Table 3 and are based on information from the household survey conducted with target and subject households. The majority of households (59%) report being relocated to their current apartment.²¹ The average age of the household head is 48.1 years, and half are male. Households moved into their residence around 2004, with relocated households moving in around 2003 and not relocated households moving around 2006. Relocated households moved from shorter buildings (2.4 floors, p-value < 0.001) with fewer apartments per floor (1.5 apt per floor, p-value < 0.001) than not relocated households.

[Table 3 about here.]

Most relocated households are owned (78%), and 83% had two or fewer housing options to choose from when they moved to their current residence. Ninety-seven percent of relocated

²¹Three survey respondents report relocation because of work assignment. Seven percent report a previous forced relocation, i.e., before the most recent move.

households were offered housing, and 57% were offered cash.²²

5.2 As good as random assignment to built environment

We oversampled communities with relocated residents to increase the chance that residents did not choose the built environment in which they live, and thus, we could plausibly interpret estimates of the effect of the built environment on helping behavior as causal. We provide evidence supporting this interpretation.

During the rapid redevelopment of Shanghai, residents from demolished communities were relocated to new communities and may have been relocated together. A concern then is that these residents bring with them their social networks and are more likely to help each other in the absence of any effects of the built environment. Also, even though we randomly assigned subject-target pairs within a building, there may be a correlation between our spatial proximity treatment assignment (i.e. next door or next floor) and previous relationships. We examine these issues.

First, we examine whether households in our sample moved from the same community (see Table A.3). Of the subject-target pairs in our sample, 8.5% share the same previous address.²³ Importantly and relevant for our analysis, subject-target pairs from the same previous address are no more likely to live next door to each other or on the next floor, nor is there a correlation between sharing a previous address and the built environment of their current residence. We conclude that while a small proportion of subject-target pairs moved from the same location, this is uncorrelated with the built environment and our spatial proximity treatments.

Next, we examine sharing of previous and current residential characteristics between

²²We have very little data on the amount of the offer.

²³There might be unobservables (e.g., friends might try to move to the same apartment building even if they did not share the same previous address), however, this is unlikely to be correlated with the built environment or spatial proximity.

subject-target pairs that could be alternative pathways to helping behavior. In other words, subject-target pairs may share other characteristics that influence helping behavior that are correlated with spatial proximity or the built environment. We find no evidence for this (Table A.3). Neither spatial proximity nor the built environment is correlated with how long the pair has lived in the building, building height of previous residence, the probability of being relocated, the income gap between the subject and target and characteristics of the workplace. The only significant correlation is between the number of apartments per floor in the current residence and sharing the same built environment in the previous residence.

Given the lack of robust correlation between the built environment and spatial proximity of apartments of our subject-target pairs with shared characteristics, we interpret the built environment and spatial proximity as good as random. In our main analysis, we do include a robust set of controls to account for the location of the target’s apartment in the building and the subject’s baseline characteristic. In particular, we control for the characteristics of subjects’ previous residence. Our results are not affected by the inclusion of these variables. In additional checks, we report analysis that includes controls for whether the subject-target pair moved from the same previous address and if they consider each other not strangers. We also present main results with the restricted sample of pairs in predominantly relocated buildings to isolate the effects of the as good as random built environment. All main results hold in these additional analyses.

6 Results

We now turn to the effects of spatial proximity and the built environment on helping behavior in the field experiment and social interactions with neighbors reported in survey responses. Robustness checks, taking into account selection and choice-based sampling, confirm our main results. We present analysis to examine what mediates the spatial proximity and built environment effects on helping a neighbor and explore treatment effects on social network formation.

Our main specification is as follows:

$$Y_{ij} = \beta_0 + \beta_1 E_i + \beta_2 P_i + \beta_3 B_i + \beta_4 X_i + \gamma_j + \epsilon_{ij} \quad (2)$$

where Y_{ij} is the outcome of interest, i.e. helping behavior (number of envelopes returned to target i in community j) or reported social interactions by i in community j with four different neighbors. E_i is a dummy variable for the assigned envelope treatment, P_i is a dummy variable for the assigned proximity treatment, B_i is the built environment in which i lives (number of apartments per floor), X_i is a set of controls for i , including floor of i 's apartment, building height, relocation status, etc., and γ_j are fixed effects for community j . The main outcome variable for each set of analyses is noted in the text and table notes. The type of regression used and full set of controls are listed in the table notes. The permutation and randomization p-values in Table 4 use different samples due to the nature of the tests. These are noted in table notes.

6.1 Treatment effects on helping

Of the 338 envelopes delivered to subjects' mailboxes, and the same number of envelopes delivered directly to the target's mailbox, we have responses on how many envelopes were received from 258 target households.²⁴ The remaining 80 targets were not at home when the survey team tried to reach them. While there is attrition in our sample (24%), we do not find evidence of differential attrition. The probability a target responds to the receipt of envelopes is uncorrelated with treatment assignments or building characteristics (Table A.1). In addition, the survey team made multiple attempts to reach the target. Attrition levels across treatments are similar as the number of visits per household increases (Figure A.5). As noted by Behaghel et al. (2015), unbiased estimates for *always responders* can be obtained if attrition across treatments is similar for some levels of the survey recovery effort.

²⁴A flow diagram of number of envelopes delivered, surveys completed and envelopes reported received for each treatment is in Figure A.4.

We conclude that attrition is not a major source of bias in our study.

Targets reported the number of envelopes they received. Recall that we delivered an envelope directly in the target’s mailbox to account for false negative reports. Thus, the target should report having received one envelope if the subject did not return the misdelivered envelope or two envelopes if the subject returned the envelope. In practice, responses were 0, 1 or 2 envelopes (27%, 35%, 38%).²⁵ We account for false negative reports, i.e. zero envelopes, in the analysis.

[Table 4 about here.]

We first present treatment effects of our experiment on helping behavior. Table 4 reports our main results using OLS regressions and the specification in equation 2 but omitting the built environment, B_i . These estimates are treatment effects of the envelope type and spatial distance between the subject and target on the number of reported envelopes. Both the treatment envelope and the spatial distance were exogenously assigned by us, so the estimates show causal effects of these variables on helping behavior. The omitted envelope treatment is China Post (T1), and the omitted spatial distance is the apartment next door. The table provides permutation and randomization p-values for the spatial distance and envelope treatments following the approach presented in Section 4.²⁶ All estimates of treatment effect significance produce very similar results. We conclude that our results are robust to treatment assignment that avoids interference and not due to small samples.

The results show that signaling urgency by spending extra on courier service delivery (T2) reduces the number of envelopes returned, but this is not statistically significant. Labeling the envelope “urgent” (T3) has no significant effect. Raising the cost to return the envelope by mismatching the address and resident name (T4) significantly reduce the number of envelopes

²⁵Reports of zero envelopes could be due to inattention, throwing the envelope away before opening it and household members other than the survey respondent handling the mail.

²⁶Re-randomization includes both spatial distance and envelope treatments as both need to be included to estimate correctly-adjusted p-values as derived in Section 4.

returned by 0.47. This latter result demonstrates that participants did pay attention to what was on the envelope. Increasing spatial distance between neighbor residences, by being located on different floors, significantly decreases the number of envelopes returned (-0.21). The effect is nearly half the size of mismatched information. These findings show that being the neighbor on a different floor reduces the willingness to help a fellow building resident, even when the cost to help either neighbor is the same.

6.2 Built environment effects

Having established the causal effect of spatial distance between neighbors on helping behavior, we investigate if built environment features affect willingness to help. Two built features that vary across buildings in our setting are building height and apartments per floor. Building height and apartments per floor are highly correlated (0.61). We focus on apartments per floor and control for building height in the regression.²⁷ The top panel of Figure 2 illustrates the effects of spatial distance and apartments per floor on helping behavior using our main specification. Figure 2(a) is based on Probit regressions that restrict estimations to the subsample that reported receiving at least one envelope. This is a sample that excludes false negatives. Figure 2(b) is based on Logit regressions that include all observations and use our design to identify non-response separately from lack of prosocial behavior.²⁸ Each set

²⁷Building height and apartments per floor are not completely interchangeable in the analysis. To see the effects of taller buildings on behavior, we would need to fix the floor on which the subject's apartment is located and vary the building height. We are not powered to do this analysis.

²⁸To account for non-response bias, we estimate a bivariate logit model whereby we assume that non-response probability is independent of treatment assignment. We estimate the model by maximum likelihood under the assumption that the probability of non-response is independent of the probability of receiving the envelope. Estimates are not affected if we allow non-response rates to vary by treatment assignment.

of regressions includes a rich set of controls for subject characteristics and target’s location of residence in the building. We assess whether the estimates are due to endogeneity of the built environment by systematically excluding buildings with a relatively small number of relocated households. The estimates are robust to these restrictions.

[Figure 2 about here]

The top panel of Figure 2 shows that an extra apartment per floor has a comparable effect to living on the same floor. One more apartment per floor reduces the likelihood of helping a neighbor by 18 percentage points. Living in the apartment next door to the neighbor who needs help increases the likelihood of helping by 16 percentage points. These effects maintain if we control for the subject-target pair not being strangers, having moved from the same previous residence and additional demographics.²⁹ The estimated effects increase in size as the proportion of relocated households in the building goes up.

The positive effect of living on the same floor on willingness to help a neighbor could be because spatial proximity makes it more likely neighbors will know one another. We use responses to the household survey on social interactions between neighbors to explore the effect of proximity and built environment. Both the target and subject were asked separately if four different neighbors living next door on the same floor or on floors above or below in the building, including the subject-target pair, were strangers or if they nodded or smiled at the neighbor. The bottom panel of Figure 2 illustrates the effects of proximity and apartments per floor on these reported social interactions, using our main specification. Both figures in the panel are based on Probit regressions. Figure 2(c), shows the effect of the built environment on reporting a neighbor is a stranger, and Figure 2(d), shows the effects on reporting that they nod or smile at the neighbor.

²⁹Results with controls for not being strangers and sharing same previous residence are reported in Columns (4)-(5) and (9)-(10) in Table A.4. Results controlling for age and sex of the subject are reported in Table A.5, with reported coefficients on relocation status. These latter estimates are based on fewer observations than in Table A.4 because of missing data.

The results show that living next door increases positive social interactions. A neighbor living next door is 32 percentage points less likely to be considered a stranger and 27 percentage points more likely to be smiled at. This holds for all households and in the restricted samples of buildings with at least 50% of residents being relocated. The number of apartments per floor increases the likelihood of reporting a neighbor is a stranger and decreases the likelihood of smiling at the neighbor, however, these are not statistically significant across all specifications.

6.3 Robustness

Before moving to mediation analysis, we present robustness tests of the effect of the built environment on helping behavior. In particular, we investigate if the results are sensitive to selection and survivor bias.

In an ideal experiment, participants would have been randomly relocated to, and made to stay in, different built structures. Section 5.2 presents evidence consistent with the assignment to the built environment to be as good as random, but people do move in and out of buildings. The built environment might be correlated with the preferences of those who choose to move in and those who choose to move out after relocation. The first behavior leads to selection bias and the second to survivor bias. Since we cannot assume that the preferences of those who chose to move in are identical to those who left, a concern is that our results are not robust.

