Demand for Urban Exploration: Evidence from Nairobi*

Joshua T. Dean, Gabriel Kreindler, Oluchi Mbonu †

November 1st, 2024

Abstract

Growing cities in low- and middle-income countries offer increased market access, yet taking advantage of urban opportunities requires that residents explore their surroundings. This is not always the case. In a sample of 800 casual workers in Nairobi, the median person commutes 7.8 km but has never been to half the neighborhoods at most 75 minutes from where they live. We next study the barriers to exploring urban locations. We offer short-term employment to workers in our sample and experimentally induce familiarity by training participants in either familiar or unfamiliar locations. We measure willingness to work in different locations across the city. Participants need to be paid more to work in a neighborhood that is unfamiliar at baseline. The premium is equivalent to 3.5 km of distance or to 108 Ksh (22% of the median daily wage), and this is fully offset after one visit. Participant beliefs about labor market opportunities and safety in unfamiliar neighborhoods are initially worse on average, but converge after one visit. We find little evidence that risk aversion can explain these results. We use additional job choice elicitation methods to show that participants only partially anticipate that one visit will eliminate the familiarity premium, and to show that unfamiliar neighborhoods are less "top of mind." Participants return on their own to the unfamiliar neighborhoods where they were trained to search for work and for other non-work reasons, and they are more likely to show up for a different paid opportunity 2-4 months after the intervention. Our results suggest that one-time exploration frictions are an important component of urban mobility costs in cities like Nairobi.

*We are grateful to the Busara Center for Behavioral Economics. We thank Emily Breza, Fiery Cushman, Ed Glaeser, Rema Hanna, Melanie Morten, Gautam Rao, Devin Pope, Alex Imas, Avner Strulov-Shlain for useful conversations. Martin Serramo, David Quirroz Barrios and Rafael Tiara provided exceptional research assistance. The project has IRB approval from Harvard University and AMREF. It is registered on the AEA Registry (AEARCTR-0011205). We gratefully acknowledge funding from the JPAL Jobs and Opportunity Initiative, Harvard University, The Becker-Friedman Development Economics Research Fund, and The University of Chicago.

†Dean: University of Chicago (joshua.dean@chicagobooth.edu), Kreindler: Harvard University (gkreindler@g.harvard.edu), Mbonu: Harvard University (ombonu@g.harvard.edu).

1 Introduction

The rapid urbanization of low and middle-income countries has been seen as an opportunity for increased market access (Bryan, Glaeser, and Tsivanidis 2020). To date research in economics has argued that for these benefits to occur, transportation in cities must be quick, convenient, safe, and affordable (Lall, Henderson, and Venables 2017; Borker 2024; Davis 2021; Tsivanidis 2023). However, in addition to having the ability to cheaply access different parts of the city, individuals must also be willing to explore different areas in the city. In this project, we use a randomized experiment in Nairobi to study people's willingness to explore new urban locations.

We first document that individuals have substantial gaps in their experience with neighborhoods near their homes. The median participant in our sample commutes 7.8 km to search for work or to work; however, the median participant has never been to one in two of the neighborhoods within 75 minutes away where they live. They can remember a landmark for one in three such neighborhoods. This is not due to a lack of awareness; the median participant has heard of 94% of the neighborhoods in this radius.

We then study the demand for exploring unfamiliar neighborhoods in a real-stakes setting by offering short-term employment opportunities to a sample of casual workers. First, we use these jobs to experimentally induce familiarity. We train participants on how to do the employment task in three neighborhoods over three days. In the treatment group, participants are trained in unfamiliar neighborhoods. These are neighborhoods where participants report at baseline to have never previously been. In the control group, participants are trainined in familiar neighborhoods.

Second, we measure willingness to work in different locations across the city in three ways to study different aspects of how of familiarity affects demand. In order to improve realism we use a task that transparently both needs to be done in a particular location and where it is plausible that we would want the task done in a variety of locations: measuring air quality.

We first consider how familiarity affects willingness to accept a job when it is offered. To do so we confront half the sample on the morning of every work day with binary choices between jobs for the day in different locations while randomly varying duration and compensation. Our first result is that participants need to be paid more to work in a neighborhood that is unfamiliar at baseline. The premium is equivalent to 3.44 km of distance or to 108 Ksh (22% of the median daily wage).

Our next result is that experimentally inducing a visit to a previously unfamiliar neighborhood during training is sufficient to completely eliminate this familiarity premium. In order to account for the potential that any training neighborhood may be more salient, we compare the willingness to work in a training neighborhood that was unfamiliar at baseline with a training neighborhood that was familiar at baseline. We find that above and beyond the salience effect of training, visiting an unfamiliar neighborhood once is equivalent to increasing the wage by 113 Ksh or bringing the neighborhood 3.64 km closer, fully offsetting the baseline familiarity premium. In a counterfactual exercise, we find that moving everyone in our sample to universal familiarity would increase a standard measure of commuter market access by the equivalent of reducing all distances in Nairobi by 17%. We also note that when job outcomes are uncertain, the wage differential that would induce someone to be willing to search in an unfamiliar neighborhood are amplified.

We next show that participants are on average pessimistic about unfamiliar neighborhoods, and these imbalances disappear after an experimentally induced visit to a previously unfamiliar neighborhood. We ask the entire sample at baseline for their beliefs about the employment opportunities for themselves and people like them in the neighborhoods and the safety of traveling to and being in the neighborhood. We find that individuals are more negative across all outcomes for unfamiliar neighborhoods. However, when we ask individuals the same questions after having visited for training, there is no difference between neighborhoods that were initially familiar or unfamiliar: one visit is sufficient to close the beliefs gap. We also show that there is a significant degree of alignment in how participants rate neighborhoods, including after an experimentally induced visit to an unfamiliar neighborhood.

We then examine three potential barriers to exploration. First we consider whether our results can be explained by risk aversion. While we cannot rule this out completely, we believe three pieces of evidence weigh against this explanation. First, the experiment has minimal risk in that employment and compensation are guaranteed and subjects are only in the location for an hour or two during the middle of the day. We also find little evidence of first-time navigation costs such as longer travel time. Second, we find that almost no subjects update negatively about safety after visiting the neighborhood. Finally, we use the Chernozhukov et al. (2023) machine-learning method of searching for heterogeneous treatment effects to see if any subjects were less willing to return to neighborhoods after visiting them. We find that the point estimates of all treatment effects are positive and we can reject any significant negative effects.

We then study two additional potential barriers to exploration: anticipation and consideration. We ask whether individuals anticipate that their willingness to work the neighborhood will change after one visit. If individuals do not fully anticipate how visiting a location will lower the cost of re-visiting it, this may lead to under exploration. We pose the same binary job choice questions on training days, after participants have been informed of where they will train but before they have actually visited the location. We find that individuals incompletely anticipate the benefits. On average, subjects make choices as if the treatment effect will be approximately 60% as large as the truth.

We also measure how familiarity with a neighborhood affects how likely someone is to think about that place as an option. If subjects are less likely to consider working in unfamiliar locations, they may be less likely to explore above and beyond what we would expect based on the preferences we measure in our previous "closed" elicitation where we ask about specific locations. To assess this possibility we use an "open" elicitation for half of the sample where subjects must generate potential employment locations rather than being confronted with binary choices. Specifically, subjects are told that there are jobs available in different neighborhoods in Nairobi, and asked to tell us where they would most like to work, followed by where they would next like to work, and so on. Subjects are told that we will assign them to the highest neighborhood on their list where a job is available. Estimating discrete choice models on this ranked data reveals a larger role for familiarity than in the "closed" elicitation. We estimate a model with both preferences and memory costs, and find that baseline familiarity and the training significantly affect memory costs. Our results show that people struggle to consider unfamiliar neighborhoods as potential places to work.

Finally, we use a combination of strategies to assess whether the training intervention caused individuals to be persistently more likely to re-visit neighborhoods that were initially unfamiliar. First, we measure whether participants in our sample take up paid opportunities two to four months after the intervention. We invite each participant to a short survey, varying the amount they will get paid and the neighborhood where they need to show up. This is a take-it-or-leave-it offer, allowing us to measure extensive margin decisions. We repeat this type of exercise six times per individual, varying the neighborhood and wage each time. We find that baseline neighborhood familiarity strongly predicts show-up (all else equal), and the experimentally induced familiarity is positive, of a similar size, yet somewhat imprecisely estimated.[1] These results suggest that exposure to a baseline unfamiliar neighborhood has persistent effects in the medium run.

In our endline survey and SMS data, across both prompted and unprompted measures, we find that individuals return to the initially unfamiliar neighborhoods where they were trained for a variety of reasons, including searching for work and non-work reasons, covering shopping, fun/leisure, healthcare, and errands. These results are a revealed preference

[1] These results hold for wage offers $\leq$ 500 KSH. For wages $\geq$ 600 KSH, show-up is flat around 80% and none of the job and neighborhood attributes predicts show-up. This ceiling effect on show-up was noted during the field work and the wage distribution was lowered to eliminate it. For completeness we present the full set of results, but focus on the wage levels where job offer attributes affect decisions.

argument that participants found exposure to these neighborhoods valuable.

Overall, we document that individuals strongly prefer not to work in neighborhoods that they have never visited and they hold pessimistic beliefs on average about these places, but that one visit is sufficient to eliminate the preference and belief gap. We also show that exploration may be additionally hampered by incomplete anticipation of the reduced cost of revisiting a neighborhood and the way individuals form consideration sets. Together this suggests that individuals' demand for exploration is an important component of urban mobility costs and urban accessibility.

Our paper is related to work on urban agglomeration in developing countries (Duranton 2015; Chauvin et al. 2017; Lall, Henderson, and Venables 2017; Bryan, Glaeser, and Tsivanidis 2020), to work that measures and unpacks transportation costs in these contexts (Borker 2021; Vitali 2024; Kreindler et al. 2023; Grosset-Touba 2024; Tang 2024; Jalota and Ho 2024), and to research on policies that reduce transportation costs (Habyarimana and Jack 2015; Gonzalez-Navarro and Turner 2018; Tsivanidis 2023; Balboni et al. 2021; Zárate 2022). Our contribution is to focus on a new component of urban mobility costs, the one-time cost of exploring a location for the first time.

Migration is another spatial choice that often requires exploration. Our paper is also related to work on subsidizing migration, the persistent effects of migration, and on information frictions in migration (Bryan, Chowdhury, and Mobarak 2014; Okunogbe 2024; Baseler 2023; Porcher 2022; Wiseman 2023; Porcher, Morales, and Fujiwara 2024; McKenzie 2024).

Our paper also relates to work studying the interplay between transportation and labor and education markets in cities in developing countries (Franklin 2018; Abebe et al. 2021; A. Banerjee and Sequeira 2023; Carranza et al. 2022; Abel et al. 2019; Belot, Kircher, and Muller 2019; Agte et al. 2024).

Finally, our paper builds on a diverse literature in economics, finance and psychology studying attention and experimentation (Scharfstein and Stein 1990; A. V. Banerjee 1992; Eyster and Rabin 2010; Hanna, Mullainathan, and Schwartzstein 2014; Larcom, Rauch, and Willems 2017; Malmendier 2021; Zajonc 1968; Montoya et al. 2017)

2 Context and Participant Sample

Our study takes place in Nairobi, a city of 4.4 million people that doubled in size over the past two decades (KNBS 2019). Much of this growth has happened in an unplanned fashion. For example, informal settlements now house about 70% of the population (Gachanja et al. 2023). This growth has also resulted in a city that is poorly served by its transportation infrastructure. Avner and Lall (2016) estimate that a more efficient layout could double residents' labor market access. This is typical of the rapid and disconnected growth found in many African cities (Lall, Henderson, and Venables 2017).

2.1 Sample: Casual, Underemployed Workers

We recruited study participants in three low-income neighborhoods (Kibera, Kawangware, and Viwandani, see Figure 1) on a rolling basis between October 2023 and January 2024. Surveyors recruited participants in person, and participants who passed the initial filter were later invited to a study venue for two days of additional surveys. The two surveys included demographic and employment questions, urban familiarity questions (discussed in section 3.1), and several measures of spatial ability. In order to limit selection into the study sample based on willingness to explore, each of the three home neighborhoods where we recruited had its own study venue located within the neighborhood.

A participant was eligible for the study if they were older than 18 years old, lived in one of our three study neighborhoods, were casual workers actively searching for work and stated that they were available to work every day for the next 7 days. We define a casual worker as someone who does not work as a permanent employee – but rather does short-term daily contract jobs. A participant is classified as actively searching for work if they searched for work for at least three days in the past two weeks. Participants who did not show up for the two initial surveys at the study venue within a day of the invitation date were excluded from the sample. Some participants with very high or very low familiarity of Nairobi, based on their responses in the first baseline survey, were also dropped because we could not randomize them appropriately to the familiar or unfamiliar training locations. We discuss these eligibility requirements in section 4.

We collected information from 1704 participants during the in-person recruitment. Of these participants, we invited 1600 who were eligible to participate in a baseline survey. Among these eligible participants, 1168 attended the first baseline survey, where 915 were eligible for the study intervention and thus invited to a second baseline survey. Of the 915 eligible at Baseline 1, 831 attended Baseline 2, and 799 of them began the first training day.

Sample Descriptive Analysis Table A.2 reports descriptive stats for the experimental study sample. The median participant is a rural migrant who has been living in Nairobi for 13 years. Three quarters are women. This sample is considerably under-employed. Participants have searched for work an average of 6.6 days in the past two weeks and only worked half that. They also travel a significant distance to work or search for work, with a median commute distance of 7.8 km. The most common job occupations for women involve working for other households doing laundry, cleaning or cooking (Table A.1). Occupations for men are more heterogeneous and typically manual and semi-skilled work such as carpenter, mason, factory work and electrician work.

Table A.3 shows that referrals are the most common way that this group searches for and finds employment, yet participants also rely significantly on spatial search strategies such as going door to door, traveling to "hiring spots", and visiting potential employers to submit resumes.

3 Neighborhood Familiarity Patterns

Throughout this paper, we say a participant is *familiar* with a neighborhood if they report having ever been to that neighborhood in the past. In this section we first discuss how we measured familiarity in our sample, and then present several stylized facts.

3.1 Measuring Familiarity

We partition Nairobi into the 61 neighborhoods with commonly used and recognized names. We coordinated a mapping team of Busara Center employees with in-depth knowledge of Nairobi and asked them to generate the neighborhood names and boundaries and to seek input from field guides in various neighborhoods when necessary. Figure 1 displays the resulting neighborhoods. We tested and piloted the neighborhood names extensively to make sure they are broadly and reliably recognized by the population which we sample. The mapping team also generated 341 sub-neighborhoods within these 63 neighborhoods. Most of our analysis uses the main neighborhoods.

For each home neighborhood (Kibera, Kawangware, and Viwandani), we elicit familiarity from a set of neighborhoods which are within 75 minutes walking or by transit (whichever is shortest) from the study venue in that neighborhood, based on data from Google Maps (Figure A.1 plots the set of neighborhoods we ask about for each home neighborhood). This led to lists of 33, 30, and 31 neighborhoods for the three home neighborhoods, respectively.

In the first baseline survey, we ask participants about all neighborhoods in this list, randomizing the order. We initially loop over all neighborhoods and ask two types questions for each neighborhood:[2]

1. Have you ever been to the neighborhood of X?
2. (if "yes") When was the last time you went to the neighborhood of X?
2. (if "no") Have you ever heard of the neighborhood of X?

[2]This design ensures that each participant answers the same number of questions regardless of their answer to the first question, which avoids incentives to misreport in order to change the duration of the survey.

We use responses to the first question as our main measure of familiarity due to high test-retest reliability in piloting. We then collect additional data on each neighborhood. We ask participants if they know how to get to X, and ask them to tell us a location, landmark or road in the neighborhood of X.

3.2 Neighborhood Familiarity Patterns

We begin with a descriptive analysis that shows that participants in our sample have significant "spatial familiarity gaps." We focus on familiarity patterns for neighborhoods that are objectively accessible, within our sample of neighborhood at most 75 minutes away from the respondent's home neighborhood.

Figure 2 displays the CDF for our main measure of familiarity. It shows that the median participant has never visited around half of the 30-33 neighborhoods in the sample. A quarter of participants have visited less than 40% of the neighborhoods.

Table 1 reports results from different measures of familiarity and varying the sample of neighborhoods around the respondent's home neighborhood. The first three columns reports results for the sample of neighborhood within 75 minutes (like in Figure 2), and the last three columns further restrict to only neighborhoods within 7.8 kilometers of the participant's home neighborhood, which is the median distance that participants report traveling in order to work or search for work.

The table highlights several results. First, participants have heard of almost all of the neighborhoods. The average rate is 92% and the median participant has heard of 94% of neighborhoods within 75 minutes. These numbers are slightly higher for the second (smaller) sample of neighborhoods. This shows that unfamiliarity is not driven by confusion about the names we use to refer to neighborhoods.

Second, familiarity is low and even lower for more demanding definitions of familiarity. Focusing on the second column, the median participant reports ever having been to *or passed by* 63% of the neighborhoods, and even been to 52% (Figure 2 uses this definition). We later also ask participants if they know how to get to the neighborhood and whether they can list a landmark for the neighborhood, regardless of whether they have ever been there. The median respondent only has 43% and 34% of the neighborhoods (around 13 and 10 neighborhoods, respectively) that they have ever visited and that they know how to reach, or that they can give a landmark for, respectively.

Third, familiarity increases as we restrict to the sample of neighborhoods closer than the median commute distance, which is 7.8km. This restricts the sample of neighborhoods to 10-16 neighborhoods that are closest to the home neighborhood. The median participant has visited 69% of these neighborhoods, although they can also name a landmark for only 48% of the neighborhoods in this restricted sample. Note that half of our respondents travel further than this cutoff on a regular basis to work or search for work, so the results in the last three columns of Table 1 offer an upper bound on familiarity of objectively accessible neighborhoods.

To provide a more granular view of familiarity patterns, Figure 3 displays the levels of familiarity for each participant from Kibera. Each row is one of the 30 neighborhoods and each column is a respondent, with both axes sorted by average level of familiarity. A cells is blue if the respondent has been to the row neighborhood and black if they have not.

The key takeaway is that there is a significant amount of idiosyncratic variation in familiarity patterns. There are few neighborhoods that are very familiar or very unfamiliar, and vice-versa there are few participants who know most or none of the neighborhoods. This both is suggestive that familiarity patterns are not driven by consensus views of neighborhood attributes and enables us to include neighborhood fixed effects in our estimation of the premium later.

To evaluate whether this variation is driven by individual-neighborhood match characteristics, we fit a series of four random forest models that flexibly predict whether individual i is familiar with neighborhood j. Figure 4 presents ROC curves and the area under the ROC curves, which can be interpreted as the probability that the model correctly ranks which of two individual-neighborhood pairs is more likely to be familiar than another. We begin with the standard gravity predictors of neighborhood fixed effects and distance from home neighborhood. These predictors do significantly better than a chance ranking of 0.5, giving an 86% chance of correctly ranking the pairs. We then add individual characteristics, neighborhood characteristics and both at the same time. These additions (and their interactions allowed by the random forest) do nothing to improve our accuracy. This suggests that, as far as we can observe, individual-neighborhood match characteristics are not important drivers of familiarity.

To understand more about what types of people are more likely to have higher levels of familiarity, Table A.5 regresses the average level of familiarity at the individual-level on several demographic indicators and an index of spatial ability. We find that men, those that are older, those with more years of education or years living in Nairobi, and those with higher spatial skills have greater average levels of familiarity. In general this heterogeneity is meaningful, but not overwhelming. For example, men are 0.3 SD more familiar than women on average and an additional 10 years in Nairobi is associated with an increase of 0.1 SD.

Table A.6 studies whether the respondent sharing an ethnicity with the plurality of the population of a neighborhood is an important driver of familiarity. We regress an indicator for whether an individual has ever been to a neighborhood on distance from their home and an indicator for whether they share the same ethnicity as the plurality of the neighborhood. While shared ethnicity alone is predictive, this effect vanishes after controlling for distance. In addition to being insignificant, the point estimates in columns 3 and 4 suggests that sharing the same ethnicity as the plurality of the neighborhood is approximately equivalent to a neighborhood being between a sixth and a third of a kilometer closer.

4 Experimental Overview

In order to experimentally identify the effect of familiarity, we want to cause individuals to travel to unfamiliar locations. In order to focus on the effect of familiarity, we need to avoid inducing experimenter demand effects, and we need to be able to control for the increased salience of places where the respondent has recently been.

Air Pollution Measurement Jobs. To accomplish these objectives, we offer short-term job opportunities to participants in our sample in different locations throughout the city. In these jobs, participants are asked to collect data on air quality in a specific neighborhood. These are jobs that obviously require being in a specific location, and they make it plausible to vary the location where the participant is working. We use these jobs for two purposes. First, we subtly induce familiarity with certain neighborhoods by randomizing whether training for these jobs takes place in familiar or unfamiliar locations. Second, we elicit participant choices over jobs. Given that we compensate participants for the work they do, these choices are real-stakes.

To complete the task individuals wear a backpack used in Berkouwer and Dean 2024 shown in Figure A.3 that contains a PM 2.5 or CO sensor and a smartphone from a rented venue in their home neighborhood to an assigned location. After arriving at the location, participants use an app on the phone to begin data collection and confirm they are in the correct place. They remain outdoors for one to two hours and then return to the study venue. In order to ensure subjects were not simply unable to get to unfamiliar neighborhoods, all participants were offered paper directions to the locations.

