PEDL Research Papers

This research was partly or entirely supported by funding from the research initiative Private Enterprise Development in Low-Income Countries (PEDL), a Foreign, Commonwealth & Development Office (FCDO) funded programme run by the Centre for Economic Policy Research (CEPR).

MATCHING WITH THE RIGHT ATTITUDE: THE EFFECT OF MATCHING FIRMS WITH REFUGEE WORKERS

FRANCESCO LOIACONO† AND MARIAJOSE SILVA-VARGAS§

Job Market Paper
This version: February 3, 2023
– Latest version here –

Abstract

How to integrate disadvantaged workers such as immigrants and refugees into host-country labor markets is a pressing global question. Refugees may be prevented from entering local labor markets because employers have misperceptions or discriminatory attitudes about refugees' skills and little incentive to gather information to correct these misperceptions or change their attitudes. This has motivated the design of several labor market policies aimed at reducing firms' cost of gaining information about disadvantaged workers to improve these workers' chances of employment and, ultimately, labor market efficiency. In this paper, we use a randomized experiment in Uganda – one of the five largest refugee-host countries in the world – to study the short- and longer-run impact on local firms' willingness to hire refugees after being provided with a skilled refugee worker for free for one week. We find that treated firms hire three times as many refugees than firms in the control group eight months after the experiment. Data collected immediately after the experiment further show, consistent with a simple Bayesian learning model, exposure to a refugee led firm managers to update their beliefs about refugees' skills in general. Yet, in the short-term, firms' willingness to hire refugees, proxied by their willingness to offer a short-term job with a (generic) refugee, did not change on average. To investigate mechanisms for why exposure caused some firms to update their beliefs about refugees' skills, and be willing to hire them, while others became less inclined to do so, we use a causal forest approach to estimate treatment heterogeneity. The algorithm identifies two predictors: employers' initial attitudes toward refugees and refugee workers' attitudes toward locals. We use these results to explore the importance of matching attitudes by estimating the variation in the treatment effect across four groups of employer-refugee pairs, distinguished by the attitude of the employer toward refugees and the attitude of the refugee toward locals. In line with a literature in social psychology, we find that positive matches, i.e., firms with a positive attitude toward refugees who were (randomly) matched with a refugee with positive attitudes toward locals, resulted in a substantial increase in firms' willingness to hire a (generic) refugee worker, while negative matches decrease firms' willingness to hire. Finally, we show that the treatment heterogeneity documented in the short-run, also helps explain the longer run results in real-world hiring. Our findings have important policy implications. Short-term exposure interventions can result in longer-term increases in employment for disadvantaged groups, but the size of this effect depends on the initial match quality.

†IIES, Stockholm University, Job Market Candidate. francesco.loiacono@iies.su.se.
§University of Maastricht. m.silvavargas@maastrichtuniversity.nl.
We are deeply grateful to our supervisors: Tessa Bold, Jon de Quidt, Andreas Madestam, and Jakob Svensson, for their continuous support, supervision, patience and encouragement. We are also grateful to Emanuele Colonnelli, Eliana La Ferrara, Nilesh Fernando, Edoardo Teso and Anna Tompsett. We thank IPA's Peace & Recovery Fund, PEDL, JPAL Jobs Opportunity Initiative, SurveyCTO, the Mannerfelt and the Siamon Foundations for funding. Josh Bwiira and Apollo Tumusiime provided outstanding research assistance. We thank James Naboth Ahimbisibwe, Robert Hakiza, Paul Kithima, Julie Ngabirano, and Fred Wanakina for their support throughout the project. This study was approved by the Uganda National Council for Science and Technology (protocol SS 5039), and the Mildmay Uganda Research Ethics Committee (Protocol 0503-2019). This study is registered in the AEA RCT Registry under the unique identifying number: AEARCTR-0006437. All errors are our own.

1. Introduction

Immigrants, especially refugees, constitute one of the world's most vulnerable populations. Among other things, they are more likely to be unemployed, leading to a loss of potential talent and a cost to society. The integration of refugees into the labor market can fail for a number of reasons. Refugees may lack the necessary human capital. They may also face entry barriers, because their abilities and skills are largely unknown to the employers, who may perceive them as low, and refugees' culture and norms may differ from those of the destination country, thus increasing the risk that negative attitudes affect the interaction between native employers and refugee workers. With a sufficiently large native labor supply, an individual firm has little incentive to gather information to correct these misperceptions, even if all firms would benefit from a more skilled labor force. This has motivated the design of several labor market policies, including internships and hiring subsidies, aimed at reducing firms' cost of gaining information about disadvantaged workers, such as refugees, to improve their chances of employment and ultimately labor market efficiency.

In this paper, we use a randomized experiment in Uganda to study the short- and longer-run impact on native owned and managed firms' willingness to hire refugees after being provided with a skilled refugee worker for free for one week. Uganda is an ideal setting to investigate the labor market integration of refugees. Not only is it one of the largest refugee-host country in the world, but refugees are allowed to move freely within the country and look for jobs, thus allowing us to focus on the importance of intergroup contact in the workplace.

To this end, we began by testing the practical skills of a sample of 552 refugees in the manufacturing and services sectors in Kampala, the capital of Uganda. We chose sectors typically associated with regular employment, as in Alfonsi et al. (2020) and Bandiera et al. (2021), including tailoring, food processing, hairdressing, and other light manufacturing and service sectors. About 70% of the refugees in our sample have work experience in at least one of these sectors. On average, they have almost 5 years of experience in the tested occupations. We ran the test in collaboration with the Directorate of Industrial Training, the agency established by the Ministry of Education to be in charge of the vocational education curriculum in Uganda, and two large refugee-led NGOs based in Kampala.

After completing the tests, we randomly paired each refugee worker with a sample of Ugandan employers, stratifying on the occupation of the refugee. Treated firms were subsidized to offer a week-long internship for free to the paired refugee worker whereas control firms were not. We find a large effect: treated firms hire almost three times as many refugees as firms in the control group. To explain the result, we use a simple Bayesian learning framework, where native employers have downward-biased prior beliefs about refugees' skills (because of inexperience). The model predicts that the internship would, on average, lead to positive belief updating about refugees' skillset and increased labor demand for refugees. Consistent with the model, we first show, using the refugee test data, that native managers do indeed have negatively biased priors regarding the skills of the refugee workers at baseline. We then turn to the short-term outcomes of the experiment. We show, consistent with the prediction from the simple Bayesian model, that exposure to a refugee worker through the one-week internship leads firm managers to update their beliefs about refugees' general skills. Yet, firms' willingness to hire a new refugee does not increase on average.

To investigate the mechanisms for why exposure to a refugee worker caused some firms to update their beliefs about refugees' skills, and be more willing to hire them, while others, if anything, became less inclined to do so, we take an agnostic empirical approach and estimate the conditional average treatment effect using a causal forest algorithm (Athey and Wager (2019); Wager and Athey (2018); Davis and Heller (2017)). The method allows us to determine which baseline characteristics are significantly more likely to be associated with heterogeneous treatment effects in the data. The algorithm identifies two predictors: employers' initial attitudes toward refugees, in terms of how supportive they are towards the labor market integration of refugee workers, and refugee workers' attitudes toward locals, in terms of how disenfranchised refugees feel with respect to native Ugandans. We explore the importance of the initial attitudes in the employer-refugee match by estimating the variation in the treatment effect across four groups, distinguished by the attitude of the employer toward refugees and the attitude of the refugee they are matched with toward locals.

We find that firms with a positive attitude toward refugees who are (randomly) matched with a refugee with positive attitudes toward locals, substantially increase their willingness to hire a (generic) refugee worker a week after the experiment ended. In particular, treated firms are 12.3 pp (or 17% at the mean) more willing to hire a refugee compared to the control group. By contrast, firms with negative attitudes toward refugees who are matched with refugees with similar negative attitudes toward locals decrease their willingness to pay by 19.7 pp (equivalent to a 28% decrease). We interpret these findings through the lens of work in social psychology. While Allport (1954) classical contribution on contact theory predicts that intergroup contact should improve the attitudes of the majority in-group (the firms) and increase the willingness to interact with members of the out-group (the refugees), more recent research emphasizes that the intergroup contact can be either positive or negative (Dijker (1987)). Specifically, negative contacts make inter-group differences more salient, inducing a general avoidance of future contact (Paolini et al. (2010); Barlow et al. (2012); Meleady and Forder (2019)). The quality of the interaction therefore affects firms' willingness to hire workers from the minority group going forward (Lepage (2022)).

Finally, and crucially, we find that the one-week exposure intervention had a substantial impact on actual hirings, with a larger effect in the sub-group of firms that initially had a positive attitude toward refugees and were (randomly) matched with a refugee with positive attitudes toward natives. The effect we estimate can be interpreted as an externality: a match with a refugee with a positive attitude toward locals increases the firm's willingness to hire refugees in general, especially so when the firm manager's initial attitudes toward refugees are also positive. Attitudes are complementary and reinforce the effect of contact on the workplace.

Taken together, our findings have important policy implications. We show that a short-term exposure intervention can result in longer-run increases in employment for an especially vulnerable group like refugees, but that the size of the effect depends on the initial match quality.

We contribute to three strands of literature. First, we relate to work studying the effects of active labor market policies in reducing the entry barriers for disadvantaged workers. Some interventions improve firms' access to information about the quality of job seekers (Bassi and Nansamba (2022); Carranza et al. (2022)), or help workers make their skills more accessible to the employers (Pallais (2014); Abebe et al. (2021); Abel et al. (2020)), or adjust workers' and employers' expectations (Bandiera et al. (2021); Abebe et al. (2022)). By contrast, our intervention targets firms' demand for workers from a disadvantaged group.

Second, we connect to the literature on programs using intergroup contact to foster the integration between different groups (Paluck and Green (2009); Broockman and Kalla (2016); Scacco and Warren (2018); Rao (2019); Mousa (2020); Lowe (2021); Bursztyn et al. (2021); Corno et al. (2022)). Unlike previous research, we experimentally vary intergroup contact and exposure in the workplace.

Third, our paper links to the growing body of work on the labor market integration of refugees and forcibly displaced people (Battisti et al. (2019); Arendt et al. (2021); Fasani et al. (2021); Fasani et al. (2022); see Becker and Ferrara (2019) for a review). While a large majority of papers in this literature focus on rich economies, few studies take place in low- or middle-income economies (Caria et al. (2020); Blair et al. (2022); Baseler et al. (2022)). Furthermore, rigorously evaluated randomized control trials in this area are rare (Schuettler and Caron (2020)). We contribute to this literature by designing and evaluating a labor market experiment in a large refugee-host low-income country, where refugees are legally allowed to seek employment.