We present a bounds approach and a control function approach to test robustness and show our main results hold under both approaches (see Appendix E for description). Using a bounds approach, we confirm that our estimates in the subsamples with mostly relocated participants are robust to selection/survivor bias (see Table A.7). One more apartment per floor reduces the probability of returning the envelope by 12 to 19 percentage points. The same results hold using a control function approach. One more apartment per floor reduces the likelihood of returning the envelope by 14 to 26 percentage points, depending on the

specification (Table A.9 and Figure A.6).

6.4 Mediation

Having established robustness of our main results, we explore whether the effects of proximity and the built environment on helping behavior are mediated by social interactions between neighbors.

To do so, we construct an indicator for subject-target familiarity. The indicator is equal to one if neither the subject nor the target declare the other as a stranger. The variable is recorded as missing if we do not have either person’s response. This variable is likely to misclassify friendships between subjects and targets. It is well-known that misclassification in covariates leads to biased estimators (e.g., Hausman, 2001). While we could bound the magnitude of the marginal effect of friendship on the probability of returning an envelope using Bollinger (1996)’s approach, the bounds are too wide to be informative.³⁰ Also, we do not have feasible instruments as in Mahajan (2006).

Misclassification likely biases the effect of friendship on helping downwards. To assess the potential magnitude of the problem, we exogenously impose a coefficient value on friendship and evaluate how the other coefficients vary. The idea is based on the reduced-form approach suggested by Chernozhukov and Hansen (2008). It relies on the assumption that the effect of being next door on helping is completely mediated by the fact that next-door neighbors are more likely to be familiar with one another due to spatial proximity. If this is true, we would expect that under the true coefficient on being familiar, not the biased ones due to misclassification, the effect of next-door neighbor would disappear. If familiarity does not mediate the effect of being a next-door neighbor, we will not find a parameter for which the effect of proximity vanishes. This procedure will overestimate the effect of friendship if there are omitted mediators (see Heckman et al., 2013). While this might be the case, we note

³⁰Using a linear regression for the specification of column (3) in Table A.4, we find the marginal effect of friendship can be between 0.04 and 27.04.

that familiarity and spatial proximity are highly correlated. The probability of reporting a neighbor is not a stranger increases when the target and subject are next-door neighbors (0.61 v. 0.90, p-value < 0.001).

[Figure 3 about here]

The top panel in Figure 3 shows the effect of setting the marginal effect of familiarity at different values. The regressions use the specification of Column (6) in Table A.4. The x-axis is the fixed marginal effect of familiarity, and the y-axis is the marginal effect of the reported variable at that fixed value of x. Figure 3(a) shows that the effect of being a next-door neighbor decreases as the coefficient on familiarity increases. If familiarity is the only mediator, the marginal effect of familiarity has to be larger than 0.40 to completely mediate the effect of being a next-door neighbor on helping behavior. Importantly, Figure 3(b) shows that the effect of apartments per floor is not mediated by familiarity at all. We conclude that while familiarity is a potential important mediating factor for the effect of spatial proximity on helping a neighbor, the built environment directly impacts behavior.

6.5 Social network structure and the build environment

We examine whether the built environment might have also affected social network features. Our experiment was not explicitly designed to test this, so this analysis is exploratory.

To maximize contrast, we concentrate on buildings with four or fewer apartments per floor and buildings with more than four apartments per floor.³¹ We treat apartments as nodes and define a directed edge, or a link between two neighbors, if a survey respondent does not declare a neighbor a stranger. We look at two characteristics of networks: the distribution of edges per node and triadic transitivity between apartments. We exploit that our household survey asked network questions for a random set of four neighbors. Thus,

³¹Three-quarters of our survey sample live in a building with four or fewer apartments per floor. Sixty-two percent live in a building with three or fewer apartments per building. The results presented are qualitatively similar using either cut of the data.

we observed social interactions with neighbors on a random subsample of their complete networks.

The bottom panel in Figure 3 illustrates these two aspects of social networks. Figure 3(c) shows the distribution of directed edges between a respondent and four randomly chosen neighbors. Respondents in buildings with fewer apartments per floor are 67% more likely to report links with all four neighbors than those in buildings with more apartments per floor (p-value = 0.1171). However, the overall degree distribution over those four neighbors is not statistically different between buildings with four or fewer apartments per floor and buildings with more than four apartments per floor (Fischer exact p-value=0.352).

Figure 3(d) shows the triadic closure of networks by apartments per floor. We follow Jackson (2008) (Section 2.2.3) to examine the transitivity of triples, i.e., the likelihood that three neighbors share connections.³² In particular, we look at the number of directed edges between node i and j as a function of the number of directed edges from i and j to a third node k . We can construct 107 such triples in our dataset.³³ Evidence of transitivity is significant in buildings with four or fewer apartments per floor (Fischer exact p-value = 0.001) but not in buildings with more than four apartments per floor (Fischer exact p-value = 0.497), albeit we have relatively few triples in this condition (28).

While a common reason for clustering is homophily, we find little evidence that neighbors are more similar in buildings with more apartments per floor (see Section 5.2). We, of course, cannot discard the existence of unobserved similarity in some building structures. Another explanation for the observed clustering is spatial proximity. Some evidence consistent with this is the larger impact of living next door on social interactions compared to living one floor above (Figure 2, bottom panel).

³²Because we do not have information on complete networks, we cannot distinguish between clustering and “support” in our data (Jackson et al., 2012).

³³Triples are constructed by asking both i and j in the household survey about social interactions with each other and a common neighbor, k .

In sum, the built environment affects helping behavior, social interactions and the structure of social networks. This latter effect is more evident on the intensive margin, rather than the extensive margin. On the extensive margin, buildings with fewer apartments per floor increase bilateral familiarity and consequently the number of connections a person has. On the intensive margin, buildings with fewer apartments per floor increase triadic closures. In other words, buildings with fewer apartments per floor foster more familiarity between two neighbors and more triadic clustering. Consequently, social networks are more dense with familiar links, thus increasing the willingness to help a neighbor. This is consistent with evidence that network structures that provide social pressure can support informal cooperation ([Jackson et al., 2012](#)).

7 Conclusions

Social connections, trust and interactions with others have long been argued as essential for a well functioning society. Urban design and built space affect how we move around, who we encounter and ultimately the where, how and type of human interactions we have. Associations have been found between the built environment and health, travel and social connections, but these are in environments that people have chosen to be in. So, it is difficult to disentangle whether effects are due to the environment itself or the type of people who chose that environment. Our study establishes the causal effect of residential built environment characteristics, such as apartments per floor and building height, as well as spatial proximity between neighbors, on low-cost helping behavior.

To do this, we conducted a field experiment in residential communities in Shanghai, China where many residents were involuntarily relocated due to rapid redevelopment in the city. Residents ended up in buildings with a range of apartments per floor and short and tall buildings. We confirm an as good as random assignment to the built environment and thus interpret our findings as causal. Our behavioral measure of helping a neighbor is the return of a misdelivered piece of mail. Mailboxes are centrally located, so the cost of putting the

envelope in the correct slot is very low and similar for all neighbors. One of our treatment arms varies the spatial distance between the person who should have gotten the mail and the subject who received the misdelivered mail to see how spatial proximity within a building affects willingness to help. When two residents are located one floor apart, rather than living on the same floor, the likelihood of returning the letter decreases by 16 percentage points. The built environment also has a sizable effect. Adding one more apartment per floor to the building also decreases the likelihood of helping by 18 percentage points.

Our findings largely confirm the impact of spatial proximity of neighbors on social interactions found in correlational studies and further show the independent causal impact of the built environment on fostering the willingness to help a neighbor. An additional insight we add to this literature is that narrow buildings create more dense social networks which then can support informal cooperation (as discussed in [Jackson et al., 2012](#)). We conclude that small spatial barriers in the built environment, such as residing on different floors of a building and more residents per floor, can have a profound effect on social interactions and the willingness to engage in small, low-cost acts of kindness that help a neighbor.

More broadly, if buildings can be thought of as mini societies, architectural and urban design that enhance social interactions and contain the number of immediate neighbors will produce communities within buildings that are more civic minded. Our study did not examine helping behavior across buildings, but within the same housing community, to see how this might extend to larger urban design. This is an interesting avenue to explore, especially given the current movement in urban design to create large car-free spaces in cities in which people can safely move around and interact with one another.

References

Susan Athey, Dean Eckles, and Guido W. Imbens. Exact p-values for network interference.

Journal of the American Statistical Association, 113(521):230–240, 2018.

AUE. Atlas of urban expansion – shanghai, 2023.

- Daniel C. Batson and Adam A. Powell. Altruism and prosocial behavior. In Theodore Millon and Melvin J. Lerner, editors, *Handbook of Psychology*, volume 5 of *Personality and Social Psychology*, pages 463–484. John Wiley and Sons, 2003.
- Luc Behaghel, Bruno Crepon, Marc Gurgand, and Thomas Le Barbanchon. Please call again: Correcting nonresponse bias in treatment effect models. *Review of Economics and Statistics*, 97(5):1070–1080, DEC 2015.
- Stephen Bochner, Robert Duncan, Elizabeth Kennedy, and Fred Orr. Acquaintance links between residents of a high rise building: An application of the “small world” method. *The Journal of Social Psychology*, 100(2):277–284, 1976.
- Page Bollen. Familiar strangers: Repeated casual contact and intergroup relations in urban south africa. Technical report, Working Paper, 2023.
- Christopher Bollinger. Bounding mean regressions when a binary regressor is mismeasured. *Journal of Econometrics*, 73(2):387–399, AUG 1996.
- Kevin J Boudreau, Tom Brady, Ina Ganguli, Patrick Gaule, Eva Guinan, Anthony Hollenberg, and Karim R Lakhani. A field experiment on search costs and the formation of scientific collaborations. *Review of Economics and Statistics*, 99(4):565–576, 2017.
- Victor Chernozhukov and Christian Hansen. The reduced form: A simple approach to inference with weak instruments. *Economic Letters*, 100(1):68–71, Jul 2008.
- Arza Churchman. Disentangling the concept of density. *Journal of planning literature*, 13(4):389–411, 1999.
- Deborah A Cohen, Sanae Inagami, and Brian Finch. The built environment and collective efficacy. *Health & place*, 14(2):198–208, 2008.
- Christine M Coussens, John Porretto, and Lovell Jones. *Rebuilding the Unity of Health and*

- the Environment: The Greater Houston Metropolitan Area: Workshop Summary*. National Academies Press, 2005.
- Zoe Cullen and Ricardo Perez-Truglia. The old boys' club: Schmoozing and the gender gap. *American Economic Review*, 113(7):1703–1740, 2023.
- Jennifer Day and Robert Cervero. Effects of residential relocation on household and commuting expenditures in shanghai, china. *International journal of urban and regional research*, 34(4):762–788, 2010.
- Ebbe B Ebbesen, Glenn L Kjos, and Vladimir J Konečni. Spatial ecology: Its effects on the choice of friends and enemies. *Journal of Experimental Social Psychology*, 12(6):505–518, 1976.
- Nancy Eisenberg, Tracy L Spinrad, and Amanda S Morris. Prosocial development. In M. E. Lamb and R. M. Lerner, editors, *Handbook of Child Psychology and Developmental Science, Socioemotional Processes*, volume 3, pages 610–656. John Wiley and Sons, 5th edition, 2013.
- Benjamin Enke, Ricardo Rodriguez-Padilla, and Florian Zimmermann. Moral universalism: Measurement and economic relevance. *Management Science*, 68(5):3590–3603, 2022.
- Gary W Evans. The built environment and mental health. *Journal of urban health*, 80: 536–555, 2003.
- Reid Ewing and Robert Cervero. Travel and the built environment: A meta-analysis. *Journal of the American planning association*, 76(3):265–294, 2010.
- Maxime Felder, Guillaume Favre, Marina Tulin, and Petros Koutsolampros. Acquaintances or familiar strangers? how similarity and spatial proximity shape neighbour relations within residential buildings. *Housing, Theory and Society*, 40(5):642–659, 2023.

