Beyond boundaries

Doctoral dissertation

BEYOND BOUNDARIES: INTEGRATING REFUGEES AND CONSOLIDATING FARMLAND. ESSAYS IN EXPERIMENTAL AND DEVELOPMENT ECONOMICS

Mariajose Silva Vargas

2024

Cover: Kee Bar - Nakivale Refugee Settlement © Mariajose Silva Vargas
Production: ProefschriftMaken ||*www.proefschriftmaken.nl*
ISBN: 978-94-6469-733-9
UNU-MERIT, Maastricht University Dissertation Series № 302

BEYOND BOUNDARIES: INTEGRATING REFUGEES AND CONSOLIDATING FARMLAND. ESSAYS IN EXPERIMENTAL AND DEVELOPMENT ECONOMICS

Dissertation

to obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus,
Prof. dr. Pamela Habibović,
in accordance with the decision of the Board of Deans,
to be defended in public on
Wednesday 24 January 2024, at 16:00 hours

by

Mariajose Silva Vargas

Supervisor
Prof. dr. ir. Eleonora Nillesen

Co-Supervisor
Prof. dr. Jonathan de Quidt (Queen Mary University of London)

Assessment Committee
Prof. dr. Melissa Siegel (Chair)
Prof. dr. Martina Björkman Nyqvist (Stockholm School of Economics)
Prof. dr. ir. Erwin Bulte (Wageningen University)
Dr. David McKenzie (World Bank)

For Chapter 1 and Chapter 2 we gratefully acknowledge the financial support provided [in part] by the UK Foreign, Commonwealth & Development Office, awarded through Innovation for Poverty Action's Peace & Recovery Program, by PEDL, JPAL Jobs Opportunity Initiative, SurveyCTO, the Mannerfelt and the Siamon Foundations.

For Chapter 3 we gratefully acknowledge the financial support of the Centre for Market Design at the University of Melbourne, the ARC Future Fellowship Research Grant FT190100630, STICERD, and Handelsbanken's Research Foundations, grant nos B2014-0460:1, BF17-0003, and P2017-0243:1.

Para mi mamá Machi y mi hermana Clow. Che viaggio ragazze!

Contents

Abstracts

Chapter 1: Contact in the Workplace and Social Cohesion: Experimental Evidence from Uganda. Social cohesion is a driver of trust among members of the same community and consequently, it is key to local economic development. A high influx of outsiders such as refugees might disrupt this cohesion, as the arrival of foreigners may change social relations. Therefore, how to construct social cohesion in refugee-host countries is both desirable and necessary for policy. We conduct a randomized control trial with refugee job seekers and native workers in locally owned and managed firms in Uganda. We measure social cohesion through a compound measure incorporating attitudes, implicit and explicit biases, and behaviors in real and hypothetical activities. Does inter-group contact in the workplace promote social cohesion between people from two different communities? Our sets of findings are two. First, while implicit bias increases, explicit bias decreases for both groups. Second, both groups of workers improve their behaviors towards the opposite group, but in a slightly different way: while local workers want to have more refugee business partners, refugee workers want to be more employed by Ugandan firms. These findings underscore the role of workplace-based contact in developing social cohesion among people from different communities.

Chapter 2: Matching with the Right Attitude: the Effect of Matching Firms with Refugee Workers. 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.

Chapter 3: Market Design for Land Trade: Evidence from Uganda and Kenya. Agriculture in low-income countries is characterized by misallocation of land across farmers, and fragmentation – the separation of farms into smaller plots – which increases costs and limits the use of increasing returns technologies. We argue that a carefully-designed set of trading rules can improve outcomes by addressing these problems, and test this using two lab-in-the-field experiments with smallholder farmers in Uganda and Kenya. First, with survey data, we document that agricultural land markets are thin, prone to exposure risk, and suffer from coordination frictions. These characteristics typically hamper decentralized trade. Market design may improve outcomes by thickening markets, finding chains, and enforcing conditional contracts, but right-in-theory designs may be unfamiliar and hard for farmers to understand. Our first experiment, conducted in Uganda, simulates status quo land markets and confirms severe inefficiency. In a second phase of the experiment we find large efficiency gains from a simple market design improvement, a centralization intervention that brings farmers together to trade at a set time. We then test whether designs that are more tailored to the land trade problem, but potentially harder to understand, can further improve outcomes. Our second experiment, in Kenya, finds that a computerized package exchange, which allows traders to specify a sequence of conditional trades as a single transaction, performs particularly well. Our results suggest that improved market design can reduce market frictions and lead to important productivity gains among smallholder farmers.

Synopsis

This dissertation comprises three self-contained essays. The first two chapters explore the role of the private sector in fostering social cohesion and labor market integration between native firms, workers, and skilled refugee workers. Chapter 1 asks if work contact between native workers and refugee workers can lead to stronger social cohesion between groups. Chapter 2 studies the effect of contact in the workplace with a refugee worker on native employers' willingness to hire refugees. Instead, Chapter 3 focuses on land markets: it examines if well-designed land trading rules can alleviate challenges in low-income countries' agriculture by addressing inefficiencies and fragmentation.

Despite the diversity across the chapters, a common theme prevails as they collectively investigate strategies to transcend social and physical boundaries in low-income countries. Furthermore, all the chapters employ experimental methodologies – including randomized controlled trials and lab-in-the-field experiments – along with primary data to address important research questions and fill research gaps in the development economics field. These studies were conducted within the contexts of Uganda and Kenya, both of which are burgeoning economies in East Africa.

Chapter 1: Contact in the Workplace and Social Cohesion: Experimental Evidence from Uganda.

with Francesco Loiacono

Social cohesion is a key factor for growth and development, especially in countries with high levels of diversity. However, forced displacement can threaten this cohesion by disrupting and changing social relations in host countries. According to the UNHCR, there are currently around 110 million forcibly displaced people, with 36.4 million being refugees. Researchers, governments, and international organizations in refugees' host countries are therefore interested in understanding what policies or programs can enhance social cohesion with conflict-affected populations. This question is of particular importance for low- and middle-income countries, that host three-quarters of the world's refugees.

This chapter delves into the impact of workplace contact on enhancing social cohesion between native and refugee workers in Uganda, the largest refugee-host country in Africa. We run a randomized controlled trial where 377 refugees and 273 local workers were matched with each other and randomly placed into a control group and three types of work contact treatments: (i) a direct contact treatment where refugees completed a 1-week internship at a local firm; (ii) an indirect contact treatment where participants watch a video documentary showing the daily interactions of a refugee and a Ugandan working together in a firm in Kampala; (iii) both.

To assess the impact of work contact on social cohesion between refugees and local workers, we define a new compound measure of social cohesion that comprises several dimensions: implicit bias against the out-group is the first outcome variable, gauged using two implicit association tests (IATs) targeting distinct dimensions of bias: general bias and work-related bias. Explicit bias is the second outcome, determined by amalgamating explicit stereotypes and negative attitudes into an index. Behavioral outcomes, both real-world and hypothetical, constitute the final set of measures.

Our sets of results are two. First, we find an overall positive impact of work contact on social cohesion. Work contact decreases explicit bias both among local and refugee workers. At the same time, implicit bias increases among both groups of workers, and the increase is significant for local workers. Second, actual behaviors move in the same direction as explicit bias, however: treated local workers are more willing to have a refugee business partner in a hypothetical scenario, while more refugees are willing to work in a similar internship program in the future, especially with Ugandan firms. This effect is large as it is equivalent to a 90% increase over the mean. We also find that treated refugees are less willing to have any partner in a hypothetical business scenario. Together, we interpret these results as evidence that through work contact, refugee workers learn that they can look for salaried jobs in established firms instead of becoming self-employed.

The fact that implicit bias increases while explicit bias and behaviors improve is intriguing. We provide suggestive evidence regarding local workers' increase in implicit bias as not being driven by negative work contact but rather by the fear of increased job competition: through work contact with a refugee, local workers learn that refugee workers are more skilled than they initially believed. Second, an increase in implicit bias does not translate into discriminatory behavior, as the effect on the behavioral outcomes is positive.

Our study also makes several methodological contributions. First, we measure biases and behavioral change with contact both for the majority group (i.e. the local workers) and the minority group (i.e. the refugees). Second, by measuring both implicit bias and explicit bias and reported behaviors, we can use the latter to interpret the former and thus contribute to the discussion on how to measure and interpret implicit bias through implicit association tests (IATs). Third, we take seriously the possibility of experimenter demand effects and design a number of safeguards to protect against them, such as collecting behavioral measures, matching enumerators and respondents by nationality and eliciting respondents' beliefs about the study's purpose.

Chapter 2: Matching with the Right Attitude: the Effect of Matching Firms with Refugee Workers.

with Francesco Loiacono

Immigrants, notably refugees, represent a highly vulnerable global population, often grappling with unemployment, leading to untapped potential and societal costs. Integrating refugees into the labor market faces challenges due to factors like inadequate human capital, entry barriers, and cultural differences affecting employer perceptions. The high offer of native labor supply discourages firms from gathering the necessary information to counter their biases, spurring labor market policies like internships and hiring subsidies to enhance refugee employment prospects and overall labor market efficiency by hiring skilled workers.

In this chapter, we study the effect of reducing demand-side frictions to hire a refugee worker by running a randomized control trial in Uganda. The country is the ideal setting to explore refugee labor market integration. As Africa's primary refugee-hosting nation, Uganda upholds an open policy that grants refugees unrestricted movement within its borders, facilitating employment opportunities. The experiment focuses on assessing both short- and long-term effects on native firms' willingness to hire refugees, following a one-week internship by skilled refugees. The firms are the same as in Chapter 1, but in this chapter, we focus on a different sample: firm owners and managers.

Our findings reveal that firms exhibiting positive attitudes towards refugees, when (randomly) paired with refugees holding positive attitudes towards locals, significantly increase their willingness to hire a (generic) refugee worker within a week after the experiment's conclusion. Conversely, firms with negative attitudes towards refugees, matched with refugees sharing similar negative attitudes towards locals, experience a decrease in willingness to hire.

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 in 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.

Chapter 3: Market Design for Land Trade: Evidence from Uganda and Kenya.

with Gharad Bryan, Jonathan de Quidt, Tom Wilkening and Nitin Yadav

Inefficient land allocation reduces productivity in low-income countries' agriculture. Farms are small and fragmented, despite the fact that labor and total factor productivity increase with farm size. Land is also misallocated – there is substantial heterogeneity in farmer productivity but almost no correlation between farmer productivity and land holding. These inefficiencies suggest large unrealized gains from trade, a claim borne out by quantitative and experimental analyses. We argue that improved market design can help unlock these gains by creating trading rules tailored to address key frictions in the land market, and note that it is likely complementary to other institutions, such as property rights, that are more often emphasized.

In the tradition of the market design literature, our exploration takes place within the confines of a simplified, lab-in-the-field, environment. We run the different experiments in Uganda and Kenya. We believe our designs capture the key constraints that market design can address, and abstract from problems that are best addressed elsewhere. For example, we do not allow for risk of fraud. While fraud may be an important part of what constrains land trade, we think that this is best addressed through complementary policies, rather than directly in the market design. The upshot of this is that our estimates of the impact of market design should be seen as conditional on getting other institutions right. Whether real-world gains would be larger or smaller depends on the extent of complementarity between market design and other programs, and it may well be that market design is a strong complement to other interventions.

We build our argument in three steps. First, we document, with a survey of smallholder farmers in Uganda, that decentralized land trade—based on buyers and sellers bargaining over individual plots—is likely to be inefficient. Farmers believe their trading environment has key characteristics that the market design literature predicts will inhibit trade. Second, we provide lab-in-the-field evidence that decentralized trade is indeed inefficient. We build a stylized representation of the environment, consistent with our survey evidence, in which real farmers trade fictitious land titles with strong financial incentives. We let them trade for a week without any formal trading rules, and show that the final allocation is far from efficient. Third, we show that specifying rules targeting the market frictions we highlight improves outcomes in the same land trade game. We tested interventions that range from simple, easy-to-understand rules that would be expected to facilitate trade in a wide range of problems, to rules that are highly tailored to the land trade problem but potentially difficult for our target audience to understand. Overall, we find that increasingly tailored rules, despite their increasing complexity, consistently improve efficiency without increasing inequality.

Samenvatting

Deze proefschrift bestaat uit drie op zichzelf staande essays. De eerste twee hoofdstukken onderzoeken de rol van de private sector bij het bevorderen van sociale cohesie en arbeidsmarktintegratie tussen lokale bedrijven, werknemers en gekwalificeerde vluchtelingen. Hoofdstuk 1 onderzoekt of werkcontact tussen lokale werknemers en gekwalificeerde vluchtelingen kan leiden tot sterkere sociale cohesie tussen groepen. Hoofdstuk 2 onderzoekt het effect van contact op de werkvloer met een vluchteling op de bereidheid van lokale werkgevers om vluchtelingen aan te nemen. Daarentegen richt Hoofdstuk 3 zich op landmarkten: het onderzoekt of goed ontworpen regels voor de handel in land de uitdagingen in de landbouw van landen met een laag inkomen kunnen verlichten door inefficiënties en fragmentatie aan te pakken.

Ondanks de diversiteit tussen de hoofdstukken, is er een gemeenschappelijk thema: gezamenlijk onderzoeken ze strategieën om sociale en fysieke grenzen in lage-inkomenslanden te slechten. Bovendien maken alle hoofdstukken gebruik van experimentele methodes, waaronder gerandomiseerde gecontroleerde experimenten, samen met primaire gegevens om belangrijke onderzoeksvragen aan te pakken en hiaten in het onderzoeksveld van de ontwikkelingseconomie op te vullen. Deze studies zijn uitgevoerd binnen de context van Oeganda en Kenia, beide opkomende economieën in Oost-Afrika.

Hoofdstuk 1: Contact in de Werkplek en Sociale Cohesie: Experimenteel Bewijs uit Oeganda.

met Francesco Loiacono

Sociale cohesie is een sleutelfactor voor groei en ontwikkeling, vooral in landen met een grote verscheidenheid aan nationaliteiten. Echter, gedwongen ontheemding kan deze cohesie bedreigen door verstoring en verandering van sociale relaties in gastlanden. Volgens de UNHCR zijn er momenteel ongeveer 110 miljoen gedwongen ontheemde mensen, waarvan 36.4 miljoen vluchtelingen zijn. Onderzoekers, regeringen en internationale organisaties in gastlanden zijn daarom geïnteresseerd in het begrijpen van welke beleidsmaatregelen of programma's de sociale cohesie met door conflicten getroffen bevolking kunnen versterken. Deze vraag is met name belangrijk voor lage- en middeninkomenslanden, die driekwart van 's werelds vluchtelingen herbergen.

Dit hoofdstuk analyseert de impact van contact op de werkvloer op het verbeteren van de sociale cohesie tussen inheemse en vluchtelingenwerkers in Uganda, het grootste gastland voor vluchtelingen in Afrika. We voeren een gerandomiseerde gecontroleerde proef uit waarbij 377 vluchtelingen en 273 lokale werkers aan elkaar werden gekoppeld en willekeurig werden ingedeeld in een controlegroep en drie soorten werkplekcontactinterventies: (i) een "direct contact" interventie waarbij vluchtelingen een stage van 1 week volgden bij een lokaal bedrijf; (ii) een "indirect contact" interventie waarbij deelnemers een video documentaire bekijken die de dagelijkse interacties toont van een vluchteling en een Oegandees die samenwerken in een bedrijf in Kampala; (iii) beide.

Om de impact van werkplekcontact op sociale cohesie tussen vluchtelingen en lokale werkers te beoordelen, definiëren we een nieuwe samengestelde maatstaf voor sociale cohesie die verschillende dimensies omvat: impliciete vooringenomenheid tegen de buitenstaanders is de eerste uitkomstvariabele, gemeten met behulp van twee IAT's die zich richten op onderscheidende dimensies van vooringenomenheid: algemene vooringenomenheid en werkgerelateerde vooringenomenheid. Expliciete vooringenomenheid is de tweede uitkomst, bepaald door expliciete stereotypen en negatieve attitudes samen te voegen tot een index. Gedragsuitkomsten, zowel in de echte wereld als hypothetisch, vormen de laatste reeks maatregelen.

Onze resultaatsverzameling bestaat uit twee delen. Allereerst constateren we over het algemeen een positieve invloed van werkcontact op sociale samenhang. Werkcontact vermindert zowel de expliciete vooroordelen bij lokale werknemers als bij vluchtelingen. Tegelijkertijd neemt de impliciete vooringenomenheid toe bij beide groepen werknemers, waarbij de toename significant is bij lokale werknemers. Ten tweede volgen daadwerkelijke gedragingen dezelfde richting als expliciete vooroordelen: behandelde lokale werknemers zijn meer bereid om een vluchteling als zakelijke partner te hebben in een hypothetisch scenario, terwijl meer vluchtelingen bereid zijn om in de toekomst deel te nemen aan een soortgelijk stageprogramma, vooral met Oegandese bedrijven. Dit effect is aanzienlijk, aangezien het equivalent is aan een toename van 90% boven het gemiddelde. We vinden ook dat behandelde vluchtelingen minder bereid zijn om in een hypothetisch zakelijk scenario met een willekeurige partner samen te werken. Samengevoegd interpreteren we deze resultaten als bewijs dat vluchtelingen door werkcontact leren dat ze naar salarisbanen kunnen zoeken bij gevestigde bedrijven in plaats van zelfstandig ondernemer te worden.

Het feit dat de impliciete vooringenomenheid toeneemt terwijl de expliciete vooringenomenheid en gedragingen verbeteren, is intrigerend. We leveren suggestief bewijs over de toename van de impliciete vooringenomenheid bij lokale werknemers, die niet wordt aangedreven door negatief werkcontact, maar eerder door de angst voor verhoogde concurrentie op de arbeidsmarkt: via werkcontact met een vluchteling leren lokale werknemers dat vluchtelingen vaardiger zijn dan ze aanvankelijk dachten. Ten tweede vertaalt een toename van impliciete vooringenomenheid zich niet in discriminerend gedrag, aangezien het effect op de gedragsresultaten positief is.

Onze studie levert ook verschillende methodologische bijdragen. Ten eerste meten we veranderingen in vooringenomenheden en gedragingen bij contact zowel voor de meerderheidsgroep (d.w.z. de lokale werknemers) als voor de minderheidsgroep (d.w.z. de vluchtelingen). Ten tweede, door zowel impliciete vooringenomenheid als expliciete vooringenomenheid en gerapporteerde gedragingen te meten, kunnen we de laatste gebruiken om de eerste te valideren en zo bijdragen aan de discussie over hoe impliciete vooringenomenheid moet worden gemeten en geïnterpreteerd via impliciete associatietests (IAT's). Ten derde nemen we de mogelijkheid van experimentatorvraageffecten serieus en ontwerpen we verschillende waarborgen om hiertegen te beschermen, zoals het verzamelen van gedragsmaatregelen, het matchen van enquêteurs en respondenten op nationaliteit en het bevragen van respondenten over hun overtuigingen over het doel van de studie.

Hoofdstuk 2: De Juiste Houding bij het Koppelen: Het Effect van het Koppelen van Bedrijven met Vluchtelingenwerkers.

met Francesco Loiacono

Immigranten, met name vluchtelingen, vertegenwoordigen een zeer kwetsbare mondiale bevolking die vaak te maken heeft met werkloosheid, wat leidt tot onbenut potentieel en maatschappelijke kosten. Het integreren van vluchtelingen in de arbeidsmarkt kent uitdagingen vanwege factoren zoals ontoereikend menselijk kapitaal, toegangsdrempels en culturele verschillen die van invloed zijn op de percepties van werkgevers. Het overvloedige aanbod van binnenlandse arbeid weerhoudt bedrijven ervan om de nodige informatie te verzamelen om hun vooroordelen tegen te gaan, wat arbeidsmarktbeleid stimuleert zoals stages en loonkostensubsidies om de arbeidskansen van vluchtelingen te verbeteren en de algehele efficiëntie van de arbeidsmarkt te vergroten door geschoolde werknemers aan te nemen.

In dit hoofdstuk bestuderen we het effect van het verminderen van een (te) lage arbeidsvraag om een vluchtelingenwerker aan te nemen door middel van een gerandomiseerde controleproef in Uganda. Het land is de ideale omgeving om de integratie van vluchtelingen op de arbeidsmarkt te verkennen. Als primair gastland voor vluchtelingen in Afrika handhaaft Uganda een open beleid dat vluchtelingen onbeperkte bewegingsvrijheid binnen zijn grenzen biedt, wat werkgelegenheidskansen vergemakkelijkt. Het experiment richt zich op het beoordelen van zowel korte- als langetermijneffecten op de bereidheid van lokale bedrijven om vluchtelingen aan te nemen, na een stage van één week door vaardige vluchtelingen. Onze steekproef van lokale bedrijven is hetzelfde als in Hoofdstuk 1, maar in dit hoofdstuk richten we ons op een andere steekproef: eigenaren en managers van bedrijven.

Onze bevindingen tonen aan dat bedrijven die positieve houdingen hebben ten opzichte van vluchtelingen, wanneer ze (willekeurig) worden gekoppeld aan vluchtelingen met positieve houdingen ten opzichte van lokale bevolking, hun bereidheid om binnen een week na afloop van het experiment een (algemene) vluchtelingenwerker aan te nemen significant vergroten. Omgekeerd ervaren bedrijven met negatieve houdingen ten opzichte van vluchtelingen, gekoppeld aan vluchtelingen die soortgelijke negatieve houdingen ten opzichte van de lokale bevolking delen, een afname in de bereidheid om aan te nemen.

Tot slot, en cruciaal, vinden we dat de interventie die slechts één week duurde een aanzienlijke impact had op daadwerkelijk aannemen van vluchtelingen, met een groter effect in de subgroep van bedrijven die aanvankelijk een positieve houding hadden ten opzichte van vluchtelingen en (willekeurig) werden gekoppeld aan een vluchteling met positieve houdingen ten opzichte van de inheemse bevolking. Het effect dat we schatten, kan worden geïnterpreteerd als een externe factor: een match met een vluchteling met een positieve houding ten opzichte van lokale mensen vergroot de bereidheid van het bedrijf om in het algemeen vluchtelingen aan te nemen, vooral wanneer de oorspronkelijke houding van de bedrijfsmanager ten opzichte van vluchtelingen ook positief is. Houdingen zijn complementair en versterken het effect van contact op de werkplek.

Samengevat hebben onze bevindingen belangrijke beleidsimplicaties. We laten zien dat een korte interventie van slechts één week kan leiden tot toename van werkgelegenheid voor een bijzonder kwetsbare groep zoals vluchtelingen op de lange termijn, maar dat de omvang van het effect afhangt van de aanvankelijke kwaliteit van de “match” tussen bedrif en vluchteling.

Hoofdstuk 3: Marktdesign voor Grondhandel: Bewijs uit Uganda en Kenia.

met Gharad Bryan, Jonathan de Quidt, Tom Wilkening en Nitin Yadav

Efficiënte toewijzing van land verhoogt de productiviteit in de landbouw van landen met een laag inkomen. Boerderijen zijn klein en versnipperd, ondanks het feit dat arbeid en totale factorproductiviteit toenemen met de grootte van de boerderij. Land wordt ook verkeerd toegewezen: er is aanzienlijke heterogeniteit in de productiviteit van boeren, maar bijna geen correlatie tussen de productiviteit van boeren en het grondbezit. Deze inefficiënties duiden op grote ongerealiseerde winsten uit handel, een claim die wordt ondersteund door kwantitatieve en experimentele analyses. Wij stellen dat een verbetering van het marktontwerp deze winsten kan ontsluiten door handelsregels te creëren die zijn afgestemd op belangrijke wrijvingen op de grondmarkt, en wij merken op dat dit waarschijnlijk complementair is aan andere instellingen, zoals eigendomsrechten, die vaker worden benadrukt.

In de traditie van de literatuur over marktontwerp vindt ons onderzoek plaats binnen de kaders van een vereenvoudigde, laboratoriumomgeving in het veld. We voeren de verschillende experimenten uit in Uganda en Kenia. We geloven dat onze ontwerpen de belangrijkste beperkingen vastleggen die marktontwerp kan aanpakken, en abstraheren van problemen die elders beter kunnen worden aangepakt. Zo staan we bijvoorbeeld geen risico op fraude toe. Hoewel fraude mogelijk een belangrijk onderdeel is van wat landhandel beperkt, denken we dat dit beter kan worden aangepakt via aanvullend beleid, in plaats van direct in het marktontwerp. Het resultaat hiervan is dat onze schattingen van de impact van marktontwerp gezien moeten worden als afhankelijk van het goed krijgen van andere instellingen. Of de winsten in de echte wereld groter of kleiner zouden zijn, hangt af van de mate van complementariteit tussen marktontwerp en andere programma's, en het kan heel goed zijn dat marktontwerp een sterke aanvulling is op andere interventies.

We bouwen ons betoog op in drie stappen. Eerst documenteren we, met een enquête onder kleine boeren in Uganda, dat gedecentraliseerde landhandel - gebaseerd op onderhandelingen tussen kopers en verkopers over individuele percelen - waarschijnlijk inefficiënt is. Boeren geloven dat hun handelsomgeving belangrijke kenmerken heeft die in de literatuur over marktontwerp worden gezien als belemmering voor de handel. Ten tweede leveren we bewijs dat gedecentraliseerde handel inderdaad inefficiënt is. We bouwen een gestileerde voorstelling van de omgeving, in overeenstemming met onze bevindingen van de enquête, waarin echte boeren fictieve kavels verhandelen met sterke financiële prikkels. We laten hen een week lang handelen zonder formele handelsregels, en tonen aan dat de uiteindelijke toewijzing verre van efficiënt is. Ten derde laten we zien dat het specificeren van regels die gericht zijn op fricties in de markt die we benadrukken, de resultaten in hetzelfde landhandelspel verbetert. We hebben interventies getest die variëren van eenvoudige, gemakkelijk te begrijpen regels die naar verwachting de handel in een breed scala van problemen zouden vergemakkelijken, tot regels die sterk zijn afgestemd op het probleem van de landhandel, maar mogelijk moeilijk te begrijpen zijn voor onze doelgroep. Over het algemeen vinden we dat steeds meer op maat gemaakte regels, ondanks hun toenemende complexiteit, de efficiëntie consistent verbeteren zonder de ongelijkheid te vergroten.

1

Chapter 1: Contact in the Workplace and Social Cohesion: Experimental Evidence from Uganda

This paper has been jointly written with Francesco Loiacono

Chapter 1: Table of Contents

1 Introduction

Social cohesion is a key factor for growth and development, especially in countries with high levels of diversity (Easterly et al. 2006; Munshi 2011). However, forced displacement can threaten this cohesion by disrupting and changing social relations in host countries (De Berry and Roberts 2018). According to the UNHCR, there are currently around 110 million forcibly displaced people, with 36.4 million being refugees. Researchers, governments and international organizations in refugees' host countries are therefore interested in understanding what policies or programs can enhance social cohesion with conflict-affected populations. This question is of particular importance for low- and middle-income countries, host to three-quarters of the world's refugees.

This paper uses an experiment in Uganda, the largest refugee-hosting country in Africa and the fifth worldwide, to study if work contact between refugees and local workers can increase social cohesion. We randomly match 377 refugees and 273 local workers to work for the same company. We then randomly group these couples into a control arm and three types of 'work contact' treatments: (i) a *"direct"* contact treatment where refugees complete a 1-week internship at a local firm; (ii) an *"indirect"* contact treatment where participants watch a video documentary showing the daily interactions of a refugee and a Ugandan working together at a firm in the capital city Kampala; (iii) a combination of both treatments. To assess the impact of work contact (both direct and indirect) on social cohesion between refugees and local workers, we define a new compound measure of social cohesion that comprises several dimensions: implicit bias, explicit stereotypes, attitudes, and behaviors both in hypothetical and real-world scenarios.

Our sets of results are two. First, we find an overall positive impact of work contact on social cohesion. Work contact decreases explicit bias both among local and refugee workers. At the same time, implicit bias increases among both groups of workers, and the increase is significant for local workers. Second, actual behaviors move in the same direction as explicit bias, however: treated local workers are more willing to have a refugee business partner in a hypothetical scenario, while more refugees are willing to work in a similar internship program in the future, especially with Ugandan firms. This effect is large as it is equivalent to 90% increase over the mean. We also find that treated refugees are less willing to have any partner in a hypothetical business scenario. Together, we interpret these results as evidence that through work contact, refugee workers learn that they can look for salaried jobs in established firms instead of becoming self-employed.

The fact that implicit bias increases while explicit bias and behaviors improve is intriguing. We provide suggestive evidence regarding local workers' increase in implicit bias as not being driven by negative work contact but rather by the fear of increased job competition: through work contact with a refugee, local workers learn that refugee workers are more skilled than they initially believed. Second, an increase in implicit bias does not translate into discriminatory behavior, as the effect on the behavioral outcomes is positive.

Our study also makes several methodological contributions. First – and in contrast to much of the existing literature – we measure biases and behaviors change with contact both for the majority group (i.e. the local workers) and the minority group (i.e. the refugees). Second, by measuring both implicit and explicit bias and actual behaviors, we can use the latter to interpret the former and thus contribute to the discussion on how to measure and interpret implicit bias through implicit association tests (IATs). Third, we take seriously the possibility of experimenter demand effects and design a number of safeguards to protect against them: (i) we collect behavioral measures, which we expect to be less subject to demand effects; (ii) we match enumerators and respondents by nationality as we expect participants to be more willing to admit biases when paired with an enumerator of the same nationality; (iii) we elicit respondents' beliefs about the study's purpose at the end of the program, to assess whether the purpose of this study was obfuscated to the participants.

This paper relates to four bodies of work in different fields. First, to the vast literature on the reduction of prejudice using contact theory (Bursztyn et al. 2023; Corno et al. 2022; Lowe 2021; Mousa 2020; Okunogbe 2023; Rao 2019; Scacco and Warren 2018). These papers study activities that promote direct contact between different groups. Our contribution is to promote contact using a different activity, namely work, which is arguably the most important in the daily lives of adults. Second, to the recent literature of the integration of refugees in low- and middle-income countries (Bahar et al. 2021; Caria et al. (2023)). These papers focus on the impact of labor or governmental programs on labor market outcomes. Our contribution is to explore if a labor market program can also promote social cohesion between refugees and locals. Third, to the literature on post-conflict reconstruction using employment programs (Blattman and Annan 2016). Many governments and donors use job programs to promote peace between different groups, but there is little empirical evidence that they actually do (Verwimp et al. 2019). Thus, we will provide empirical evidence on the role of an employment program in building social cohesion in a conflict-affected community. Finally, to the literature on implicit bias measurement and interpretation (Cunningham and De Quidt 2023). Implicit bias has been mostly measured using implicit association tests, but in the psychology field there is an open debate about the validity of the tool (Singal 2017): some studies have found that it measures empathy or exposure to stereotypes (Andreychik and Gill 2012; Uhlmann et al. 2006) and there is no evidence to support the claim that the IAT score is related to discriminatory behavior (Paluck et al. 2020). In economics, IATs are increasingly used, but there is some evidence that implicit bias does not always translate into prejudiced behavior (Alesina et al. 2018). We contribute to this literature by collecting several outcomes together with IATs, so as to understand their relation to behavior and what it might be capturing.

The remainder of the paper is organized as follows. Section 2 summarizes the literature on contact theory and on the measurement and use of implicit attitudes. Section 3 describes the context and provides some stylized facts. Section 4 details the conceptual framework, experimental design and sampling. Section 5 describes the main outcomes of the paper and provides some descriptive statistics. Section 6 outlines the specification used in the analysis, reports and discusses the results of the experiment. Section 7 concludes the paper.

2 Literature Review

In this section we review the current literature on the contact hypothesis, the interventions aimed at reducing prejudice, and the measurement of implicit bias.

The contact hypothesis, proposed by Allport in 1954, is a theory that explains how contact between different groups can reduce prejudice and discrimination. The hypothesis states that under certain conditions, direct contact between individuals from different groups can increase mutual understanding and reduce prejudice (Allport 1954).

According to Allport, there are four essential conditions for contact to be effective in reducing prejudice: (1) there is equal status between majority and minority groups; (2) contact is endorsed by institutional support, laws or custom; (3) groups work for a common goal; and (4) there is intergroup cooperation. Allport believed that these conditions would allow people from different groups to see each other as individuals rather than as members of a particular group. This would help to reduce stereotypes and prejudice by breaking down the social barriers that exist between different groups.

The contact hypothesis has been widely used and studied to promote social cohesion or to reduce prejudice in a variety of fields (Bertrand and Duflo 2017). In experimental economics and political science, direct contact has been studied using natural experiments or randomized controlled trials in specific settings: in sports (Mousa 2020; Lowe 2021) in education (Rao 2019; Scacco and Warren 2018) in locations such as university rooms (Carrell et al. 2015; Corno et al. 2022) or neighbourhoods (Bursztyn et al. 2023; Okunogbe 2023).

Most of the above-mentioned papers collected explicit attitudes and behavioral outcomes, finding similar patterns in the results. Mousa (2020) found that Christians assigned to play on a soccer team with Muslim teammates were more likely to engage in tolerant behaviors toward Muslim teammates up to 6 months after the intervention ended, yet the tolerant behavior did not generalize to other Muslims or to attitudes. Lowe (2021) instead found that adversarial contact in cricket teams in India had a mixed effect and, in some cases, negative, supporting Allport's rule that members from different groups need to have contact while working towards common goals, in order to achieve positive outcomes. In education, Scacco and Warren (2018) found that randomly assigning Muslims and Christians to computer classes reduced the tendency to discriminate in behavioral games, but not in prejudiced self-reported attitudes.

Corno et al. (2022) is one of the few studies that collected three types of outcomes: explicit attitudes, implicit bias, and behaviors. They studied the random allocation of white and black students in rooms at a South African university. They find that exposure to a roommate from a different race reduces the implicit bias of white students, improved the academic ability of black students, and improved explicit attitudes and friendship patterns for white students. They conclude that their results are encouraging because South Africa has a deep history of prejudice and conflict between groups, and it should be more difficult to reduce prejudice in such a context.

In the psychology field, Paluck et al. (2020) run a meta-analysis of interventions to reduce prejudice. They find that there is enthusiasm about implicit bias and its reduction, but there is no clear evidence that implicit bias reduction is correlated or leads to a reduction in prejudiced behavior. In their analysis, only two experiments measured implicit and behavioral change. Also, they conclude that attitudes and behaviors, although correlated, diverge. Interventions seem to be more effective at changing behavior than attitudes.

Regarding the measurement of implicit attitudes, implicit bias has been mostly measured using implicit association tests (IATs). The IATs are psychological tools that capture biases using *categorization tasks* (Greenwald and Banaji 1995). In the socio-psychological literature, there is a wide discussion regarding the IAT validity and interpretation (Singal 2017). Mainly, the discussion deals with two points: if the IAT actually measures prejudice and if the IAT's score is a predictor of discriminatory behavior.

The co-creators of the IAT showed in a meta-analysis that the IAT correlated with discriminatory behavior (Greenwald et al. 2009). Nevertheless, the meta-analysis has been criticized for including studies and outcomes that do not actually measure discriminatory behavior (Singal 2017). Another meta-analysis has shown the contrary, that the IAT score does not translate into prejudiced behavior (Oswald et al. 2013). However, this analysis has also been criticized for including studies with small sample sizes (Corno et al. 2022).

Both opponents and proponents agree that the evidence is very thin, especially experimental evidence. A recent meta-analysis by Paluck et al. (2020) concluded that only 2 experimental studies included both implicit and behavioral outcomes. Also, they concluded that there appears to be no correlation between implicit and explicit stereotypes, confirming "the notion that the two measures gauge distinct (and largely unrelated) response tendencies".

Regarding the discussion about what the IAT is actually measuring, there have been studies that show alternative explanations. For instance, Uhlmann et al. (2006) have shown in an experimental study that the IAT is measuring familiarity with negative stereotypes regarding a specific group. Others have shown that the IAT might be measuring empathy towards a group (Andreychik and Gill 2012).