Participant Timeline and Randomization. Figure A.4 presents an overview of the study timeline from the participant perspective.

The first two days consist of baseline survey data collection. On the first day, participants are invited to the study venue, where they participate in the first baseline survey providing demographic information and answering the familiarity questions discussed in section 3.1.

On the second day, participants return to the study venue and complete the second baseline survey, which collects more demographic data, data on networks, more detailed employment data, data on self-reported spatial ability, and beliefs about the labor market and safety of neighborhoods. Participants who complete this second survey form the analysis sample for the experiment.

We then randomize participants into completing the job training in familiar neighborhoods (control group) or unfamiliar neighborhoods (treatment group). We cross-randomize the method by which we will elicit their work location preferences, which we discuss below. We stratify the randomization by home neighborhood and familiarity level (above- and below median).

The main intervention takes place over the next three days. Participants begin the day at the study venue where they answer a short survey, then get trained on how to complete the air quality task by a field guide, visiting a different neighborhood each day. The participant returns to the study venue at the end of each day to return the air pollution measurement equipment and answer a short survey.

Participants work unaccompanied for the following three days. They again start and end the day at the study venue in their neighborhood, where they answer short surveys before and after the job.

Participants are later contacted by SMS and by phone for endline data collection at least a month after they begin training. (We discuss this study component in section 7.)

Target Neighborhoods. After the first baseline survey, we use an algorithm to select 10 neighborhoods for each participant, which we henceforth refer to as "target neighborhoods." We select 6 familiar neighborhoods and 4 unfamiliar neighborhoods for each person. We then randomly designate 3 familiar and 3 unfamiliar neighborhoods as "main" familiar and "main" unfamiliar neighborhoods, respectively. Figure A.5 plots the target neighborhoods of one participant to illustrate.

Participants will complete their training on the air pollution job in neighborhoods from one of these groups, depending on their treatment assignment. Control participants will visit the three target main familiar neighborhoods, while treatment participants will visit the three target main unfamiliar neighborhoods.

The remaining *familiar* neighborhoods allow us to identify the salience induced by visiting an already familiar neighborhood. The remaining *unfamiliar* neighborhood allows us to identify any potential spillover effect of increased willingness to visit unfamiliar neighborhoods.

We select the ten target neighborhoods to minimize spatial spillovers between categories, to keep neighborhoods in each category close to each other, to balance distance from home to each category, and to prioritize nearby neighborhoods. For each participant, given their pattern of familiar and unfamiliar neighborhoods, we run an optimization algorithm to minimize the weighted sum of several cost components. First, we penalize spatial spillovers between the main familiar, other familiar, and all unfamiliar neighborhoods. We code such a spillover as happening when a 500-meter buffer around the Google Maps walk or transit route to a neighborhood j intersects a neighborhood k. Second, we penalize distance between neighborhoods within each group. Keeping neighborhoods within each group close to each other enables participants in the treatment group to develop deeper familiarity with a new area of a city. Third, we penalize differences between groups in the average distance from the home neighborhood to neighborhoods in that group. Finally, we penalize longer distance from the home neighborhood to each neighborhood. We drop participants with less than four unfamiliar neighborhoods, and those with less than six familiar neighborhoods because we are unable to choose the correct number of target neighborhoods for them.[3]

Inducing Familiarity using Job Training. On each of the three training days, participants visited their assigned neighborhoods and received training on how to perform the air pollution task. Each participant was accompanied by a field guide.[4] Field guides used a prespecified route to reach the assigned location, and they were instructed to point out to the participant landmarks along the route and generally ensure that the participant understands the route (in case they need to return there). Once in the neighborhood, they instructed the participant how to collect data (stay outdoors, stay in vicinity of major roads, stand at the same location or move around as they prefer, and how to use smartphone app). Field guides then left the participant to continue their task in the neighborhood. Each participant was informed that they would return on their own from the training neighborhood. Participants were paid 700 Ksh for each training day, and given 200 Ksh for transportation.

In order to minimize experimenter demand effects, participants were told that the purpose of the study was to "to better understand how casual workers travel in Nairobi and how this affects their search for work opportunities." We also did not make explicit that we were randomizing training locations based on familiarity, instead telling participants "The neighborhood you travel to with the field officer will be chosen randomly." Surveyors and field guides were also not informed of individual participants' treatment status and field

[3]Starting with a sample of 1168 participants who showed up at Baseline 1, we excluded 119 because they had six or fewer familiar neighborhoods, 133 because they had four or fewer unfamiliar neighborhoods, and one participant who had fewer than four unfamiliar and six familiar neighborhoods. Consequently, we invited 915 participants to continue in the study.

[4]Field guides were recruited by Busara from each home neighborhood and trained for this study.

guides were not informed of the design or purpose of the experiment.

Compliance with the training assignment was almost universal (Table A.9). On average, 98.7% of participants who showed up for their first day complied with their treatment assignments and completed the three days of training. However, there is a small imbalance between treatment and control group, with 0.5-1.3 p.p. more participants in the treatment group refusing to travel to their assigned neighborhoods, with this difference growing over time. All analyses are based on intention to treat.

We re-measure familiarity when participants return from the training to the study venue. We ask each participant if they had already been to the assigned neighborhood before the day of the training. Table A.10 displays the results. In the control group, 88% of respondents report having visited that neighborhood before the training day, which means that 12% had never been there, despite reporting in the baseline survey that they had. In the treatment group, 34% of participants report having already traveled to the assigned neighborhood, despite reporting in the baseline survey that they had never been. Anecdotally, this is primarily due to either disagreement about the name of a location or visiting the location reminding the treatment group of a time that they had visited.

In the rest of this paper, we report intent to treat (ITT) results using the baseline familiarity measure. Depending on the source of discrepancy between the original data collection and the revised measure, and based on the desired notion of familiarity, it may be appropriate to use the treatment assignment to instrument for familiarity. This is roughly equivalent to inflating our later estimates by a factor of 1.85 (which is one divided by the first stage effect of 0.54). For example, this would be the case if the source of discrepancy is that participant make idiosyncratic mistakes when they respond to our baseline survey. However, to be conservative, we do not compute the IV in our analyses.

Eliciting Job Preferences and Attention. We randomize the way we elicit job choices during the three employment days that follow the training period in order to study the relative importance of familiarity in shaping preferences of where to work and the formation of consideration sets.

In the first elicitation method, we use "structured" binary choices where individuals choose between two potential jobs given information about their location, duration, wage, etc. These structured choices allow us to price the disutility of traveling to an unfamiliar location, but directly confront the individuals with the possibility of working in unfamiliar neighborhoods.

However, people may also be less likely to consider unfamiliar neighborhoods in the first place. To measure this effect, in the second, "unstructured", elicitation method we tell participants that jobs are available in different neighborhoods across Nairobi and ask them for their most preferred location, the next most preferred location and so on. If it is the case that unfamiliar neighborhoods are less likely to enter individuals' consideration sets, we should expect that respondents rank unfamiliar neighborhoods even lower than we would expect based on their preferences alone.

The other additional barrier we study is whether individuals anticipate how their utility of visiting a neighborhood will change after having visited it once. If individuals fail to anticipate that the first visit will reduce the cost of subsequent visits, they may underestimate the benefits of exploration. To study this we ask participants in the structured group to make the same choices on the training days as they will on the employment days. Participant make these choices on a training day in the morning, before they have visited the neighborhood but after they have learned that they are about to. If they under-anticipate the change in their utility, we would expect these participants to be more willing to work in the unfamiliar neighborhoods after having actually visited than they are immediately before visiting.

5 Impact of Neighborhood Familiarity: Results

5.1 The Revealed Preference for Familiar Neighborhoods

Job Choice Elicitation. We begin by discussing results for the half of our study sample where we elicit job preferences in a structured manner. After the three days of training, participants are invited to the study venue for three additional days of employment. Each morning, they answer a series of binary choice questions about the air pollution job they will perform that day. Each question presents two variants of the job and asks the respondent which one they prefer for that day. For each option, we randomly vary the neighborhood where the job will take place – selecting from among the ten target neighborhoods for that respondent – and other job attributes, the wage, job duration, upfront cash to finance transportation, and how much of the compensation is in the form of a risky bonus payment.

Participants are informed that one of the questions will be randomly picked and implemented. To ensure and check comprehension, before answering this type of question for the first time, the surveyor goes through a practice session with the participant. [5]

Participant responses to these real-stakes questions allow us to estimate preferences for working different types of neighborhoods, the effects of the training, and the impact of in-

[5]Respondents also answer a similar set of questions on the morning of each training day. Questions on training day $t \in \{1, 2, 3\}$ refer to jobs on day $t+3$, and these questions have an equal chance of being selected to be implemented. That is, each question asked on a "work" day has an equal probability of 1/14 to be selected. There is a 50% chance that the selected question is one of those seven asked three days prior.

ducing familiarity via training. By design, each question forces the respondent to consider working in the two neighborhoods. We interpret this as meaning that we estimate preferences, and we return to the issue of whether unfamiliar neighborhoods are less likely to enter consideration sets in section 6.3, using a different elicitation method for the other half of participants.

Job Choice Model and Estimation. We use the binary choices between potential jobs to estimate a random utility model.[6] In order to estimate the impact of our experiment on the familiarity premium of working in an unfamiliar neighborhood, we make the following assumptions:

1. Visiting a familiar neighborhood during the training does not affect the utility of working in any other target neighborhood.
2. Visiting an unfamiliar neighborhood during the training does not affect the utility of working in any familiar target neighborhood.

These two assumptions state that there are no cross-neighborhood spillovers in the utility of working in a neighborhood involving familiar neighborhoods. The first assumption says that when the neighborhood visited during training was already familiar at baseline, no other target neighborhoods are affected in terms of their utility of working there. Note that this is not a statement on the probability of deciding to work in one of these neighborhoods, because visiting a familiar neighborhood may make it *relatively* more or less attractive than other familiar neighborhoods. The second assumption states that training visits to unfamiliar neighborhoods do not affect the utility of working in non-visited familiar neighborhoods. Note that these assumptions do not preclude participants in the treatment group becoming more open to unfamiliar neighborhoods in general. Rather, we focus on assuming that they have "made up their minds" about familiar neighborhoods.

One example of how these assumptions might fail is when a trip to a training neighborhood exposes participants to other target neighborhoods. This is why our algorithm to minimize the possibility of these spatial spillovers.

Given these assumptions, the utility of working in non-visited familiar neighborhoods is unaffected by either treatment assignment. This provides a stable utility benchmark and allows us to identify both the effect of increased neighborhood salience due to training (comparing visited and non-visited familiar neighborhoods) and the additional familiarity premium (comparing this salience effect to the effect of visiting an unfamiliar neighborhood).

We further parameterize this relationship and assume that the utility of job offer $j \in$

[6] We also allow participants to turn down both jobs. Due to the underemployed sample and the competitive wages, this was extremely rare.

$\{1, 2\}$ in target neighborhood n, for individual i is given by

$$\begin{aligned} u_{ijn} = &\beta^F BaselineFamiliar_{in} + \\ &\underbrace{\beta^V Train_{in}}_{\substack{\text{(experimental)}\\ \text{training effect}}} + \underbrace{\beta^U Train_{in} \times (1 - BaselineFamiliar_{in})}_{\text{(experimental) familiarity effect}} + \\ &\underbrace{\gamma X_{ij}}_{\substack{\text{randomized}\\ \text{job attributes}}} + \mu_n + \beta^D d_{in} + \epsilon_{ijn} \end{aligned} \tag{1}$$

where $BaselineFamiliar_{in}$ is a dummy for i being familiar with neighborhood n at baseline, μ_n are neighborhood fixed effects, d_{in} is distance from the study venue to n, and the set of job covariates X_{ij} includes the wage, duration, whether a portion of the compensation is a risky potential bonus, and any amount paid up front to reduce liquidity constraints. The terms ϵ_{ijn} are idiosyncratic preference shocks.

The term $Train_{in}$ measures whether i was trained in neighborhood n. This includes control participants trained in baseline familiar neighborhoods, and treatment individuals trained in baseline unfamiliar neighborhoods. Hence, β^V gives the effect of training while β^U provides the additional effect of having been trained in an unfamiliar location.

We estimate equation (1) using a logit model, assuming that ϵ_{ijn} are distributed iid according to an extreme value of type 1 distribution with parameter 1. Table A.11 shows that results are similar when using a linear probability model.

Table 2 shows the results. The first column omits neighborhood fixed effects. The second column adds home by destination neighborhoods fixed effects, which makes distance drop out. The final column uses neighborhood fixed effects, which allows the distance coefficient to be identifies because of variation in distance to a neighborhood based on the home neighborhood. The large overall scale of coefficients indicates that the logit model explains a large share of participant decisions. Results are generally similar across specifications, so here we only discuss the results from the third column.

Baseline familiarity with a neighborhood has a large positive impact on the utility of working there (0.86). The magnitude is equivalent to 108 Kenyan shillings (Ksh) of additional compensation, which we obtain by dividing the baseline familiarity coefficient by the expected compensation coefficient (0.80). This is a large amount given that the median daily wage outside our the experiment is 500 Ksh.

Another way to benchmark the cross-sectional familiarity premium is with respect to distance. Working in a neighborhood that is familiar at baseline is valued equally to working in an unfamiliar neighborhood that is similar in all respects except 3.4 kilometers closer.

The causal effect of gaining familiarity due to the training also has a large positive effect on the utility of working there (0.91). This effect is in addition to the general effect of training in a neighborhood, which is also positive but smaller (0.30), and which applies both to familiar and unfamiliar neighborhoods.

The cross-sectional familiarity premium and the the experimental familiarity premiums have very similar magnitudes. This means that visiting the neighborhood once recently is sufficient to close the familiarity gap.

As another benchmark, we conduct a counterfactual exercise and compute how much commuter market access increases if familiarity levels go up for everyone. Specifically, we assume the existence of jobs with the same expected wage and duration in all neighborhoods, and then use our utility estimates to compute the expected utility gain from making all neighborhoods familiar. This increased utility is equivalent to shrinking all distances by 800 meters or 17%.

5.2 Impact on Beliefs

We have shown that visiting an unfamiliar neighborhood during training erases the initial wage premium that participants require to work in such a neighborhood. We now investigate to what extent the visit affects beliefs about the visited neighborhood (and route) characteristics. Beliefs may change either because the visit is informative, for example if participants learn about the neighborhood during their visit, or because of a potential bias affecting attitudes towards unfamiliar places.

We collect data on beliefs for the main six target neighborhoods (three familiar and three unfamiliar) during the second baseline survey for all participants. After each training day, when participants return to the study venue, we collected the same beliefs questions again referring to the neighborhood that was just visited.

We measure five dimensions of beliefs about labor market potential and safety. For each neighborhood, we ask participants about the likelihood of finding a daily or casual work opportunity in that neighborhood, both for the average person in the same home neighborhood, and for the respondent themselves. Asking both questions allows us to distinguish if a participant believes that a neighborhood that is unfamiliar does not offer good job opportunities in general or specifically for them. We next ask whether the pay is good conditional on finding a job. We then ask whether the trip to the neighborhood is safe, and whether the neighborhood itself is safe. We record responses on a likert-like scale and code responses on a scale from 1 to 5, where 5 is the best outcome. For each question, we also ask how confident the respondent is in their answer. See Appendix A.3 for precise question wording.

For respondent i, a neighborhood n where i was trained, and data collection time $t = 0$ (baseline 2) or $t = 1$ (after training), we estimate:

$$\text{Belief}_{int} = \gamma_n + Post_t + Pre_t \times Treated_i + Post_t \times Treated_i + \epsilon_{int}.$$

$Belief_{int}$ is the belief or certainty rating from 1 to 5, for one of the five belief outcomes described above.

Respondents in the control group are asked about familiar neighborhoods, while those in the treatment group are asked about unfamiliar neighborhoods. We interact the treatment dummy with pre- and post- dummies to compare how beliefs vary across familiarity and across time. The home neighborhood times destination neighborhood fixed effects γ_n captures all common aspects of a neighborhood, and ensures that we compare the same neighborhood that is familiar at baseline for one participant and unfamiliar at baseline for another participant.

Table 3 shows the results. At baseline, individuals are more pessimistic about unfamiliar neighborhoods, as indicated by the $Pre_t \times Treated_i$ coefficients. For example, respondents rate unfamiliar neighborhoods 0.18 SD lower on their likelihood of finding a causal job in the neighborhood, and 0.13 SD lower for the likelihood for an average resident of their neighborhood. We see no differences for wages. Travel safety to an unfamiliar neighborhood is rated 0.4 SD lower, while the neighborhood's safety is 0.15 SD lower.

After the training, these differences vanish. All $Post_t \times Treated_i$ coefficients are smaller in magnitude, close to zero, and never statistically significant. This means that on average, a single in-person visit eliminates the initial imbalance in ratings about these neighborhoods.

Table A.17 reports results on respondents' level of confidence. The results are in the expected direction. Respondents are less confident about unfamiliar neighborhoods at baseline, they are more confident in general after the training, and confidence differences between familiar and unfamiliar neighborhoods are no longer statistically significant after the training.

The training visits also have overall average effects across all neighborhoods. The $Post_t$ coefficient indicates that respondents update downward about job finding probabilities immediately after the visit, but update upward on safety. With only a pre-post comparison it is hard to interpret this coefficient (for example it could be due to mood differences between the end of a day of training and during the baseline, or it could be due to overly positive initial evaluations of familiar neighborhoods). Thus we focus on the treatment and control differences at both time points and include this coefficient for completeness.

We next assess whether this convergence in average beliefs is driven by increased agreement or increased randomness in responses. To do so, we estimate the mean pre-belief for each neighborhood among those who were familiar prior to the experiment and regress post beliefs on these average familiar priors.

Because we have relatively few observations per neighborhood, we estimate the priors for each neighborhood with a Bayesian partial-pooling random effects model. We assume rating by an individual i for neighborhood n, y_{in}, has an ordered logit likelihood where the latent rating is determined by a neighborhood level random effect μ_n. We estimate the model with the following priors for the random effects:

$$\begin{aligned} \mu_n &\sim N(0, \sigma_n^2) \\ \sigma_n^2 &\sim \text{InvGamma}(0.01, 0.01) \end{aligned}$$

And we use improper flat priors on the cut points for the ordinal logit. We estimate this model using the STATA Metropolis-Hastings algorithm with a burn-in period of 15,000 and a sample of 10,000 after thinning by retaining every 10th draw. We then compute the average prior belief for each neighborhood using the posterior means of the model. We then estimate the following regression

$$\text{Belief}_{in} = Treat_i + Prior_n + Treat_i \times Prior_n + \epsilon_{in}$$

with standard errors clustered at the neighborhood level.

Figures 5a and 5b presents the results for pre-visit and post-visit beliefs, respectively. The first result is that the beliefs we elicit have content: the Average Familiar Priors estimated on a subsample of participants strongly predict the beliefs of other participants. Second, in the post period, these results are always very similar for both those in the treatment group who had never been to the neighborhood before and control participants who had. This suggests that not only does one visit close the average beliefs gap between the two groups, it does so by leading the two groups to rate the individual neighborhoods in the same way. For the pre-beliefs, the Average Familiar Priors predict very well the beliefs of other participants familiar with the same neighborhood, while for those unfamiliar with the neighborhood, the results vary based on the outcome.

6 Barriers to Exploration

6.1 Risk Aversion

One potential explanation for this unwillingness to visit new neighborhoods is concern about risk. While we cannot completely rule out this explanation of our results, we believe several pieces of evidence weigh against it. First, there is relatively little potential for risk in the experiment. Individuals are only in the locations for an hour or two, and their payment and employment status are guaranteed. Thus, for risk aversion to explain the premium we estimate, it must be over a non-employment risk that is likely to occur in a short period.

Second, we examine whether some participants update very negatively about the neighborhoods that they visit. We do not find evidence for this. Figure A.8 displays the belief transition matrix for each respondent, their 3 visited neighborhoods, and for each of our five beliefs measures and their average. A key result is that there is little mass significantly below the diagonal. For the safety measures, almost all participants rate the trip to the neighborhood and the neighborhood itself as "Safe" or "Very Safe" and almost no participants who update downward significantly.

Finally, we examine whether visiting neighborhoods makes any individuals less willing to return. We estimate heterogeneous treatment effects by neighborhood and individual using the machine learning methods proposed by Chernozhukov et al. (2023). This method allows us to systematically search for any group who may have been less willing to return to a neighborhood after visiting it. Specifically we focus on structured choices between a non-target familiar and an unfamiliar location and regress an indicator for choosing to work in the unfamiliar location on treatment status. We then estimate the Grouped Average Treatment Effects for both a median and a quartile split. Figures 6 and A.9 present the results. We find that across all groups the treatment effect point estimates are positive and we can reject large negative effects.

Taken together we believe that this suggests risk aversion cannot explain our results.

Fixed Costs of Exploration. We find little evidence that first-time travel costs are large enough to explain the familiarity premium. In Table 4 we regress the travel time it takes participants to reach a neighborhood where they work, on baseline familiarity and experimentally induced visits. We do not find evidence that it takes longer to reach a neighborhood when the participant has never been there before. Training has a small effect on shorter times, but we do not find any additional effect for training in unfamiliar neighborhoods. Table 5 repeats the exercise for self-reported measures of navigation such as getting lost and asking for directions. One for one outcomes, which measures if participants reported the trip being more difficult than they expected, do we find an effect for experimentally induced familiarity. Overall, we find little support that first-time navigation costs could explain the familiarity premium documented earlier.