- Leon Festinger, Stanley Schachter, and Kurt Back. Social pressures in informal groups; a study of human factors in housing. 1950.
- Jan Gehl. *Cities for people*. Island press, 2013.
- Robert Gifford. The consequences of living in high-rise buildings. *Architectural science review*, 50(1):2–17, 2007.
- Edward L Glaeser and Bruce Sacerdote. The social consequences of housing. *Journal of housing economics*, 9(1-2):1–23, 2000.
- Edward L Glaeser, David I Laibson, Jose A Scheinkman, and Christine L Soutter. Measuring trust. *The quarterly journal of economics*, 115(3):811–846, 2000.
- Peter Hall. On representatives of subsets. *Journal of the London Mathematical Society*, 10: 26–30, 1935.
- William D Hamilton. The genetical evolution of social behaviour. ii. *Journal of theoretical biology*, 7(1):17–52, 1964.
- Glenn W Harrison and John A List. Field experiments. *Journal of Economic Literature*, 42(4):1009–1055, 2004.
- Jerry Hausman. Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *Journal of Economic Perspectives*, 15(4):57–67, FALL 2001.
- Shenjing He. New-build gentrification in central shanghai: demographic changes and socio-economic implications. *Population, Space and Place*, 16(5):345–361, 2010.
- James Heckman, Rodrigo Pinto, and Peter Savelyev. Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, 103(6):2052–2086, Oct 2013.

- Bill Hillier. *Space is the machine: a configurational theory of architecture*. Space Syntax, 2007.
- Guido W Imbens and Donald B Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.
- Matthew O Jackson. *Social and Economic Networks*. Princeton University Press, 2008.
- Matthew O Jackson, Tomas Rodriguez-Barraquer, and Xu Tan. Social capital and social quilts: Network patterns of favor exchange. *American Economic Review*, 102(5):1857–1897, 2012.
- Jane Jacobs. *The Death and Life of Great American Cities*. Random House Press, 1961.
- Ade Kearns, Elise Whitley, Phil Mason, and Lyndal Bond. ‘living the high life’? residential, social and psychosocial outcomes for high-rise occupants in a deprived context. *Housing Studies*, 27(1):97–126, 2012.
- Dennis Krebs. Empathy and altruism. *Journal of Personality and Social psychology*, 32(6):1134, 1975.
- Kevin J Krizek. The complex role of urban design and theoretical models of physical activity. *Progressive Planning*, 157:28–29, 2003.
- Stephen Leider, Markus M Möbius, Tanya Rosenblat, and Quoc-Anh Do. Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics*, 124(4):1815–1851, 2009.
- Si-ming Li and Yu-ling Song. Redevelopment, displacement, housing conditions, and residential satisfaction: a study of shanghai. *Environnement and Planning A*, 41:1090–1108, May 2009.
- Aprajit Mahajan. Identification and estimation of regression models with misclassification. *Econometrica*, 74(3):631–665, MAY 2006.

- David Marmaros and Bruce Sacerdote. How do friendships form? *The Quarterly Journal of Economics*, 121(1):79–119, 2006.
- Soumya Mazumdar, Vincent Learnihan, Thomas Cochrane, and Rachel Davey. The built environment and social capital: A systematic review. *Environment and Behavior*, 50(2): 119–158, 2018.
- Stanley Milgram. The familiar stranger: An aspect of urban anonymity, 1977.
- NBSC. China statistical yearbook 2022, 2022.
- Theodore M. Newcomb. Varieties of interpersonal attraction. In A Cartwright, D & Zander, editor, *Group dynamics: Research and theory*, pages 104–119. 2nd edition, 1960.
- Oscar Newman. *Defensible Space: Crime Prevention through Urban Design*. Macmillan Press, 1972.
- Oscar Newman. *Defensible space: People and design in the violent city*. Architectural Press, 1973.
- Martin A Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563, 2006.
- Mia A Papas, Anthony J Alberg, Reid Ewing, Kathy J Helzlsouer, Tiffany L Gary, and Ann C Klassen. The built environment and obesity. *Epidemiologic reviews*, 29(1):129–143, 2007.
- Robert D Putnam. *Bowling alone: The collapse and revival of American community*. Simon and schuster, 2000.
- Donald B Rubin. Bayesian inference for causal effects: The role of randomization. *The Annals of statistics*, pages 34–58, 1978.
- Arianna Salazar Miranda and Matthew Claudel. Spatial proximity matters: A study on collaboration. *Plos one*, 16(12):e0259965, 2021.

- SCG. Construction in shanghai 1991-1995. Scientific Popularization Press, Shanghai, 1996.
- Wenbiao Sha and Xianqiang Zou. The political economy of eminent domain: The economic and political effects of housing demolition in china. Technical report, SSRN 4237524, 2023.
- Mi Shih. The evolving law of disputed relocation: Constructing inner-city renewal practices in shanghai, 1990–2005. *Internation Journal of Urban and Regional Research*, 34:350–364, June 2010.
- Oddvar Skjaeveland and Tommy Garling. Effects of interactional space on neighbouring. *Journal of environmental psychology*, 17(3):181–198, 1997.
- SSB. Shanghai statistical yearbook 2007, 2007.
- SSB. Shanghai statistical yearbook 2010. <https://tjj.sh.gov.cn/tjnj/tjnj2010e.htm>, 2010.
URL <https://tjj.sh.gov.cn/tjnj/tjnj2010e.htm>.
- Shiwei Sun and Yuchuan Deng. A study of the role of urban planning in shanghai. *Urban Planning Forum*, 2:31–39, 1997.
- Maferima Touré-Tillery and Ayelet Fishbach. Too far to help: The effect of perceived distance on the expected impact and likelihood of charitable action. *Journal of personality and social psychology*, 112(6):860, 2017.
- Robert L Trivers. The evolution of reciprocal altruism. *The Quarterly review of biology*, 46(1):35–57, 1971.
- Barry Wellman. The community question: The intimate networks of east yorkers. *American journal of Sociology*, 84(5):1201–1231, 1979.
- Jean D Wineman, Felichism W Kabo, and Gerald F Davis. Spatial and social networks in organizational innovation. *Environment and behavior*, 41(3):427–442, 2009.

- Furlong Wu. Residential relocation under market-oriented redevelopment: the process and outcomes in urban china. *Geoforum*, 35((4)):453–470, July 2004a.
- Furlong Wu. Intraurban residential relocation in shanghai: modes and stratification. *Environment and Planning A: Economy and Space*, 36:7–25, January 2004b.
- Zuhai Yao and Qiang Jiang. Housing and civil infrastructure construction. *Social Development Bluebook of Shanghai 2002*, pages 94–107, 2002.
- Renee Zahnow, Min Zhang, and Jonathan Corcoran. The girl on the bus: Familiar faces in daily travel and their implications for crime protection. *Annals of the American Association of Geographers*, 111(5):1367–1384, 2021.
- Jingjing Zhong, Wenting Liu, Buqing Niu, Xiongbin Lin, and Yanhua Deng. Role of built environments on physical activity and health promotion: A review and policy insights. *Frontiers in Public Health*, 10:950348, 2022.
- Jieming Zhu. From land use right to land development right: Institutional change in china’s urban development. *Urban Studies*, 41(7):1249–1267, June 2004.

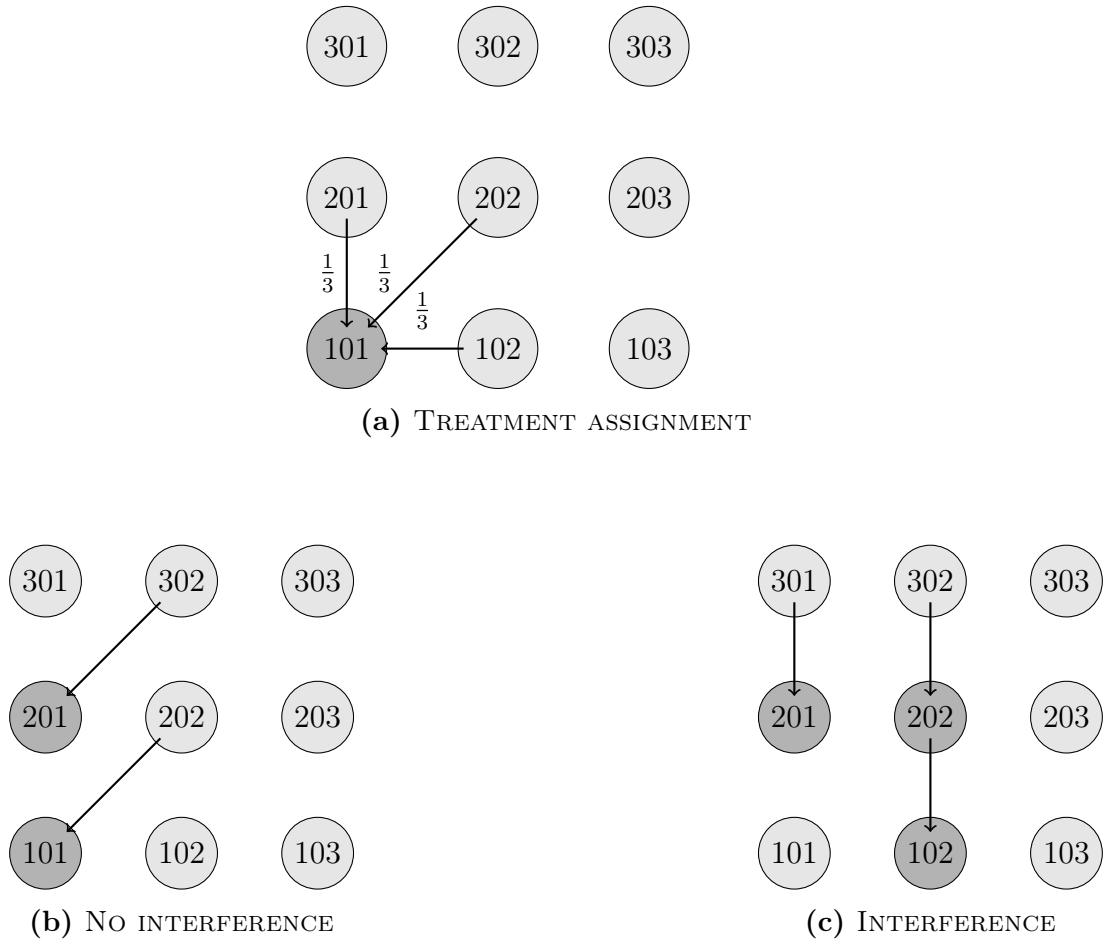
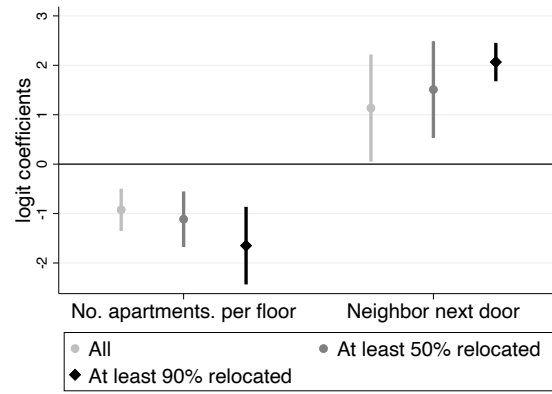
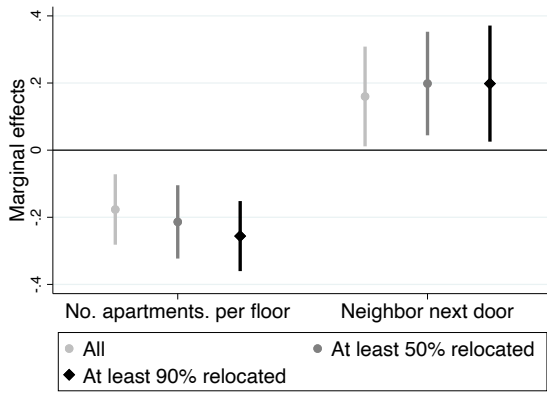


