The Effect of Exposure: Evidence from Spatial Choices in Nairobi*

Joshua T. Dean, Gabriel Kreindler, Oluchi Mbonu †

September 7th, 2025

Abstract

How much do people dislike opportunities they have never been exposed to, and why? We study this in the context of how exposure with locations affects how casual workers in Nairobi make work location decisions. We offer short-term employment opportunities across the city and experimentally induce exposure to locations by training participants in either familiar or unfamiliar neighborhoods. We find that participants are willing to travel 3.5 km further or take a pay cut worth 22% of the median daily wage to avoid working in a location never visited before. This differential is fully offset after one visit to an unfamiliar neighborhood. These results are inconsistent with sorting and neighborhood quality risk channels, and we find little evidence of first-time navigation costs or exploration risk. The results are most consistent with one-time psychological exposure costs or incorrect beliefs about the size of the fixed costs. Using a separate elicitation, we show that participants are also initially less likely to spontaneously consider working in an unfamiliar neighborhood and a single visit closes part of this gap. Our results suggest that exposure is an important component of urban mobility in cities with rapid unplanned growth like Nairobi.

*We are grateful to the Busara Center for Behavioral Economics, especially to Pauline Wanjeri, Jennifer Adhiambo, Jane Atieno, Joshua Omare, Grace Muniu and Alice Wambui. We thank Emily Breza, Fiery Cushman, Ed Glaeser, Rema Hanna, Elisa Macchi, Frank Schilbach, Matt Turner, Sendhil Mullainathan, 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: World Bank Development Research Group (ombonu@worldbank.org).

1 Introduction

Economists typically assume that individuals make decisions about whether to take advantage of opportunities by considering the quality of the opportunity and the cost of accessing it. However, evidence from psychology and marketing suggests that prior exposure to options may also shape choices both through preferences and the formation of consideration sets (Zajonc 1968; Bornstein 1989; Montoya et al. 2017; Goldstein and Gigerenzer 2002; Gigerenzer and Goldstein 2011; Allport 1954; Schank 1982).

In this paper, we study how prior exposure with urban neighborhoods shapes work location choices in Nairobi, and its effects on preferences and the formation of consideration sets. Spatial choices in a rapidly growing city are a useful domain to study the importance of exposure for two reasons. First, spatial choices govern how people access urban opportunities (jobs, schools, goods and services), and this is particularly relevant in low and middle-income countries where cities are key drivers of economic opportunity (Bryan, Glaeser, and Tsivanidis 2020).[1] Second, prior exposure may be particularly relevant for spatial choices. Being in a location is a vivid experience relative to other forms of exposure (e.g., seeing a product in a supermarket), and evidence from neuroscience and urban studies suggests that people build mental representations of space based on their past experiences (Lynch 1960; O'Keefe and Dostrovsky 1971; Fyhn et al. 2004).

We proceed in six steps. First, we document large gaps in individuals' exposure to nearby neighborhoods. Second, by making short-term job offers, we estimate the revealed cost of unfamiliarity and how it changes with a single exposure. Third, we use a benchmark model of a Bayesian expected utility maximizer to explore information and fixed costs as possible mechanisms and look for evidence in the data. Fourth, we test whether individuals anticipate how exposure will change their preferences. Fifth, we measure the persistence of the exposure effects after the experiment. Finally, we study how past and experimental exposure affect the formation of consideration sets when choosing where to work in an open-ended elicitation where subjects come up with a ranked list of possible work locations.

In the first step, we document that individuals in our setting have substantial gaps in their exposure to neighborhoods near their homes. In our sample of casual workers living in poor neighborhoods, the median participant commutes 7.8 km to work or to search for work; however, the median participant has never been to one in two of the neighborhoods within 75 minutes of where they live. They can recall a landmark for one in three such neighborhoods. This is not due to a lack of awareness; the median participant has heard of

[1] Our focus on the role of prior exposure complements work that focuses on how access to opportunities depends on quick, convenient, cheap, and safe transportation (Lall, Henderson, and Venables 2017; Borker 2024; Davis 2021; Tsivanidis 2023).

94% of the neighborhoods in this radius. The familiarity patterns also appear inconsistent with sorting based on neighborhood quality. Almost all neighborhoods have intermediate levels of average familiarity and match-specific observables have no predictive power.

In the second step, we quantify the revealed costs of working in an unfamiliar neighborhood by offering short-term employment opportunities in a real-stakes setting. Participants are offered jobs to measure air quality in different neighborhoods across Nairobi – a task that requires being physically present in a specific location and plausibly justifies assignments to a variety of locations.

To experimentally induce exposure, we train participants on how to complete the employment task in three neighborhoods over three days. Participants in the treatment group are trained in neighborhoods they report never having been to at baseline (unfamiliar neighborhoods), while participants in the control group are trained in familiar neighborhoods.[2]

To isolate the effects of exposure on preferences, separately from consideration, we present half of the participants with a series of job choices in the three days following training. Each day, these participants make seven binary choices between pairs of short-term job offers. The air pollution monitoring task is held fixed across all job offers, but we randomly vary job location, duration, and compensation. Each morning, after participants make their choices, we randomly select one of the offer pairs and assign the option the participant chose as their job for the day. This design allows us to quantify in a real-stakes setting how participants trade off familiarity with a location separately from distance, job compensation and other amenities.

Our first main result is that participants prefer working in familiar locations. Participants are willing to travel an additional 3.5 km of distance, work over an hour longer, or sacrifice 112 Ksh of salary (22% of the median daily wage) to avoid working in a neighborhood they have never visited. This effect is driven by a preference for neighborhoods that are the most familiar at baseline, defined as places where participants can name a landmark or report knowing how to get there. This *baseline familiarity premium* suggests that individuals may dislike visiting unfamiliar locations, but may also reflect unobserved, heterogeneous neighborhood preferences (sorting).

Our second main result is that experimentally inducing a single exposure during training to a previously unfamiliar neighborhood is sufficient to completely eliminate this familiarity premium. We find that above and beyond the main effect of training, having visited an unfamiliar neighborhood once is equivalent to increasing the wage by 109 Ksh or bringing

[2]We choose potential treatment and control neighborhoods for each individual. This provides us with a clear counterfactual training assignment for each participant and allows us to choose the neighborhoods to minimize potential spatial spillovers.

the neighborhood 3.4 km closer, fully offsetting the baseline familiarity premium.[3]

In our third step, we consider what might drive this effect of exposure and the baseline familiarity premium. We begin with a benchmark model of a Bayesian agent with correct priors who maximizes expected utility. Under these assumptions, we can rule out information and sorting as explanations for our results, and rule in a fixed cost of exposure. We model exposure as both revealing information about the utility of working in a neighborhood and incurring a fixed cost.[4] We first show that exposure to an initially unfamiliar neighborhood on average changes the agent's utility to work there by the value of the fixed cost. This follows from the law of iterated expectations. Absent fixed costs, the expected utility elicited ex-ante is the same as the average utility elicited after exposure, when some or all of the uncertainty is resolved. This relies on correct priors and risk-neutrality *over utilities*, and it means that uncertainty over unfamiliar neighborhood characteristics does not imply an exposure effect in this model.[5] We then show that the average baseline familiarity premium equals a "sorting" term plus the fixed cost.

Our experimental result that the familiarity premium and exposure effects are equal rules out sorting, consistent with the descriptives above showing seemingly idiosyncratic patterns of past exposure. That is, participants on average do not appear to sort into becoming familiar with neighborhoods that they are more willing to work in. We next consider what fixed costs, "real" or psychological, could explain our results before turning to the potential role of non-standard beliefs.

First, we find little evidence of first-time *navigation* costs. We show that all else equal, participants are willing to work in a familiar neighborhood over an unfamiliar one even if it means traveling 3.4 km farther or working more than an hour longer. Thus, in order for navigation costs to explain the premium, we would expect travel times to differ by similar magnitudes. Instead, using a Heckman selection correction model with the other randomized job attributes as instruments, we find small differences in travel times based on baseline familiarity and no effect of the exposure treatment on travel time.

Second, we consider whether the fixed cost may be due to risks associated with first-time exploration. Two pieces of evidence suggest it is unlikely to account for our findings. First, the experiment entails minimal risk: employment and compensation are guaranteed,

[3]To estimate the main effect of training in a neighborhood, we use the fact that control participants are trained in baseline familiar neighborhoods, which we compare to a different set of familiar neighborhoods not visited during training.

[4]We make no additional restrictions of how familiarity emerges in the first place. For example, our model allows that agents choose which locations to become familiar with based on noisy signals.

[5]We take this result to the data in a specification without neighborhood fixed effects, so as not to condition on the realized (true) value of the utility for a given neighborhood. Our results are very similar with or without neighborhood fixed effects, consistent with an exposure effect driven by a fixed cost.

and participants spend only one to two hours in the assigned neighborhood during daylight hours. Second, we do not observe any adverse events in our sample, suggesting that if risks do exist, they are rare. While we cannot rule this out completely, we conduct a back of the envelope calculation showing that for a risk rare enough to not occur in our sample, and still explain the familiarity premium, its associated cost needs to be equivalent to at least three months of wage income. Finally, in order to be consistent with the premium disappearing after one visit, any kind of exploration risk would have to fully disappear after one visit. Thus, for exploration risk to explain the result it would have to be rare, severe and fleeting.

While we do not find evidence that actual exploration risk or navigation costs fully explain the familiarity premium, incorrect beliefs about these quantities could play a role. Our findings are also consistent with psychological fixed costs like processing fluency (Jacoby and Dallas 1981; Jacoby and Whitehouse 1989; Winkielman et al. 2003). In this account, exposure to a neighborhood helps individuals learn how to process the neighborhood's stimuli, leading to increased processing fluency. This increase in fluency may then improve individuals' willingness to work in a neighborhood.

In our fourth step, we consider whether deviations from correct beliefs about the experience of working in a neighborhood can explain our results. For example, if individuals hold overly pessimistic priors, they will update positively after exposure. This mechanism is closely related to the original affective conditioning explanation for the mere exposure effect (Zajonc 1968). We find some evidence for this channel: participants' beliefs about unfamiliar neighborhoods are more pessimistic before exposure but converge with those for familiar neighborhoods afterwards.

To explore this further, we assess whether individuals change their choices over future work locations after learning they will visit an unfamiliar neighborhood in the interim. This timing of information revelation helps us distinguish between participants who hold inaccurate beliefs and those who anticipate a fixed cost of exposure. If results are driven by incorrect beliefs, the announcement of future information should not change choices because it does not affect information at the moment when the choice is made. In contrast, if fixed costs are the mechanism shaping choice, agents should alter their choices. The training location announcement informs agents that, by the time the employment day arrives, they will have already paid any fixed costs. Thus, before learning the training location, agent's should factor these fixed costs into their decisions, but after the announcement they should ignore these fixed costs as sunk.

To test this, we use the same binary job choice questions during training days to elicit preferences over jobs three days later. We do this after they have been informed of where they will train on that day but before they have actually visited that location. We find that

individuals partially anticipate the benefits of exposure. On average, subjects make choices as if the treatment effect will be approximately 64% as large as what we observe, with a 95% confidence interval of [0.3.1.0]. Individuals partly anticipate a fixed cost of exposure, and this suggests that incorrect priors are unlikely to explain the entire effect.[6]

In our fifth step, we find some evidence that exposure to an initially unfamiliar neighborhood has persistent effects on individuals' willingness to return. We measure persistence in three ways: cross-sectionally, with a different job offer, and by observing travel after the experiment. First, we cross-sectionally investigate the persistence by separately estimating the impact of familiarity on choices based on the time that has passed since the respondent last visited the location. We find that the effect of familiarity is relatively stable for locations visited within the last three years. Second, 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 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 imprecisely estimated because of ceiling effects at high wages.[7] These results suggest that exposure to a baseline unfamiliar neighborhood may have a persistent effect in the medium run. We also use an endline survey and SMS data to measure where participants travel outside of our study. 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 argument that participants found exposure to these neighborhoods valuable.

In our sixth and final step, we consider how prior exposure affects the consideration of locations as employment options. If individuals form mental representations of their urban surroundings based on their past exposures (Lynch 1960; O'Keefe and Dostrovsky 1971;

[6]Another explanation is that workers expect that if they learn during training that the neighborhood is a worse place to work in than their outside option, they can later opt out of showing up for work. To test this possibility we study whether the exposure effect is sometimes negative. Negative beliefs updating is very rare, and using the method in Chernozhukov et al. 2023 we can reject moderate negative treatment effects of exposure.

[7]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, including distance, predicts show-up. When the ceiling effect on show-up was detected during field work, we lowered the entire wage distribution. For completeness, we present the full set of results, but focus on the wage levels where job offer attributes affect decisions and the rounds with the reduced wage offer distribution.

Fyhn et al. 2004; Schank 1982; Simon and Gilmartin 1973; Mack and Palmeri 2024; M. K. Ho et al. 2022; Thalmann, Souza, and Oberauer 2019; Bordalo et al. 2025), they may be less likely to include unfamiliar locations in their consideration sets when deciding where to travel.

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 the surveyor 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.[8]

Revealed preference estimates based on this ranked data show a larger role for baseline familiarity than in the "closed" elicitation, which we interpret as evidence that beyond its effect on preferences, familiarity affects consideration. The effect of a single exposure to a location partially undoes the open elicitation familiarity premium. We also estimate a model with both preferences and consideration costs, and find that baseline familiarity and the training significantly affect consideration costs. Our results show that people struggle to consider unfamiliar neighborhoods as potential places to work.

Our paper builds on several areas of psychology. First, we build on an extensive literature studying the mere exposure effect and its underlying mechanisms (Zajonc 1968; Bornstein 1989; Montoya et al. 2017). This literature predominantly focuses on repeated and controlled exposure for short durations (including subliminal exposures) to a variety of specific stimuli in the laboratory. We extend this literature by focusing on a longer duration, single exposures to holistic environments. We also relate to literature demonstrating an aversion to unfamiliar individuals, and that individuals can fail to anticipate how much they will enjoy conversing with strangers (Allport 1954; Epley and Schroeder 2014; Schroeder, Lyons, and Epley 2022). We extend this literature by examining how prior exposure affects spatial preferences. Finally, we also draw inspiration from a diverse literature on how past experience can shape mental representations and how these representations can be used in planning and decision making (Schank 1982; Simon and Gilmartin 1973; Mack and Palmeri 2024; M. K. Ho et al. 2022; Thalmann, Souza, and Oberauer 2019; Bordalo et al. 2025). We extend this literature by focusing on a naturalistic, high-stakes task and demonstrating the separate contributions of experience through preferences and formation of consideration sets.

Our paper also builds on a diverse literature in economics and finance studying the importance of experience and experimentation for learning (Gittins 1979; Scharfstein and Stein

[8]We view this method as a bridge between methods that elicit preferences from explicit choice sets and methods that measure what is top of mind more directly Haaland et al. 2024.

1990; A. V. Banerjee 1992; Eyster and Rabin 2010; Hanna, Mullainathan, and Schwartzstein 2014; Malmendier 2021; Foster and Rosenzweig 1995; Weitzman 1979) by demonstrating an additional effect of exposure in addition to potential scope for resolving uncertainty.[9] Our work is also related to a large literature in marketing that studies the mechanisms and effects of advertising (e.g. Shapiro 2018; Shapiro, Hitsch, and Tuchman 2021; Keller 1987; Milgrom and Roberts 1986; Kihlstrom and Riordan 1984; Nelson 1970; Nelson 1974; Sahni and Nair 2020; Ackerberg 2003; Ackerberg 2001).

Our paper has implications to realizing the benefits of urban agglomeration in developing countries (Duranton 2015; Chauvin et al. 2017; Lall, Henderson, and Venables 2017; Bryan, Glaeser, and Tsivanidis 2020), complementing previous work that measures and unpacks transportation costs in these contexts for different groups (Borker 2021; Vitali 2024; Kreindler et al. 2023; Grosset-Touba 2024; Tang 2024; Jalota and L. Ho 2024; Cook, Currier, and Glaeser 2024), and research on policies that reduce transportation costs (Habyarimana and Jack 2015; Gonzalez-Navarro and Turner 2018; Tsivanidis 2023; Balboni et al. 2021; Zárate 2022; A. Banerjee and Sequeira 2023; Franklin 2018; Abebe et al. 2021; Field and Vyborny 2022). Policies to encourage exposure may be additional tools, beyond reducing access costs, to increasing effective access to opportunities.