The remainder of the paper is organized as follows. Section 2 describes the context and the samples of refugee workers and Ugandan employers. Section 3 details the experimental design and test the randomization protocol. Section 4 outlines the specification used in the analysis as well as describes the main outcomes of the paper. Section 5 reports the results of the experiment. Section 6 discusses the results. Finally, Section 7 concludes the paper.

2. Institutional setting and Samples

In this section, we explain why Uganda is the most well-suited environment where to ask our research question. First, we describe the institutional environment of Uganda as a refugee-host country. Second, we describe in details our data and how we selected the participants to our experiment.

2.1. **Institutional setting.**

The refugee policy. Uganda is the largest refugee host-country in Africa and one of the 5 largest in the world. Uganda opened its borders to 7,000 refugees from Poland already during the Second World War (Lwanga-Lunyiigo (1993)). Since then, it has always supported an open-door policy. Today, Uganda is considered to be one of the most welcoming refugee-host country in the world.[1] Currently, it hosts approximately 1.5 million refugees, the majority of

[1]"As Rich Nations Close the Door on Refugees, Uganda Welcomes Them", *New York Times*, 2018.

whom comes from South Sudan, the Democratic Republic of Congo, Somalia, and Burundi.[2] The Ugandan Refugees Act 2006 and its subsequent amendment in 2010 allow refugees to move freely within the country. Refugees can look for employment opportunities, and share access to education, health, and other basic services with the local communities. As shown by the Center for Global Development Uganda has one of the most open policies towards refugees' rights, both de jure and de facto, and at similar levels than many OECD countries (Ginn et al. (2022)).

While the great majority of the refugees live in settlements, shared with the host communities and located in rural areas, approximately 8.5% are registered as dwellers of Kampala. Our experiment takes place in this city, as it hosts 44 percent of all business establishments and almost 50 percent of non-agricultural jobs in Uganda (Sladoje et al. (2019)), and therefore the location where most of the skilled refugees belonging to our sample look for employment opportunities (Figure 1, Panel A). Approximately 70% of the population of refugees residing in Kampala are of working age (aged 18-59). Overall, approximately 15% of the total refugees of working age reside in Kampala (Panel B).

The latest national household survey conducted in 2018 shows that 56% of Ugandans aged 15 to 65 have a job, while the unemployment rate is equal to 11%. Conversely, refugees' unemployment rate is more than three times as large as the natives' one.

2.2. **Samples.**

Refugees. Our sample of interest is composed by skilled refugee job-seekers living in Kampala. To the best of our knowledge, there are no publicly available datasets on individual refugees' characteristics and their location in Uganda. Therefore, we created a collaboration with two local refugee-led NGOs, which have access to a wide population of refugees in Kampala. Thanks to their assistance, we listed 1,088 refugees with the following characteristics: they are not already in permanent employment and are actively searching for jobs at the time of our data collection. Furthermore, we required them to own employable skills in vocational sectors.[3]

In order to verify their skills, we invited a sample of 977 refugees to do a test, and 552 showed up.[4] In partnership with the Directorate of Industrial Training (DIT) and a large vocational institute in Kampala, we organized one examination week during the second half of April 2021. During this week, DIT official examiners tested all the refugees that showed up among those whom we invited, using the DIT's national curriculum.

[2]https://data.unhcr.org/en/country/uga, portal accessed in December 2022.

[3]At listing, we asked them to list the three most important skills they think they possess and would be ready to be tested on. Figure 2 shows the list of most preferred skills, by whether the refugees attended or not the test.

[4]We dropped refugees who revealed not to be skilled enough to pass a practical skills test, such as the one we were offering as well as refugees who were skilled in sectors that did not reach a critical number for the test to take place (5). Compared to the refugees who did not show up at the test, our sample is composed of more experienced and skilled workers, who were both more motivated to get an internship at a local firm and were also more willing to accept a lower wage. Futhermore, they are more likely to have learnt their skills outside Uganda (see Figure 5).

The test focused on the practical skills of the workers and varied in length, depending on the occupation chosen by the candidate. For instance, hairdressers were asked to perform hair style on a client, chefs to prepare and serve a beef stew, tailors to produce a short-sleeved shirt, and so forth. Table A1 in the Appendix shows which skill was tested for each occupation.

Examiners, who are typically experts in each specific sector, scored the performance of each candidate on a 0 to 100 basis, following the national guidelines provided by the DIT. Candidates who score at least 65 successfully pass the test. In the Appendix, we show a picture of an example of a testing day (Figure 3). Of the 552 refugees that showed up at the test, only 11 people failed the exam, and therefore did not obtain a certificate. For this reason, we drop these workers and focus on the ones who passed the test (541). Due to a second wave of covid, we paused the project until September 2021. However, we successfully tracked 527 of the original sample (see our detailed timeline in Figure 4).

In order to compare refugees in Kampala with natives residing in the same city, we use data from the latest Ugandan National Household Survey, conducted in 2018 by the National Bureau of Statistics (UBOS) in collaboration with the World Bank. Table A2 shows that refugees in our sample are more likely to be unemployed and less likely to have a job at baseline, compared to Ugandans living in Kampala, in spite of being more educated. Conditional on being employed, they also earn significantly less than natives.

Firms. To construct a sample of employers, we sampled and conducted a baseline survey with 1,196 firms active in selected sectors in Kampala, using a random walk sampling procedure.[5] Figure 6 maps the location of the firms who belong to our baseline sample. Of these, 535 fulfilled the two criteria for inclusion into our sample: they were owned and/or managed by a Ugandan and they were willing to hire a refugee worker, at least for free, for a period of one week.

Our intervention consisted in matching 325 firms to host an internship of one week with one refugee worker. The remaining 210 compose our control group of firms who did not match with a refugee worker. To assess the impact of the intervention, we conduct two follow up surveys. A first one took place about a month after the matching intervention. For this interview, we tracked 525 firms (attrition is balanced between treatment and control, see Table A3, columns 1 and 3). For the second one, which took place approximately 8 months after the intervention, we collected longer term follow-up data using phone calls from the 474 firms we managed to reached to. Table A3 assesses attrition at the second follow-up in columns 2 and 4.

Our sample of firms is positively selected compared to the average firms in similar sectors in Kampala, along different dimensions. Table A4 compares the characteristics of the firms belonging to our sample and the ones of firms interviewed in the Manpower survey conducted by UBOS in 2016. Our firms are slightly larger, both in terms of employees and revenues. They are more likely to be owned by higher educated people and are more likely to keep

[5]We randomly select a set of neighboring parishes for each day of data collection, based on the Uganda Census of Businesses conducted in 2010. The team leader chooses a landmark and select randomly the directions the data collectors are supposed to take to look for respodents. We halted the data collection for a week in October following three terror attacks in the city of Kampala, and we resumed when the situation normalized.

management books. Additionally, they have been operating for a longer period of time. These differences are not surprising, as our firms stated that they are willing to expand in the near future, whereas the representative firm in the Manpower survey is significantly less likely to plan to hire new workers in the future.

3. Experimental Design

The main aim of the experiment is to increase firms' demand for refugees by changing their beliefs about refugees' skills. The treatment we study is one short-term, fully subsidized internship with one skilled refugee worker. This section has two parts. First, we describe in details the implementation of the experiment, that is how we selected the sample of firms and how we assigned employers to treated and control groups. Second, we outline a simple conceptual framework that we will use to guide the interpretation of the results of the experiment.

3.1. **Protocol.** We begin by randomly pairing refugees and employers, conditional on the occupation of the refugee worker (see Figure 7 for a summary of the randomization design). For example, refugee cooks match with owners of restaurants, beauticians and hairdressers with owners of beauty salons and coiffeurs, and so on. In each pair, during the baseline survey of the firms, employer i evaluates the CV of refugee j, which we constructed using the information from the baseline of the refugee workers. The couple is pre-assigned to a treatment or a control group. Table 1 reports results from a balance test of characteristics between treated and control firms in the full sample (Panel A) and in the exposed sample (Panel B), where the exposed sample is composed of the firms whose treatment actually took place (see below for a more detailed discussion).

Whether a firm enters a treatment group first depends on the employer's "willingness-to-hire" the refugee worker. In a nutshell, while their status is randomly determined, employers express their preferences to select into the treatment. To elicit this willingness, we measure the employers' willingness-to-pay (WTP) using a variation of the Becker-DeGroot-Marschak (BDM) elicitation method called "Multiple Price List" (Becker et al. (1964); Burchardi et al. (2021)). The method consists of a series of take-it-or-leave-it offers, where the price offered to pay increases at each step. We inform the employers that the "price" have already been decided and is in a sealed envelope which the team would open at the end of the elicitation procedure. We do not inform them about the distribution of this price, but we tell them that the price is between 0 and $100,000 UGX$.

We start by eliciting the employer's WTP to hire a hypothetical local Ugandan worker. For this purpose, we show a CV of one hypothetical worker, a man or a woman, possessing the same characteristics of the real refugee worker randomly assigned to the firm (Figure 8). We carefully explain that the worker is hypothetical, inviting the employer to imagine that a worker like the one we are showing is looking for a job at the firm (see script in the Appendix). We use this measure to introduce the respondent to the concept of WTP to hire a worker. We teach the employer the concept of a "random wage" and we make sure that the procedure is clear, by asking comprehension questions at the end of each elicitation. We do not vary the

Table 1. Randomization balance

	Treatment			Control			
Variable	N	Mean	SD	N	Mean	SD	Diff.
Panel A: Full sample							
Employer's education (years)	325	10.853	3.530	210	10.773	3.625	-0.112
Employer is a woman	325	0.563	0.497	210	0.581	0.495	-0.063**
Age of the employer	325	34.929	8.572	210	33.848	8.027	1.665**
Firm age	325	7.640	6.659	210	8.086	6.627	-0.321
Firm is formal	325	0.182	0.386	210	0.190	0.394	-0.015
Has a vacancy	325	0.449	0.498	210	0.371	0.484	0.077*
Desires expand in the future	325	0.852	0.355	210	0.871	0.336	-0.033
Employees at baseline	325	2.434	3.137	210	2.581	3.169	0.216
Ever offered internships (any worker)	325	0.646	0.479	210	0.552	0.498	0.087**
Ever hired a migrant	325	0.351	0.478	210	0.376	0.486	-0.022
Ever hired a refugee	325	0.178	0.383	210	0.171	0.378	0.005
Met refugee job seekers, past month	324	0.145	0.353	210	0.181	0.386	-0.025
Beliefs about refugees' test score	325	65.052	14.501	210	62.705	16.013	2.126
Law should allow refugees' employm.	325	0.923	0.267	210	0.924	0.266	0.006
Locals should have priority to jobs	325	3.388	1.249	210	3.305	1.299	0.104
WTP at baseline	325	17.077	20.486	210	16.881	17.646	0.916
Panel B: Exposed sample							
Employer's education (years)	182	10.799	3.684	210	10.773	3.625	-0.189
Employer is a woman	182	0.582	0.495	210	0.581	0.495	-0.040
Age of the employer	182	35.253	8.767	210	33.848	8.027	2.035**
Firm age	182	7.742	6.546	210	8.086	6.627	-0.347
Firm is formal	182	0.181	0.386	210	0.190	0.394	-0.009
Has a vacancy	182	0.423	0.495	210	0.371	0.484	0.068
Desires expand in the future	182	0.863	0.345	210	0.871	0.336	-0.016
Employees at baseline	182	2.615	3.497	210	2.581	3.169	0.425
Ever offered internships (any worker)	182	0.643	0.480	210	0.552	0.498	0.093*
Ever hired a migrant	182	0.357	0.480	210	0.376	0.486	-0.014
Ever hired a refugee	182	0.198	0.399	210	0.171	0.378	0.034
Met refugee job seekers, past month	182	0.137	0.345	210	0.181	0.386	-0.029
Beliefs about refugees' test score	182	64.390	14.241	210	62.705	16.013	1.455
Law should allow refugees' employm.	182	0.934	0.249	210	0.924	0.266	0.019
Locals should have priority to jobs	182	3.429	1.276	210	3.305	1.299	0.104
WTP at baseline	182	17.445	20.724	210	16.881	17.646	1.235

order of the CVs. That is, all the employers first evaluate the profile of the hypothetical worker before the one of the real worker.