Figure 1: POSSIBLE ASSIGNMENTS

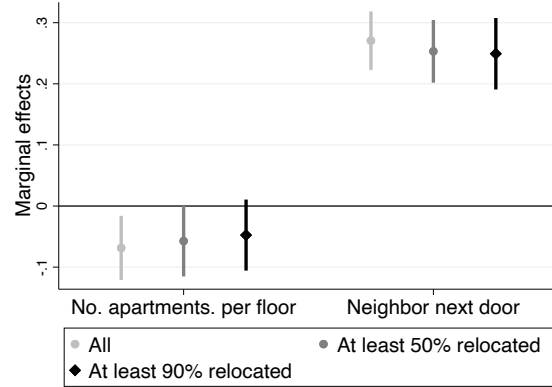
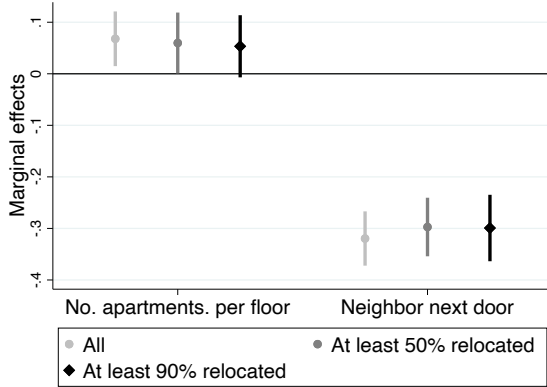
Note: Target apartment is in dark grey. Potential subject apartments are in light grey. The probability of treatment assignment is $\frac{1}{3}$.



(a) Probit (excluding false negatives)

(b) Logit (accounting for false negatives)

Top panel: Behavioral measure of helping behavior



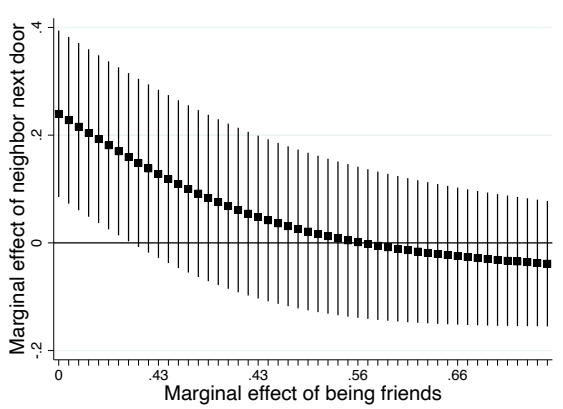
(c) Neighbor is a stranger

(d) Nods and smiles at neighbor

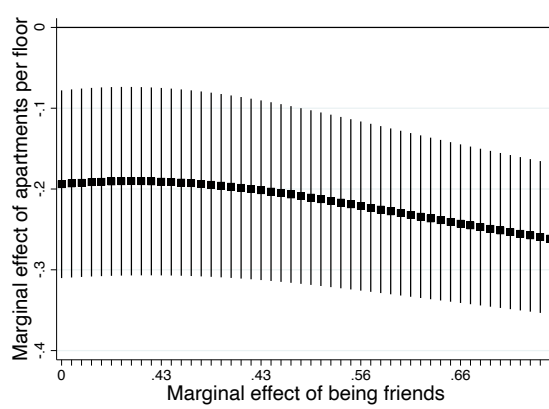
Bottom panel: Survey measures of quality of social interactions

Figure 2: Helping behavior and social interactions

Notes: Top panel dependent variable is whether the envelope was returned to the target. Figure (a) reports coefficients from Probit regressions ($N \leq 189$) in Columns (1)-(3) and Figure (b) from Logit regressions ($N \leq 258$) in Columns (6)-(8) in Table A.4. ‘All’ is all households, and ‘At least 50% [90%]’ relocated includes buildings with at least 50% [90%] relocated households. Error bars are 90-percent confidence intervals. Bottom panel dependent variable for Figure (c) is 1 if neighbor is ‘non-friendly’ or ‘a stranger’ and for Figure (d) is 1 if respondent says that he/she nods and smiles at a neighbor. Regression coefficients from Probit regressions in Table A.6.

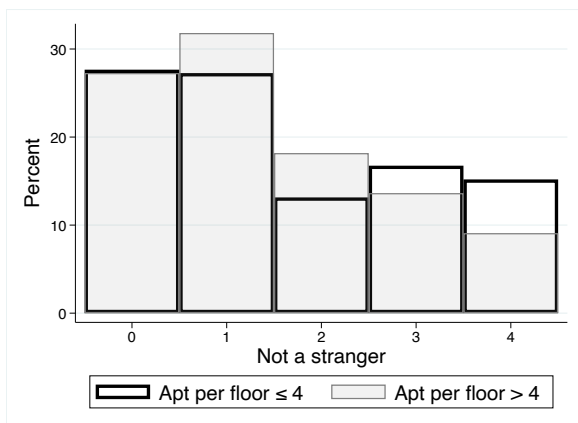


(a) Neighbor next door

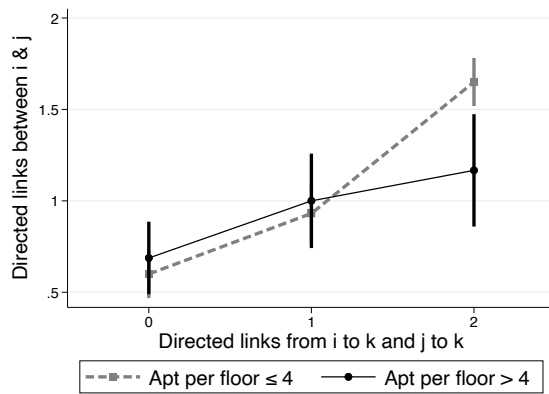


(b) Apartments per floor

Top panel: Mediation



(c) Degree distribution



(d) Triadic closure

Bottom panel: Network effects by spatial distance and apartments per floor

Notes: Figure (c) shows the distribution of connections across the four neighbors asked about in the household survey. Figure (d) shows the average number of links between i and j when the links between i and k and j and k equal 0, 1 or 2.

Figure 3: Mediation and Network effects

Table 1: Envelope treatments

Treatment	Description	Households
T1	Regular China Post	92
T2	Courier service	88
T3	Courier service + urgent stamp	85
T4	Target's name, subject's address	73
Placebo	No envelope delivered	107
Total		445

Notes: The table reports the number of target households assigned to each treatment. Each household in T1-T4 was treated with two envelopes, one deposited in the target mailbox to test for false negatives and one deposited in the subject mailbox to test for helping behavior. Households in the Placebo treatment did not receive an envelope. Placebo households serve as a test for false positives.

Table 2: Building characteristics of sample communities

Variable	Observations	Mean	S.D.	Min	Max
Number of floors	429	7.74	3.94	5	24
Apartments per floor	429	2.39	1.10	2	8
Number of entrances	429	1.12	0.33	1	2
Number of elevators	429	0.33	0.67	0	2
% with elevator(s)	429	20.98	40.76	0	100
Number of stairwells	429	1.01	0.11	1	2
With shop on first floor	429	6.06	23.89	0	100
Stairwell has window	429	68.76	46.40	0	100
Year built	429	2000.38	5.14	1988	2009
% relocated households	429	76.23	29.88	0	100

Table 3: Household characteristics of sample

Variable	Relocated	Not relocated	Total
Percent of sample	59%	41%	100%
Age	52.0	42.6	48.1
Male	49.0%	51.7%	50.1%
Year moved	2003.2	2006.2	2004.4
Previous residence height	2.4	5.3	3.6
New residence height	12.3	11.4	11.9
Previous residence apts/floor	1.5	3.1	2.2
New residence apts/floor	4.1	3.4	3.8
# housing options			
One	45%		
Two	38%		
Three	11%		
Four or more	6%		
Offered housing	97%		
Offered cash	57%		
Owner	78%		
Observations	249	174	423

Notes: Ownership and housing options questions were only asked of relocated households.