The exposure effect we demonstrate here offers a possible micro-foundation for several results showing people are reluctant to explore new locations or travel routes and persistence when people are exposed to such locations (Bryan, Chowdhury, and Mobarak 2014; Larcom, Rauch, and Willems 2017; Okunogbe 2024; McKenzie 2024). One particularly relevant example is Brady and Mcnulty (2011) which studies the effect of consolidated polling places in Los Angeles during the 2003 gubernatorial recall election. The authors use a matching design to show that being assigned a new polling place increases the probability of not voting by approximately 2% regardless of the change in travel time to the new polling place. Our paper is particularly relevant for recent work that proposes aversion to rural to urban migration falls with sustained experience in the city (Lagakos, Mobarak, and Waugh 2023). The exposure effect is also an additional type of spatial friction, complementing work that studies information frictions in commuting, migration and trade (Vitali 2024; Baseler 2023; Porcher 2022; Wiseman 2023; Porcher, Morales, and Fujiwara 2024). Beyond spatial domains, our results may also be relevant for other decisions such as occupation choice and technology adoption.

[9]Biased beliefs about neighborhood quality may explain some of our results, but the anticipation results mentioned above exclude the possibility that they are the entire explanation.

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 based on a canvasing exercise, 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 canvased 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.1 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.2). Occupations for men are more heterogeneous and typically include 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 61 neighborhoods. Most of our analysis uses the main neighborhoods.

For each home neighborhood (Kibera, Kawangware, and Viwandani), we elicit participant familiarity with neighborhoods that 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:[10]

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?

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 ever 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

[10]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.

list a landmark for the neighborhood, regardless of whether they have ever been there. The median respondent only has ever been and knows how to reach 43% of neighborhoods and can name a landmark for only 34% (around 13 and 10 neighborhoods, respectively). The lower levels of familiarity for these more demanding definitions of familiarity are important because we later show in Section 5.1 that these measures affect participants' willingness to work in a neighborhood.

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 cell 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 is suggestive that familiarity patterns are not driven by consensus views of neighborhood attributes.

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) essentially do nothing to improve our accuracy.[11] These

[11]We also analyze whether Nairobi residents partially sort to destination neighborhoods based on ethnicity, and whether this explains the familiarity patterns we document here. 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

results suggest that, as far as we can observe, individual-neighborhood match characteristics are not important drivers of exploration.

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.

4 Experimental Overview

In order to experimentally identify the effect of exposure, we want to cause individuals to travel to unfamiliar locations without 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.[12] 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.

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.

[12]Neither type of sensor displays the pollution level, which rules out participants learning about neighborhood pollution levels from the task.

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 at the study venue in the participant's home neighborhood. On the first day, participants respond to 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 and show up the next day form the analysis sample for the experiment.

We randomize participants who finish the second survey 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). Tables A.7 and A.8 report randomization balance comparisons.

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. The training takes place in the target neighborhood, and participants visit 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 5.3. Two to four months after the training, they receive invitations for a separate survey in various neighborhoods. We analyze their show-up decisions to these invitations in section 5.3.

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 (e.g. accidentally inducing a visit to an unfamiliar neighborhood while traveling to a familiar one), 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.[13]

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.[14] 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

[13]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.

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

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 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 participants 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 "closed" 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, "open", 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 final elicitation we use is design to assess whether individuals anticipate how their utility of visiting a neighborhood will change after having visited it once. We use this variation later to assess the degree to which biased beliefs about neighborhood quality can explain our results. To study this we ask participants in the closed elicitation 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 anticipate the change in their utility, we would expect these participants to be more willing to work in the unfamiliar neighborhoods before having actually visited.

5 Impact of Neighborhood Exposure: 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 "closed" 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.[15]

Participant responses to these real-stakes questions allow us to estimate preferences for working in different types of neighborhoods, the effects of the training, and the impact of inducing 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, 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.[16]

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 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.

[15]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.

[16]We also allow participants to turn down both jobs. Due to the underemployed sample and the competitive wages, this was extremely rare. We study extensive-margin decisions for work opportunities in section 5.3.

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 is designed 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 effect of exposure (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 $j \in \{1, 2\}$ in target neighborhood n, for individual i is given by

$$u_{ijn} = \underbrace{\beta^F Familiar_{in}}_{\substack{\text{Baseline} \\ \text{familiarity} \\ \text{premium}}} + \underbrace{\beta^V Train_{in}}_{\substack{\text{Effect of} \\ \text{training} \\ \text{in} \\ \text{familiar}}} + \underbrace{\beta^U Train_{in} \times Unfam_{in}}_{\text{Effect of exposure}} + \underbrace{\beta^D d_{in}}_{\text{Distance}} + \underbrace{\gamma X_{ijn}}_{\substack{\text{job} \\ \text{attributes} \\ \text{(e.g.} \\ \text{wage)}}} + \underbrace{\epsilon_{ijn}}_{EV(1)} \quad (1)$$

where $Familiar_{in}$ is an indicator for i being familiar with neighborhood n at baseline, d_{in} is distance from the study venue to n, and the set of randomly allocated job covariates X_{in} 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. The term $Unfam_{in}$ is defined $(1 - Familiar_{in})$, so it is switched on only for treatment group participants and the (baseline unfamiliar) neighborhoods where they are trained in. Hence, β^V gives the effect of training while β^U provides the additional effect of having been trained in an unfamiliar location.

Benchmark Model. In order to structure the interpretation of these coefficients we now present a benchmark model for how a Bayesian expected utility maximizer's choices would be affected by a one-visit exposure.

We consider an expected-utility maximizer agent i. For each agent, when they enter the experiment, the set of all neighborhoods $\mathcal{N}$ is partitioned into baseline unfamiliar and baseline familiar such that $F_i \cup U_i = \mathcal{N}$. We assume that the agent i receives utility u_{in} from working in neighborhood n. We assume that if the neighborhood is familiar, that the agent knows the utility they would receive. We further assume they have a correct prior p_i over

u_{in} for unfamiliar neighborhoods.[17]

We further assume that exposure (training) in an unfamiliar neighborhood reveals information about u_{in} and that the agent forms a Bayesian posterior q_{in}. We also allow for a fixed cost c_i during first time visiting a new neighborhood.

Thus, in the experiment for a given individual i in location n, our choice experiment elicits the following valuation

$$Y_{in} = \mathbb{1}(n \in U_i)[D_{in}\mathbb{E}_{q_{in}}[u_{in}] + (1 - D_{in})(\mathbb{E}_{p_i}[u_{in}] - c_i)] + \mathbb{1}(n \in F_i)u_{in}$$

Where D_{in} is an indicator that person i was induced to visit neighborhood n by the experiment.

Through the independence of treatment assignment and repeated application of the law of iterated expectation it follows that our baseline familiarity premium coefficient identifies

$$\begin{aligned}
\beta^F &= \mathbb{E}_{(i,n)}[Y_{in}|n \in F_i, D_{in} = 0] - \mathbb{E}_{(i,n)}[Y_{in}|n \in U_i, D_{in} = 0] \\
&= \mathbb{E}_i[\mathbb{E}[u_{in}|i, n \in F_i]] - \mathbb{E}_i[\mathbb{E}_{p_i}[u_{in}|i, n \in U_i] - c_i] \\
&= \underbrace{\mathbb{E}_i[\mathbb{E}[u_{in}|i, n \in F_i] - \mathbb{E}_{p_i}[u_{in}|i, n \in U_i]]}_{\text{Sorting}} + \underbrace{\mathbb{E}_i[c_i]}_{\text{Fixed Costs}}
\end{aligned}$$

That is, the baseline familiarity premium identifies the difference in expected utilities between familiar and unfamiliar neighborhoods and the fixed costs of visiting the unfamiliar neighborhoods. This then implies that the effect of exposure is the average fixed cost:

$$\begin{aligned}
\beta^U &= \beta^F - (\mathbb{E}_{(i,n)}[Y_{in}|n \in F_i, D_{in} = 1] - \mathbb{E}_{(i,n)}[Y_{in}|n \in U_i, D_{in} = 1]) \\
&= \beta^F - (\mathbb{E}_i[\mathbb{E}[u_{in}|i, n \in F_i]] - \mathbb{E}_i[\mathbb{E}[\mathbb{E}_{q_{in}}[u_{in}]|i, n \in U_i]]) \\
&= \mathbb{E}_i[c_i]
\end{aligned}$$

Where the last line follows because the assumed accuracy of priors over neighborhoods implies that $\mathbb{E}[\mathbb{E}_{q_{in}}[u_{in}]|i, n \in U_i] = \mathbb{E}_{p_i}[u_{in}|i, n \in U_i]$.

In this model, uncertainty about the utility of unfamiliar neighborhoods does not contribute to the familiarity premium. This happens because we work directly with the utility of working in a given neighborhood and, by definition, expected utility maximizing agents are risk-neutral over utilities.

[17] We do not impose other restrictions on how the sets U_i and F_i emerged. For example, our model covers a situation where previously, the agent decided which neighborhoods to visit based on noisy signals about these neighborhoods. We can also extend the model to allow the prior to be conditional on known characteristics X_n of the neighborhood.

Results. 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. The sample includes nine target neighborhoods for each participant, the six main familiar and unfamiliar neighborhoods, and the three other familiar neighborhoods.

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 identified 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 first column.

Baseline familiarity with a neighborhood has a large positive impact on the utility of working there (0.85). The magnitude is equivalent to 112 Kenyan shillings (Ksh) of additional compensation, which we obtain by dividing the baseline familiarity coefficient by the expected compensation coefficient (0.76). This *baseline familiarity premium* is a large amount given that the median daily wage outside our the experiment is 500 Ksh. We can also benchmark the cross-sectional familiarity premium with respect to distance and job duration. 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 3.54 kilometers closer or having to work for more than an hour less.

These effects suggest that familiarity increases an individual's willingness to work in a neighborhood; however, because baseline familiarity is not randomly assigned, these results could also be driven by unobserved preference heterogeneity. To asses this possibility, we now turn to the exposure experimentally induced by training.

Our first main result is that the single exposure during training is sufficient to erase this familiarity premium. Training in an unfamiliar location increases the utility of working there (0.83), equivalent to increasing the wage by 109 Ksh or bringing the neighborhood 3.46 km closer. This effect is in addition to the general effect of training in a neighborhood, which is also positive but smaller (0.29), and which applies both to familiar and unfamiliar neighborhoods.

We find similar effects of familiarity, both cross-sectional and experimental, depending on whether the neighborhood is close or far to the participant's home neighborhood (Table A.12). We find that familiarity matters more for the male participants in our sample, even after adjusting for a slightly higher effect of job duration and for wages, but that the single exposure erases the premium for both (Table A.13).

How do different threshold for familiarity affect these results? In Table A.14, we separate the $Familiar_{in}$ measure in equation (1) based on whether the participant reported at baseline knowing how to get there, or not, and based on whether they were able to mention a landmark in the neighborhood. For both definitions, the familiarity premium is larger for more familiar neighborhoods.

Do participants trained in unfamiliar neighborhoods become more willing to work in other unfamiliar neighborhoods? To study this, we add to the estimation sample the tenth target neighborhood, which is unfamiliar and never included in the training, and estimate a separate coefficient for this neighborhood interacted with treatment. Table A.15 shows the results. A positive coefficient indicates positive spillovers to other unfamiliar neighborhoods. In column 3, we find a positive yet imprecisely estimated coefficient of 0.27 with a standard error of 0.18.

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 "closed" 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.

5.2 Mechanisms

In this section we consider what mechanisms may drive this effect of exposure. Under the benchmark model these results imply that fixed costs must explain the effect of exposure that we observe. We begin by considering two sources of potential "real" fixed costs and alternative psychological explanations. We then turn to whether deviations from the benchmark beliefs model can explain our results.

First-time Navigation Costs. We find little evidence that first-time travel costs are large enough to explain the familiarity premium. Individuals are willing to travel an additional 3.5 km and work for more than an additional hour in order to work in a familiar neighborhood. Thus, for first-time navigation costs to explain the premium, we would expect travel times to increase by a similar amount. This does not appear to be the case.

In Table 6 we regress the travel time it takes participants to reach a neighborhood where

they work, on baseline familiarity and experimentally induced visits using a Heckman (1976) two-step probit model to control for selection into working in an unfamiliar neighborhood where we use the other randomly allocated job attributes, like compensation, as instruments.

It takes participants slightly longer to reach neighborhoods where they have been before, but we can reject an increase of greater than 8 minutes in both directions. Training has a small effect on travel times, but we do not find any additional effect for training in unfamiliar neighborhoods. Table 6 also includes self-reported measures of navigation such as getting lost and asking for directions. Only 6% of respondents report getting lost during the trips and this is not differential by the familiarity status of the neighborhood. We do find a small decrease in stating that the trip was "more difficult than expected" after training. But overall, we find little support that first-time navigation costs could explain the large familiarity premium documented earlier.

Exploration Risk. Another potential source of fixed-costs is if visiting a new neighborhood for the first time is risky. 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. We also do not observe any participants in our experiment encountering adverse experiences. Thus, for risk aversion to explain the premium we estimate, it must be a rare, non-employment risk that can occur in a short time period and that vanishes after one visit. The results in the previous section already speak to this issue by showing that the risk of getting lost along the way — a plausible proxy for more significant risks — is not statistically higher when traveling to unfamiliar neighborhoods.

One concern might be risks that are sufficiently rare that we would not observe them in our experiment. To assess this possibility, we can do a back of the envelope calculation of how bad the event would have to be in order to justify the premium given the low probability. Specifically, suppose every time the agent explores there's a probability p_u of a bad outcome such that

$$p_u[u(good) - u(bad)] = 109$$

We observe 243 trips to unfamiliar locations on employment days among the control group with zero adverse outcomes. That means that we can reject, with 95% confidence, that $p_u > 1.2\%$. In order to justify a premium of 109 Ksh, this would imply that the difference in utilities between the good and bad events must be at least 9,083 Ksh or three months of wage income for our respondents.

Psychological Fixed Costs. While we believe the evidence suggests there are not real exploration risks or navigation costs experienced by the participants, it is possible that participants hold incorrect beliefs about these quantities. For example, agents might be overly pessimistic about how long it will take them to reach a location for the first time. This could then be consistent with the rapid updating either if individuals learn their beliefs were incorrect or only hold the distorted beliefs regarding first-time navigation costs.

Our results are also consistent with psychological fixed costs like processing fluency (Jacoby and Dallas 1981; Jacoby and Whitehouse 1989; Winkielman et al. 2003). In this account, when exposed to a neighborhood, individuals learn how to process the neighborhood's stimuli leading to increased processing fluency. This increase in fluency then may improve individuals' willingness to work in a neighborhood.

Non-standard Beliefs We next consider whether deviations from correct beliefs can explain our results. For example, if individuals hold overly pessimistic priors they will update positively after exposure. This mechanism is closely related to the original affective conditioning explanation for the mere exposure effect (Zajonc 1968).

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.

Table 3 shows the results. At baseline, individuals are on average more pessimistic about unfamiliar neighborhoods, as indicated by the negative $Pre_t \times Treated_i$ coefficients. For example, respondents rate unfamiliar neighborhoods 0.24 SD lower on their likelihood of finding a causal job in the neighborhood, and 0.25 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 0.44 SD lower, while the neighborhood's safety is 0.17 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 at the 5% threshold. The probability of others finding a job remains significant at the 10% level, but the coefficient still shrinks substantially. This means that on average, a single in-person visit eliminates the initial imbalance in ratings about these neighborhoods.

We examine whether some participants update very negatively about the neighborhoods that they visit and 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 there are almost no participants who update downward significantly.

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 use a split-sample approach to estimate the mean pre-belief for each neighborhood among those who were familiar prior to the experiment. Specifically, we randomly split the sample into 10 groups and for each group we predict familiar prior beliefs using the remaining 90% of the data. We then regress post beliefs for other participants 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, $y_{in} \in \{1, 2, 3, 4, 5\}$, 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 steps 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 the held out 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. For those unfamiliar with the neighborhood, the results are more variable, based on the outcome, while for four out of five outcomes we cannot reject a coefficient equal to 1. These results show that participants have agreement over how different neighborhoods compare on the dimensions in the beliefs data, and this also holds after a single visit to an unfamiliar neighborhood.

Do Subjects Anticipate How Their Preferences Change After Exposure? To explore the importance of biased beliefs as a mechanism further we assess whether individuals change their choices for future exployment after learning they will visit a neighborhood in the interim. Individuals who want to have accurate beliefs should not change their actions in response to learning they will soon acquire more information, even if they suffer from having

incorrect priors.