6.2 Incomplete Anticipation

We showed that in order to be induced to work in an unfamiliar neighborhood, study participants require a familiarity premium that is equivalent to being paid around 20% of a daily wage, or 3.4 km shorter distance. We also showed that a one-time exposure via training in an unfamiliar neighborhood eliminates these differences.

While these costs are large relative to daily outcomes, a rational individual that contemplates whether to visit an unfamiliar neighborhood would weight these one-time costs against a stream of potential future benefits, taking into account that travel costs decrease on average after the first visit. However, individuals may fail to anticipate that a location that is currently unfamiliar to them will feel familiar after visiting once. If this is true, individuals might under-explore relative to an efficient level.

We now analyze whether individuals anticipate the change in their preferences that visiting a neighborhood will induce. To study this question, in the structured elicitation arm we elicited choices during the three training days. On training day d, surveyors asked participants about their employment preferences for employment day d – that is, for employment three days in the future. Crucially, respondents received these questions *after* individuals learn where they will train on day d, but *before* they have actually visited the location. Study participants received the exact same seven questions that they they later received on the corresponding employment days (day $d + 3$). These questions were incentivized, respondents were informed that each question has an equal chance of being selected to be implemented three days later.

We analyze these choices by comparing how much a participant actually visiting a neighborhood changes their willingness to work there compared to learning that they are about to visit it.

We estimate a similar logit model as previously, but we now include the choices made on training days when individuals knew they were about to visit the neighborhood, but had not yet done so. The utility of working in neighborhood n is given by

$$
\begin{aligned}
u_{ijn} = & \beta^F BaselineFamiliar_{in} + \\
& \underbrace{\beta^V Train_{in}}_{\text{training effect}} + \underbrace{\beta^U Train_{in} \times (1 - BaselineFamiliar_{in})}_{\text{familiarity effect}} + \\
& \underbrace{\beta^{AV} AnticipTrain_{in}}_{\substack{\text{anticipation} \\ \text{training effect}}} + \underbrace{\beta^{AU} AnticipTrain_{in} \times (1 - BaselineFamiliar_{in})}_{\text{anticipation familiarity effect}} + \\
& \underbrace{\gamma X_{ij}}_{\substack{\text{randomized} \\ \text{job attributes}}} + \mu_n + \beta^D d_{in} + \epsilon_{ijn}
\end{aligned}
\tag{2}
$$

The coefficients β^{AV} and β^{AU} capture how participants value training that is about to happen in neighborhood n, in general and the additional effect for unfamiliar neighborhoods, respectively. The corresponding variables are switched on in the choices that participants make on training days, specifically for the neighborhood that will be visited that day.

We estimate (2) using a binary logit model. Table 6 shows the results. In column 1, we use the pooled data on training days and employment days. We find $\beta^U = 0.78$ and $\beta^{AU} = 0.45$. The ratio β^{AU}/β^U of the two coefficients provides a measure of to what degree individuals incompletely anticipate the utility change. The ratio is 0.58 and the p-value for equality is 0.01. In column 2, we only use data from the training period. The advantage of this specification is that all choices made during this period are for a time in the future, whereas in column 1 we also include choices made for the same day. In column 2, we leverage variation in realized familiarity induced by training on previous days. We find quite similar results and an even lower ratio of $\beta^{AU}/\beta^U = 0.51$. In column 3 we bring back the employment questions but restrict to the first question on each day, which by design always included the training neighborhood as one of the options. Because the questions we ask of training day d are completely identical to those asked three days later (on employment days $d+3$), in this specification we leverage within-individual and within-question variation, as the same neighborhood switches from anticipation to realized. Unfortunately, due to the significantly lower amount of data, the results are much noisier, and we do not detect a significantly positive familiarity premium β^U, nor can we reject equality between β^U and β^{AU}.

Our results are consistent with a failure to completely anticipate the utility changes induced by visiting. Another possibility is that this attenuation is driven by risk aversion. For example, if there are a few very bad neighborhoods and participants are reluctant to accept the risk of potentially committing to working in one, this would attenuate the treatment effect. While we cannot rule this explanation out, two factors weigh against it. First, participants' belief data does not suggest that individuals believe there are many very bad neighborhoods. For example, no neighborhood has a median rating of "Unsafe" or "Very Unsafe" in the baseline 2 data. Second, the commitment is only to stand in the neighborhood for 1-2 hours in the middle of the day and the compensation does not depend on the location attributes, which minimizes the scope for some common risks. Either of these explanations will reduce the amount that individuals explore. However, attenuate due to risk aversion would not be inefficient.

6.3 The Effect of Familiarity on Consideration Sets

So far, we discussed results from preferences elicited from study participants by confronting them with specific neighborhood options. This method plausibly measures respondent preferences for working in different types of neighborhoods, but it shuts down any role that familiarity may have on how likely people are to consider a neighborhood as an option, to begin with. In this section, we study this additional barrier to exploration, whether unfamiliar neighborhoods are less likely to be considered as potential places to work.

To measure how familiarity affects consideration sets, for half of our sample we elicit choices by asking respondents an open-ended question about locations where they would like to work. For each of the three employment days, when workers show up at the study venue in the morning, a surveyor tells them that air pollution monitoring jobs are available in some neighborhoods and not others that day, and that availability is random, with each neighborhood having an independent 1 in 4 chance of a job available. The wage and duration are the same for all neighborhoods, and we randomize these at the participant by day level. The surveyor then asks the participant to report an ordered list of neighborhoods where they would be willing to work in, and the surveyor then records the neighborhoods listed by the respondent, one by one. After the list is complete, the surveyor then goes to check which neighborhoods are available that day, and the respondent is assigned to work that day in the first available neighborhood in their list.

Under this elicitation method, respondents have an incentive to report neighborhoods in decreasing order of preference of performing the air pollution job in that location. However, because participants need to come up with the neighborhoods they want to rank, ease of consideration may also play a role. Our hypotheses are that under the open-ended elicitation, unfamiliar neighborhoods are ranked lower than we would expect based on preferences alone, and training in an unfamiliar neighborhood has a larger effect on ranking that neighborhood higher in the list than we would expected based on preferences alone.

To illustrate how the elicitation method affects the types of neighborhoods that participants report, we first use the model that we estimated on the structured choice data, which covers the other half of our sample of participants, to predict how often one of the three target main unfamiliar neighborhoods should be the first choice provided by an individual in the unstructured condition based on preferences alone. Figure 8 shows these results, focusing on the first day. In the control group, the preference model predicts that individuals will list these neighborhoods twice as often as they actually do (5% vs 2.4%). This is consistent with individuals being less likely to consider an unfamiliar neighborhood when asked open-ended questions.

In the treatment group, both the model and the data show that respondents are much more likely to report a target main unfamiliar neighborhood, and individuals actually list the unfamiliar neighborhoods more often than the model would predict (17% vs 15%). Overall, this suggests that in addition to having a preference to avoid unfamiliar neighborhoods, unfamiliarity also impedes consideration in less structured choice environments. We find similar results when we look at the likelihood of mentioning any unfamiliar neighborhood (not necessarily one of the three target unfamiliar ones) as the first choice (Figure A.13). In this case, the model overestimate in the control group is even higher.

We provide additional preliminary evidence on these decisions by estimating a multinomial logit model for the the k-th choice on the list using the open-ended elicitation data. Table 7 reports the results. In the first column, we copy the results from the third column from Table 2, based on the half the sample with closed elicitation. Column 2 reports the results of a multinomial logit on the top choice on day 1 of open elicitation. The key result is that familiarity (both baseline and experimental) matter more for open elicitation than when we confront participants with options. The training effect "Visited Any" is also more important, highlighting the salience effect of training in a given neighborhood in the past few days. Distance matters to a similar degree as in the first column. The other four columns repeat this exercise for the second, third, fourth and fifth ranked neighborhoods on the list. We find that 380 (96%) of participants rank at least five neighborhoods. The coefficients for baseline familiarity converges to a value close to that in preferences, while the overall training effect remains large. The additional effect of training in an unfamiliar neighborhood falls, although the sum "Visited Any" plus 'Visited Unfamiliar" remains above that from column 1 for the 2nd and 3rd choices. The coefficient on distance remains stable.

6.3.1 A Two-Self Model with Memory Costs

We now set up a simple model of the process by which respondents list neighborhoods in our open elicitation task. The model includes preferences over jobs similar to those outlined previously in equation (1), but also leaves room for recall costs that may differ based on neighborhood characteristics.

Each respondent has two selves. The "memory" self has access to all neighborhoods and their utilities, but for each neighborhood faces a cost to transmit this neighborhood to the "action" self. This actor optimally chooses which neighborhood to transmit. The action self simply tells the surveyor the neighborhood that they receive.

The memory self for agent i knows preferences

$$u_{ij} = \alpha^U \mathbf{X}_{ij} + \varepsilon_{ij}$$

for each neighborhood j where $\mathbf{X}_{ij}$ is a vector of neighborhood characteristics, as in equation (1). We assume that ε_{ij} has variance equal to 1.

For the k-th neighborhood to be ranked, the memory self incurs (negative) transmission cost

$$c_{ijk} = \alpha^C \mathbf{X}_{ij} + \nu_{ijk}$$

where ν_{ij} are iid shocks with variance $= \sigma_\nu \geq 0$.

The neighborhood ranked k-th on the list carries weight $\lambda_k \leq 1$. Neighborhoods further down the list are less likely to be relevant as some a previous neighborhood is likely to have a job available. Given that in reality we allow each neighborhood to have a job available with probability 0.25, the correct weights are $\lambda_k = \frac{1}{4} \cdot \left(\frac{3}{4}\right)^{k-1}$. However, we do not impose this at this time.

We further assume that at each step when the memory self is asked to transmit a neighborhood, they act myopically and send[7]

$$j_k^* \in \arg\max_j \lambda_k u_{ij} + c_{ijk} \quad \text{or "stop" (assumed net utility } v_{i0k})$$

Estimation and Results. We estimate this model using maximum likelihood based on the first 15 options ranked by the participant (or fewer if they stopped earlier). We fix the preference parameters over baseline familiarity, distance, training, and training in an unfamiliar location, to the values we estimated using binary choices (Table 2, column 3). We estimate the other parameters, including how participants value the job wage and duration (which vary at the participant level), the subjective job success probability, all the cost coefficients and the variance of idiosyncratic cost shocks.

Table 8 reports the results. We estimate that memory costs are lower for baseline familiar neighborhoods with a coefficient of 0.73. Having trained in a neighborhood has a large effect

[7]The model where the memory self is fully forward looking and optimizes over the entire list poses significant complications due to the combinatorial nature of that problem.

on memory costs, with a coefficient of 0.74, while the additional effect of having trained in a neighborhood that was unfamiliar at baseline is also positive and significant but smaller, with a coefficient of 0.37 $[0.19, 0.69]$. Distance also matters for memory, but the coefficient is lower than for preferences (-0.1 for memory costs compared to -0.25 for preferences).

Overall, the results from the open elicitation method show that above and beyond the fact that people dislike to work in unfamiliar neighborhoods, they are also less likely to consider these places as options in the first place.

7 Persistence

We finally consider whether this utility premium induced by visiting a neighborhood once persists. This matters because if maintaining familiarity requires re-visiting neighborhoods, individuals effectively have a constraint on the number of places they can be familiar with at once. On the other hand, if the premium persists, exploration can expand the set of familiar neighborhoods.

We first examine the cross-sectional evidence by estimating our logit equation based on equation (1), but splitting the cross-sectional familiarity term by how recently the individual reported having visited. Table A.12 shows that the familiarity premium is relatively stable up to neighborhoods that the respondent visited within the last 3 years.

We organized new work opportunities and invited study participants to them, around 2-4 months after the intervention. These are take-it-or-leave-it offers that let us learn about the extensive margin elasticity. (Our previous results mostly measure *where* a respondent prefers to work on a given day, and how much money they are willing to give up for these preferences.) For each of these opportunities, participants in our study received a phone call inviting them to a short (5-minute) survey that takes place in a given neighborhood two days later. The topic of the survey is commuting in Nairobi, which makes it natural for us to invite participants to different neighborhoods. The wages were randomized, and the neighborhood where the participant was invited was selected randomly from among the nine target neighborhoods for that person. Surveyors recorded show-up two days later. The entire procedure was then repeated, and each participant was invited to six different neighborhoods.

This setup allows us to estimate equation (1) using the show-up data and a binary logit model. Table 9 shows the results.

We split the results by whether the random wage was above 500 KSH or below, because for high wages, showup plateaus around 80% (Figure A.14) and none of the neighborhood or job characteristics, including compensation, affect show-up.[8]

For wages below 500 KSH, we see that baseline familiarity strongly affects show-up for the survey, with a magnitude similar to that we found for the air pollution jobs. The experimental effect of training in an unfamiliar neighborhood is positive and of a similar magnitude, although the estimates are at noisy, significant at 10% in one of the two specifications. Distance and wages also matter for show-up.

These results suggest that the intervention has lasting effects on participants' willingness to return to baseline unfamiliar neighborhoods, at least in the medium term.

Do participants return to visited neighborhoods? We next examine whether individuals return to the neighborhoods on their own as reported through SMS and phone surveys.[9]. Table 10 reports the results. Individuals do revisit the neighborhoods that they visited during training, including those that were unfamiliar at baseline.[10]

We find that respondents return to initially unfamiliar neighborhoods where they trained, namely, the sum of "Visited Any" and "Visited Unfamiliar" is positive and significant for all columns except the third. They return for a variety of reasons, including searching for work, as well as non-work reasons such as shopping, leisure, health, and errands. We do not find any effect on working in unfamiliar neighborhoods, although our estimates are noisy so we cannot reject a meaningful effect relative to the baseline mean. Furthermore, participants did search for work in these neighborhoods (column 2). Such outcomes may take time to realize.

Overall, these results outline a revealed preference argument that study participants found it worthwhile to return to the neighborhoods that they had never visited before our study.

[8] We analyzed average show-up in the entire sample (not separately by treatment) after initially launching these invitations with wages randomized between 500 and 1,000 KSH. After observing the high show-up rate, we reduced the distribution of wages, ultimately to between 100 and 400 KSH to avoid a ceiling effect on showup.

[9] We also attempted to measure this using a GPS tracking app, but were only able to obtain data for 15% of the sample.

[10] The results in Table 10, columns 1-4, are from directly asking respondents about whether they have recently visited the target neighborhoods, and if so for what purpose. The SMS survey reported in columns 5 and 6 is "unprompted," asking participants if they've worked or searched for work the day before, and if so, where (in their own words). We then code their response to our neighborhood list. We also collect travel information in less prompted ways and present the results in Table A.19. In the first three columns, we ask participants to tell us where they have been for different purposes, but without asking about any particular neighborhoods. These results are qualitatively similar to those in Table 10, but are less precise. The last two columns report trips as measured by our GPS tracking app. Unfortunately, we were only able to obtain this data for 15% of the sample, but we include the results for completeness.

8 Discussion

Growing cities in low- and middle-income countries offer the potential for increased market access, yet this requires that residents explore their surroundings. This is not always the case. In a sample of 800 casual workers in Nairobi, we've shown that the median person commutes 7.8 km but has never been to 1 in 5 of neighborhoods within that distance. We then showed that being experimentally induced to visit a neighborhood once is equivalent to bringing it 3.5 km closer or to increasing the wage by 112 Ksh (22% of the median daily wage), and this is equivalent to the full cross-sectional premium. We also show that after one visit, participants' initially pessimistic beliefs about the labor market opportunities and safety of unfamiliar neighborhoods converge after one visit. Finally, we show that there are at least two additional potential barriers to exploration: people incompletely anticipate the utility change and unfamiliar neighborhoods are less likely to be salient.

Policies to encourage exploration may be cost-effective ways of increasing effective market access. Additionally, policy makers may be able to decrease the cost of exploration by investing in making their cities more "readable". This is a concept from urban studies that notes that the logic of some cities is easy to infer (e.g. NYC) while others is difficult (e.g. Boston). Rapid growth in low and middle income contexts may lead to cities that are particularly difficult to "read". Investing in infrastructure such as clear signage and addresses may help alleviate this problem.

References

Abebe, Girum et al. (2021). "Anonymity or distance? Job search and labour market exclusion in a growing African city". In: *The Review of Economic Studies* 88.3, pp. 1279–1310.

Abel, Martin et al. (2019). "Bridging the intention-behavior gap? The effect of plan-making prompts on job search and employment". In: *American Economic Journal: Applied Economics* 11.2, pp. 284–301.

Agte, Patrick et al. (2024). *Search and Biased Beliefs in Education Markets*. Tech. rep. National Bureau of Economic Research.

Avner, Paolo and Somik V. Lall (2016). "Matchmaking in Nairobi: The role of land use". In: *World Bank Policy Research Working Paper* 7904.

Balboni, Clare et al. (2021). "Could Gentrification Stop the Poor from Benefiting from Urban Improvements?" In: *AEA Papers and Proceedings*. Vol. 111. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, pp. 532–537.

Banerjee, Abhijit and Sandra Sequeira (2023). "Learning by searching: Spatial mismatches and imperfect information in Southern labor markets". In: *Journal of Development Economics* 164, p. 103111.

Banerjee, Abhijit V (1992). "A simple model of herd behavior". In: *The quarterly journal of economics* 107.3, pp. 797–817.

Baseler, Travis (2023). "Hidden income and the perceived returns to migration". In: *American Economic Journal: Applied Economics* 15.4, pp. 321–352.

Belot, Michele, Philipp Kircher, and Paul Muller (2019). "Providing advice to jobseekers at low cost: An experimental study on online advice". In: *The review of economic studies* 86.4, pp. 1411–1447.

Berkouwer, Susanna B. and Joshua T. Dean (2024). "Cooking, health, and daily exposure to pollution spikes". In: *Mimeo*. URL: https://sberkouwer.github.io/BerkouwerDean_Health.pdf.

Borker, Girija (2021). "Safety First: Perceived Risk of Street Harassment and Educational Choices of Women". In: *Policy Research Working Paper;No. 9731*.

— (2024). "Understanding the constraints to women's use of urban public transport in developing countries". In: *World Development* 180, p. 106589.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak (2014). "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh". In: *Econometrica* 82.5, pp. 1671–1748.

Bryan, Gharad, Edward Glaeser, and Nick Tsivanidis (2020). "Cities in the developing world". In: *Annual Review of Economics* 12, pp. 273–297.

Carranza, Eliana et al. (2022). "Job search and hiring with limited information about workseekers' skills". In: *American Economic Review* 112.11, pp. 3547–3583.

Chauvin, Juan Pablo et al. (2017). "What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States". In: *Journal of Urban Economics* 98, pp. 17–49.

Chernozhukov, Victor et al. (Oct. 23, 2023). *Fisher-Schultz Lecture: Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments, with an Application to Immunization in India*. arXiv: 1712.04802[econ,math,stat]. URL: http://arxiv.org/abs/1712.04802 (visited on 06/21/2024).

Davis, Lucas W (2021). "Estimating the price elasticity of demand for subways: Evidence from Mexico". In: *Regional Science and Urban Economics* 87, p. 103651.

Duranton, Gilles (2015). "Growing through cities in developing countries". In: *The World Bank Research Observer* 30.1, pp. 39–73.

Eyster, Erik and Matthew Rabin (2010). “Naive herding in rich-information settings”. In: *American economic journal: microeconomics* 2.4, pp. 221–243.

Franklin, Simon (2018). “Location, search costs and youth unemployment: experimental evidence from transport subsidies”. In: *The Economic Journal* 128.614, pp. 2353–2379.

Gachanja, James et al. (Feb. 2023). *Urban Economic Growth in Africa: A Case Study of Nairobi City County, Kenya*. Brookings Africa Growth Initiative.

Gonzalez-Navarro, Marco and Matthew A Turner (2018). “Subways and urban growth: Evidence from earth”. In: *Journal of Urban Economics* 108, pp. 85–106.

Grosset-Touba, Florian (2024). “Complementarities in Labor Supply: How Joint Commuting Shapes Work Decisions”. In: *Working Paper*.

Habyarimana, James and William Jack (2015). “Results of a large-scale randomized behavior change intervention on road safety in Kenya”. In: *Proceedings of the national Academy of Sciences* 112.34, E4661–E4670.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein (2014). “Learning through noticing: Theory and evidence from a field experiment”. In: *The Quarterly Journal of Economics* 129.3, pp. 1311–1353.

Jalota, Suhani and Lisa Ho (2024). “What Works for Her? How Work-from-Home Jobs Affect Female Labor Force Participation in Urban India”. In: *Working Paper*.

KNBS (2019). *2019 Kenya Population and Housing Census*. Kenya National Bureau of Statistics. URL: https://www.knbs.or.ke/?wpdmpro=2019-kenya-population-and-housing-census-volume-i-population-by-county-and-sub-county#.

Kreindler, Gabriel et al. (2023). *Optimal Public Transportation Networks: Evidence from the World's Largest Bus Rapid Transit System in Jakarta*. Tech. rep. National Bureau of Economic Research.

Lall, Somik V., J Vernon Henderson, and Anthony J Venables (2017). *Africa's cities: Opening doors to the world*. World Bank Publications.