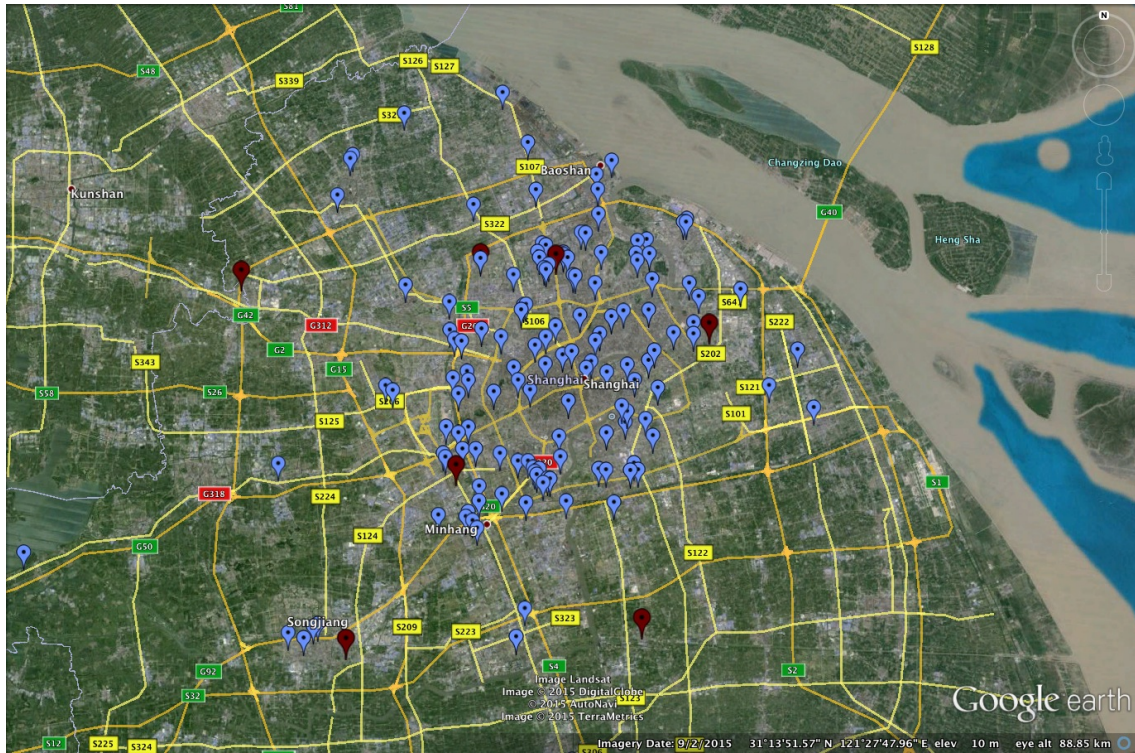
Table 4: Treatment effects of proximity and envelopes on number of envelopes returned

	Model 1	(1)	(2)	(3)	(4)	Model 2	(1)	(2)	(3)	(4)
Express (T2)	-0.212	0.130	0.160	0.242	0.137	-0.212	0.130	0.160	0.242	0.137
Urgent Express (T3)	-0.012	0.930	0.982	0.854	0.929	-0.012	0.930	0.982	0.854	0.929
Urgent Express - Mismatch (T4)	-0.474	0.001	0.002	0.002	0.001	-0.474	0.001	0.002	0.002	0.001
Apartment right above	-0.249	0.037	0.046	0.034	0.042					
Apartment above right	-0.153	0.236	0.242	0.192	0.323					
Apartment above						-0.208	0.048	0.046	0.044	0.042

Notes: OLS regressions. The dependent variable is the number of envelopes, out of 2, returned (avg 1.11). Fixed effects for community, apartment’s floor, and building height are included. The sample includes 258 target apartments and excludes results from the placebo treatment. Model 1 and Model 2 columns report coefficients. (1) is the standard p-value. (2) is the Naive permutation p-value obtained by ignoring interference (n=258). (3) is the Greedy permutation p-value obtained by excluding targets that cannot be assigned to three subjects due to interference (n=241). (4) is the Re-randomization p-value obtained by creating artificial experiments that do not violate Remark 1 (n=258) (see Section 4 and Appendix D).

Online Appendices for: Castillo, Marco, Ragan Petrie and Rong Rong, 2025, “My neighbor next floor: the built environment and social preferences,” *Review of Economics and Statistics*

A Additional figures and tables



Notes: Sampled communities in maroon, randomly sampled communities that fit sampling criteria in blue

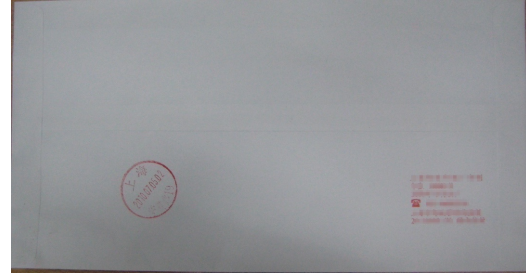
Figure A.1: Map of seven sample communities



Figure A.2: Mailboxes



Front



Back

Panel A: China Post (T1 & T4)



Front

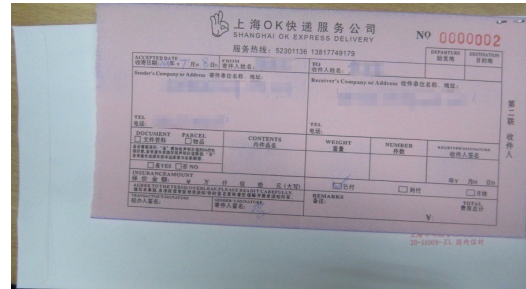


Back

Panel B: Express courier (T2)



Front



Back

Panel C: Express courier urgent (T3)

Figure A.3: Treatment envelopes

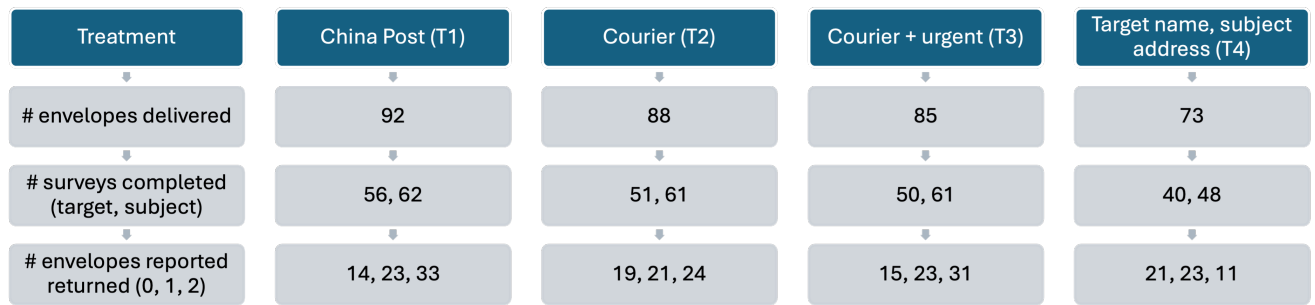


Figure A.4: Flow diagram of number of envelopes delivered, surveys completed and envelopes reported received by treatment

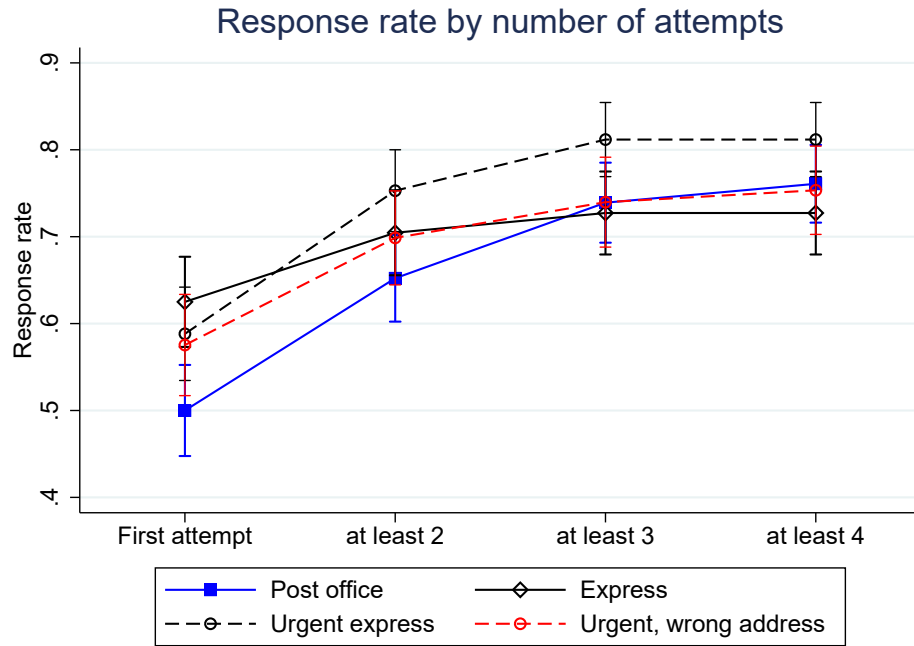


Figure A.5: Attrition

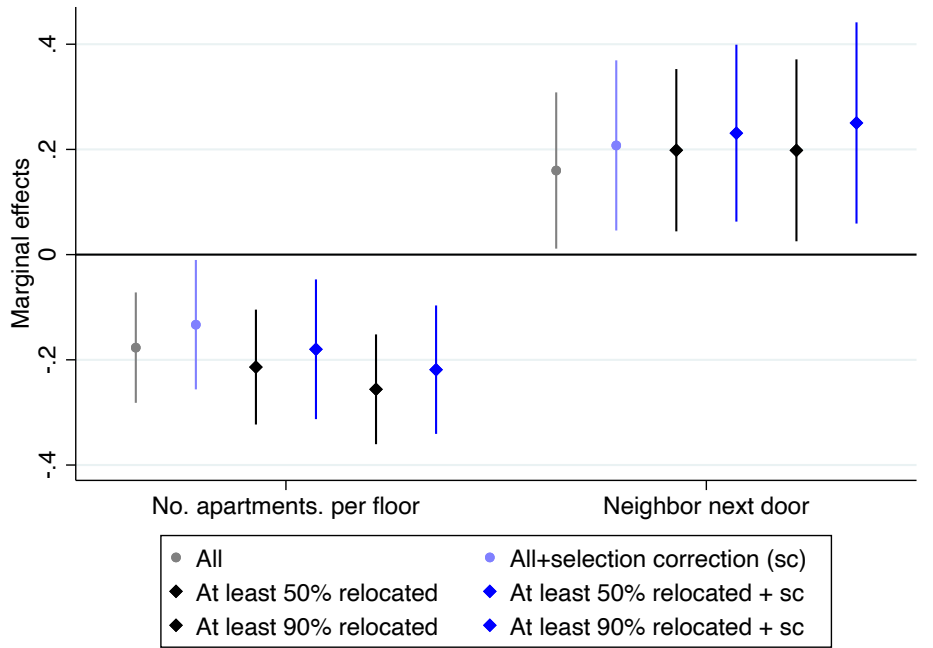


Figure A.6: Behavioral measure of helping behavior and selection bias

Notes: Figure shows regression coefficients from Probit regressions (excluding false negatives) from Columns (1)-(6) in Table A.9. Regressions control for envelope treatment assignment, community, floor of target’s apartment \times building height, subject’s relocation status, length of residence in the building, monthly expenditures, receiving only one offer for relocation, and height and apartments per floor of previous dwelling. All includes all households. At least 50% [90%] relocated only includes buildings with at least 50% [90%] relocated households. Estimates exclude cases reporting having received no envelopes ($N \leq 189$).