To do so, in the "closed" elicitation arm we elicited choices during the three training days. On training day $d \in \{1, 2, 3\}$, surveyors asked participants about their employment preferences for employment day $d + 3$ – that is, for employment three days in the future. Crucially, each respondent received these questions *after* learning 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 later received on the corresponding employment days $d+3$. 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 and report how much a participant who has already visited a neighborhood changes their willingness to work there, compared to the effect of only 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

$$\begin{aligned} u_{ijn} = & \beta^F Familiar_{in} + \\ & \underbrace{\beta^V Train_{in}}_{\text{training effect}} + \underbrace{\beta^U Train_{in} \times Unfamiliar_{in}}_{\text{exposure effect}} + \\ & \underbrace{\beta^{AV} AnticipTrain_{in}}_{\substack{\text{anticipation} \\ \text{training effect}}} + \underbrace{\beta^{AU} AnticipTrain \times Unfamiliar_{in}}_{\text{anticipation exposure effect}} + \\ & \underbrace{\gamma X_{in}}_{\substack{\text{randomized} \\ \text{job attributes}}} + \beta^D d_{in} + \epsilon_{ijn} \end{aligned} \tag{2}$$

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 equation (2) using a binary logit model. Table 7 shows the results. We use the pooled data on training days and employment days. We find $\beta^U = 0.767$ and $\beta^{AU} = 0.483$. 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 0.63 and the p-value for equality is 0.026. The fact that subjects partially anticipate being more willing to return after visiting, suggests that non-standard beliefs like incorrect priors cannot fully explain our results. These results are consistent with subjects who anticipate a fixed cost of exploration.

The training announcement makes them realize that this cost will be incurred during the training.

In column 3, 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 3, we leverage variation in realized familiarity induced by training on previous days. We find quite similar results with a ratio of $\beta^{AU}/\beta^{U} = 0.51$. In column 4 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}.

5.3 Persistence

We next consider whether the increased willingness to visit an unfamiliar neighborhood induced by exposure persists. We analyze persistence using a cross-sectional analysis, and two medium-term follow-up exercises, one experimental, and the other based on survey data on visits.

Cross-sectional persistence. 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.18 shows that the familiarity premium is relatively stable up to neighborhoods that the respondent visited within the last 3 years. While cross-sectional, this pattern suggests that the effects of exposure are likely to persist.

Show-Up to Medium-run Task Invitation. We next examine whether participants remain more likely to return to neighborhoods they visited during the experiment when invited back. Approximately 2-4 months after the intervention, we organized new work opportunities in different neighborhoods and invited study participants to attend. Each invitation was to a short (5-minute) survey on commuting in Nairobi, scheduled to take place in a given neighborhood two days later. Because the survey focused on commuting, inviting participants to different neighborhoods did not appear artificial.

Invitations were issued in two rounds, with participants receiving three invitations per round, spread over about 10 days. The survey neighborhood, and the offered wages were randomized.[18] Surveyors recorded whether the respondent showed up on the designated survey day. After initially launching these invitations with wages randomized between 500 and 1,000 KSH we observed show up rates in excess of 80%. To mitigate ceiling effects, we then reduced the distribution of wages, eventually settling on a distribution between 100 and 400 KSH.

We estimate equation (1) using these show-up data in a binary logit model, including fixed effects for the survey date, randomization strata, and invitation sequence (i.e., whether this was the participant's 1st, 2nd, etc., successful invitation). To deal with ceiling effects we take two approaches. First, we split the results by whether the wage was above 500 KSH or below using the entire sample. This is because for wages higher than 500KSH, show-up plateaus around 80% (Figure A.11) and none of the neighborhood or job characteristics, including distance and wage, affect show-up. Second, we present the results after restricting the sample to the rounds of invites where wages were randomized between 100 and 400 KSH.

Table 4 reports the results. Results are similar for both wages below 500 KSH and the rounds after wages were reduced. We see that baseline familiarity strongly predicts show-up for the survey. The experimental effect of training in an unfamiliar neighborhood is positive and around two thirds of the baseline effect, albeit insignificant. Unfortunately, the reduced effective sample size due to the ceiling effects significantly limits our power. With our estimated standard errors, the minimum detectable effect is larger than the baseline familiarity premium. Distance and wages also matter for show-up in these subsamples. These logit coefficients are across the board somewhat smaller than those estimated in the main experiment (see Table 2), which is consistent with a higher dispersion of idiosyncratic shocks for this extensive margin decision.

Do participants return to visited neighborhoods? Finally, we examine whether individuals return to the neighborhoods on their own as reported through SMS and phone surveys.

Table 5 reports the results. The outcomes in columns 1-4 are from directly asking respondents about whether they have visited the target neighborhoods in the past two weeks, 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

[18]Neighborhoods were randomly selected from the participants' nine target neighborhoods (three target familiar, three target unfamiliar, and three non-target familiar), and the invitation order of neighborhood type was stratified across participants, separately for the two rounds. Neighborhoods used in the first round were excluded from the second round.

(in their own words). We then code their response to our neighborhood list.[19]

We find that respondents return to initially unfamiliar neighborhoods when they are trained there. This effect is captured by the sum of the "Visited Any" and "Visited Unfamiliar" coefficients, which is positive and significant for all columns except the third.[20]

Participants 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.[21] Furthermore, participants did search for work in these neighborhoods (column 2). Such outcomes may take time to realize.

Individuals do revisit the neighborhoods that they visited during training, including those that were unfamiliar at baseline.

6 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 exposure may have on how likely people are to consider a neighborhood as an option, to begin with. In this section, we study 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 begin to report an ordered list of neighborhoods where they would be willing to work in, and the surveyor then records the neighborhoods

[19] 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 5, 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.

[20] Unlike our previous results, the coefficient on "Visited Unfamiliar" is smaller than that on "Baseline Familiar," which means that while participants return to these places, they do not treat them exactly the same as other familiar neighborhoods.

[21] We can only reject a +0.025 effect on the probability of working in a specific visited unfamiliar neighborhood, relative to a mean of 0.069.

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, respondent choices reflect a mix of preferences and consideration. 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 that 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 compare our previous estimates of a binary logit model on the binary choice elicitation data with estimates from a multinomial logit model on the top choice on the first day of the open elicitation ranked list. These estimates should be identical if consideration has no impact on choices and if preferences over working in different neighborhoods follow a multinomial logit model.

We find that familiarity and experimental exposure affect choices more in the open elicitation than in they do for binary choices. The left side of Figure 8 shows preferences that we estimated previously using the binary choice data (equation (1)). Neighborhood familiarity is valued as much as the neighborhood being 3.4 km closer, and the impact of exposure is equivalent to bringing the neighborhood 3.6 km closer. Focusing on the top choice in the open elicitation, shown on the right side of the same figures, participants rank a familiar neighborhood as if it is 7.8 km closer, and this is statistically higher than for binary choices. The impact of exposure implied by their ranking choices is equivalent to bringing the neighborhood 5.5 km closer, although we cannot reject equality with binary choices at conventional levels. Table A.22 shows the full multinomial logit estimates for the k-th choice, $k = 1, \ldots, 5$, which reveal that these differences between binary choices and the top choice on the first day in open elicitation are driven by higher coefficients for familiarity and exposure, rather than differences for distance.

6.1 A Two-Self Model with Memory Costs

We next 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. The memory self 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

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

for each neighborhood n where $\mathbf{X}_{in}$ is a vector of neighborhood characteristics, including baseline familiarity and training indicators, and a constant term. We use exactly the same specification as in equation (1) except that we use more compact notation, and we also include a constant term. We assume that ε_{in} has variance equal to 1, and we normalize the utility of not working to $u_{i0} = \varepsilon_{i0}$.

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

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

where ν_{ijk} are iid shocks with variance $= \sigma_\nu \geq 0$. (Again, $\mathbf{X}_{in}$ includes a constant term.) We normalize the memory cost associated with stopping, that is, not ranking any other neighborhoods, to $c_{i0k} = \nu_{i0k}$.

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 objective 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[22]

$$j_k^* \in \arg\max_j \lambda_k u_{in} + c_{ijk} \quad \text{or "stop" (assumed net utility } \lambda_k \varepsilon_{i0} + \nu_{i0k})$$

Estimation and Results. We estimate this model using maximum likelihood based on the first K options ranked by the participant (or fewer if they stopped earlier), varying $K \in \{5, 10, 15, 20\}$. We fix the preference parameters over baseline familiarity, distance, training, and training in an unfamiliar location, to the values we estimated using binary

[22]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.

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 baseline familiar neighborhoods are easier to remember, with a memory cost coefficient of 2.53. Having trained in a neighborhood has a large effect on memory costs, with a coefficient of 2.02, 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 1.08 and a bootstrapped 95% confidence interval of $[0.37, 2.13]$. Distance also matters (-0.4 for memory costs compared to -0.25 for preferences). We estimate a relatively high memory cost dispersion $\sigma \approx 2$, and a subjective probability λ of around half that a given neighborhood has a job on a given day (compared to the true $\lambda = 0.25$). The relative magnitudes of the first four coefficients are broadly stable to changing the specification, the number of ranked choices used in estimation, and the day when choices are elicited (Tables A.23 and A.24).

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

7 Discussion

When trying to understand whether individuals will take advantage of opportunities economists typically focus on quality and access costs. This is true of the literature studying the rapidly growing cities in low- and middle-income countries, where research has focused on the role of infrastructure and safety improvements in mediating access to opportunities.

In a sample of 799 casual workers in Nairobi, we've shown that exposure is also a key driver of choice and that past exposure can be partial. We then showed that being experimentally induced to visit a neighborhood once is equivalent to bringing it 3.45 km closer or to increasing the wage by 109 Ksh (22% of the median daily wage), and this is equivalent to the full cross-sectional premium. We showed that this effect is most consistent with psychological fixed costs or incorrect beliefs.

Because one visit is sufficient to erase the premium and these effects appear to persist, policies to encourage exposure 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. the grid structure of New York City) while for others it is difficult (e.g. Boston). This may be particularly useful if the

psychological fixed costs are driven by processing fluency, and rapid growth in low and middle income contexts may lead to cities that are particularly difficult to "read."

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.

Ackerberg, Daniel A. (2001). "Empirically Distinguishing Informative and Prestige Effects of Advertising". In: *The RAND Journal of Economics* 32.2. Publisher: [RAND Corporation, Wiley], pp. 316–333. ISSN: 0741-6261. DOI: 10.2307/2696412. URL: https://www.jstor.org/stable/2696412 (visited on 05/25/2025).

— (2003). "Advertising, learning, and consumer choice in experience good markets: an empirical examination". In: *International Economic Review* 44.3. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-2354.t01-2-00098, pp. 1007–1040. ISSN: 1468-2354. DOI: 10.1111/1468-2354.t01-2-00098. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-2354.t01-2-00098 (visited on 05/25/2025).

Allport, Gordon W (1954). "The nature of prejudice". In: *Reading/Addison-Wesley*.

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.

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.

Bordalo, Pedro et al. (Feb. 2025). *A Cognitive Theory of Reasoning and Choice*. DOI: 10.3386/w33466. URL: https://www.nber.org/papers/w33466 (visited on 06/09/2025).

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

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

Bornstein, Robert F. (Sept. 1989). "Exposure and affect: Overview and meta-analysis of research, 1968–1987". In: *Psychological Bulletin* 106.2. Publisher: American Psychological Association, pp. 265–289. ISSN: 0033-2909. DOI: 10.1037/0033-2909.106.2.265. URL: http://proxy.uchicago.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=pdh&AN=1990-00422-001&site=ehost-live&scope=site (visited on 06/09/2025).

Brady, Henry E. and John E. Mcnulty (Feb. 2011). "Turning Out to Vote: The Costs of Finding and Getting to the Polling Place". In: *American Political Science Review* 105.1, pp. 115–134. ISSN: 1537-5943, 0003-0554. DOI: 10.1017/S0003055410000596. URL: https://www.cambridge.org/core/journals/american-political-science-review/article/turning-out-to-vote-the-costs-of-finding-and-getting-to-the-polling-place/6CB07495AFC81E3251CC01A60A10B8FB (visited on 06/03/2025).

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.

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).

Cook, Cody, Lindsey Currier, and Edward Glaeser (2024). "Urban mobility and the experienced isolation of students". In: *Nature Cities* 1.1, pp. 73–82.

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.

Epley, Nicholas and Juliana Schroeder (2014). "Mistakenly seeking solitude". In: *Journal of Experimental Psychology: General* 143.5. Place: US Publisher: American Psychological Association, pp. 1980–1999. ISSN: 1939-2222. DOI: 10.1037/a0037323.

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

Field, Erica and Kate Vyborny (Apr. 28, 2022). *Women's Mobility and Labor Supply: Experimental Evidence from Pakistan*. Rochester, NY. DOI: 10.2139/ssrn.4095705. URL: https://papers.ssrn.com/abstract=4095705 (visited on 09/04/2025).

Foster, Andrew D and Mark R Rosenzweig (1995). "Learning by doing and learning from others: Human capital and technical change in agriculture". In: *Journal of political Economy* 103.6, pp. 1176–1209.

Franklin, Simon (2018). "Location, Search Costs and Youth Unemployment: Experimental Evidence from Transport Subsidies". In: *The Economic Journal* 128.614, pp. 2353–2379.

Fyhn, Marianne et al. (2004). "Spatial representation in the entorhinal cortex". In: *Science* 305.5688, pp. 1258–1264.

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

Gigerenzer, Gerd and Daniel G. Goldstein (Feb. 2011). "The recognition heuristic: A decade of research". In: *Judgment and Decision Making* 6.1, pp. 100–121. ISSN: 1930-2975. DOI: 10.1017/S1930297500002126. URL: https://www.cambridge.org/core/product/identifier/S1930297500002126/type/journal_article (visited on 01/30/2025).

Gittins, John C (1979). "Bandit processes and dynamic allocation indices". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 41.2, pp. 148–164.

Goldstein, Daniel G. and Gerd Gigerenzer (2002). "Models of ecological rationality: The recognition heuristic." In: *Psychological Review* 109.1, pp. 75–90. ISSN: 1939-1471, 0033-295X. DOI: 10.1037/0033-295X.109.1.75. URL: https://doi.apa.org/doi/10.1037/0033-295X.109.1.75 (visited on 01/30/2025).

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*.

Haaland, Ingar K et al. (May 2024). *Understanding Economic Behavior Using Open-ended Survey Data*. Working Paper 32421. National Bureau of Economic Research. DOI: 10.3386/w32421. URL: http://www.nber.org/papers/w32421.

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.

Heckman, James J. (Oct. 1976). "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models". In: *Annals of Economic and Social Measurement, Volume 5, number 4*. NBER, pp. 475–492. URL: https://www.nber.org/books-and-chapters/annals-economic-and-social-measurement-volume-5-number-4/common-structure-statistical-models-truncation-sample-selection-and-limited-dependent-variables-and (visited on 01/30/2025).

Ho, Mark K. et al. (June 2, 2022). "People construct simplified mental representations to plan". In: *Nature* 606.7912, pp. 129–136. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-022-04743-9. URL: https://www.nature.com/articles/s41586-022-04743-9 (visited on 05/05/2025).

Jacoby, Larry L and Mark Dallas (1981). "On the Relationship Between Autobiographical Memory and Perceptual Learning". In: *Journal of Experimental Psychology: General* 110.3, pp. 306–340.

Jacoby, Larry L and Kevin Whitehouse (1989). "An Illusion of Memory: False Recognition Influenced by Unconscious Perception". In: *Journal of Experimental Psychology: General* 118.2, pp. 126–135.

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*.

Keller, Kevin Lane (1987). "Memory Factors in Advertising: The Effect of Advertising Retrieval Cues on Brand Evaluations". In: *Journal of Consumer Research* 14.3. Publisher: Oxford University Press, pp. 316–333. ISSN: 0093-5301. URL: https://www.jstor.org/stable/2489494 (visited on 05/25/2025).

Kihlstrom, Richard E. and Michael H. Riordan (1984). "Advertising as a Signal". In: *Journal of Political Economy* 92.3. Publisher: University of Chicago Press, pp. 427–450. ISSN: 0022-3808. URL: https://www.jstor.org/stable/1837226 (visited on 05/25/2025).

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.

Lagakos, David, Ahmed Mushfiq Mobarak, and Michael E Waugh (2023). "The welfare effects of encouraging rural–urban migration". In: *Econometrica* 91.3, pp. 803–837.

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 (May 2017). "The Benefits of Forced Experimentation: Striking Evidence from the London Underground Network*". In: *The Quarterly Journal of Economics* 132.4, pp. 2019–2055. ISSN: 0033-5533. DOI: 10.1093/qje/qjx020. eprint: https://academic.oup.com/qje/article-pdf/132/4/2019/30637871/qjx020_supp.pdf. URL: https://doi.org/10.1093/qje/qjx020.

Lynch, Kevin (1960). *The Image of the City*. ISSN: 00218529. Cambridge, Massachusetts: MIT Press. DOI: 10.2307/427643.