Subsequently, we elicit each employer's WTP twice for one, randomly chosen, refugee worker.[6] We elicit the first WTP right after showing a document with the profile of the candidate for a one-week internship.[7] Furthermore, we tell employers that they can hire the worker at any time in the 4 days following the interview. Firms who are not willing to hire the matched refugee worker report different reasons, with more than half mentioning lack of work as a reason why they are not interested in hiring the refugee (see Figure 9). Conditional on the employer's WTP being positive or equal to 0, we then conduct a new WTP elicitation. After this first elicitation, the research team communicates to a subset (165) of the treated employers that the refugee worker pursued a certificate of vocational skills. To measure whether the certificate affects employers' WTP to hire the worker, we elicit it a second time. We do not show the remaining employers any additional information about the refugee worker. However, our field officers make a more flexible offer to all employers, thus providing the firms with the chance to hire the worker in the next 8 days. See Figure 11 for the original experimental design.

At the end of the second elicitation, we extract a "random wage", W, from a sealed envelope. The random wage determines the outcome of the exercise. Specifically, if $P \geq W$, the employer can hire the refugee worker, otherwise she cannot. In practice, though, we have full control of the randomization procedure and extract only two prices: $W = 0$UGX and $W = 100,000$UGX.[8] Figure 12 shows the (inverted) demand function for a refugee worker in our sample.

If the firm in the treated couple is not interested in hiring the refugee worker we propose (i.e., if the WTP for that specific worker is below 0), we randomly assign the refugee worker to a new firm.[9] The employers with a negative WTP select out of the experiment. We re-iterate the process until we obtain the WTP for all treated refugees.

Approximately 45% of the 1,196 firms interviewed at baseline have a non-negative WTP to hire a refugee worker (see Figure 13). The remaining firms are either not interested in hiring any worker (approximately 35%) or interested in hiring a worker only if Ugandan (about 20%).

Finally, we facilitate the meeting of the treated firm-refugee pair. Field officers set appointments a few days before the agreed starting day of the internship. The team meets the refugee workers at a pre-specified location, which is at walking distance from the firms they are supposed to work for. Importantly, while setting the appointments, the team does not share

[6] Since we have more firms than refugees, multiple employers in the control group may see the profile of the same refugee

[7] The document is a one-page CV containing basic demographic information (a picture of the worker, gender, age, current address and years since moved to Kampala), years of work experience in the selected occupation and knowledge of languages (see Figure 10).

[8] Extensive pilot suggested that the 100,000UGX wage was an unreasonable price for an internship of only one week in the Ugandan SME context.

[9] Younger refugees and refugees who report to speak a better English are more likely to match earlier compared to the rest. By "matching earlier" we mean that the employer(s) they are paired to are more likely to report a non-negative WTP. Both refugees assigned to treated couples and those assigned to control ones are matching with a similar success rate. For more details, see Figure 14.

any information about the firm with the refugee worker. This means that the decision of the refugee worker to show up at the appointment does not depend on the characteristics of the firm. In other words, whenever a refugee shows up at the appointment, the firm takes up the treatment, i.e. the internship takes place. If the refugee fails to show up, the internship does not take place.

When invited to the introductory meeting at a pre-specified location nearby the firm's premises, about 56% of the refugees came. As a consequence, about half of the firms assigned to the treatment group were actually treated (in the sense of receiving a refugee intern). Unsurprisingly, though, conditional on area fixed effects, the sample of firms which receives the worker is balanced in terms of random assignment and has similar characteristics to the sample of firms which did not receive the worker (see Table 2). In section 9 we discuss the determinants of take-up among refugee workers.

3.2. **Conceptual Framework.** In this subsection we provide a simple conceptual framework to interpret the experiment and guide the interpretation of the results. The experiment investigates how exposure, based on observing one refugee for one week, affects the employer's beliefs about refugees' abilities and her willingness to hire new refugees. Suppose that the worker's output contains information regarding the refugee group's mean ability, θ and an individual component ε: $a = f(\theta, \varepsilon)$. If hired by the employer, the worker can produce a signal regarding her ability: $s = a$. The employer cannot observe the average group component, but has some prior beliefs about it. Given her inexperience with refugee workers, the employer's prior is biased: $m_0 < \theta$. The employer's willingness to hire a refugee is a function of the initial beliefs about θ. Furthermore, her utility depends on the expected marginal profit from hiring one refugee. Suppose, finally, that firms' profits depend on the worker's output. Given these assumptions, exposure should have a clear impact: first, it affects the employer's beliefs. Specifically, it should increase them on average towards the true θ. Consequently, exposure should increase, on average, the employer's willingness to hire new refugees.

Guided by this framework, we turn to the data and test the following two hypotheses: working together increases their demand for new refugees and it improves employers' beliefs.

4. Outcomes and Specification

We measure employers' beliefs regarding the ability of refugee workers at baseline by asking the following question: *"Workers can undertake a modular assessment on some specific skills. The assessment, called "Non-Formal", tests workers' practical skills in specific occupations. At the end of each assessment, they can receive a modular transcript issued by the Directorate of Industrial Training. The modular assessment reports a score associated to the performance of the worker during the test. The score ranges between 0 and 100. The threshold to pass the test is 65. Suppose a refugee job seeker, whom you do not know, does this test for the first time. What is the score you would expect him or her to achieve?"*. We elicit the employers' beliefs about Ugandan workers by asking the following question: *"Suppose a typical Ugandan job seeker, whom you do not know, does this test for the first time. What is the score you would expect him or her to achieve?"*. We randomize the order of the questions such that some employers

Table 2. Firms' take-up of the internships

Variable	Match	No match	Control	p(Matched=No)	N
Employer's education (years)	10.799	10.921	10.773	0.761	535
	(3.684)	(3.336)	(3.625)		
Employer is a woman	0.582	0.538	0.581	0.074	535
	(0.495)	(0.500)	(0.495)		
Age of the employer	35.253	34.517	33.848	0.343	535
	(8.767)	(8.329)	(8.027)		
Firm age	7.742	7.510	8.086	0.688	535
	(6.546)	(6.821)	(6.627)		
Firm is formal	0.181	0.182	0.190	0.769	535
	(0.386)	(0.387)	(0.394)		
Has a vacancy	0.423	0.483	0.371	0.417	535
	(0.495)	(0.501)	(0.484)		
Desires expand in the future	0.863	0.839	0.871	0.548	535
	(0.345)	(0.369)	(0.336)		
Employees at baseline	2.615	2.203	2.581	0.243	535
	(3.497)	(2.602)	(3.169)		
Ever offered internships (any worker)	0.643	0.650	0.552	0.929	535
	(0.480)	(0.479)	(0.498)		
Ever hired a migrant	0.357	0.343	0.376	0.944	535
	(0.480)	(0.476)	(0.486)		
Ever hired a refugee	0.198	0.154	0.171	0.255	535
	(0.399)	(0.362)	(0.378)		
Beliefs about refugees' test score	64.390	65.895	62.705	0.423	535
	(14.241)	(14.832)	(16.013)		
Law should allow refugees' employm.	0.934	0.909	0.924	0.275	535
	(0.249)	(0.288)	(0.266)		
Locals should have priority to jobs	3.429	3.336	3.305	0.438	535
	(1.276)	(1.216)	(1.299)		
WTP at baseline	17.445	16.608	16.881	0.966	535
	(20.724)	(20.242)	(17.646)		

Note: Successful matches (*Match*): 182 firms; Not successful matches *No match*: 143 firms; Control group: 210 firms. First, second and third columns report group means. Fourth column reports p-value of a t-test of equality of coefficients between the group of *Match* and *No match* firms from a linear regression where Variable y is regressed over an indicator equal to 1 for *Match* firms, an indicator equal to 1 for *No match* firms, strata and area fixed effects.

get to see first the question about refugee job-seekers and then the one about Ugandans, and vice versa. We can compare the employers' beliefs with the actual scores obtained by the refugee workers. We can additionally compare their beliefs regarding Ugandan workers to a non-random sample of Ugandan workers who took the same test in the last 2 years at the same test center we worked with. The exact scores are not available, but we use the midpoint of the bins used by the DIT to provide a final result on the test.

Our initial hypothesis is that local employers have biased beliefs about the ability of refugee workers. We measure their beliefs by asking them to rate the skills of a hypothetical refugee worker taking a practical skills test such as the one we conducted with our refugee workers. We ask the employers to tell us what score they expect the worker to obtain, on a range between 0 and 100.

Our main outcome of interest is the number of refugees hired after the experiment. We measure this outcome using the last follow-up, conducted eight months after the intervention. We capture this outcome by asking the following question: "Have you offered work on probation to any refugee worker since January 2022. And if yes, to how many?". To explore the mechanism of the experiment, we use the data from the short-term follow-up, collected approximately one month after the intervention. Specifically, we first measure the immediate impact of exposure on the demand for refugees, eliciting the employers' WTP to hire a new, hypothetical, refugee worker.[10]

We chose characteristics of the hypothetical refugee to be desirable for a new worker to come and look for a job at these businesses: the worker is 26 years old (which is equal to the average age of the workforce employed by firms in our sample), has 4 years of experience (twice as much as the average worker in the sample, and equal to the average number of years of experience of the refugees in the sample), and resides in Kampala since 2020. Furthermore, he or she has good knowledge of both English and Luganda (with a self-reported rating of 4 on a scale between 1 and 5). The gender of the new worker depends on that of the previously shown refugee: the firms who had been shown the profile of a man get to see a new male worker, whereas the ones who evaluated the profile of a woman at baseline get to see a new female worker.