Larcom, Shaun, Ferdinand Rauch, and Tim Willems (2017). “The benefits of forced experimentation: Striking evidence from the London underground network”. In: *The Quarterly Journal of Economics* 132.4, pp. 2019–2055.

Malmendier, Ulrike (2021). “Experience effects in finance: Foundations, applications, and future directions”. In: *Review of Finance* 25.5, pp. 1339–1363.

McKenzie, David (2024). “Fears and Tears: Should More People Be Moving within and from Developing Countries, and What Stops this Movement?” In: *The World Bank Research Observer* 39.1, pp. 75–96.

Montoya, R Matthew et al. (2017). “A re-examination of the mere exposure effect: The influence of repeated exposure on recognition, familiarity, and liking.” In: *Psychological bulletin* 143.5, p. 459.

Okunogbe, Oyebola (2024). “Does exposure to other ethnic regions promote national integration? Evidence from Nigeria”. In: *American Economic Journal: Applied Economics* 16.1, pp. 157–192.

Porcher, Charly (2022). “Migration with costly information”. In: *Working Paper*.

Porcher, Charly, Eduardo Morales, and Thomas Fujiwara (2024). *Measuring Information Frictions in Migration Decisions: A Revealed-Preference Approach*. Tech. rep. National Bureau of Economic Research.

Scharfstein, David S and Jeremy C Stein (1990). “Herd behavior and investment”. In: *The American economic review*, pp. 465–479.

Tang, Yulu (2024). “To Follow the Crowd? Benefits and Costs of Migrant Networks”. In: *Working Paper*.

Tsivanidis, Nick (2023). “Evaluating the impact of urban transit infrastructure: Evidence from Bogota’s Transmilenio”. In: *University of California, Berkeley*.

Vitali, Anna (2024). “Consumer search and firm location: Theory and evidence from the garment sector in Uganda”. In: *Working Paper*.

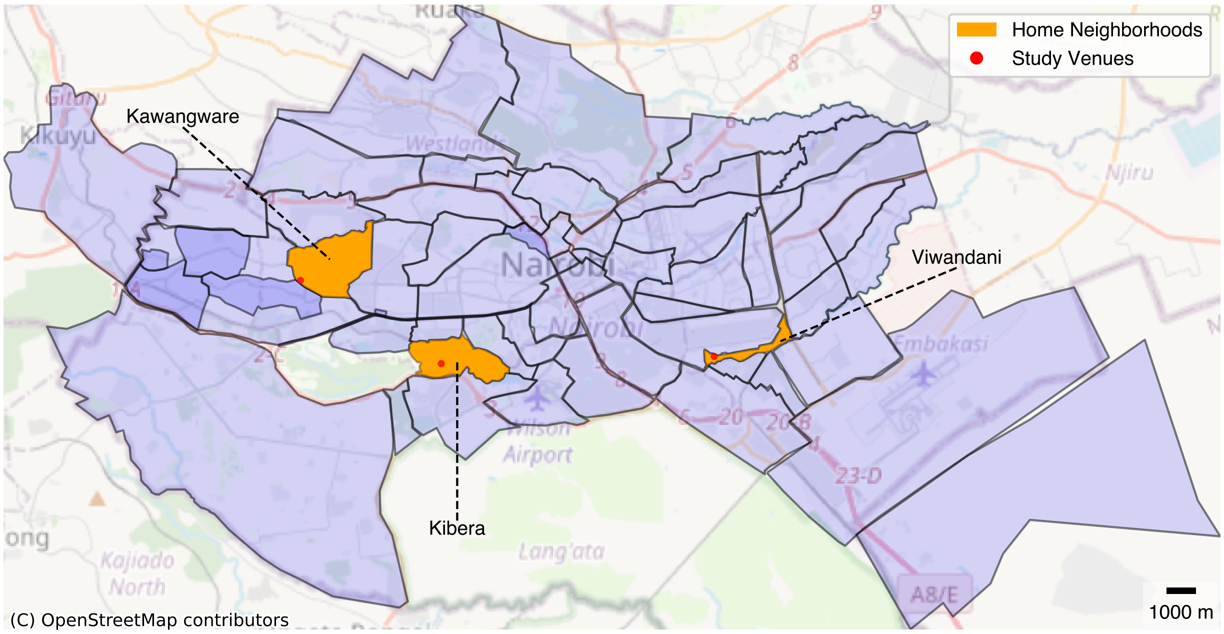
Wiseman, Eleanor (2023). “Border trade and information frictions: Evidence from informal traders in kenya”. In: *Working Paper*.

Zajonc, Robert B (1968). “Attitudinal effects of mere exposure.” In: *Journal of personality and social psychology* 9.2p2, p. 1.

Zárate, Román D (2022). “Spatial misallocation, informality, and transit improvements: Evidence from mexico city”. In: *Working Paper*.

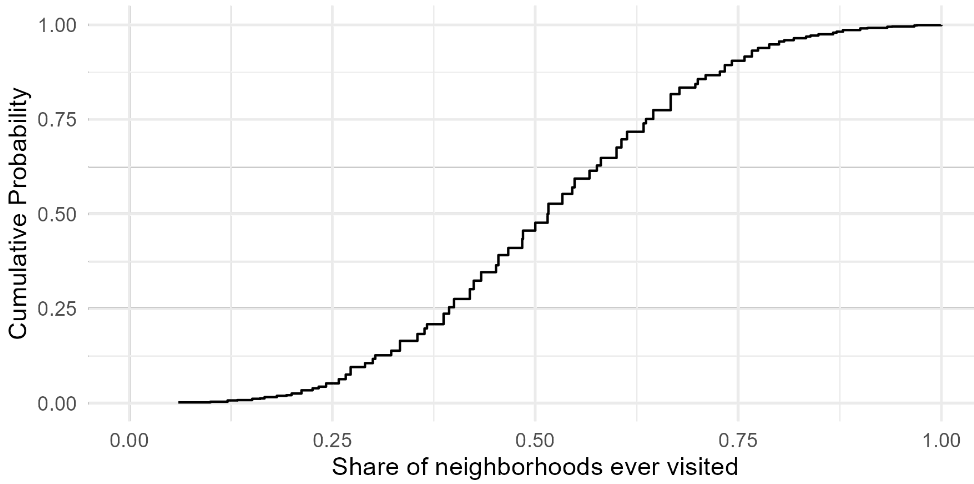
Figures and Tables

Figure 1: Division of Nairobi Into Neighborhoods



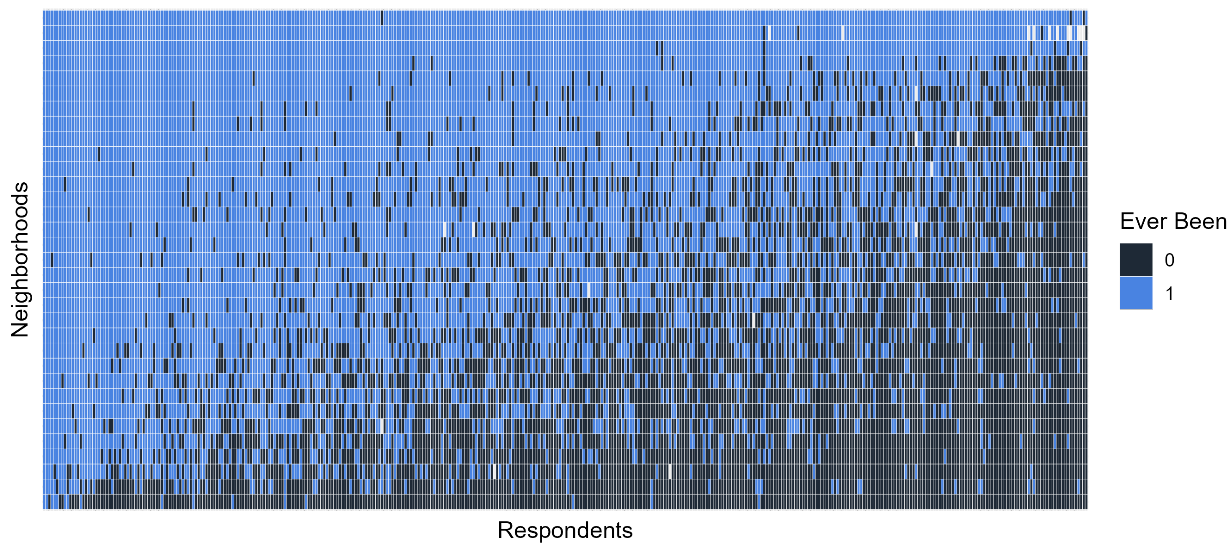
Notes: This figure shows the partition of the main neighborhoods in Nairobi. The orange polygons represent the home neighborhoods—Kibera, Kawangware, and Viwandani—where participants were recruited. The location of the study venues are highlighted in red. The remaining 58 neighborhoods are represented by light blue polygons.

Figure 2: CDF of Familiarity Within 75 minutes from Home



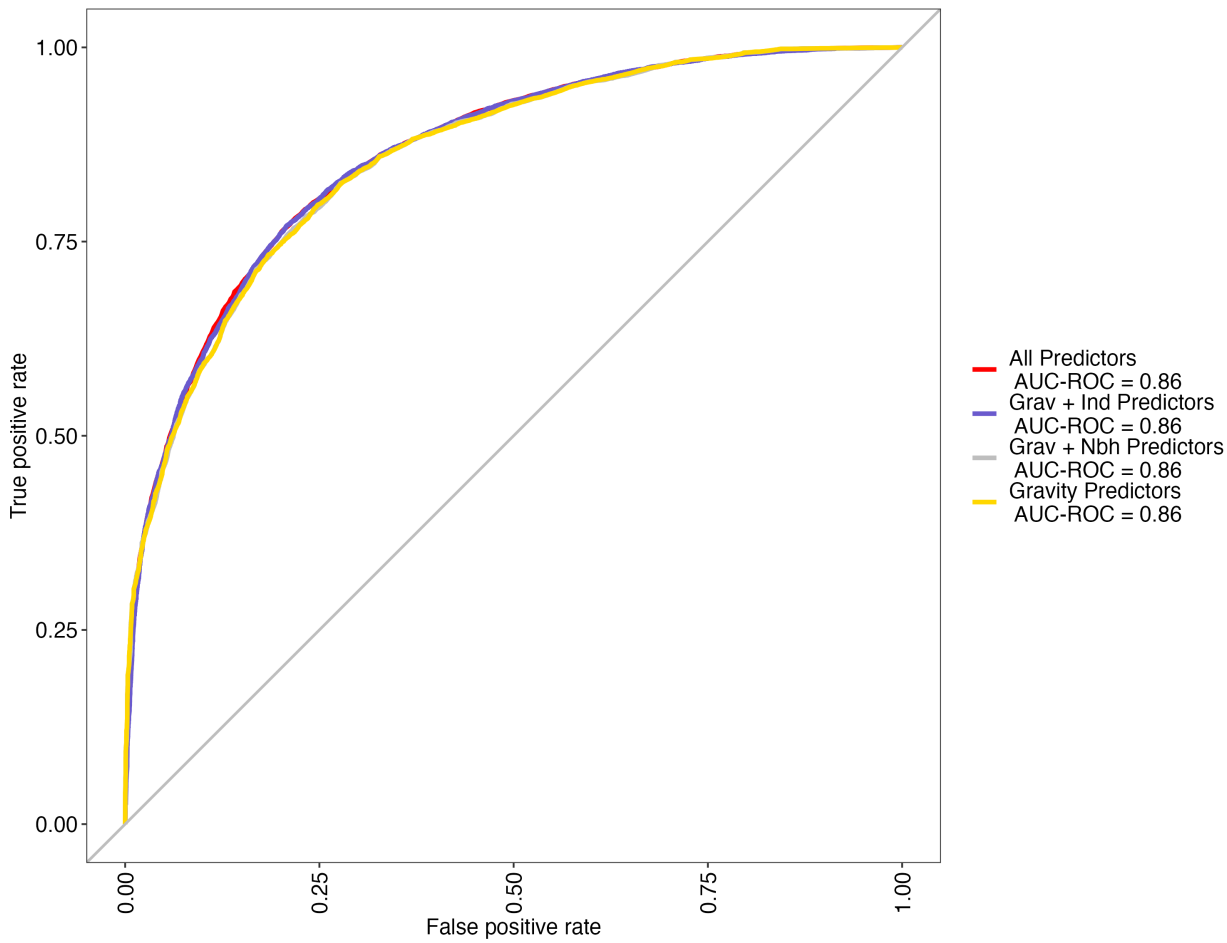
Notes: This figure plots the cumulative distribution function for our main measure of familiarity. The sample of respondents includes all 1168 participants who completed the first baseline survey. The sample of neighborhoods is restricted to within 75 minutes (walking or by transit) of the respondent's home neighborhood. This amounts to approximately 30 neighborhoods for each of our three home neighborhoods. The X axis lists the share of neighborhoods N that a given respondent answers yes to the question "Have you ever been to the neighborhood of N".

Figure 3: Significant Idiosyncratic Variation in Familiarity (Kibera)



Notes: This figure plots the entire familiarity matrix for all study participants from Kibera who completed the first baseline, for all 30 neighborhoods that they were asked about. A cell is blue if the (column) respondent has ever been to the (row) neighborhood, and black otherwise. Neighborhood and respondents are ordered by average familiarity. (White indicates missing data.)

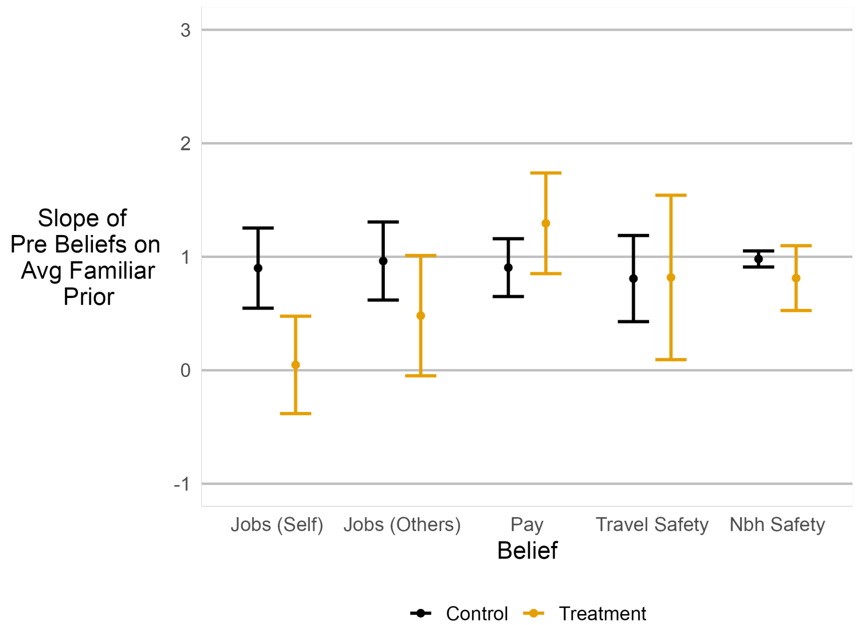
Figure 4: No Evidence of Match-Specific Familiarity Patterns



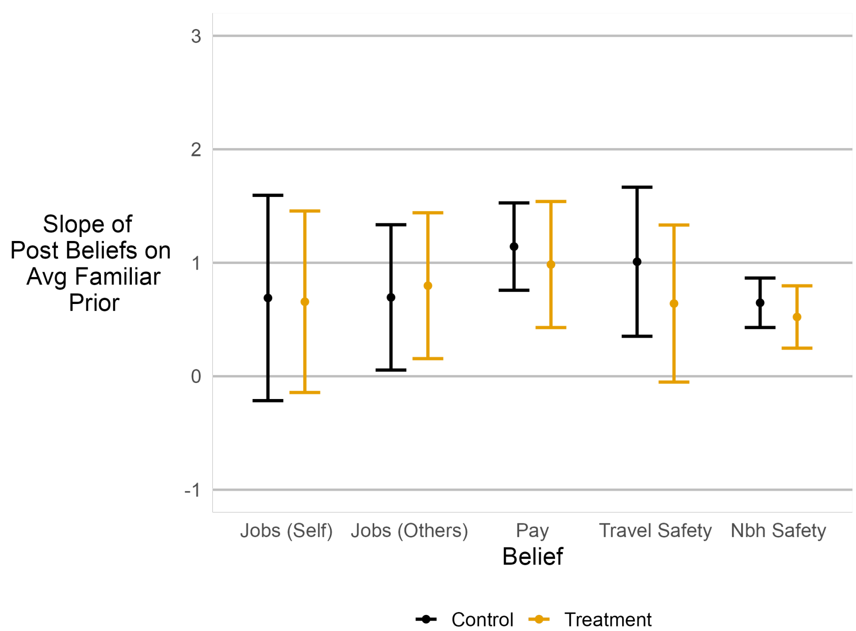
Notes: The figure shows the area under the ROC curve for four different random forest models predicting whether a given individual is familiar with a given neighborhood. While the standard gravity model predictors of distance and neighborhood fixed effects lead to a prediction substantially better than chance, adding additional predictors such as individual and neighborhood characteristics, does not improve the prediction. This suggests match-specific components are not important in predicting familiarity patterns. See Figure A.16 for further results using logit-lasso models.

Figure 5: Relationship Between Average Familiar Priors and Posteriors by Treatment

(a) Pre-Visit Beliefs

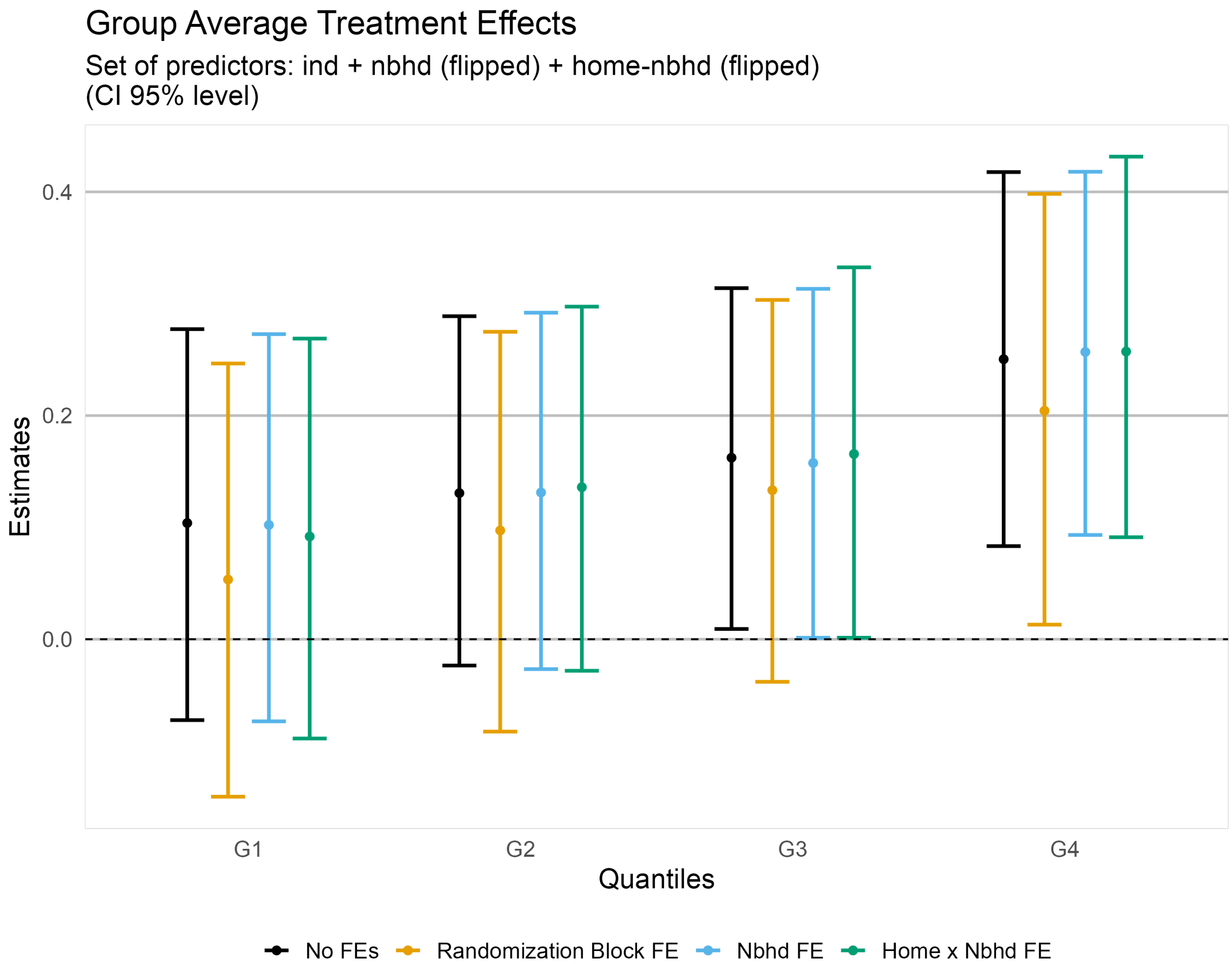


(b) Post-Visit Beliefs



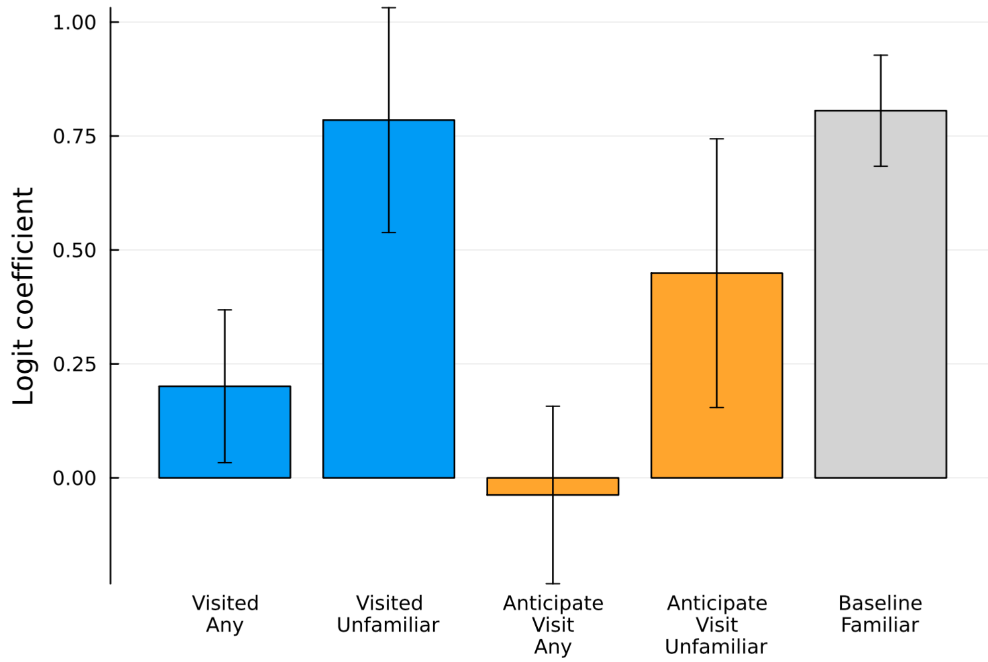
Notes: The figure presents the results comparing beliefs between different respondents, before and after the training intervention. We first construct for each neighborhood and each outcome a predicted rating based on priors from respondents familiar with the respective neighborhood, using a Bayesian model and 10 cross folds. In panel (5a) each point represents the estimated slope of the prior beliefs (pre-visit) on these ratings. In (5b) we repeat the exercise with posterior beliefs (post-visit). Standard errors are clustered at the neighborhood level in all regressions.