Mack, Michael L. and Thomas J. Palmeri (June 25, 2024). "Discrimination, Recognition, and Classification". In: *The Oxford Handbook of Human Memory, Two Volume Pack: Foundations and Applications*. Ed. by Michael J. Kahana and Anthony D. Wagner. Oxford University Press, p. 0. ISBN: 978-0-19-091798-2. DOI: 10.1093/oxfordhb/9780190917982.013.30. URL: https://doi.org/10.1093/oxfordhb/9780190917982.013.30 (visited on 03/07/2025).

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.

Milgrom, Paul and John Roberts (1986). "Price and Advertising Signals of Product Quality". In: *Journal of Political Economy* 94.4. Publisher: University of Chicago Press, pp. 796–821. ISSN: 0022-3808. URL: https://www.jstor.org/stable/1833203 (visited on 05/25/2025).

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.

Nelson, Phillip (1970). "Information and Consumer Behavior". In: *Journal of Political Economy* 78.2. Publisher: The University of Chicago Press, pp. 311–329. ISSN: 0022-3808. URL: https://www.jstor.org/stable/1830691 (visited on 05/25/2025).

— (1974). "Advertising as Information". In: *Journal of Political Economy* 82.4. Publisher: University of Chicago Press, pp. 729–754. ISSN: 0022-3808. URL: https://www.jstor.org/stable/1837143 (visited on 05/25/2025).

O'Keefe, John and Jonathan Dostrovsky (1971). "The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat." In: *Brain research*.

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.

Sahni, Navdeep S and Harikesh S Nair (May 1, 2020). "Does Advertising Serve as a Signal? Evidence from a Field Experiment in Mobile Search". In: *The Review of Economic Studies* 87.3, pp. 1529–1564. ISSN: 0034-6527. DOI: 10.1093/restud/rdz053. URL: https://doi.org/10.1093/restud/rdz053 (visited on 05/25/2025).

Schank, Roger C. (1982). *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*. New York, NY: Cambridge University Press. ISBN: 0-521-24858-2.

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

Schroeder, Juliana, Donald Lyons, and Nicholas Epley (May 2022). "Hello, stranger? Pleasant conversations are preceded by concerns about starting one." In: *Journal of Experimental Psychology: General* 151.5, pp. 1141–1153. ISSN: 1939-2222, 0096-3445. DOI: 10.1037/xge0001118. URL: https://doi.apa.org/doi/10.1037/xge0001118 (visited on 06/09/2025).

Shapiro, Bradley T. (Feb. 2018). "Positive Spillovers and Free Riding in Advertising of Prescription Pharmaceuticals: The Case of Antidepressants". In: *Journal of Political Economy* 126.1. Publisher: The University of Chicago Press, pp. 381–437. ISSN: 0022-3808. DOI: 10.1086/695475. URL: https://www.journals.uchicago.edu/doi/full/10.1086/695475 (visited on 05/25/2025).

Shapiro, Bradley T., Günter J. Hitsch, and Anna E. Tuchman (2021). "TV Advertising Effectiveness and Profitability: Generalizable Results From 288 Brands". In: *Econometrica* 89.4. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA17674, pp. 1855–1879. ISSN: 1468-0262. DOI: 10.3982/ECTA17674. URL: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA17674 (visited on 05/25/2025).

Simon, Herbert A and Kevin Gilmartin (July 1, 1973). "A simulation of memory for chess positions". In: *Cognitive Psychology* 5.1, pp. 29–46. ISSN: 0010-0285. DOI: 10.1016/0010-0285(73)90024-8. URL: https://www.sciencedirect.com/science/article/pii/0010028573900248 (visited on 04/29/2025).

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

Thalmann, Mirko, Alessandra S. Souza, and Klaus Oberauer (Jan. 2019). "How does chunking help working memory?" In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 45.1, pp. 37–55. ISSN: 1939-1285, 0278-7393. DOI: 10.1037/xlm0000578. URL: https://doi.apa.org/doi/10.1037/xlm0000578 (visited on 04/29/2025).

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*.

Weitzman, Martin L. (1979). "Optimal Search for the Best Alternative". In: *Econometrica* 47.3. Publisher: [Wiley, Econometric Society], pp. 641–654. ISSN: 0012-9682. DOI: 10.2307/1910412. URL: https://www.jstor.org/stable/1910412 (visited on 08/01/2023).

Winkielman, Piotr et al. (2003). "The hedonic marking of processing fluency: Implications for evaluative judgment". In: *The psychology of evaluation: Affective processes in cognition and emotion*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, pp. 189–217. ISBN: 978-0-8058-4047-6.

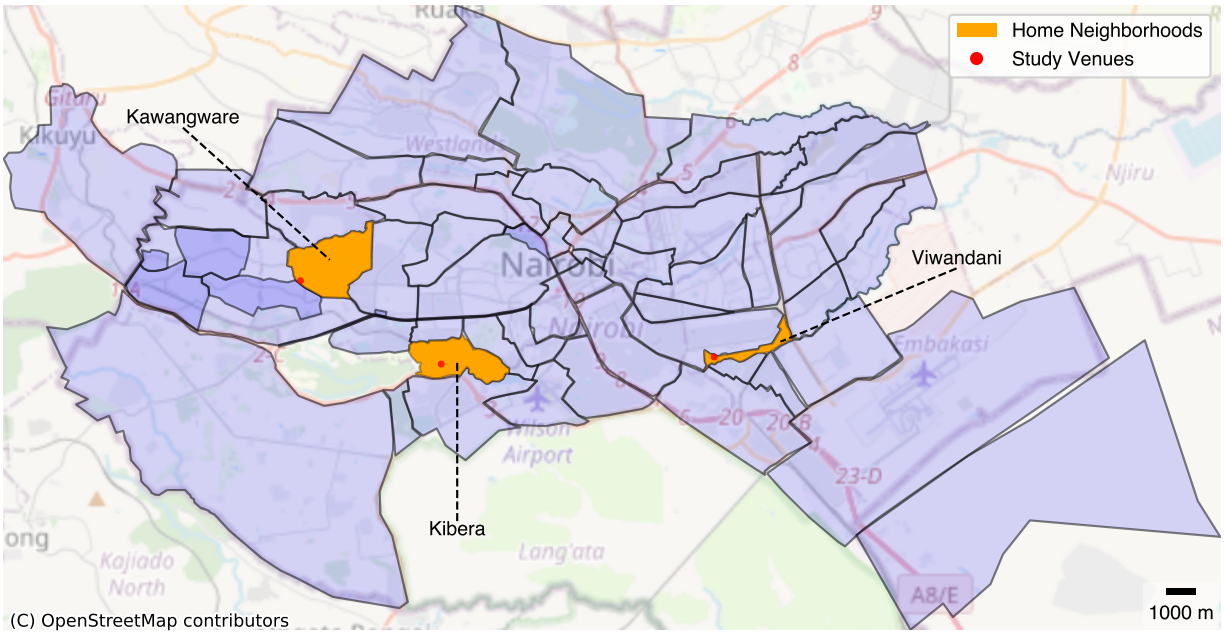
Wiseman, Eleanor (2023). "Border trade and information frictions: Evidence from informal traders in kenya". In: *Working Paper*.

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*.

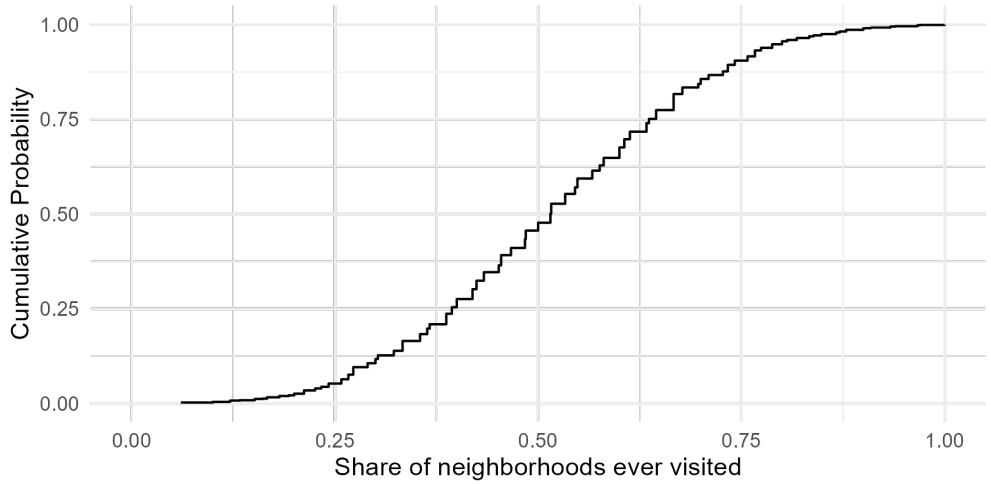
Figures and Tables

Figure 1: Division of Nairobi Into Neighborhoods



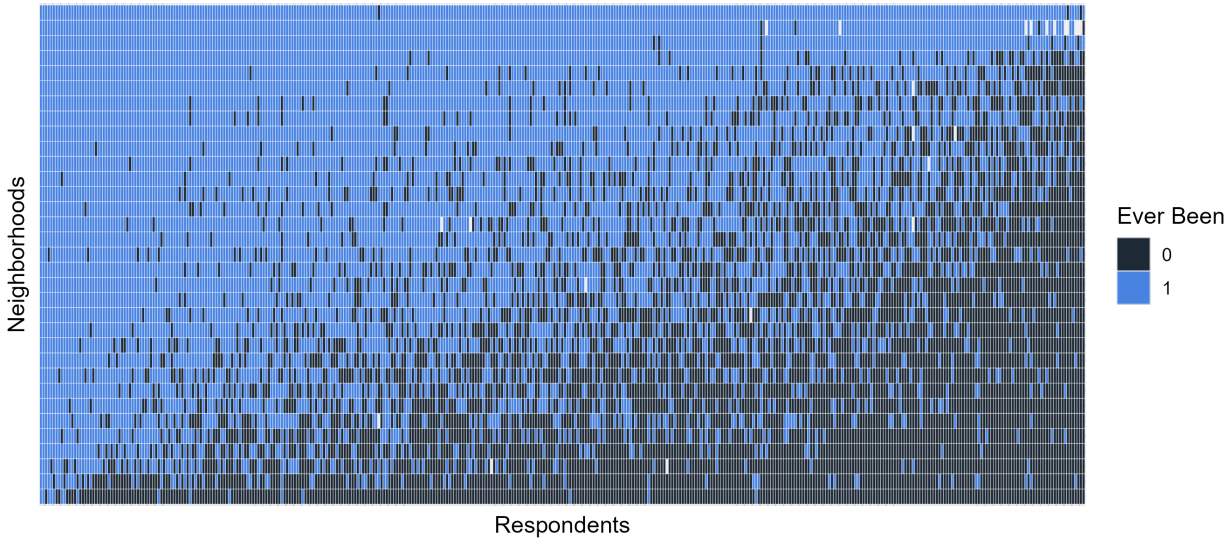
Notes: This figure shows the partition of the main neighborhoods in Nairobi. The orange polygons represent the home neighborhoods—Kibera, Kawangware, and Viwandani—where participants were recruited. The location of the study venues are highlighted in red. The remaining 58 neighborhoods are represented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

Figure 2: CDF of Familiarity Within 75 minutes from Home



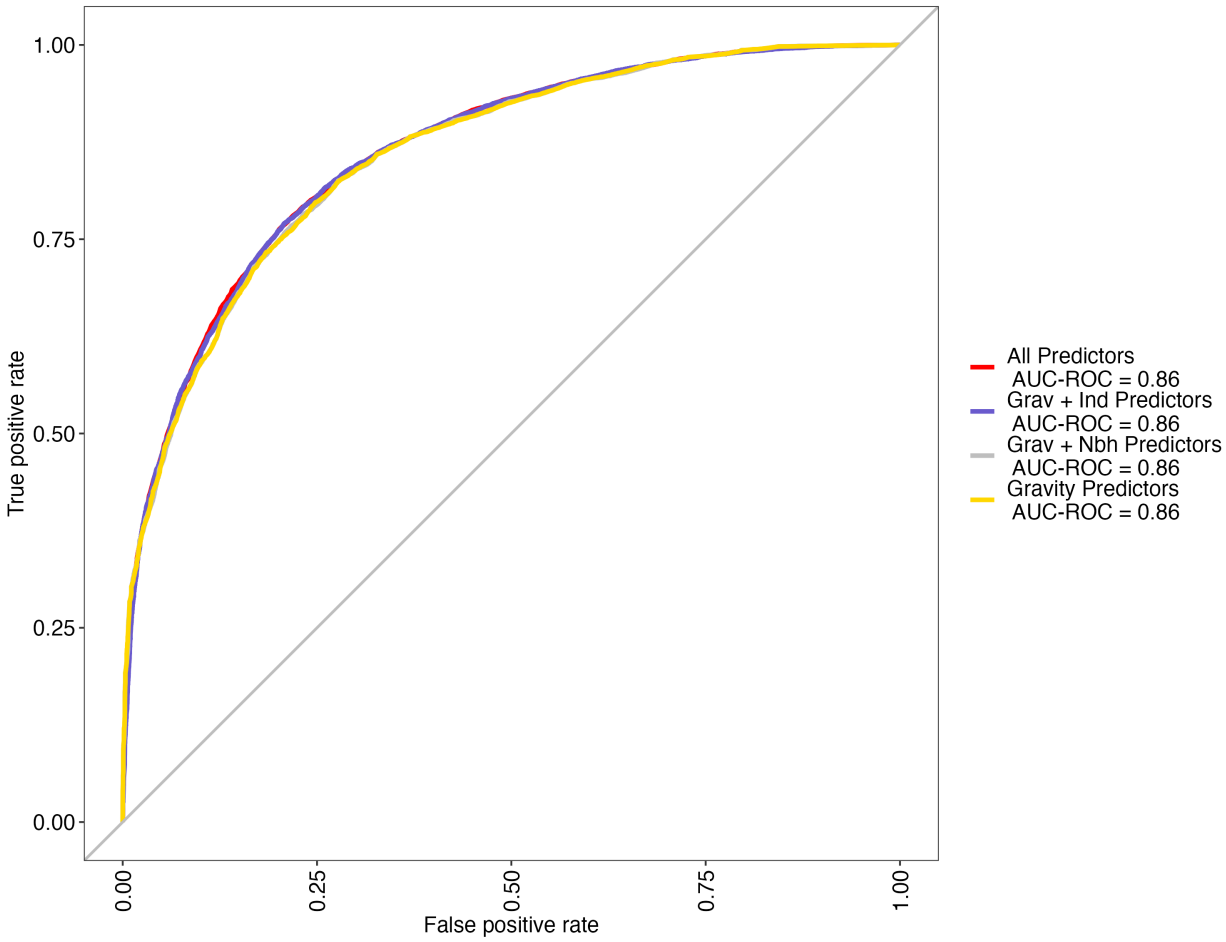
Notes: This figure plots the cumulative distribution function for our main measure of familiarity, across participants. The sample of respondents includes all 1,168 participants who completed the first baseline survey. The sample of neighborhoods is restricted to within 75 minutes (shortest among walking 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.)