Since not all employers are willing to hire a refugee worker at the first follow-up, either because their WTP is now negative (i.e. they require a positive amount of money to hire the worker) or they are simply no longer interested in refugees, we create a dummy variable equal to 1 if the firm is willing to hire the new refugee worker at least for free.

Finally, we measure employers' beliefs using self-reported ratings between 1 and 5 to different statements regarding skills of refugees: the employer's beliefs about the hard (e.g. theoretical abilities, practical skills and actual unit-performance at work) and the soft skills (e.g. time management, team work and work ethics) of a generic refugee worker who may come and look for a job in the future; and beliefs regarding how trustworthy and respectful refugee workers are.[11]

In order to study whether the intervention had any impact on these outcomes, we run the following specification:

[10]Employers were not initially aware that the profile was the one of a hypothetical worker, but we revealed it soon after the elicitation exercise was complete.

[11]We chose this set of skills after extensive piloting exercises with firms similar to the ones belonging to our sample. Specifically, we asked pilot firms to rank workers' skills in order of importance for the success of a business like their own.

$$y_{i1} = \beta_0 + \beta_1 T_i + y_{i0} + X_i'\delta + \varepsilon_i, \tag{4.1}$$

where T_i is a dummy equal to 1 for firms assigned to the treatment group and X_i is a matrix of the randomization strata (the occupations of the refugee workers). In some specifications, it includes area fixed effects, to reflect the imperfect compliance caused by the refugees not showing up at the internships. We discuss the issue of imperfect compliance more in details in Section 9. In a nutshell, refugee workers living further away from where the firms are located are significantly less likely to show up at the internship. Whenever possible, we control for the baseline value of the outcome y or its pre-intervention one (therefore, we run an ANCOVA). Standard errors are clustered at the refugee level, to reflect the experimental design whereby the same refugee might have been shown to multiple firms.[12]

4.1. **Initial beliefs.** Figure 15 shows two things. First, employers' believe that Ugandan job seekers are significantly better than refugee ones. While on average employers believe that Ugandans score 70, they believe that refugees do not pass the test, by assigning an average score of 63. Second, their beliefs are biased downwards, and this is particularly true in the case of the refugee workers. Our refugee workers' actual score on the test is equal to 84.

Taken together, these findings show that Ugandan employers have biased beliefs regarding the ability of refugee workers, and this thus supports the initial hypothesis of our conceptual framework.

4.2. **The internship.** A total of 182 internship took place, but we successfully tracked 179 firms at the first follow-up. The median duration of the internship was 7 days, in line with what employers and workers agreed on. During the internship, employers assigned workers both simple and complex tasks (where complexity is measured using a self-reported scale between 1 and 5 collected for each firm-specific task listed at baseline). About 40% of the employers paid their interns on average 19,000UGX (about 4.5USD) for the full week (even if the worker in most cases had not asked to). On average, each intern worked for 7 hours a day and managers at the firm spent more than 5 hours supervising the intern every day. The employers did not think that the supervision was too complex (rated on average 2.5 on a scale between 1 and 5), and communication was not difficult either (on average rated 3). Firms seem quite satisfied with the experience (a median rating equal to 4). Overall, two thirds of the firms who did the internship are willing to re-hire the same worker. About 7 workers were hired (or 3.9% of the total number of interns). The vast majority of employers (70%), finally, recommended or would recommend the worker to another firm (Table 3).

[12] In the original study design, before eliciting their WTP to hire the refugee worker, we showed a subsample of the treated firms the refugee's certificate of skills obtained after the test. The results on the two treatment arms are both positive and significant, but not statistically distinguishable one from another. In the Appendix, we report the original design in Figure 11. Furthermore, we re-run specification 4.1 using two dummies instead of one, and report the results in Tables A13 and A14:

$$y_{i1} = \beta_0 + \beta_1 T_1 + \beta_2 T_2 + y_{i0} + X_i'\delta + \varepsilon_i$$

Table 3. Descriptives of the internships

	Mean	Median	SD	Min	Max	N
Agreed days of internship	7.419	7	2.994	1	30	179
Completed days of internship	5.324	7	2.847	1	14	179
Internship was extended	0.101	0	0.302	0	1	179
Hours worked by intern each day	7.331	8	2.637	0	12	179
Intern asked to be paid	0.078	0	0.269	0	1	179
Intern was paid during internship	0.425	0	0.496	0	1	179
Intern total payment ('000UGX)	19.730	10	21.113	0	140	74
Max tasks difficulty	3.229	3	1.116	1	5	179
Intern supervised by manager	0.911	1	0.286	0	1	179
Daily firm-hours spent in supervision	5.771	5	4.135	0	20	179
Supervised more than other workers	0.571	1	0.497	0	1	133
Rate how demanding superv. this worker	2.553	2	1.250	1	5	179
How hard communic. [1=Easy, 5=Hard]	3.335	3	1.302	1	5	179
Rate overall experience with worker	3.564	4	1.227	1	5	179
Rate relationship with other employees	3.632	4	1.228	1	5	133
WTP re-hire same, non-neg.	0.676	1	0.469	0	1	179
Intern was hired	0.039	0	0.194	0	1	179
Exchanged phone numbers	0.363	0	0.482	0	1	179
Intern recommended to other firms	0.134	0	0.342	0	1	179
Would recommend worker to other firms	0.709	1	0.455	0	1	179

Note: This data comes from the sample of treated firms whose internship took place (N=182), less of employers whom we did not manage to track at follow-up 1.

Taken together, these descriptive statistics show that the internships were short but intense, with the worker present at the business premises for 7 hours, 5 of which the employer spent them supervising the worker. Among those firms with at least 1 employee, the employer spent more time supervising the intern than any other employee.

5. Results

This section focuses on the main results of our study. Here, we report the effect of the treatment on our core outcomes: number of refugees hired and learning.

We report two separate sets of results. In the first, using the full sample of firms, we show the results of the experiment, that is, the intention to treat. In the second, using the sample of exposed firms, we study the effect of exposure.[13]

[13]The core reason to conduct a separate analysis is given by the fact that firms which were promised a worker who never showed up at the appointment may have had a negative effect on firms' beliefs regarding refugees. In fact, the firms were also contacted on the day of the appointment. Therefore, once the refugee worker failed to show up, the firms' complaints were unhappy with the research firm and the refugees. Examples of comments are *"[...] He was also disappointed with us not giving him a worker"; "He is not happy with us because he told us to match the worker on the day he had agreed with us which was Saturday but up to know he is still waiting for her and no response is getting"; "The*

5.1. **The intention-to-treat effect of the experiment and the effect of exposure.** In this section, we begin by showing the effect of the experiment and the effect of exposure on the demand for new refugees among firms. We will then move to the mechanisms.

Table 4 reports the results of equation 4.1 on the first outcome of interest: total number of refugees hired. We measure this outcome approximately 8 months after the end of the intervention.

TABLE 4. Number of refugees hired

	Dep. var.: Num. refugees hired		
	(1)	(2)	(3)
Panel A: Average Treatment Effect			
Assigned-to-treat	0.069**	0.067**	0.063**
	(0.032)	(0.032)	(0.030)
	[0.029]	[0.034]	[0.034]
N. Firms	474	474	474
Mean Control	0.048	0.048	0.048
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L
Panel B: Effect of exposure			
Exposed	0.079**	0.073**	0.073**
	(0.034)	(0.034)	(0.033)
	[0.021]	[0.035]	[0.028]
N. Firms	343	343	343
Mean Control	0.048	0.048	0.048
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the refugee level in parenthesis. P-values reported in square brackets. Controls: 15 strata (refugees' occupations: tailor, cook, hair-dresser, domestic electrician, craft maker, painter, baker, motorvehicle mechanic, barber, beautician, hotel staff, plumber, carpenter, leather designer, bricklayer, electronics technician, welder and waiter) and 6 area fixed effects (dummies identifying the location of the business premises: Central Kampala, Nakawa, Kawempe, Rubaga, Makindye and Wakiso). Column 3 runs a post-double lasso, always including strata fixed effects but letting the lasso choose among: area fixed effects, gender and age of the employer, and a dummy equal to 1 if the firm has ever offered internships to any worker.

Table 4 shows that a short-term intervention, such as an internship of one week, increases significantly the number of refugees hired by firms, compared to the control group. Panel A

firm owner was very disappointed with the worker who was given a place for internship but didn't show up for work".. In a nutshell, we cannot instrument exposure with the offer of the treatment, because it is not a valid instrument.

shows the intention-to-treat effect of the experiment, using the full sample. Panel B focuses on the effect of exposure, dropping firms who were not treated because the refugee worker did not show up for the internship. The comparison between the coefficients in both panels shows that the effect is concentrated among the exposed sample only, as expected. Furthermore, the effect is very large and equal to almost tripling the total number of refugees hired. Both the specifications with and without area fixed effects (columns 1 and 2) yield virtually the same results.

In order to take into account possible confounding factors arising from unbalanced covariates between samples, column 3 runs a double-selection lasso linear regression, letting lasso choose which covariates should enter the regression. We include covariates that did not balance at baseline (a dummy equal to 1 if the firm has ever offered internships and age of the respondent) and that, similarly to the area where the business premises are located, differ in the exposed sample only (gender of the respondent). Column 3 yields exactly the same coefficients as columns 1 and 2, reflecting the fact that lasso chooses only one of the area fixed effects (i.e. a dummy equal to 1 if the business premises are located in the division of Nakawa, which is the division located the futherst away (i.e. approximately 7km) for the refugee-host districts of Makindye and Rubaga). In the Appendix, Table A5 repeat equation 4.1 using alternatively a Poisson and a Tobit regression and shows that the result is robust to different specifications.

In order to explore the mechanisms, we use the short-term follow-up and investigate whether firms update their beliefs regarding the skills of refugees and whether this affects firms' demand for hypothetical refugees right after exposure.

5.2. **Mechanisms.** We explore the mechanisms of the experiment by studying the effect of treatment on self-reported scales rating refugees' skills. We then analyze how the program affects firms' demand to hire a new refugee about a month after the internship is completed.