Figure 6: Quartiles of Heterogeneous Treatment Effects



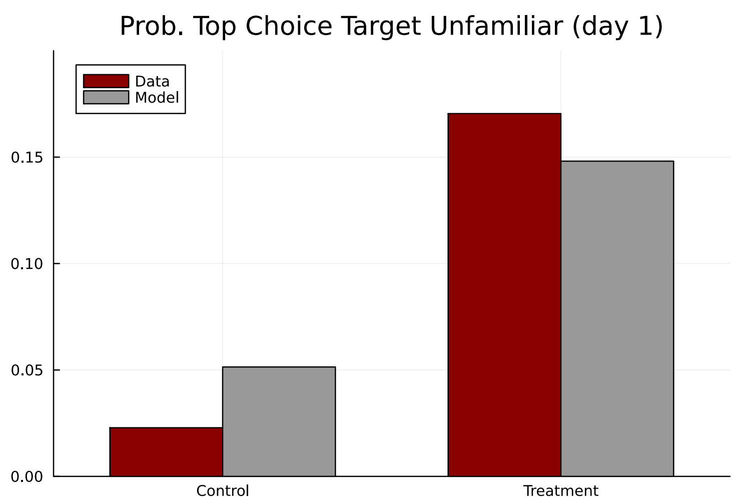
Notes: The figure shows four quartiles of the Group Average Treatment Effects estimated with the Chernozhukov et al. (2023) method. To employ the method we restrict to structured choices between an non-target familiar neighborhood and an unfamiliar neighborhood. We then use as the outcome an indicator for choosing to work in an unfamiliar neighborhood.

Figure 7: The Impact of Realized and Anticipated Familiarity on Job Choices



Notes: This figure plots the logit coefficients from Table 6, column 2.

Figure 8: Comparison of First Choices Predicted by Structured and Realized Data



Notes: This figure reports the share of respondents in the unstructured elicitation arm who mention as their top neighborhood choice one of the three target main unfamiliar neighborhoods on the first day, in the control group and in the treatment groups. Red bars indicate the empirical frequencies, and the gray bars indicate predicted shares based on the logit utility model estimated in Table 2 using the structured elicitation arm.

Table 1: Participants Have Significant Familiarity Gaps

	Share of Neighborhoods Visited					
Nbhd sample:	< 75min			< 7.8km		
	p25	p50	mean	p25	p50	mean
Measure of familiarity:						
Heard of	0.89	0.94	0.92	0.90	0.97	0.93
Ever been OR passed by	0.50	0.63	0.62	0.62	0.79	0.74
Ever been	0.39	0.52	0.52	0.50	0.69	0.66
Ever been + knows get there	0.33	0.43	0.45	0.40	0.56	0.56
Ever been + gave landmark	0.24	0.34	0.35	0.33	0.48	0.46

Notes: This table reports statistics for the share of neighborhoods that a participant is familiar with.

Table 2: Revealed Preference Estimates of the Familiarity Premium

	(1)	(2)	(3)
Baseline Familiar	0.85***	0.76***	0.86***
	(0.120)	(0.121)	(0.118)
Visited Any	0.29**	0.33**	0.30**
	(0.109)	(0.103)	(0.105)
Visited Unfamiliar	0.83***	0.93***	0.91***
	(0.211)	(0.205)	(0.199)
Distance (km)	-0.24***		-0.25***
	(0.016)		(0.019)
Job duration (hrs)	-0.58***	-0.62***	-0.59***
	(0.062)	(0.061)	(0.059)
$\mathbb{E}$ Compensation (100 KSH)	0.76***	0.83***	0.80***
	(0.027)	(0.029)	(0.027)
Cash Upfront (100 KSH)	0.06	0.06	0.07
	(0.045)	(0.046)	(0.042)
Bonus (100 KSH)	-0.21***	-0.22***	-0.22***
	(0.010)	(0.011)	(0.010)
N	6,756	6,756	6,756
Home $\times$ neighborhood FE		Yes	
Neighborhood FE			Yes

Notes: This table reports the results of the logit estimation of the "structured" choice elicitation where individuals chose between two potential job offers which varied based on location, duration, total compensation, the amount offered in advance to ease liquidity constraints and the amount that depended on a risky bonus. "Visited Any" is an indicator equal to one if a participant trained in the neighborhood, while "Visited Unfamiliar" is constructed similarly but only switched on for participants in the treatment group, for whom the neighborhood is always unfamiliar at baseline. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table 3: Average Beliefs Converge After One Visit

	Dependent variable:				
	Find Job	Find Job (Others)	Pay is Good	Travel Safety	Overall Safety
	(1)	(2)	(3)	(4)	(5)
Post	-0.140^{***}	-0.193^{***}	-0.048	0.148^{***}	0.251^{***}
	(0.040)	(0.039)	(0.038)	(0.034)	(0.039)
Pre $\times$ Treated	-0.182^{***}	-0.125^{**}	-0.018	-0.368^{***}	-0.146^{**}
	(0.061)	(0.059)	(0.063)	(0.060)	(0.060)
Post $\times$ Treated	0.004	0.053	0.018	0.026	0.034
	(0.057)	(0.059)	(0.053)	(0.045)	(0.046)
Mean	3.4	3.4	3.6	4.0	3.8
SD	1.0	1.0	1.1	0.9	1.0
Home $\times$ Neighborhood FE	Yes	Yes	Yes	Yes	Yes
Observations	4,401	4,393	4,293	4,404	4,339

Notes: This table regresses beliefs about attributes of the visited neighborhoods elicited at baseline two and after visiting on an indicator for the time point and an interaction with treatment status. Those in the treated group are rating unfamiliar neighborhoods while those in the control are rating familiar. All outcomes are rated on a likert scale from 1 to 5 with 5 being the most positive outcome. The table shows that while beliefs are initially more negative for unfamiliar neighborhoods, the gap closes after visiting. Table A.17 repeats the analysis using belief confidence as the outcome. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table 4: Does it Take Longer to Reach Unfamiliar Neighborhoods?

	Surveys	Task App		
Dep. Var (hours):	Round Trip	To Job	From Job	Job Duration
	(1)	(2)	(3)	(4)
Baseline familiar	0.07	0.01	0.01	0.00
	(0.08)	(0.04)	(0.05)	(0.05)
Visited Any	-0.09	-0.08**	-0.06*	0.08*
	(0.06)	(0.03)	(0.03)	(0.05)
Visited Unfamiliar	0.06	0.05	0.04	-0.08
	(0.10)	(0.05)	(0.06)	(0.07)
Control Mean	3.92	1.24	1.30	1.47
N	1,159	1,133	1,098	1,089
Home $\times$ neighborhood FE	Yes	Yes	Yes	Yes

Notes: This table examines travel times (in hours) between home and job destination neighborhoods on employment days. The sample includes only participants in the structured group who completed the task on each employment day. Column 1 dependent variable is the round-trip duration, calculated using the time difference between the pre-employment and post-employment surveys. Columns 2 to 4 considers the timestamps reported by participants on the Task-App: column 2 is the travel time from home to the job destination neighborhood; column 3 is the travel duration from the job destination to the home location neighborhood; column 4 is the total time spent by the participant in the job destination neighborhood.

Table 5: Is it More Difficult to Navigate to Unfamiliar Neighborhoods?

Dep. Var (Post-Visit):	Get Lost There (1)	Get Lost Back (2)	Ask Directions There (3)	Ask Directions Back (4)	Diff. Level Std. (5)	More Difficult Than Expected (6)
Baseline familiar	0.00	0.00	-0.09**	-0.06***	0.02	-0.07**
	(0.02)	(0.00)	(0.04)	(0.02)	(0.09)	(0.03)
Visited Any	-0.04***	0.00	-0.17***	-0.01	-0.15*	-0.07***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.08)	(0.02)
Visited Unfamiliar	0.00	0.01	-0.07	-0.02	-0.06	-0.10***
	(0.03)	(0.01)	(0.05)	(0.02)	(0.12)	(0.03)
Control Mean	0.06	0	0.31	0.05	0.15	0.18
N	1,159	1,159	1,159	1,159	1,158	1,158
Home × neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table replicates Table 4 using using participants' navigation experiences reported in the post-employment surveys. The dependent variables in columns 1 and 2 are indicators confirming if participants got lost during their trips to the job destination and home location neighborhood, respectively. Columns 3 and 4 are indicators confirming if participants asked for directions on their trips to the job destination and home neighborhood, respectively. Column 5 shows a difficulty index, standardized to a mean of zero and a standard deviation of one, based on a Likert scale where the lowest level (1) indicates that the participant found the trip to the job destination easier than initially expected, and the highest level (5) indicates it was a lot more difficult than initially expected. Column 6 takes the value of one if the participant considered the trip to the job destination difficult or a lot more difficult than initially expected, and zero otherwise.

Table 6: The Impact of Realized and Anticipated Familiarity on Job Choices

	(1)	(2)	(3)
Baseline Familiar	0.81***	0.79***	0.72***
	(0.069)	(0.076)	(0.218)
β^V Visited Any	0.20*	0.01	0.68***
	(0.086)	(0.115)	(0.198)
β^U Visited Unfamiliar	0.78***	0.80***	0.26
	(0.138)	(0.184)	(0.363)
β^{AV} Anticipate Visit Any	-0.04	-0.05	0.03
	(0.097)	(0.086)	(0.190)
β^{AU} Anticipate Visit Unfamiliar	0.45**	0.41**	0.29
	(0.154)	(0.147)	(0.342)
N	13,658	6,902	2,137
P-value $\beta^U = \beta^{AU}$	0.01	0.02	0.43
Sample: Training	Yes	Yes	Yes
Sample: Employment	Yes		Yes
Sample: only Q1			Yes
Neighborhood FEs	Yes	Yes	Yes
Job Attribute Controls	Yes	Yes	Yes

Notes: This table reports binary logit estimation results of equation (2). Column 1 pools all the choice data for the three training days, and for the three employment days. column 2 only uses the training days data, leveraging variation in realized familiarity induced by training on previous training days (e.g. if respondent i visits neighborhood n on day 1 of training, then $Training_{in} = 1$ for choices made on the later days. Column 3 uses training and employment data but restricts to the first question each day. Coefficients on expected compensation, liquidity, bonus and distance are included but not reported to save space. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table 7: Multinomial Logit of k-th Choice in the Open-Elicitation

	(1)	(2)	(3)	(4)	(5)	(6)
Baseline Familiar	0.86***	2.20***	1.31***	1.33***	1.10***	1.26***
	(0.118)	(0.309)	(0.193)	(0.193)	(0.145)	(0.164)
Visited Any	0.30**	1.00***	0.98***	1.19***	1.11***	0.99***
	(0.105)	(0.135)	(0.177)	(0.172)	(0.185)	(0.199)
Visited Unfamiliar	0.91***	1.53***	0.73*	0.75*	0.22	0.23
	(0.199)	(0.373)	(0.320)	(0.315)	(0.304)	(0.371)
Distance (km)	-0.25***	-0.28***	-0.26***	-0.23***	-0.18***	-0.19***
	(0.019)	(0.021)	(0.020)	(0.023)	(0.019)	(0.021)
Obs.	6,756	10,537	10,146	9,705	9,277	8,741
Binary Choices	Yes					
Open Elicitation		Yes	Yes	Yes	Yes	Yes
Rank in the List		1	2	3	4	5
Respondents	409	391	391	389	387	380

Notes: This table reports multinomial results of the k-th choice on the list in the “open” job elicitation, on day, for $k = 1, \ldots, 5$. For comparison, the first column repeats the results from Table 2, column 3, based on binary choices between options that we ask about.

Table 8: Estimates of Memory Costs from "Open" Elicitation

	Utility	**Cost**
Baseline Familiar	0.86	0.73 [0.57, 1.11]
Visited Any	0.30	0.74 [0.57, 1.14]
Visited Unfamiliar	0.91	0.37 [0.19, 0.69]
Distance	-0.25	-0.10 [-0.16, -0.08]
Wage (100 KSH)	0.07 [0.05, 0.11]	
Duration (hours)	-1.38 [-6.10, 4.52]	
Constant	-27.57 [-43.50, -15.77]	-1.46 [-2.44, -1.18]
Shock parameter σ	1.00	0.75 [0.60, 1.10]
Subjective Job Probability λ	0.18 [0.14, 0.20]	
Participants	391	
Observations	10,831	

Notes: This table reports maximum likelihood estimates of the two-self memory model from section 6.3. We fix the first four preference parameters to those estimated in Table 2 column 3, and estimate the remaining parameters, including the memory cost parameters, using the ranked neighborhood data from the "open" preference elicitation. Individual-level bootstrapped 95% confidence intervals in parentheses.

Table 9: Show-up 2-4 Months After Intervention

		(1)	(2)
Baseline Familiar	$\times$ (Wage $\leq$ 500)	0.90***	0.85***
		(0.26)	(0.27)
Visited Any	$\times$ (Wage $\leq$ 500)	0.08	0.14
		(0.24)	(0.26)
Visited Unfamiliar	$\times$ (Wage $\leq$ 500)	0.84*	0.78
		(0.48)	(0.49)
Distance (km)	$\times$ (Wage $\leq$ 500)	-0.33***	
		(0.05)	
Wage	$\times$ (Wage $\leq$ 500)	1.2***	1.2***
		(0.09)	(0.09)
(Wage $>$ 500)		2.5	4.6***
		(1.6)	(1.4)
Baseline Familiar	$\times$ (Wage $>$ 500)	0.006	-0.39
		(0.56)	(0.61)
Visited Any	$\times$ (Wage $>$ 500)	0.68	0.64
		(0.66)	(0.71)
Visited Unfamiliar	$\times$ (Wage $>$ 500)	-0.23	-0.05
		(1.0)	(1.1)
Distance (km)	$\times$ (Wage $>$ 500)	-0.09	
		(0.09)	
Wage	$\times$ (Wage $>$ 500)	0.30*	0.35*
		(0.17)	(0.18)
Neighborhood FE		Nbhd	Home $\times$ Nbhd
Observations		2,648	2,634

Notes: this table reports estimates using equation (1) for the show-up outcome for the travel survey invitations 2-4 months after the intervention. The outcome is a dummy for whether the participant showed up. Wages above 500 KSH led to high show-up that is insensitive to any job and neighborhood characteristic, including further wage increases (see Figure A.14).

Table 10: People Re-visit the Neighborhoods From the Study

	Endline Survey (prompted)				SMS (unprompted)	
	Any Trip	Search Work	Work	Other	Visited	Num Visits
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline Familiar	0.146***	0.062***	0.028***	0.073***	0.010**	0.062***
	(0.017)	(0.011)	(0.010)	(0.013)	(0.004)	(0.019)
Visited Any	0.079***	0.037***	0.025**	0.020	0.017***	0.092***
	(0.020)	(0.014)	(0.012)	(0.015)	(0.005)	(0.028)
Visited Unfamiliar	−0.005	0.006	−0.021	0.023	−0.004	−0.028
	(0.035)	(0.023)	(0.021)	(0.025)	(0.009)	(0.044)
Mean	0.267	0.113	0.069	0.112	0.027	0.148
SD	0.442	0.317	0.253	0.316	0.098	0.48
Visited Any + Visited Unfam	0.074	0.043	0.004	0.043	0.013	0.064
[p-value]	[0]	[0.001]	[0.683]	[0.003]	[0.005]	[0.002]
Observations	6,927	6,927	6,927	6,927	5,163	5,163

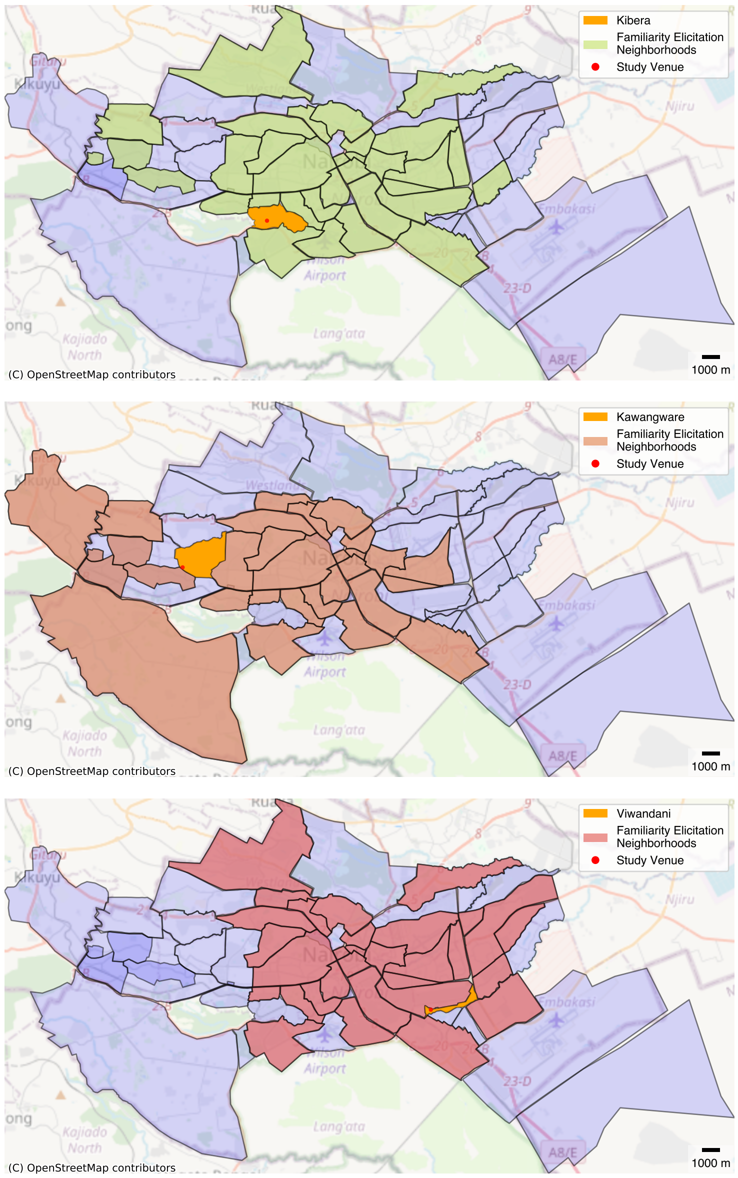
Notes: This table regresses whether individuals report returning to the target neighborhoods on an indicator for being familiar at baseline, whether any training occurred in the neighborhood, and whether the interaction of training and whether the neighborhood was unfamiliar at baseline. Columns 1-3 are trips measured using the over-the-phone endline survey while columns 4 and 5 are visits measured by the high-frequency SMS. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

A Appendix - For Online Publication

Contents

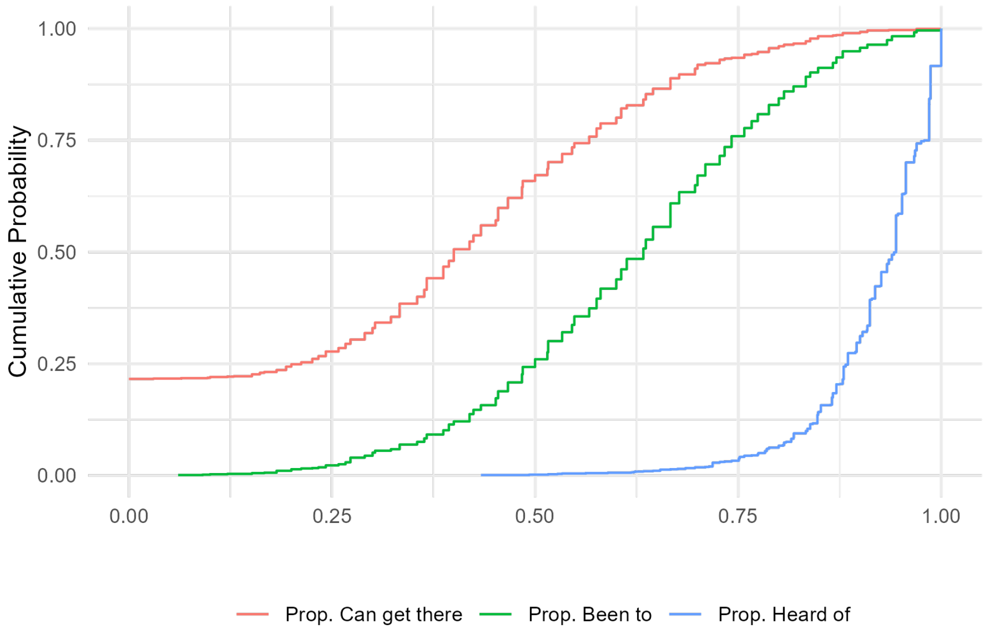
A.1 Appendix Figures

Figure A.1: Familiarity Elicitation Neighborhoods



Notes: This figure replicates Figure 1 for each home neighborhood and highlights the set of main neighborhoods for which we elicited familiarity. We elicited familiarity in 33 neighborhoods of Kibera, 30 in Kawangware, and 31 in Viwandani.