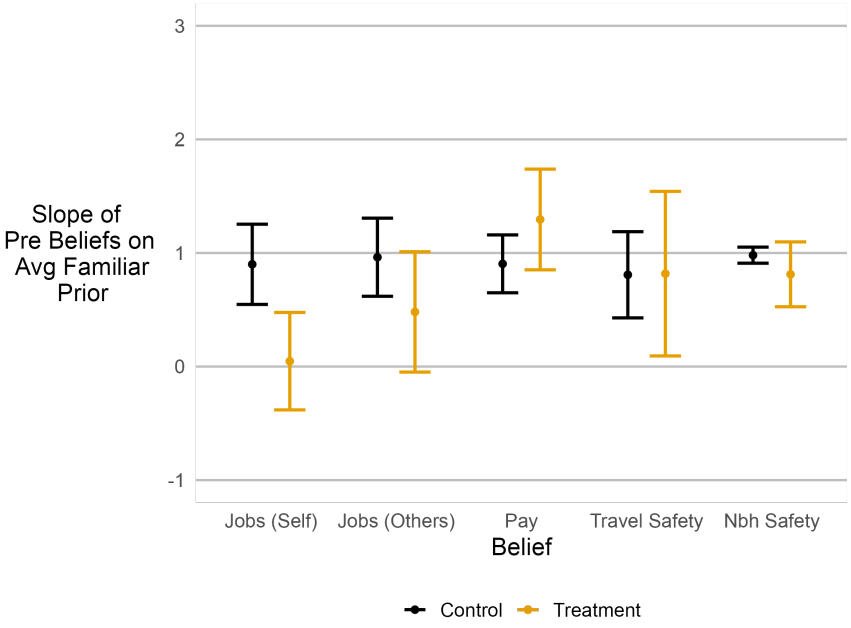
Figure 4: No Evidence of Match-Specific Familiarity Patterns



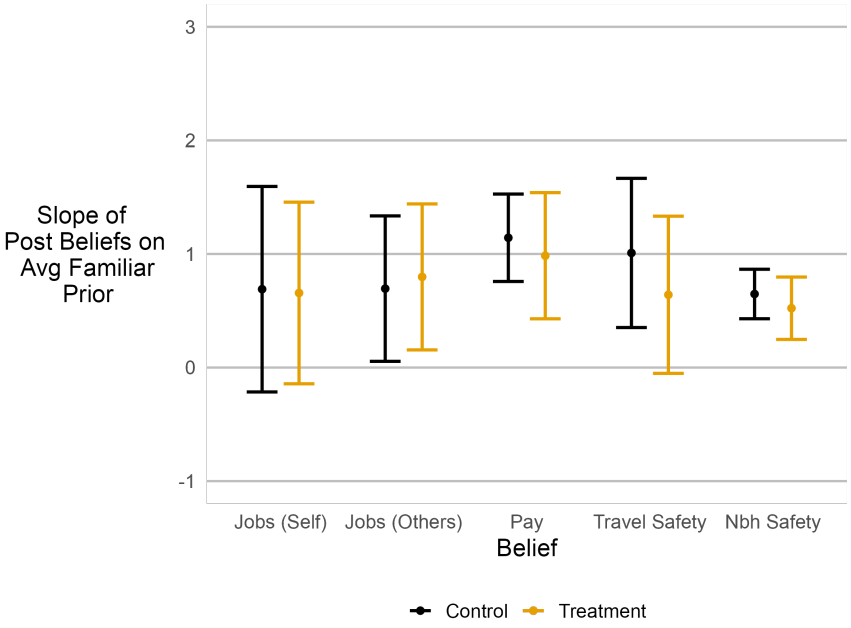
Notes: The figure shows the ROC curves for four different random forest models predicting whether a given individual is familiar with a given neighborhood. While the standard gravity model predictors of distance and neighborhood fixed effects lead to a prediction substantially better than chance (yellow line), adding additional predictors such as individual and neighborhood characteristics (gray, blue, and red lines), does not meaningfully improve the prediction. This suggests match-specific components are not important in predicting familiarity patterns. See Figure A.13 for further results using logit-lasso models.

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

(a) Pre-Visit Beliefs

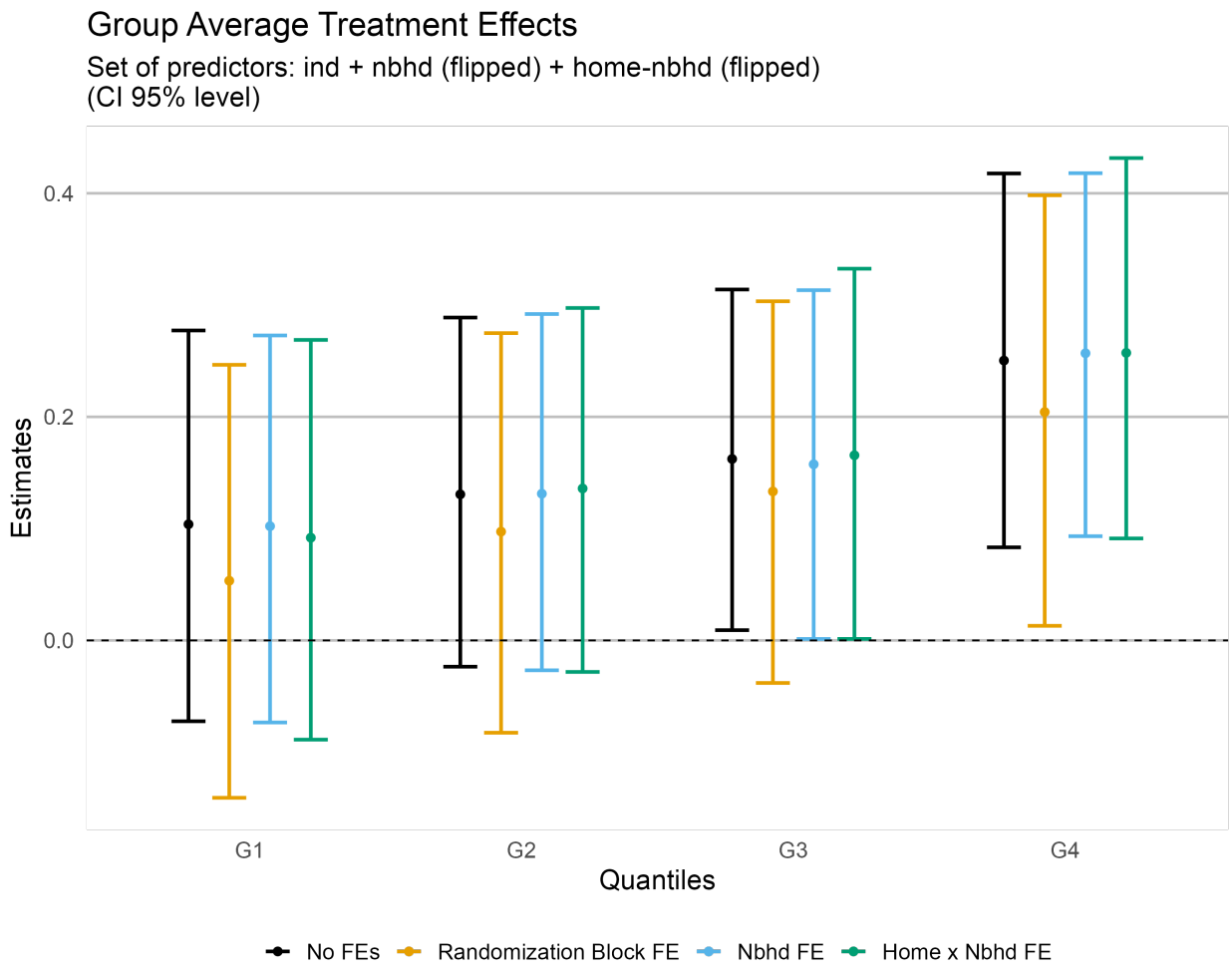


(b) Post-Visit Beliefs



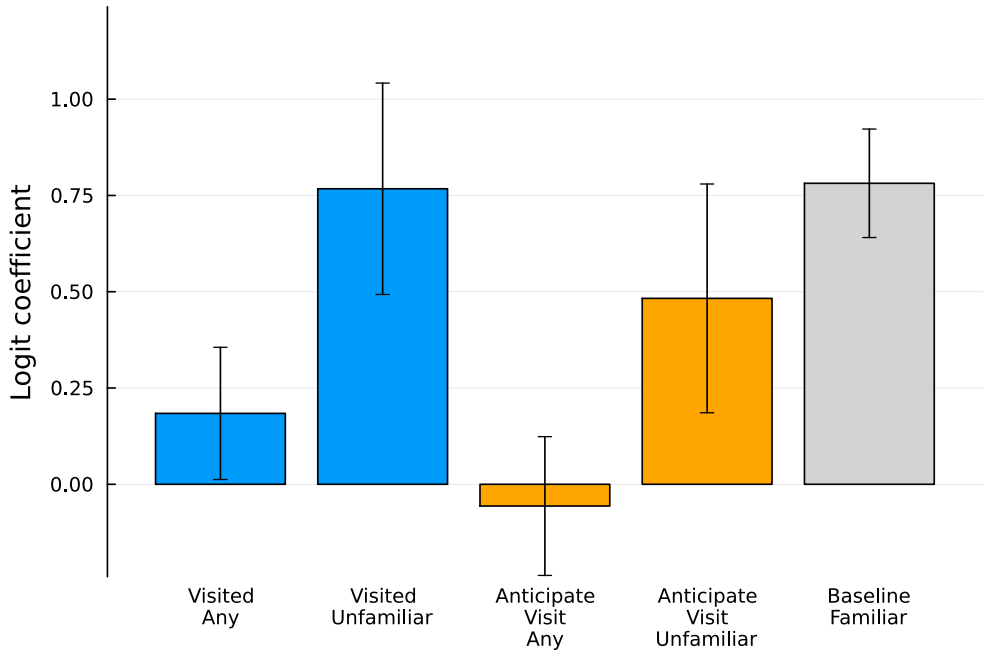
Notes: The figure presents the results comparing beliefs between different respondents, before and after the training intervention. We first construct for each neighborhood and each outcome a predicted rating based on priors from respondents familiar with the respective neighborhood, using a Bayesian 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). In both exercises, the outcome and the ratings are measured on different samples of participants. Standard errors are clustered at the neighborhood level in all regressions.

Figure 6: Quartiles of Heterogeneous Treatment Effects



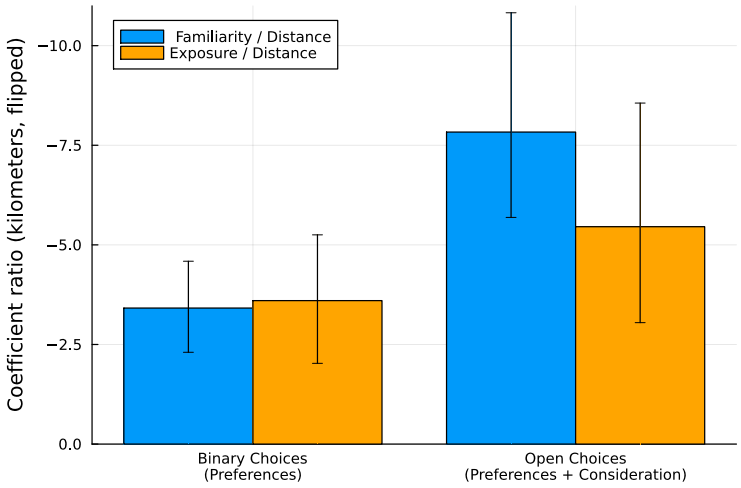
Notes: The figure shows four quartiles of the Group Average Treatment Effects, for the impact of training in an unfamiliar neighborhood, on willingness to later work in that neighborhood, estimated with the Chernozhukov et al. (2023) method. To employ the method we restrict to closed elicitation choices 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 7, column 2.

Figure 8: Open Elicitation and the Revealed Role of Familiarity and Exposure



Notes: This figure reports the coefficient on familiarity and experimental exposure, relative to the coefficient on distance. The Y axis measures the distance equivalent effect of familiarity or experimental exposure. Estimates on the left come from the binary logit specification (1) and those on the right from a multinomial logit specification of the top choice on the first day in the ranked data. Figure A.15 repeat the exercise with the results for the top choice on days 2 and 3 of open elicitation.

Table 1: Participants Have Significant Familiarity Gaps

Nbhd sample:	Share of Neighborhoods Visited					
	< 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 × neighborhood FE		Yes	
Neighborhood FE			Yes

Notes: This table reports the results of the logit estimation of the "closed" 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)
Constant	3.418***	3.427***	3.580***	4.030***	3.844***
	(0.037)	(0.035)	(0.040)	(0.035)	(0.036)
post	−0.134***	−0.189***	−0.055	0.140***	0.237***
	(0.039)	(0.039)	(0.037)	(0.033)	(0.037)
pre:treated	−0.247***	−0.249***	−0.050	−0.393***	−0.170***
	(0.056)	(0.053)	(0.058)	(0.053)	(0.053)
post:treated	−0.077	−0.092*	−0.012	0.003	0.024
	(0.051)	(0.052)	(0.047)	(0.040)	(0.041)
Mean	3.418	3.427	3.58	4.03	3.844
SD	1.022	0.996	1.047	0.887	0.975
pre x treat = post x treat, p-val	0.003	0.004	0.492	0.000	0.001
Observations	4,584	4,575	4,468	4,587	4,518
Note:					*p<0.1; **p<0.05; ***p<0.01

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. In the analysis sample, those in the treated group are rating unfamiliar neighborhoods while those in the control are rating familiar neighborhoods. 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.16 repeats the analysis using belief confidence as the outcome. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table 4: Impact of Familiarity and Experimental Exposure on Medium-run Job Acceptance

		All Invites			Second Round Invites		
		(1)	(2)	(3)	(4)	(5)	(6)
Baseline Familiar	× (Wage ≤ 500)	0.39***	0.32**	0.31**	0.47***	0.44***	0.46***
		(0.12)	(0.13)	(0.13)	(0.14)	(0.17)	(0.16)
Visited Any	× (Wage ≤ 500)	0.13	0.17	0.14	0.19	0.26*	0.21
		(0.12)	(0.12)	(0.12)	(0.14)	(0.15)	(0.15)
Visited Unfamiliar	× (Wage ≤ 500)	0.24	0.24	0.25	0.27	0.17	0.26
		(0.22)	(0.22)	(0.22)	(0.26)	(0.28)	(0.27)
Distance (km)	× (Wage ≤ 500)	-0.15***		-0.16***	-0.14***		-0.13***
		(0.02)		(0.02)	(0.02)		(0.03)
Wage	× (Wage ≤ 500)	0.53***	0.54***	0.54***	0.57***	0.60***	0.59***
		(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
(Wage > 500)		1.7**	3.0***	1.8***			
		(0.69)	(0.61)	(0.70)			
Baseline Familiar	× (Wage > 500)	0.09	-0.25	-0.06			
		(0.25)	(0.27)	(0.26)			
Visited Any	× (Wage > 500)	0.09	0.07	0.11			
		(0.25)	(0.27)	(0.26)			
Visited Unfamiliar	× (Wage > 500)	-0.07	0.02	-0.08			
		(0.41)	(0.44)	(0.43)			
Distance (km)	× (Wage > 500)	-0.01		-0.03			
		(0.03)		(0.04)			
Wage	× (Wage > 500)	0.06	0.07	0.07			
		(0.07)	(0.07)	(0.07)			
N		3,960	3,939	3,960	1,814	1,795	1,803
Survey Day FE		Yes	Yes	Yes	Yes	Yes	Yes
Invite Sequence FE		Yes	Yes	Yes	Yes	Yes	Yes
Randomization Strata FE		Yes	Yes	Yes	Yes	Yes	Yes
Home × Neighborhood FE			Yes			Yes	
Neighborhood FE				Yes			Yes

Notes: This table reports binary logit estimates using a version of equation (1) for the show-up outcome for the travel survey invitations 2-4 months after the intervention. The outcome is an indicator for whether the participant showed up. We include survey date, invite sequence, and randomization strata fixed effects. We interact all coefficients with an indicator for wage larger than 500 KSH due to persistently high show-up rates above that level (see Figure A.11). Columns 4 - 6 report results from the second round of invitations where all wages were below 500 KSH

Table 5: People Re-visit the Neighborhoods From the Study

	Endline Survey (prompted) In the last two weeks				SMS (unprompted) Yesterday	
	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
Visited Any + Visited Unfam	0.074	0.043	0.004	0.043	0.013	0.064
(se)	(0.02)	(0.013)	(0.011)	(0.014)	(0.005)	(0.021)
[p-value]	[0.000]	[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 1-4 are trips measured using the over-the-phone endline survey while columns 5 and 6 are visits measured by the high-frequency SMS. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table 6: Selection Model For Duration and Navigation to Unfamiliar Neighborhoods

	Task App (hours)		Navigation (Post-Visit)				
	To Job	From Job	Get Lost There	Get Lost Back	Ask Directions There	Ask Directions Back	More Difficult Than Expected
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Baseline familiar	-0.06*	-0.07	-0.01	0.00*	-0.10***	-0.04**	-0.09***
	(0.04)	(0.04)	(0.02)	(0.00)	(0.04)	(0.02)	(0.03)
Visited Any	-0.07**	-0.02	-0.05***	0.00	-0.18***	-0.01	-0.07***
	(0.03)	(0.04)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Visited Unfamiliar	0.05	-0.02	0.01	0.01	-0.07	-0.03	-0.09***
	(0.06)	(0.06)	(0.02)	(0.01)	(0.04)	(0.02)	(0.03)
Distance (km)	0.09***	0.10***	0.00	0.00	0.01*	-0.00	0.00
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Cons	0.51***	0.40***	0.07**	0.00	0.26***	0.07***	0.14***
	(0.05)	(0.07)	(0.03)	(0.00)	(0.06)	(0.02)	(0.04)
Control Mean	1.23	1.30	0.06	0.00	0.31	0.05	0.18
N	2,291	2,257	2,318	2,318	2,318	2,318	2,317