In economics, IATs are increasingly used as a proxy for prejudice, but very few papers use it as an outcome. Beaman et al. (2009) find that a quota to reserve political seats for women in the local government in India does not improve the implicit or explicit distaste for female leaders, and it actually improves the relative explicit preference for male leaders. Yet, it improves some behavioral outcomes in hypothetical scenarios, such as female leader effectiveness. Alesina et al. (2018), find that while math teachers with stronger implicit bias grade immigrant students lower grades than local students, literature teachers do not act upon their implicit bias.

3 Context

Due to conflicts, economic, political, and climate instability, the number of displaced people has increased in many regions around the world. According to the UNHCR, there are currently around 110 million forcibly displaced people. Of those, 36.4 million are refugees.[1] Three-quarters of the world's refugees are hosted by low- and middle-income countries. Uganda is the fifth-largest refugee host country in the world and the first in Africa, currently hosting around 1.5 million refugees. The country has been praised worldwide for its progressive refugee policy: refugees have freedom of movement and have the right to live and work outside the settlements. Refugees can choose where to register. They can choose between settlements situated in rural areas or to become urban dwellers, by going to live in major cities such as the capital Kampala. Conditional on choosing to live in a settlement, refugees can also receive a plot of land to cultivate and receive aid. However, some frictions to their integration remain. If refugees decide to leave the settlements, they do not receive aid; when needed, getting an official work

[1] https://www.unhcr.org/refugee-statistics/, accessed November 2023.

permit is difficult; some employers hesitate to hire refugees because they are unsure about the laws and policies; and refugees are less likely to be employed than Ugandans and often accept jobs below their skills and education level (Loiacono and Silva-Vargas 2019; Loiacono and Silva-Vargas 2023). Additionally, a continuous influx of refugees is posing new dilemmas and open questions, particularly, what the best policy to promote social cohesion would be to avoid conflict and stimulate economic growth in the host country.

3.1 Study pilot and stylized facts

Between 2019 and 2020, we collected two rounds of pilot data with 421 urban refugees and 401 local firm owners in two large cities in Uganda. Our pilot data show some interesting insights: 83% of refugees in Kampala believe that Ugandans are not trustworthy, and 70% have low levels of generalized trust. Yet, only 42% report low levels of trust towards refugees of their same nationality. Around half of the sample believes that Ugandans are prejudiced towards refugees and rate their interactions with Ugandans negatively. Very few refugees have established work contact with Ugandans, suggesting that refugees are segregated with respect to the local communities. Only 20% were paid employees at the time of the interview, and out of this, 40% had a Ugandan employer. Finally, even if living in such urban settings and not in isolated settlements, only 16% report that they have zero weekly economic interactions with locals. We believe that these findings provide suggestive evidence that meaningful work contact between refugees and local peers is limited.

We observe similar trends among Ugandan respondents. Out of 401 firm owners, 46% have economically interacted with refugees; 86% report that they do not trust refugees while only 15% say the same for Ugandans of their own ethnicity; 69% believe that hosting refugees does not help the country economically and socially and 68% believe that hosting refugees creates more competition for opportunities in the country. Surprisingly, although Uganda is the fifth host country in terms of the number of refugees, only 7% of the firms reported ever having hired a refugee, and a substantial number of firms do not know the refugee policy: 61% do not know that refugees can live outside the settlements, and 59% do not know that refugees can work anywhere in Uganda.

We find interesting correlations. Refugees that have more interactions per week with locals have more trust towards Ugandans, report that Ugandans are less prejudiced towards refugees, and rate their interaction with locals more positively (Figure 1). For Ugandans, if firm owners ever interacted economically with refugees, they trust refugees more, they are less likely to agree that refugees create competition with Ugandan workers, and they agree more that refugees help the country economically and socially.

These correlations are consistent with the contact hypothesis and are suggestive evidence that economic interactions, more positive attitudes and beliefs towards the out-group are correlated. Therefore, working together not only could yield economic returns but also could improve social outcomes.

Finally, at the end of 2020, we run qualitative discussions with both groups (61 refugees and 120 locals) to understand what type of actions or activities show integration between both groups in urban areas. Ugandans agreed that the main activities to promote cohesion are to support refugee businesses and to work together, while for refugees, important activities are to attend social and religious gatherings and work together.

4 Research Design

In this section we describe the experimental design, introducing a conceptual framework that motivates our intervention and specifying the hypothesis we want to test.

4.1 Conceptual framework and hypotheses

The contact hypothesis is a widely proposed theory for reducing prejudice. Developed by the sociologist Gordon Allport in 1954, the original theory states that contact between different groups may reduce prejudice if four conditions are met: (1) there is equal status between majority and minority groups; (2) contact is endorsed by institutional support, laws, or custom; (3) groups work for a common goal; and (4) there is intergroup cooperation (Allport 1954). However, direct contact can also lead to negative outcomes, such as an increase in prejudice, due to misunderstandings (Paluck et al. 2019) or because an individual has never seen examples of in-group members positively interacting with out-group members. In this case, they do not know how to positively approach the new contact experience.

Exposure to role models with whom a member of a group can identify with can also be a powerful method to induce attitude and behavioral change (Riley 2022; Bernard et al. 2015; DellaVigna and La Ferrara 2015). In this sense, observing in-group role models positively interacting with out-group members can induce an indirect or vicarious contact experience, which has been shown to be an effective method to reduce prejudice (Murrar and Brauer 2018).

Following the contact and role model theories, our project tests if work contact, promoted through direct or indirect contact in the workplace, can improve social cohesion outcomes for both groups. Direct contact consists of refugees and local workers working together directly in a firm, while indirect contact is promoted by showing a video documentary that portrays a refugee and a Ugandan national working together in a firm in Kampala.

The direct contact respects Allport's four conditions. First, to respect the equal status condition, we focus on firm workers from two groups - refugees and locals - that work on similar tasks within a firm. This eliminates any potential hierarchy difference between the employees. For institutional support, we focus only on firms that are willing to participate in the program, thus endorsing the contact between employees. The third and fourth conditions are respected because workers work for the same firm and in the same department, and thus, cooperate towards common goals.

Regarding the role model requirement: that people need to identify with the person they are observing, the video documentary shows relatable and real characters from both groups: a Ugandan worker from Kampala – to relate to local workers – and a refugee worker. We avoid mentioning the nationality of the refugee, in order to make him relatable to all refugee respondents. The documentary is in English with subtitles in the 6 languages spoken by our respondents.[2]

Based on the conceptual framework, we test our main hypothesis that contact in a work setting has a positive effect on social cohesion between refugee workers and local workers (see Section 5 for a detailed description of our measures of social cohesion).

4.2 Experimental design and randomization

Our main treatment is work contact which includes direct and indirect contact at the workplace. We begin by describing what we define "direct" contact. Then we move on to explain the "indirect" contact treatment. The summary of our research design is shown in figure A.1.

4.2.1 Direct contact

In order to promote direct contact, we run a job placement program that assists displaced populations in finding jobs in Uganda. The program provided a one-week internship to skilled refugees at Ugandan firms that were willing to participate. Refugee workers are skilled in vocational occupations such as carpentry, tailoring, and hairdressing (Loiacono and Silva-Vargas 2023). In order to match refugees with firms, we first tested refugees' skills. The test is an official exam run by the Directorate of Industrial Training, the agency established by the Ministry of Education to be in charge of the vocational education curriculum in Uganda. We also had the support of two

[2]The languages are: Luganda, Swahili, Frech, Kinyarwanda, Kirundi, English

large refugee-led NGOs based in Kampala to organize the skills testing. Refugees who passed the test were randomly matched to firms in the same sector as the refugees' occupation. We offered a subsidy to refugees for the one-week internship.[3] Half of the subsidy was paid upon beginning the internship, the other half upon completing it. Local workers are employees already working at the firm. Section 4.3 explains the sampling procedure, that is how local workers were selected.[4]

4.2.2 Indirect contact

The indirect contact took place through a video documentary that we shot in Kampala in March 2021. The video is a short 4-minute documentary about relatable and real-life characters from both groups: Elvis Zani, a Ugandan worker from Kampala – to relate to local workers – and Paul Kithima, an urban refugee worker in Kampala. We avoid mentioning the nationality of the refugee worker to make the main character relatable to all refugees belonging to any nationality. Both workers work together in permaculture.[5] We chose this specific case as we wanted the characters to work in a sector that does not belong to the direct contact treatment, in order to avoid any priming effect.

In the video, both characters talk about their experience working together, what they learned from each other, and what they think about refugees and Ugandans collaborating in the workplace. The video also has a musical background without lyrics that was piloted and it is relatable to all nationalities. Moreover, the video is in English with subtitles in 6 languages (the languages spoken by our respondents) and respondents could decide in which language they wanted the subtitles to be in. The video was piloted with both groups in June 2021 in order to make sure the main message was transmitted, and no other factors were seen as major points. Figure A.2 shows a snapshot of the video.[6]

The placebo video was shown to people that were not assigned to the indirect intervention. The placebo video is a 3-min YouTube video that shows animals in the East African Savannah. We chose this placebo video because we needed something all na-

[3] They were offered 50,000UGX, that is approximately 15USD. This subsidy was substantial and equal to about 85% of the monthly median earnings of the refugees.

[4] Chapter 2 provides a more detailed description of the program as its focus is the impact of internships on firms' willingness to hire refugees

[5] Permaculture is a holistic design system and philosophy that uses principles of ecology and sustainability to create sustainable human settlements and agricultural systems. It emphasizes the use of local resources and the integration of different elements to create a self-sustaining system (Permaculture Research Institute, 2021).

[6] The intervention video can be seen in this link https://www.youtube.com/watch?v=8zTT0VbgKJo

tionalities could relate to, that was not in any specific language and that would not create any particular emotion related to work or contact between groups.[7]

For the analysis, we consider 1 treatment group, which comprises respondents that randomly received direct contact, indirect contact, or both. The control group is composed of refugee workers that are not matched to any firm and local workers that are not matched to work together with refugees, and workers that watch the placebo video. Finally, picture A.3 shows a Ugandan and a refugee participant during the matching phase.

4.2.3 Randomization

There were 3 randomization stages. The first one randomized refugees and firms into direct contact or in the control group. We randomized pairs of firms and refugees working in the same sector: if the refugee was a hairdresser, she was matched to a beauty saloon, etc. The pair was randomly assigned to direct contact following a specific procedure as described in Chapter 2. The second randomization cross-randomized refugees and local workers into indirect contact or in the control group. Finally, due to the cross-randomization, some respondents received both interventions.

4.2.4 Logistics

In order to match refugees and firms, we took refugees to the assigned firms for their first day of work. We organized different groups – according to the location of firms – and gave detailed instructions on the phone to the refugees on how to reach us at a pre-specified landmark, close to the business premises. We instructed enumerators on different tasks to perform during that day: (i) check attendance; (ii) show treatment video individually; (iii) take refugees to the assigned firm; (iv) introduce them to the firm owner; (v) pay refugees the first part of their subsidy. After the week, we sent the rest of the subsidy to refugees by mobile money.

The video was shown individually to respondents by the enumerators using tablets. For refugees, it happened during the "job placement day", when respondents gathered in groups in order to go to the firms they were assigned to. Local workers watched the video soon after the baseline survey.

[7] The placebo video can be seen in this link https://www.youtube.com/watch?v=GBrfomUQXI0. For the version we showed participants, we deleted the beginning where countries are mentioned.

4.2.5 Measures to minimize confounds

In order to avoid spillover, priming and experimenter demand effects we followed several methods.

1. We reduced experimenter demand effects by matching refugee enumerators with the refugee sample – matched by nationality – and Ugandan enumerators with locals. We tested this in our pilot and found that respondents changed answers to some sensitive questions when interviewed by people from different groups compared to their own group.

2. By design, the social cohesion purpose of the study is obfuscated to participants: participants know they are part of a job program, which is about employment assistance, and it is presented in this way to respondents in the consent forms. To confirm this, we elicit refugee workers' beliefs about the study purpose at the end of the program and confirmed that they believed the program was only about job assistance.

3. In order to reduce priming effects regarding the video treatment, main outcomes are collected a week after showing the video to local workers, and six months after for the refugee sample, thus any short-term priming effects are no longer relevant.

4. In order to avoid any effect on social cohesion due to the single act of showing a video to some respondents in front of other respondents, we showed a placebo video to those not assigned to the video treatment.

5. Lastly, to avoid spillover effects regarding the content of the video, we told respondents that the information of the video was confidential and that is why it was shown individually.

4.3 Sampling

In this subsection we describe more in detail the procedure we followed to sample our participants.

4.3.1 Refugees

.

With the collaboration of refugee leaders and refugee-led organizations, we composed a database of 1,088 skilled refugees who were (i) job seekers, (ii) were not looking for jobs but were interested in applying to one if possible, or (iii) were not in permanent employment. We set an appointment and approached the respondents with two messages:

first, to ask some questions regarding their skills and work experience; and second, to explain what the research program was and get consent for it.

The listing was conducted between February and April 2021. From this list, 1,019 refugees agreed to be registered for the program. The first part of the program took place between April 19th and April 24th and consisted of testing refugees on their skills. A final number of 537 refugee workers successfully passed the test of skills. After the skills testing, refugees were invited to participate in the baseline and reminded that some could receive a one-week of internship offer. For our final sample, we had to drop out refugees that never found a match (N=126).[8] Furthermore, we had an attrition at endline of 24 refugees. Our final sample is composed by 377 refugee workers.[9]

4.3.2 Local workers

In June 2021, we conducted a listing survey with firms in Kampala, active in sectors that match the occupations of refugee workers. Using the Uganda Census of Establishment Data 2010, the team of enumerators was assigned to different parishes daily and was instructed to interview all the firms that fell within a sector of interest. Enumerators were instructed to (i) look for the owner, the manager, or any employee with faculty to make managerial decisions; and (ii) the owner must be a Ugandan national.

Due to COVID-19 in the country and two terror attacks, the activities stopped and resumed between September-October 2021, when new firms were recruited. A total of 1,196 firms were recruited but only 536 were willing to hire a refugee. To select local workers, the sampling procedure was: (i) if the firm had only one worker, we interviewed that worker; (ii) if the firm had more than one worker, we asked the owner or manager of the firm which workers were most likely to work in close contact with a new employee.

Since not all the firms in the sample had at least one worker, our final sample of local workers is 273. These are the workers present at baseline and endline. If the worker changed between the two surveys, we kept the baseline answers of the baseline worker, but use the endline replies of the new worker. For this reason, our results are representative of all local workers in the firm, and not of the individual local worker.

[8] That is, firms in the sample in Chapter 2 were not interested in hiring these refugees

[9] During the skills testing some refugees were dropped because they lost interest after registration or because they did not have any of the skills among the ones listed by the program. During the week of skills testing, 402 refugees did not show up. Of the 548 that showed up for the exam dates, 11 people did not pass the test.

4.4 Interventions and follow-ups

The matching of refugees and firms and the 1-week internship happened in October 2021. Soon after the internship, we carried out the endline of firms and local workers between November and December 2021. The endline of refugees happened between July and August 2022. A timeline of data collections and project implementation is reported in figure A.4.

4.5 Covid-19

Uganda has a high informal sector, employing around 80% of the population. In this sector, people cannot work remotely and the informal economy is essential for the daily livelihoods of the majority of the population. Moreover, due to COVID-19 lockdowns, 95% of employees were let go and some moved back to agriculture (Alfonsi et al. 2021). Yet, urban refugees do not have plots of land in the country, and therefore, they have remained in the cities facing higher levels of hunger and unemployment. According to the World Bank, refugees in Uganda will need higher assistance in order to avoid a poverty trap due to COVID-19 (Aramanov et al. 2021). Appendix C describes the COVID-19 prevention plan that we followed during our activities.

5 Data

In this section we describe the data we collect and detail how we use them in our analysis.

5.1 Outcome variables

We collect data at baseline and endline. Due to the design of the project, for local workers, there was around 1 month between baseline and endline data collection. For refugee workers, there are around 6 months between baseline and endline.
We have four main types of outcomes to capture the most important dimensions of social cohesion: implicit bias, explicit stereotypes, attitudes, and behaviors. A meta-analysis of contact projects to reduce prejudice found that behavioral change is not accompanied by attitudinal change, few studies capture both dimensions, and there is almost no evidence that implicit bias is related to discriminatory behavior (Paluck et al. 2020). For this reason, we collect implicit bias, explicit stereotypes, attitudes towards the out-group, and hypothetical behaviors, which are collected both at baseline and endline. Real incentivised behavior is collected only at endline. We specify each component and outcome below.

5.1.1 Implicit bias: Implicit Association Tests measurement

Implicit association tests (IATs) are psychological tools that capture biases using "categorization tasks" (Greenwald and Banaji 1995). A series of stimuli is shown on the screen, and the respondent must sort them into two categories. The main assumption is that the stronger the association a respondent makes between a stimulus and a group (in our case, refugee or local), the faster they make these associations.

We followed the "classic" IAT design with seven rounds (Greenwald et al. 2003). Two initial training rounds to practice sorting stimuli into two categories of the same concept (stimuli into positive or negative or into refugee or local). A "stereotypical" pairing where stimuli from all concepts are shown. Respondents categorize these stimuli into the two concepts "stereotypically" combined on the same side of the screen: e.g. for local workers, refugees and negative are on one side, local and positive on the other (Figure A.9 in the appendix B). Another training round, where respondents practice swapping left and right for one category. Finally, the "non-stereotypical" pairing: concepts are "not stereotypically" combined: e.g. for local workers, refugees and positive are now on the same side, local and negative on the other. We go a step further and randomize the order of the "stereotypical" and "non-stereotypical" rounds.

Faster associations reflect higher implicit associations between the concepts. For example, if a respondent responds faster when refugees and negative are on the same side, she associates refugees with negative stereotypes. The final IAT score is the normalized difference in response times between the "stereotypical" and "non-stereotypical" groups. A higher score is a proxy for more implicit bias.

We run two IATs: one to measure implicit biases towards the out-group's work characteristics ("Work IAT") and one to measure implicit biases towards the out-group's general, stereotypical characteristics ("General IAT"). The reason why we use two different IATs is that, in our context, a co-worker could be differently biased towards the work abilities of the out-group member but not in general against them, and/or viceversa. For example, a person can be unbiased towards refugees as neighbors or friends because he or she implicitly believes refugees have positive general characteristics (such as being friendly or generous). Yet, the same person can be biased towards refugees as co-workers because he or she implicitly believes refugees have fewer skills. The order of the two IATs was randomized in each survey.

The words for the Work and General IATs were selected for two main reasons: first, they were piloted extensively with Ugandan workers and refugee workers to capture words that would reflect mostly one of each context (work or general). Second, the words could be translated in 6 languages (5 for refugees and 2 for locals). The words we use for each IAT are specified in Figure A.9 in the appendix B.

For the main analysis, we construct an index averaging the two IATs scores. We refer to this index as "Implict Bias". Figure 2a and 2b show the density of the Work and General IAT at baseline respectively. The IATs are coded so that higher values denote more implicit bias. The pattern that emerges is that locals are implicitly more biased than refugees: the mean of the General IAT for locals is 5 times higher than the one for refugees (which is close to zero). For Work IAT, the mean for locals is double. K-Smirnov tests show that the distributions differ significantly one from each other.

5.1.2 Explicit bias: Explicit stereotypes and attitudes

To measure explicit stereotypes, we directly ask the respondent to rank the same stimuli shown in the two IATs (Figure A.9 in appendix B) related to the out-group using a 7-points Likert-scale. For attitudes, we ask respondent if they agree with a series of statements related to culture, trust, safety, intermarriage, job collaboration, and perceived discrimination. We ask the same statements for local and refugee workers. Again, respondents could select any answer using a 7-points Likert-scale. Attitudes were selected after collecting pilot data and focus group discussions with refugees and locals where we directly asked them which attitudes were signals of integration of refugees in the country. Appendix B lists the statements.
We randomize the order of the explicit stereotypes, attitudes and IATs. For the main analysis, we create an index that combines explicit stereotypes and attitudes. We use the GLS weighting procedure as described in Anderson (Anderson 2008).[10] We refer to this index as "Explicit Bias".

5.1.3 Behaviors

For this dimension, we collect evidence of two types of behaviors: real and hypothetical. For real behavior we ask refugee workers that if they would like to participate in a similar program (the internship at firms) in the future, they can send a SMS to a telephone number. In the SMS, they need to specify if they would like to work with a Ugandan firm or with a refugee firm.[11]
For hypothetical behaviors, we ask a question at baseline and endline. We elicit respondents' willingness to work with an out-group member in the future. The question asks to imagine a hypothetical scenario where respondents can start a new business.

[10] We use the command `swindex` in Stata (Schwab et al. 2020).
[11] The SMS outcome was only asked to the refugee sample.

They can choose the number of business partners and their nationality. All questions are reported in the appendix B.

5.1.4 Correlations and IAT interpretation

In our baseline data, we find two opposite trends. First, for locals, we observe that the combined IAT index is not correlated with explicit bias: while the point estimate is negative, it is not significantly different from zero.[12] When we look at the correlations with Work IAT and General IAT separately, only the Work IAT is significantly and negatively correlated with the full explicit bias index (Figure A.5 in appendix A). Instead for refugees, we observe the opposite. The IAT is positive and significantly correlated with explicit bias, both the full IAT index as well as the General and Work IAT separately.

These correlations are suggestive evidence that the IAT might be measuring different things for different groups. The main motive why this might be happening is that refugees might feel more open to explicitly expressing their bias (i.e. talking with refugee enumerators), while locals might feel judged expressing their implicit bias explicitly, as they are the majority group and "hosts" of a vulnerable group (even if talking with local enumerators).

Another possible explanation could be that refugees, who have had several interactions with the out-group while living in Uganda, had experienced more difficulties with the out-group. Therefore, the IAT is indeed measuring general bias towards the out-group. Table A.1 shows correlations between the full IAT and some single variables that compose the explicit bias index. The table suggests that the IAT is correlated to negative attitudes: believing that Ugandans are less friendly, that intermarriage is not good, that their culture is not similar and trusting more refugees. Instead for local workers, Table A.2 shows that the IAT is correlated to positive attitudes: trust more refugees, and that refugees are more likely to be serious and diligent (coefficient of the variable "Unserious").

When we look at behavioral outcomes that were collected also at baseline, we observe that locals' combined IAT is negative and significantly correlated with altruism (Table A.2), while for refugees is the opposite (Table A.1). Refugees' IAT is also positively correlated with reciprocity. As for business partners in a hypothetical scenario, only the refugees' IAT is negatively correlated with having any business partner.

[12]We evaluate baseline correlations using the following specification: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, where y_i is the combined index for explicit bias, and x_i are three different covariates: i) the combined IAT, ii) the work-related IAT, and iii) the general IAT.

These correlations can be suggestive evidence that refugees that have experienced more difficulties in the country increase their altruism and reciprocity towards the most vulnerable groups. Indeed, the refugees' IAT is correlated to more trust towards other refugees.

5.2 Descriptive statistics

In this study, our sample is composed by 650 employees working in firms situated in Kampala, Uganda. Of these, 273 are local workers, e.g. native people born in Uganda, and 377 are refugee, mostly of Congolese nationality. Since we pooled the treatments into one for the main analysis, our sample in each group is as follows: 236 local workers and 288 refugee workers in the treatment group; 37 local workers and 89 refugee workers in control group.

Tables 1 and 2 report summary statistics at baseline for the main outcomes of interest and some controls for local and refugees workers, respectively. The tables report the control group (column 1), the pooled treatment group (column 2), and the full sample (column 3). The last column is the difference in means between the control and pooled treatment group.

Almost all outcome variables and controls are balanced between groups for both samples.[13] On average, the sample of refugee workers is older than local workers: 34 vs 24 years old. Refugee workers have more experience in the sector compared to local workers: 57.5 vs 42.5 months. Yet refugee workers have lower English and Luganda (self-reported) scores. Females are the majority in both samples. Furthermore, local workers are more biased at baseline than refugees: their General IAT score is 5 times higher than refugees' one, while their Work IAT score is almost double. Also, 45% of local workers said that they would like an out-group business partner, while 85% of refugee stated the same.

6 Analysis

6.1 Empirical strategy

To estimate the effect of our pooled treatment on social cohesion, we pool the two samples. We do so in order to estimate the effect of the treatment on both refugee and

[13] There are two variables that are statistically different between control and treatment groups for local workers: if they want a business partner from the same group (local) or any business partner. Including the baseline value of these variables does not change the results.

local workers' outcomes, and assess if the effect is different between the two samples. Therefore, our total sample is composed by 650 workers.
For the analysis of implicit and explicit bias, we pool the indices together into one variable called "Bias Index". We do so to jointly test across regression equations, increasing the statistical power while reducing the concern of multiple hypothesis testing. Therefore, each observation is repeated twice, one for implicit bias and one for explicit bias. Our total sample in this case would be 1300.

For the analysis of implicit and explicit bias we run the following specification:

$$\begin{aligned} BiasIndex_{i1} = &\beta_1 T \times local \times implicit + \beta_2 T \times local \times explicit + \\ &+ \beta_3 T \times refugee \times implicit + \beta_4 T \times refugee \times explicit \\ &+ \beta_5 local \times implicit + \beta_6 local \times explicit + \\ &+ \beta_7 refugee \times implicit + \beta_8 refugee \times explicit + \\ &+ \alpha BiasIndex_{i0} + X_i' \delta + \varepsilon_i \end{aligned} \tag{1}$$

where $BiasIndex_{i1}$ takes the value from the implicit or explicit bias measures for worker i at follow-up. The first four explanatory variables are dummies equal to 1 for workers assigned to the treatment and the other two conditions. For example, $T \times local \times implicit$ is a dummy equal to 1 for local workers assigned to the treatment group interacted with the index for implicit bias, and so on. We include four indicators, such as $local \times implicit$, that are dummies equal to 1 when the worker is local and the index is implicit, and so on for the rest of the indicators. We control for the baseline value of the outcome $BiasIndex$ and X_i' is a matrix of the randomization strata (the occupations of the refugee workers). This specification uses robust standard errors.
Our coefficients of interest are β_1, β_2, β_3, and β_4. A positive value indicates an increase in bias. Both the implicit and explicit were normalized from 0 to 1 for comparison. We also run a test of equality between the local and refugee workers to test if the coefficients are equal to each other.
The second specification is as follows:

$$y_{i1} = \beta_1 Treatment_i + \beta_2 Local_i + \beta_3 Treatment_i \times Local_i + \alpha y_{i0} + X_i' \delta + \varepsilon_i \tag{2}$$

where y_{i1} is the outcome of worker i at follow up. $Treatment_i$ equals to 1 if the worker was assigned to either direct or indirect contact. $Local$ is a dummy equal to 1 if the worker is Ugandan. $Treatment \times Local$ is an interaction term. When possible, we control for the baseline value of the outcome y and X_i' is a matrix of the randomization strata.

Our coefficients of interest are β_1 and β_3. The former shows whether the treatment had a significant effect for refugee workers, the latter whether the effect for locals was different from that for refugees. To get the effect for locals, we sum β_1 and β_3 and run the test $\beta_1 + \beta_3 = 0$.
The third specification is as follows:

$$y_{i1} = \beta_1 Treatment_i + X_i' \delta + \varepsilon_i \tag{3}$$

where y_{i1} is the outcome of worker i at follow up. $Treatment_i$ equals to 1 if the worker was assigned to the treatment. X_i' is a matrix of the randomization strata. Our coefficient of interest is β_1 which shows whether the treatment had a significant effect for the workers.

6.2 The Impact of Contact on Implicit and Explicit Bias

Table 3 contains our first results on the effect of exposure to work contact for local and refugee workers, compared to workers that do not work with the out-group. The dependent variable is explicit and implicit bias, using specification 1. Column 1 reports estimated coefficients for refugees and local workers separately on an aggregate index of bias. In the Appendix, Tables A.3 and A.4 report also the results with the treatments separately.
First, for local workers, we find that exposure to work contact has a significant effect on both implicit and explicit bias. We see that explicit bias significantly decreases by 0.056 standard deviations, while implicit bias significantly increases by 0.084 standard deviations. For refugee workers, we see a similar trend: explicit bias significantly decreases by 0.039 standard deviations, while implicit bias also increases, but it is not significant. Moreover, we run tests of equality and we cannot reject that the coefficients between locals and refugees are the same.
These results go against other experimental work that has found that contact reduces implicit bias. For instance, Corno et al. (2022) found that sharing a room at a university reduces the implicit bias of the majority group. Yet, we believe our context is different since work contact provides a layer of competition between workers. In fact, columns 2 and 3 show the effect of work contact on work and general bias separately. We see that the effect is significant for local workers' work implicit bias. The effect on general implicit bias is positive for refugee and local workers, but not significant. Therefore, the increase in implicit bias is led by work-induced implicit bias.

6.3 The Impact of Contact on Hypothetical and Real Behaviors

Table 4 contains our second set of results. The dependent variable is partners in a hypothetical business scenario, using specification 2. In the Appendix, Tables A.3 and A.4 report also the results with the treatments separately.

Working together for one week or watching a video where two workers from different groups collaborate, increases the local workers' willingness to have a business partner from the out-group, an effect of almost 17 percentage points, which is around 42 percent increase over the mean. For refugee workers work contact decreases their willingness to work with a local but it is not significant. Yet, the refugee worker's willingness to work with any partner significantly decreases by 5 percentage points. We can reject that the effect is the same between groups in these two variables.

Table 5 shows the final set of results using specification 3. We asked refugee workers to send an SMS if they are interested in working in a similar internship matching program in the future, and the nationality of the potential employer. We find that the effect is significant and positive: treated refugees are 11 percentage points more willing to work in a similar program compared to the control group, which implies an effect of 90% over the mean. Moreover, the effect is significant and positive for their willingness to work with a Ugandan firm, but not for a refugee firm. In the Appendix, Table A.5 reports also the results with the treatments separately.

6.4 Robustness checks

In this subsection we perform some robustness checks to our main analysis. We begin by studying whether results are affected by attrited workers at endline. To do so we bound our estimates using the Kling and Liebman sensitivity bounds (Kling and Liebman 2004). In this analysis we ask what the results would have been if "unfound" workers differ by 0.25 s.d. from those who are found.

Table A.6 shows that our main results fall within the 0.25SD bounds, as the point estimates are virtually unchanged.

The second robustness check is shown in Table A.7. For this robustness check we construct the explicit index using the Principal Component Analysis method and choosing the first component. The refugee workers' explicit bias goes down and it is significant, as in our main analysis. The local workers' explicit bias is not significant in this analysis but the coefficient is negative, as in our main analysis. Also, we cannot reject the null hypothesis that they are not different from each other.

Finally, we run another robustness check by adding more variables to the explicit index for locals. The variables were collected only for local workers because of the content of the question tailored on the local population. We asked workers to what extent they agree with the following questions: (i) Working with a refugee will reduce my productivity, (ii) Refugees increase job competition in the country. Finally, we asked: (iii) Think about the current law on refugees in Uganda, do you think that the law should allow all refugees to work anywhere in Uganda? Possible answer was between Yes or No. Table A.8 shows the results. We find the exact same result as in our main analysis: the explicit bias reduces, with a coefficient slightly larger: -0.062 and the same significance level of 10%.

6.5 Discussion

Contact on the workplace improve refugee and local workers' attitudes towards each other. We believe that the results of our study demonstrate two key points. First, it appears that treated local workers are keen to work with refugee workers in the future due to the high level of skill that the refugee workers possess, as the sample of refugee workers was selected based on their good skills. This is not what the local workers initially expected. As shown by Figure A.6, the locals' original beliefs regarding the refugees' experience in the sector was 32 months, but in reality, this sample of refugee workers had an average experience of 57.5 months, almost double the local workers' original expectations. Additionally, the average experience of the local workers was 42.5 months, which is lower than that of the refugees. Additionally, Figure A.7 demonstrates that local workers typically have less education compared to refugees.

The local workers' exposure to the refugees' skills may also contribute to an increase in implicit bias. As they work alongside these highly skilled workers, they may develop an implicit fear of job competition. Our regression analysis (Equation 2) was performed with single outcome variables relating to job competition and work collaboration, taking into consideration that these views may not align with the workers' implicit feelings. Figure A.8 shows that the impact of work contact on these variables is not statistically significant. However, the coefficient for the view that refugees increase job competition is positive, while the coefficients for negative views on work collaboration are negative. Despite this implicit fear or bias, local workers still express a desire to collaborate with refugee workers in future business ventures.

Second, treated refugee workers are less interested in starting a business with any partners, but they are more interested in employed work, particularly in Ugandan firms. This provides suggestive evidence that treated refugee workers have learned about the advantages of working for an established company rather than becoming entrepreneurs, which is instead a common coping strategy among refugees in Kampala to improve their livelihoods and sustain themselves.

7 Conclusion

Using a randomized controlled trial with local and refugee workers, this paper examines the impact of work contact on social cohesion in Uganda, the largest refugee-hosting country in Africa. The study shows that work contact can improve social cohesion by reducing explicit biases and increasing positive behaviors. Local workers have increased implicit biases, but this does not lead to discriminatory behavior. In fact, they are more willing to work with refugee workers in hypothetical business scenarios, likely due to the higher skills of the refugees compared to their own. For refugee workers, explicit biases significantly decrease, and their willingness to participate in similar job programs in the future increases, particularly with Ugandan firms. This study highlights the importance of work contact as a means of promoting social cohesion in refugee-hosting countries and adds to the existing literature on the integration of refugees into society and the economy.

Regarding the measurement and interpretation of implicit bias, this study supports the socio-psychological literature that states that implicit and explicit biases are distinct and largely unrelated. Although implicit bias increases, explicit bias decreases. This study also provides evidence on the relationship between implicit bias and prejudiced behavior, showing that implicit bias does not necessarily lead to discriminatory behavior. Alesina et al. (2018) found that math teachers with strong implicit biases tended to give lower grades to immigrant students compared to local students, while literature teachers did not act on their implicit biases.

Our study is an important contribution to the discussion, particularly in light of the recent surge in workplace implicit bias trainings aimed at mitigating prejudiced behavior. Our findings suggest that these interventions may not be optimal for achieving the desired outcome, thereby warranting careful consideration and potential reevaluation of prevailing approaches in addressing workplace discrimination.

Regarding the external validity of our findings. It is important to note that our findings may not be applicable to all refugee-hosting countries. Our two samples are plausibly representative of workers and refugees in urban areas, who tend to have higher skills compared to those in rural areas due to better access to education and employment opportunities. However, in some countries, refugees may be required to stay in settlements and may not have the same access to skills development and employment opportunities as those in urban areas. In future studies, a different type of contact in the workplace could be tailored to study what aspects of work collaboration increase social cohesion.

References

Alesina, A., Carlana, M., La Ferrara, E., and Pinotti, P. (2018). Revealing stereotypes: Evidence from immigrants in schools. Working paper, National Bureau of Economic Research.

Alfonsi, L., Bassi, V., Manwaring, P., Ngategize, P., Oryema, J., Stryjan, M., and Vitali, A. (2021). The impact of covid-19 on ugandan firms. Technical report, International Growth Center (IGC) Uganda.

Allport, G. W. (1954). *The nature of prejudice*. Garden City, NJ Anchor, 1954.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, 103(484):1481–1495.

Andreychik, M. R. and Gill, M. J. (2012). Do negative implicit associations indicate negative attitudes? social explanations moderate whether ostensible ânegativeâ associations are prejudice-based or empathy-based. *Journal of Experimental Social Psychology*, 48(5):1082–1093.

Aramanov, A., Beltramo, T., Waita, P., and Nobuo, Y. (2021). Covid-19 socioeconomic impact worsens for refugees in uganda. Technical report, World Bank Blog.

Bahar, D., Ibáñez, A. M., and Rozo, S. V. (2021). Give me your tired and your poor: Impact of a large-scale amnesty program for undocumented refugees. *Journal of Development Economics*, 151:102652.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful women: does exposure reduce bias? *The Quarterly journal of economics*, 124(4):1497–1540.