Figure A.2: Levels of Familiarity Within 75 minute Commuting Distance

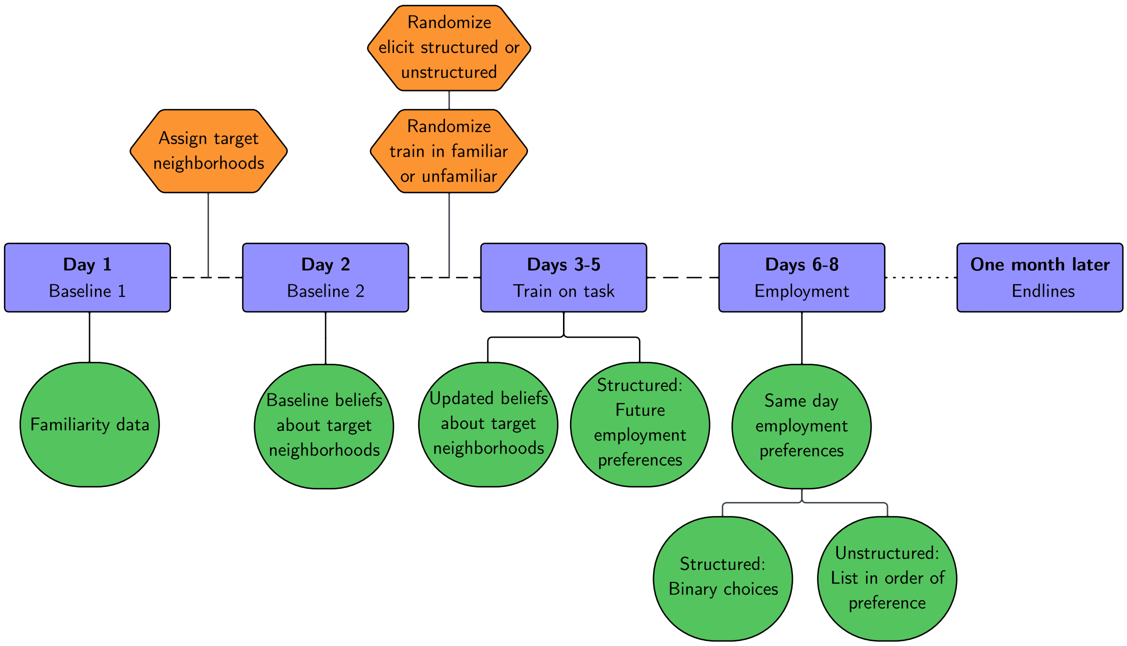


Notes: This figure replicates Figure 2 with a more expansive sample of neighborhoods – all neighborhoods within 75 min transit or walking distance.

Figure A.3: Backpacks Used in Employment Task

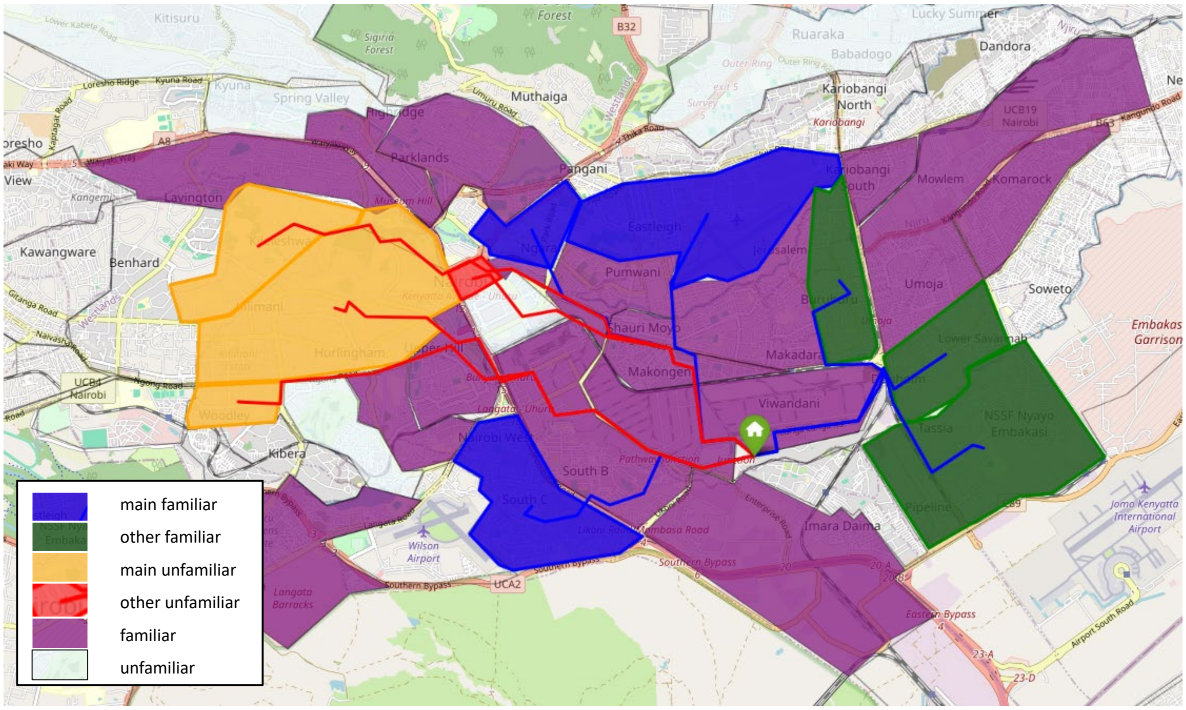
Notes: This figure shows a picture of an air pollution backpack used by study participants during training and employment days.

Figure A.4: Timeline of Study



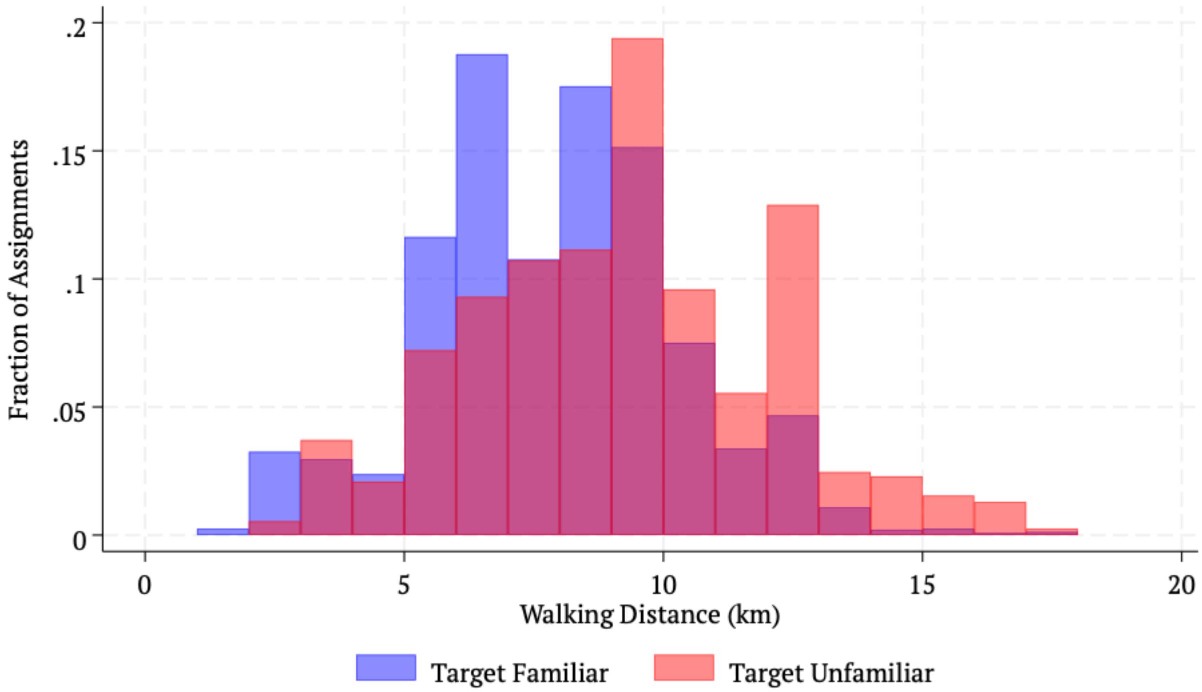
Notes: This figure shows the timing outline of the study. In the first two days, participants provided information about their neighborhood familiarity, demographics, beliefs about target neighborhoods, employment and job search status, and spatial ability. Over the next three days, participants were asked to collect air pollution data in either familiar or unfamiliar neighborhoods and received training on how to do so. Upon returning, they were asked about belief updating and familiarity with the neighborhood visited. Participants assigned to the "structured" group were elicited about their job preferences one day at a time for the upcoming employment days. On the last three days, participants worked in their top-preferred neighborhood, as elicited through either an "unstructured" survey or "structured" binary choices—Section 4 discusses the "structured" surveys in more detail. They concluded the employment days with a survey about trip duration, directions, and familiarity checks. After a month, participants were contacted by phone to inquire about their employment and job search status, spatial ability, and neighborhood trip patterns through SMS.

Figure A.5: Example of 10 Target Neighborhoods for One Participant



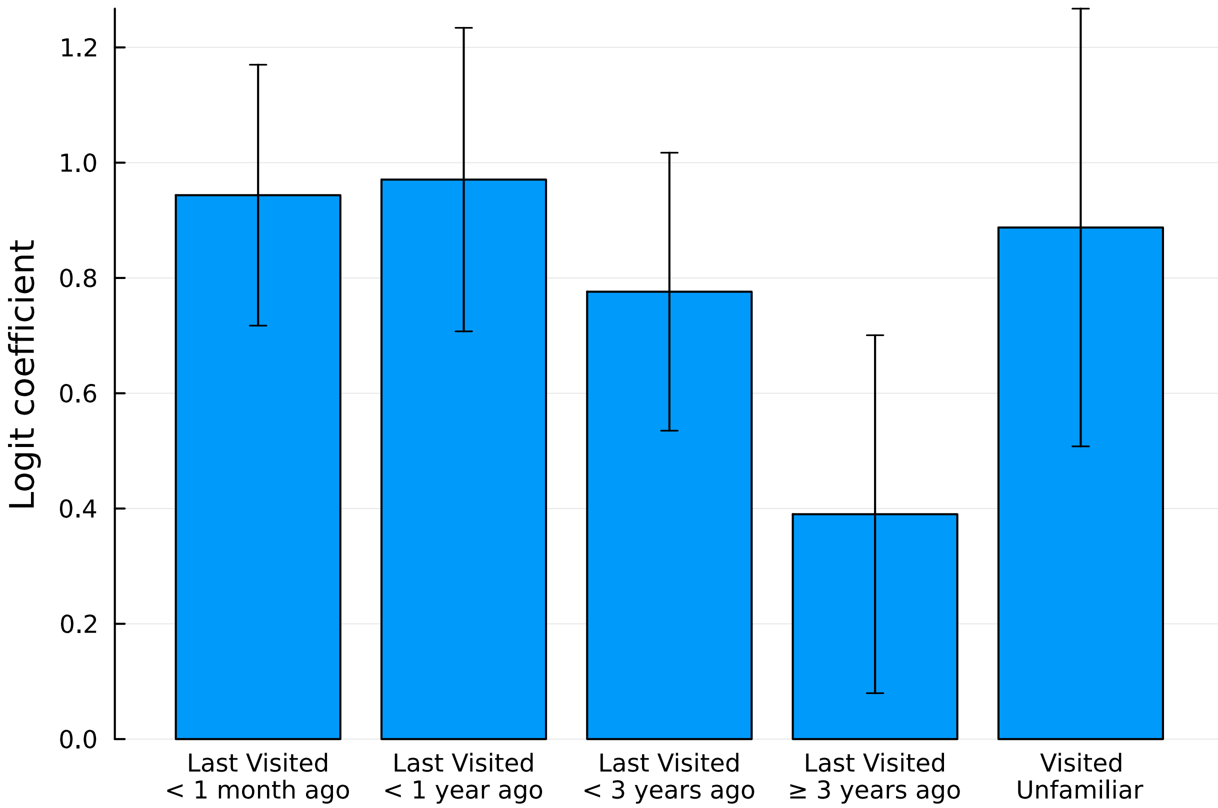
Notes: This map plots the ten target neighborhoods for one participant from the home neighborhood of Viwandani. Non-target neighborhoods are shaded in purple if they had ever visited at baseline or white if they had not. Ambiguous neighborhoods are not shaded. The four target categories are shared in different colors: blue for the three main familiar neighborhoods, green for the three other familiar neighborhoods, yellow for the three main unfamiliar neighborhoods, and red for the single other unfamiliar neighborhood. The walk or transit route to each neighborhood is also plotted.

Figure A.6: Overlap in Distances to Home Neighborhood Between Familiar and Unfamiliar Assignments



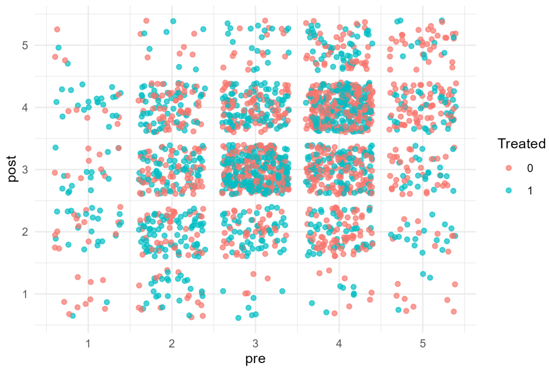
Notes: This figure plots the distribution of walking distances to home neighborhood in kilometers by target familiar and unfamiliar neighborhoods. The sample includes 4796 observations, split as 799 participants each assigned 3 target familiar and 3 target unfamiliar neighborhoods. See Table A.4 for further results on the distance differences.

Figure A.7: Familiarity Lasts 3 Years

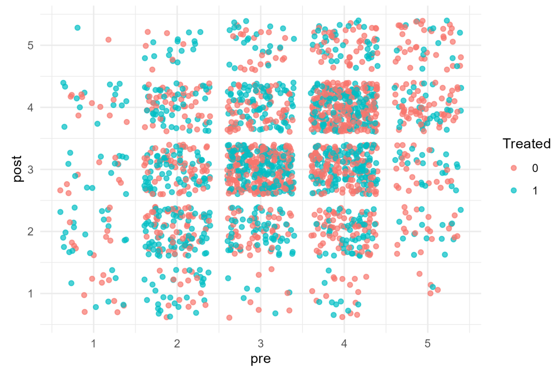


Notes: This figure plots the coefficients of Table A.12 column 2.

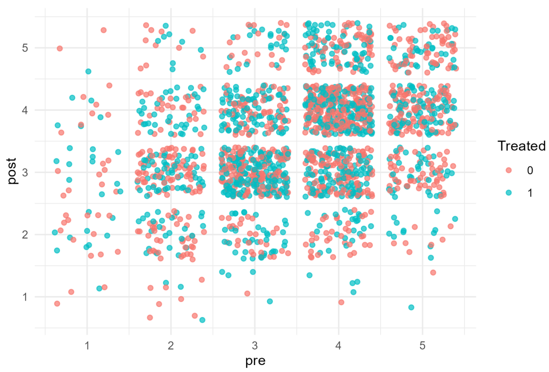
Figure A.8: Beliefs Transition Matrices



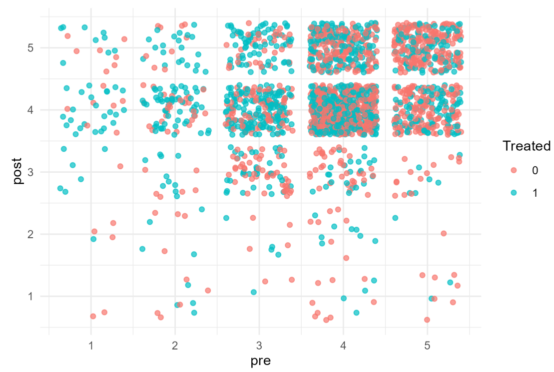
(A) Find Job



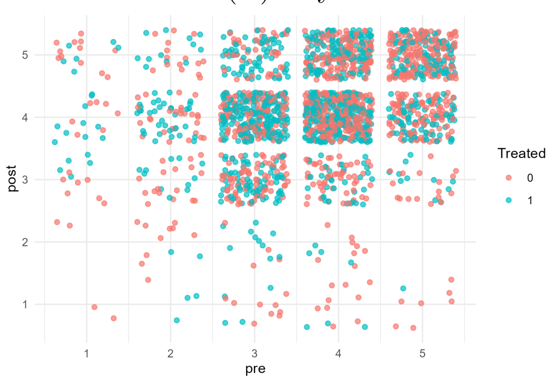
(B) Find Job - Others



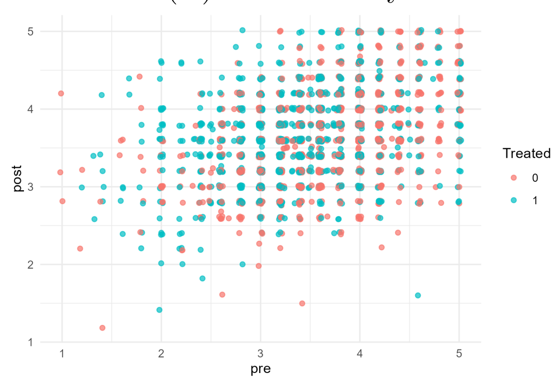
(C) Pay



(D) Travel Safety



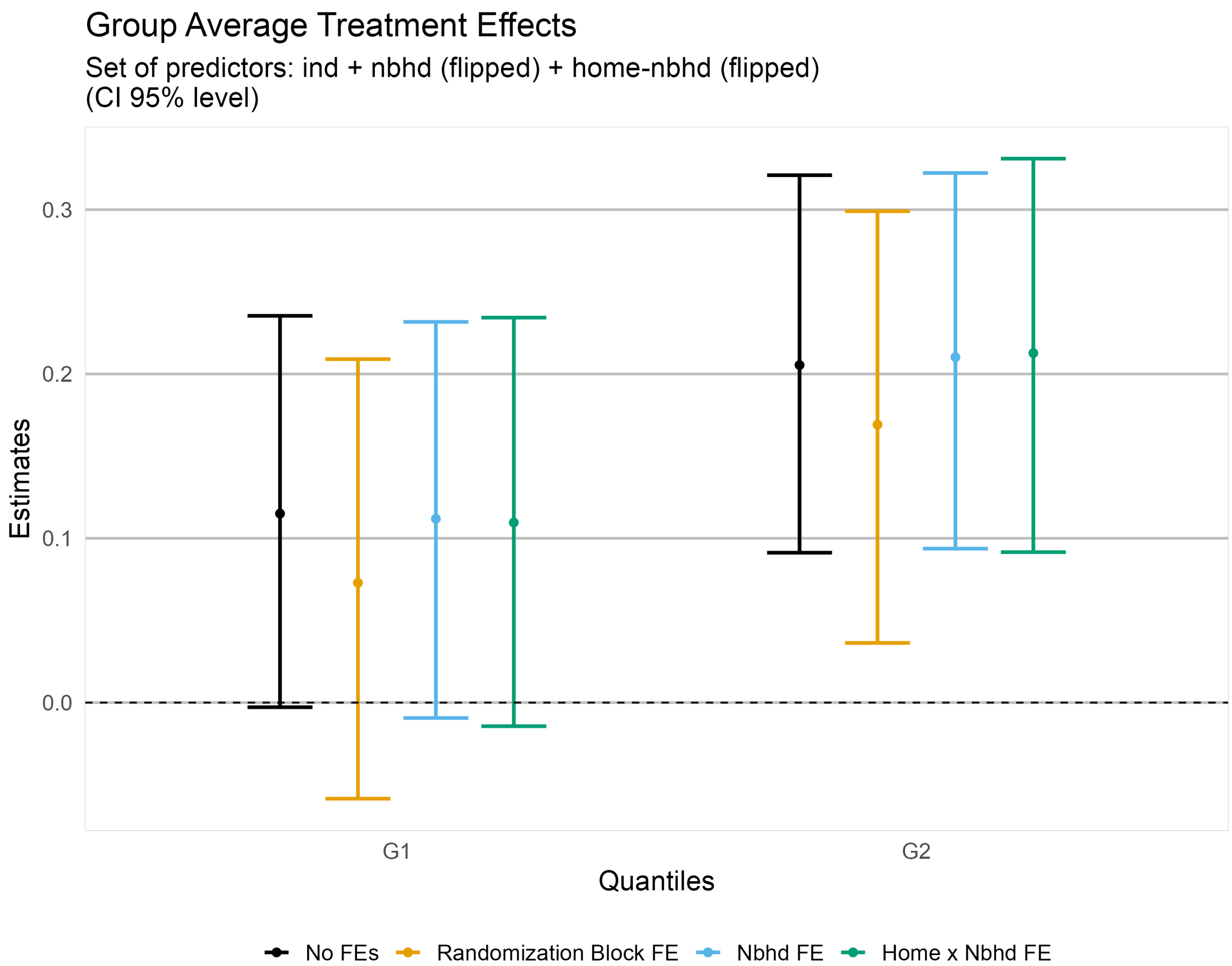
(E) Overall Safety



(F) Average (All Measures)