Notes: This table presents estimates from a Heckman (1976) two-step probit model applied to the work-choices bundles over the employment days. The sample includes participants in the closed elicitation group who completed the job during the employment period. In columns 1 to 6, odd-number columns refer to the trip to the job neighborhood, and even-number columns refer to the round trip to the home location. The dependent variables in columns 1 and 2 represent travel duration (in hours). Columns 3 and 4 use indicators showing whether participants got lost during their trips, while columns 5 and 6 use whether participants asked for directions. Column 7 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. Bootstrap standard errors with 1000 replications in parenthesis. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

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

	(1)	(2)	(3)	(4)
Baseline Familiar	0.78***	0.81***	0.79***	0.72***
	(0.071)	(0.069)	(0.076)	(0.218)
β^V Visited Any	0.18*	0.20*	0.01	0.68***
	(0.088)	(0.086)	(0.115)	(0.198)
β^U Visited Unfamiliar	0.77***	0.78***	0.80***	0.26
	(0.149)	(0.138)	(0.184)	(0.363)
β^{AV} Anticipate Visited Any	-0.06	-0.04	-0.05	0.03
	(0.095)	(0.097)	(0.086)	(0.190)
β^{AU} Anticipate Visited Unfamiliar	0.48**	0.45**	0.41**	0.29
	(0.159)	(0.154)	(0.147)	(0.342)
N	13,658	13,658	6,902	2,137
P-value $\beta^U = \beta^{AU}$	0.026	0.012	0.020	0.434
Sample: Training	Yes	Yes	Yes	Yes
Sample: Employment	Yes	Yes		Yes
Sample: only Q1				Yes
Neighborhood FEs		Yes	Yes	Yes
Job Attribute Controls	Yes	Yes	Yes	Yes

Notes: This table reports binary logit estimation results of equation (2). Columns 1 and 2 pool all the choice data for the three training days, and for the three employment days. Column 3 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 $VisitedAny_{in} = 1$ for choices made on training days $d \in \{1, 2\}$. Column 4 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 8: Estimates of Memory Costs from "Open" Elicitation

	Utility (α^U)	**Cost** (α^C)
Baseline Familiar	0.86	2.53
		[1.67, 4.02]
Visited Any	0.30	2.02
		[1.34, 3.12]
Visited Unfamiliar	0.91	1.08
		[0.37, 2.13]
Distance	-0.25	-0.40
		[-0.62, -0.28]
Constant		2.49
		[1.19, 5.42]
Cost variance σ		1.98
		[1.40, 2.90]
Subjective λ		0.50
		[0.36, 0.66]
Observations		54,155
Participants		391
Top K choices		5

Notes: This table reports maximum likelihood estimates of the two-self memory model from section 6. 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 on the first day of employment. We use the first (up to) 5 ranked neighborhoods, including the decision to stop ranking. We parametrize $\lambda_k = \lambda(1-\lambda)^{k-1}$ and estimate the subjective job probability λ. Individual-level bootstrapped 95% confidence intervals in parentheses.

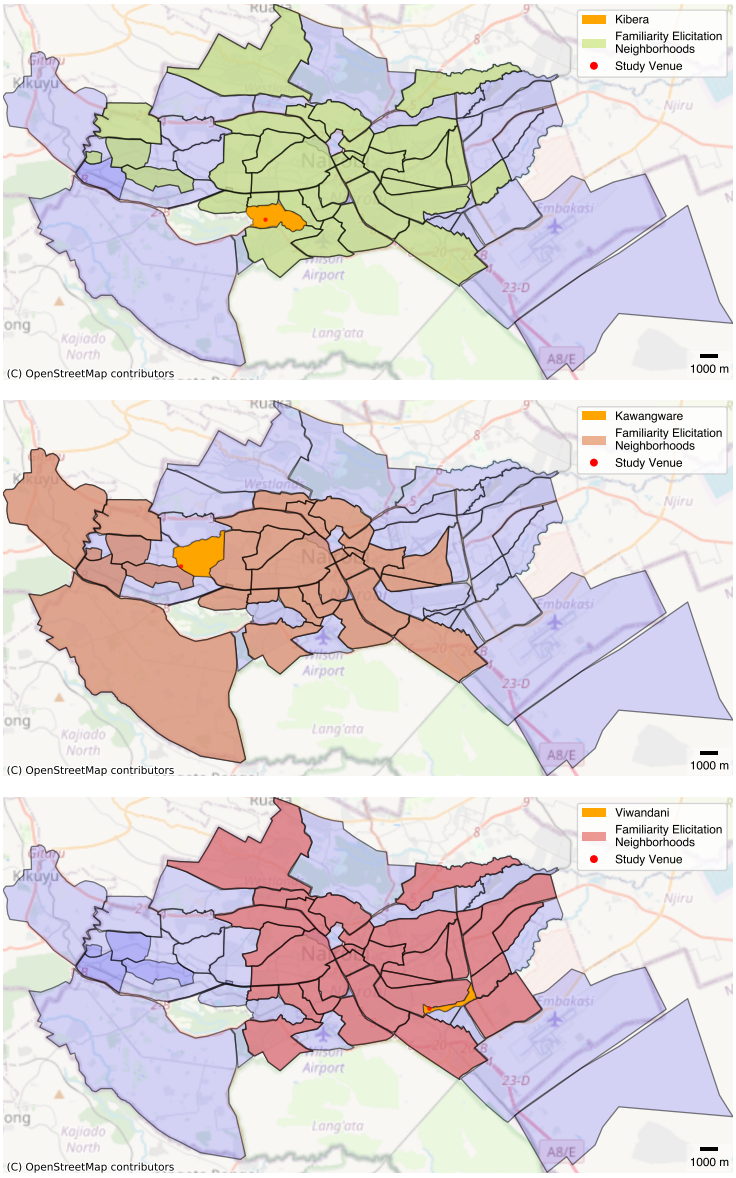
A Appendix - For Online Publication

Contents

A.1 Appendix Figures 57
A.2 Appendix Tables 72
A.3 Measuring Beliefs 90

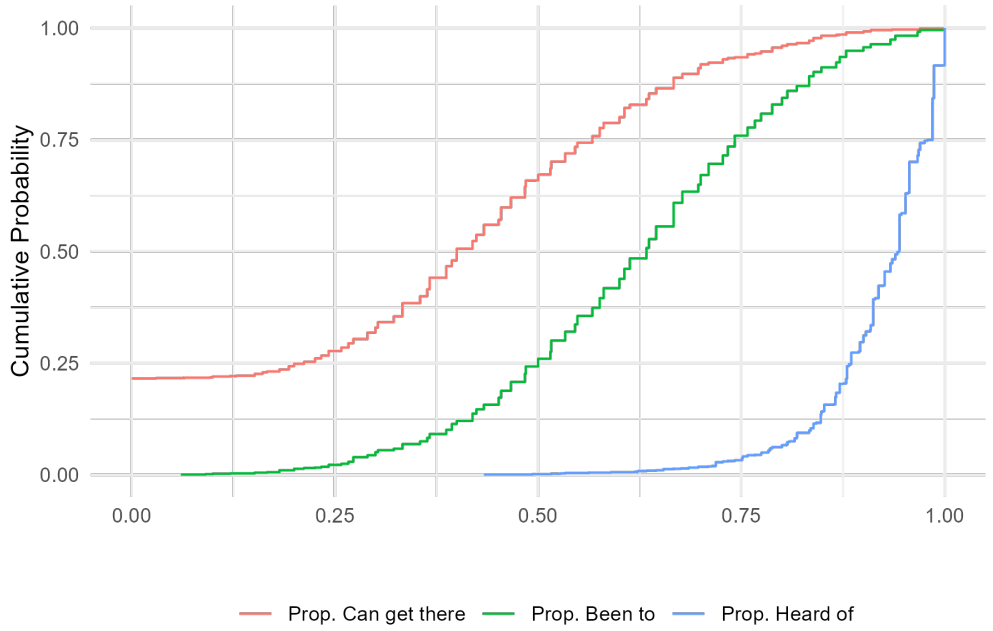
A.1 Appendix Figures

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. For each home neighborhood, these neighborhoods were reachable in at most 75 minutes from the study venue. A small number of neighborhoods were dropped due to safety or other concerns.

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



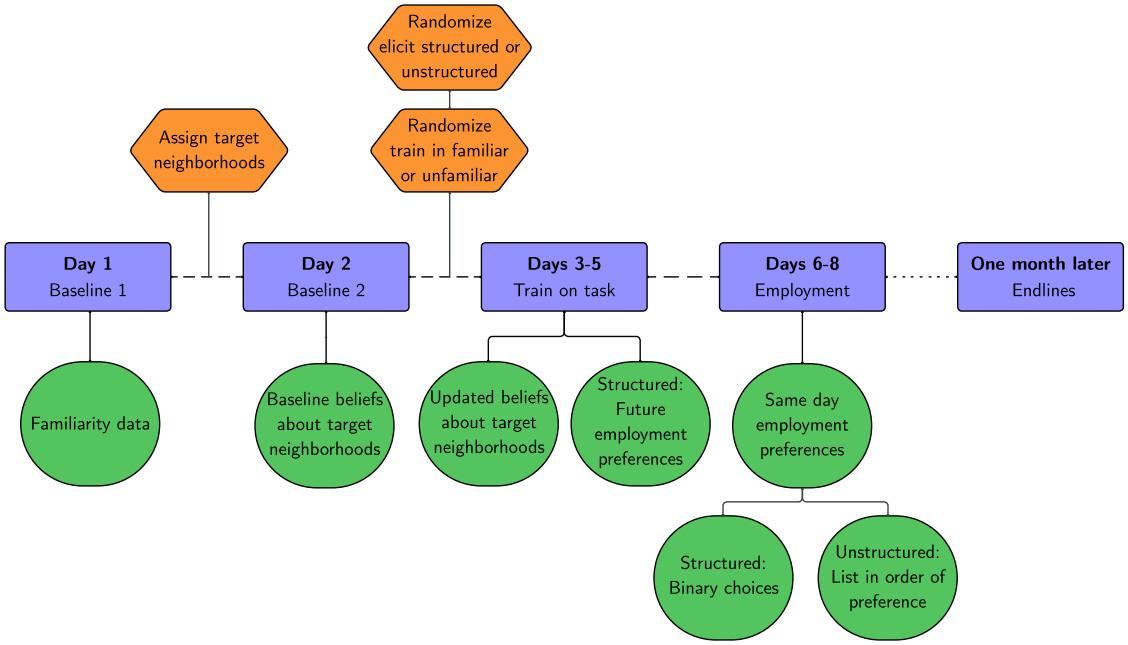
Notes: This figure replicates Figure 2 with three different measures of familiarity.

Figure A.3: Backpacks Used in Employment Task



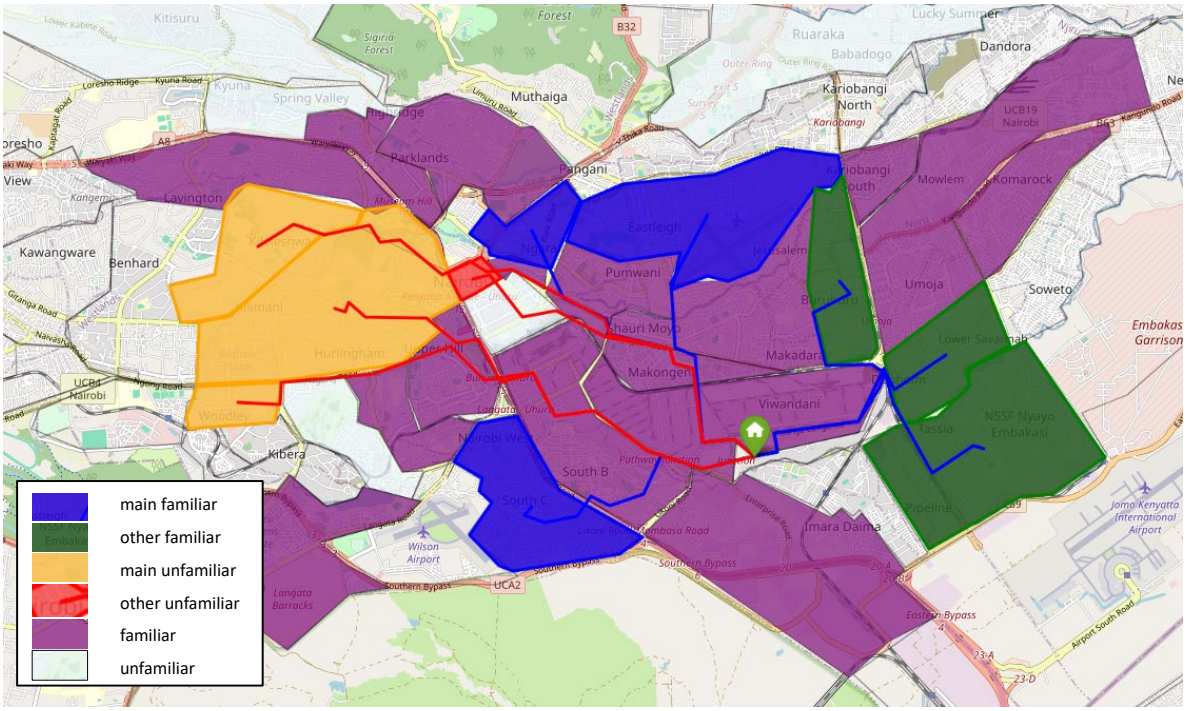
Notes: This figure shows a picture of an air pollution backpack used by study participants during training and employment days.

Figure A.4: Timeline of Study



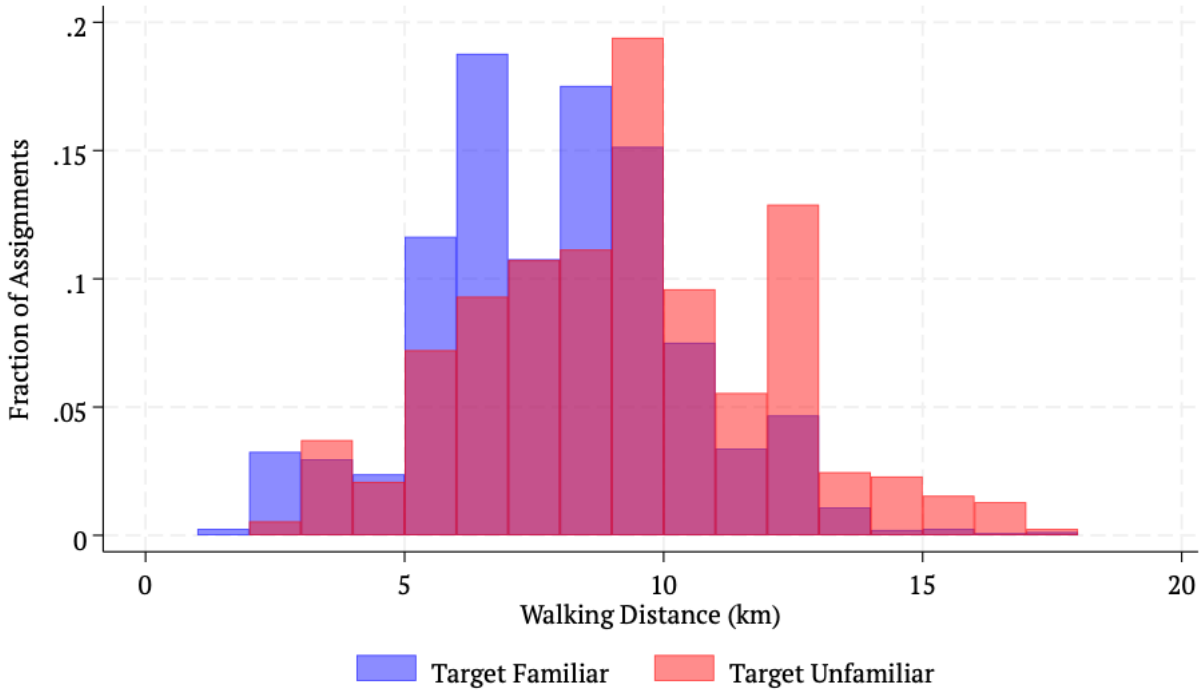
Notes: This figure shows the timing outline of the study. In the first two days, participants provided information about their neighborhood familiarity, demographics, beliefs about target neighborhoods, employment and job search status, and spatial ability. Over the next three days, participants were 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 "closed" elicitation group were asked 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 "open" survey or "closed" binary choices—Section 4 discusses the "closed" 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