Bernard, T., Dercon, S., Orkin, K., and Seyoum Taffesse, A. (2015). Will video kill the radio star? assessing the potential of targeted exposure to role models through video. *The World Bank Economic Review*, 29(suppl_1):S226–S237.

Bertrand, M. and Duflo, E. (2017). Field experiments on discrimination. *Handbook of economic field experiments*, 1:309–393.

Blattman, C. and Annan, J. (2016). Can employment reduce lawlessness and rebellion? a field experiment with high-risk men in a fragile state. *American Political Science Review*, 110(1):1–17.

Bursztyn, L., Chaney, T., Hassan, T., and Rao, A. (2023). The immigrant next door. *American Economic Review, Forthcoming*.

Caria, A. S., Gordon, G., Kasy, M., Quinn, S., Shami, S., and Teytelboym, A. (2023). An adaptive targeted field experiment: Job search assistance for refugees in jordan. *Journal of the European Economic Association*.

Carrell, S. E., Hoekstra, M., and West, J. E. (2015). The impact of intergroup contact on racial attitudes and revealed preferences. Working paper, National Bureau of Economic Research.

Corno, L., La Ferrara, E., and Burns, J. (2022). Interaction, stereotypes, and performance: Evidence from south africa. *American Economic Review*, 112(12):3848–75.

Cunningham, T. and De Quidt, J. (2023). Implicit preferences. Working paper.

De Berry, J. P. and Roberts, A. J. (2018). Social cohesion and forced displacement: a desk review to inform programming and project design. Technical report, The World Bank.

DellaVigna, S. and La Ferrara, E. (2015). Economic and social impacts of the media. In *Handbook of media economics*, volume 1, pages 723–768. Elsevier.

Easterly, W., Ritzen, J., and Woolcock, M. (2006). Social cohesion, institutions, and growth. *Economics & Politics*, 18(2):103–120.

Greenwald, A. G. and Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, 102(1):4.

Greenwald, A. G., Nosek, B. A., and Banaji, M. R. (2003). Understanding and using the implicit association test: I. an improved scoring algorithm. *Journal of personality and social psychology*, 85(2):197.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., and Banaji, M. R. (2009). Understanding and using the implicit association test: Iii. meta-analysis of predictive validity. *Journal of personality and social psychology*, 97(1):17.

Loiacono, F. and Silva-Vargas, M. (2019). Improving access to labor markets for refugees: Evidence from uganda. Technical report, International Growth Center (IGC) Uganda.

Loiacono, F. and Silva-Vargas, M. (2023). Matching with the right attitude: The effect of matching firms with refugee workers. Working paper.

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review*, 111(6):1807–44.

Mousa, S. (2020). Building social cohesion between christians and muslims through soccer in post-isis iraq. *Science*, 369(6505):866–870.

Munshi, K. (2011). Labor and credit networks in developing economies. In *Handbook of Social Economics*, volume 1, pages 1223–1254. Elsevier.

Murrar, S. and Brauer, M. (2018). Entertainment-education effectively reduces prejudice. *Group Processes & Intergroup Relations*, 21(7):1053–1077.

Okunogbe, O. M. (2023). Does exposure to other ethnic regions promote national integration? evidence from nigeria. *American Economic Review - Forthcoming.*

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., and Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: a meta-analysis of iat criterion studies. *Journal of personality and social psychology*, 105(2):171.

Paluck, E. L., Green, S. A., and Green, D. P. (2019). The contact hypothesis reevaluated. *Behavioural Public Policy*, 3(2):129–158.

Paluck, E. L., Porat, R., Clark, C. S., and Green, D. P. (2020). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, 72.

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in delhi schools. *American Economic Review*, 109(3):774–809.

Riley, E. (2022). Role models in movies: the impact of queen of katwe on students' educational attainment. *Review of Economics and Statistics*, pages 1–48.

Scacco, A. and Warren, S. S. (2018). Can social contact reduce prejudice and discrimination? evidence from a field experiment in nigeria. *American Political Science Review*, 112(3):654–677.

Schwab, B., Janzen, S., Magnan, N. P., and Thompson, W. M. (2020). Constructing a summary index using the standardized inverse-covariance weighted average of indicators. *The Stata Journal*, 20(4):952–964.

Singal, J. (2017). Psychology's favorite tool for measuring racism isn't up to the job. *The Cut*.

Uhlmann, E. L., Brescoll, V. L., and Paluck, E. L. (2006). Are members of low status groups perceived as bad, or badly off? egalitarian negative associations and automatic prejudice. *Journal of Experimental Social Psychology*, 42(4):491–499.

Verwimp, P., Justino, P., and Brück, T. (2019). The microeconomics of violent conflict. *Journal of Development Economics*, 141:102297.

8 Figures and Tables

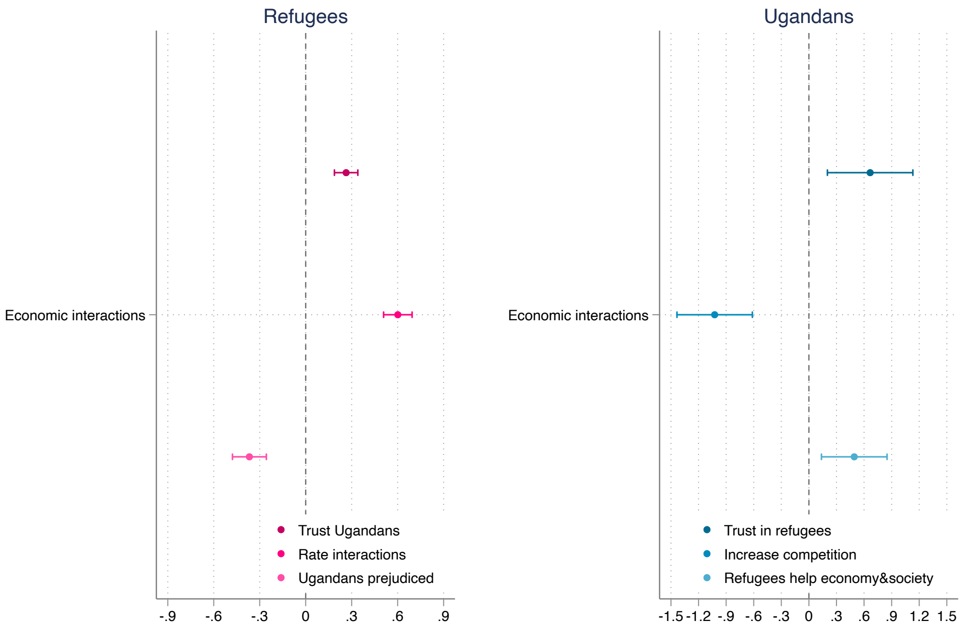
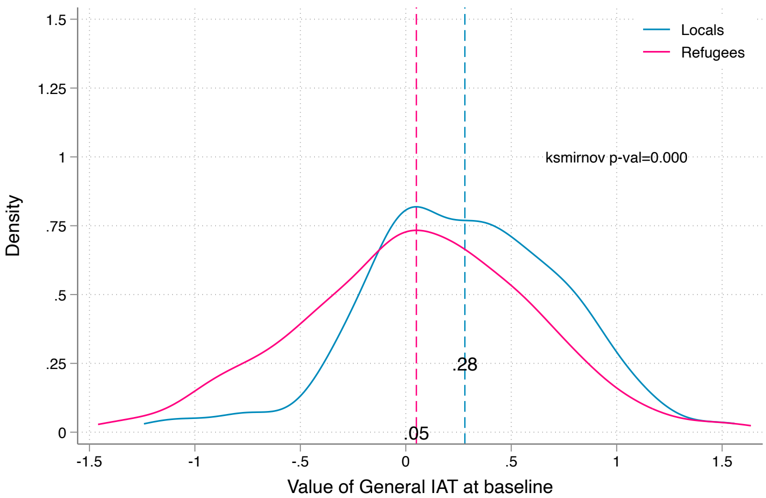


Figure 1: Interactions with outgroups and social cohesion

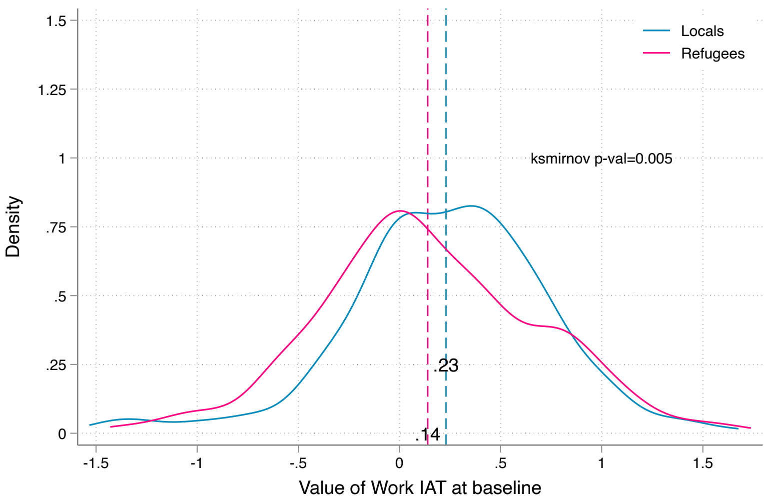
Notes: This graph uses the data from a pilot survey in the cities of Kampala and Mbarara, targeting approximately 400 refugees and 400 locals. It plots the coefficients from the following regression: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, where y_i is number of days the refugee/firm owner has economic interactions with the out-group (Ugandans and refugees, respectively), and x_i is one of the three social cohesion covariates asked to the refugees sample: *"Rate on a scale between 0 and 10, where 0 means Not at all and 10 Very much: i) how much do you trust Ugandans; ii) your interactions with Ugandans"; iii) how much you think Ugandans are prejudiced against refugees"* or those asked to the locals sample: *"Rate on a scale between 0 and 10, where 0 means Not at all and 10 Very much: i) how much do you trust refugees; ii) refugees increase job competition in the country; iii) refugees help the country economically and socially"*. Standard errors are robust.

Figure 2: Distribution of Implicit Association Test scores

(a) General IAT



(b) Work IAT



Notes: This graph depicts the distribution of the IAT score, where a higher score means that there is higher bias against the outgroup. Panel A plots the density of the scores from the Implicit Association Test (IAT) with general words concerning quality of the outgroup. Panel B shows the same distribution for the score on the IAT with work-related words.

Table 1: Descriptive and balance checks: local workers

		(1)		(2)		(3)	T-test
		Control		Pooled Treatment		Full sample	Difference
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)
Age	37	25.243	236	24.186	273	24.330	1.057
		(1.287)		(0.455)		(0.430)	
English score	37	0.605	236	0.565	273	0.570	0.040
		(0.041)		(0.016)		(0.015)	
Luganda score	37	0.709	236	0.700	273	0.702	0.009
		(0.031)		(0.014)		(0.012)	
Female dummy	37	0.568	236	0.644	273	0.634	-0.077
		(0.083)		(0.031)		(0.029)	
Experience in months	37	49.297	236	41.441	273	42.505	7.857
		(6.831)		(2.815)		(2.604)	
General IAT	37	0.240	236	0.282	273	0.276	-0.042
		(0.071)		(0.031)		(0.029)	
Work IAT	37	0.338	236	0.214	273	0.231	0.124
		(0.079)		(0.033)		(0.031)	
Implicit bias index	37	0.289	236	0.248	273	0.254	0.041
		(0.058)		(0.023)		(0.021)	
Explicit bias index	37	0.000	236	0.074	273	0.064	-0.074
		(0.164)		(0.069)		(0.064)	
Outgroup business partner	37	0.351	236	0.466	273	0.451	-0.115
		(0.080)		(0.033)		(0.030)	
Same business partner	37	1.000	236	0.924	273	0.934	0.076*
		(0.000)		(0.017)		(0.015)	
Any business partner	37	1.000	236	0.928	273	0.938	0.072*
		(0.000)		(0.017)		(0.015)	

Note: This table shows descriptive statistics and balance across treatment and control groups in the sample of local workers. English and Luganda scores are an average of self reported measures of reading, writing, speaking and listening of the different languages. Experience in months is the experience in the sector. General and Work IATs are the scores obtained after completing the implicit association tests. Implicit bias index is the average of the two IATs. Explicit bias index is contructed using the GLS Anderson weighting procedure combining negative attitudes and explicit negative stereotypes. Business partners variables are dummies indicating if they wanted a partner or not. The value displayed for t-tests are the differences in the means across the groups. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively.

Table 2: Descriptive and balance checks: refugee workers

	(1)		(2)		(3)		T-test
	Control		Pooled Treatment		Full sample		Difference
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)
Age	89	33.989	288	33.642	377	33.724	0.346
		(1.005)		(0.614)		(0.525)	
English score	89	0.449	288	0.414	377	0.422	0.035
		(0.030)		(0.016)		(0.014)	
Luganda score	89	0.382	288	0.374	377	0.376	0.008
		(0.032)		(0.016)		(0.014)	
Female dummy	89	0.618	288	0.667	377	0.655	-0.049
		(0.052)		(0.028)		(0.025)	
Experience in months	89	61.373	288	56.254	377	57.462	5.120
		(9.028)		(4.847)		(4.268)	
General IAT	89	0.099	288	0.039	377	0.053	0.060
		(0.055)		(0.033)		(0.028)	
Work IAT	89	0.187	288	0.122	377	0.138	0.065
		(0.051)		(0.033)		(0.028)	
Implicit bias index	89	0.139	288	0.083	377	0.096	0.056
		(0.048)		(0.028)		(0.024)	
Explicit bias index	89	-0.000	288	0.011	377	0.009	-0.011
		(0.106)		(0.060)		(0.052)	
Same business partner	89	0.955	288	0.934	377	0.939	0.021
		(0.022)		(0.015)		(0.012)	
Outgroup business partner	89	0.809	288	0.872	377	0.857	-0.063
		(0.042)		(0.020)		(0.018)	
Any business partner	89	0.978	288	0.972	377	0.973	0.005
		(0.016)		(0.010)		(0.008)	

Note: This table shows descriptive statistics and balance across treatment and control groups in the sample of refugee workers. English and Luganda scores are an average of self reported measures of reading, writing, speaking and listening of the different languages. Experience in months is the experience in the sector. General and Work IATs are the scores obtained after completing the implicit association tests. Implicit bias index is the average of the two IATs. Explicit bias index is contructed using the GLS Anderson weighting procedure combining negative attitudes and explicit negative stereotypes. Business partners variables are dummies indicating if they wanted a partner or not. The value displayed for t-tests are the differences in the means across the groups. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively.

Table 3: The Effect of Contact on Implicit and Explicit Bias

	(1) Bias Index	(2) Work bias	(3) General bias
T × Local × Implicit	0.084**	0.083**	0.040
	(0.035)	(0.037)	(0.029)
	[0.017]	[0.026]	[0.162]
T × Local × Explicit	-0.056*	-0.053	-0.025
	(0.033)	(0.035)	(0.026)
	[0.093]	[0.122]	[0.322]
T × Refugee × Implicit	0.021	0.011	0.016
	(0.024)	(0.024)	(0.029)
	[0.377]	[0.650]	[0.594]
T × Refugee × Explicit	-0.039**	-0.046**	-0.041**
	(0.019)	(0.023)	(0.020)
	[0.043]	[0.049]	[0.040]
Observations	1200	1172	1170
Mean DV	0.480	0.445	0.470
Mean DV Local Implicit Bias	0.460	0.426	0.534
Mean DV Refugee Implicit Bias	0.405	0.460	0.476
Mean DV Local Explicit Bias	0.533	0.443	0.448
Mean DV Refugee Explicit Bias	0.531	0.441	0.448
H_0 : T × Local × Implicit=Refugee	0.134	0.105	0.547
H_0 : T × Local × Explicit=Refugee	0.659	0.853	0.623

Notes: This table reports results from specification 1. Explicit bias index is constructed using the GLS Anderson weighting procedure combining negative attitudes and explicit negative stereotypes. Implicit bias is an average of Work IAT score and General IAT score. Both indices are normalized 0 to 1 for comparisson. An increase means more prejudice. Control for refugees strata (refugees' occupations). Robust standard errors in parenthesis and p-values in brackets. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. The sample is not 1300 because we have 56 missing IATs at baseline and 44 missing IATs at endline.

Table 4: The Effect of Contact on Desired Hypothetical Business Partners

	(1) Out-group	(2) Same group	(3) Any partner
Treated	-0.063	-0.027	-0.051**
	(0.044)	(0.034)	(0.025)
	[0.155]	[0.436]	[0.044]
Local	-0.393***	-0.009	-0.054
	(0.092)	(0.053)	(0.048)
	[0.000]	[0.869]	[0.261]
Treated × Local	0.231**	0.059	0.094*
	(0.098)	(0.058)	(0.054)
	[0.019]	[0.316]	[0.083]
Observations	650	650	650
Mean DV	0.722	0.921	0.952
Mean DV Locals	0.405	0.919	0.919
Mean DV Refugees	0.854	0.897	0.966
Treated + Local × Treated	0.168	0.032	0.043
H_0: Treated + Treated × Local=0	0.055	0.501	0.362

Notes: This table reports results from specification 2. The outcome variables are dummies indicating if respondents want a business partner or not. Control for refugees strata (refugees' occupations). Robust standard errors in parenthesis and p-values in brackets. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively.

Table 5: SMS sent by refugee workers

	(1) Sent SMS	(2) SMS Ugandan	(3) SMS Refugee
Treated	0.113**	0.062**	0.052
	(0.044)	(0.025)	(0.038)
	[0.010]	[0.016]	[0.174]
Observations	377	377	377
Mean DV	0.124	0.034	0.090

Notes: This table reports results from specification 3. Sent SMS outcome is a dummy indicating if refugee workers sent a SMS to participate in similar future internship programs. SMS for Ugandan and for refugee firm indicate what type of firm the worker would like to work in future interventions. Control for refugees strata. Robust standard errors in parenthesis and p-values in brackets. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively.

A Appendix A: Figures and extra analyses

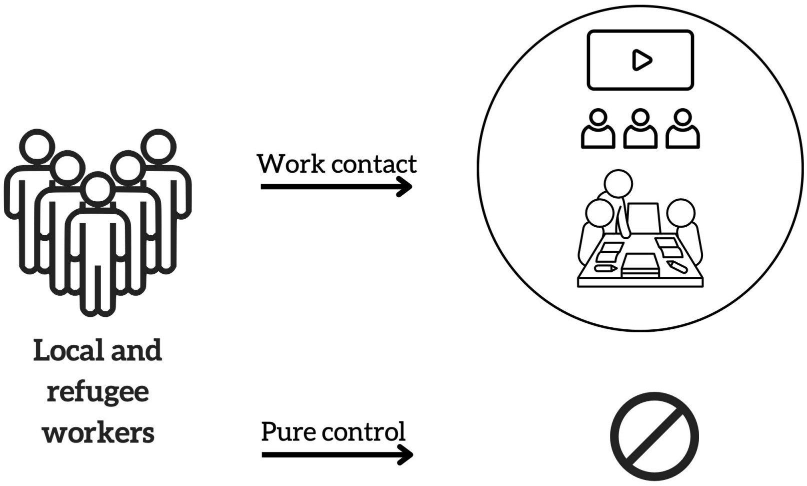


Figure A.1: RCT Design

Notes: This graph shows the experimental design. Couples of refugee and Ugandan workers are assigned to either a "Work contact" treatment arm or to a pure control group, where no contact takes place.

Figure A.2: Snapshot of the video documentary

Notes: Elvis and Paul collaborating on a permaculture project. © Mariajose Silva-Vargas

Figure A.3: Example of contact on the workplace

Notes: Sifa, a Congolese worker (on the right) working for Mariam, a Ugandan firm owner (on the left). © Mariajose Silva-Vargas

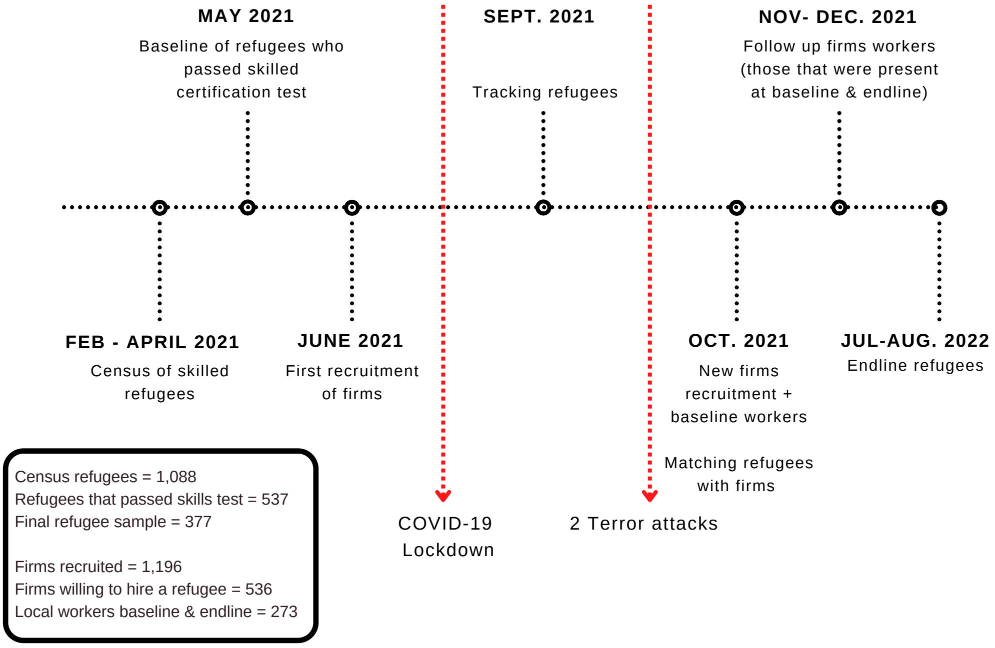
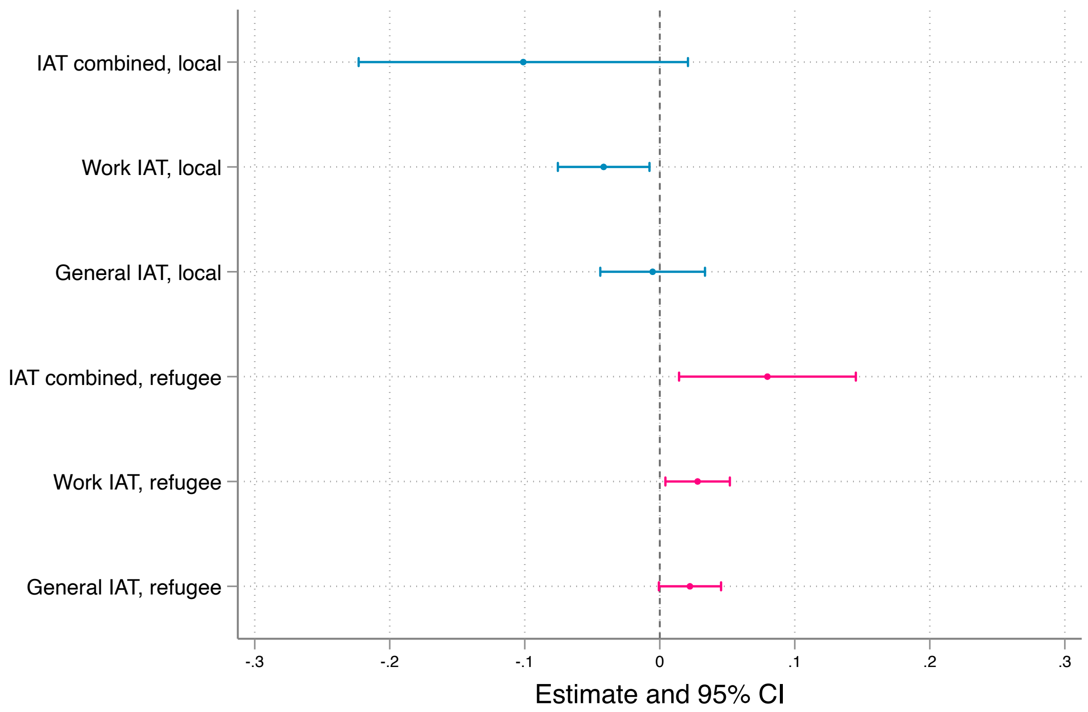


Figure A.4: Timeline

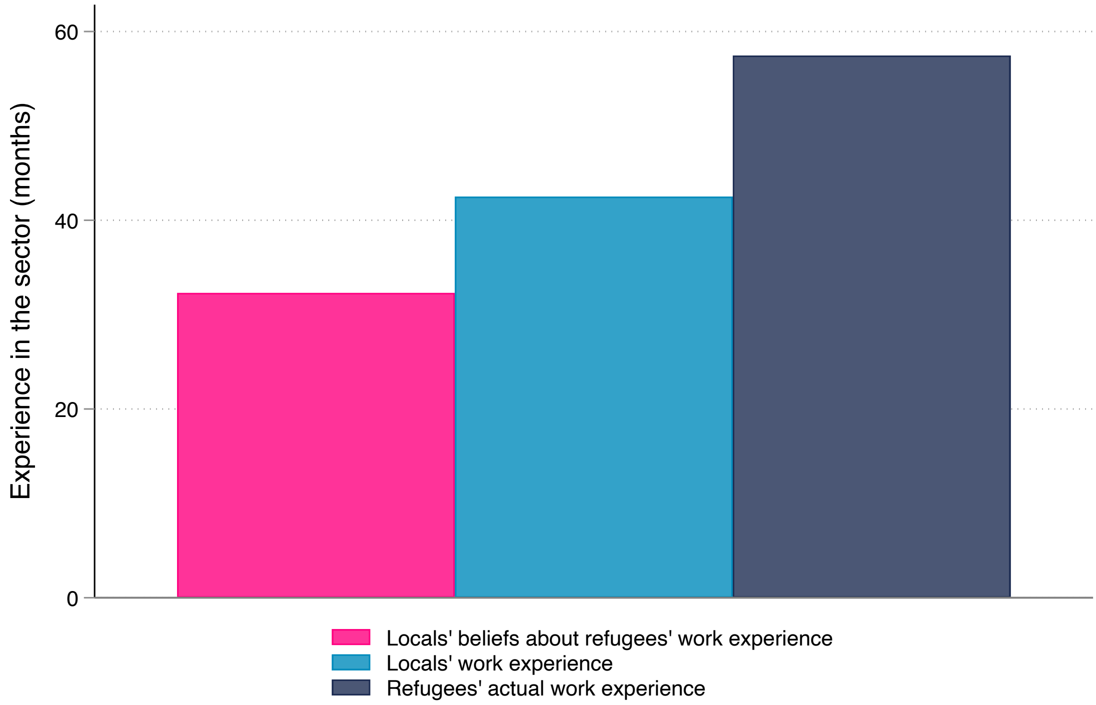
Notes: Timeline of data collection and project implementation, following the events from Chapter 2. Box on the bottom left of the picture details number of participants to the experiment.

Figure A.5: Correlations Between Explicit and Implicit Bias



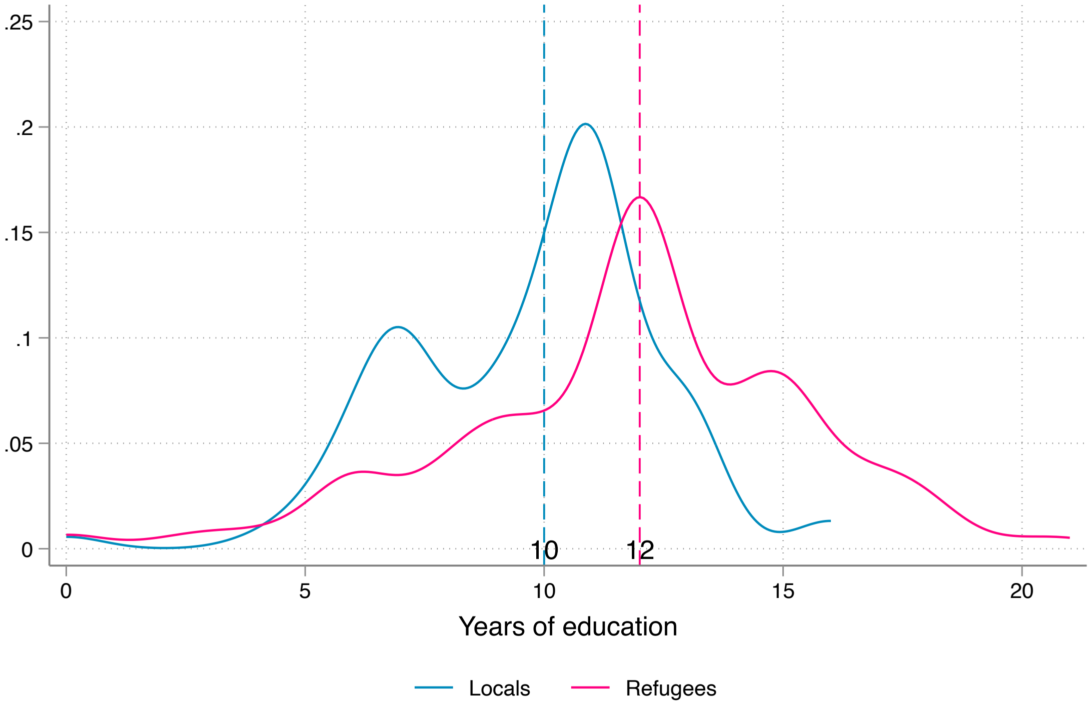
Notes: This graph plots the coefficients from the following regressions: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, where y_i is the combined explicit bias index constructed as described in Section 5.1.4 and each x_i are individual controls in different regressions (the average (combined) IAT score, the work-related IAT and the general stereotypes IAT). Standard errors are robust.

Figure A.6: Work-relate experience in the sector in months



Notes: This graph plots number of month of work experience. First bar to the left represents Ugandan workers' beliefs about refugee workers number of months of work-related experience. Second bar shows the actual number of months of work-related experience in the sample of local workers. Finally, third bar plots the actual number of work-related experience in the sample of the refugee workers.

Figure A.7: Years of education



Notes: This graph plots the distribution of years of education of local and refugee workers.

Figure A.8: Effect on locals' job views

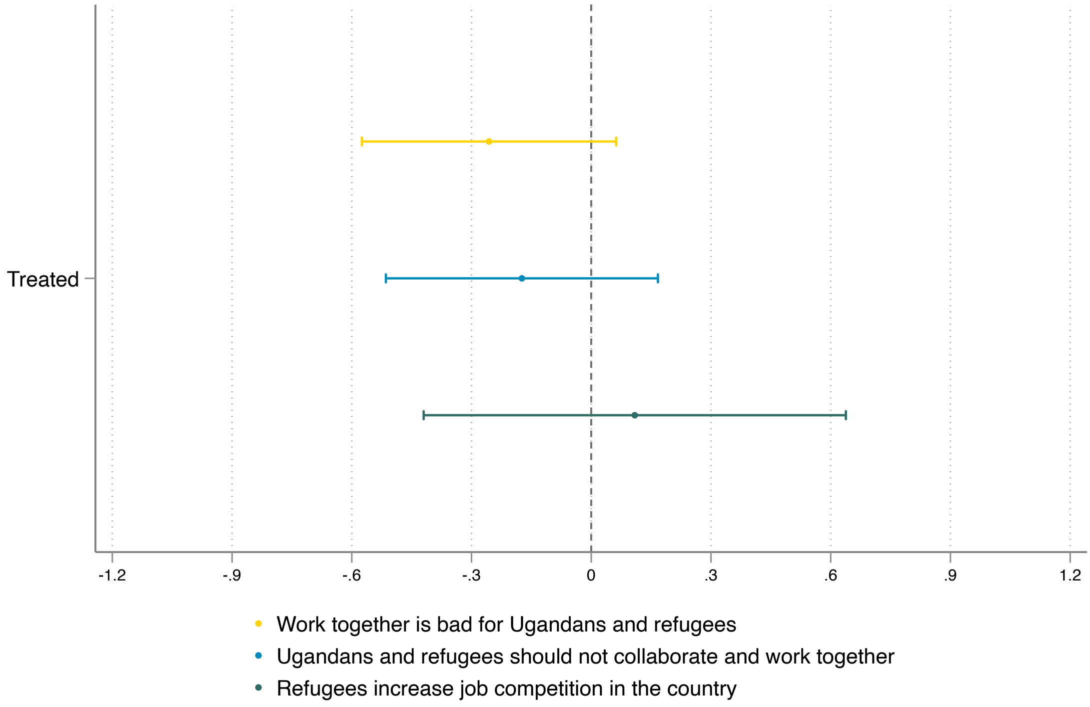


Table A.1: Correlations Between IAT, Statements and Demographics: Refugees

Variable	IAT Combined	Mean	N
Friendly	-0.973	4.549	377
	(0.433)**		
Unserious	0.416	4.218	377
	(0.394)		
Intermarriage is a good thing	-1.042	3.653	377
	(0.519)**		
Trust refugees - same nationality	0.729	3.899	377
	(0.432)*		
Trust refugees - other nationality	0.657	3.714	377
	(0.410)		
Trust Ugandans	-0.099	3.472	377
	(0.400)		
Similar culture	-1.797	3.777	377
	(0.521)***		
Altruism	6,494.922	17,074.271	377
	(3,046.029)**		
Reciprocity	2,070.151	7,750.663	377
	(1,128.958)*		
Refugee partner	-0.047	0.939	377
	(0.070)		
Ugandan partner	0.087	0.857	377
	(0.098)		
Business partners	-0.063	0.973	377
	(0.037)*		
Age	-3.178	33.724	377
	(2.635)		
English language index refugee	0.009	0.422	377
	(0.073)		
Luganda language index	0.016	0.376	377
	(0.077)		
Female	-0.025	0.655	377
	(0.074)		
Experience in months	-15.272	57.462	377
	(19.137)		
Years living in Uganda	-0.753	6.676	377
	(1.036)		

Notes: This table reports the coefficients from the following regression using the refugee subsample: $y_i = \beta_0 + \beta_1 IAT_i + \varepsilon_i$, where y_i is each variable underlining the full bias index (reported in each row) and IAT_i is the individual average IAT score between the general stereotypes-IAT and the work-related IAT. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively.

Table A.2: Correlations Between IAT, Statements and Demographics: Locals

Variable	IAT Combined	Mean	N
Friendly	0.814	4.927	273
	(0.594)		
Unserious	-1.270	2.729	273
	(0.599)**		
Intermarriage good thing	-0.139	4.901	273
	(0.705)		
Trust Ugandans - same ethnic	0.637	4.216	273
	(0.669)		
Trust Ugandans - other ethnic	0.565	4.044	273
	(0.577)		
Trust refugees	1.144	3.802	273
	(0.553)**		
Similar culture	0.269	3.656	273
	(0.782)		
Altruism	-15,318.500	21,681.318	273
	(5,830.996)***		
Reciprocity	-2,195.148	6,846.154	273
	(1,716.264)		
Refugee partner	-0.139	0.451	273
	(0.198)		
Ugandan partner	0.053	0.934	273
	(0.096)		
Business partners	0.055	0.938	273
	(0.096)		
Age	4.826	24.330	273
	(2.311)**		
English language	0.097	0.570	273
	(0.094)		
Luganda language	0.153	0.702	273
	(0.085)*		
Female	-0.035	0.634	273
	(0.132)		
Experience in months	14.057	42.505	273
	(13.932)		

Notes: This table reports the coefficients from the following regression using the native workers subsample: $y_i = \beta_0 + \beta_1 IAT_i + \varepsilon_i$, where y_i is each variable underlining the full bias index (reported in each row) and IAT_i is the individual average IAT score between the general stereotypes-IAT and the work-related IAT. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively.