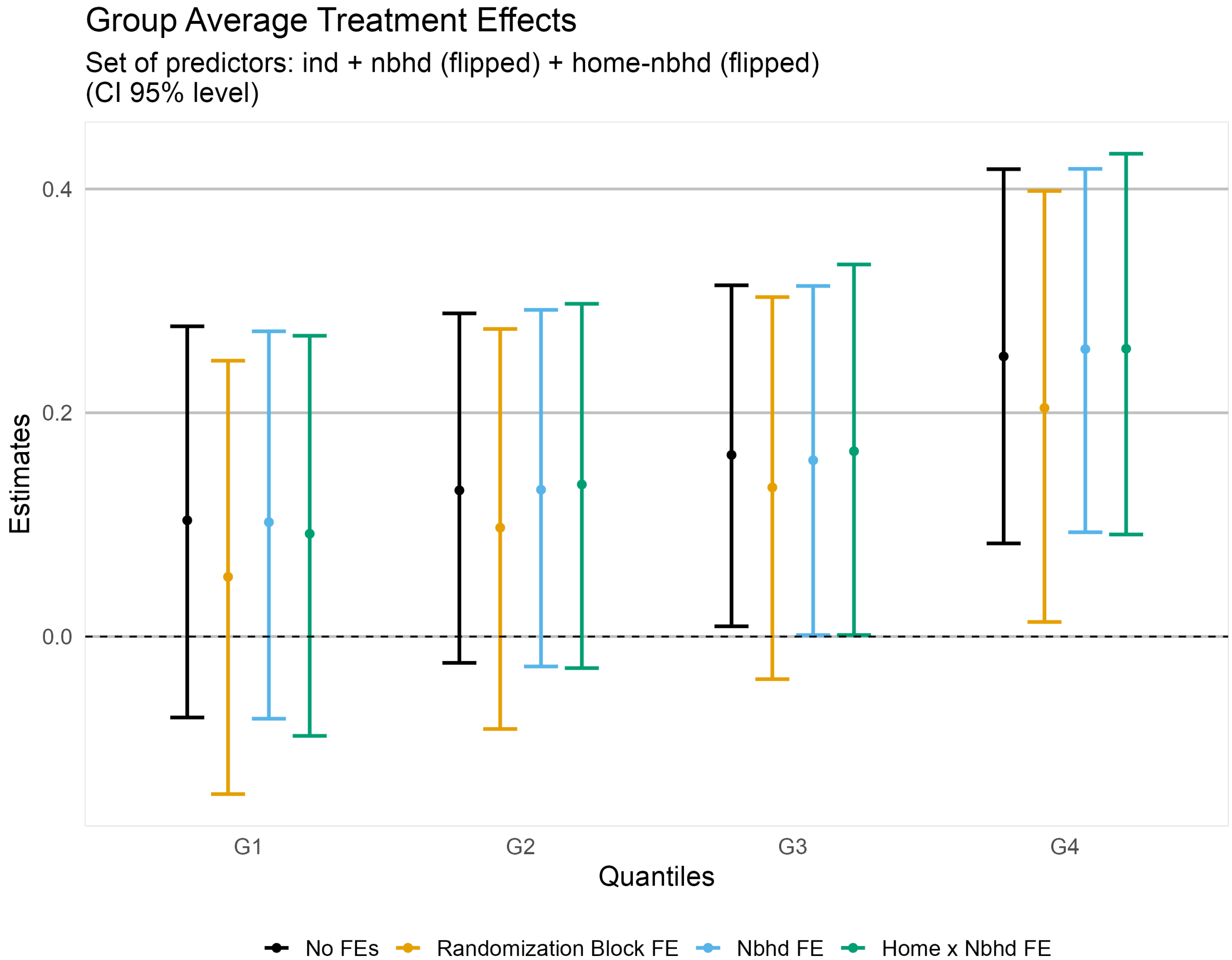
Notes: each graph plots beliefs pre- and post-intervention on the X- and Y-axis, respectively. Higher values correspond to better outcomes. See Section A.3 for the precise questions underlying each measure.

Figure A.9: Median Split of Heterogeneous Treatment Effects



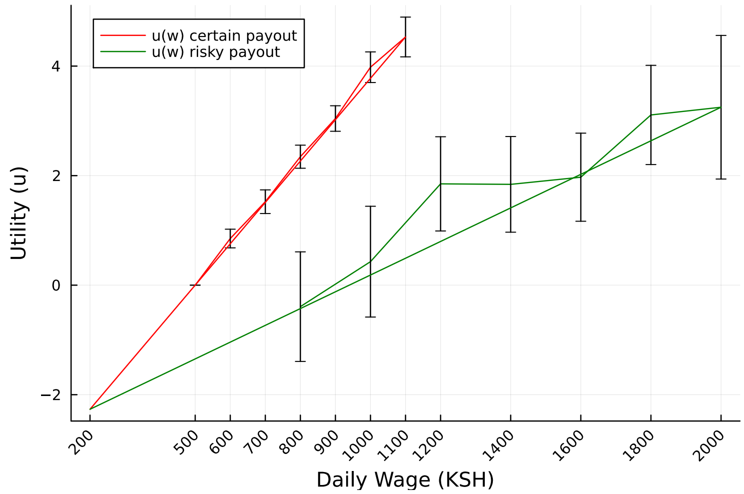
Notes: The figure shows two quartiles of the Group Average Treatment Effects estimated with the Chernozhukov et al. (2023) method. To employ the method we restrict to structured choices between an non-target familiar neighborhood and an unfamiliar neighborhood. We then use as the outcome an indicator for choosing to work in an unfamiliar neighborhood.

Figure A.10: Quartile Splits of Heterogeneous Treatment Effects



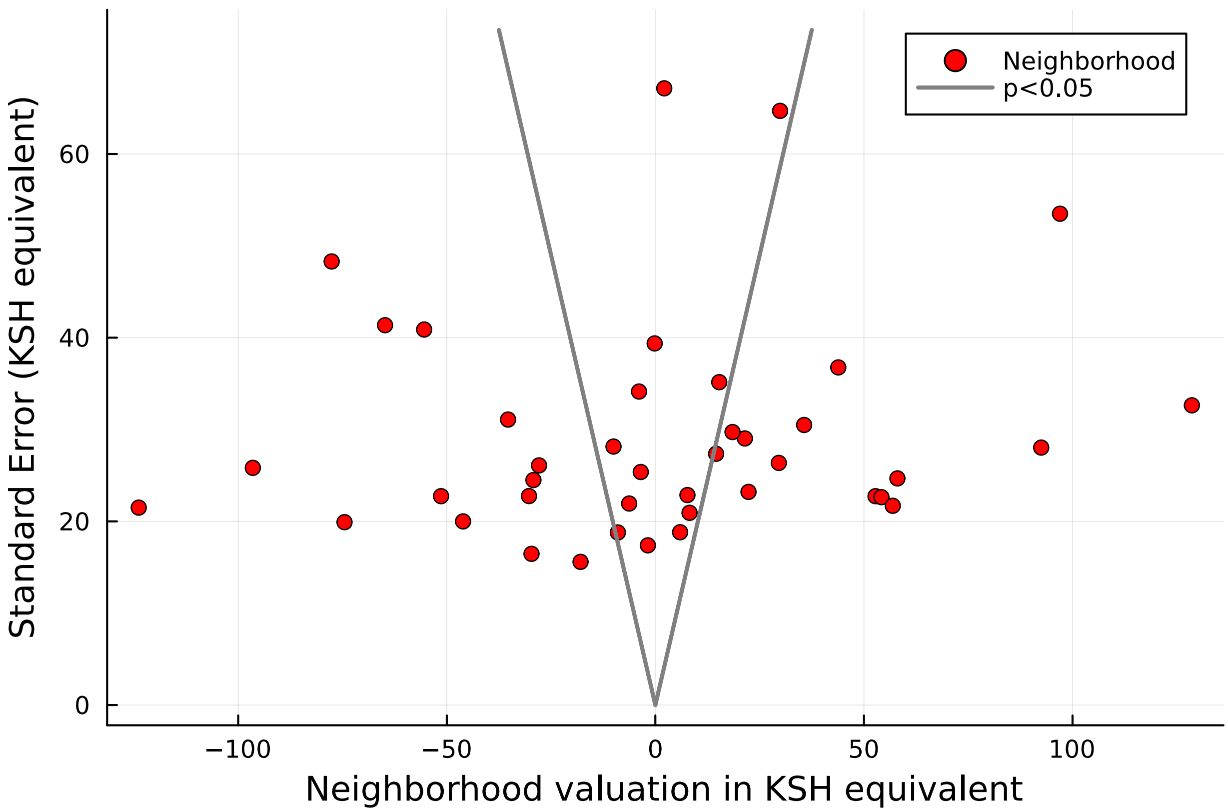
Notes: This figure replicates Figure A.9 using four quartiles of the Group Average Treatment Effects.

Figure A.11: Linear Preferences over Daily Wages



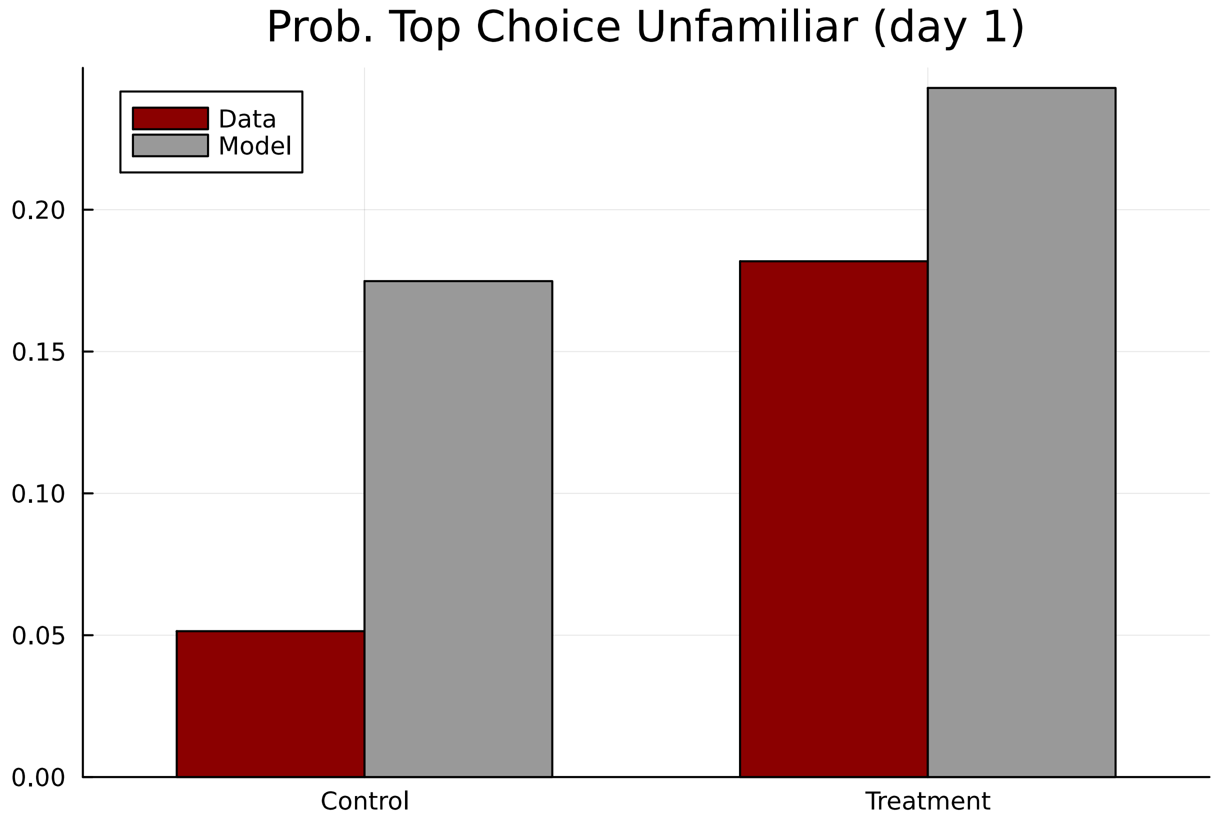
Notes: **(NOTES REQUIRED)**. Risky payout jobs: 200 KSH for sure and 50%-50% change of bonus.

Figure A.12: Distribution of Neighborhood Fixed Effects



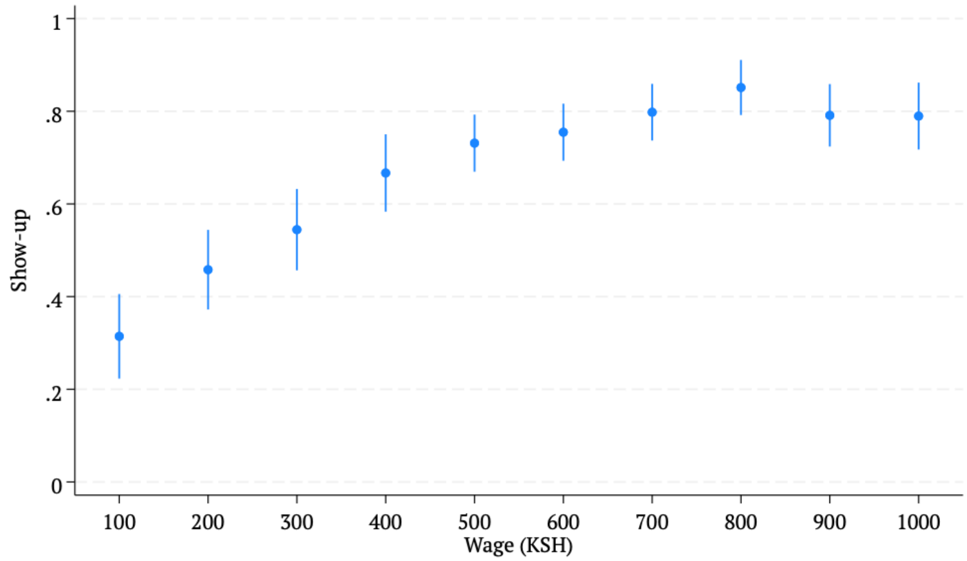
Notes: **(NOTES REQUIRED)**.

Figure A.13: Comparison of First Choices Predicted by Structured and Realized Data: Any Unfamiliar Neighborhood



Notes: this figure replicates Figure 8 where the Y axis measures the share of respondents who mention *any baseline unfamiliar* neighborhood as their top choice.

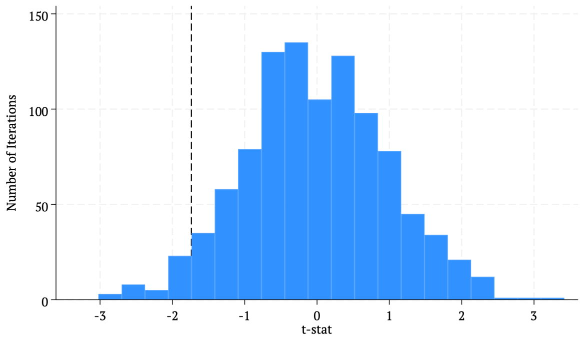
Figure A.14: Show-up 2-4 Months After Intervention: High Wage Plateau



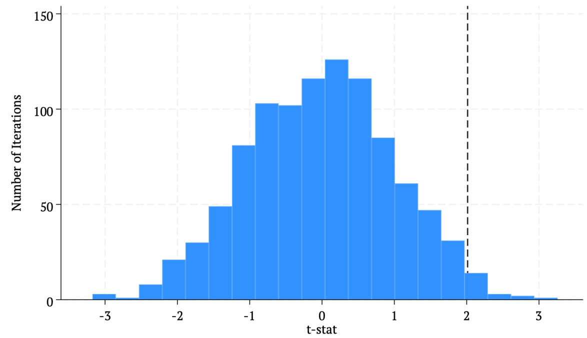
Notes: This figure plots average show-up for the travel survey task by (random) wage offer.

Figure A.15: Randomization Inference Results for Imbalances in the Structured Sample

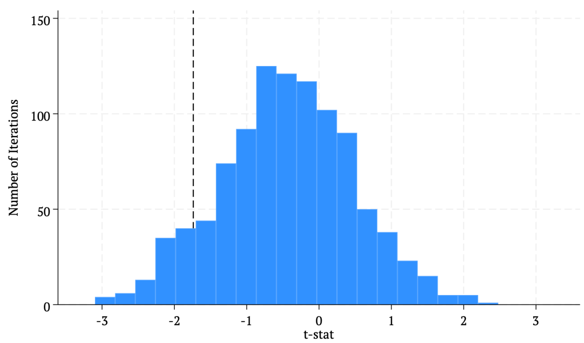
(a) Age



(b) Enrolled in School

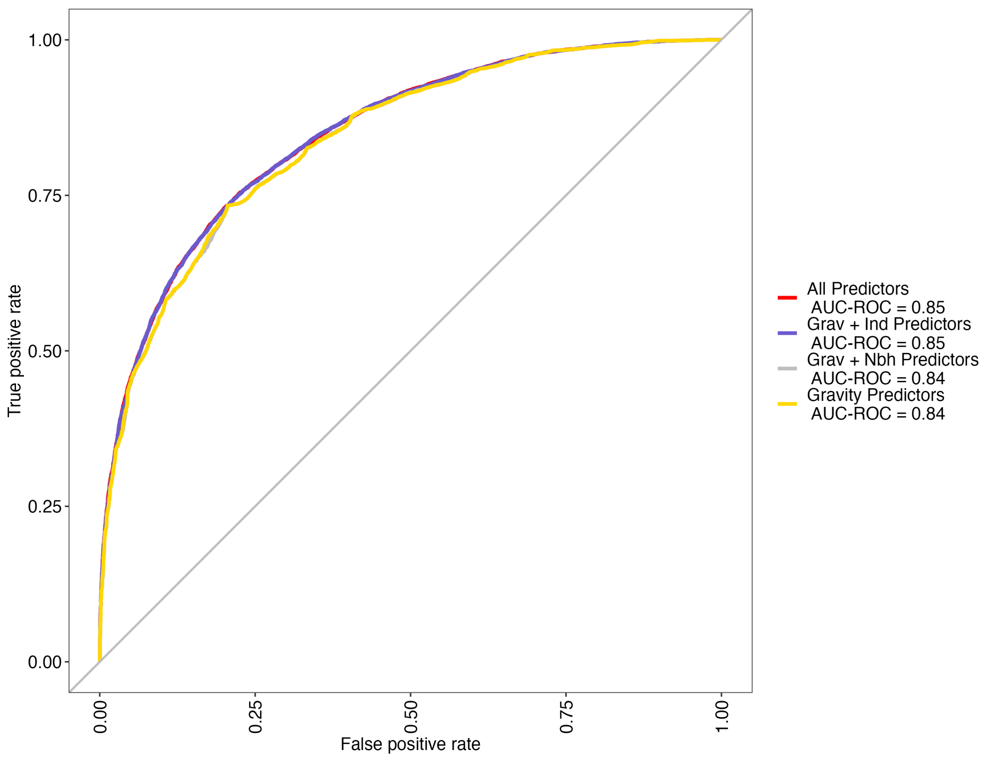


(c) Days Searched for Work Last 2 Weeks



Notes: This figure plots the distribution of the t-statistics from mean difference tests between treatment and control groups across 1000 simulations of treatment assignment. The sample includes only participants in the structured group, focusing on variables with evidence of imbalance (See Table A.8 for further results). The dotted lines represent the observed t-statistics for the mean differences in the structured sample.

Figure A.16: No Evidence of Match-Specific Familiarity Patterns



Notes: This figure replicates Figure 4 results using logit-lasso models.