Table 5 reports the results on employers' beliefs. We focus on our preferred specification, controlling for area fixed effects. In the Appendix, we replicate this table removing area fixed effects (Table A6) and re-running a new specification using a post-double lasso procedure (Table A7). We find that, on average, the assignment to treatment does not have any effect on employers' learning (Panel A). This is expected given the null or negative effect of some refugees' lack of compliance. Using the exposed sample to determine the effect of exposure, we find that employers update their beliefs: exposure makes them more likely to report a higher rate on refugees' skills, especially soft skills (Panel B). In the Appendix, we show the effect of exposure on each individual skill we ask a rating for (Table A8). Exposed employers are also more likely to rate refugees as trustworthy and showing more respect in the workplace. In column 5, we summarize the effect on learning computing the average standardized effect of the learning outcomes, averaging the effects in columns 1 to 4, estimating a seemingly unrelated regression system

$$Y = [I_n \otimes T]\beta + \mu \tag{5.1}$$

where Y is a vector of n beliefs outcomes and the square matrix $I_n \otimes T$ collects the Kronecker product of the identity matrix and the treatment assignment vector. Following Kling et al. (2004) and Nyqvist et al. (2019), we collect the estimated coefficient $\hat{\beta}_n$ of the treatment effect on outcome n and standardize it by the standard deviation $\hat{\sigma}_n$ from the control group in outcome n to obtain the standardized coefficient $\tilde{\beta} = \frac{1}{n}\sum_{n=1}^{N} \frac{\hat{\beta}_n}{\hat{\sigma}_n}$ reported in column 5 of Table 5. The coefficient is positive and highly significant, suggesting that the internships worked in updating the beliefs of the treated employers.

Table 5. Learning

	Dependent variable:				
	(1) Hard skills	(2) Soft skills	(3) Trust	(4) Respect	(5) Avg. std. effect
Panel A: Full sample					
Assigned to Treatment	0.011	0.126	0.175*	0.094	0.102
	(0.100)	(0.104)	(0.102)	(0.101)	(0.084)
	[0.913]	[0.228]	[0.088]	[0.353]	[0.228]
N. Firms	525	525	525	525	525
Mean Control	-0.000	0.000	0.000	0.000	
Area FE	Yes	Yes	Yes	Yes	
Regr.	OLS	OLS	OLS	OLS	
Panel B: Exposed sample					
Exposed	0.103	0.269**	0.366***	0.197*	0.234**
	(0.118)	(0.123)	(0.114)	(0.119)	(0.098)
	[0.382]	[0.030]	[0.001]	[0.099]	[0.017]
N. Firms	385	385	385	385	385
Mean Control	-0.000	0.000	0.000	0.000	
Area FE	Yes	Yes	Yes	Yes	
Regr.	OLS	OLS	OLS	OLS	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the refugee level in parenthesis. P-values reported in square brackets. Controls: 15 strata (refugees' occupations: tailor, cook, hair-dresser, domestic electrician, craft maker, painter, baker, motorvehicle mechanic, barber, beautician, hotel staff, plumber, carpenter, leather designer, bricklayer, electronics technician, welder and waiter) and 6 area fixed effects. Indices are computed following Anderson (2008), using the following underlying covariates: theoretical skills, practical skills and speed for the index on hard skills (Column 1); work ethics, time management and team work ability for the index on soft skills (Column 2).

Our conceptual framework predicts that employers learn and are therefore more willing to hire new refugees, already right after the experiment. We test this prediction, analyzing the effect of exposure on the firms' willingness to hire a new refugee approximately one month after the internship took place. We interpret this measure as the immediate reaction of firms to the internship program.

For this purpose, we show the profile of a new hypothetical refugee worker at follow-up 1 (see Figure 16). By construction, the new profiles have the same characteristics for all firms (treated and control) in the sample, therefore we can isolate the effect of treatment only.

Not all firms in our sample report a non-negative willingness-to-pay (i.e., some firms are not willing to hire the new worker for any price, including for free). Employers no longer willing to hire a refugee report different reasons, but the greatest majority in both the exposed and the control groups say they do not have enough work or space to accommodate a new refugee (see Figure 17). Very few claims that the refugee is not skilled enough. About a similar percentage report to have been disappointed with the refugee workers. For this reason, our main outcome of interest is a dummy variable equal to 1 if the firm says it is willing to hire the new worker at least for free.[14] While 71% of firms in the control group are willing to hire the worker at a price of 0UGX, we find that treated firms are not more willing to hire a new refugee worker. Table 6 shows that the treatment effect is essentially zero, i.e., we find no evidence that treatment in the full sample (Panel A) or in the group of exposed firms (Panel B) increases firms' demand for a new refugee worker. The estimated standard errors are small, and range between .04 and .049. Notice that the point estimate in the full sample is more than 5 times larger in magnitude than the point estimate in the exposed sample, thus suggesting that there are firms who are considerably more negative than control ones, among employers whose internship did not take place. This is true regardless of the specification we use (columns 1 to 3).

In the Appendix, we report the curves for the demand of a new refugee by treatment status, imputing 0s for the firms with a non-positive willingness to pay. Figure 18 shows that the demand does not shift differentially across the groups, with no difference between the full sample and the exposed one. Table A9 replicates Table 6 using as an outcome the willingness to pay to hire the refugee worker, imputing missing values with zeros. We fail to reject the null hypothesis of an effect of the experiment. One may worry that the reason why we do not find a significant average effect is because treated firms satisfied their demand for workers significantly more than control firms. To check this we investigate whether treated firms are less likely to have a vacancy at follow-up 1. We do not find evidence for this in the exposed sample (see Table A10, whereas there is a significant decrease in the number of firms who say to have an open vacancy in the full sample, which seems to be driven by the firms that do not match with the promised refugee worker.)

In order to investigate what drives some firms to increase their demand while some others if anything decrease it, we take an agnostic approach, run a causal forest algorithm and let the data tell us which covariates are more likely to predict heterogeneous treatment effects. This method will allow us to detect unanticipated results, exploring multiple dimensions of heterogeneity, limiting the risks of p-hacking, especially when the heterogeneity analysis is not pre-specified (Davis and Heller (2017)).

5.3. **Causal Forest.** Causal forest is a machine learning method that allows to predict the heterogeneity in the causal treatment effect. More precisely, it estimates the Conditional

[14]Another reason not to use WTP for the new refugee is that firms may have learnt that refugees would accept a low wage, and therefore are willing to pay a lower wage to hire the worker.

Table 6. Willingness to hire a new worker

	Dep. var.: WTP≥ 0		
	(1)	(2)	(3)
Panel A: Full sample			
Assigned to Treatment	-0.017	-0.021	-0.019
	(0.041)	(0.041)	(0.041)
	[0.688]	[0.610]	[0.644]
N. Firms	525	525	525
Mean Control	0.709	0.709	0.709
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L
Panel B: Exposed sample			
Exposed	0.003	-0.004	-0.003
	(0.048)	(0.049)	(0.048)
	[0.955]	[0.938]	[0.953]
N. Firms	385	385	385
Mean Control	0.709	0.709	0.709
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the refugee level in parenthesis. P-values reported in square brackets. Controls: 15 strata (refugees' occupations: tailor, cook, hair-dresser, domestic electrician, craft maker, painter, baker, motorvehicle mechanic, barber, beautician, hotel staff, plumber, carpenter, leather designer, bricklayer, electronics technician, welder and waiter) and 6 area fixed effects. Both specifications include the baseline value of the WTP.

Average Treatment Effect (CATE), that is the average treatment effect conditional on a vector of baseline covariates:

$$\tau(X) = \mathrm{E}[Y_{1i} - Y_{0i} | X = x]$$

where Y is the outcome of interest and X is a vector of baseline observables. This method emerged with the theoretical work of Athey and Imbens (2016) and Wager and Athey (2018), and the empirical application of the algorithm in Athey and Wager (2019) and Davis and Heller (2017), Davis and Heller (2020). Since then, empirical papers using experiments adopted the causal forest algorithm to investigate heterogeneity in the data (for instance, Carlana et al. (2022); Athey et al. (2021)).

First, we run the algorithm on the exposed sample of 385 observations. Given the small sample size, we train the algorithm growing a large number of trees (200,000). This procedure should guarantee that the confidence intervals are accurately estimated and is recommended by the creators of the algorithm to obtain stable estimates.[15] Furthermore, we use the so-called "honest approach": we split the training sample in half, with only half of the observations used to grow a tree and the other half used to estimate the treatment effect in each leaf, in mutually exclusive sets. As the covariates fed into the causal forest, we choose firms', workers' and matches' characteristics that may affect firms' willingness to hire a new worker. Using our rich data from both the employers' and the refugees' surveys, we construct indices using the first factor from a factor analysis. For each index, we create a dummy equal to 1 if the individual observation has a value larger than the median. Therefore, employers with an index value larger than the median display a high prevalence of the concept represented by the index. We include the following firm- and employer-specific variables, refugee-specific variables and match-specific variables: the employers' experience with hiring a migrant; a dummy equal to 1 if the employer belongs to the major ethnic group of the Baganda; attitudes towards labor market integration of refugees; the perceived cost of learning about refugees' skills; the willingness to expand their businesses; management quality; current size (in terms of number of employees, number of tasks and number of business premises); a dummy equal to 1 if the firm's sector is manufacturing; and beliefs regarding the skills of the matched worker; the workers' ability; attitudes towards Ugandans and Ugandan culture; knowledge of languages; their experience with Ugandan employers in the past; age; country of origin; finally, we include a dummy equal to 1 if the worker lives in the same neighborhood where the business premises are located and if the employer and the worker have the same gender. We describe each variable in more details in the Appendix.

Second, we compute the out-of-bag predicted CATE estimate, that is, the predictions produced by the algorithm using trees that do not include observation i. We use it to identify what covariates are associated with heterogeneity in the treatment effect.

Third, once we have obtained the individual predictions, we split the training sample into two groups with respect to the median: observations with a high predicted CATE, belonging to the top 50%, and those with low predicted CATE, belonging to the bottom 50%.

Finally, we investigate what characteristics are associated with high predicted CATE using two different methods: first, we run a balance test across the two different groups of observations, and correcting the p-value of equality using the method suggested in List et al. (2019). Second, we use a doubly-robust estimator to compute the best linear projector of $\tau(X)$ (Chernozhukov et al. (2018)).

Table 7 reports the results of the balance test. In the Appendix, Table A11 reports the results from the best linear projector estimation. There are only two characteristics surviving the correction of the p-values, and therefore significantly associated with a heterogeneous predicted CATE: the employer's attitudes and the refugee's attitudes.

Finally, Figure 19 depicts a heatmap of the predicted CATE across bins of the indices of refugee's attitudes and firm's attitudes. It shows that the better the initial attitudes of both the firm and the refugee, the more positive is the firm's predicted CATE (colder colors). And viceversa, the worse their initial attitudes, the lower the predicted CATE (warmer colors).

[15]The resulting `excess.error` is negligible and equal to $2.79e-07$.