Table A.3: Outcomes among locals, all treatments separately

	Bias		Business partner		
	(1)	(2)	(3)	(4)	(5)
	Implicit	Explicit	Out-group	Same group	Any partner
Video + exposure	0.076**	-0.284	0.118	0.053	0.054
	(0.038)	(0.225)	(0.104)	(0.055)	(0.055)
	[0.045]	[0.209]	[0.259]	[0.338]	[0.325]
Only exposure	0.068*	-0.242	0.244**	0.034	0.035
	(0.037)	(0.210)	(0.099)	(0.052)	(0.052)
	[0.069]	[0.249]	[0.014]	[0.517]	[0.507]
Only video	0.067*	-0.305	0.109	0.004	0.037
	(0.039)	(0.229)	(0.108)	(0.058)	(0.054)
	[0.091]	[0.184]	[0.313]	[0.951]	[0.496]
Observations	271	273	273	273	273
Mean DV	0.470	-0.000	0.405	0.919	0.919
H_0: Video=Exposure=Video+Exposure	0.930	0.933	0.152	0.518	0.849

Notes: Explicit bias index is contructed using the GLS Anderson weighting procedure combining negative attitudes and explicit negative stereotypes related to general attitudes. Implicit bias is an average of Work IAT score and General IAT score. Partners are dummies indicating if respondents want a partner or not. For local workers, we had an extra treatment: only video, due to a mistake in coding on the survey on the tablet. Control for refugees strata. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. Robust SE. p-values in square brackets.

Table A.4: Outcomes among refugees, all treatments separately

	Bias		Business partner		
	(1)	(2)	(3)	(4)	(5)
	Implicit	Explicit	Out-group	Same group	Any partner
Video + exposure	0.033	-0.381***	-0.087	-0.092*	-0.090**
	(0.030)	(0.147)	(0.059)	(0.050)	(0.040)
	[0.268]	[0.010]	[0.140]	[0.065]	[0.026]
Only exposure	0.011	-0.173	-0.070	-0.005	-0.042
	(0.024)	(0.124)	(0.047)	(0.035)	(0.027)
	[0.634]	[0.164]	[0.136]	[0.882]	[0.123]
Observations	333	377	377	377	377
Mean DV	0.409	-0.000	0.854	0.921	0.966
H_0: Exposure=Video+Exposure	0.396	0.093	0.756	0.056	0.225

Notes: Explicit bias index is contructed using the GLS Anderson weighting procedure combining negative attitudes and explicit negative stereotypes related to general attitudes. Implicit bias is an average of Work IAT score and General IAT score. Partners are dummies indicating if respondents want a partner or not. For local workers, we had an extra treatment: only video, due to a mistake in coding on the survey on the tablet. Control for refugees strata. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. Robust SE. p-values in brackets.

Table A.5: SMS sent by refugee workers, all treatments separately

	Sent SMS	SMS for Ugandan firm	SMS for refugee firm
	(1)	(2)	(3)
Video + exposure	0.072	0.036	0.036
	(0.056)	(0.034)	(0.047)
	[0.193]	[0.287]	[0.444]
Only exposure	0.120***	0.060**	0.059
	(0.046)	(0.028)	(0.040)
	[0.010]	[0.033]	[0.135]
Observations	377	377	377
Mean DV	0.124	0.034	0.090
H_0: Exposure=Video+Exposure	0.371	0.484	0.600

Notes: Sent SMS outcome is a dummy indicating if refugee workers sent a SMS to participate in similar future interventions. SMS for Ugandan and for refugee firm indicate what type of firm the worker would like to work in future interventions. Control for refugees strata. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. Robust SE. p-values in brackets.

Table A.6: The Effect of Contact on Implicit and Explicit Bias, Kling and Liebman Bounds

	Bias Index		Work bias		General bias	
	(1) Lower 0.25SD	(2) Upper 0.25SD	(3) Lower 0.25SD	(4) Upper 0.25SD	(5) Lower 0.25SD	(6) Upper 0.25SD
T × Local × Implicit	0.070**	0.070**	0.073**	0.073**	0.031	0.032
	(0.035)	(0.035)	(0.036)	(0.036)	(0.029)	(0.029)
	[0.044]	[0.043]	[0.045]	[0.044]	[0.280]	[0.274]
T × Local × Explicit	-0.052*	-0.052	-0.043	-0.042	-0.026	-0.025
	(0.032)	(0.032)	(0.034)	(0.034)	(0.025)	(0.025)
	[0.100]	[0.102]	[0.203]	[0.207]	[0.297]	[0.303]
T × Refugee × Implicit	0.015	0.015	0.004	0.004	0.017	0.017
	(0.023)	(0.023)	(0.023)	(0.023)	(0.028)	(0.028)
	[0.515]	[0.507]	[0.870]	[0.859]	[0.542]	[0.535]
T × Refugee × Explicit	-0.037*	-0.036*	-0.038*	-0.038*	-0.032*	-0.032*
	(0.019)	(0.019)	(0.022)	(0.022)	(0.019)	(0.019)
	[0.053]	[0.054]	[0.083]	[0.085]	[0.096]	[0.098]
Observations	1254	1254	1248	1248	1243	1243
Mean DV	0.481	0.481	0.443	0.443	0.466	0.466
Mean DV Local Implicit Bias	0.470	0.470	0.432	0.432	0.541	0.541
Mean DV Refugee Implicit Bias	0.409	0.409	0.465	0.465	0.470	0.470
Mean DV Local Explicit Bias	0.529	0.529	0.435	0.435	0.448	0.448
Mean DV Refugee Explicit Bias	0.528	0.528	0.433	0.433	0.439	0.439
H_0 : T × Local × Implicit=Refugee	0.186	0.185	0.110	0.110	0.716	0.715
H_0 : T × Local × Explicit=Refugee	0.674	0.676	0.900	0.902	0.837	0.835

Notes: This table reports results from specification 1, constructing bounds following Kling and Liebman (2004). Explicit bias index is constructed using the GLS Anderson weighting procedure combining negative attitudes and explicit negative stereotypes related to work. Implicit bias is the Work IAT. Both indices are normalized 0 to 1 for comparison. An increase means more prejudice. Control for refugees strata (refugees' occupations). Robust standard errors in parenthesis and p-values in brackets. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively.

Table A.7: The Effect of Contact on Implicit and Explicit Bias, Principal Component

	(1)
	Bias Index
T × Local × Explicit	-0.023
	(0.035)
	[0.516]
T × Local × Implicit	0.089**
	(0.035)
	[0.011]
T × Refugee × Explicit	-0.047**
	(0.024)
	[0.045]
T × Refugee × Implicit	0.021
	(0.024)
	[0.375]
Observations	1200
Mean DV	0.426
Mean DV Local Implicit Bias	0.460
Mean DV Refugee Implicit Bias	0.405
Mean DV Local Explicit Bias	0.407
Mean DV Refugee Explicit Bias	0.439
H_0 : T × Local × Implicit=Refugee	0.107
H_0 : T × Local × Explicit=Refugee	0.561

Notes: This table reports results from specification 1. Explicit bias index is constructed using Principal Component Analysis (PCA) combining negative attitudes and explicit negative stereotypes. Implicit bias is an average of Work IAT score and General IAT score. Both indices are normalized 0 to 1 for comparison. An increase means more prejudice. Control for refugees strata (refugees' occupations). Robust standard errors in parenthesis and p-values in brackets. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively.

Table A.8: Explicit bias among locals

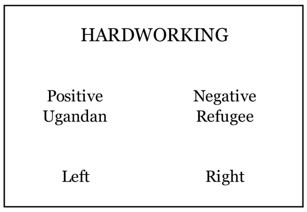
Variable	Treated	Mean	N
Bias Index (norm. 0-1)	-0.056 (0.034)	0.511	273
Refugees are hospitable	-0.041 (0.228)	5.161	273
Refugees are friendly	0.345 (0.237)	5.282	273
Refugees are peaceful	-0.065 (0.243)	5.275	273
Refugees are kind	0.236 (0.250)	5.168	273
Refugees are trustwhorty	0.202 (0.286)	4.835	273
Refugees are honest in business	0.226 (0.270)	5.029	273
Refugees are professional	-0.307 (0.227)	4.967	273
Refugees are hardworking	-0.139 (0.211)	5.326	273
Refugees are trouble makers	0.144 (0.231)	2.597	273
Refugees are dangerous	-0.151 (0.209)	2.531	273
Refugees are jealous	0.300 (0.280)	2.740	273
Refugees are dirty	-0.055 (0.274)	2.648	273
Refugees are thieves	-0.264 (0.266)	2.454	273
Refugees are lazy	-0.034 (0.280)	2.755	273
Refugees are corrupt	-0.047 (0.239)	2.505	273
Refugees are unserious	0.222 (0.212)	2.670	273
Intermarriage is a good thing	-0.076 (0.240)	5.183	273
I trust refugees	-0.005 (0.271)	3.908	273
Working together helps both groups	0.237 (0.203)	5.711	273
Refugees and locals should work more together	0.127 (0.210)	5.780	273
I would feel safe with refugees as neighbors	0.262 (0.263)	4.916	273
I see myself similar to a refugee	0.192 (0.345)	4.103	273
Refugees' culture is different from mine	-0.246 (0.218)	5.139	273
Refugees discriminate towards Ugandans	-0.511 (0.255)**	3.689	273
I often feel anxious around refugees	-0.177 (0.285)	3.158	273
People should marry from same nationality	0.375 (0.303)	3.656	273
Refugees increase job competition in the country	0.145 (0.324)	4.667	273
Law should allow refugees to work	0.013 (0.051)	0.912	273
Working with a refugee will reduce my productivity	0.210 (0.258)	2.593	273

Notes: In this table, we replicate specification 3 changing the way we construct the bias index and adding separately each variable used to construct the index. Explicit bias index is contructed using the GLS Anderson weighting procedure combining negative attitudes and explicit negative stereotypes related to general attitudes. For robustness check, we include to the index variables that were collected only for local workers because of the content of the question (i.e. were not appropriate for refugee workers). The variables are: "Working with a refugee will reduce my productivity", "Refugees increase job competition in the country" and "Think about the current law on refugees in Uganda. Do you think that the law should allow all refugees to work anywhere in Uganda?". Control for refugees strata. Robust standard errors in parenthesis and p-values in brackets. ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively.

B Appendix B: Outcomes survey questions and IAT

1. IAT and explicit stereotypes I am going to ask you how well each of the following words describes most Ugandans/refugees living in Kampala. Please answer using a scale between 1-7 where 1 means "It does not describe them at all" and 7 means "It describes them extremely well"

Figure A.9: IAT screen and stimuli list



IAT Screen

	General IAT	**Work IAT**		**General IAT**	**Work IAT**
Positive words	Hospitality Kindness Friendship Peaceful	Trust Hardworking Honest in business Professional	**National concepts**	Entebbe Jinja Domestic Ugandan	National Rolex Chapati Luganda Ugandan Cranes
Negative Words	Danger Jealous Trouble maker Dirty	Laziness Corruption Thief Unserious	**Refugee concepts**	Resettlement Non-native UNHCR Refugee Camp	Displaced person Foreign Migrant Urban refugee

2. Attitudes Now I will read series of statements about Ugandans and refugees in Kampala. Please indicate how much you agree with the statement. You can choose any number from 1 that means "I do NOT agree at all," and a 7 means "Agree totally".

   - I believe intermarriage between refugees and Ugandans is a good thing
   - I would advise my family and my refugee/Ugandan friends that they should only marry people from the same nationality
   - Ugandans/refugees' culture is different from my own culture
   - I see myself similar to a Ugandan/refugee
   - Ugandans/refugees discriminate towards refugees/Ugandans
   - I would feel safe having Ugandans/refugees as neighbours in the same compound
   - I assume that in general, Ugandans/refugees have only the best intentions
   - Work between Ugandans and refugees is good for both groups
   - I often feel anxious around Ugandans/refugees
   - Ugandans and refugees should collaborate and work more together

3. Hypothetical behaviors

   - Imagine you start a new business, and you can choose between different business partners that have a lot of experience in the sector. How many partners between 0 and 6 would you choose?
   - Of these, how many would be refugees?

- Of these, how many would be Ugandans?
- Now, suppose you are given 100,000UGX in cash and you have to decide how to spend it in the next 30 days. You can allocate it for: education expenses (for your or your family), food and consumption, health, leisure (including games, festivals, ceremonies, sport, clothing), savings and investments in a business. How would you allocate it? Education Food and home consumption Health Leisure Savings Investments in business

4. Real behavior

   - We would like to know your interest in future projects that might give you the possibility to be matched with Ugandan or refugee firms in Kampala. If you are interested, you can register by sending an SMS to the phone number we will give you. In the message, you need to include (1) your full name, (2) the ID number we will give you and (3) your preference between being matched to a Ugandan firm with Ugandan employees or a refugee firm with refugee employees (include only one preference). Please only register yourself, not other people! All firms are the same in terms of wages and hours worked.

C Appendix C: Covid Prevention Plan

Covid-19 prevention plan. Due to the COVID-19 pandemic during the period of our study, we followed several guidelines. These guidelines were based on recommendations from the World Health Organization, the Ugandan National Council for Science and Technology, Innovations for Poverty Action and IDinsight. We proceeded with writing a COVID-19 Risk Management Plan, and a COVID-19 guideline for training and data collection. In summary, we provided field officers hand sanitizers, face masks and instructed them to maintain a 1.5m of distance from the respondents. Moreover, each morning the team leader measured the temperature of field officers using an infrared thermometer and checked their health status. If a field officer had a fever of 37.5 or more, or showed signs of illness such as runny nose, cough or sneezes, he or she was sent home. Additionally, they had the right to interrupt the interview if the respondent refused to observe the SOPs. Since our study implied that people meet other people in person, we also followed guidelines to ensure that participants were safe. During the day we matched firms and workers, we instructed enumerators to measure participants' temperatures, to check their health status, and to check that they were following all the SOPs properly (using masks and sanitizers). If a respondent had a fever of 37.5 or more, or showed signs of illness such as runny nose, cough or sneezes, he or she was asked to go home.

D Appendix D: Ethical approvals

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). Moreover, this study is registered in the AEA RCT Registry under the unique identifying number: AEARCTR-0007238. All errors are our own.

2

Chapter 2: Matching with the Right Attitude: the Effect of Matching Firms with Refugee Workers

This paper has been jointly written with Francesco Loiacono

Chapter 2: Table of Contents

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

2.1.1 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 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 A.1, Panel A). Approximately 70% of the population of refugees residing in Kampala are of working age (aged 18-59).

[1]"As Rich Nations Close the Door on Refugees, Uganda Welcomes Them", *New York Times*, 2018.
[2]https://data.unhcr.org/en/country/uga, portal accessed in November 2023.

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

2.2.1 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.

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 A.1 in the Appendix shows which skill was tested for each occupation.

[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 A.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 A.4).

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 A.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 1).

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 A.2 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.

2.2.2 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 2 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 A.3, columns 1 and 3). For the second one, which took place approximately 8 months after the intervention, we collected longer term follow-up

[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.

data using phone calls from the 474 firms we managed to reached to. Table A.3 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 A.4 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 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 experiment consists of two parts. We use the first part to elicit firms' willingness to hire a refugee worker. In the second part of the experiment, we random assign firms to a treatment and a control group. 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

To elicit the employers' willingness to hire a refugee worker and to randomly allocate firms to treatment and control, we randomly pair refugees and employers, according to the occupation of the refugee worker (see Figure 3 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.

To elicit the employer's willingness-to-pay (WTP) for the paired refugee, we use 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 $100,000 UGX$.

In each firm-refugee pair, we begin by showing each employer the CV of a hypothetical Ugandan worker, to make sure that the employer understands the concept of WTP. 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 A.5). 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 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 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 for the randomly paired refugee worker and we do so twice, varying the level of information shared with the employer.[6] We elicit the first WTP right after showing a document with the profile of the candidate for a one-week internship. 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 A.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 A.6). 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.[7] 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.

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

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

[7]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 A.10.

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 A.8 for the original experimental design.

Approximately 45% of the 1,196 firms interviewed at baseline have a non-negative WTP to hire a refugee worker (see Figure 4). The remaining firms are either not interested in hiring any worker (approximately 35%) or interested in hiring a worker only if Ugandan (about 20%).

We use the second elicitation to allocate approximately 60% of the sample of firms to the treatment group. To do so, we extract a "random wage", W, from a sealed envelope. The random wage determines the outcome of the exercise. Specifically, if $WTP \geq W$, 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 A.9 shows the (inverted) demand function for a refugee worker in our sample.

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 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). 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 1). In section E we discuss the determinants of take-up among refugee workers.

Table 2 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.

[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.

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

In this section we briefly introduce our outcomes of interest. The goal of the experiment is to study whether exposure to one refugee changes firms' demand for refugees and the employers' beliefs regarding the ability of refugee workers.

Our initial hypothesis is that local employers have biased beliefs about the ability of refugee workers. At baseline, we measure employers' beliefs regarding the ability of refugee workers by asking what the employer thinks a refugee worker would score on the DIT practical skills test.[9] Additionally, we measure whether the employers'

[9] *"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?"*

beliefs vary if the worker were Ugandan.[10] 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.

Figure A.11 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.

Furthermore, 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]

Our main outcome of interest is the demand for refugee workers. We measure this using two outcomes. First, the number of refugees hired after the experiment. We measure this outcome using the last follow-up, conducted eight months after the intervention.[12] Second, during the short-term follow-up, collected approximately one month after the intervention. Specifically, we elicit the employers' WTP to hire a new, hypothetical, refugee worker, and we do so in a more nuanced, non-incentive compatible fashion.[13]

[10] *"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 get to see first the question about refugee job-seekers and then the one about Ugandans, and vice versa.

[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.

[12] 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?"

[13] 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. 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

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

$$y_{i1} = \beta_0 + \beta_1 T_i + y_{i0} + X_i^{'}\delta + \varepsilon_i, \tag{1}$$

where y_{i1} is one of our main outcomes of interest (the demand for new refugees and the beliefs regarding refugees' abilities). More specifically, in one of the measures for the demand for refugees, we use a dummy variable equal to 1 if the firm is willing to hire the new refugee worker at least for free, 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. 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 E. 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.[14]

4.1 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

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.

[14]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 A.8. Furthermore, we re-run specification 1 using two dummies instead of one, and report the results in Tables A.13 and A.14:

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

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).

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.[15]

[15]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 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.

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 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 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 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 A.5 repeat equation 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 A.6) and re-running a new specification using a post-double lasso procedure (Table A.7). 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 A.8). 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

$$Y = [I_n \otimes T]\beta + \mu \tag{2}$$

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.

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 A.12). 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 A.13). 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.[16] 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 A.14 shows that the demand does not shift differentially across the groups, with no difference between the full sample and the exposed one. Table A.9 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 A.10, 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)).

[16] 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.

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

$$\tau(X) = E[Y_{1i} - Y_{0i}|X = x] \tag{3}$$

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.[17] 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

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

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 A.11 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 A.15 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).

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:

$$y_{i1} = \beta_0 + \beta_1 TxPositive + \beta_2 TxMixed + \beta_3 TxNegative + X_i^{'}\delta + \varepsilon_i \qquad (4)$$

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.[18] 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 4. 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 Potscher (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 4 and report the results in Appendix Table A.12. 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 A.16 reports the p-values of β_1 and β_3, as well as the p-value from the test of equality between the 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

[18]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.

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 2 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 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 A.17 to A.23). 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 (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 A.24 reports the results of this survey. Panel 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 A.15 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 A.16 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 A.17 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 A.18 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.[19] 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 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

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

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., Caria, S., Fafchamps, M., Falco, P., Franklin, S., and Quinn, S. (2021). Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City. *The Review of Economic Studies*, 88(3):1279–1310.

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

Abel, M., Burger, R., and Piraino, P. (2020). The Value of Reference Letters: Experimental Evidence from South Africa. *American Economic Journal: Applied Economics*, 12(3):40–71.

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., and Vitali, A. (2020). Tackling Youth Unemployment: Evidence From a Labor Market Experiment in Uganda. *Econometrica*, 88(6):2369–2414.

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

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

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

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

Athey, S. and Wager, S. (2019). Estimating Treatment Effects with Causal Forests: An Application. *Observational Studies*, 5(2):37–51.

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

Barlow, F. K., Paolini, S., Pedersen, A., Hornsey, M. J., Radke, H. R. M., Harwood, J., Rubin, M., and Sibley, C. G. (2012). The Contact Caveat: Negative Contact Predicts Increased Prejudice More Than Positive Contact Predicts Reduced Prejudice. *Personality and Social Psychology Bulletin*, 38(12):1629–1643.

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

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

Bassi, V. and Nansamba, A. (2022). Screening and Signalling Non-Cognitive Skills: Experimental Evidence from Uganda. *The Economic Journal*, 132(642):471–511.

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

Becker, G. M., Degroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3):226–232.

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

Blair, C. W., Grossman, G., and Weinstein, J. M. (2022). Forced Displacement and Asylum Policy in the Developing World. *International Organization*, 76(2):337–378.

Broockman, D. and Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science*, 352(6282):220–224.

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

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

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

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

Carranza, E., Garlick, R., Orkin, K., and Rankin, N. (2022). Job Search and Hiring with Limited Information about Workseekersâ Skills. *American Economic Review*, 112(11):3547–3583.

Chassang, S., Padro i Miquel, G., and Snowberg, E. (2012). Selective trials: A principal agent approach to randomized controlled experiments. *American Economic Review*, 102(4):1279–1309.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68.

Corno, L., La Ferrara, E., and Burns, J. (2022). Interaction, Stereotypes, and Performance: Evidence from South Africa. *American Economic Review*, 112(12):3848–3875.

Davis, J. M. and Heller, S. B. (2017). Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs. *American Economic Review*, 107(5):546–550.

Davis, J. M. and Heller, S. B. (2020). Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs. *The Review of Economics and Statistics*, 102(4):664–677.

Dijker, A. J. M. (1987). Emotional reactions to ethnic minorities. *European Journal of Social Psychology*, 17(3):305–325.

Fasani, F., Frattini, T., and Minale, L. (2021). Lift the Ban? Initial Employment Restrictions and Refugee Labour Market Outcomes. *Journal of the European Economic Association*, 19(5):2803–2854.

Fasani, F., Frattini, T., and Minale, L. (2022). (The Struggle for) Refugee integration into the labour market: evidence from Europe. *Journal of Economic Geography*, 22(2):351–393.

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

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

Leeb, H. and Potscher, B. M. (2005). Model selection and inference: Facts and fiction. *Econometric Theory*, 21(01).

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

List, J. A., Shaikh, A. M., and Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22(4):773–793.

Lowe, M. (2021). Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration. *American Economic Review*, 111(6):1807–1844.

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

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(2):127–154.

McKeown, S. and Dixon, J. (2017). The âcontact hypothesisâ: Critical reflections and future directions: Critical reflections and future directions. *Social and Personality Psychology Compass*, 11(1):e12295.

Meleady, R. and Forder, L. (2019). When contact goes wrong: Negative intergroup contact promotes generalized outgroup avoidance. *Group Processes & Intergroup Relations*, 22(5):688–707.

Mousa, S. (2020). Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq. *Science*, 369(6505):866–870.

Nyqvist, M. B., Guariso, A., Svensson, J., and Yanagizawa-Drott, D. (2019). Reducing Child Mortality in the Last Mile: Experimental Evidence on Community Health Promoters in Uganda. *American Economic Journal: Applied Economics*, 11(3):155–192.

Pallais, A. (2014). Inefficient Hiring in Entry-Level Labor Markets. *American Economic Review*, 104(11):3565–3599.

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

Paolini, S., Harwood, J., and Rubin, M. (2010). Negative Intergroup Contact Makes Group Memberships Salient: Explaining Why Intergroup Conflict Endures. *Personality and Social Psychology Bulletin*, 36(12):1723–1738.

Rao, G. (2019). Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools. *American Economic Review*, 109(3):774–809.

Scacco, A. and Warren, S. S. (2018). Can Social Contact Reduce Prejudice and Discrimination? Evidence from a Field Experiment in Nigeria. *American Political Science Review*, 112(3):654–677.

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

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

Wager, S. and Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

8 Figures and Tables

Figure 1: Timeline

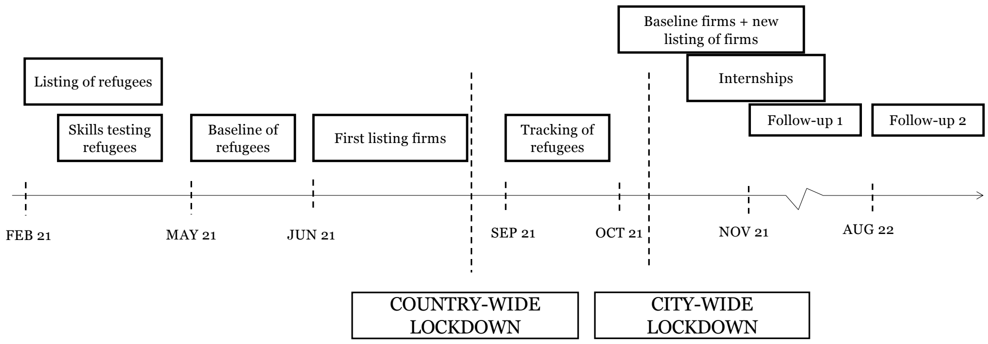
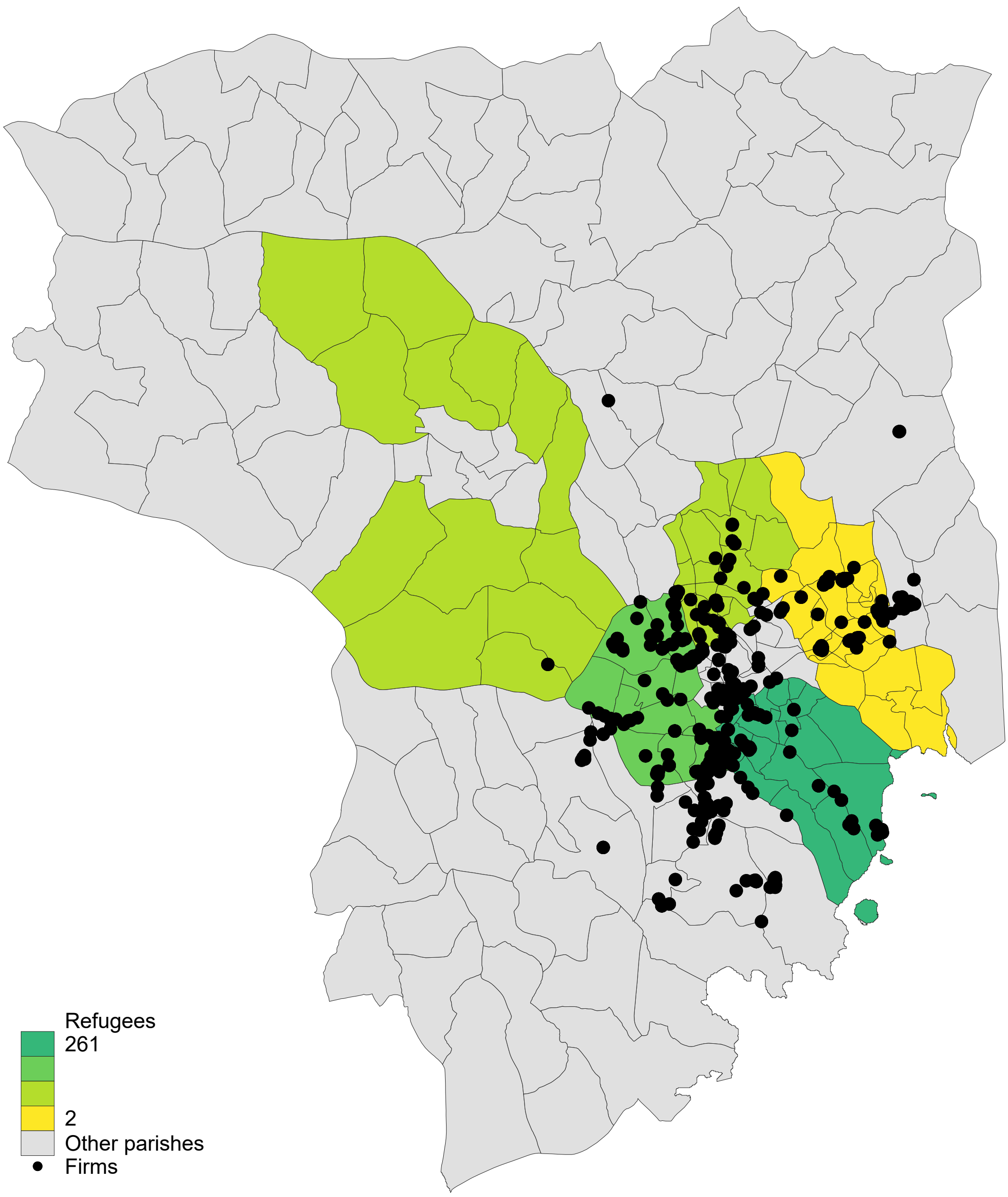


Figure 2: Map of firms' location



Notes: Each black dot is a firm who is willing to hire our refugee worker. We color parishes according to the number of residing refugees from our sample.

Figure 3: Design

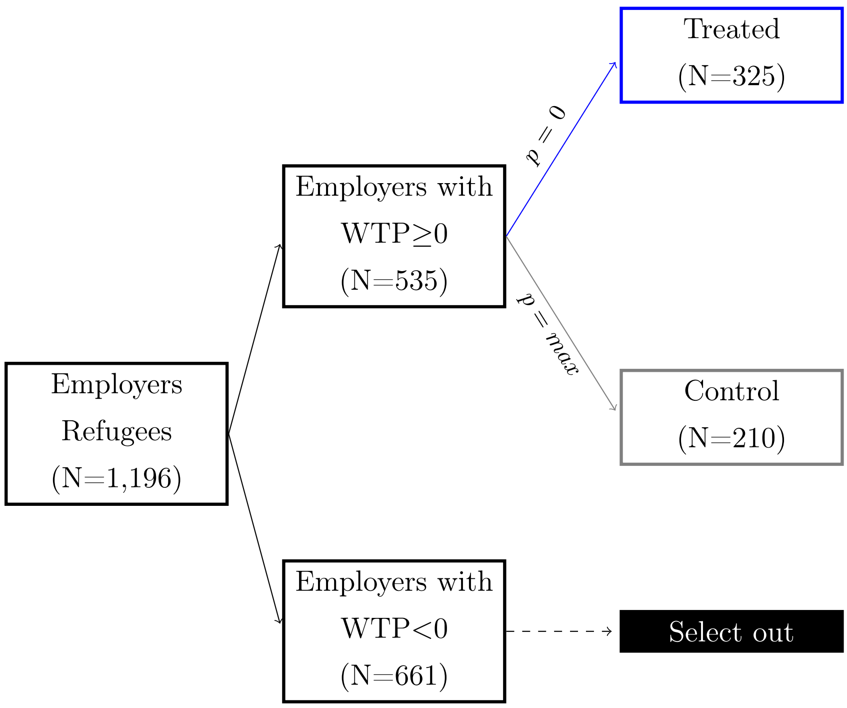
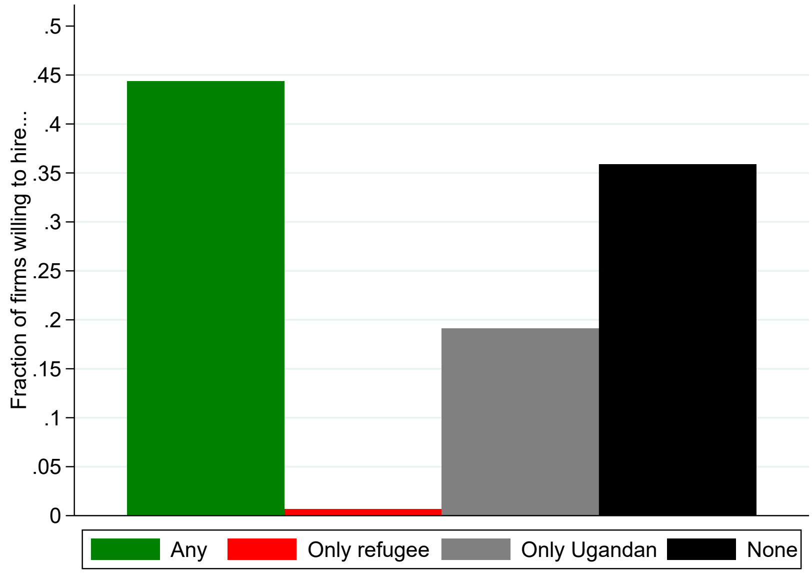


Figure 4: Firms' willingness to hire at baseline



Notes: The green bar represents the percentage of firms who are willing to hire (at least for free) both the hypothetical local worker and the real refugee worker; the red one the percentage of firms who are willing to hire only the real refugee worker; the gray one are firms who are interested only in the local worker; finally, the black bar is the percentage of firms who are not interested in neither of the two workers.

Table 1: Firms' take-up of the internships

Variable	Match	No match	Control	p(Matched=No)	N
Employer is a woman	0.582	0.538	0.581	0.074	535
	(0.495)	(0.500)	(0.495)		
Firm age	7.742	7.510	8.086	0.688	535
	(6.546)	(6.821)	(6.627)		
Revenues past month, M-UGX	1.569	2.047	2.036	0.118	535
	(2.004)	(3.345)	(2.651)		
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)		
Num. of rooms in business premises	1.159	1.182	1.176	0.375	535
	(0.788)	(0.738)	(0.876)		
Number of firms' tasks	3.308	3.350	3.476	0.893	535
	(1.484)	(1.637)	(1.599)		
Manufacturing sector	0.346	0.343	0.314	0.006	535
	(0.477)	(0.476)	(0.465)		
Ever offered internships	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)		
Supports refugees' empl. rights	0.934	0.909	0.924	0.275	535
	(0.249)	(0.288)	(0.266)		
Jobs to locals first	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)		

Notes: 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.

Table 2: Randomization balance

	Treatment			Control			
Variable	N	Mean	SD	N	Mean	SD	Diff.
Panel A: Full sample							
Employer is a woman	325	0.563	0.497	210	0.581	0.495	-0.063**
Firm age	325	7.640	6.659	210	8.086	6.627	-0.321
Revenues past month, M-UGX	325	1.780	2.684	210	2.036	2.651	-0.035
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
Num. of rooms in business premises	325	1.169	0.765	210	1.176	0.876	0.024
Number of firms' tasks	325	3.326	1.551	210	3.476	1.599	-0.073
Manufacturing sector	325	0.345	0.476	210	0.314	0.465	-0.020*
Ever offered internships	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
Beliefs about refugees' test score	325	65.052	14.501	210	62.705	16.013	2.126
Supports refugees' empl. rights	325	0.923	0.267	210	0.924	0.266	0.006
Jobs to locals first	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 is a woman	182	0.582	0.495	210	0.581	0.495	-0.040
Firm age	182	7.742	6.546	210	8.086	6.627	-0.347
Revenues past month, M-UGX	182	1.569	2.004	210	2.036	2.651	-0.234
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
Num. of rooms in business premises	182	1.159	0.788	210	1.176	0.876	0.006
Number of firms' tasks	182	3.308	1.484	210	3.476	1.599	-0.025
Manufacturing sector	182	0.346	0.477	210	0.314	0.465	-0.039**
Ever offered internships	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
Beliefs about refugees' test score	182	64.390	14.241	210	62.705	16.013	1.455
Supports refugees' empl. rights	182	0.934	0.249	210	0.924	0.266	0.019
Jobs to locals first	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

Notes: Balance test of characteristics between treated and control firms. The exposed sample are firms whose treatment took place.

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

Notes: 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.

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

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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 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	

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

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

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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.

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

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

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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.

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

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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.

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

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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.