Table A.1: Response to surveys

	(1)	(2)
Express (T2)	-0.166 (0.348)	-0.147 (0.350)
Urgent Express (T3)	0.300 (0.375)	0.295 (0.376)
Urgent Express (T4)	-0.120 (0.372)	-0.128 (0.375)
Building has lift		-0.348 (0.412)
Apartments per floor		0.090 (0.132)
Relocated household		-0.148 (0.338)
Observations	338	338
R2	0.026	0.028

Notes: Standard errors in parentheses. Community level dummies included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table A.2: Balancedness across target-subject treatment assignments

Variable	(1) Next door	(2) Next floor up	(3) Next floor right	(4) 0 v 1	(5) 0 v 2	(6) 1 v 2
Age (target)	44.649 (16.659)	44.077 (13.914)	42.771 (16.534)	-0.573 (2.246)	-1.878 (2.604)	1.305 (2.402)
Male (target)	0.392 (1.085)	0.407 (0.494)	0.386 (1.243)	0.015 (0.124)	-0.006 (0.181)	0.021 (0.143)
Building height	12.062 (6.500)	12.637 (7.204)	11.271 (6.138)	0.576 (1.000)	-0.790 (0.996)	1.366 (1.075)
Apartments per floor	3.742 (2.446)	4.033 (2.622)	3.471 (2.295)	0.291 (0.370)	-0.271 (0.374)	0.562 (0.395)
Male (subject)	0.508 (0.504)	0.606 (0.492)	0.593 (0.496)	0.098 (0.087)	0.085 (0.092)	0.013 (0.091)
Age (subject)	48.969 (16.254)	49.017 (15.315)	46.427 (11.716)	0.048 (2.840)	-2.541 (2.636)	2.589 (2.560)
HH size (subject)	3.083 (1.381)	3.383 (1.091)	3.538 (1.056)	0.300 (0.227)	0.455* (0.235)	-0.155 (0.204)
Year move (subject)	2,004.621 (4.348)	2,003.404 (4.648)	2,004.941 (4.333)	-1.217 (0.839)	0.320 (0.833)	-1.538* (0.868)
Monthly expenses (subject)	2,596.370 (1,597.584)	3,025.040 (1,689.977)	2,960.935 (1,782.082)	428.670 (322.382)	364.564 (338.046)	64.105 (354.398)
People known (subject)	2.053 (3.956)	3.412 (5.628)	2.133 (4.294)	1.360 (0.911)	0.080 (0.802)	1.280 (0.985)
Relocated (subject)	0.581 (0.497)	0.554 (0.502)	0.540 (0.503)	-0.027 (0.092)	-0.041 (0.095)	0.014 (0.098)
Observations	97	91	70	188	167	161

Notes: The three possible target-subject treatment assignments are Next door, Next floor up and Next floor right. Columns (1)-(3) list averages across each covariate. Columns (4)-(6) report differences in means between Next door and Next floor up, Next door and Next floor right and Next floor up and Next floor right, respectively. Standard errors are in parentheses.

Table A.3: Characteristics shared by subject-target pairs

Panel A: By treatment

	Year moved	Origin	B. height	No. apts	Displaced	Inc. gap	Job descrip.	Employer	Workplace
Next floor	0.056 (0.089)	0.048 (0.063)	0.117 (0.109)	-0.017 (0.114)	-0.008 (0.104)	-0.046 (0.179)	-0.010 (0.071)	0.011 (0.058)	0.046 (0.071)
Next floor right	-0.147* (0.088)	-0.024 (0.064)	0.168 (0.107)	-0.067 (0.112)	-0.108 (0.104)	-0.009 (0.183)	-0.079 (0.073)	-0.060 (0.058)	-0.021 (0.070)
Constant	0.261*** (0.061)	0.077* (0.045)	0.216*** (0.076)	0.267*** (0.079)	0.408*** (0.070)	-0.147 (0.123)	0.170*** (0.050)	0.135*** (0.040)	0.141*** (0.049)
R2	0.041	0.012	0.023	0.005	0.010	0.001	0.015	0.015	0.008
F-test	2.742*	0.671	1.302	0.192	0.639	0.036	0.697	0.850	0.454

Panel B: By apartments per floor

	Year moved	Origin	B. height	No. apts	Displaced	Inc. gap	Job descrip.	Employer	Workplace
Apts. per floor	0.088 (0.086)	-0.001 (0.062)	-0.001 (0.139)	0.299** (0.139)	0.138 (0.087)	0.243 (0.157)	0.060 (0.066)	0.023 (0.052)	0.104 (0.066)
Constant	-0.064 (0.299)	0.249 (0.198)	0.666* (0.361)	0.509 (0.354)	0.434 (0.286)	-0.915 (0.603)	0.063 (0.212)	0.137 (0.182)	-0.106 (0.232)
R2	0.367	0.430	0.461	0.658	0.516	0.568	0.476	0.420	0.359
Pairs	131	117	112	88	129	106	93	118	118
Mean	0.229	0.085	0.312	0.239	0.372	-0.165	0.142	0.119	0.149

Notes: Dependent variable equals one if both subject and target share the characteristic listed in each column. Year moved is year moved into current residence, Origin is previous address, B. height is previous residence building height, No. apts is previous residence number of apartments per floor, Displaced is having been relocated to current residence, Job descrip. is description of current job, Employer is current employer and Workplace is address of current workplace. Inc. gap is the difference in (log) monthly expenditure between target and subject. Number of pairs and mean of dependent variable listed in bottom rows of Panel B are the same for Panel A.

Table A.4: Envelope returned

	Probit regressions (marginals) (excluding false negatives)					Logit regressions (accounting for false negatives)				
	(1) All	(2) 50% rell.	(3) 90% rell.	(4) w/nstrang	(5) w/s. orig.	(6) All	(7) 50% rell.	(8) 90% rell.	(9) w/nstrang	(10) w/s. orig.
Apt. per floor	-0.177*** (0.064)	-0.214*** (0.066)	-0.256*** (0.063)	-0.194*** (0.071)	-0.171*** (0.064)	-0.924*** (0.260)	-1.114*** (0.342)	-1.649*** (0.477)	-0.968** (0.385)	-0.917*** (0.237)
Same floor	0.160* (0.090)	0.198** (0.094)	0.198* (0.105)	0.241** (0.108)	0.161* (0.089)	1.136* (0.659)	1.509** (0.595)	2.066*** (0.236)	1.607 (1.206)	1.138* (0.663)
Relocated (subject)	0.303** (0.143)	0.337** (0.158)	0.412** (0.193)	0.421** (0.170)	0.297** (0.142)	1.784** (0.801)	1.712* (0.976)	2.080 (1.574)	2.209 (1.592)	1.768** (0.764)
Year of move (subject)	-0.008 (0.015)	-0.014 (0.015)	-0.004 (0.016)	-0.003 (0.015)	-0.007 (0.014)	-0.001 (0.072)	-0.052 (0.073)	-0.030 (0.068)	0.002 (0.075)	-0.001 (0.072)
Observations	116	112	83	103	116	245	218	161	231	245

Notes: Dependent variable is whether envelope was returned. Regressions control for envelope treatment assignment, community, floor of target's apartment \times building height, subject's relocation status, length of residence in the building, monthly expenditures, receiving only one offer for relocation, and height and apartments per floor of previous dwelling. Controlling for not stranger ("w/nstrang") includes all households and controls for whether the target or subject say they are not strangers. Controlling for same origin ("w/s. orig.") includes all households and controls for whether the target and subject moved from the same previous residence. Any missing data is completed with average at the building level. All includes all households. At least 50% [90%] relocated only includes buildings with at least 50% [90%] relocated households. The left-hand side panel estimates exclude cases reporting receiving no envelopes ($N \leq 189$). The right-hand side panel accounts for misclassification using the rate of false negatives estimated from envelopes delivered by the authors ($N \leq 258$). The model excludes the possibility of false positives since the placebo treatment refutes that case. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table A.5: Envelope returned

	(1)	(2)	(3)	(4)
	Probit regressions (marginals)		Logit regressions	
	(excluding false negatives)		(accounting for false negatives)	
	All	50%	90%	w/friends
	b/se	b/se	b/se	b/se
Subject relocated	0.628*** (0.141)	1.183*** (0.377)	9.747 (7.097)	6.169*** (2.118)
Year of move	-0.015 (0.017)	0.012 (0.029)	-0.354 (0.300)	-0.079 (0.132)
Apartments per floor	-0.303*** (0.081)	-0.962** (0.404)	-3.816*** (1.464)	0.066 (0.714)
Same floor	0.234** (0.100)	0.224 (0.511)	3.517* (1.828)	3.391 (2.864)
Male	0.332*** (0.115)	-0.686 (0.461)	4.740* (2.435)	4.979* (2.988)
Age	0.014*** (0.004)	-0.003 (0.012)	0.181*** (0.057)	0.187** (0.073)
Apt per floorXMale		0.560** (0.272)		-0.759 (0.619)
Apt per floorXAge		0.009* (0.005)		-0.032** (0.015)
Same floorXMale		-0.198 (0.211)		-2.312 (2.241)
Same floorXAge		0.004 (0.009)		-0.025 (0.034)
Observations	81	81	193	193

Notes: Table reports regressions from columns (1) and (6) in Table A.4 and includes the sex and age of the respondent in the subject's household who answered the household survey. Regressions control for envelope treatment assignment, community, floor of target's apartment \times building height, subject's relocation status, length of residence in the building, monthly expenditures, receiving only one offer for relocation, and height and apartments per floor of previous dwelling. Any missing data is completed with average at the building level, except for subject sex or age. These data are not completed. Thus, observations are dropped due to missing data and correlation with outcomes. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table A.6: Survey measures of quality of social interactions

	All	50% rell.	90% rell.	All	50% rell.	90% rell.
No. apartments per floor	0.068** (0.032)	0.060* (0.036)	0.053 (0.037)	-0.069** (0.032)	-0.057 (0.035)	-0.048 (0.035)
Same floor	-0.320*** (0.032)	-0.297*** (0.035)	-0.299*** (0.039)	0.271*** (0.029)	0.253*** (0.031)	0.249*** (0.036)
Relocated (respondent)	-0.035 (0.051)	-0.023 (0.053)	-0.070 (0.062)	-0.007 (0.050)	-0.022 (0.052)	0.023 (0.056)
Year of move (respondent)	0.019*** (0.006)	0.020*** (0.007)	0.016** (0.007)	-0.012* (0.006)	-0.013* (0.007)	-0.011 (0.007)
Observations	1276	1151	845	1269	1147	856
Respondents	329	297	217	327	296	220

Notes: Table reports marginals of Probit regressions. Regressions control for community, floor of respondent's apartment \times building height, relocation status, length of residence in building, monthly expenditures, receiving only one offer for relocation, and height and apartments per floor of previous dwelling. Any missing data is completed with average at the building level. Stranger equals 1 if respondent says that a neighbor is non-friendly or a stranger and 0 otherwise. Nod and smile equals 1 if respondent says that he/she nods and smiles at a neighbor. Each respondent was asked to respond to these questions for four randomly chosen neighbors. Standard errors are clustered at the respondent level. The survey asked other questions about social interactions, but they suffer from high item non-response. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table A.7: Bounds on the effects of apartments per floor and living on the same floor on returned envelopes

	APARTMENTS PER FLOOR		SAME FLOOR	
	Lower bound	Upper bound	Lower bound	Upper bound
Probit(All)	-0.163 [-0.207,-0.128]	-0.046 [-0.099,0.038]	0.252 [0.188,0.321]	0.245 [0.185,0.309]
Probit(≥ 50)	-0.193 [-0.230,-0.157]	-0.089 [-0.152,-0.001]	0.291 [0.220,0.366]	0.242 [0.187,0.300]
Probit(≥ 90)	-0.194 [-0.216,-0.160]	-0.117 [-0.184,-0.011]	0.255 [0.198,0.314]	0.232 [0.157,0.313]
Logit(all)	-0.849 [-1.048,-0.668]	-0.334 [-0.761,0.346]	1.442 [1.138,1.741]	2.272 [1.743,2.875]
Logit(≥ 50)	-0.891 [-1.061,-0.669]	-0.655 [-1.392,0.251]	1.565 [1.179,1.908]	2.287 [1.739,2.880]
Logit(≥ 90)	-0.995 [-1.203,-0.755]	-1.431 [-2.418,-0.171]	1.666 [1.328,2.011]	2.578 [1.842,3.449]

Notes: Table reports lower and upper bounds on coefficient estimates on apartments per floor and residing on the same floor using the specifications in Table A.4 Columns (1)-(3) and (6)-(8). Lower and upper bounds are reported in each row and 90-percent confidence intervals reported in square brackets.