Table 7. Causal forest, balance table

Variable	Low CATE	High CATE	Diff.	MHT pval
Ever hired a migrant	0.383	0.344	-0.040	0.976
Owner is Muganda	0.705	0.635	-0.069	0.818
Employer's attitudes	0.642	0.839	0.196	0.000
Firm's beliefs	0.430	0.552	0.122	0.192
Employer's perceived cost of learn.	0.528	0.490	-0.039	0.970
Firm's expansion plan	0.269	0.286	0.017	0.918
Firm's quality	0.446	0.521	0.075	0.825
Firm's size	0.523	0.474	-0.049	0.975
Refugee's ability	0.534	0.469	-0.065	0.908
Refugee's attitudes	0.052	0.865	0.813	0.000
Refugee's knowledge of languages	0.161	0.104	-0.056	0.731
Manufacturing sector	0.316	0.339	0.022	0.953
Refugee ever employed by Ugandan	0.275	0.250	-0.025	0.972
Refugee's age	33.565	34.323	0.758	0.951
Refugee is Congolese	0.912	0.849	-0.063	0.499
Employer+worker live in same neigh	0.109	0.120	0.011	0.750
Employer+worker same gender	0.829	0.792	-0.037	0.963

5.4. **Why would the employer's attitudes matter?** To understand why attitudes matter, we return to the conceptual framework and extend it to include the role of first the employer's attitudes, and then to additionally include the role of the worker's attitudes. First, to understand what attitudes means in our context, we begin by explaining how we constructed the indices (see Appendix for a full description). To construct the attitudes of the employers, we construct a dummy equal to 1 if the answer to the following statements are not "Agree" or "Strongly agree": *"When jobs are scarce, Ugandans should have more right to a job than refugees"*. Furthermore, we construct a dummy equal to 1 if the answer to the following question is positive: *"Do you think that refugees should be allowed to work in Uganda?"*. Finally, we run a factor analysis and extract the first factor. Therefore, by attitudes of the employer we mean their attitudes towards labor market integration of refugees. A positive employer is someone who encourages labor market integration of refugees.

On possible way to interpret the role of attitudes among employers is the following. The supervision of a worker is costly. Additionally, an employer devoting time to a worker in probation will have to reduce her attention to more profitable activities. This is likely to be happening in mSMEs like those in our sample, where managers do not fully delegate responsibilities to other workers (Bassi et al. (2022)). Suppose that employers have to exert efforts to learn about the skills of refugees, and that the higher their efforts, the more they will learn. An employer chooses her efforts weighting the benefit of learning about the productivity of refugees (which is a function of the prior beliefs) and the cost of exerting efforts (c). Suppose also that how much efforts an employer exerts depend on her initial attitudes towards refugees, δ. That is, employers with more open views about refugees are more likely to exert more efforts than those with less open views. Conversely, employers with negative views (e.g. those that have a very high value of δ) will be less likely to exert efforts, and will therefore be less likely to learn. These two assumptions together now predict the creation of two groups of employers. Positive ones will exert more efforts to learn and are going to learn more about refugees. Consequently, their willingness to hire a refugee will increase, given that on average initial beliefs are biased. On the contrary, negative employers are less likely to exert efforts and to learn. Therefore, their willingness to hire a refugee should not change as compared to the control group.

5.5. **The role of refugees' attitudes.** The causal forest algorithm predicts that the workers' attitudes are associated with heterogeneous effects in the demand for new refugees among employers. We construct refugee's attitudes as follows. First, we construct dummies equal to 1 if the refugee worker agrees or strongly agrees with the following statements: "*Ugandans' culture is different from my own culture*", "*Ugandans discriminate towards refugees*", "*I assume that in general, Ugandans have only the best intentions*", "*Work between Ugandans and refugees is good for both groups*". We interpret the first factor from a factor analysis on these variables as the sense of belonging that refugees feel in Uganda. A positive refugee is one that feels a tighter cultural proximity to Ugandans and perceives to be more integrated.

In what follows we conceptualize why these attitudes matter. Suppose that refugees' attitudes affect the efforts at work. Refugees with positive attitudes are more likely to exert efforts at work. This affects employer's learning, who therefore update more on refugees' skills as compared to an employer in control. The opposite happens when a refugee with negative attitudes matches with an employer, who in turns does not learn or learn to a much smaller extent as compared to the control group.

This extended framework produces two additional predictions:

1) Employers with positive attitudes matching with workers with positive attitudes exert more efforts to learn about refugees, learn more because the worker is more motivated on the job and therefore learn more about refugees' skills. Given that exposure is a positive experience, the employer's attitudes improve even more, and become more positive. As a consequence, her willingness to hire new refugees uniquivocally increases.
2) Employers with negative attitudes matching with workers with negative attitudes do not learn as much. Given that the exposure is also a negative experience, the employer may become even more negative against refugees. As a result, her willingness to hire a refugee may decrease.
3) What happens in the two mixed groups with opposite attitudes is instead ambiguous. Two different forces are at play: refugees' efforts on the job and employers' efforts on learning. Given that neither of the two prevails, the total effect on learning and the demand for new refugees may not be different from zero.

These predictions are supported by the social psychology literature as well. Specifically, these studies have established the opposite role of positive versus negative contact. Allport (1954) had already warned that the "wrong kind of contact" could exacerbate the perceived differences between groups, "prompting an increase in negative emotions and stereotypes"

(McKeown and Dixon (2017)). More recently, empirical work has shown the polarizing effects of positive versus negative contact (Barlow et al. (2012); Paolini et al. (2010)). Reconciling a learning model with social psychology theories on the effect of contact could help explain our results.

We estimate the effect of exposure across the different groups using the following specification:

$$y_{i1} = \beta_0 + \beta_1 TxPositive + \beta_2 TxMixed + \beta_3 TxNegative + X_i^{'}\delta + \varepsilon_i \tag{5.2}$$

where $TxPositive$ is a an indicator for treated positive employers that matched with a positive refugee, $TxNegative$ an indicator for treated negative employers that matched with a negative refugee, and $TxMixed$ is an indicator variable for treated negative (positive) employers that matched with a positive (negative) refugee. Each coefficient tells us the effect of treatment among a specific match. A test of equality between coefficients tells us whether the effect is significantly different across these groups.[16] Finally, the matrix of controls X_i contains strata, area fixed effects and a dummy equal to 1 if the firm is positive towards refugees in all specifications. In some specifications it includes variables that are unbalanced between full and exposed sample, as well as those unbalanced at baseline randomization, using a post-double lasso linear regression.

Table 8 reports the results of equation 5.2. Positive matches are more likely to cause an increase in the willingness to hire a new refugee worker. The increase varies between 11.5pp and 12.3pp, depending on the specification. In other words, exposure increases the number of employers interested in hiring a new refugee by approximately 17%. Viceversa, when the match is negative, the employer's willingness to hire a new worker reduces by approximately 19pp to 19.7pp, i.e. a reduction of approximately 28%. When testing the equality of coefficients β_1 and β_3, we can reject the null hypothesis that they are equal to each other. The effect on mixed matches is small and not distinguishable from zero.

These results are robust to the method we use to estimate the effect of exposure. Ignoring model selection may lead to invalidate inference (Leeb and Pötscher (2005)). In a nutshell, the finite-sample properties of post-model-selection estimators may not be similar to the respective asymptotic distributions. While it is not yet theoretically clear whether standard errors are not correctly specified once we run a regression post-causal forest, we acknowledge that there are some methods designed to take care of this issue. We therefore use a doubly-robust estimator to re-estimate equation 5.2 and report the results in Appendix Table A12. These results are stronger than the ones reported in column 3 of Table 8. Now, positive matches increase the employers' willingness to hire of about 20pp, that is more than 28% over the mean, while negative matches decrease it by almost 28pp, that is more than 39%. Finally, Figure 20 reports the p-values of β_1 and β_3, as well as the p-value from the test of equality between the

[16]There are two mixed groups, one where the employer has positive attitudes and the refugee worker has negative attitudes, and another one where the opposite is true. Since our conceptual framework predicts that the effect is ambiguous in both these groups, we merge them in one group.

TABLE 8. Short-term demand for refugees by employer's and worker's initial attitudes

	Dependent variable: WTP≥ 0		
	(1)	(2)	(3)
β_1: TxPosit.	0.123*	0.115*	0.123*
	(0.064)	(0.066)	(0.065)
	[0.055]	[0.081]	[0.058]
β_2: TxMixed	-0.014	-0.020	-0.020
	(0.058)	(0.059)	(0.058)
	[0.808]	[0.730]	[0.726]
β_3: TxNegat.	-0.190*	-0.192*	-0.197*
	(0.111)	(0.112)	(0.111)
	[0.089]	[0.089]	[0.076]
p($\beta_1 = \beta_2$)	0.059	0.064	0.046
p($\beta_1 = \beta_3$)	0.014	0.018	0.011
p($\beta_2 = \beta_3$)	0.134	0.148	0.134
N. Firms	385	385	385
Mean Control	0.709	0.709	0.709
Area FE	No	Yes	Yes
Controls	No	No	Yes
Regr.	OLS	OLS	PDS-L

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the refugee level in parenthesis. P-values reported in square brackets. Controls: 15 strata (refugees' occupations: tailor, cook, hair-dresser, domestic electrician, craft maker, painter, baker, motorvehicle mechanic, barber, beautician, hotel staff, plumber, carpenter, leather designer, bricklayer, electronics technician, welder and waiter) and 6 area fixed effects (dummies identifying the location of the business premises: Central Kampala, Nakawa, Kawempe, Rubaga, Makindye and Wakiso). All the specification controls for a dummy equal to 1 if the employers' attitudes are positive (i.e. have an index with value above median). Column 3 runs a post-double lasso, always including strata fixed effects but letting the lasso choose among: area fixed effects, gender and age of the employer, and a dummy equal to 1 if the firm has ever offered internships to any worker.

two coefficients, using randomized-based inference (RBI). The RBI p-values are in line with the ones from the main regressions.

Importantly, the heterogeneous effect across groups of attitudes shows up 8 months after, with real hiring being different across the three groups. Table 9 shows that the effect is concentrated around the group of employers with positive attitudes who match with refugees with positive attitudes. Across the usual three specifications the coefficient is stable but becomes more noisy as we add controls and the p-value in the last column is equal to 0.096.