Table 11: Learning, single components

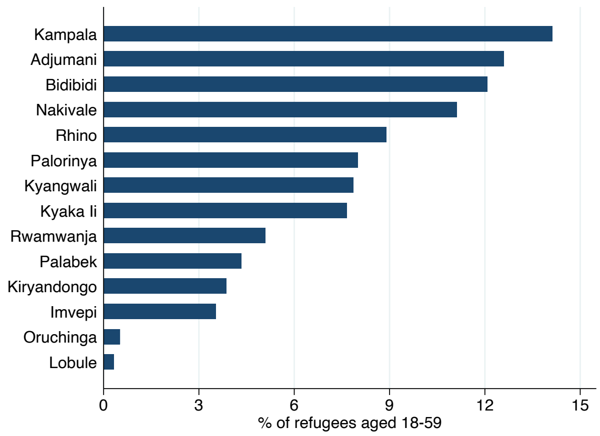
	Hard skills			Soft skills			Trust			Respect		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
β_1	0.403**	0.410**	0.435**	0.433**	0.419**	0.437**	0.668***	0.680***	0.665***	0.352**	0.375**	0.368**
	(0.178)	(0.181)	(0.174)	(0.179)	(0.183)	(0.177)	(0.165)	(0.168)	(0.164)	(0.164)	(0.167)	(0.161)
	[0.024]	[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]
$\beta_1 = \beta_2$	0.045	0.039	0.034	0.361	0.339	0.386	0.018	0.017	0.020	0.196	0.192	0.246
$\beta_1 = \beta_3$	0.033	0.022	0.022	0.300	0.283	0.314	0.053	0.045	0.060	0.340	0.331	0.377
$\beta_2 = \beta_3$	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

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

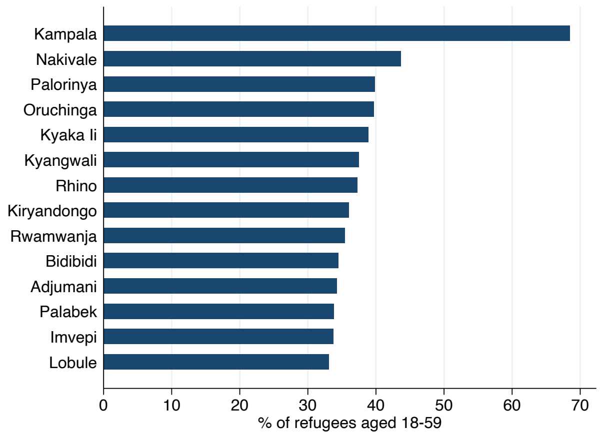
A Appendix A: Figures and extra analyses

Figure A.1: Refugees in Uganda

(a)

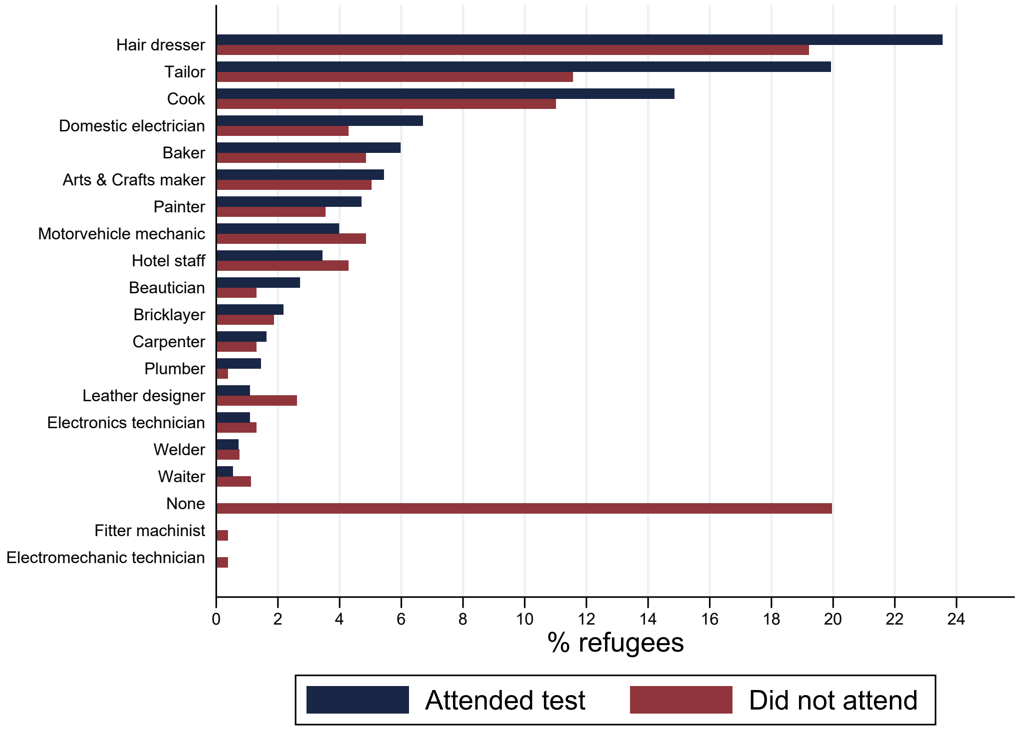


(b)



Notes: We use the latest available data from the office of the UNHCR in Uganda. Panel (A) shows the distribution of working-age refugees across each registered place of residence of refugees. Panel (B) reports the percentage of working-age refugees within each settlement.

Figure A.2: Refugees' skills, by attendance to the test

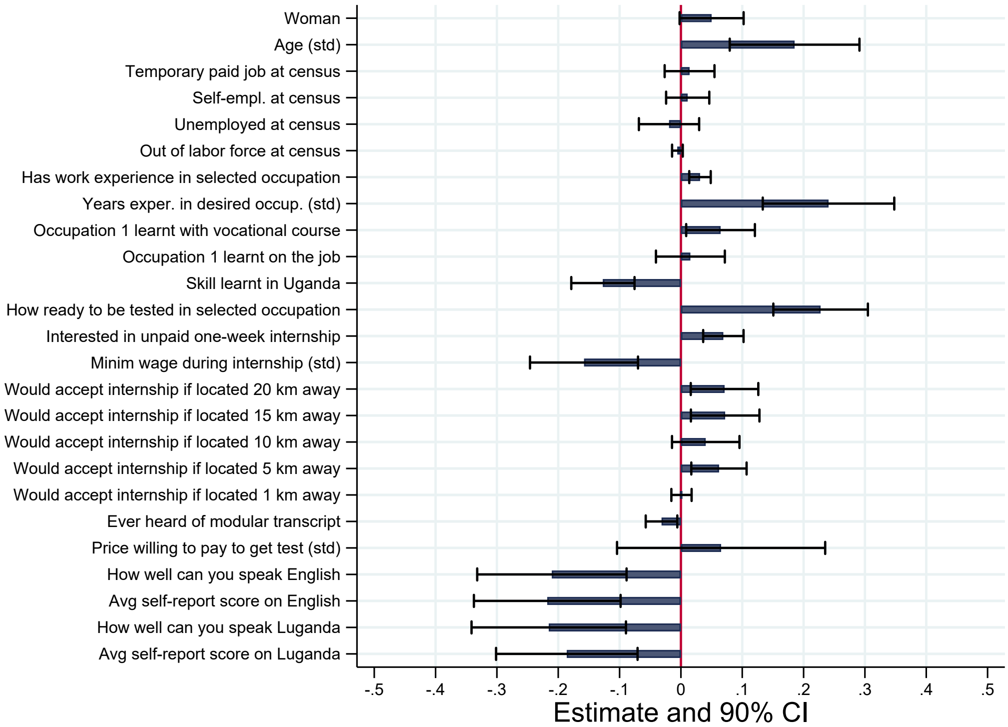


Notes: We invited 977 out of the 1,088 refugees we listed. Refugees that showed at the test were 552 (dark blue bars). Those who did not were 425 (red bars). Among those who did not show up, 111 were not invited to the test, either because they said they did not have any skill to be tested, or because the occupation group did not reach 5 components (as requested by the school that administered the test). These were refugees that declared to be skilled as fitter machinists or electromechanic technicians.

Figure A.3: DIT testing

Notes: Example of testing day with a group of refugee tailors. An official examiner controls quality of work (e.g. a short-sleeved shirt).© Mariajose Silva-Vargas

Figure A.4: Refugees who attended the test vs those who did not



Notes: Each dot is a coefficient from a single regression comparing the characteristics of refugees who showed up at the internship with those who did not.

Figure A.5: CVs Ugandan workers

(a)

John Sabiti

Tel: 0772 608515

Resident: Kampala, Nsambya, since: 2015

Age: 34

Expertise: cook

Years of experience as a cook: 8

Gender: Male

Nationality: Ugandan

Knowledge of English (self-reported scale 1-5):

Reading:	*Speaking:*	*Writing:*	*Listening:*
3=Moderately well	2=Not well	3=Moderately well	3=Moderately well

Knowledge of Luganda (self-reported scale 1-5):

Reading:	*Speaking:*	*Writing:*	*Listening:*
4=Well	4=Well	4=Well	4=Well

(b)

Dorcas Mandela

Tel: 0772 608515

Resident: Kampala, Masajja, since: 2016

Age: 36

Expertise: cook

Years of experience as a cook: 10

Gender: Female

Nationality: Ugandan

Knowledge of English (self-reported scale 1-5):

Reading:	*Speaking:*	*Writing:*	*Listening:*
3=Moderately well	2=Not well	2=Not well	3=Moderately well

Knowledge of Luganda (self-reported scale 1-5):

Reading:	*Speaking:*	*Writing:*	*Listening:*
3=Moderately well	3=Moderately well	1=Not at all	3=Moderately well

Notes: The CVs of the hypothetical local workers are filled with the same information found in the CVs of the refugee workers. The names in the CVs are chosen not to flag any particular Ugandan ethnicity. The two pictures are chosen not to indicate any tribal affiliation (and were modified from real images). The pictures, the (fake) names and the (non-existing) phone numbers do not vary across employers.

Figure A.6: Reasons why firms are not interested in the matched refugee

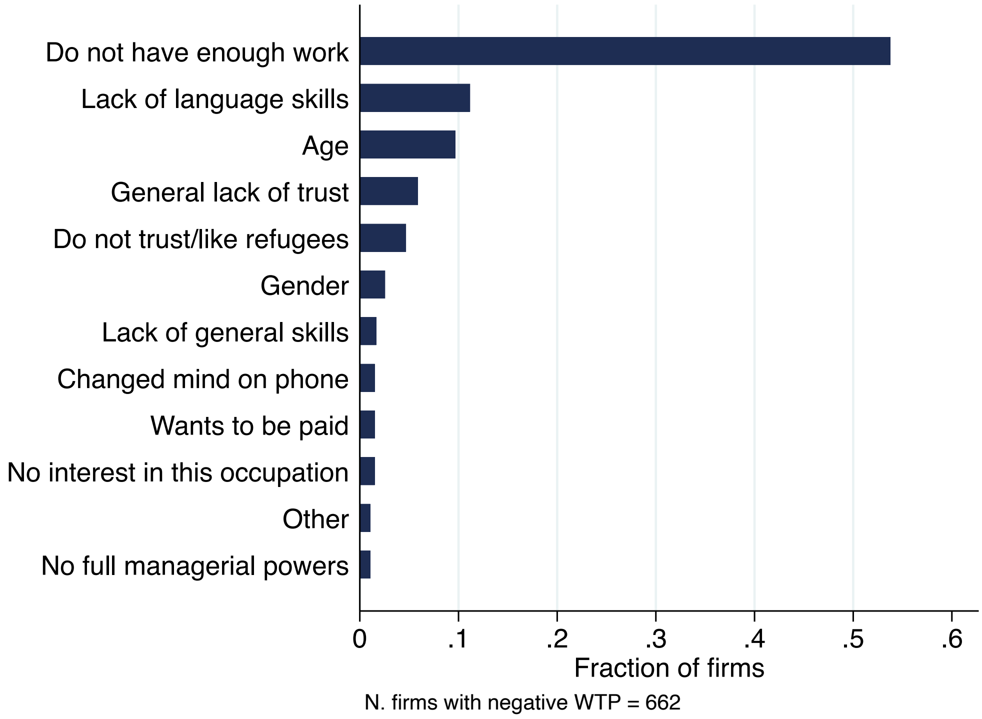
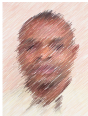


Figure A.7: CVs refugee worker

(a)

Wisdom Karungu



Tel: 0772 608515

Resident: Kampala, Nsambya, since: 2015

Age: 34

Expertise: cook

Years of experience as a cook: 8

Gender: Male

Nationality: Congolese

Knowledge of English (self-reported scale 1-5):

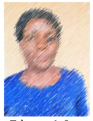
Reading: 3=Moderately well	*Speaking:* 2=Not well	*Writing:* 2=Not well	*Listening:* 3=Moderately well

Knowledge of Luganda (self-reported scale 1-5):

Reading: 3=Moderately well	*Speaking:* 3=Moderately well	*Writing:* 1=Not at all	*Listening:* 3=Moderately well

(b)

Noella Kabale



Tel: 0772 608515

Resident: Kampala, Masajja, since: 2016

Age: 36

Expertise: cook

Years of experience as a cook: 10

Gender: Female

Nationality: Congolese

Knowledge of English (self-reported scale 1-5):

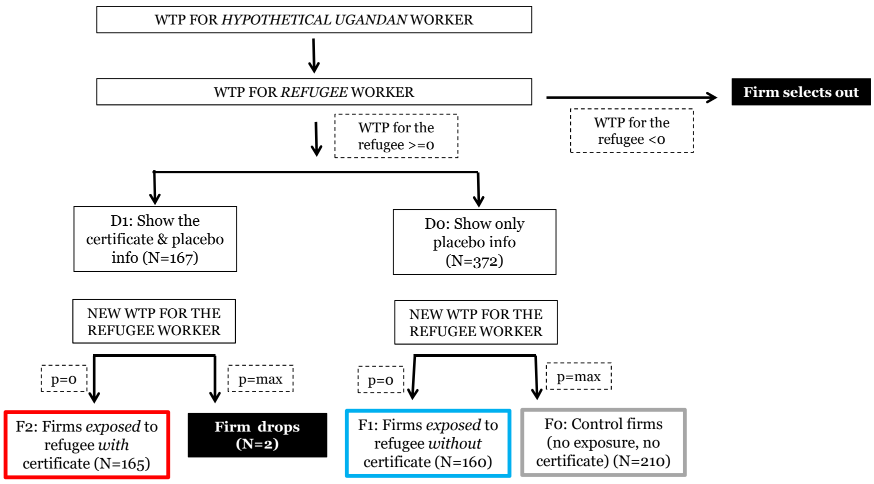
Reading: 3=Moderately well	*Speaking:* 2=Not well	*Writing:* 2=Not well	*Listening:* 3=Moderately well

Knowledge of Luganda (self-reported scale 1-5):

Reading: 3=Moderately well	*Speaking:* 3=Moderately well	*Writing:* 1=Not at all	*Listening:* 3=Moderately well

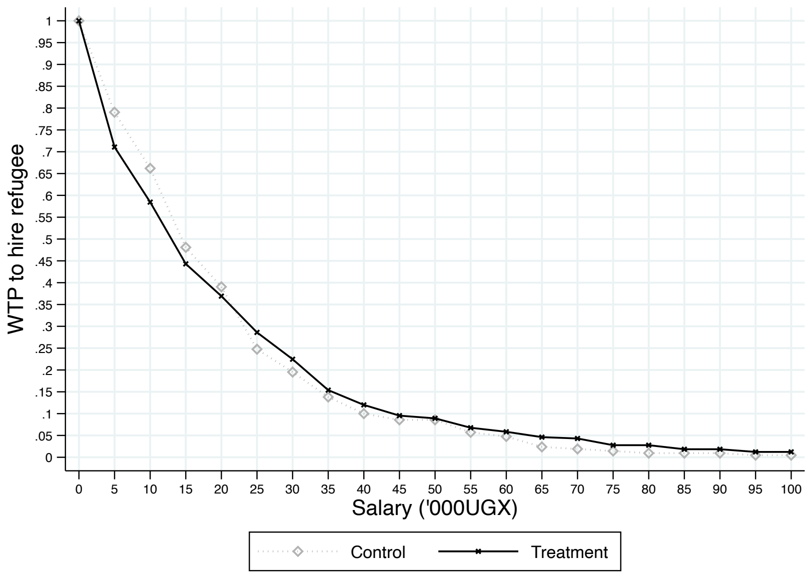
Notes: The two CVs contain information on the real refugee workers that are randomly pair to each employer (the names and phone numbers shown in this figure are not real). The CVs are filled with the information that the refugee respondents shared at baseline and for which we have consent to share with the employers.

Figure A.8: Original design



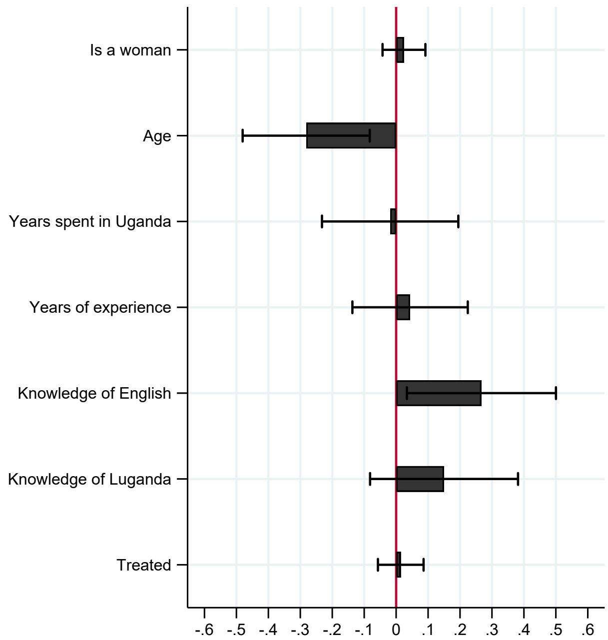
Notes: In the original design we show a subgroup of employers the certificate obtained by the matched refugee worker. We drop two employers belonging to the D1 arm to guarantee the incentive compatibility of the BDM mechanism (that is, to guarantee that the likelihood of "winning" the lottery of the random price is strictly lower than 1). The WTP is elicited twice. In the first elicitation we let the employer know that the hiring would happen in 4 days time. In the second elicitation we provide a weakly desireable increase in the terms of the hiring, letting the employer know that the hiring would happen in 8 days from the baseline.

Figure A.9: WTP curves at baseline



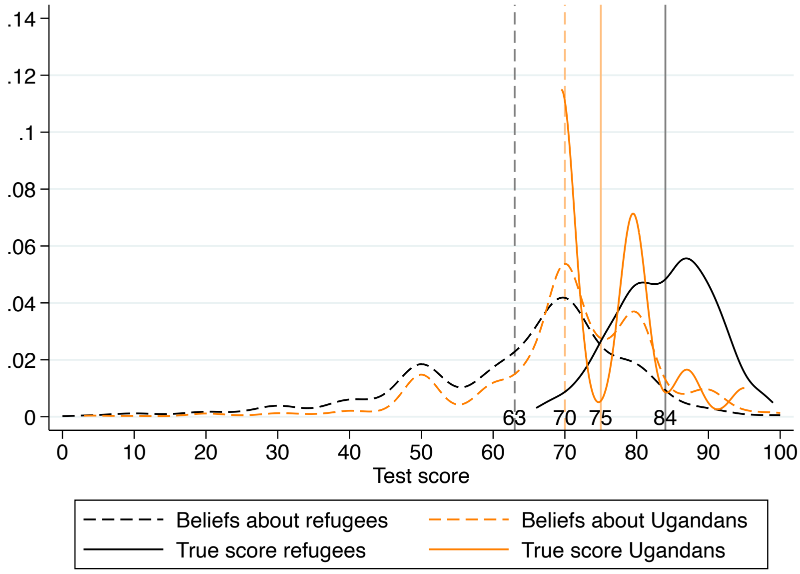
Notes: The figure plots the CDF of the WTP to hire a refugee worker at baseline. The gray line is the demand among control firms. The black one is the demand among treated who are assigned to treatment. The dark blue line excludes firms for which the internship took place.

Figure A.10: Refugees' matching success rate



Notes: $N = 527$ refugees. Coefficients from a linear regression on refugee's characteristic x on "average success rate", where this rate is computed as the average number of firms whose WTP is non-negative. Additional controls: occupation fixed effects. Standard errors clustered at the refugee level.

Figure A.11: Firms' beliefs about refugees' ability



Notes: Full baseline sample with 1,204 firms. Dashed lines represent the employers' beliefs (i.e. slef-reported score they think the job-seeker obtained). Solid lines represent the true scores. Black lines refer to the refugee workers, orange ones refer to Ugandans. Notice that Ugandans' scores may not be fully comparable to the ones of refugees, as the sample we use here to capture "real" scores is composed by typically younger and less experienced students.

Figure A.12: CVs new (hypothetical) refugee worker

(a)

Jamii Ndoli

Tel: 0773882694

Resident: Kampala, Makindye , since: 2020

Age: 26

Expertise: cook

Years of experience as a cook: 4

Gender: Man

Nationality: Congolese

Knowledge of English (self-reported scale 1-5):

Reading:	*Speaking:*	*Writing:*	*Listening:*
4=Well	4=Well	4=Well	4=Well

Knowledge of Luganda (self-reported scale 1-5):

Reading:	*Speaking:*	*Writing:*	*Listening:*
4=Well	4=Well	4=Well	4=Well

(b)

Christelle Bahati

Tel: 0773882694

Resident: Kampala, Makindye , since: 2020

Age: 26

Expertise: cook

Years of experience as a cook: 4

Gender: Woman

Nationality: Congolese

Knowledge of English (self-reported scale 1-5):

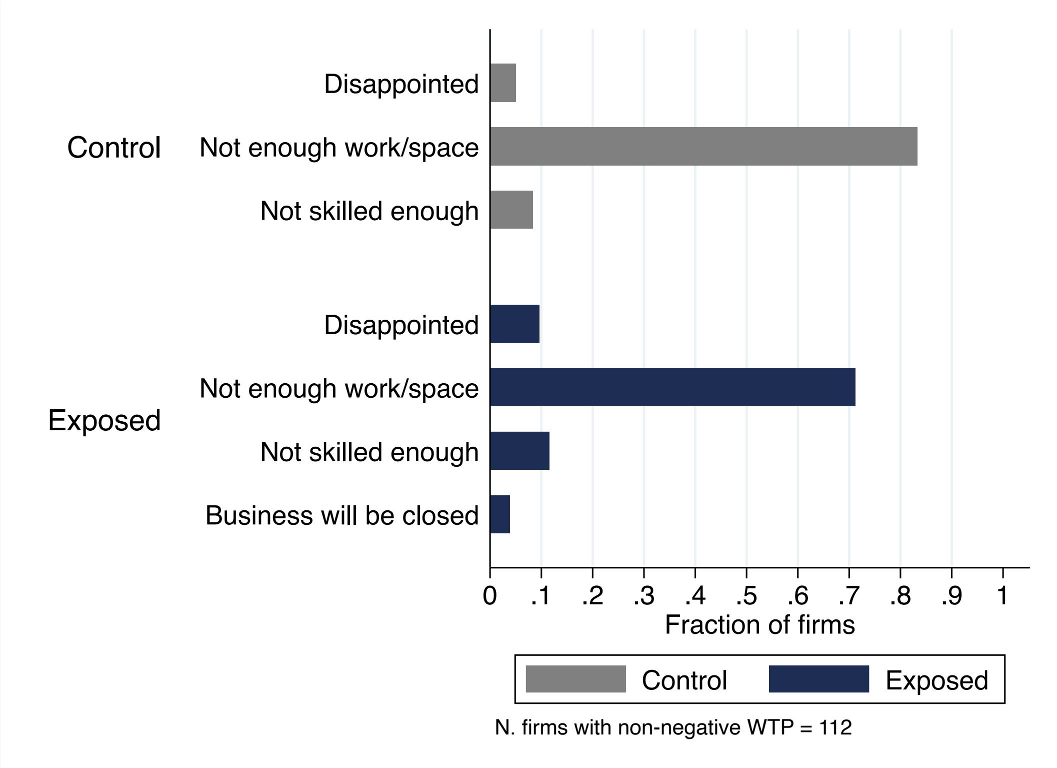
Reading:	*Speaking:*	*Writing:*	*Listening:*
4=Well	4=Well	4=Well	4=Well

Knowledge of Luganda (self-reported scale 1-5):

Reading:	*Speaking:*	*Writing:*	*Listening:*
4=Well	4=Well	4=Well	4=Well

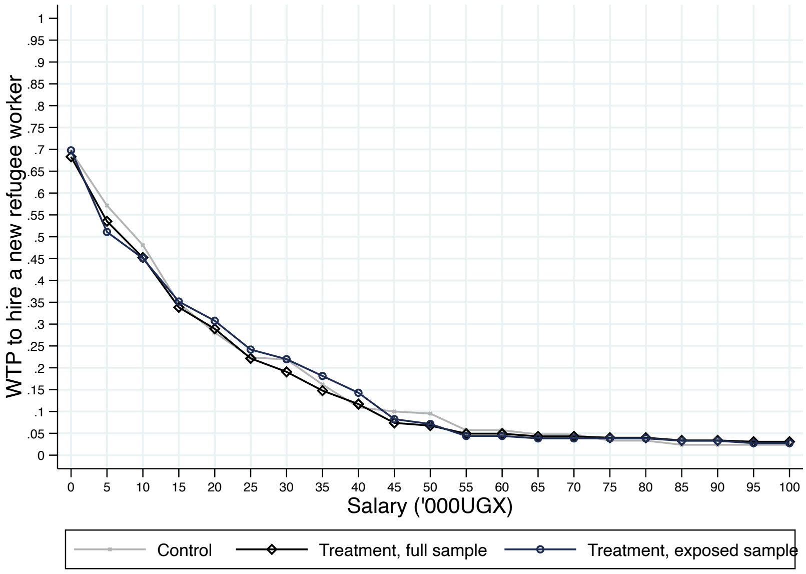
Notes: The two hypothetical CVs are constructed using desireable characteristics and used the pictures and the names of two real workers.

Figure A.13: Reasons for not being willing to hire a new refugee



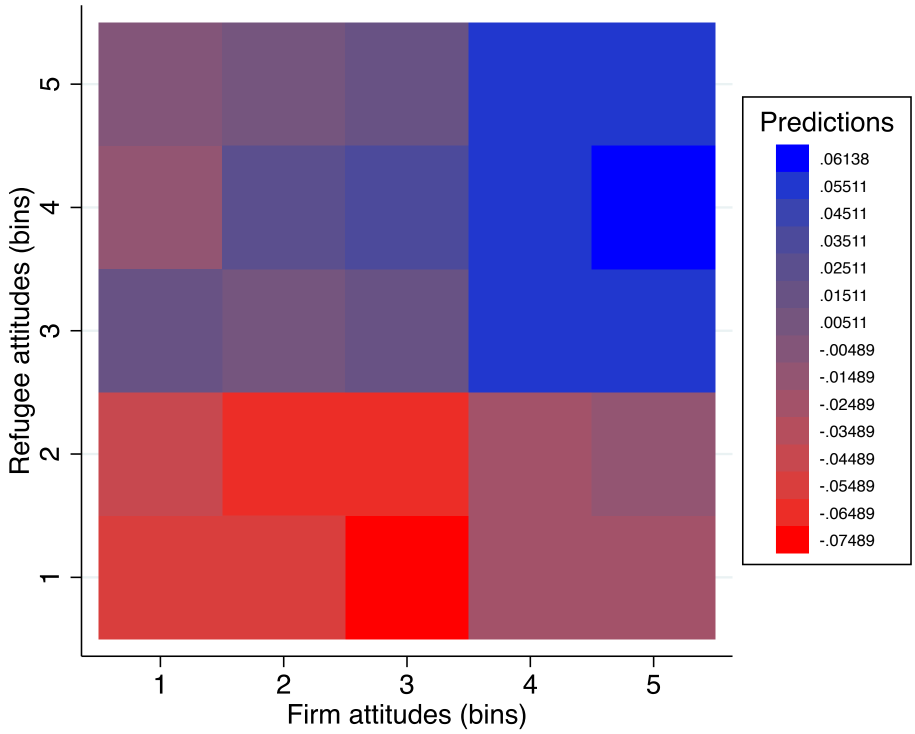
Notes: The graph reports the fraction of firms not willing to hire the new hypothetical refugee, by treatment status. The total number of firms reached at follow-up 1 is 385.

Figure A.14: WTP curves at follow-up 1



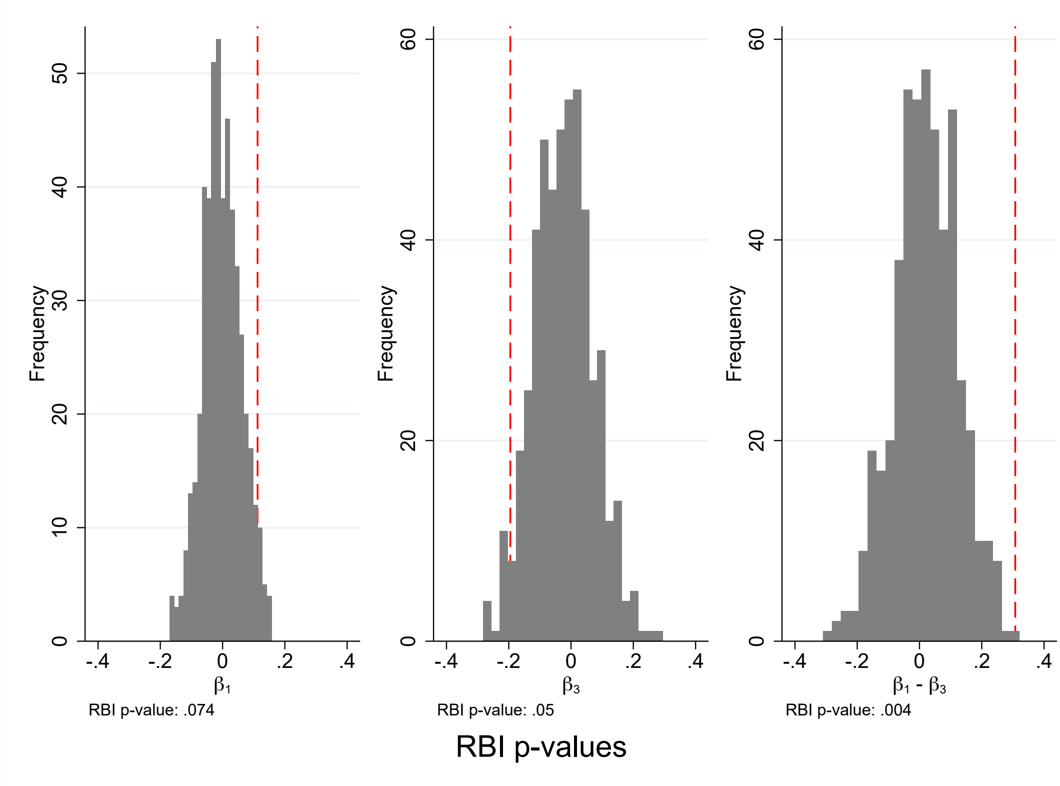
Notes: The figure plots the CDF of the WTP to hire a refugee worker at follow-up 1. The gray line is the demand among control firms. The black curve with diamonds corresponds to the demand of firms assigned to treatment. The dark blue line with circles excludes firms for which the internship did not take place.

Figure A.15: Predicted CATE and attitudes



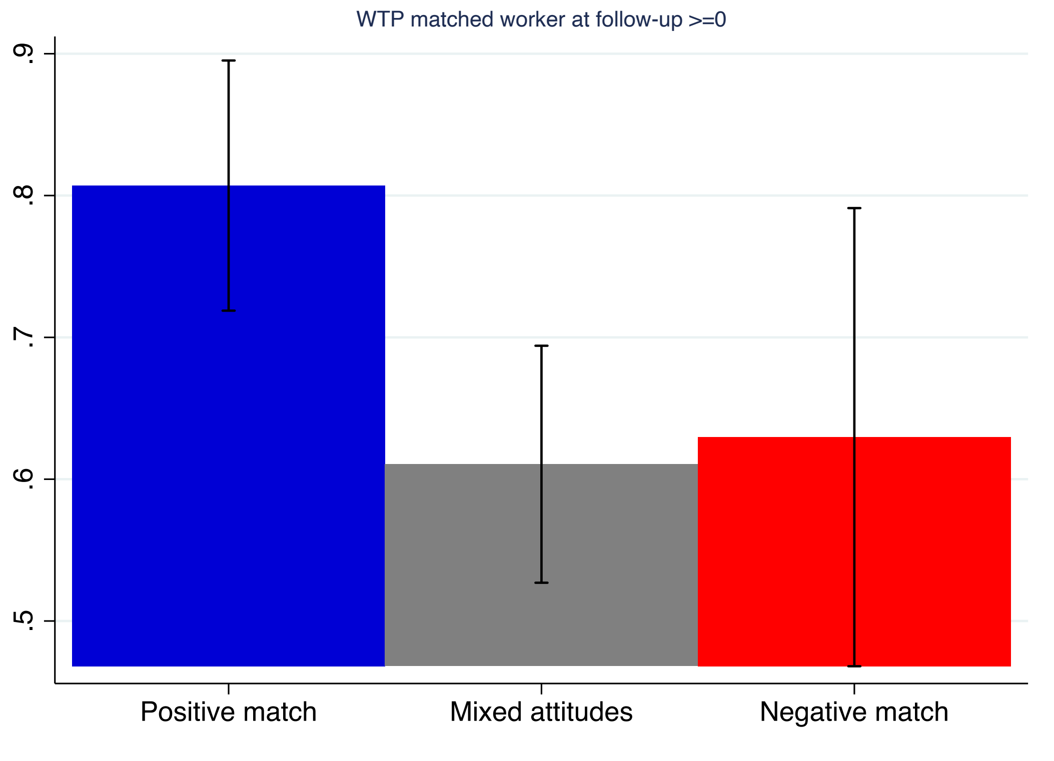
Notes: The heatmap plots predicted CATE across quartiles of the index of attitudes of both the employer (X-axis) and the refugee worker (Y-axis). The colder the color (i.e. the closer to blue), the more positive the effect on WTP to hire a new refugee worker. Viceversa, the warmer the color (i.e. the closer to red) the lower the predicted effect on WTP.

Figure A.16: Randomization-based inference



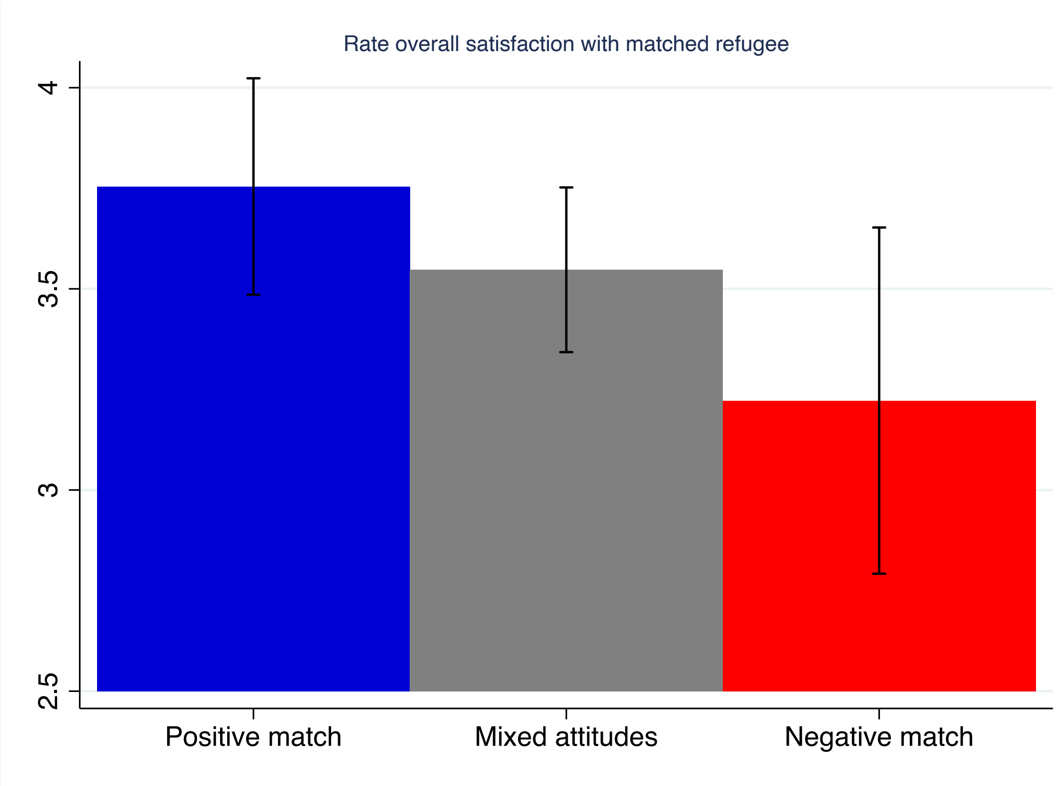
Notes: The first graph from the left reports the distribution of the values of β_1, computed using 500 simulations. The middle graph reports the distribution of β_3. Finally, the last graph is the distribution of the t-test of equality between β_1 and β_3. Each RBI p-value is reported below.