Table A.8: Selection equation into staying in current location as a function of building characteristics

	Stayed
Apartments per floor	-0.406*** (0.095)
Building height	0.009 (0.033)
Apartments per floor×Building height	0.020*** (0.007)
community 15×Building height	1.163*** (0.360)
community 266×Building height	-0.015 (0.232)
community 471×Building height	0.194 (0.236)
community 703×Building height	-0.471** (0.193)
community 833×Building height	-0.063 (0.267)
community 1164×Building height	-0.212 (0.247)
community 1250×Building height	-0.427 (0.356)
community 266	3.008*** (1.043)
community 471	2.980*** (1.010)
community 703	4.561*** (1.011)
community 833	2.551** (1.092)
community 1164	4.217*** (1.070)
community 1250	3.321*** (1.220)
Relocated	0.088 (0.153)
Monthly expenses	-0.000** (0.000)
Apartment floor	0.106*** (0.030)
Constant	-1.743* (0.913)
Observations	622

Notes: Information on previous and current dwelling location and characteristics are used to estimate a model of selection into staying in current location. Weights are used to estimate a Probit regression following the method of [Xie and Manski \(1989\)](#) to obtain consistent estimates with choice-based sampling. * p<0.10, ** p<0.05, *** p<0.010

Table A.9: Effect of built environment controlling for selection

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	50%	50%	90%	90%
Apts per floor	-0.177*** (0.064)	-0.142** (0.071)	-0.214*** (0.066)	-0.186** (0.077)	-0.256*** (0.063)	-0.227*** (0.071)
Same floor	0.160* (0.090)	0.201** (0.096)	0.198** (0.094)	0.226** (0.100)	0.198* (0.105)	0.245** (0.115)
Inverse Mill's ratio		0.413 (0.374)		0.287 (0.384)		0.384 (0.368)
Observations	116	116	112	112	83	83

Notes: Columns (1), (3) and (5) in this table reproduce corresponding columns from Table A.4. Columns (2), (4) and (6) run the specification to the left of the column and control for selection. All regressions control for envelope treatment assignment, community, floor of target's apartment \times building height, subject's relocation status, length of residence in the building, monthly expenditures, receiving only one offer for relocation, and height and apartments per floor of previous dwelling. Any missing data is completed with average at the building level. All includes all households. At least 50% [90%] relocated only includes buildings with at least 50% [90%] relocated households. The inverse Mill's ratio is calculated using estimates from Table A.8.

B Experimental materials

B.1 Letter content of misdelivered mail

Dear resident,

We are researchers from economics department of Shanghai Jiao Tong University. We are running a research project in your community testing the service quality of various delivery companies. We want to know the operational efficiency and delivery accuracy of those private companies. This letter is one of our testing letters.

Please keep this mail until our research team worker come and pick it up from you. We will pay 10 Yuan to express our thanks for your participation. It is crucial that you keep this information confidential since other members of your community might be part of the study and sharing information will make it hard for us to evaluate the outcomes of the study.

Here is your code to redeem the 10 Yuan cash payment: _____ This code is matched with your apt number _____. This coupon is not transferable and can be only redeemed by you. Also, you might receive other letters like this during the study, please keep them all to redeem for cash payment. Don't forget to keep this information confidential until the study is done.

You only need to keep this letter to redeem your cash payment. We will be coming to your door in about 1 week.

If you have questions about the project, please contact us by phone, text message or email. If we are not available immediately, we will get back to you as soon as possible. XYZ (phone: 111111111, email: xyz@ggg.edu) ABC (phone: 222222222, email: abc@nnn.com)

Thanks for your help.

Economics Department at Shanghai Jiao Tong University, Urbanization project July 2010

B.2 Household survey

1. How many people live in this apartment (#)? Who are they?

2. Is this your only apartment or household? Do you own a car?

3. How much is the average monthly expenses for your family (Yuan/mo)?

4. Employment:
 - a. Where do you work? _____
 - b. What is your position? _____
 - c. Where is the location of your workplace? _____

5. When did you move in this apartment (Month,Year)?

6. Did you move to your current residence because of public demolition and resettlement (yes/no)?
 - a. At the time of the announcement of demolition, how many options did the Housing Bureau offer you?

 - b. Did you wait to make your decision (yes/no)? How many more options did you get by waiting?

 - c. For each option offered (*including current residence*), fill in the following table:

	Option 1	Option 2	Option 3	Option 4	Option 5
Rent or buy?					
Address					
# floors					
# apts/floor					
Size (m ²)					
Cash Offered					
Cash Paid					
On what floor apartment was located					
Direction					

apartment faced (south, north, east, west)					
# entrances					
# elevators					
Shops on 1 st floor (y/n)?					
Entrance facing main street (y/n)?					
Stairwell with window (y/n)?					

7. Did you move to your current residence because of housing assignment by working unit (yes/no)?

a. How many apartments did you get to choose from?

8. I am going to read to you several reasons for moving to your current residence,

a. Tell me which ones apply (*mark all that apply*):

<input type="checkbox"/> Close to workplace	<input type="checkbox"/> Close to downtown
<input type="checkbox"/> Reasonable price	<input type="checkbox"/> Community security
<input type="checkbox"/> Friendly neighbors	<input type="checkbox"/> Good public transportation
<input type="checkbox"/> Good outdoor environment	<input type="checkbox"/> Good apartment design
<input type="checkbox"/> Good community amenities (gym, playground, etc)	<input type="checkbox"/> Close to family/friends
<input type="checkbox"/> Did not look at other options	<input type="checkbox"/> Liked that apartment faced south/east/west/north
<input type="checkbox"/> Wanted to meet new people	<input type="checkbox"/> Spouse liked the apartment
<input type="checkbox"/> Other not mentioned:	

b. Of all the reasons that apply, what was the most important, 2nd most important and 3rd most important reason for moving (*mark 1, 2, 3 next to the top three reasons*)?

9. What was the exact address of your previous permanent residence (not previous temporary residence)?

10. Did you move to your previous residence because of public demolition and resettlement (yes/no)?

C Theory

We present a model of establishing social relationships to develop intuition on how the built environment might impact helping behavior towards neighbors. The built environment determines the spatial distance between the spaces people occupy, i.e. it could be distance between apartments in residential buildings or offices in non-residential buildings. Establishing a social relationship with a neighbor implies a higher likelihood of helping.

There are many models that can predict our results, and many others can predict the opposite. The evidence from our study shows that the comparative statics point in the direction of our proposed model. Some alternative models are discussed at the end of this section.

The model shows that if establishing social relationships is costly, and the built environment adds to that cost, the built environment will impact relationship building. Costly social relationships lead people to be choosier when more alternatives are available. This implies that, abstracting from possible search costs, social relationships should be of higher value whenever more people, or options, are available.

Let the value of a relationship between i and j be $v_{i,j} \sim F([0, \infty[)$, for some distribution F . Let the value of solitude be equal to $c > 0$. We interpret c as the outside option of friends living elsewhere, but it can also represent the cost of getting to know someone. Let N be the number of neighbors i has and let j be any such neighbor. To highlight the effect of friendship selection, we assume that a person selects the one best available neighbor to befriend conditional on $v_{i,j} > c$. This is equivalent to assuming that the cost of a second friendship is infinity.³⁴ To simplify, we assume a friend is chosen at random among those

³⁴The model can be extended to a larger number of friends. This adds complications, but the intuition is the same provided there is a bound on the number of friends.

who are acceptable, thus the probability that i selects j as a friend is:³⁵

$$Pr(\{i, j\}) = \sum_{k=1}^{k=N} \frac{1}{k} C(N-1, k-1) (1-F(c))^k F(c)^{N-k} \quad (1)$$

The expression is the average probability that j is chosen among those who are acceptable. $C(N, k)$ is the combinations of k out of N . Let $p = 1 - F(c) = 0.5$ be the probability that any neighbor is acceptable. In this case, if there is only one neighbor, then $Pr(\{i, j\}) = 0.5$. If i has 5 neighbors instead, we have that $Pr(\{i, j\}) \approx 0.19$. So, even if there are no costs to making friends, the chances of a friendship decrease as i can be more selective. This does not mean that j will not have a friend. In this model, the chance that j is the best choice for some person is $N \times Pr(\{i, j\})$. As groups grow in size, the chances any person has a friend increase, but the chances of two individuals being friends with one another decrease.

The model yields several predictions.³⁶

Prediction 1: The larger the number of potential friends (N) the lower is the probability that two randomly chosen people become friends. This implies that a built environment with more people in a geographical space, i.e. more apartments or offices per building floor, would reduce the chance two people know each other.

Prediction 2: Higher costs (c) reduce the probability that two randomly chosen people are friends. This implies that people living on different floors of a building should be less likely to know each other than two people living on the same floor (*ceteris paribus*).

Prediction 3: Reported familiarity with neighbors decreases with the number of people per building floor. This follows from Prediction 1.

³⁵The probability that j is the only acceptable neighbor is $F(c)^{N-1}(1-F(c))$. The probability that there are two acceptable neighbors and j is one of them is $(N-1)F(c)^{N-2}(1-F(c))^2$. The probability that there are m acceptable neighbors and j is one of them is $C(N-1, k-m)F(c)^{N-m}(1-F(c))^m$. We follow the convention that $C(N, 0) = 1$ to arrive at equation (1). $\frac{1}{k}$ is included in the formula to obtain the average probability of being selected among those who are acceptable.

³⁶If we relax the assumption that a friend is chosen at random among those who are acceptable, another prediction is that the strength of the relationships should be stronger the more selective friendships are.

We note that the model has two effects that determine outcomes: the barrier effect (not being on the same building floor, from different backgrounds, different types) and the selection effect (becoming choosier as alternatives increase). These effects could oppose one another. That is, even if the costs are lower for those living on the same building floor, a larger number of people living on the same floor allows people to be more selective.

Other models of establishing social relationships produce similar predictions. For instance, consider friendship is a public good with value equal 1 that requires at least one person to initiate, or produce, it at cost $c < 1$. Friendship then mimics a volunteer's dilemma game. The symmetric Nash equilibrium of the game is to initiate a friendship with probability $p = 1 - c^{\frac{1}{N-1}}$, which decreases with N and c . This model predicts that relationships are less likely to be produced due to a free-riding problem rather than selection.

D Rerandomization test derivation and procedure

We present the proof of Remark 1 and the derivation and procedure of our randomization test.