A.2 Appendix Tables

Table A.1: Top 10 Participant Jobs Over the Last Two Weeks

Women		Men	
Occupation	Share	Occupation	Share
Laundry	62%	Carpenter/Mason	33%
Cleaner	14%	Industrial/factory worker	9.6%
Househelp	9.8%	Electrician	6%
Washing dishes/utensils	8.2%	Cleaner	5.4%
Cook	7.3%	Small Business	4.8%
Salon	6.6%	Mechanic	4.2%
Waiter	3.9%	Cook	3.6%
Small Business	3.7%	Carrying luggage	3.6%
Industrial/factory worker	2.7%	Plumber	3.6%
Sales person	2.7%	Boda boda operator	3%

Notes: This table lists the top 10 most frequent jobs performed by participants in the last two weeks, categorized by gender. The sample consists of 605 participants—438 women and 167 men—who worked at least one day during the two weeks prior to the intervention days. The percentage frequencies are derived from a multiple-choice format question, so each row represents the share of the total gender-specific sample.

Table A.2: Sample characteristics

	Mean	Std. Dev.
Female	0.74	0.43
Age	29.4	7.6
Education Years	10.7	2.7
Enrolled in School	0.04	0.19
Workdays Last 2 Weeks	3.1	2.9
Days Searched for Work Last 2 Weeks	6.6	3.4
Married	0.46	0.50
Years in Nairobi	15.2	9.3
Resided Outside Nairobi	0.67	0.47
Observations	799	

Notes: This table reports basic statistics on participant characteristics. The sample includes only those participants who attended the first training day.

Table A.3: Casual Workers Use Spatial Search Strategies

	Job Search Strategies			
	In Last Two Weeks:		**Ever Found Work:**	
	Women	Men	Women	Men
Travel to other nbhd	0.64	0.60		
Door to door	0.36	0.20	0.51	0.29
Hiring spot	0.38	0.35	0.43	0.38
Drop CV	0.16	0.19	0.16	0.24
Ask people I know/employer	0.88	0.91		
In person referral			0.77	0.90
Receive call (referral)			0.72	0.73
Online	0.15	0.27	0.04	0.11
Observations	617	213	570	199

Notes: This table shows the share of participants who reported using specific spatial job search strategies, categorized by gender. The first two columns include participants who searched for work in the last two weeks prior to the first training day. The next two columns include participants who were contacted through phone calls at the endline. The shares should be interpreted relative to the sample size stated at the bottom of each column.

Table A.4: Familiar vs Unfamiliar Neighborhoods

	Distance (km)
	(1)
Target Familiar	-1.36***
	(0.09)
Constant	9.29***
	(0.04)
Individual FEs	Yes
Observations	4794

Notes: This table compares the distance to the home neighborhood across all target familiar and unfamiliar neighborhoods. The sample includes 4796 observations, split as 799 participants each assigned 3 target familiar and 3 target unfamiliar neighborhoods. The dependent variable is the walking distance to the home neighborhood in kilometers. Standard errors in parenthesis are clustered at the individual level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Table A.5: Correlates of Individual-level Average Familiarity

	Average "Ever Been"	
	(1)	(2)
Female	-0.053^{***}	-0.056^{***}
	(0.011)	(0.011)
Age	0.003^{***}	0.003^{***}
	(0.001)	(0.001)
Years of education	0.012^{***}	0.012^{***}
	(0.002)	(0.002)
Years in Nairobi	0.002^{***}	0.003^{***}
	(0.0005)	(0.0005)
Spatial Ability Idx	0.014^{***}	0.018^{***}
	(0.005)	(0.005)
N	827	827
Outcome SD	0.176	0.176
Surveyor FEs	No	Yes
Adjusted R^2	0.140	0.209

Notes: This table reports the correlation between average familiarity at the participant level and participant characteristics. The sample includes participants who completed Baseline 2. Standard errors in parenthesis are clustered at the individual level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.6: Ethnicity is Not A Large Factor in Determining Familiarity

	Participant i "Ever Been" to neighborhood j				
	(1)	(2)	(3)	(4)	(5)
Distance (km)	-0.045^{***}		-0.045^{***}	-0.056^{***}	
	(0.001)		(0.001)	(0.001)	
Same Ethnicity		0.063^{***}	-0.007	0.016	0.016
		(0.017)	(0.015)	(0.015)	(0.015)
Individual FEs	Yes	Yes	Yes	Yes	Yes
Neighborhood FEs				Yes	
Home x Neighborhood FEs					Yes
Observations	13,331	13,331	13,331	13,331	13,331
Adjusted R^2	0.165	0.062	0.165	0.347	0.384

Notes: This table reports the correlation between participant familiarity and the main ethnicity of the neighborhood. The sample consists of participants who completed both Baseline 2 and the phone surveys at the endline. Participants were asked about their ethnic identity during the endline survey. A belief question about the largest ethnic group in the neighborhood was added in baseline 2 for later batches. About 50% of baseline 2 participants were asked this ethnic belief question. A neighborhood's dominant ethnicity is defined as the most common ethnic group reported by respondents familiar with that neighborhood. *Same Ethnicity* is an indicator for whether a respondent's ethnicity matches the neighborhood's dominant ethnicity. *Distance (km)* represents the walking distance in kilometers, from the study venues to the specific neighborhood centroid, as estimated from the Google Maps API. Standard errors in parentheses are clustered at the individual level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.7: Balance on Main Study Sample

	Obs	Treatment Mean	Control Mean	P-value
Female	799	0.72	0.77	0.23
Age	799	29.11	29.74	0.55
Education Years	798	10.63	10.71	0.67
Enrolled in School	799	0.05	0.03	0.40
Workdays Last 2 Weeks	799	3.02	3.14	0.50
Days Searched for Work Last 2 Weeks	799	6.51	6.61	0.64
Married	799	0.46	0.47	0.91
Years in Nairobi	799	15.24	15.14	0.52
Resided Outside Nairobi	799	0.65	0.69	0.28
Joint P-value				0.84

Notes: This table reports the mean test differences between participants assigned to treatment and control groups. The sample includes only those participants who attended the first training day. The last column reports the p-value for the test of treatment in a regression including randomization block fixed effects, and the joint p-value is a test of the joint equality of all listed treatment and control mean characteristics.

Table A.8: Balance on Structured Sample

	Obs	Treatment Mean	Control Mean	P-value
Female	400	0.72	0.75	0.31
Age	400	28.54	30.31	0.08
Education Years	399	10.69	10.79	0.54
Enrolled in School	400	0.05	0.00	0.04
Workdays Last 2 Weeks	400	2.96	3.05	0.69
Days Searched for Work Last 2 Weeks	400	6.32	6.87	0.08
Married	400	0.39	0.48	0.30
Years in Nairobi	400	14.88	16.08	0.27
Resided Outside Nairobi	400	0.67	0.66	0.71
Joint P-value				0.25

Notes: This table replicates Table A.7, restricted to participants in the "structured" group. See Figure A.15 for randomization inference results on age, school enrollment, and days searched for work last 2 weeks.

Table A.9: Show-up for Air Pollution Jobs

	Training (1)	Employment (2)	Endline (3)
Treated	-0.007*	-0.022**	-0.001
	(0.004)	(0.010)	(0.014)
Day 2	-0.005**		
	(0.003)		
Day 3	-0.013***		
	(0.004)		
Day 5		-0.010***	
		(0.004)	
Day 6		-0.023***	
		(0.005)	
Control Mean	0.998	0.983	0.963
Observations	2397	2397	799

Notes: This table shows the show-up rates at different stages of the study timeline. Columns 1 and 2 are panels of 799 participants—who attended the first training day—over the three training and employment days, respectively. Column 3 includes 617 participants who were reached by phone survey calls at the endline. All regressions include fixed effects for participant batch assignments in the study Standard errors in parentheses are clustered at the individual level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.10: Familiarity Check After Job Training

	(1) Ever Been
Treated	-0.54***
	(0.02)
Constant	0.88***
	(0.01)
Observations	2385

Notes: This table checks the familiarity reported by participants after job training and their treatment assignment. The sample is an unbalanced panel of 799 participants over the three training days. The dependent variable takes a value of one if, after returning to the study venue in the home neighborhood from a standard training day, the participant had previously visited the assigned neighborhood; it is assigned a value of zero otherwise. Standard errors in parenthesis are clustered at the individual level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.11: Revealed Preferences Estimates of Familiarity Premium - OLS

	(1)	(2)	(3)	(4)
Baseline Familiar	0.099***	0.097***	0.081***	0.096***
	(0.014)	(0.014)	(0.014)	(0.014)
Visited Any	0.039**	0.039**	0.042**	0.040**
	(0.014)	(0.014)	(0.013)	(0.013)
Visited Unfamiliar	0.093***	0.096***	0.098***	0.099***
	(0.026)	(0.027)	(0.025)	(0.025)
Distance (km)	-0.029***	-0.029***		-0.030***
	(0.002)	(0.002)		(0.002)
Job duration (hrs)	-0.069***	-0.071***	-0.072***	-0.071***
	(0.007)	(0.007)	(0.007)	(0.007)
$\mathbb{E}$ Compensation (KSH)	0.102***	0.101***	0.102***	0.101***
	(0.002)	(0.002)	(0.002)	(0.002)
Cash Upfront (KSH)	0.007	0.007	0.008	0.008
	(0.006)	(0.006)	(0.006)	(0.006)
Bonus (KSH)	-0.028***	-0.028***	-0.028***	-0.028***
	(0.001)	(0.001)	(0.001)	(0.001)
N	6,756	6,756	6,756	6,756
R^2	0.470	0.504	0.527	0.518
Within-R^2		0.470	0.495	0.486
Person FE		Yes	Yes	Yes
Home $\times$ neighborhood FE			Yes	
Neighborhood FE				Yes

Notes: This table replicates results from Table 2 but uses a linear probability model. Standard errors in parenthesis are clustered at the individual level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.12: Cross-Sectional Premium Persists Up to Three Years Since Last Visit

	(1)	(2)
Baseline Familiar	0.86***	
	(0.125)	
Visited < 1 month ago		0.94***
		(0.130)
Visited < 1 year ago		0.97***
		(0.136)
Visited < 3 years ago		0.78***
		(0.134)
Visited ≥ 3 years ago		0.39*
		(0.169)
Visited Any	0.30**	0.31**
	(0.109)	(0.106)
Visited Unfamiliar	0.91***	0.89***
	(0.213)	(0.215)
N	6,756	6,756
Job Attribute Controls	Yes	Yes
NBH FE	Yes	Yes

Notes: This table presents the logit of on the "structured" choices shown in Table 2 but splitting the baseline familiar coefficient based on how long ago the individual last visited the neighborhood. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.13: Preferences for Familiar Nbhds By Distance

	(1)	(2)	(3)
Baseline Familiar × Close	0.89***	0.77***	0.88***
	(0.134)	(0.137)	(0.142)
Baseline Familiar × Far	0.83***	0.75***	0.85***
	(0.122)	(0.134)	(0.141)
Visited Any × Close	0.39**	0.29*	0.29*
	(0.134)	(0.129)	(0.138)
Visited Any × Far	0.18	0.36**	0.31*
	(0.138)	(0.125)	(0.137)
Visited Unfamiliar × Close	1.00***	0.98***	1.01***
	(0.243)	(0.233)	(0.258)
Visited Unfamiliar × Far	0.70**	0.88***	0.82**
	(0.240)	(0.223)	(0.255)
N	6,756	6,756	6,756
Home × neighborhood FE		Yes	
Neighborhood FE			Yes

Notes: This table replicates results from Table 2 accounting for proximity heterogeneity. The proximity interaction term is defined as travel durations below or above the median (roughly 1 hour).

Table A.14: Preferences for Familiar Nbhds By Gender

	(1)	(2)
Baseline Familiar	0.74***	1.42***
	(0.122)	(0.219)
Visited Any	0.38**	0.10
	(0.130)	(0.212)
Visited Unfamiliar	0.72**	1.54***
	(0.235)	(0.436)
Distance (km)	-0.26***	-0.25***
	(0.022)	(0.032)
Job duration (hrs)	-0.51***	-0.84***
	(0.061)	(0.132)
$\mathbb{E}$ Compensation (KSH)	0.78***	0.93***
	(0.034)	(0.064)
Cash Upfront (KHS)	0.07	0.02
	(0.052)	(0.085)
Bonus (KSH)	-0.22***	-0.24***
	(0.011)	(0.022)
Controls	Yes	Yes
N	4,952	1,804
NBH FE	Yes	Yes
Sample	Women	Men

Notes: This table replicates results from Table 2 for the women (column 1) and men (column 2) samples.

Table A.15: Familiarity Premium Concentrated For Strong Familiarity

	(1)	(2)	(3)
Baseline Familiar		0.88***	
		(0.117)	
Baseline Familiar × Know Get There	0.90***		
	(0.116)		
Baseline Familiar × Not Know Get There	0.22		
	(0.177)		
Not Baseline Familiar × Know Get There		0.16	
		(0.153)	
Baseline Familiar × Landmark			1.00***
			(0.120)
Baseline Familiar × No Landmark			0.54***
			(0.124)
Visited Any	0.30**	0.29**	0.31**
	(0.110)	(0.110)	(0.111)
Visited Unfamiliar	0.83***	0.85***	0.82***
	(0.211)	(0.207)	(0.212)
N	6,756	6,756	6,756
Job Attribute Controls	Yes	Yes	Yes

Notes: This table replicates results from Table 2 considering strong familiarity at baseline.

Table A.16: Spillovers To Other Unfamiliar Neighborhood

	(1)	(2)	(3)
Baseline Familiar	0.77***	0.70***	0.77***
	(0.103)	(0.098)	(0.106)
Visited Any	0.29**	0.32***	0.29**
	(0.106)	(0.089)	(0.097)
Visited Unfamiliar	0.75***	0.87***	0.84***
	(0.202)	(0.164)	(0.193)
Unfamiliar Not Visited × Treated	0.25	0.24	0.27
	(0.173)	(0.163)	(0.184)
Distance (km)	-0.24***		-0.25***
	(0.014)		(0.017)
Job duration (hrs)	-0.58***	-0.61***	-0.59***
	(0.052)	(0.055)	(0.052)
$\mathbb{E}$ Compensation (KSH)	0.75***	0.80***	0.78***
	(0.024)	(0.026)	(0.025)
Cash Upfront (KHS)	0.08*	0.09*	0.09*
	(0.040)	(0.039)	(0.043)
Bonus (KSH)	-0.21***	-0.22***	-0.21***
	(0.010)	(0.010)	(0.010)
N	8,183	8,183	8,183
Home × neighborhood FE		Yes	
Neighborhood FE			Yes

Bootstrapped standard errors, clustered at the individual level.

Notes: This table replicates results from Table 2 to explore spillover effects on the unfamiliar neighborhoods not visited by the treated group.

Table A.17: Confidence on Beliefs After One Visit

	Dependent variable:				
	Find Job	Find Job - Others	Pay	Travel Safety	Overall Safety
	(1)	(2)	(3)	(4)	(5)
Post	0.131^{***}	0.081^{*}	-0.044	0.145^{***}	0.163^{***}
	(0.042)	(0.042)	(0.038)	(0.034)	(0.038)
Pre $\times$ Treated	-0.101	-0.116^{*}	-0.141^{**}	-0.164^{***}	-0.125^{**}
	(0.063)	(0.065)	(0.060)	(0.058)	(0.056)
Post $\times$ Treated	-0.075	-0.010	-0.027	0.048	0.006
	(0.053)	(0.057)	(0.052)	(0.042)	(0.045)
Mean	4.1	4.1	4.3	4.4	4.3
SD	1.0	1.1	0.9	0.9	0.9
Home $\times$ neighborhood FE	Yes	Yes	Yes	Yes	Yes
Observations	4,555	4,544	4,534	4,545	4,524

Notes: This table replicates estimates from Table 3 using participant confidence in their reported beliefs after one visit. All outcomes are rated on a Likert scale from 1 to 5, with 1 being not confident at all and 5 being the most confident. Standard errors in parenthesis are clustered at the individual level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.18: Beliefs "Don't Know" After One Visit

	Dependent variable: Responded "Don't Know"				
	Find Job	Find Job - Others	Pay	Travel Safety	Overall Safety
	(1)	(2)	(3)	(4)	(5)
Post	-0.031^{***}	-0.035^{***}	-0.030^{***}	-0.023^{***}	-0.031^{***}
	(0.007)	(0.007)	(0.008)	(0.006)	(0.008)
Pre × Treated	0.072^{***}	0.082^{***}	0.079^{***}	0.097^{***}	0.104^{***}
	(0.015)	(0.015)	(0.017)	(0.015)	(0.017)
Post × Treated	-0.016^{***}	-0.013^{**}	-0.014	-0.015^{***}	-0.019^{**}
	(0.006)	(0.005)	(0.010)	(0.005)	(0.007)
Mean	0.03	0.04	0.05	0.02	0.04
SD	0.2	0.2	0.2	0.2	0.2
Home × neighborhood FE	Yes	Yes	Yes	Yes	Yes
Observations	4,587	4,587	4,587	4,587	4,587

Notes: This table replicates estimates from Table 3 using participant unknownness as the dependent variable. The dependent variable is a dummy that takes a value of one if a participant does not know the attributes being asked about. Standard errors in parenthesis are clustered at the individual level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.19: Impact on Job Search and Work - Unprompted Measures

	Endline - (unprompted)				Smartphone	
	Any Trip	Search for Work	Work	Other	Visited	Num Visits
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline Familiar	0.024***	0.016***	0.003	0.007**	0.016**	0.340*
	(0.007)	(0.006)	(0.002)	(0.003)	(0.006)	(0.186)
Visited Any	0.020**	0.023**	−0.001	−0.003	−0.004	−0.192
	(0.010)	(0.009)	(0.003)	(0.004)	(0.008)	(0.211)
Visited Unfamiliar	−0.008	−0.018	0.005	0.008	0.007	0.312
	(0.015)	(0.014)	(0.005)	(0.007)	(0.014)	(0.411)
Mean	0.035	0.03	0.003	0.005	0.011	0.222
SD	0.185	0.169	0.059	0.071	0.058	1.315
Visited Any + Visited Unfam	0.012	0.005	0.004	0.005	0.003	0.12
[p-value]	[0.082]	[0.377]	[0.203]	[0.202]	[0.688]	[0.729]
Observations	6,896	6,896	6,896	6,896	941	941

Notes: The table regresses whether we observe individuals revisiting target neighborhoods through two unprompted measures on indicators for whether the neighborhood was familiar at baseline, whether the individual trained in the neighborhood during the experiment and whether the trained neighborhood was unfamiliar at baseline. Columns 1-4 show the results from asking participants open-ended questions about where they have recently visited, while columns 5 and 6 include trips measured by the GPS tracking application. Standard errors in parenthesis are clustered at the individual level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.20: Impact on Job Search and Work - Number of Neighborhoods Visited

	Endline				SMS	Smartphone
	Any Trip	Search for Work	Work	Other	Any Trip	Any Trip
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	−0.081	−0.072	0.001	−0.019	−0.001	0.179
	(0.060)	(0.055)	(0.025)	(0.031)	(0.025)	(0.159)
Constant	1.023***	0.777***	0.135***	0.199***	0.940***	1.820***
	(0.043)	(0.039)	(0.018)	(0.023)	(0.019)	(0.101)
Observations	769	769	768	769	553	105

Notes: The table regresses the total number of neighborhoods individuals report visiting across different modes of elicitation on treatment status. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.21: Impact on Job Search and Work - Number of Unfamiliar Neighborhoods Visited

	Endline				SMS	Smartphone
	Any Trip	Search for Work	Work	Other	Any Trip	Any Trip
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.0004	0.0003	0.003	0.003	0.073***	−0.120
	(0.017)	(0.015)	(0.006)	(0.007)	(0.025)	(0.091)
Constant	0.054***	0.044***	0.005	0.008*	0.085***	0.270***
	(0.012)	(0.010)	(0.004)	(0.004)	(0.014)	(0.076)
Observations	769	769	768	769	451	71

Notes: The table regresses the number of unfamiliar neighborhoods visited on treatment status for different elicitation methods. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A.22: No Differential Contact Rate for Job Offer

	Call answered	
	(1)	(2)
Constant	0.89***	
	(0.01)	
Treated	-0.005	-0.004
	(0.02)	(0.02)
Date FE	No	Yes
Observations	4,482	4,482

Notes: The table examines differential response to the job offers between treatment and control groups.

A.3 Measuring Beliefs

In our surveys, we ask the following beliefs questions to all participants:

1. Think about an average person who lives in your home neighborhood. If this person goes to X to find daily or casual work opportunities, they are likely to find one.
    - 1=Strongly disagree, 2=Disagree, 3=Neutral, 4=Agree, 5=Strongly agree.
2. If you go to X to find daily or casual work opportunities, you are likely to find one.
    - 1=Strongly disagree, 2=Disagree, 3=Neutral, 4=Agree, 5=Strongly agree.
3. If you find a daily or casual work opportunity in X, the pay is good.
    - 1=Strongly disagree, 2=Disagree, 3=Neutral, 4=Agree, 5=Strongly agree.
4. If you travel alone, how safe is the trip from your home to X?
    - 1=Very Unsafe, 2=Unsafe, 3=Neutral, 4=Safe, 5=Very Safe.
5. How safe do you think X is?
    - 1=Very Unsafe, 2=Unsafe, 3=Neutral, 4=Safe, 5=Very Safe.

After each question, we also ask:

- How confident are you in the above answer?
    - 1=Not confident at all, 2=Slightly confident, 3=Somewhat confident, 4=Fairly confident, 5=Completely confident.