Figure A.17: Internship data: WTP to hire the same worker



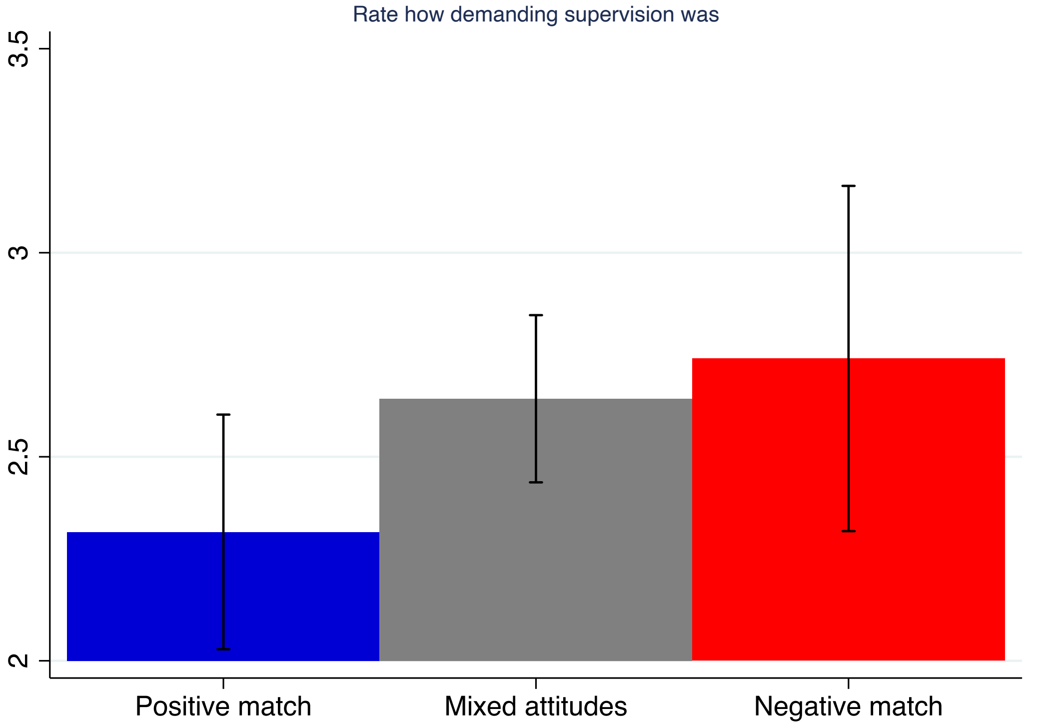
Notes: Group means and 95% confidence intervals around the mean. Answers reported by the employers who successfully matched with 1 refugee worker.

Figure A.18: Internship data: rate overall experience



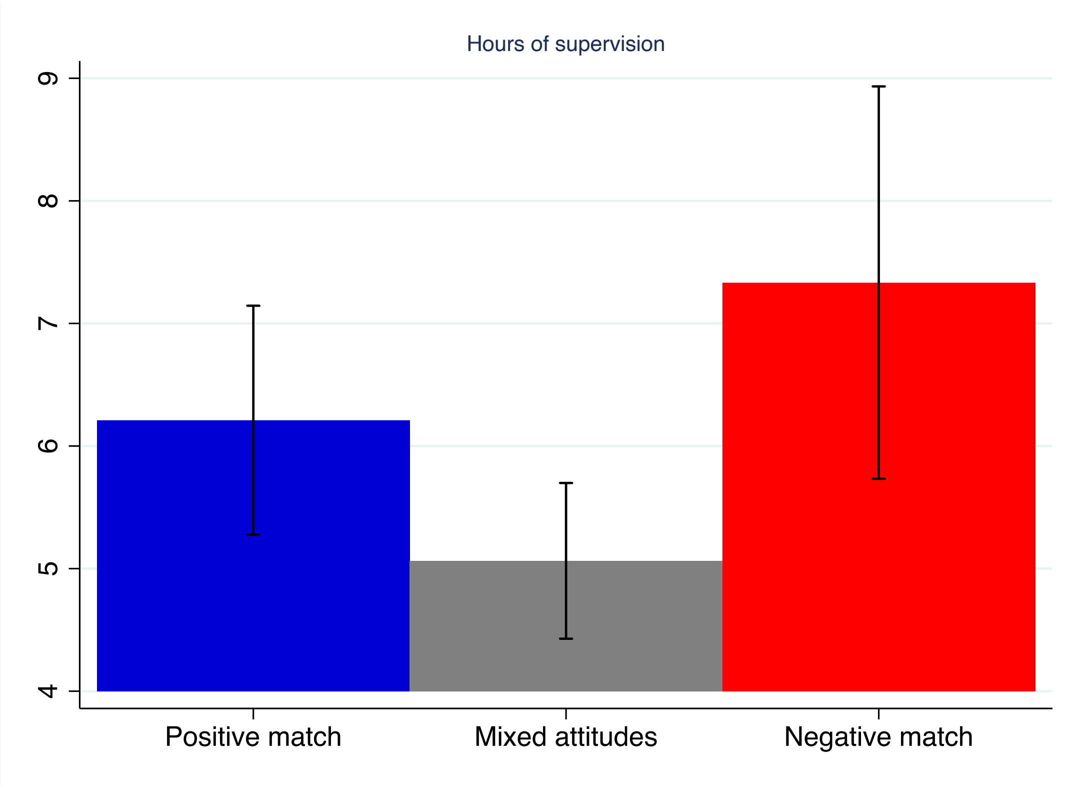
Notes: Group means and 95% confidence intervals around the mean. Answers reported by the employers who successfully matched with 1 refugee worker.

Figure A.19: Internship data: rate how demanding was supervision



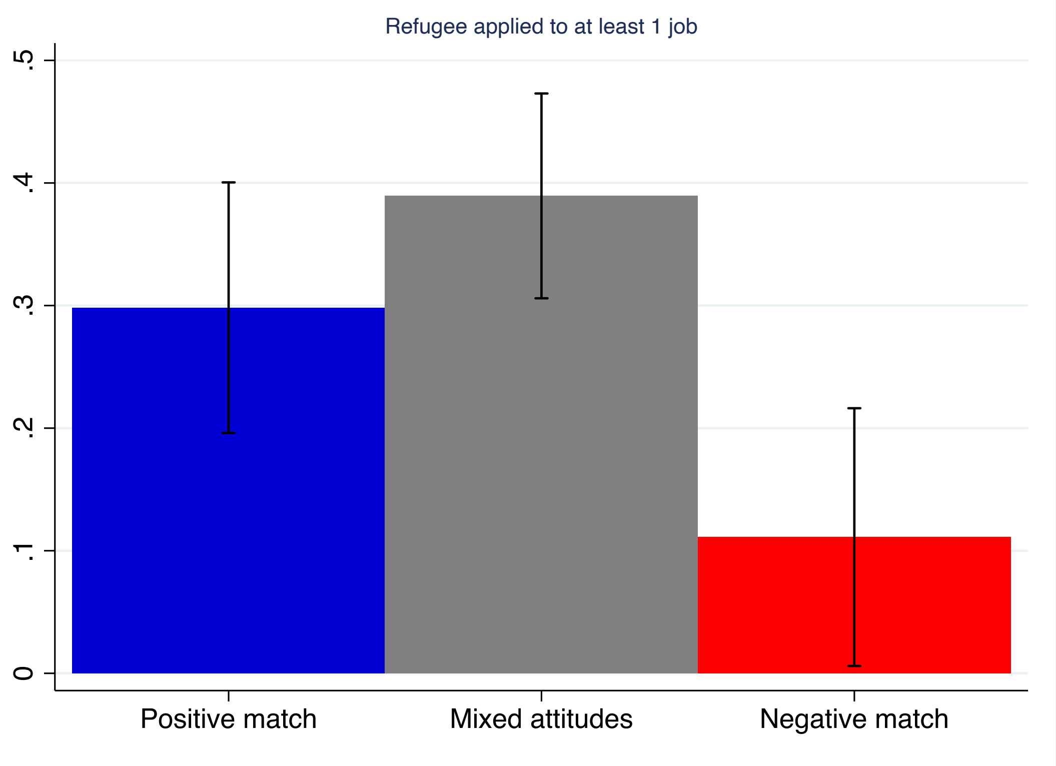
Notes: Group means and 95% confidence intervals around the mean. Answers reported by the employers who successfully matched with 1 refugee worker.

Figure A.20: Internship data: hours of supervision



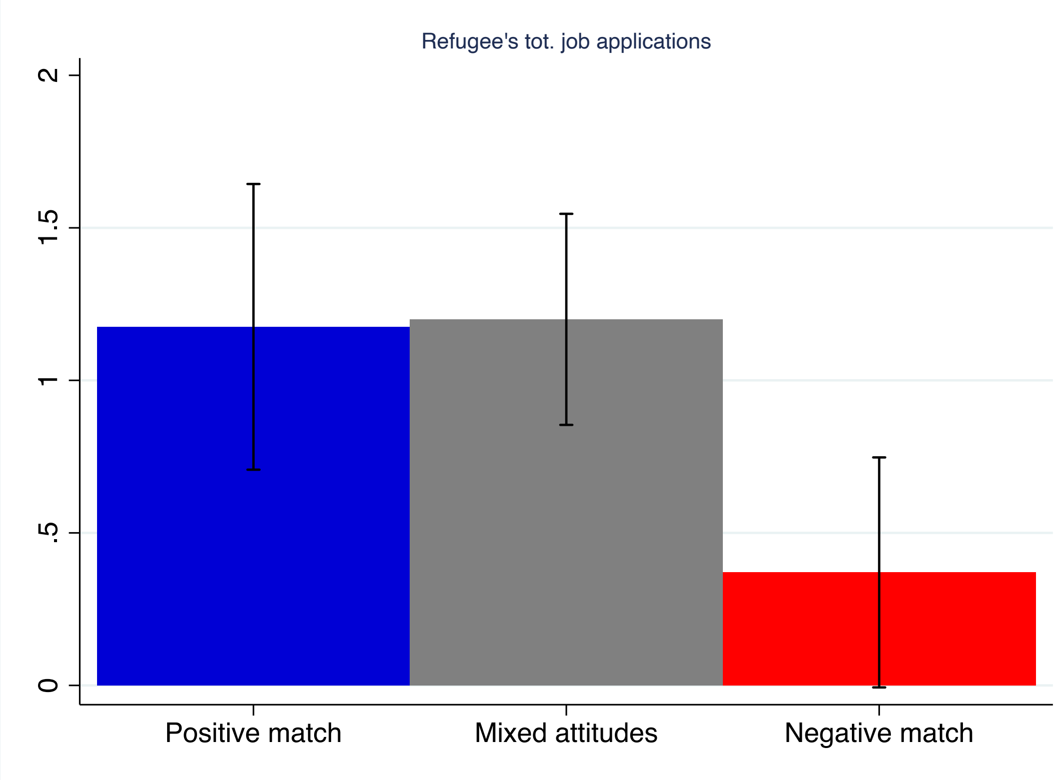
Notes: Group means and 95% confidence intervals around the mean. Answers reported by the employers who successfully matched with 1 refugee worker.

Figure A.21: Internship data: worker was actively looking for jobs before the experiment



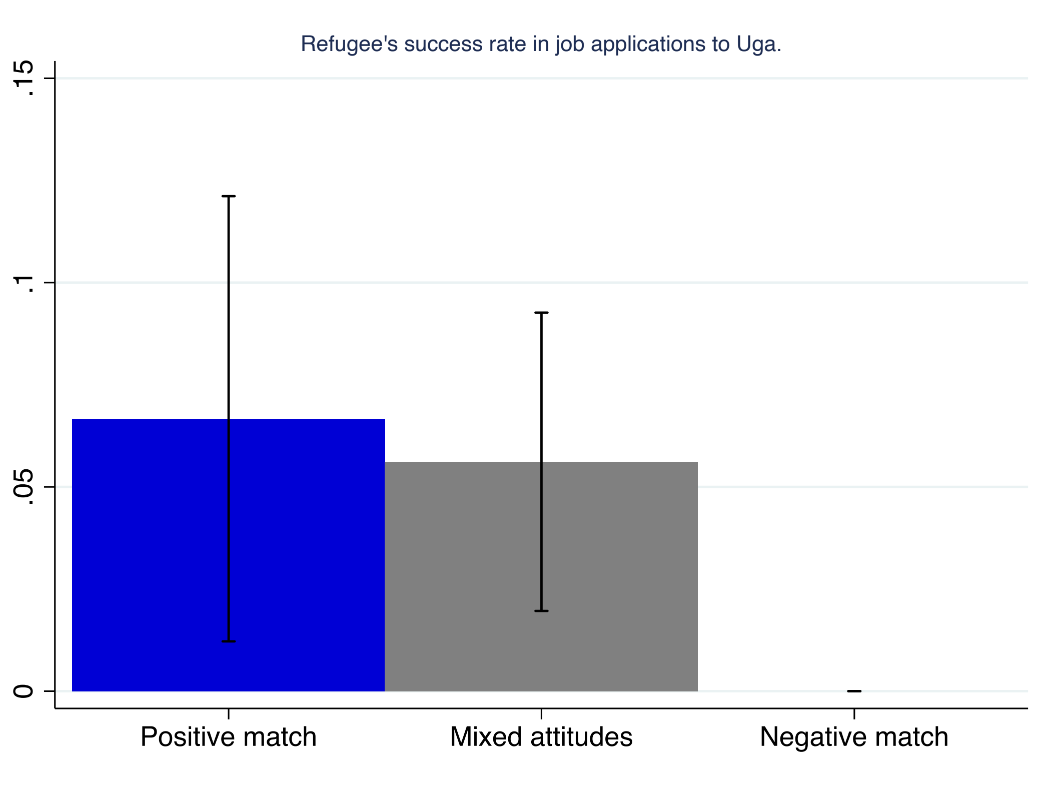
Notes: Group means and 95% confidence intervals around the mean. Answers reported by the refugee worker who matched with 1 of the firms.

Figure A.22: Internship data: worker was actively looking for jobs before the experiment at Ugandan firms



Notes: Group means and 95% confidence intervals around the mean. Answers reported by the refugee worker who matched with 1 of the firms.

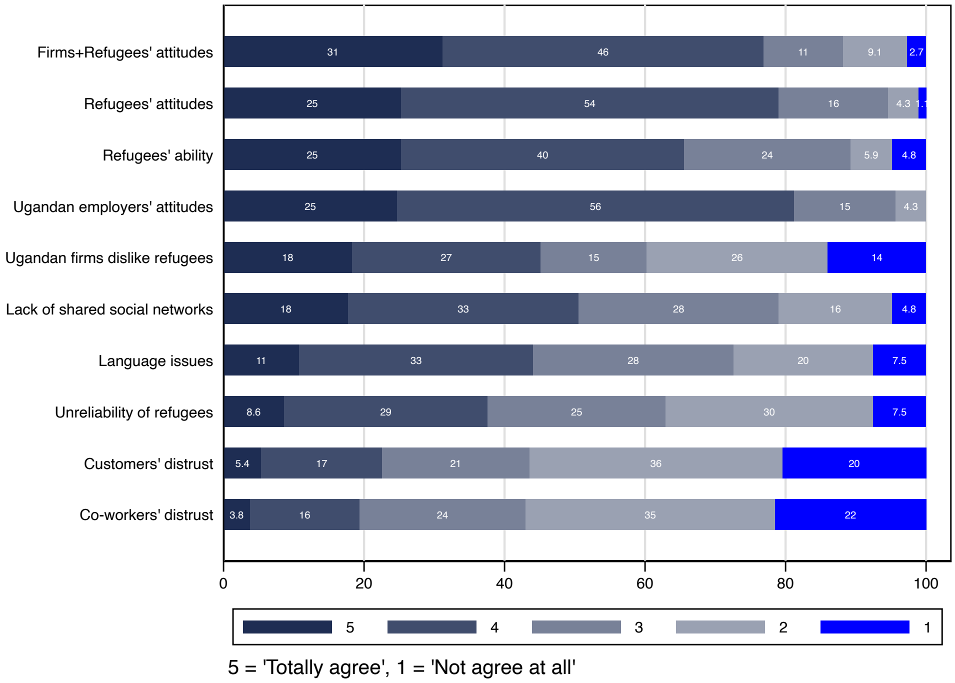
Figure A.23: Internship data: worker's job finding rate with Ugandan firms



Notes: Group means and 95% confidence intervals around the mean. Answers reported by the refugee worker who matched with 1 of the firms.

Figure A.24: Opinions about refugees and work together with them

(a) Distributions of opinions regarding working with refugees



(b) Percentage of control employers who agree or strongly agree for each statement

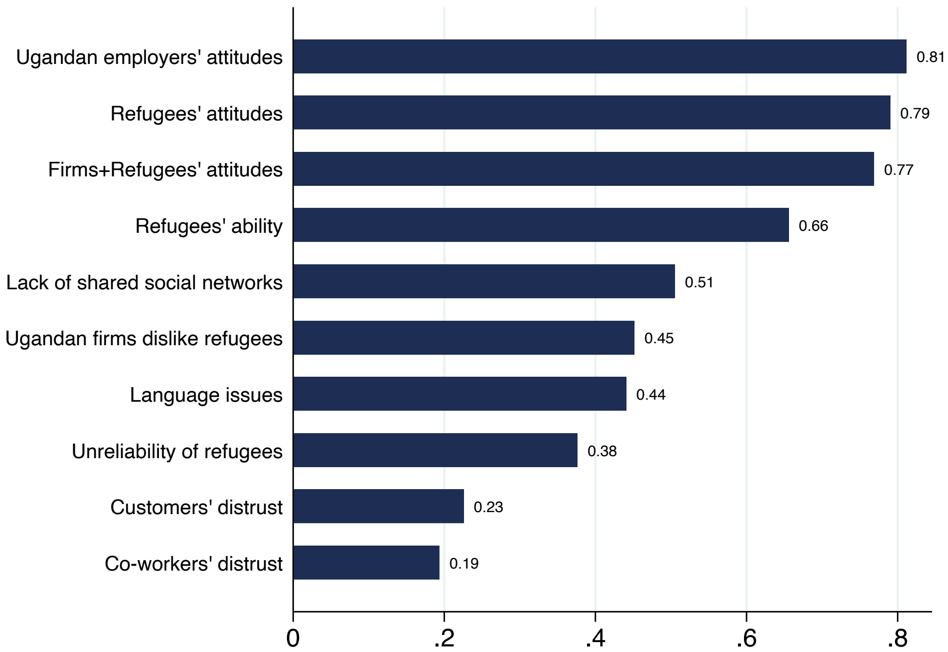


Table A.1: Skills tested for each occupation

Occupation	**Tested skill**
Baker	Bake a loaf of diabetic bread
Barber	Conduct a marines hair cut
Bead artist	Make beaded earrings
Beautician	Apply make-up on a client
Brick layer	Construct a header bond with attached stretcher
Carpenter	Make a small wooden chair
Cook	Cook rice pilao with beef stew
Domestic electrician	Wire and install two lamps in full conduit work
Electronics technician	Replace jack pin and mouth piece of a phone
Hairdresser	Twist style
Hairdresser	Cornrow style
Hotel receptionist	Make reservations and reserve a room for a guest
Hotel room attendant	Service a hotel room
Knitter	Make a long-sleeved sweater
Leather designer	Make a pair of men sandals
Motorvehicle mechanics	Repair car brakes
Painter	Paint interior walls of a medium-size room
Plumber	Fit and connect pipes
Tailor	Make a casual short-sleeved shirt
Waitron	Perform table food service and customer care
Weaver	Weave a table cloth
Welder	Make a small metallic window

Table A.2: Comparing refugees in the sample with locals in Kampala

	UNRHS			Baseline survey			
	N	Mean	SD	N	Mean	SD	Diff
High. educ.: None	601	0.02	0.14	527	0.01	0.10	-0.010
High. educ.: Primary	601	0.73	0.44	527	0.11	0.32	-0.617***
High. educ.: Secondary	601	0.23	0.42	527	0.88	0.33	0.644***
Employed	714	0.56	0.50	527	0.48	0.50	-0.079***
Unemployed	714	0.11	0.32	527	0.16	0.37	0.047**
Out of labor force	714	0.32	0.47	527	0.36	0.48	0.033
Monthly earnings	247	620.59	1108.03	255	301.54	294.08	-319.046***

Notes: this table compares the characteristics of our sample of refugees with a representative sample of Ugandans living in Kampala from the most recent wave of the National Household Survey conducted in Uganda (2018). The sample of Ugandans contain 714 individuals living in Kampala, and the sample of refugees is composed by 527 individuals.

Table A.3: Attrition at follow up

	Full sample		Exposed sample	
	(1)	(2)	(3)	(4)
	Follow-up 1	Follow-up 2	Follow-up 1	Follow-up 2
Treated	0.004	-0.010	0.005	-0.041
	(0.011)	(0.030)	(0.013)	(0.036)
Control	0.981	0.886	0.981	0.886
Firms	525	474	385	343

Table A.4: Comparing firms in the sample with other firms in Kampala

	Manpower survey			Baseline survey			
	N	Mean	SD	N	Mean	SD	Diff
Respondent is a woman	730	0.56	0.50	535	0.57	0.50	0.010
Age of the respondent	723	33.65	10.46	535	34.50	8.37	0.852
Education: None	726	0.04	0.20	535	0.02	0.14	-0.021**
Education: Primary	726	0.33	0.47	535	0.19	0.39	-0.144***
Education: Secondary	726	0.42	0.49	535	0.46	0.50	0.032
Education: Vocational	726	0.11	0.31	535	0.21	0.41	0.107***
Education: University	726	0.06	0.24	535	0.10	0.30	0.041**
Firm age	723	4.80	6.31	535	7.81	6.64	3.010***
Keeps accounting books	723	0.36	0.48	535	0.64	0.48	0.286***
Employees at baseline	723	1.99	2.33	535	2.49	3.15	0.497***
Revenues past month, M-UGX	720	0.86	2.58	499	1.88	2.77	1.017***
Expects future increase in size	730	0.10	0.31	535	0.86	0.35	0.756***

Notes: this table compares the characteristics of our sample of firms with a representative sample of firms in Kampala. The sample of 730 firms comes from the Manpower survey conducted in 2016 in Uganda.

Table A.5: Number of refugees hired, using a Poisson and a Tobit model

	Poisson		*Tobit*	
	(1)	(2)	(3)	(4)
Panel A: Average Treatment Effect				
Assigned-to-treat	0.867*	0.824*	0.857*	0.818*
	(0.447)	(0.445)	(0.481)	(0.464)
	[0.052]	[0.064]	[0.076]	[0.079]
N. Firms	474	474	474	474
Mean Control	0.048	0.048	0.048	0.048
Area FE	No	No	No	No
Panel B: Effect of exposure				
Exposed	0.926**	0.897*	0.945**	0.922**
	(0.426)	(0.472)	(0.400)	(0.424)
	[0.030]	[0.057]	[0.019]	[0.030]
N. Firms	343	343	343	343
Mean Control	0.048	0.048	0.048	0.048
Area FE	No	Yes	No	Yes

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

Table A.6: Learning, no area fixed effects

	Dependent variable:				
	(1) Hard skills	(2) Soft skills	(3) Trust	(4) Respect	(5) Avg. std. effect
Panel A: Full sample					
Assigned to Treatment	0.015	0.145	0.171*	0.077	0.102
	(0.100)	(0.102)	(0.099)	(0.101)	(0.083)
	[0.877]	[0.156]	[0.083]	[0.443]	[0.215]
N. Firms	525	525	525	525	525
Mean Control	-0.000	0.000	0.000	0.000	
Area FE	No	No	No	No	
Regr.	OLS	OLS	OLS	OLS	
Panel B: Exposed sample					
Exposed	0.111	0.291**	0.363***	0.175	0.235**
	(0.114)	(0.119)	(0.110)	(0.117)	(0.095)
	[0.334]	[0.015]	[0.001]	[0.137]	[0.014]
N. Firms	385	385	385	385	385
Mean Control	-0.000	0.000	0.000	0.000	
Area FE	No	No	No	No	
Regr.	OLS	OLS	OLS	OLS	

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

Table A.7: Learning, post-double lasso

	Dependent variable:			
	(1) Hard skills	(2) Soft skills	(3) Trust	(4) Respect
Panel A: Full sample				
Assigned to Treatment	0.018	0.149	0.176*	0.084
	(0.096)	(0.096)	(0.093)	(0.096)
	[0.851]	[0.120]	[0.060]	[0.381]
N. Firms	525	525	525	525
Mean Control	-0.000	0.000	0.000	0.000
Area FE	Yes	Yes	Yes	Yes
Regr.	PDS-L	PDS-L	PDS-L	PDS-L
Panel B: Exposed sample				
Exposed	0.125	0.300***	0.369***	0.210*
	(0.112)	(0.116)	(0.108)	(0.112)
	[0.267]	[0.010]	[0.001]	[0.062]
N. Firms	385	385	385	385
Mean Control	-0.000	0.000	0.000	0.000
Area FE	Yes	Yes	Yes	Yes
Regr.	PDS-L	PDS-L	PDS-L	PDS-L

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

Table A.8: Learning, individual components of hard and soft skills

	Dependent variable:					
	(1) Theory	(2) Practice	(3) Performance	(4) Time mgmt	(5) Team work	(6) Work ethics
Panel A: Full sample						
Assigned to Treatment	0.090	-0.063	-0.010	0.079	0.166	0.115
	(0.096)	(0.097)	(0.101)	(0.096)	(0.107)	(0.099)
	[0.347]	[0.516]	[0.918]	[0.411]	[0.120]	[0.246]
N. Firms	525	525	525	525	525	525
Mean Control	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Regr.	OLS	OLS	OLS	OLS	OLS	OLS
Panel B: Exposed sample						
Exposed	0.170	0.018	0.071	0.148	0.330**	0.269**
	(0.110)	(0.115)	(0.120)	(0.115)	(0.128)	(0.113)
	[0.125]	[0.877]	[0.552]	[0.198]	[0.011]	[0.017]
N. Firms	385	385	385	385	385	385
Mean Control	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Regr.	OLS	OLS	OLS	OLS	OLS	OLS

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

Table A.9: Willingness to pay to hire refugee

	Dep. var.: WTP (UGX)			*Tobit*	
	(1)	(2)	(3)	(4)	(5)
Panel A: Full sample					
Assigned to Treatment	-900.839	-987.741	-787.734	-2289.439	-2444.197
	(2015.344)	(2051.914)	(1955.189)	(3412.439)	(3406.924)
	[0.655]	[0.631]	[0.687]	[0.503]	[0.473]
N. Firms	525	525	525	525	525
Mean Control	15097.087	15097.087	15097.087	15097.087	15097.087
Area FE	No	Yes	Yes	No	Yes
Regr.	OLS	OLS	PDS-L		
Panel B: Exposed sample					
Exposed	-95.055	-453.374	-343.429	-1633.583	-2115.165
	(2375.682)	(2455.806)	(2354.275)	(4029.148)	(4064.149)
	[0.968]	[0.854]	[0.884]	[0.685]	[0.603]
N. Firms	385	385	385	385	385
Mean Control	15097.087	15097.087	15097.087	15097.087	15097.087
Area FE	No	Yes	Yes	No	Yes
Regr.	OLS	OLS	PDS-L		

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

Table A.10: Probability of having an open vacancy

	Dep. var.: Has a vacancy		
	(1)	(2)	(3)
Panel A: Full sample			
Assigned to Treatment	-0.093**	-0.091**	-0.085*
	(0.042)	(0.042)	(0.044)
	[0.028]	[0.032]	[0.053]
N. Firms	525	525	525
Mean Control	0.403	0.403	0.403
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L
Panel B: Exposed sample			
Exposed	-0.074	-0.063	-0.057
	(0.048)	(0.049)	(0.051)
	[0.122]	[0.196]	[0.266]
N. Firms	385	385	385
Mean Control	0.403	0.403	0.403
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

Table A.11: Best Linear Projector of CATE

	Best Linear Projector of CATE			
	Beta	SE	t-stat	p-value
Intercept	-.47	.356	-1.32	.187
Refugee's ability	-.035	.104	-.334	.739
Refugee's attitudes	.259	.106	2.446	.015
Refugee knowledge of languages	-.158	.167	-.941	.347
Refugee's age	-.001	.006	-.161	.872
Refugee is Congolese	.042	.162	.257	.798
Refugee ever employed by Ugandan	-.039	.128	-.307	.759
Employer's attitudes	.244	.118	2.075	.039
Firm's size	.021	.106	.202	.84
Firm's quality	0	.098	-.003	.997
Firm's beliefs	.028	.107	.264	.792
Firm's perceive cost of learning	-.044	.098	-.448	.655
Firm's expansion plan	-.051	.102	-.498	.619
Employer ever employed migrant	.033	.107	.312	.755
Manufacturing sector	.085	.119	.711	.477
Owner is Muganda	.111	.103	1.074	.284
Employer+refugee live same area	-.226	.154	-1.464	.144
Employer+worker same gender	.173	.132	1.314	.19

Notes: Best Linear Projector estimated using r-command `blp` from the Generalized Random Forest package `grf`. The only two variables with p-values less than 5% are refugee's attitudes (p-val = 0.015) and employer's attitudes (p-val = 0.039)

Table A.12: Doubly-robust estimator post-causal forest

	Doubly-robust estimators			
	Beta	SE	Lower CI (95%)	Upper CI (95%)
TxPositive	.2	.087	.03	.37
TxMixed	-.053	.065	-.179	.074
TxNegative	-.278	.128	-.53	-.027

Notes: Robust standard errors. We produce these estimates using the r-command `average_treatment_effect` from the Generalized Random Forest package `grf`

Table A.13: Learning, original treatments

	Dependent variable:				
	(1) Hard skills	(2) Soft skills	(3) Trust	(4) Respect	(5) Avg. std. effect
Panel A: Full sample					
Assigned to T2	-0.055	0.062	0.210*	0.176	0.098
	(0.115)	(0.116)	(0.117)	(0.120)	(0.096)
	[0.631]	[0.594]	[0.073]	[0.142]	[0.307]
Assigned to T1	0.082	0.194	0.137	0.006	0.105
	(0.122)	(0.127)	(0.122)	(0.116)	(0.101)
	[0.501]	[0.128]	[0.260]	[0.956]	[0.298]
N. Firms	525	525	525	525	525
Mean Control	-0.000	0.000	0.000	0.000	
p(T2=T1)	0.272	0.299	0.551	0.162	
Area FE	Yes	Yes	Yes	Yes	Yes
Regr.	OLS	OLS	OLS	OLS	
Panel B: Exposed sample					
Exposed+Certificate	0.029	0.200	0.442***	0.308**	0.245**
	(0.140)	(0.147)	(0.135)	(0.144)	(0.118)
	[0.834]	[0.176]	[0.001]	[0.034]	[0.038]
Exposed only	0.185	0.346**	0.282*	0.073	0.222*
	(0.151)	(0.158)	(0.144)	(0.147)	(0.125)
	[0.221]	[0.029]	[0.051]	[0.619]	[0.076]
N. Firms	385	385	385	385	385
Mean Control	-0.000	0.000	0.000	0.000	
p(T2=T1)	0.362	0.418	0.323	0.165	0.872
Area FE	Yes	Yes	Yes	Yes	Yes
Regr.	OLS	OLS	OLS	OLS	

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

Table A.14: Willingness to hire new refugees, original treatments

	Dep. var.: WTP ≥ 0		
	(1)	(2)	(3)
Panel A: Full sample			
Assigned to T2	-0.007	-0.009	-0.011
	(0.048)	(0.048)	(0.049)
	[0.880]	[0.848]	[0.820]
Assigned to T1	-0.026	-0.034	-0.026
	(0.049)	(0.050)	(0.050)
	[0.594]	[0.496]	[0.601]
N. Firms	525	525	525
Mean Control	0.709	0.709	0.709
p(T2=T1)	0.712	0.636	0.775
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L
Panel B: Exposed sample			
Exposed+Certificate	-0.023	-0.023	-0.026
	(0.057)	(0.058)	(0.059)
	[0.693]	[0.684]	[0.656]
Exposed only	0.030	0.018	0.023
	(0.059)	(0.061)	(0.058)
	[0.614]	[0.770]	[0.698]
N. Firms	385	385	385
Mean Control	0.709	0.709	0.709
p(T2=T1)	0.433	0.536	
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

Table A.15: Number of refugees hired, excluding matched refugee

	Dep. var.: Num. refugees hired		
	(1)	(2)	(3)
Panel A: Average Treatment Effect			
Assigned-to-treat	0.065**	0.064**	0.058**
	(0.031)	(0.031)	(0.030)
	[0.039]	[0.044]	[0.048]
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	0.077**	0.072**	0.068**
	(0.034)	(0.034)	(0.032)
	[0.024]	[0.036]	[0.034]
N. Firms	343	343	343
Mean Control	0.048	0.048	0.048
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

Table A.16: Number of refugees hired, by location of business premises

	Dep. var.: Num. refugees hired		
	(1)	(2)	(3)
Panel A: Areas with high concentration of refugees			
Exposed	0.080*	0.079*	0.090**
	(0.041)	(0.041)	(0.042)
	[0.052]	[0.052]	[0.031]
N. Firms	218	218	218
Mean Control	0.035	0.035	0.035
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L
Panel B: Areas with low concentration of refugees			
Exposed	0.032	0.019	0.030
	(0.053)	(0.062)	(0.049)
	[0.539]	[0.763]	[0.539]
N. Firms	125	125	125
Mean Control	0.068	0.068	0.068
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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). Areas with high concentration of refugees are the following: Makindye and Rubaga.

Table A.17: Number of Ugandans hired

	Dep. var.: Num. Ugandans hired		
	(1)	(2)	(3)
Panel A: Average Treatment Effect			
Assigned-to-treat	0.106	0.089	0.106
	(0.116)	(0.120)	(0.118)
	[0.359]	[0.458]	[0.367]
N. Firms	474	474	474
Mean Control	0.398	0.398	0.398
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L
Panel B: Effect of exposure			
Exposed	0.092	0.056	0.098
	(0.128)	(0.137)	(0.130)
	[0.474]	[0.684]	[0.451]
N. Firms	343	343	343
Mean Control	0.398	0.398	0.398
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

Table A.18: Attitudes

	Dep. var.: Attitudes towards refugees		
	(1)	(2)	(3)
Panel A: Average Treatment Effect			
Assigned-to-treat	0.023	0.020	0.022
	(0.025)	(0.025)	(0.024)
	[0.354]	[0.419]	[0.362]
N. Firms	474	474	474
Mean Control	0.657	0.657	0.657
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L
Panel B: Effect of exposure			
Exposed	0.007	0.003	0.006
	(0.031)	(0.031)	(0.031)
	[0.833]	[0.919]	[0.853]
N. Firms	343	343	343
Mean Control	0.657	0.657	0.657
Area FE	No	Yes	Yes
Regr.	OLS	OLS	PDS-L

Notes: ***, **, *, indicate significance at the 1%, 5%, and 10% levels respectively. 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).