To explain these findings we take some additional steps. We first investigate whether a similar pattern shows up in the other outcomes that characterize the relationship between the firm and the worker. Using the average standardized coefficients constructed following 5.1 we find that the average positive effect of the exposure is concentrated among the positive matches (Table 10). Table 11 explores the components of learning. We find that this is especially true for

Table 9. Real hiring of refugees by employer's and worker's initial attitudes

	Dependent variable: Number of refugees hired		
	(1)	(2)	(3)
β_1: TxPosit.	0.110*	0.102*	0.095*
	(0.060)	(0.060)	(0.057)
	[0.065]	[0.090]	[0.096]
β_2: TxMixed	0.056	0.056	0.060
	(0.043)	(0.043)	(0.045)
	[0.192]	[0.187]	[0.186]
β_3: TxNegat.	0.081	0.064	0.080
	(0.083)	(0.084)	(0.079)
	[0.328]	[0.451]	[0.313]
p($\beta_1 = \beta_2$)	0.440	0.505	0.610
p($\beta_1 = \beta_3$)	0.781	0.715	0.879
p($\beta_2 = \beta_3$)	0.795	0.940	0.831
N. Firms	343	343	343
Mean Control	0.048	0.048	0.048
Area FE	No	Yes	Yes
Controls	No	No	Yes
Regr.	OLS	OLS	PDS-L

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the refugee level in parenthesis. P-values reported in square brackets. Controls: 15 strata (refugees' occupations: tailor, cook, hair-dresser, domestic electrician, craft maker, painter, baker, motorvehicle mechanic, barber, beautician, hotel staff, plumber, carpenter, leather designer, bricklayer, electronics technician, welder and waiter) and 6 area fixed effects (dummies identifying the location of the business premises: Central Kampala, Nakawa, Kawempe, Rubaga, Makindye and Wakiso). All the specification controls for a dummy equal to 1 if the employers' attitudes are positive (i.e. have an index with value above median). Column 3 runs a post-double lasso, always including strata fixed effects but letting the lasso choose among: area fixed effects, gender and age of the employer, and a dummy equal to 1 if the firm has ever offered internships to any worker.

firms' beliefs regarding hard skills and how trustworthy refugees are (a Wald test of equality of coefficients rejects the null of the coefficients being the same). The magnitude of the coefficients also suggests that the effects are stronger when the match is positive (for instance, the effect among positive matches is between 1.7 and 7 times as large for the hard skills and the soft skills, respectively, 6 times as large for trust and approximately 3 times as large for respect).

Second, we use the data from the internships and show suggestive evidence that the quality of exposure depends on the initial attitudes of both the employer and the worker (figures 21 to 27). These figures report the averages across the three groups of attitudes of different internship's outcomes, as well as different refugees' characteristics.

When the match is positive, employers are significantly more willing to once more hire the same worker, and they rate the overall experience higher compared to firms in negative matches. Furthermore, firms with positive matches found it less demanding to supervise the worker

Table 10. Learning

	Dependent variable: Avg. std. eff.		
	(1)	(2)	(3)
β_1: TxPosit.	0.464***	0.471***	0.458***
	(0.138)	(0.141)	(0.144)
	[0.001]	[0.001]	[0.001]
β_2: TxMixed	0.137	0.135	0.133
	(0.124)	(0.126)	(0.126)
	[0.272]	[0.285]	[0.289]
β_3: TxNegat.	0.059	0.048	0.043
	(0.178)	(0.177)	(0.177)
	[0.741]	[0.785]	[0.808]
p($\beta_1 = \beta_2$)	0.049	0.043	0.054
p($\beta_1 = \beta_3$)	0.067	0.055	0.060
p($\beta_2 = \beta_3$)	0.698	0.663	0.651
N. Firms	385	385	385
Area FE	No	Yes	Yes
Controls	No	No	Yes
Regr.	OLS	OLS	OLS

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the refugee level in parenthesis. P-values reported in square brackets. Controls: 15 strata (refugees' occupations: tailor, cook, hair-dresser, domestic electrician, craft maker, painter, baker, motorvehicle mechanic, barber, beautician, hotel staff, plumber, carpenter, leather designer, bricklayer, electronics technician, welder and waiter) and 6 area fixed effects.

(although not significantly). These findings suggest that the internship went significantly better in the group of employers that matched with positive initial attitudes with workers with positive attitudes.

Furthermore, refugees in the positive matches are also more likely to have been looking for jobs prior to the experiment, applying to more positions and being more successful with Ugandan employers (albeit not significantly so). Higher job offer rates from Ugandan employers among refugees in positive matches also suggest that these refugees may have already had better experiences with Ugandan employers in the past. These second set of findings suggests that refugees with positive attitudes matching with the positive employers were also more motivated in providing a better signal of their ability to their employer during the internship.

Finally, we use our longer term follow-up phone survey to collect the employers' views on some challenges regarding employing refugees, and use it as evidence supporting the mechanisms of our experiment. We ask employers belonging to the control group to what extent they agree with a series of statements, using a scale between 1 and 5. We report the results in two different ways. First, we show the distribution of the ratings for each statement. Then, we rank each statement in terms of the percentage of firms which agree or strongly agree with them (rates equal to 4 and 5 respectively). Figure 28 reports the results of this survey. Panel

Table 11. Learning, single components

	Hard skills			Soft skills			Trust			Respect		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
β_1	0.403**	0.410**	0.435**	0.433**	0.419**	0.437**	0.668***	0.680***	0.665***	0.352**	0.375**	0.368**
	(0.178)	(0.181)	(0.174)	(0.179)	(0.183)	(0.177)	(0.165)	(0.168)	(0.164)	(0.164)	(0.167)	(0.161)
	[0.024]	[0.024]	[0.012]	[0.016]	[0.023]	[0.014]	[0.000]	[0.000]	[0.000]	[0.033]	[0.025]	[0.022]
β_2	-0.012	-0.019	0.001	0.242	0.218	0.256*	0.215	0.219	0.227*	0.102	0.122	0.147
	(0.147)	(0.149)	(0.146)	(0.151)	(0.156)	(0.150)	(0.138)	(0.141)	(0.137)	(0.150)	(0.151)	(0.146)
	[0.937]	[0.900]	[0.993]	[0.112]	[0.163]	[0.087]	[0.120]	[0.120]	[0.097]	[0.499]	[0.421]	[0.314]
β_3	-0.147	-0.173	-0.148	0.125	0.100	0.140	0.174	0.164	0.189	0.084	0.102	0.126
	(0.192)	(0.190)	(0.194)	(0.242)	(0.242)	(0.243)	(0.203)	(0.205)	(0.203)	(0.236)	(0.234)	(0.229)
	[0.443]	[0.363]	[0.444]	[0.605]	[0.679]	[0.563]	[0.392]	[0.425]	[0.351]	[0.722]	[0.665]	[0.581]
1=2	0.045	0.039	0.034	0.361	0.339	0.386	0.018	0.017	0.020	0.196	0.192	0.246
1=3	0.033	0.022	0.022	0.300	0.283	0.314	0.053	0.045	0.060	0.340	0.331	0.377
2=3	0.545	0.489	0.502	0.659	0.654	0.661	0.854	0.806	0.866	0.944	0.935	0.935
N	385	385	385	385	385	385	385	385	385	385	385	385
Area	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Regr.	OLS	OLS	PDS-L	OLS	OLS	PDS-L	OLS	OLS	PDS-L	OLS	OLS	PDS-L

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the refugee level in parenthesis. P-values reported in square brackets. Controls: 15 strata (refugees' occupations: tailor, cook, hair-dresser, domestic electrician, craft maker, painter, baker, motorvehicle mechanic, barber, beautician, hotel staff, plumber, carpenter, leather designer, bricklayer, electronics technician, welder and waiter) and 6 area fixed effects. Indices are computed following Anderson (2008), using the following underlying covariates: theoretical skills, practical skills and speed for the index on hard skills (Columns 1 to 3); work ethics, time management and team work ability for the index on soft skills (Columns 4 to 6).

A shows the distribution of the ratings for each statement and we summarize each statement into the core mechanism we are exploring. Panel B instead ranks each mechanism according to the percentage of firms which agree or strongly agree with each statement. We find that at least 80% of firms agree or strongly agree that refugees' and firms' attitudes (and both of them at the same time) are a relevant factor explaining why firms may not hire refugees. There is also a consistent percentage of firms who believe or strongly believe that refugees need more training before being given a job. Only half of the firms claim that it is hard to give a job to a refugee job-seeker because Ugandan employers do not share the same social networks with them. Overall, we interpret these results as supportive of the main mechanism of our experiment. Namely, attitudes towards the out-group is a crucial factor in hiring refugees, and this idea is additionally supported by local employers.

6. Discussion

This experiment teaches what a government would need to learn if interested in affecting labor market integration of refugees involving the private sector through short-term internships.

First, just about half of all the possible employers will be interested in joining the experiment. This means that firms will be positively selected. We argue that in many encouragement design participants tend to be positively self-selected. If anything, one can interpret our RCT as a selective trial thanks to our willingness to pay to hire exercise, which reveals what firms are truly interested in trying a refugee worker (see Chassang et al. (2012) for a discussion on selective trials). Thanks to our rich data, we can characterize who these participants are. On the positive note, these are firms who are most likely going to be able to offer internships themselves once they start to learn about refugees. Very few firms have ever hired one refugee before our experiment (about 17%). Lack of experience with these workers may explain why employers have uncertain and wrong beliefs about refugees. We show that very short-term internships teach firms about the real ability of refugees and therefore can be used as a tool to integrate refugees increases firms' demand for these type of workers.

Interestingly and crucially, the effect on real hiring is not driven by the same worker we matched the firms with. Table A15 shows the same specification as in Table 4, excluding firms that mentioned that they have hired the matched refugee at some point during the 8 months after the internship. How did these firms start to hire more refugees? One possible explanation are network effects. Table A16 shows that the effect on hiring is concentrated among firms located in divisions of Kampala typically hosting refugees (Makindye and Rubaga).

Given the dimension of the firms belonging to our sample, one concern is that they reduce hiring of Ugandan workers to accommodate new refugee ones. In the Appendix, Table A17 shows that this is not the case. Treated firms are not less likely to hire Ugandans. Therefore, internships do not create displacement effects.

We find that initial attitudes drive the positive effects on real hiring, showing that initial attitudes are complementary for the success of matching. We interpret these findings through the lens of social psychology. Unlike this literature, however, we do not find effect on the attitudes and biases. Namely, attitudes do not seem to change as a result of exposure. We compute attitudes at the second follow-up using the same definition we use at baseline, constructing the index in the same way. Table A18 show that on average exposure did not change employers' attitudes.

Having access to the full cost of the matching program, we can compute the cost for each job created. First, while control firms hired a total of 10 refugees, treated firms hired 22 refugees. That is, our program helped firms to hire 11 more refugees. The program's overall cost, inclusive of wages of the field officers (1,929USD), transport and communication costs (877USD), wage subsidies (2,628USD) and management fees (978USD), amounted to 6,413USD.[17] Therefore, the total cost per job created was equal to 583USD and the total cost per firm participating to the experiment (182) was equal to 17USD, well in line with costs of other programs described in McKenzie (2017).

7. Policy Implications and Conclusions

How to improve the labor market integration of disadvantaged workers such as migrants and refugees is an open question with a huge policy implication. Their poor integration has long-term costs on the economies who host them. This is especially true in low-income country

[17] We exclude the costs associated to testing the skills of the refugees as well the costs of baseline surveys.

settings, where labor markets often do not function well and the national resources are already stretched.