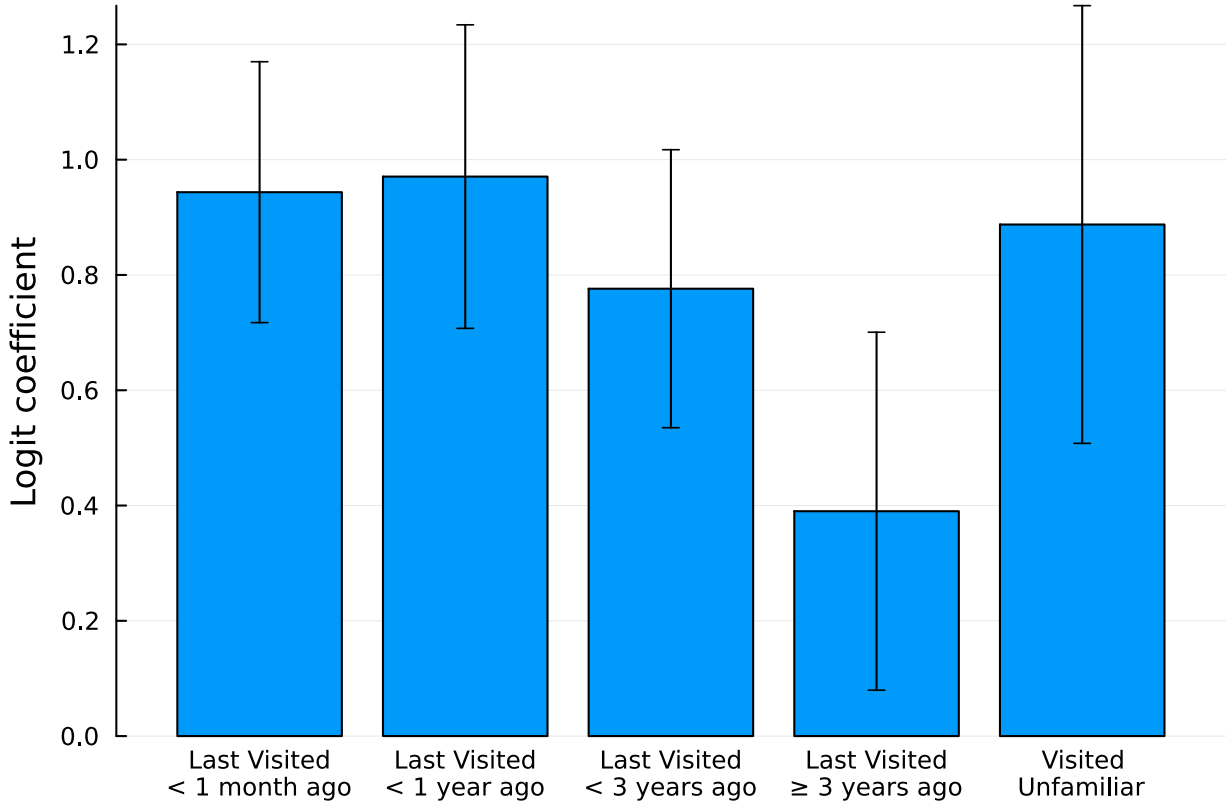
Notes: This map plots the ten target neighborhoods for one participant from the home neighborhood of Viwandani. Non-target neighborhoods are shaded in purple if they had ever visited at baseline or white if they had not. Ambiguous neighborhoods are not shaded. The four target categories are shared in 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



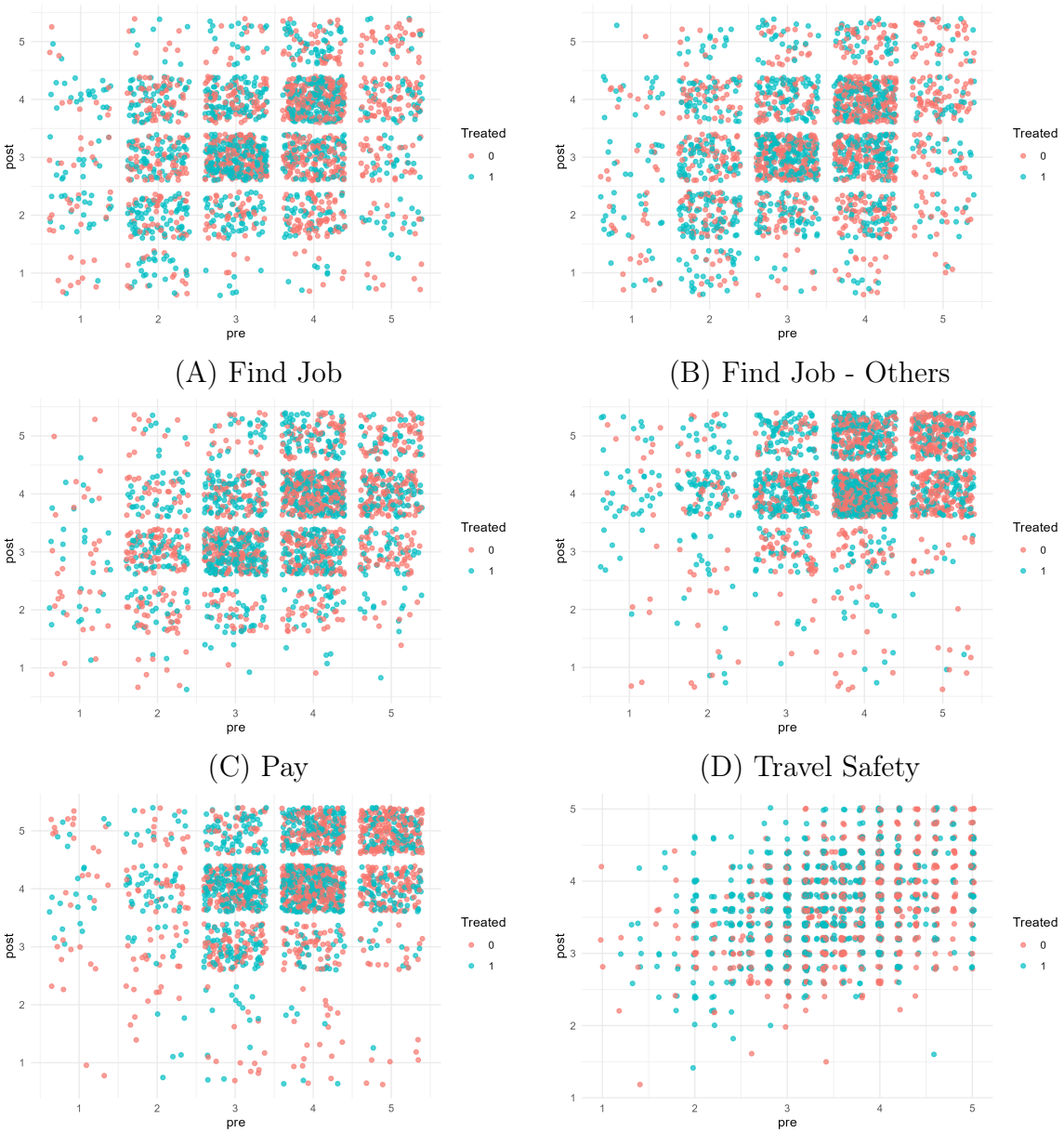
Notes: This figure plots the distribution of walking distances to home neighborhood in kilometers by target familiar and unfamiliar neighborhoods. The sample includes 4796 observations, split as 799 participants each assigned 3 target familiar and 3 target unfamiliar neighborhoods. See Table A.4 for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

Figure A.7: Familiarity Lasts 3 Years



Notes: This figure plots the coefficients of Table A.18 column 2.

Figure A.8: Beliefs Transition Matrices



(A) Find Job

(B) Find Job - Others

(C) Pay

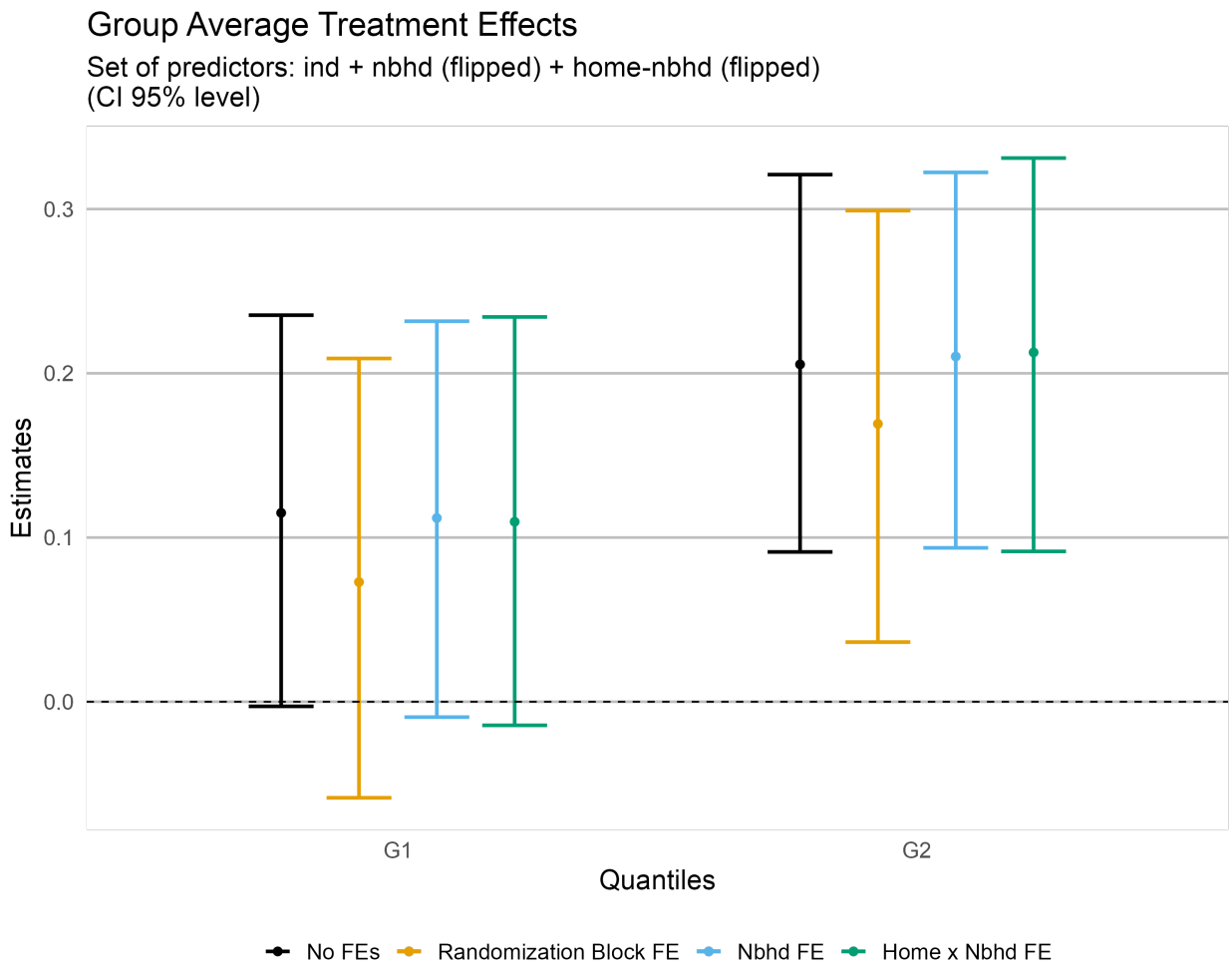
(D) Travel Safety

(E) Overall Safety

(F) Average (All Measures)

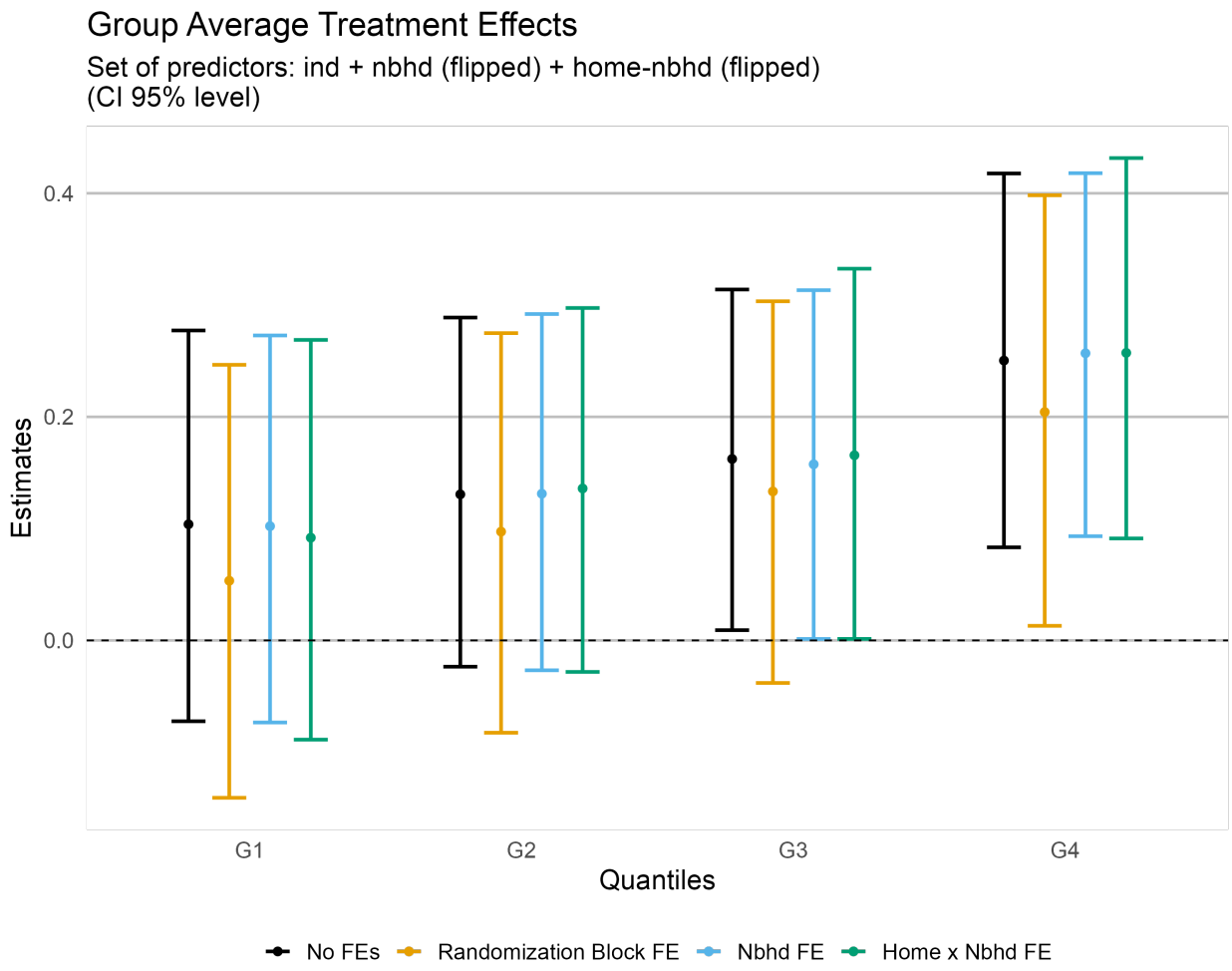
Notes: each graph plots beliefs pre- and post-intervention on the X- and Y-axis, respectively. Higher values correspond to better outcomes. See Section A.3 for the precise questions underlying each measure.