B Appendix B: Script WTP

Introduction to WTP

The purpose of the exercise that will follow is to understand what is your "Willingness To Pay" for some workers. What we mean by this is the most that you would be willing to pay to hire a worker. Please, keep in mind that there are no right or wrong answers. We will just ask some questions to check your understanding.

Before moving on with the explanation, I would like you to think about the following situation: imagine a job seeker come to look for a job at your firm. Usually, after getting some information on her, you might already have in mind what you would be willing to pay to hire her. In other words, you might think about what is the maximum price at which you would still hire the worker. Since you do not know the salary at which she would be willing to work for you, the salary you think about is usually your own valuation of the worker. Talking to her, you learn about the actual salary she wants to receive and you decide whether to hire her or not. Your decision will depend on the salary the worker is willing to accept: if the salary is higher than your valuation, you will not hire the worker. If instead the salary is equal or lower than your valuation, you will hire her.

We will ask you to form your own valuation about the maximum salary you would pay for one worker looking to work for you for one week of probation. This worker is hypothetical, i.e. s/he does not exist, although his/her characteristics are very similar to the types of workers we have interviewed few months ago.

After you have thought about this salary, we will present you a list of 21 possible salaries for this worker for one week of work and we will ask you whether you would be willing to pay each possible salary for her. The salaries range from 0 UGX to 100,000 UGX and increase by 5,000 UGX each time. For example we will ask "Would you be willing to hire this worker for one week under probation if you have to pay her a salary of 10,000UGX?"; "Would you be willing to hire this worker for one week under probation if you have to pay her a salary of 15,000UGX?"; and so on.

Once you have answered all these questions, you will be given an envelope with a price like this one [Enumerator: show the envelope]. This price is between 0 and 100,000UGX. The price has been randomly selected by the computer and **I DO NOT KNOW IT, NEITHER I COULD CHANGE IT**.

If the maximum salary you agreed to pay in the 21 possible options is higher than the number in the envelope, you will get the worker for a probation period of one week, by agreeing to pay the salary you see in the envelope. Therefore, imagine this worker will start to work for you: at the end of the week, she will expect you to pay the agreed salary. If the maximum salary you agreed to pay is lower than the price in the envelope, you will not be able to work with this job-seeker.

Given the mechanism, it is in your best interest to be truthful, meaning to accept to pay salaries up to the maximum amount you are willing to pay for the worker. In this way you will never pay more than the maximum value the worker has for you and you could end up paying less.

Moreover, the price you stated will affect your chance of hiring the worker but might not be the price you will actually pay. The price you will pay is fixed and your valuation will not change it.

Remember that this worker is hypothetical. However, it is important to us that you take the choices seriously, and do your best to give us the answer you would give if they were real workers.

Multiple Price List

- Show hypothetical candidate.
- Would you be willing to hire this worker for one week under probation, starting up to 4 days from now, if you have to pay her a salary of 0UGX?
  - If no: Are you sure you **don't** want to hire this worker even if for free?
  - If sure: You said you are not willing to hire this worker even if for free. Can you tell us why?
- If yes: Would you be willing to hire this worker for one week under probation, starting up to 4 days from now, if you have to pay her a salary of 5,000UGX?
- Are you sure you **don't** want to hire this worker for 5,000UGX?
- ...
- ...Would you be willing to hire this worker for one week under probation, starting up to 4 days from now, if you have to pay her a salary of 100,000UGX?

Comprehension checks

- Final wage the respondent agrees to pay is X UGX.
- Suppose that the price in the envelope is: $X - 5,000$. What will happen?
  - If no: Enumerator, explain respondent the procedure one more time and ask this question again.
  - After you do this, ask again: Suppose that the price in the envelope is: $X - 5,000$. What will happen?
  - If no: Enumerator, explain respondent the procedure one more time and ask this question again.
  - Is the procedure clear now?
- Suppose that the price in the envelope is: $X + 5,000$. What will happen?
  - If yes: Enumerator, explain respondent the procedure one more time and ask this question again.
  - After you do this, ask again: Suppose that the price in the envelope is: $X + 5,000$. What will happen?
  - Enumerator, explain respondent the procedure one more time and ask this question again.
  - Is the procedure clear now?

C Appendix C: Outcomes

- The firm's beliefs about the hard and soft skills of a generic refugee worker who may come and look for a job in the future (measured only at endline), using Likert scales between 1 and 5, and aggregated in an index following Anderson (2008):
  - Think about this worker's theoretical skills (e.g.theoretical skills that are relevant to work in a firm like yours). On a scale between 1 and 5, where 1="Not at all competent" and 5="Very competent", how competent do you think this person will be?
  - Think about this worker's practical skills (e.g., technical skills that can be applied to work in a firm like yours). On a scale between 1 and 5, where 1="Not at all competent" and 5="Very competent", how competent do you think this person will be?
  - Think about this worker's performance at work (e.g., in terms of units serviced, quantity, pieces completed, etc.). On a scale between 1 and 5, where 1="Terrible" and 5="Excellent", how do you think he will perform?
  - Think about this worker's time management ability (i.e., the ability of completing an assigned task meeting a deadline). On a scale between 1 and 5, where 1="Terrible" and 5="Excellent", how do you think he will perform?
  - Think about this worker's team work ability (i.e., the ability of working in a team with other employees). On a scale between 1 and 5, where 1="Terrible" and 5="Excellent", how do you think he will perform?
  - Think about this worker's work ethics (i.e., discipline and hard-work abilities). On a scale between 1 and 5, where 1="Terrible" and 5="Excellent", how do you think he will perform?
- The firm's WTP to hire a new, hypothetical refugee worker (measured at follow-up 1). We show the profile of the same worker to all the firms, changing only the gender and the occupation of the worker, to being exactly the same as the ones of the worker already proposed at baseline.

D Appendix D: Indices and variables used in the causal forest

- Firm owner belongs to majority ethnic group in Uganda (i.e. is a Muganda)
- Index on firms' attitudes, constructed using the first factor in a factor analysis including i) whether the firm agrees or strongly agrees with the sentence: "When jobs are scarce, Ugandans should have more rights to a job than refugees", and ii) whether the firm believes that the law should not allow refugees to work in Uganda
- Firm's initial beliefs about the hard and soft skills of the matched worker
- Firm's perceived cost of learning about the quality of a refugee, constructed using the first factor of a factor analysis taking into account i) days it would take to learn about the refugee worker's skills (both hard and soft), and ii) beliefs that refugees fail at tests such as the one on practical skills provided by the DIT
- Firm's willingness to expand, with a first factor of a factor analysis including whether the firm has a vacancy and whether the firm expects to increase its size in the next 5 years
- Firm's quality, constructed using a index including whether the owner owns the business premises, the owner's years of education, a dummy equal to 1 if the firm is formal and pays taxes to the local revenues authority, if keeps accounting books, has a separate bank accounts from the owner's one, and whether it advertises its products or services
- Firm's size at baseline, using an index including number of employees at baseline, number of tasks performed in the firm and number of rooms in the business premises
- An indicator equal to 1 if the firm is active in a manufacturing sector (arts and crafts; bakery; carpentry; leather works; metal works; tailoring)
- An indicator equal to 1 if the firm has ever employed a migrant

- Refugee's ability, constructed taking the first factor from a factor analysis feeded with: test score, experience, education and cognitive skills.
- Refugee's attitudes, constructed using a first factor analysis considering the following dimensions: i) whether the refugee worker perceived to be discriminated in Uganda; ii) whether she thinks that Ugandans are not trustworthy; iii) whether she thinks that working together does not help; and iv) whether she feels distant and different from Ugandans
- Refugee's experience with working with Ugandans, using a dummy equal to 1 if the refugee worker has ever worked for a Ugandan employer.
- Refugee's knowledge of English and the most important local languge, Luganda
- Refugee's age
- An indicator equal to 1 if the refugee worker is a Congolese (the majority in our sample)
- An indicator equal to 1 if the refugee worker and the firm live in the same narrowly defined neighborhood
- An indicator equal to 1 if the refugee worker and the firm are of the same gender

E Appendix E: Take up

In section 3 we described the random assignment of the firm-refugee matches. We find that not all of the refugees show up at the appointment. In this section, we describe the characteristics of the refugees who did not take up the offer and compare them to those of the refugees who showed up at the appointments.

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). We can investigate whether any observable characteristics correlates with the likelihood of matching, both at the refugee and the firm level. Using the rich data collected at baseline from both samples, we run the following specification in the sample of refugees matched with treated firms:

$$y_j = \gamma_0 + \gamma_1 Matched_j + X_j^{\prime}\delta + \varepsilon_j, \tag{5}$$

where the coefficient of interest, γ_1, correlates characteristic y_j with a dummy equal to 1 if the refugee worker j showed up at the meeting with the firm. The specification uses robust standard errors and controls for strata fixed effect, that is the occupation of the refugee worker.

Results from 5 are reported in Table A.19. We find that refugees who took up the offer are more likely to be self-employed, relatively more wealthy, have smaller households and larger household income per capita. Importantly, they are also less likely to be unemployed. Furthermore, refugees show up at internships in locations mostly populated by refugees, such as Rubaga, and less so in places further away from the neighborhoods where they usually live, such as Wakiso, Kawempe, and Nakawa. The offer of 50,000UGX in two installments as a reimbursement to undertake the internship may not have been enough to incentivize refugees to take up the offer. This payment, approximately equal to 15USD, is 33% larger than the median starting salary paid the firms in our sample, and equal to about 85% of the median *monthly* refugee earnings in our sample before the experiment. Importantly, we report that unemployed refugees have significantly less savings than employed ones (40,000UGX versus 133,000UGX). These results suggest that the offer attracted refugees who could afford to travel more regularly to the firms where the internship took place, and points that liquidity constraints are important factors hindering refugees' labor market integration, similar to the findings in Caria et al. (2020).

Table A.19: Refugees' take up of the internships

	Not matched			Matched			
	n	mean	sd	n	mean	sd	Diff
Refugee worker is a woman	136	0.69	0.46	182	0.68	0.47	0.014
Age of the refugee worker	136	32.10	10.62	182	34.26	10.29	2.188*
Refugee worker is Congolese	136	0.84	0.37	182	0.87	0.33	0.030
Years living in Uganda	136	6.65	4.05	182	6.83	3.81	0.250
Years of education	136	11.34	3.92	182	11.71	3.59	0.271
Work experience (years)	136	4.41	7.38	182	4.85	6.25	0.530
English speaking level	136	2.71	1.23	182	2.68	1.05	-0.079
Luganda speaking level	136	2.78	1.25	182	2.65	1.15	-0.162
Index on refugee attitude	136	-0.01	0.61	182	0.04	0.53	0.036
Tot. num. adults	136	3.41	2.27	182	2.84	1.65	-0.565**
Tot. num. children	136	2.55	1.93	182	2.60	1.99	0.046
HH inc./adult('000UGX)	136	122.94	150.03	182	156.75	139.87	29.235
Index food security, May 21	136	0.85	0.26	182	0.81	0.30	-0.039
Remittances received('000UGX)	136	40.00	153.02	182	63.79	174.13	23.799
Tot. savings, Sept 21	136	69.76	143.88	182	98.74	206.12	27.392
Has received relief aid, Sept 21	136	0.18	0.39	182	0.21	0.41	0.032
Life satisfaction, 1-10	136	2.28	1.51	182	2.15	1.51	-0.122
Ever employed by Ugandan	136	0.30	0.46	182	0.27	0.45	-0.030
Was employed by someone, Sept 21	136	0.10	0.31	182	0.10	0.30	-0.004
Was self-employed, Sept 21	136	0.32	0.47	182	0.41	0.49	0.086
Unemployed, past 7 days, Sept 21	136	0.23	0.42	182	0.13	0.34	-0.096**
Out of labor force, past 7 days, Sept 21	136	0.35	0.48	182	0.36	0.48	0.015
Hours worked past 7 days	136	18.10	22.51	182	20.46	20.47	2.262
Total earnings, past 30 days, Sept 21	136	130.13	219.25	182	191.66	261.33	56.698**
Looked for jobs, past 30 days, Sept 21	136	0.32	0.47	182	0.31	0.47	-0.008
Hours spent looking for jobs, Sept 21	136	1.82	6.84	182	3.43	10.09	1.541
Willing internship if 20km away	114	0.39	0.49	152	0.47	0.50	0.072
Willing internship if 15km away	114	0.46	0.50	152	0.49	0.50	0.041
Willing internship if 10km away	114	0.61	0.49	152	0.59	0.49	0.000
Willing internship if 5km away	114	0.86	0.35	152	0.81	0.39	-0.047
Willing internship if 1km away	114	1.00	0.00	152	0.95	0.21	-0.048***
Interested in unpaid one-week internship	136	0.91	0.28	182	0.95	0.23	0.032
Minimum wage for one-week internship	132	2651.52	10690.64	180	2833.33	17151.15	-173.448
Internship located in Central	136	0.11	0.31	182	0.16	0.37	0.040
Internship located in Kawempe	136	0.14	0.35	182	0.09	0.29	-0.051
Internship located in Makindye	136	0.26	0.44	182	0.27	0.44	0.008
Internship located in Nakawa	136	0.12	0.32	182	0.05	0.22	-0.065**
Internship located in Rubaga	136	0.32	0.47	182	0.41	0.49	0.101*
Internship located in Wakiso	136	0.05	0.22	182	0.02	0.13	-0.033

Notes: this table compares the characteristics of the refugee workers who showed up at the internship versus those who did not. The variables come from the baseline survey with the sample of refugees. Each row is an individual dependent variable from the following specification: $y_j = \gamma_0 + \gamma_1 Matched_j + X_j^{'}\delta + \varepsilon_j$, where $Matched_j$ is a dummy equal to 1 if the refugee worker showed up for the internship and X_j is a matrix of strata FE. Each regression uses robust standard errors.

3

Chapter 3: Market Design for Land Trade: Evidence from Uganda and Kenya

This paper has been jointly written with Gharad Bryan, Jonathan de Quidt, Tom Wilkening and Nitin Yadav

Chapter 3: Table of Contents

1 Introduction

Inefficient land allocation reduces productivity in low-income country agriculture. Farms are small and fragmented (Adamopoulos and Restuccia, 2014; Ali et al., 2015; Deininger et al., 2016), despite the fact that labor and total factor productivity increase with farm size (Foster and Rosenzweig 2022; Aragón et al. 2022b). Land is also misallocated – there is substantial heterogeneity in farmer productivity but almost no correlation between farmer productivity and land holding (e.g., Chen et al. 2022b). These inefficiencies suggest large unrealized gains from trade, a claim borne out by quantitative and experimental analyses.[1] We argue that improved market design can help unlock these gains by creating trading rules tailored to address key frictions in the land market, and note that it is likely complementary to other institutions, such as property rights, that are more often emphasized (e.g., Besley and Ghatak 2010).

We build our argument in three steps. First, we document, with a survey of smallholder farmers in Uganda, that decentralized land trade—based on buyers and sellers bargaining over individual plots—is likely to be inefficient. Farmers believe their trading environment has key characteristics that the market design literature predicts will inhibit trade. Second, we provide lab-in-the field evidence that decentralized trade is indeed inefficient. We build a stylized representation of the environment, consistent with our survey evidence, in which real farmers trade fictitious land titles with strong financial incentives. We let them trade for a week without any formal trading rules, and show that the final allocation is far from efficient. Third, we show that specifying rules targeting the market frictions we highlight improves outcomes in the same land trade game. We tested interventions that range from simple, easy-to-understand rules that would be expected to facilitate trade in a wide range of problems, to rules that are highly tailored to the land trade problem but potentially difficult for our target audience to understand. Overall, we find that increasingly tailored rules, despite their increasing complexity, consistently improve efficiency without increasing inequality.

Our work complements prior studies that have established a set of missing markets and institutions that inhibit trade, focusing largely on property rights.[2] Despite the

[1] Acampora et al. (2022) provide experimental rental subsidies, and find positive returns that exceed the payments, consistent with unrealized gains from trade. While showing an improvement, they do not attempt to estimate the total set of gains available, nor to understand what proportion of surplus is unlocked by their subsidy. Our goal is to design a mechanism that can unlock a large proportion of the available gains. Several studies estimate gains from land reallocation using quantitative models. Estimated returns vary widely (from about 20% to over 300%), but are typically positive (e.g. Chari et al. 2020; Chen et al. 2022a,b; Bolhuis et al. 2021; Adamopoulos et al. 2022; Britos et al. 2022). See Gollin and Udry (2021) and Aragón et al. (2022a) for a discussion of some of the empirical challenges in this literature.

[2] de Soto 2000; Deininger and Feder 2001; Deininger and Jin 2006; Goldstein and Udry 2008; Field 2007; Besley and Ghatak 2010; Galiani and Schargrodsky 2010, 2011; Fenske 2011; de Janvry et al. 2015;

success of this literature, and the wide spread adoption of programs such as land titling, recent work suggests that these innovations are necessary but not sufficient to achieve efficiency. Property rights reforms such as rental market liberalization and titling programs have been only partially effective in realizing gains from trade (Ali et al., 2014; Bolhuis et al., 2021; Chen et al., 2022a; Chari et al., 2020), and historical studies show that initial misallocation of land can persist for decades even in environments where property rights are well established (Bleakley and Ferrie, 2014; Smith, 2019; Finley et al., 2021). Perhaps recognizing this, governments in many countries have opted for centrally-planned land consolidation and reallocation programs, even when other institutions are well functioning (see e.g., Hartvigsen, 2014). We believe that addressing the frictions we identify can help unlock the benefits of land titling, and that our market designs, which rely on voluntary trade and leverage farmers' own preferences and information, may be preferable to governmental consolidation programs in settings with low state capacity, low trust in government, and risk of expropriation.

In the tradition of the market design literature (e.g., Roth 2002), our exploration takes place within the confines of a simplified, lab-in-the-field, environment. We believe our designs capture the key constraints that market design can address, and abstract from problems that are best addressed elsewhere. For example, we do not allow for risk of fraud. While fraud may be an important part of what constrains land trade, we think that this is best addressed through complementary policies, rather than directly in the market design.[3] The upshot of this is that our estimates of the impact of market design should be seen as conditional on getting other institutions right. Whether real world gains would be larger or smaller depends on the extent of complementarity between market design and other programs, and it may well be that market design is a strong complement to other interventions.[4]

Lawry et al. 2016; Agyei-Holmes et al. 2020 all consider property rights. Other factors include incomplete credit markets (Eswaran and Kotwal, 1986; Binswanger and Rosenzweig, 1986; Binswanger and Elgin, 1988; Carter and Mesbah, 1993), low-quality maps (Libecap and Lueck, 2011; D'Arcy et al., 2021), and culture (Platteau, 2015).

[3]Other concerns that we think are real but do not address include: dispute and uncertainty of titling (best addressed through the legal system, but also perhaps not that important e.g., Ali et al. 2014), savings constraints that impede sales because land is the preferred store of value (best addressed through improved financial access), gains from fragmentation due to risk mitigation (e.g., McCloskey 1975, best addressed by insurance markets, and likely less important as technology improves Foster and Rosenzweig 2022), and cultural considerations such as taboos surrounding trade of ancestral land, which may evolve as more trade takes place.

[4]For example, farmers in many places may choose to keep fragmented farms in order to insure against local weather events even though such fragmentation is inefficient. A first-order benefit of better insurance is therefore the efficiency gains that can come from consolidating land into larger plots. We show that this consolidation is difficult without properly designed markets. As such, much of the potential gains from an insurance market intervention may be lost without also considering how to better facilitate trade.

Section 2 presents our survey evidence from a sample of 1,404 Ugandan smallholder farmers. We document five *facts*: 1. farmers believe there are increasing returns to scale at the *plot level* that are currently unrealized; 2. farmer ability and land quality are heterogeneous and complementary; 3. there are decreasing returns to scale at the *farm level* because farmers predict limits to how much land they can productively cultivate; 4. cultural constraints imply not all plots would be tradable, even in a well-functioning market; and 5. farmers have private information about gains from trade, but do not believe there is asymmetric information about land quality, i.e. no "lemons" problem.

Section 3 introduces the game that we use to study the land trade problem, which incorporates the facts in a stylized way. Figure I gives an example. Gains from trade arise from *consolidation* of multiple non-contiguous plots to a single contiguous unit, benefiting from increasing returns at the plot level (fact 1), and from *sorting*, when higher-ability farmers are assortatively matched to better-quality land (fact 2). But decreasing returns at the farm level (fact 3) imply that surplus can be lost if trade leads some farmers to hold too much land (we call these *exposure* losses because they arise due to exposure risk, as we explain below).

We use the game to illustrate how the five facts jointly imply three *frictions* predicted to impede decentralized trade: I) thin markets; II) exposure risk; and III) coordination frictions. Thin markets (where only a small number of buyers and sellers are willing to trade a particular good) exacerbate two-sided information problems á la Myerson and Satterthwaite (1983). Exposure risk is a generalization of hold-up that occurs when a sequence of trades is needed to realize a surplus, but later trades cannot be guaranteed (Goeree and Lindsay, 2019). Coordination frictions arise when efficiency requires coordinating the actions of many traders (Milgrom, 2017). Friction I is mostly due to plot-level increasing returns (fact 1), which implies a given plot's likely buyers are a small set of farmers with adjacent land. Frictions II and III arise because farmers often wish to condition current trades on future trades with other parties, leading to chains with many participants. These frictions mean that in a decentralized trade environment, farms are likely to remain fragmented and poorly sorted.

Section 5 demonstrates using a lab-in-the-field experiment that decentralized trade is indeed inefficient in a setting characterized by our five facts.[5] In the experiment, our sample of Ugandan farmers tried to consolidate and sort land in the land trade game introduced above, by freely trading hypothetical land allocations in trading periods lasting one week. Despite strong financial incentives, they realized only 23% of the

[5] We follow the engineering philosophy advocated by Roth (2002), that suggests using lab experiments to build an understanding of the environment and iterate on design. Lab-in-the-field experiments also allow us to isolate market design considerations from other contracting and property rights issues. By isolating one channel we can improve designs prior to direct interventions.

potential gains from trade. Furthermore, there is striking heterogeneity across sources: they realized around 61% of potential consolidation gains, but only 25% of sorting gains. This is important, because lack of sorting is the primary inefficiency emphasized in the quantitative literature. Participants also frequently end up with exposure losses, indicating failure to complete a planned sequence of trades. The results do not reflect a general inability to trade: participants reached over 90% efficiency in a pair of training games similar to those in Chamberlin (1948).

The inefficiency of decentralized trade opens the possibility that careful market design could improve outcomes. Design can address the three frictions by increasing thickness (friction I); enforcing conditional contracts (friction II); and helping farmers find chains (friction III). But the designer faces a trade-off between participant understanding, and the value of tailoring a design to the specific setting. Tailored designs are predicted to improve efficiency for rational agents, but they can often be complex and unfamiliar. For example, the antique market is likely thin, so antique fairs, which centralize trading at a specific time and place, likely increase efficiency and are easily understood. Centralization could also thicken the housing market (e.g., through online real estate aggregators), but houses suffer from an additional acute exposure problem: I only want to buy your house if I can sell mine. A more tailored market design allowing conditional contracts could further increase efficiency, but may in fact reduce efficiency by making transactions harder for inexperienced traders to understand. The trade-off between tailoring and familiarity is critical for our target population of low-numeracy smallholder farmers, and it is an empirical question which effect will dominate.

Finally we present direct evidence on the effectiveness of market design. We use lab-in-the-field experiments to show that improved rules can improve efficiency, and that our most-tailored designs perform particularly well. We begin with a simple market centralization intervention. After a week of decentralized trade, all participants in the above-described experiment were given a surprise opportunity to trade together in a single location for about one hour. As noted, this should thicken the market and may also help find chains and enforce conditionality, because agreements to trade will be more observable to others in the social network. We find substantial increases in efficiency, with gains coming from additional consolidation (reaching 70% of the optimum) and near-complete unwinding of exposure losses.[6] But we see zero additional sorting, and overall efficiency remains below 50%.[7]

[6] These improvements are not simply driven by additional time – we show below that the impacts of the centralization intervention dwarf those of exogenous variation in time to trade.

[7] Why did centralization work when farmers could haved done it themselves? We find 89% of villages tried to organize a central trading opportunity. In section VII we show endogenous centralization was not successful, and conjecture this shows the importance of external pressure to ensure participation.

We then study the impact of further tailoring. Section 6 introduces a second experiment, conducted in Kenya. We created a centralized, computerized land exchange on which farmers played a smaller version of the land trading game with six players and two plots each. We test three designs, increasing in how well tailored they are to the problem. First, we created a version of the classic continuous double auction (CDA), in which bidders could bid to buy or sell a single plot at a time; a general-purpose untailored design.[8] The exchange enforces one condition: e.g., "I will sell this plot conditional on receiving at least X shillings" and we refer to it as "Package-1." Our most tailored design, "Package-4," is also based on the CDA, but is a package exchange in which farmers could make offers with up to four conditions: e.g., "I will sell these two plots, conditional on receiving those two plots and paying no more than Y."[9]

Package-4 is highly tailored to our environment: reaching efficiency is possible with just one transaction per participant, potentially eliminating exposure risk, and the platform is responsible for finding chains, setting prices and enforcing all conditions. The market is also thickened because conditioning purchases on sales removes the importance of the initial endowment, leading to more potential purchasers for each plot. By facilitating wholesale realocation with a single bid, Package-4 may be particularly effective at facilitating efficient sorting. In contrast, Package-1 requires many trades to reach efficiency leaving open the exposure problem, and does little to help find chains or enforce conditions. We also examined an intermediate "Package-2" treatment that permitted swapping one plot for another; we conjectured that this would help with consolidation, but less with sorting.

Efficiency in the benchmark Package-1 condition is quite high, around 70%, indicating that farmers could understand and use the exchange platform. The most-tailored treatment, Package-4, increases efficiency by 7 percentage points relative to Package-1. Moreover, Package-4 is particularly effective at unlocking sorting gains that were

[8] Continuous double auctions have a long tradition in experimental research, dating back to the classic work of Smith (1962, 1964). Empirically, the double auction has good efficiency properties even when the number of buyers and sellers is small (e.g., Smith and Williams (1981); Smith (1990)). We chose the double auction as our starting point since it can be implemented both synchronously and asynchronously and does not require an auctioneer to start and stop bid rounds and hence could realistically be used in the village setting with low variable costs.

[9] Our algorithm is based on Goeree and Lindsay (2019). Unlike their design, we imposed XOR bidding which ensures that only one bid from each player is triggered in any transaction. We also allow communication, use a different visualization protocol, and have larger packages, tailored to our setting. While there has been considerable work on package auctions, package exchanges have attracted less attention. Combinatorial exchanges have been explored in the context of airport take-off and landing slots (Rassenti et al. 1982; Grether et al. 1989; Balakrishnan 2007), native vegetation offset permits (Nemes et al. 2008), pollution permits (Fine et al. 2017), housing (Goeree and Lindsay 2019) and spectrum reallocation (Milgrom and Segal, 2017). See Milgrom (2007) and Loertscher et al. (2015) for reviews.

not realized by any of our other interventions. Collecting findings, we find that increased tailoring always increases efficiency, a result that initially seemed unlikely in our low-numeracy setting.

More tailored designs improve efficiency, but it is also important to understand distributional effects. A particular concern is that complex rules might enable sophisticated traders, who understand the mechanism well, to profit at others' expense. We investigate this using an Atkinson index to measure (in)equality of final outcomes. Surprisingly and encouragingly we find that our more tailored designs *reduced* inequality.

Finally, Section 7 investigates some of the frictions that restrict trade. We experimentally manipulated the presence of non-tradeable plots in our first experiment (in line with cultural constraints, fact 4), and credit constraints in our second experiment. These features are predicted to worsen the thin markets and exposure problems. We find that they did not affect efficiency, but led to more unequal outcomes. We also provide evidence on second-best trading strategies used in the absence of tailored rules, and show tailoring crowds them out.

Closely related to our problem but not our setting, our package exchange was inspired by Goeree and Lindsay (2019), who study house reallocation. They propose, and experimentally verify, that a package market can help overcome exposure risk. An important precedent to our work is Tanaka (2007), which compares the performance of different land-consolidation mechanisms that might overcome exposure problems in lab experiments with US college students. We differ from Tanaka by studying a set of trading frictions including the exposure problem, and asking whether market design can work for actual farmers. Less closely related, the design literature related to land has concentrated on the land assembly problem where a single buyer wants to buy complementary plots from a number of small sellers, and in which a holdout problem occurs. Several papers explore solutions (e.g, Plassman and Tideman 2010, Kominers and Weyl 2012, Grossman et al. 2019, Sarkar 2017 and Sarkar 2022).

A few other studies have investigated efficiency in decentralized trading games in low-income country settings. Fiala (2015) and Bulte et al. (2013) find around 80-90% efficiency in trading games similar to the training games we use in our first experiment. This is slightly lower than our training games and much higher than our decentralized land trade game. Lowe (2021) conducts a game in which participants begin with mismatched goods (e.g. two left gloves) and have 4–5 days to find a partner with whom to swap; 88% succeed. This introduces complementarities so is conceptually related to our land trading game, but does not include the other frictions that we argue make efficient land trade difficult (e.g., there is no need for chains). Overall these findings are consistent with our claim that low efficiency in our decentralized land trade game is due to the frictions we identify, not a general inability to trade.

A small number of papers discuss market design for other development challenges. Many of these expect government to be the main buyer, for example, incentives for vaccines (Kremer, 2001b,a), refugee allocation (Delacrétaz et al., 2020), and antiquity protection (Kremer and Wilkening, 2015). Closer to our motivation, Hussam et al. (2022) explore how revelation mechanisms might be used to select quality borrowers.

2 Describing the land trade problem

We start with descriptive evidence from a survey of 1,404 farmers from 68 villages (LC1s) in Masaka district, Uganda. We had two aims in fielding the survey: to understand the environment and whether it has features that theory predicts will hinder trade; and to validate the design of our experimental games. We first discuss the characteristics of the sample, before documenting five key facts about the production technology and trading environment.

We began with a sample of rural villages in Masaka district, selected to ensure that agriculture was an important part of the economy.[10] In each village we asked the village leader (LC1 chairperson) for a list of households that would likely be willing to play our trading games three times over three weeks.[11] From that list, we randomly selected 22 households subject to the household cultivating some land, and deriving at least 50% of household income from farming.[12] We invited household heads but households could send another member if the head was unavailable. Table B1 describes the resulting sample, and provides comparable national averages for farm households in the World Bank's Living Standards Measurement Study (Uganda National Bureau of Statistics, 2020). Our sample is reasonably representative, but slightly more educated (7 years on average) and with higher farm incomes.

Fact 1: Consolidation Gains due to Increasing Returns at the Plot Level

Our survey evidence strongly suggests that farms are fragmented and that farmers believe that there are increasing returns at the plot level. Table I summarizes the data. Fragmentation is clearly present. Over 60% of our sample owns two or more fragmented

[10] We excluded coastal villages and those with very high or low population density, to focus on agricultural villages. We then selected a random sample of villages stratified by parish. See Appendix A for more details.

[11] We briefly described the games to the chairperson and showed some of the materials, to give them enough context for their recommendation. We did not explain the mechanics of the experiment nor any treatments.

[12] There were zero exclusions in 93% of villages. We also required participants to have access to a mobile money account (theirs or a friend's or family member's) so that they could be paid. This did not lead to any exclusions.

plots, which are on average 24 to 41 minutes' walk from one another. Inherited plots are much closer to home than those that are purchased or received as a gift, which is consistent with a poorly-functioning land market in which it is hard to purchase new plots close to existing ones.

Farmers recognize that this fragmentation is costly.[13] When directly asked, 91% of respondents stated that they would prefer a consolidated 2-acre plot to two 1-acre plots. Restricting to participants who own fragmented plots, about 88% of them believe their earnings would increase if all of their land was consolidated, and estimate an average earnings gain of 53%. We also asked all participants their beliefs about the returns to adding a third acre to an initially-consolidated 2-acre farm. If the third acre is contiguous with the others, 66% of participants believe in increasing returns to scale, while 33% believe returns would be constant. If the third acre is not contiguous, only 30% believe there would be increasing returns, 37% expect constant returns, and 33% predict decreasing returns.

Fact 2: Sorting Gains due to Farmer-Farm Complementarity

Our sample also believes there is heterogeneity in farmer ability and farm quality, and that these are complements, implying gains to sorting. Evidence is provided in Table II. 99% of farmers believe there is ability heterogeneity in the village. On average, they believe the best farmer in the village produces more than three times as much per acre as the worst, a very substantial difference. Further, 100% of participants believe total output in the village would increase if the best farmers cultivated a larger share of the village's land indicating both gains from matching, and that the current allocation of land does not maximize production.

Fact 3: Limited Agglomeration Gains due to Decreasing Returns at the Farm Level

Our sample also strongly believes that farmers face limits to the amount of land they can cultivate. Table III shows that only 60% believe they could farm more land than they have now. In addition, 99% believe that there is heterogeneity in these limits; some people could manage larger farms. On average, (assuming no hired labor) they estimate that the best farmer could manage nearly 5 acres while the worst could manage

[13] Although it doesn't appear to be the case in our setting, fragmentation can be used to mitigate risks when insurance markets are poor since weather shocks are often local. An advantage of the market design approaches that we propose is that trade is voluntary and farmers can use the tools as they see fit to adjust land holding in response to changes in auxiliary institutions. We chose an experiment where consolidation increased returns since it aligns with farmers' desires in our environment and is likely to be efficiency-enhancing in the long run once insurance markets become better developed.

less than 1 acre. Farmers do not typically believe that they are the best farmer. On average, they consider their own household's maximum capacity to be 2.5 acres, or about 1 acre per adult member.[14]

Fact 4: There Are Cultural Constraints to Trade, but Trade is Possible

Table IV confirms that a thin land market already exists. All farmers in the sample are aware of land trade occurring. Leases are more common than sales: in the last 12 months 17% of farmers leased in some land while only 8% bought.[15] The final panel shows that while there are only a small number of trades per year, these accumulate and 45% of land that is currently owned was purchased rather than inherited or gifted. We also see that farmers report making purposeful attempts to consolidate land through trade, with 24% reporting an attempted consolidation, but only 50% of these attempts are successful. Overall, the table provides direct evidence that trade is possible, although markets appear to be quite thin.[16]

However, there are cultural and institutional constraints that limit trade. Table V summarizes the evidence. From above, we know that the first-best allocation likely involves sorting the best farmers to the best land. However, the survey suggests that more than 60% of farmers think that this would be unfair or very unfair. This is a potential strong constraint to trade, and highlights the importance of appropriate compensatory transfers when implementing such an allocation. There is also strong agreement that families should not sell ancestral land (i.e. land that has been passed down in the family). However, participants are more equivocal when asked if a family should be free to trade land to improve their situation, with more than 50% suggesting this is okay. Thus there is not a strong general taboo against land trade, but some plots may not be available for trade.[17]

[14] These findings suggest limited gains from agglomeration but there may still be gains from better farmers cultivating more land and it might be efficient for some to leave farming altogether. As discussed below, our respondents are reluctant to exit farming and move to the city and there appears to be some stigma about farmers leaving the village. To avoid such sensitive issues, we did not explore this type of agglomeration in our experiments.

[15] There is asymmetry, with fewer people claiming they sold or rented out than those claiming they bought or rented in. This could reflect missing larger landlords in our sample, or plot fragmentation at the time of sale.

[16] We asked some additional questions that sought to establish whether complementary markets and institutions were functional (not shown in the table). Land registries and credit markets appear to be well functioning in the region. Most farmers know how to register land sales and there appears to be a formal way to transfer property rights. Credit markets also appear to function and 50% of farmers believe that they could borrow to purchase land. Over 70% of participants have participated in some kind of auction, suggesting familiarity with the kinds of solutions we are interested in, but these types of markets do not seem to exist for land trade.

[17] Some of the cultural constraints could potentially be alleviated by using long-term leases rather

Farming is an important part of our respondents' identity. They strongly agree that it is important to them to own land, they would like their children to remain farmers, and they are generally unwilling to migrate even if they could fetch a good price for their land. Such attitudes may be more permissive toward land reallocation between farmers than toward reallocation that moves some people off the land altogether.