Refugees face barriers to integration even if they possess experience and employable skills, and even if the local institutions support their rights to work. Local employers may have few incentives to hire a refugee, because they may believe that they are unskilled and the cost of testing a refugee is too high. We design and evaluate an experiment with the goal of facilitating employers' learning about workers from this disadvantaged group and helping refugees in signalling their skills to local employers.

We find that a short-term exposure is enough to stimulate the long-term (8 months) hiring among firms. This is especially true among those employers who experienced a positive match with their intern. The average effect on their willingness to hire a refugee worker on the short-term is not statistically different from zero, but firms on average do update their beliefs. The effect on the willingness to hire once more is positive among the employers who experienced a good match.

Additionally, it is worth noting that not all refugees assigned to an internship are willing to take up the offer. This is likely due to severe credit constraints and transportation costs: refugees living further away from the location of the internships are less likely to show up at the appointments.

These findings have two important policy implications. First, governments interested in investing resources to incentivize internships should take into account the constraints to access the program. For instance, refugees may need to be assisted with cash to move around the city and start their work engagements. Furthermore, both the local employers and the refugee workers may benefit from a preparatory training before engaging in the internship. This may assist them in adjusting their initial attitudes and improve the out-group contact experience.

Finally, this paper opens new questions relevant to the effect of initial attitudes on the employer-worker relationships. What is the outcome of exposure between employers and workers of any other group of workers with whom they have rarely interacted? Future research should investigate whether attitudes play a role regardless of the socio-economic status of the worker.

References

Abebe, G., S. Caria, M. Fafchamps, P. Falco, S. Franklin, and S. Quinn (2021): "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City," *The Review of Economic Studies*, 88, 1279–1310. 1

Abebe, G., S. Caria, M. Fafchamps, P. Falco, S. Franklin, S. Quinn, and F. Shilpi (2022): "Matching Frictions and Distorted Beliefs: Evidence from a Job Fair Experiment," *Working paper*. 1

Abel, M., R. Burger, and P. Piraino (2020): "The Value of Reference Letters: Experimental Evidence from South Africa," *American Economic Journal: Applied Economics*, 12, 40–71. 1

Alfonsi, L., O. Bandiera, V. Bassi, R. Burgess, I. Rasul, M. Sulaiman, and A. Vitali (2020): "Tackling Youth Unemployment: Evidence From a Labor Market Experiment in Uganda," *Econometrica*, 88, 2369–2414. 1

Allport, G. W. (1954): *The nature of prejudice*, Addison-Wesley. 1, 5.5

Arendt, J., I. Bolvig, M. Foged, L. Hasager, and G. Peri (2021): "Language Training and Refugees' Integration," *SSRN Electronic Journal*. 1

Athey, S., K. Bergstrom, V. Hadad, J. C. Jamison, B. Ozler, L. Parisotto, and J. D. Sama (2021): "Shared Decision-Making : Can Improved Counseling Increase Willingness to Pay for Modern Contraceptives?" *Working paper*. 5.3

Athey, S. and G. Imbens (2016): "Recursive partitioning for heterogeneous causal effects," *Proceedings of the National Academy of Sciences*, 113, 7353–7360. 5.3

Athey, S. and S. Wager (2019): "Estimating Treatment Effects with Causal Forests: An Application," *Observational Studies*, 5, 37–51. 1, 5.3

Bandiera, O., V. Bassi, R. Burgess, I. Rasul, M. Sulaiman, and A. Vitali (2021): "The Search for Good Jobs: Evidence from a Six-year Field Experiment in Uganda," *SSRN Electronic Journal*. 1

Barlow, F. K., S. Paolini, A. Pedersen, M. J. Hornsey, H. R. M. Radke, J. Harwood, M. Rubin, and C. G. Sibley (2012): "The Contact Caveat: Negative Contact Predicts Increased Prejudice More Than Positive Contact Predicts Reduced Prejudice," *Personality and Social Psychology Bulletin*, 38, 1629–1643. 1, 5.5

Baseler, T., T. Ginn, R. Hakiza, H. Ogude, and O. Woldemikael (2022): "Can Aid Change Attitudes Toward Refugees? Experimental Evidence from Uganda," . 1

Bassi, V., J. H. Lee, A. Peter, T. Porzio, R. Sen, and E. Tugume (2022): "Self-Employment within the Firm," . 5.4

Bassi, V. and A. Nansamba (2022): "Screening and Signalling Non-Cognitive Skills: Experimental Evidence from Uganda," *The Economic Journal*, 132, 471–511. 1

Battisti, M., Y. Giesing, and N. Laurentsyeva (2019): "Can job search assistance improve the labour market integration of refugees? Evidence from a field experiment," *Labour Economics*, 61, 101745. 1

Becker, G. M., M. H. Degroot, and J. Marschak (1964): "Measuring utility by a single-response sequential method," *Behavioral Science*, 9, 226–232. 3.1

Becker, S. O. and A. Ferrara (2019): "Consequences of forced migration: A survey of recent findings," *Labour Economics*, 59, 1–16. 1

Blair, C. W., G. Grossman, and J. M. Weinstein (2022): "Forced Displacement and Asylum Policy in the Developing World," *International Organization*, 76, 337–378. 1

Broockman, D. and J. Kalla (2016): "Durably reducing transphobia: A field experiment on door-to-door canvassing," *Science*, 352, 220–224. 1

Burchardi, K. B., J. de Quidt, S. Gulesci, B. Lerva, and S. Tripodi (2021): "Testing willingness to pay elicitation mechanisms in the field: Evidence from Uganda," *Journal of Development Economics*, 152, 102701. 3.1

Bursztyn, L., T. Chaney, T. A. Hassan, and A. Rao (2021): "The Immigrant Next Door: Long-Term Contact, Generosity, and Prejudice," *SSRN Electronic Journal*. 1

Caria, S., g. gordon, M. Kasy, S. Quinn, S. Shami, and A. Teytelboym (2020): "An Adaptive Targeted Field Experiment: Job Search Assistance for Refugees in Jordan," *SSRN Electronic Journal*. 1, 9

Carlana, M., E. La Ferrara, and P. Pinotti (2022): "Goals and Gaps: Educational Careers of Immigrant Children," *Econometrica*, 90, 1–29. 5.3

Carranza, E., R. Garlick, K. Orkin, and N. Rankin (2022): "Job Search and Hiring with Limited Information about Workseekers' Skills," *American Economic Review*, 112, 3547–3583. 1

Chassang, S., G. Padró i Miquel, and E. Snowberg (2012): "Selective Trials: A Principal-Agent Approach to Randomized Controlled Experiments," *American Economic Review*, 102, 1279–1309. 6

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018): "Double/debiased machine learning for treatment and structural parameters," *The Econometrics Journal*, 21, C1–C68. 5.3

Corno, L., E. La Ferrara, and J. Burns (2022): "Interaction, Stereotypes, and Performance: Evidence from South Africa," *American Economic Review*, 112, 3848–3875. 1

Davis, J. M. and S. B. Heller (2017): "Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs," *American Economic Review*, 107, 546–550. 1, 5.2, 5.3

——— (2020): "Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs," *The Review of Economics and Statistics*, 102, 664–677. 5.3

Dijker, A. J. M. (1987): "Emotional reactions to ethnic minorities," *European Journal of Social Psychology*, 17, 305–325. 1

Fasani, F., T. Frattini, and L. Minale (2021): "Lift the Ban? Initial Employment Restrictions and Refugee Labour Market Outcomes," *Journal of the European Economic Association*, 19, 2803–2854. 1

——— (2022): "(The Struggle for) Refugee integration into the labour market: evidence from Europe," *Journal of Economic Geography*, 22, 351–393. 1

Ginn, T., R. Resstack, H. Dempster, E. Arnold-Fernandez, S. Miller, M. Guerrero Ble, and B. Kanyamanza (2022): "2022 Global Refugee Work Rights Report," Tech. rep., Center for Global Development. 2.1

Kling, J. R., J. B. Liebman, L. F. Katz, and L. Sanbonmatsu (2004): "Moving to Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-Sufficiency and Health From a Randomized Housing Voucher Experiment," *Working paper*. 5.2

Leeb, H. and B. M. Pötscher (2005): "MODEL SELECTION AND INFERENCE: FACTS AND FICTION," *Econometric Theory*, 21. 5.5

Lepage, L.-P. (2022): "Experience-based Discrimination," *Working paper*. 1

List, J. A., A. M. Shaikh, and Y. Xu (2019): "Multiple hypothesis testing in experimental economics," *Experimental Economics*, 22, 773–793. 5.3

Lowe, M. (2021): "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration," *American Economic Review*, 111, 1807–1844. 1

Lwanga-Lunyiigo, S. (1993): "Uganda's long connection with the problem of refugees: From the Polish Refugees of World War II to the Present," . 2.1

McKenzie, D. (2017): "How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence," *The World Bank Research Observer*, 32, 127–154. 6

McKeown, S. and J. Dixon (2017): "The "contact hypothesis": Critical reflections and future directions: Critical reflections and future directions," *Social and Personality Psychology Compass*, 11, e12295. 5.5

Meleady, R. and L. Forder (2019): "When contact goes wrong: Negative intergroup contact promotes generalized outgroup avoidance," *Group Processes & Intergroup Relations*, 22, 688–707. 1

Mousa, S. (2020): "Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq," *Science*, 369, 866–870. 1

Nyqvist, M. B., A. Guariso, J. Svensson, and D. Yanagizawa-Drott (2019): "Reducing Child Mortality in the Last Mile: Experimental Evidence on Community Health Promoters in Uganda," *American Economic Journal: Applied Economics*, 11, 155–192. 5.2

Pallais, A. (2014): "Inefficient Hiring in Entry-Level Labor Markets," *American Economic Review*, 104, 3565–3599. 1

Paluck, E. L. and D. P. Green (2009): "Prejudice Reduction: What Works? A Review and Assessment of Research and Practice," *Annual Review of Psychology*, 60, 339–367. 1

Paolini, S., J. Harwood, and M. Rubin (2010): "Negative Intergroup Contact Makes Group Memberships Salient: Explaining Why Intergroup Conflict Endures," *Personality and Social Psychology Bulletin*, 36, 1723–1738. 1, 5.5

Rao, G. (2019): "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools," *American Economic Review*, 109, 774–809. 1

Scacco, A. and S. S. Warren (2018): "Can Social Contact Reduce Prejudice and Discrimination? Evidence from a Field Experiment in Nigeria," *American Political Science Review*, 112, 654–677. 1

Schuettler, K. and L. Caron (2020): *Jobs Interventions for Refugees and Internally Displaced Persons*, World Bank, Washington, DC. 1

Sladoje, M., G. Randolph, and L. Khan (2019): "Transforming Secondary Urban Areas for Job Creation: A Study of Uganda," Tech. rep., International Growth Center. 2.1

Wager, S. and S. Athey (2018): "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests," *Journal of the American Statistical Association*, 113, 1228–1242. 1, 5.3