Figure A.9: Median Split of Heterogeneous Treatment Effects



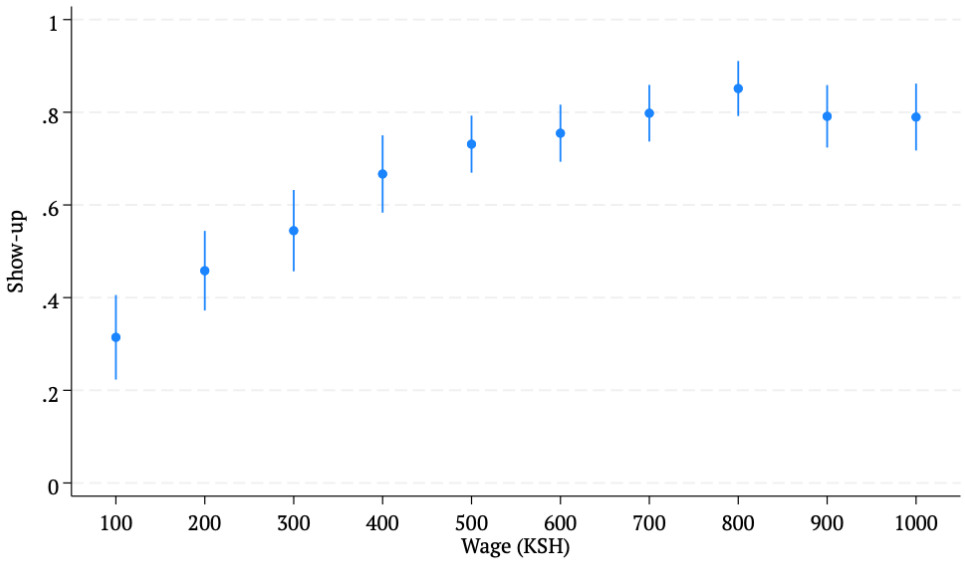
Notes: The figure shows two quartiles of the Group Average Treatment Effects estimated with the Chernozhukov et al. (2023) method. To employ the method we restrict to closed elicitation choices between an non-target familiar neighborhood and an unfamiliar neighborhood. We then use as the outcome an indicator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

Figure A.10: Quartile Splits of Heterogeneous Treatment Effects



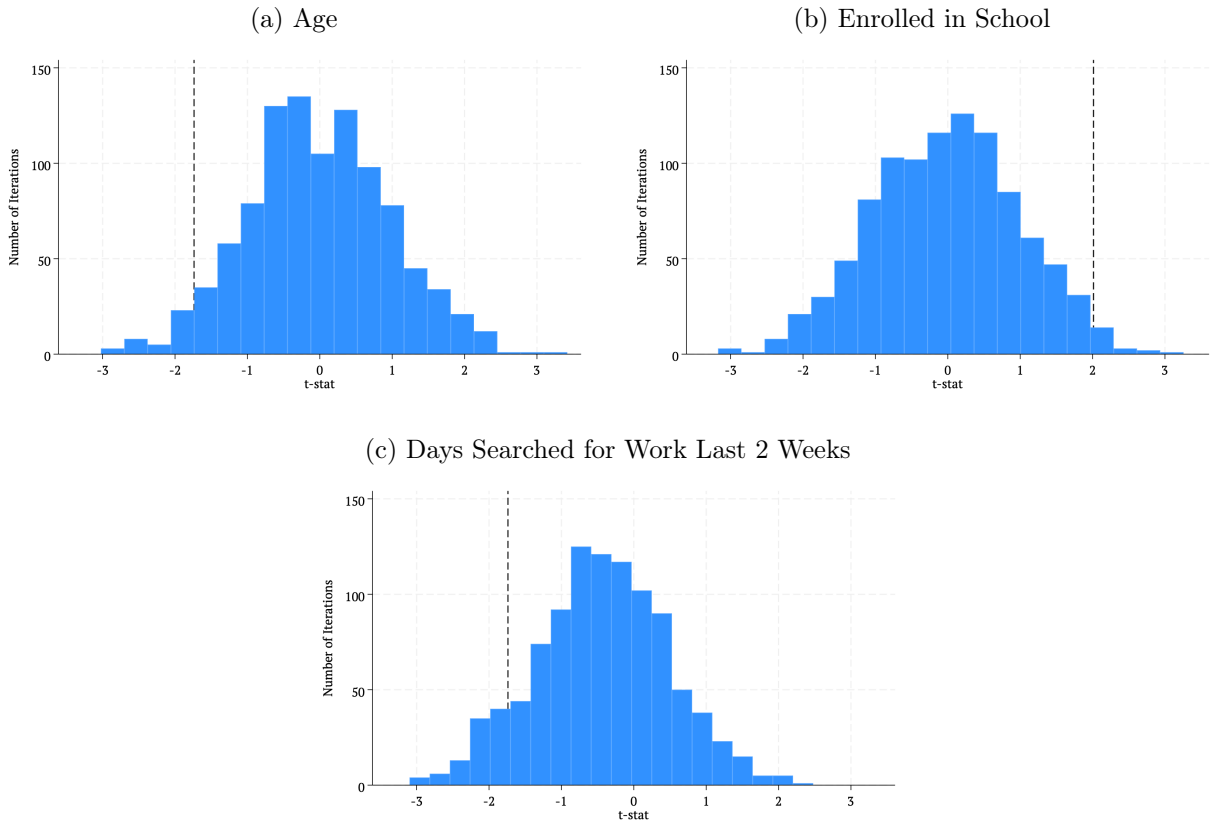
Notes: This figure replicates Figure A.9 using four quartiles of the Group Average Treatment Effects.

Figure A.11: Show-up 2-4 Months After Intervention: High Wage Plateau



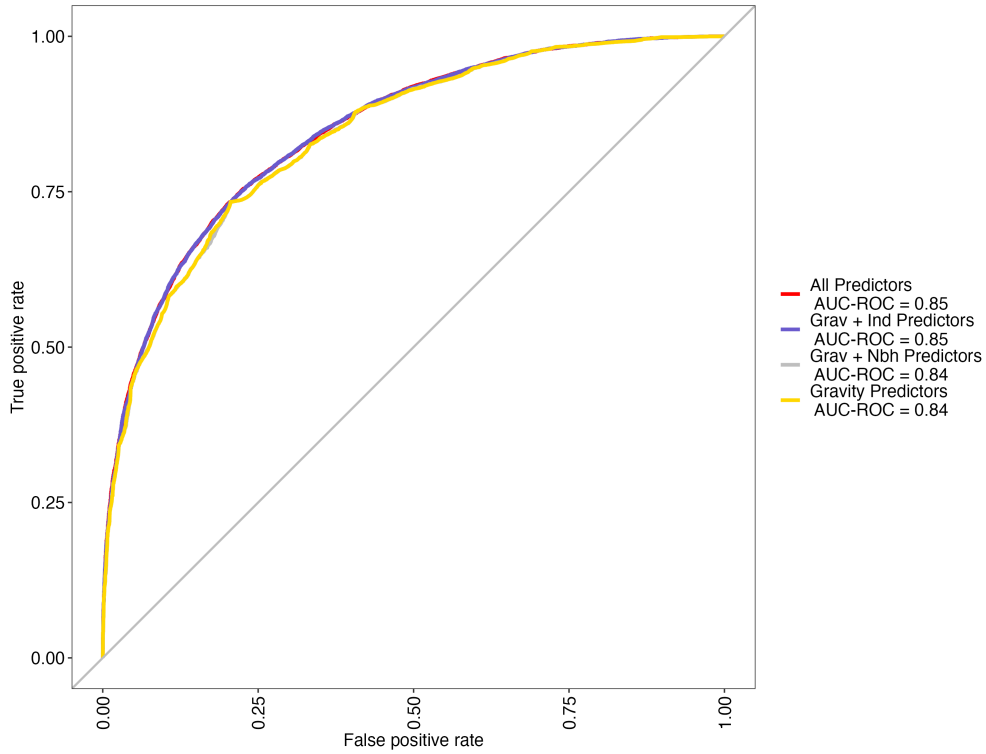
Notes: This figure plots average show-up for the travel survey task by (random) wage offer.

Figure A.12: Randomization Inference Results for Imbalances in the Closed Elicitation Sample



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 closed elicitation 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 closed elicitation sample.

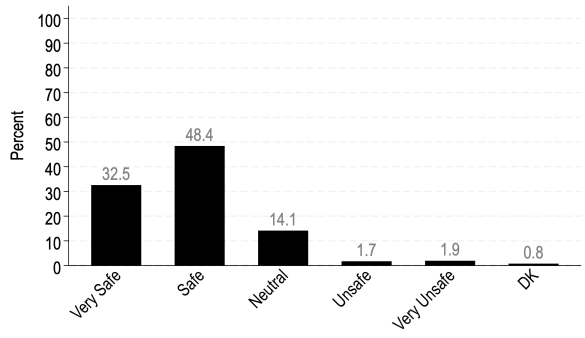
Figure A.13: No Evidence of Match-Specific Familiarity Patterns



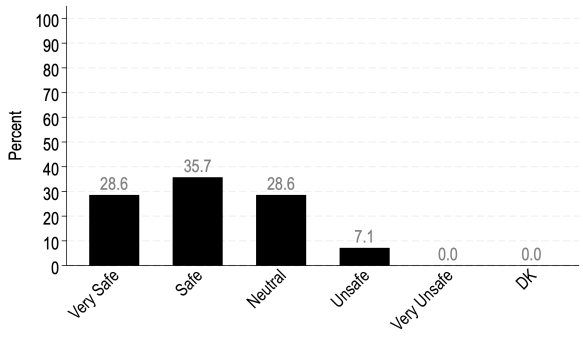
Notes: This figure replicates Figure 4 results using logit-lasso models.

Figure A.14: Safety Ratings during Training Days: Pooled Sample vs. Non-Showups

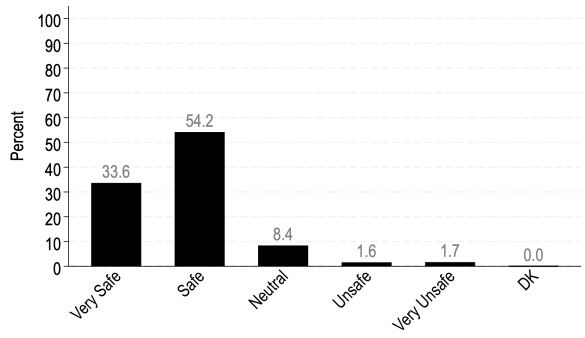
(a) General safety — All participants (Days 1–3)



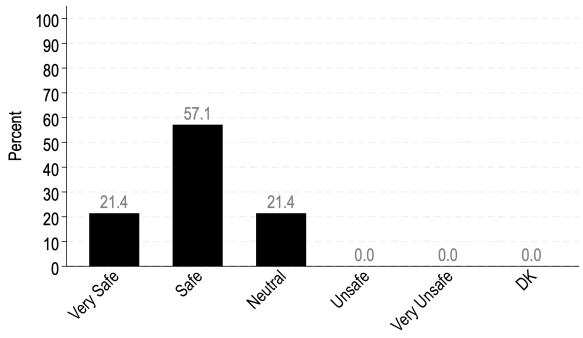
(b) General safety — Non-showups ($t-1$)



(c) Travel safety — All participants (Days 1–3)

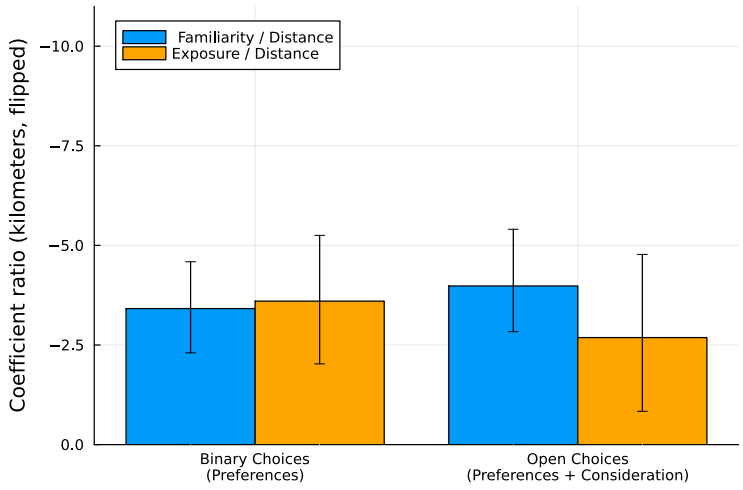


(d) Travel safety — Non-showups ($t-1$)

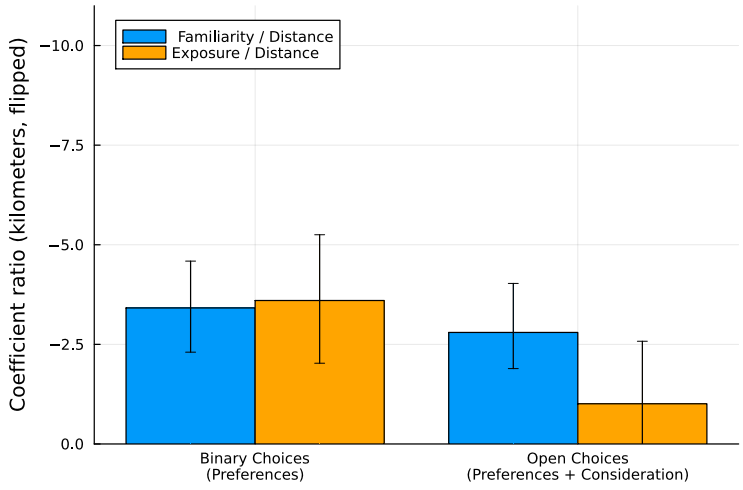


Notes: Panels show the distribution of safety ratings for being in the neighborhood (top row) and traveling to the neighborhood (bottom row). Left panels pool all participants across the three training days (2,397 observations). Right panels report ratings for non-showups (14 observations) from the training day immediately preceding their absence.

Figure A.15: Open Elicitation and the Revealed Role of Familiarity and Exposure (Days 2 and 3)



Panel (A)



Panel (B)

Notes: This figure replicates Figure 8 for the top choice on days 2 and 3 of open elicitation.

A.2 Appendix Tables

Table A.1: 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.2: 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.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 Closed Elicitation 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 "closed" elicitation group. See Figure A.12 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	Employment	Endline
	(1)	(2)	(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 × 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: 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.13: 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.14: 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.15: 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.16: 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 × Treated	−0.101	−0.116*	−0.141**	−0.164***	−0.125**
	(0.063)	(0.065)	(0.060)	(0.058)	(0.056)
Post × 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 × 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.17: 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.18: 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 $\geq$ 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 results from a logit model estimated on the "closed" elicitation 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.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 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.21: 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.

Table A.22: 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 estimates from multinomial logit models of the k-th choice on the list in the open job elicitation, on the first employment day, for $k = 1, \ldots, 5$. For comparison, the first column repeats the binary logit results from Table 2, column 3, based on binary choices between options that we ask about.

Table A.23: Estimates of Memory Costs from "Open" Elicitation: Different Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Baseline Familiar	1.89	2.28	2.53	0.88	0.62	0.58
	[1.38, 2.46]	[1.77, 3.01]	[1.67, 4.02]	[0.70, 1.08]	[0.49, 0.75]	[0.48, 0.69]
Visited Any	1.51	1.73	2.02	0.89	0.68	0.64
	[1.12, 2.01]	[1.29, 2.31]	[1.34, 3.12]	[0.73, 1.09]	[0.53, 0.86]	[0.53, 0.78]
Visited Unfamiliar	0.79	1.01	1.08	0.37	0.28	0.26
	[0.19, 1.38]	[0.41, 1.72]	[0.37, 2.13]	[0.14, 0.57]	[0.10, 0.43]	[0.08, 0.42]
Distance	-0.29	-0.33	-0.40	-0.14	-0.09	-0.09
	[-0.36, -0.22]	[-0.40, -0.27]	[-0.62, -0.28]	[-0.17, -0.11]	[-0.12, -0.08]	[-0.10, -0.07]
Constant			2.49	-0.64	-0.83	-0.84
			[1.19, 5.42]	[-0.79, -0.42]	[-0.94, -0.68]	[-0.93, -0.73]
Cost variance σ	1.49	1.70	1.98	0.86	0.67	0.64
	[1.21, 1.92]	[1.41, 2.11]	[1.40, 2.90]	[0.73, 1.03]	[0.56, 0.76]	[0.56, 0.71]
lambda	0.39		0.50	0.29	0.25	0.24
	[0.28, 0.48]		[0.36, 0.66]	[0.25, 0.33]	[0.21, 0.28]	[0.21, 0.27]
N	54,155	54,155	54,155	54,155	54,155	54,155
Top K choices	5	5	5	10	15	20

Notes: This table replicates the MLE estimates from Table 8 varying the specification and the number of ranked positions used in estimation $K \in \{5, 10, 15, 20\}$.

Table A.24: Estimates of Memory Costs from "Open" Elicitation: Days 2 and 3

	(1)	(2)	(3)
Baseline Familiar	2.53	3.08	2.90
	[1.67, 4.02]	[1.81, 6.11]	[2.01, 21.29]
Visited Any	2.02	2.82	3.07
	[1.34, 3.12]	[1.60, 5.44]	[1.97, 18.96]
Visited Unfamiliar	1.08	1.44	0.97
	[0.37, 2.13]	[0.36, 3.77]	[0.15, 9.25]
Distance	-0.40	-0.65	-0.82
	[-0.62, -0.28]	[-1.29, -0.39]	[-5.74, -0.53]
Constant	2.49	6.32	8.21
	[1.19, 5.42]	[2.69, 15.10]	[4.98, 76.26]
Cost variance σ	1.98	2.60	2.85
	[1.40, 2.90]	[1.57, 4.90]	[2.01, 18.47]
lambda	0.50	0.55	0.62
	[0.36, 0.66]	[0.31, 0.84]	[0.00, 0.73]
N	54,155	54,000	53,355
Top K choices	5	5	5

Notes: This table replicates the MLE estimates from Table 8 using data from days two and three of the open elicitation.

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.