D.1 Remark 1

Remark 1. *A necessary and sufficient condition for the sharp null hypothesis $Y(W) = Y_{obs(W)}$ for all W to hold is for no two targets in T to share a neighbor or be neighbors.*

Proof: If no two targets share a neighbor or are neighbors, then any treatment permutation is also an assignment. If, in an assignment, two targets share a neighbor or are neighbors, there is a permutation of treatments that is not an assignment (the permutation that assigns the same neighbor to two targets or the permutation that assigns a neighbor as a subject). The permutation of treatments constitutes a group of transformations of the data over which a permutation test can be conducted (see [Lehmann and Romano, 2005](#), Chapter 15).

D.2 Test derivation and procedure

We introduce some needed notation from [Athey et al. \(2018\)](#). A level set given null hypothesis H_0 , for each individual i and each treatment level w , is defined as: $V(i, w, H_0) = \{w' \in W : Y_i(w') = Y_i(w) \text{ given } H_0\}$. The level set for i given treatment w is the set of treatments w' such that, under the null hypothesis, the potential outcome for unit i is identical to the potential outcome given treatment w . An important implication of [Athey et al. \(2018\)](#) is that we can choose a subset of $\mathcal{T} \subseteq T$, called focal units, where the desired null hypothesis H_0 is sharp. In our context, we want \mathcal{T} to satisfy the following condition:

$$W = \cup_{w \in W} \{\cap_{i \in \mathcal{T}} V(i, w, H_0)\}$$

In words, we want to find a set of focal units where the sharp null hypothesis $Y(W) = Y_{obs(W)}$ holds. Since a randomization test is valid in any such set \mathcal{T} satisfying this condition,

we would like to implement the test on the largest set of focal units. We refer the reader to [Athey et al. \(2018\)](#) for proofs of these results.

From a practical perspective, finding the largest set of focal units may be computationally demanding. In our sample, 241 out of 258 targets (93%) satisfy the condition of Remark 1. However, we can exploit the fact that any set \mathcal{T} satisfying this condition produces exact p -values of the test to implement a computationally straightforward procedure. The procedure is based on a random dictator algorithm.

Randomly sort all targets in T . Initialize the algorithm by assigning the first-ranked target the three neighbors corresponding to each treatment. Remove these neighbors, or subjects, from the dataset and include the target in the set of focal units. Move to the second-ranked target and assign the three neighbors corresponding to each treatment, if available. Include the second-ranked target in the set of focal units and remove the corresponding three neighbors from the dataset. If not all neighbors are available, skip the target and repeat the procedure with the third-ranked target. Continue the procedure until no targets or their neighbors remain.

This algorithm will always include any target satisfying the conditions of Remark 1. But, it will likely include more targets by excluding some targets that, if kept, would violate Remark 1. Significantly, the choice of focal points is not based on either the original treatment assignment or the observed outcomes.³⁷ We can then calculate the desired test in this sample. This constitutes one iteration of the randomization test. The procedure is repeated multiple times by reordering targets and re-running the algorithm to produce a distribution of the test under the null hypothesis. Note that in the limit, i.e., for a sufficiently large number of replications, it will produce a weighted average of the distribution of the test over all possible sets of focal units satisfying Remark 1. Since the distribution of the test under the null hypothesis is the same for each one of these sets of focal points, the average is a consistent estimator of the desired distribution. We implement this procedure to provide

³⁷This is a requirement to obtain the results in [Athey et al. \(2018\)](#).

exact p-values of no treatment effects Table 4.

E Robustness

This section assesses the robustness of our main findings (Table 4) to issues of selection using a Bounds approach and a Control function approach. Both approaches confirm our main findings and yield similar results.

E.1 Bounds approach

The potential magnitude of the bias on the built environment estimates depends on the proportion of non-relocated participants in a building structure (see Manski, 1995). We assume non-relocated participants replace those who chose to move and whose behavior in the experiment is thus unobserved.³⁸ Our main results establish that findings are similar when we restrict our sample to those structures with a prevalence of relocated subjects (Figure 2). We stress test those results by calculating bounds on the effect of the built environment on helping behavior assuming that we do not observe if relocated residents who moved would have helped in the absence of moving.

The calculation of bounds requires knowledge of the relocation status of all participants in our study. With that knowledge, the researcher can estimate bounds on parameters by replacing the behavior of non-relocated participants with the lower and upper bounds of available choices (see Chernozhukov et al., 2013; Andrews and Shi, 2013). This approach is not feasible with our data because we do not know the relocation status for about one-third of participants in our experiment.³⁹ We provide an alternative approach that accounts for these missing data.

We first randomly assign a relocation status to each participant whose data is missing. We assume that the proportion of relocated participants does not exceed the average of the

³⁸We focus on non-relocated participants, as relocated participants do not have a choice of where to move.

³⁹Relocation status is collected by the household survey team, and helping behavior is collected by the envelope survey team. Participants may have responded to one survey but not the other. We find no evidence of differences in helping behavior based on reported relocation status.

relocation proportion observed in the household and community surveys.⁴⁰ Next, we alter the decision of any valid response⁴¹ to either report one envelope (lowest possible choice) or two envelopes (highest possible choice) received for any (observed or assigned) non-relocated participant. The responses of relocated participants remain unchanged. Any participant with missing covariate data is assigned the average characteristics of the relocated/non-relocated participants in the same building. Once this is done, we re-estimate the model and repeat this calculation 1,000 times.

Our approach does not point identify the lower and upper bound of the estimates due to missing relocation status data. Thus, we simulate the possible values. The average estimate and 90% CI of the effect of apartment per floor when reported envelopes are changed to one and two, corresponding to all six statistical models in Figure 2 and Table A.4, confirm that our estimates in the subsamples with mostly relocated participants are robust to selection/survivor bias (see Table A.7). One more apartment per floor reduces the probability of returning the envelope by 12 to 19 percentage points. This is a reassuring result, given that we modify the behavior of 15 percent of the sample in this subsample.

The reader might wonder whether we should use the minimum of the lower bound estimate and the maximum of the upper bound estimate as the relevant bound estimates, i.e., select the set of participants whose decisions have changed status to relocated to minimize/maximize estimated parameters. We do not follow this approach because it confounds missing data and selection issues. We do not have evidence against data missing at random (Table A.1). Taking the minimum and maximum values as estimates of our bounds would

⁴⁰We use the average of these two measures since they are the only sources of information on relocation status. The two measures correlate significantly ($\rho = 0.48, p\text{-value} < 0.0001$).

⁴¹This is done to calculate bounds. Thus, we generate outcomes in the counterfactual that choices were different since we cannot observe the real outcomes of those who departed. We change only valid responses, i.e. a response of having received 1 or 2 envelopes, to maintain the structure of our data. We assume that non-response bias would be similar if the original relocated dwellers had participated in the experiment. The propensity to report zero envelopes is not statistically significantly different between relocated and non-relocated participants ($p\text{-value} = 0.6162$). We remark that our Probit estimates exclude zero-envelope response and, therefore, are not affected by this choice.

be equivalent to assuming that relocation status and missing covariate data are selective.

E.2 Control function approach

A second pathway to assess the potential magnitude of the survival bias is to use a control function approach (Heckman, 1979). In this approach, we need to account for the decision to remain in a built environment that one was relocated to or chose to live in. In particular, we model the decision to stay in the current dwelling as a discrete choice:

$$Pr(S_i|Z_i) = \mathbf{1}[Z_i'\beta + u_i \geq 0] \quad (2)$$

where S_i equals one if a person voluntarily remains in a dwelling and zero otherwise, Z_i is a set of covariates, and u_i is distributed normally. The decision to return a piece of mail is modeled as:

$$Pr(R_i|X_i) = \mathbf{1}[X_i'\beta + e_i \geq 0] \quad (3)$$

where R_i equals one if a subject returns the mail and zero otherwise, X_i is a set of covariates, and e_i is distributed normally. We assume that u_i and e_i are jointly distributed normal. As in Heckman (1979), these assumptions allow us to obtain consistent estimates by including as an additional covariate a control function equalling $E[u_i|Z_i, S_i = 1]$.

We do not observe the behavior of those who moved out of a building. Thus, to implement this approach, we take advantage of the fact that our household survey collects information on an individual's previous residence location and characteristics, as well as their current residence location and characteristics. The survey also collects information on whether the individual has moved to and from their previous residence by forced relocation or choice. We thus have two decisions of whether or not to move from each respondent: from their previous residence and their current residence. We use those two decisions in the analysis. We classify relocated respondents as non-movers in both their previous and current residences. Since

they were forcibly removed from the previous address, we conclude that they did not want to move in both situations. Non-relocated respondents are classified as movers from their previous residence and non-movers at their current residence. These data allow us to estimate the effect of relocation status on the probability of staying in a dwelling.

Finally, the household survey gathers information on the location of current employment and household characteristics. We use geographic information to calculate travel distance to work from previous and current residences, assuming that employment location remains the same. We exclude distance to employment location from the outcome equation and use it to predict voluntary changes in residence. We do this to avoid identification based only on functional form restrictions.

There are two main challenges to implement our approach. First, we have a choice-based sample rather than a random population sample. Second, the architectural features of the previous residences have different supports. To address the first challenge, we follow [Xie and Manski \(1989\)](#), who show that consistent estimates of Probit models in choice-based samples can be obtained by reweighting observations based on observed outcomes. We obtain probabilities of switching residence in our sample from the distribution of relocation status. To address the second challenge, we retain the same covariates as in the selection and outcome equation but use a simpler parametric specification in the selection equation. We use information on distance to work to produce independent variation in the selection equation. Finally, we restrict our analysis to the Probit regression because correcting for selection in this case follows standard methods.

We estimate a Probit model of the decision to stay in a dwelling. The probability of staying decreases with the number of apartments per floor, but that relationship is less negative in taller buildings (Table [A.8](#)). These results confirm that the characteristics of the building can affect the likelihood of remaining in the sample. The interaction terms between distance to location of employment and community are highly significant ($\chi^2(7) = 18.96$, p-value = 0.0083) and mostly indicative that the further a person lives from their location

of employment the more likely they are to move out.

When we include the control function in the main regression specification (Table A.4), the estimated effect of apartment per floor on returning the envelope is slightly smaller, but it remains significant (Table A.9 and Figure A.6). One more apartment per floor reduces the likelihood of returning the envelope by 13 to 22 percentage points, depending on the specification. In none of the specifications do we find that the control function is significant. As discussed, this is not because the selection model does not predict the decision to stay in a dwelling.

References

- Donald WK Andrews and Xiaoxia Shi. Inference based on conditional moment inequalities. *Econometrica*, 81(2):609–666, 2013.
- Susan Athey, Dean Eckles, and Guido W. Imbens. Exact p-values for network interference. *Journal of the American Statistical Association*, 113(521):230–240, 2018.
- Victor Chernozhukov, Sokbae Lee, and Adam M Rosen. Intersection bounds: Estimation and inference. *Econometrica*, 81(2):667–737, 2013.
- James J Heckman. Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, pages 153–161, 1979.
- Erich Leo Lehmann and Joseph P Romano. *Testing statistical hypotheses*. Springer, 2005.
- Charles F Manski. *Identification problems in the social sciences*. Harvard University Press, 1995.
- Yu Xie and Charles F Manski. The logit model and response-based samples. *Sociological Methods & Research*, 17(3):283–302, 1989.