Fact 5: Private Information but No (Perceived) Adverse Selection

The information structure of the trading environment is important for determining what frictions apply and what solutions may be appropriate. A particular concern is *asymmetric* information. If potential buyers believe that sellers have private information about land quality, the market may completely unravel as in Akerlof (1970)'s lemons problem. Strikingly, our respondents do not believe this is a concern. Table VI shows that 99% of farmers believe they know how to identify the best land in the village, and farmers have multiple strategies for evaluating land quality. Farmers also do not hold strong beliefs that people sell or rent poor quality land, and do not believe that it is hard to assess the quality of other farmers' land. Thus our participants do not believe there is an adverse selection problem.

Bargaining-based frictions as in Myerson and Satterthwaite (1983) arise from two-sided private information, meaning that farmers do not know one another's potential gains from a given transaction. Is that a reasonable assumption in our setting? Table II shows that 98% of respondents believe that everyone agrees on who the best farmers are, suggesting that at least some information about gains is not private. But there is substantial heterogeneity in the predicted productivity increase from sorting, implying uncertainty of the exact gains from a given transaction. In general, it seems implausible that farmers know all of their trading counterpart's outside options, investment plans, or the value they might assign to a given plot for idiosyncratic cultural and non-pecuniary reasons.

3 An Experimental Model of the Problem

We now describe a simplified experimental model of the land trade problem, which mirrors the facts set out in section 2. This has two purposes, first, it can help us to gain a theoretical understanding of why land trade is difficult, and second it forms the basis for our experiments, which help us to understand empirically whether land trade

than sales as the basis for the transaction. Such leases would allow individuals to sort while still ensuring a flow of payments to owners of high-quality land and their descendants. For our experiments, it was easiest to describe transactions as sales. Thus, we do not address this potentially important contracting issue.

is difficult, and how better market design can improve the situation. Further details about the setup can be found in Appendix A.

Figure I describes a land trading game consisting of a map of 72 plots. Land is divided into three regions corresponding to low, medium, and high *quality*, indicated by symbols (, ,). There are 18 farmers, each described by a different color and number, and each with an initial allocation of three plots. Colors indicate *farmer ability*, which is low (blue), medium (orange), or high (green). Participants in the experiment take on the role of a farmer, and begin the game with endowments of land and game currency. Through trade, they can earn money by improving the value of their endowments.

We model increasing returns at the plot level (Fact 1) through *adjacency bonuses*. Farmers earn an additional return when they own two or three plots in the same quality region that are adjacent to one another. We model farmer-farm complementarity (Fact 2) by setting the return to a given plot equal to the product of its quality and the owner's ability. Better farmers therefore earn higher returns from any given plot. Note that while we always refer to quality, we could equivalently think of the quality dimension as capturing different plot sizes, in which case better farmers produce more from a given *quantity* of land. We model decreasing returns at the farm level (Fact 3) in a simple way by imposing that each farmer can cultivate at most three plots. If they have more they only earn the return to their best three. The combined implication of these three features is that in an efficient allocation, all farmers should hold three consolidated plots, and farmer ability and land quality should be assortatively matched.

We capture cultural constraints (Fact 4) by assuming 18 plots on the map, represented by white space, are not available for trade at any price (for example, these could represent ancestral land). We also created a version of each map without non-tradable plots in order to better understand the impact of cultural constraints of this type (See Figure II). We will refer to maps with nontradable plots as "complex," and maps without nontradable plots as "simple."

We assume that all farmers know their own value of any plot or combination of plots, i.e. there is no adverse selection (Fact 5). However, these values are private to them and in our experiments participants are asked not to share this information with other participants.

Figure Ia shows an example initial allocation, featuring fragmentation and misallocation of land. Figure Ib shows an efficient allocation.

3.1 Why is Land Trade Hard?

We now discuss why Facts 1–5 imply three *frictions* that could impede efficient land trade.

Friction I: Thin Markets and Private Information

A thin market is one with only a small number of buyers and sellers willing to trade a particular good. In our setting, decentralized markets are likely to be thin due to increasing returns at the plot level (Fact 1) and farmer-farm complementarity (Fact 2). In our maps, the complimentary between farmer and land type implies that there are only ever 6 efficient buyers for each quality of plot. In a real world market heterogeneity is likely to be more continuous, but the matching requirement created by complementarity will still reduce thickness relative to a homogeneous goods market. Increasing returns at the plot level exacerbates the problem because a given plot is of relatively low value to those who do not already have nearby plots. For instance, if farmer 16 wishes to sell her plot of low-quality land, only farmers 3 and 18 would benefit from the increasing returns, meaning the market may only contain two interested buyers. Market thinness means that competitive prices may not emerge and transactions must instead be bargained over (Rustichini et al., 1994).

Thin markets combine with two-sided private information (Fact 5) to inhibit ex-post efficient trade. For a single buyer and a single seller, Myerson and Satterthwaite (1983) show that when valuations and costs are private and drawn from continuous distributions that overlap, it is impossible to construct an incentive compatible and individually rational mechanism that generates ex-post efficient trade, without generating a deficit. This result has been shown to extend to all problems where trade is one-to-one, and to a wide array of many-to-many allocation problems (Vickrey, 1961; Gresik and Satterthwaite, 1989; McAfee, 1992; Segal and Whinston, 2016; Delacrétaz et al., 2019). While it is difficult to directly tackle the private information problem through market design, a key implication of this discussion is that farmers would do better if it were possible to thicken markets.

The problem may be worsened by the presence of nontradable plots (due to cultural constraints, Fact 4). A benchmark for identifying the extent to which private information is likely to prevent trade is to calculate the expected deficit that would be generated when using the Vickrey-Clarke-Groves (VCG, Vickrey (1961); Clarke (1971); Groves (1973)) mechanism to reallocate units. The VCG mechanism is useful since it is the cost minimizing way to induce truthful reports in a large class of problems (Williams, 1999). In our setting, the deficit that results from the VCG mechanism being run on a complex map (with nontradable plots, Figure IIa) is always larger than its paired simple version (no nontradable plots, Figure IIb).[18]

Consolidation is easier than sorting. This literature gives us an important additional prediction: it is easier to realize gains from consolidation than from sorting. The consolidation problem has strong similarities to the case of partners who own an asset in common and wish to dissolve the partnership. This is an important exception to the general inefficiency of trade with two-sided private information, and Cramton et al. (1987) and Loertscher and Waser (2019) show it is possible to reach ex-post efficiency in a variety of partnership problems. Consider a situation where farmer 16 and farmer 6 have each sold one piece of land and are negotiating over who will hold the two plots in the north-east corner of the high-quality region. We can think of this as a partnership with a single asset (the consolidated plot) and an outside option to keep the original allocation. Intuitively, the adjacency bonus means that it is common knowledge that there are gains to trade, and so efficiency is easier to achieve.[19] In contrast, trades that aim to improve sorting, for example farmer 16 purchasing the disjointed plot owned by farmer 7, are closer to the original Myerson and Satterthwaite (1983) setting. For all of our analysis we show results for overall efficiency, but also decompose them into gains from consolidation and sorting.

Friction II: Exposure Risk

Exposure risk arises when at least one party stands to make a loss if a chain of trades (a sequence of purchases and sales) is not completed (Goeree and Lindsay, 2019). A leading example is hold-out, where a trader toward the end of a chain realizes they can capture most of the gains from the chain by holding out for higher price. Exposure can be thought of as a more general form of hold-out, including such strategic behavior but also any other reason the later trade may not take place, such as an exogenous financial shock to a buyer late in the chain.

Exposure risk arises in our setting due to increasing returns at the plot level, farmer-farm complementarity, and decreasing returns at the farm level (Facts 1–3). Decreasing returns at the farm level guarantee chains will form, because a farmer that buys land

[18] The intuition is as follows. In the VCG mechanism, each individual's compensation is related to the difference in surplus that other players receive when this individual participates in the mechanism, versus when the player opts out and retains his original land. In the simple map, the adjacency bonuses of others are almost never impacted by one individual being excluded from the mechanism, because they can always trade with someone else. This is not the case in the complex map because an individual's plot may act as a bridge between two components of the network and may be necessary to assign players to contiguous sets of land. Hence, in the complex maps many individuals must receive a high compensation, so we expect a large deficit.

[19] Further, the initial ownership is exactly split, which Cramton et al. (1987) show is a case where it is always possible to construct a mechanism that allows for efficient trade.

will also need to sell land, and vice versa. Increasing returns at the plot level mean that dividing up a previously-consolidated farm in order to relocate entails a risk of not being able to reconsolidate later (e.g., Farmer 17 in Figure I). The gain to the initial buyer may not be sufficient to cover the loss of an adjacency bonus, so one party must make a loss on the first trade. Alternatively, if farmer 5 attempts to purchase the low-quality plot owned by farmer 10, but fails to form a consolidated farm, that initial transaction might be unprofitable. Turning to the role of farmer-farm complementarity, consider the case where farmer 16 wishes to first sell her low-quality plot and then buy the plot owned by farmer 6 in the north-east corner of the high-quality region. Farmer 3 is a natural initial trade partner for farmer 16, but farmer 3 has lower productivity than farmer 16 and cannot fully compensate 16 for 16's initial loss of value. Thus, farmer 16 must take a loss from this first trade, in anticipation of a gain upon completion of the chain.

We conjecture that nontradable land (due to cultural constraints, Fact 4), increases exposure risk by reducing the number of potential adjacencies. A farmer who initiates a chain thus has a smaller set of potential trading partners later on, weakening their bargaining power.

Exposure risk could play out in two ways. First, it may lead farmers to avoid initiating chains, reducing the efficiency of decentralized markets. Alternatively, some farmers may decide to take on the risk by making a first trade, but end up holding inefficiently more land than they can cultivate. This second outcome is particularly important for market design considerations, both because it reduces efficiency, but also because it could amplify inequality among potentially vulnerable individuals, which would be ethically problematic. These observations motivate us to evaluate both efficiency and inequality in our empirical analysis and to further separate out "exposure losses" when decomposing gains from trade.

Exposure risk implies that farmers may prefer to make conditional offers, which could grow into long chains involving many farmers. To illustrate, in the example above, 16 would like to sell her low quality plot to farmer 3 *conditional* on being able to purchase a high quality plot from farmer 6. This is a chain of three farmers that that will move toward an optimal allocation so long as farmer 16 can be convinced that the conditionality will be enforced. Hence, exposure risk creates demand for conditionality.

Friction III: Coordination Frictions

Increasing returns at the plot level, farmer-farm complementarity, and decreasing returns at the farm level (Facts 1–3) combine to imply that reaching efficiency likely requires long chains of trade involving many farmers. In total, to get from the initial allocation in Figure Ia to the efficient allocation in Figure Ib requires 45 plots to change hands, and many of these will be conditional trades because of the exposure problem. As highlighted by Milgrom (2017), when many complex transactions are required to reach efficiency, even small transaction costs can make it difficult for decentralized markets to function well.

Because land is inherently immovable, initial allocations or trades can create "packing" problems where an individual who begins to assemble land in the wrong place can make it impossible for others to assemble land efficiently without additional trades occurring. These problems arise even in the simplest case but are exacerbated by the presence of nontradable plots (due to cultural constraints, Fact 4) because this further reduces the number of potential efficient assignments.[20] We conjecture that this contributes to low efficiency, and may make it harder for individuals to consolidate land once others have begun to do so. We explore these issues with our complex and simple map treatments in experiment 1.

Credit Constraints Worsen the Problem

Although our survey evidence suggested that farmers have some access to credit, poorly functioning credit markets are a common feature of low-income economies (e.g., Banerjee, 2003). They can play a particularly important role in land markets, where large cash payments are required upfront and the benefits accrue mostly in the future. If our farmer cannot borrow, the only way for her to raise funds for a land purchase may be to first sell some land. But then the problem simply passes along the chain: her buyer must also raise funds and may also be constrained, and so on. This exacerbates the need for chains, potentially worsening exposure risk and coordination frictions. We explore the issue of credit constraints in experiment 2 where we vary participants' initial cash balances.

4 Experiments: Implementation and Analysis

We conducted two experiments based on the game described in Section 3. Experiment 1 was conducted in Masaka district, Uganda, in Fall 2019. It was designed to measure the efficiency of trade in a decentralized setting, the role of nontradable plots, and the effectiveness of a simple market centralization intervention. Experiment 2 was conducted in Kiambu country, Kenya, in Summer 2016. It studies the effectiveness of

[20] For example, it is not possible to consolidate the high-quality region without trading with farmer 17. If farmer 10 buys farmer 4's neighboring medium-quality plot he blocks full consolidation of the medium-quality region, and so on.

two increasingly tailored package exchanges, relative to a benchmark centralized land exchange. We also investigated the role of credit constraints. In our analysis we present experiment 1 first because although chronologically it was conducted later, it provides clear conceptual motivation for experiment 2.[21]

Our analysis of experiment 1 follows a pre-analysis plan posted to the AEA trial registry. We follow the plan closely, and highlight deviations where they occur.[22] In keeping with the approach advocated by Athey and Imbens (2017), our pre-specified regression specifications always control for the full set of strata fixed effects within which our treatment variation occurs; we specify these in the table notes to all regressions. Experiment 2 did not have a pre-analysis plan but we implement the same regression specification, decomposition, and analysis strategy, to ensure comparability. In both experiments we used blocked randomization of treatments and, within treatments, the different maps that formed the basis of the games.

Our main specifications regress efficiency, measured as the fraction of gains from trade realized by participants, on treatment dummies. We also decompose efficiency into three components: consolidation, sorting, and (avoided) losses due to exposure. Our tables report the decomposition measured in efficiency units (i.e., the three components add up to total efficiency), and also express consolidation and sorting as a percentage of the values they would take at an efficient allocation. For the decomposed analyses we report q-values (Anderson, 2008) that adjust for multiple testing across the three components.[23]

5 Experiment 1: Decentralized and Centralized Trade

Our first experiment uses the game described above and depicted in Figure I. Appendix A provides full implementation details. Here, we summarize the design, then turn to results.

[21]One of the key findings from experiment 2 was high efficiency in the benchmark treatment, which prompted our investigation of efficiency in decentralized trade in experiment 1. The reason for conducting the experiments in different locations was purely pragmatic and based on the capacity of our implementing partners to conduct the complex experimental designs we use.

[22]The plan can be found at `https://doi.org/10.1257/rct.4581`. One thing to note is that we pre-specified our analysis of simple versus complex maps (presented in Section 7) as primary hypotheses, and our analysis of the centralization intervention in Section 5 as secondary. We subsequently combined both experiments into this single paper, at which point it was narratively more logical to describe the intervention effects first. Appendix B contains additional exploratory analyses from the pre-plan.

[23]Our pre-analysis plan stated we would adjust over efficiency and consolidation. Ex post we concluded it made more sense to treat the efficiency analysis as a standalone hypothesis test, and then adjust for multiple testing when we decompose into its three components.

The distinguishing feature of this experiment is that in each week of play we gave participants seven days to trade amongst themselves within the community, without input from the research team. We therefore interpret outcomes as reflecting what is achievable under *decentralized* trade. At the end of the second week, we surprised participants with a simple *centralization* treatment; giving them additional time to trade with everyone in the same room. We argue that this is informative about the benefits of simple market design interventions that centralize the market, without imposing additional structure in the form of trading rules.

In each of 68 participating villages we recruited a group of 18 farmers to play our experimental games over a two week period, plus four reserve farmers who would step in if a participant dropped out. Recruitment and sample characteristics are as described in Section 2. We conducted three meetings, with seven days in between each.

Meeting 1 (start of week 1): we introduced the study, and played two training games. We publicly explained the main games, privately distributed participants' materials, and dismissed them, giving them seven days to trade.

Meeting 2 (start of week 2): we collected final endowments from week 1, and calculated earnings. We conducted the survey from Section 2. We privately distributed materials for a new week of play, and dismissed participants for seven more days.

Meeting 3 (final meeting): we collected endowments from week 2, then surprised participants with the opportunity to continue trading for one more hour. We measured their final endowments, conducted an exit survey, and dismissed them.

We paid participants for every game they played, plus a show-up fee for each meeting. Games involved beginning with an allocation of land and cash, plus an amount of debt. Payments were always based on the value of net assets at the end of trade, meaning the value of their landholdings, plus their final cash balance, minus their debt. Debt equalized initial net asset values, and sharpened incentives because it meant that a larger share of participants' earnings were derived from gains from trade, rather than just the value of their initial assets.

5.1 Training games

We use two training games to introduce the idea of trading fictitious land titles, demonstrate earnings calculations, and provide a benchmark against which to compare the outcomes of the main games. They were played by the main participants and reserves in each village.

Training Game 1 followed the spirit of Chamberlin (1948) and introduced the concepts of heterogeneous ability and the potential value of trade. Participants were assigned a type corresponding to their payoff from owning a single land title; holding more than one land title yielded no additional payoff. Types were uniformly spaced and land titles were initially assigned to odd-numbered types. Everyone began with the same cash balance, and debt such that initial net assets were worth 4,000 UGX ($3.05 PPP) per player. Trade was free-form but, unlike Chamberlin (1948), resale was allowed so players could buy and sell as many times as they liked. Play continued until nobody wished to trade any more. Efficiency was reached if all titles were owned by the highest-type players. Average earnings at an efficient allocation were worth 7,000 UGX ($5.34) per player, a 75% gain.

Training Game 2 introduced the concepts of heterogeneous land, the possibility of holding multiple plots, and scale limits. Participants were assigned to one of three ability types, and there were three land quality types. Players earned payoffs on their best three plots, capturing the idea of decreasing returns at the farm level. Everyone began with one title, a cash balance, and debt, with initial net assets equal to 4,000 UGX ($3.05 PPP). Efficiency was reached if all titles were owned by high types, with at most three plots each. Mean net assets at an efficient allocation were worth 6,727 UGX ($5.13 PPP), a 68% gain.

5.2 Main games

The main games used the game introduced in Section 3. For each round of play we randomly assigned each village to a "map," which consisted of an initial allocation of three land titles for each player. We used a set of randomly-generated maps where the gains from trade were divided approximately 50–50 between consolidation and sorting. In addition to their three titles, each player was given an (identical) cash balance in the form of printed paper bills, and debt such that their initial net assets were worth 14,000 UGX ($10.67 PPP).[24] Mean net assets at an efficient allocation were worth 21,940 UGX ($16.73 PPP), a 57% gain.

We explained the rules of the game in public, but told farmers their type privately, and instructed not to share this information with others. We gave them a card that showed them the value for each type of land and the adjacency bonus they would receive from farming two and three contiguous plots. We gave each participant a map of all potential plots, marking their initial allocation as well as the (initial) owners of all other plots. Figure A1 shows an example of handout materials.

Most players knew one another but we added measures to ensure they could find each other. Each participant had a player ID (from 1–18, uncorrelated with type), and was given a sheet on which they could record the other players' identities. The village

[24] In principle participants could exchange real money or non-game goods for land titles, which would affect our inequality measure but not efficiency. Nobody reported doing so.

chief (LC1 chairperson) was also given a sheet with all names and IDs; participants could consult the chief if needed. Since most people were socially connected to the chief in some way, it was natural for them to play this role.

5.3 Outcome Measures

As noted in Section 3.1, the theoretical literature on two-sided trade suggests that consolidating land may be easier than sorting. Further, the exposure problem may influence both efficiency and inequality. Our outcome measures reflect both these insights. We compute outcomes separately for each village and trading week. In the case of week 2 we compute outcomes both before and after the centralization treatment.[25]

Efficiency. Efficiency is the fraction of possible gains from trade realized. Define *surplus* as the sum of the land values and adjacency bonuses of land owners, then:

$$\text{Efficiency} = \frac{\text{Final surplus} - \text{Initial surplus}}{\text{First-best surplus} - \text{Initial surplus}}.$$

Efficiency equals 0 if no trade occurs, and 1 if a first-best final allocation is reached. Negative realizations are possible if trade decreases total surplus. Note that the final allocation of game currency does not affect efficiency (but will matter for our analysis of inequality).

Efficiency has three components: Efficiency = Exposure + Consolidation + Sorting.

Exposure. Recall that we have defined exposure as a situation where one party stands to make a loss if a chain of trades is not completed. In our setting, individuals who end up with too much or too little land could unambiguously improve their outcome with one additional trade and thus participants ending the game with too much or too little land is a direct loss attributable to exposure.[26] To compute these direct exposure losses, we identified all plots that were uncultivated in the final allocation, and reassigned these plots to individuals who owned less than three pieces of land, in such a way as to maximize total surplus. We define *exposure value* and *exposure bonuses* as the additional land value and adjacency bonuses generated by this hypothetical reassignment. The

[25] We pre-specified the construction of these outcome variables for this experiment. Appendix B shows that our results are similar when we calculate consolidation and sorting gains without first adjusting for exposure, or when focusing on the high-quality region of the map.

[26] The sharp penalty for owning more than three plots was included in the game to capture decreasing returns without avoid overwhelming participants with parameters and difficult calculations. The downside is that it can amplify losses due to exposure, relative to a smoother decreasing-returns production function.

(direct) loss due to exposure is calculated as:

$$\text{Exposure} = -\frac{\text{Exposure value} + \text{Exposure bonuses}}{\text{First-best surplus} - \text{Initial surplus}}.$$

Note that exposure is weakly negative and normalized by the same denominator as efficiency. Thus, it can be interpreted as the percentage of the overall potential gains from trade lost due to partially completed sets of trades.

Consolidation. Our measure of gains from consolidation is:

$$\text{Consolidation} = \frac{(\text{Final bonuses} + \text{Exposure bonuses}) - \text{Initial bonuses}}{\text{First-best surplus} - \text{Initial surplus}},$$

where *final bonuses* is the sum of the value of landowners' adjacency bonuse values in the final allocation and *exposure bonuses* is defined above. Exposure bonuses are added back so as to decompose gains from consolidation from gains from (avoided) exposure losses.

Sorting. Our measure of the gains achieved through sorting is:

$$\text{Sorting} = \frac{(\text{Final land value} + \text{Exposure value}) - \text{Initial land value}}{\text{First-best surplus} - \text{Initial surplus}},$$

where *final land value* is the sum of landowners' plot values and *exposure value* is the additional values that would be attained if exposed plots were reassigned.

Inequality We use a version of the Atkinson index to measure to inequality, with the auxiliary assumption of log utility over income. Letting $\{y_1, y_2, \ldots, y_n\}$ represent the earnings of each player and μ representing the mean of these values, the Atkinson index is:

$$I^A(f) = 1 - \exp\left[\sum_i \frac{(\ln y_i - \ln \mu)}{n}\right]. \tag{1}$$

A nice feature of this specification is that the Atkinson index represents the proportion of lost social welfare that is due to inequality. For example, if $I^A = 0.3$, then it would be possible to reach the same social welfare with 70% of the income, but with equally distributed payoffs. We compute the Atkinson index of players' final assets net of debt. As we set debt values to equalize initial net assets, the initial allocation features perfect equality. Ex post, around 7% of participants had negative net assets after trade, likely driven by exposure losses. A consequence is that $\ln y_i$ is not defined for these participants. Our analysis explores a range of (non-pre-specified) adjustments to deal with this unanticipated issue.

5.4 Treatment variations

Centralization intervention. After the second week of the game was completed and we had recorded the players' post-trade holdings of land and cash, we surprised them with an additional hour to continue trading, this time in a centralized location.[27] We think of this as a very simple market centralization intervention. We conjectured that centralization would have two effects. First, the task of searching for chains of trade is much less cumbersome, reducing coordination frictions and potentially increasing market thickness. Second, the presence of everyone in the same room may make it easier for people to commit to conditional trades, as any reneging will be observed by a large part of the community. Being able to enforce conditionality will tend to reduce exposure risk, increasing efficiency and improving inequality.

Nontradable plots. Our theoretical discussion suggests that cultural constraints on trade, leading to nontradable plots on the map, will reduce trading efficiency and/or increase inequality of outcomes. As explained in Section 3 we constructed each map with a complex and simple form where simple maps had no nontradable plots but identical initial payoffs and potential gains from trade. Each village played a complex map in week 1 and a simple map in week 2, or vice versa. We use variation in map complexity to explore the importance of cultural constraints, which the theory above suggests will reduce trading efficiency, or create inequality. We further conjectured that the centralization intervention would be more effective in the complex treatment, because the problems of thin markets, exposure risk, and coordination frictions are more severe in the complex treatment.

In this section of the paper we focus on the effects of the centralization intervention, and discuss the effects of nontradable plots in section 7.

5.5 Results

We begin by benchmarking efficiency in the main experiment against training games 1 and 2. This helps us to understand the relative difficulty of the land trade problem compared to more standard lab trading games. We then decompose efficiency into

[27] Since this treatment was applied in all villages at the end of week 2, an alternative explanation is that its effects come simply from more time to trade. We show below that this is unlikely to explain our findings.

consolidation gains, sorting gains, and exposure losses. As discussed in section 3.1, we might expect the consolidation problem to be easier to solve than the sorting problem.

Our first result establishes that the land trade problem is hard relative to the training games, that sorting is harder than consolidation, and that exposure risk is a significant problem.

Result 1 *Relative to the training games, efficiency in the main experiment is low. Farmers are able to capture some of the gains from consolidation but capture very little of the potential gains from sorting. Further, many farmers end up with sub-optimal amounts of land leading to large exposure losses.*

Figure III shows that efficiency in the two training games is around 90%. The high level of efficiency demonstrates that farmers understand how to trade induced-value land titles and are able to achieve high efficiency in regular trading environments.

In contrast, efficiency is below 40% in weeks 1 and 2 of the land trade game (Table VII, Panel A). This low efficiency is not due to inactivity. Around half of all plots changed hands in each week. In our exit survey around 95% of farmers reported trying to buy at least 1 plot over the course of the games, and 87% reported that they successfully traded at least one plot.[28]

Panel A of Table VII decomposes efficiency into consolidation, sorting, and exposure losses, averaged over weeks 1 and 2 of the game. Farmers perform substantially better on the consolidation than the sorting dimension, realizing 61% and 25% of the potential gains from each, respectively, suggesting that efficient consolidation may be easier than sorting. However, exposure losses are large, equalling roughly 20% of total potential gains from trade.

Next, we study how our centralization intervention affects efficiency. Based on the theoretical discussion we would predict that centralized trading reduces exposure risk since trades can be conducted synchronously.

Result 2 *The centralization intervention delivers higher efficiency and lower inequality. Efficiency gains are primarily due to the unwinding of exposure losses and improvements in consolidation. There are essentially no additional sorting gains.*

Panel B of Table VII shows that efficiency improves markedly during the short period of centralized trade at the end of week 2. Efficiency at the start of the the

[28]While we might have had anticipated learning effects, efficiency actually falls between week 1 and 2, driven primarily by larger exposure losses in week 2 (Table B3). We conjecture that, having become more comfortable with the game after week 1, players attempted more trade, moved into exposed allocations, and found themselves unable to unwind these positions.

centralized trading period was around 12%. After around 1 hour of centralized trade it had increased to 47%.

These gains come predominantly from participants unwinding their exposure losses, which fall from 32% to just 3% over the course of centralized trade. They also do substantially more consolidation. From a baseline level equal to 57% of all gains from consolidation, they increase to 70%. In contrast, the effect on sorting is a precise zero.

Panel A of Table VIII shows that the centralization intervention also reduced inequality (and this conclusion is robust to four different approaches to adjusting for negative net assets). The reduction in inequality is likely driven by the mitigation of exposure losses, which particularly hurt farmers left with too much land.

The main concern with our interpretation of the effects of centralization is that this intervention also increased the time available to trade. To rule out this explanation, we exploit plausibly exogenous variation in the morning/afternoon scheduling of village meetings that generated up to 8 hours' variation in the time each village had to trade prior to the centralization intervention. If the gains we see come from increased time we would predict that (i) villages that had more time during week 2 would do better prior to the centralization intervention, and (ii) the impact of the centralization intervention would be smaller for villages that had more time before the intervention (because the remaining realizable gains are smaller). Figure B2 graphs efficiency against the number of hours available to trade during week 2. Neither prediction holds in the data. For prediction (i) we find the relationship between efficiency and time to trade is essentially flat and comes nowhere close to the discrete jump up in efficiency during the period of centralized trade. For prediction (ii) we see that if anything the returns to the centralization intervention were larger for villages that had already had more time to trade.

Taken together, the results from our first experiment suggest that land trade is hard and that farmers often trade to inefficient allocations when trade is decentralized. Realizing gains from consolidation appears substantially easier than sorting. Bringing farmers together helps to unwind exposure losses and unlocks further consolidation gains, but has no effect on sorting. This suggest that there may be additional gains from more tailored market design interventions. We explore this possibility in our second experiment.

6 Experiment 2: Tailored Package Exchanges

Our first experiment demonstrated that the land trade problem is hard and that a simple centralization intervention can generate considerable efficiency gains on some, but not all dimensions. In particular, farmers find it hard to realize gains from sorting, and this is not improved by market centralization.

Our second experiment uses a simplified version of the same game to study more tailored market design interventions. As a benchmark we set up a centralized computer-based land exchange in which farmers can trade individual plots in a Continuous Double Auction (CDA) mechanism, which we call “Package-1.” Our more-tailored interventions add the possibility of package bids, allowing farmers to condition transactions on one another. For example in “Package-2” they can bid to sell a plot simultaneously with buying another, while “Package-4” allows for still-larger combinations. The exchange platform is responsible for finding feasible trades and setting prices. The theoretical discussion in Section 3.1 argued that more tailored interventions could improve the efficiency of trade by thickening the market for any individual plot; finding chains; and enforcing conditionality. But whether they do depends on farmers’ comprehension and behavior, and is an empirical question.

6.1 Design

We used a simplified version of the land trade game, with 6 farmers and 12 plots of land. Figure IV gives an example, and Figure C1 shows all eight maps we used. Each farmer was initially allocated two plots. We imposed that they could cultivate at most two, if they had more they received the return from their best two. Otherwise the game structure was essentially the same as experiment 1. See Appendix C for all details.[29]

We set up two labs in a town in southern Kiambu County, Kenya, and recruited land-owning farmers from a census of the local area. Table D1 provides summary statistics and compares our sample characteristics to national averages from the Kenya DHS (Kenya National Bureau of Statistics et al., 2015).

We conducted 48 experimental sessions, each consisting of 6 farmers and 8 auction rounds. At the beginning of each session, farmers were randomly assigned a computer and an enumerator or “bid assistant” whose role was to train them on the game and then assist with the computer interface.[30] An additional assistant was available in each session to pass messages between farmers. Assistants were explicitly told not to

[29]The main remaining differences were that (1) we did not target a 50-50 split of gains from consolidation and sorting (the realized split is 73/27), (2) we did not use debt, and (3) adjacency bonuses were calculated as 40% of the plot’s value – meaning higher-quality plots earned larger bonuses, whereas in experiment 1 each participant had a single fixed adjacency bonus irrespective of land quality.

[30]Bid assistants are a common feature of real-life combinatorial auctions when the target population may have difficulty with the interface, and have been used, for instance, in the auction of slot machines and taxi medallions in Australia. In our experiment the assistants read the instructions in the participant’s preferred language, answered questions, helped with calculations, and entered bids into the system. To reduce the influence that an individual bid assistant might have on the experiment, we randomized bid assistants across participants and treatments.

suggest or actively organize trades. We allow for oral communication in this experiment since we are interested in developing exchanges that can be used in conjunction with current institutions. Given that communication is a feature in our target environment we consider it an important part of our design.

After the instructions and a 15 minute practice auction, we conducted eight 10-minute auction rounds, each using a different initial allocation, in blocked random order. Participants knew their types and that there were three ability types of farmers, but not the other players' payoffs (they were assigned to a new ability type after the 4th auction). As discussed in more detail in the appendix, participants could see their current allocation and bids on their screen, and a centralized screen showed the map with labels for each current plot owner and icons indicating plots with active bids. Figure C2 shows the interface.

Participants were informed of their outcome at the end of each auction and paid for all eight auctions. Mean initial assets were worth 47 KES per auction, ($1.20 PPP) while efficient play would result in average earnings of 55 KES ($1.40 PPP), a 17% gain. Proportional gains were smaller than in experiment 1 because we did not use initial debt. The average participant earned 418 KES ($10.61 PPP, around 1.5 days' wages).

Our outcome measures of efficiency, sorting, consolidation, and inequality are identical to experiment 1, except that we have to slightly adjust how we compute sorting and consolidation because in this experiment the adjacency bonuses scale with land quality.[31]

6.2 Treatment variation

We implemented a three-by-two design where we varied the trading mechanism between sessions and varied initial cash balances between auctions within a session. Our main analysis centers on the trading mechanisms, and we discuss effects of the cash treatment in section 7.

[31] To do this, we compute consolidation gains as if the farmer's land quality stayed constant between their initial and final allocations, and then attribute the rest of their gains to sorting. Formally, let y_i denote the value associated with farmer i's two best plots (ignoring consolidation bonuses), and let $c_i \in \{0,1\}$ indicate whether these plots are fragmented ($c_i = 0$) or consolidated ($c_i = 1$). It follows that the total profit on farmer i's land is $s_i := (1 + 0.4c_i)y_i$. After some algebra, the change in surplus from a farmer's initial allocation to their final allocation can be rewritten as:

$$s_i^{final} - s_i^{initial} = \underbrace{0.4\left[c_i^{final} - c_i^{initial}\right]y_i^{initial}}_{\text{Consolidation}} + \underbrace{\left(1 + 0.4c_i^{final}\right)\left[y_i^{final} - y_i^{initial}\right]}_{Sorting}.$$

Trading Mechanisms. We consider three trading mechanisms based on the package market of Goeree and Lindsay (2019): the benchmark treatment, "Package-1," permits bids to buy or sell one plot at a time. The second treatment, "Package-2," adds the possibility of bids to buy one plot conditional on selling a second one. The third treatment, "Package-4," allows packages consisting of up to two buy orders and up to two sell orders. In our game, two buys and two sells would be sufficient for every participant to move to an efficient allocation.

A package bid specifies which plots are to be traded (e.g., buy plot 1 and sell plot 4) and a maximum willingness to pay or minimum willingness to accept. Bids are submitted sequentially and the computer searches for the existence of a set of bids where (i) supply equals or exceeds demand for all plots; (ii) only a single bid is used for each farmer; and (iii) there is a non-negative surplus of cash (i.e., total willingness to pay is weakly positive). If more than one set of bids satisfies the criteria, the computer triggers the set with the largest cash surplus. Plots that were offered but not purchased stay with their original owners, and all other plots are transferred according to the winning orders. Prices are set by dividing the surplus among the winning farmers as equally as possible subject to revealed-preference constraints generated by the bids of non-trading bidders.[32]

Bids are exclusive-OR (XOR), meaning that only one bid from a farmer could be used in any given transaction. When a farmer makes a trade, their other bids become inactive. They then have the option to reactivate any inactive bid if they like. This ensures that farmers do not accidentally buy or sell too much land, or consolidate land and then break up land, in the same cycle of transactions or in quick succession.

We describe the algorithms used to trigger trades (the winner determination problem) and to allocate surplus formally in Appendix C. Here we note that our mechanism is a near real-time auction based on Goeree and Lindsay (2019) modified to impose XOR bidding. Bids were entered by the bidding assistants through the computer interface described in the Appendix.

Credit Constraints. Section 3.1 discussed how credit constraints can exacerbate exposure risk: a chain must form when a buyer does not have enough liquid assets to compensate the seller. To mimic the effect of credit constraints we varied participants' initial cash balances across auctions. In half of all auctions they began with enough

[32] These require that a non-trading party would not prefer to be part of the transaction given their expressed bids and the realized prices. See Kwasnica et al. (2005) for a broader discussion of revealed-preference constraints. We explain our surplus division rule to participants using the logic of a farmer who has buy offers from either one or two farmers. If there is only one buyer, we split the surplus evenly between the two farmers. If there are two buyers, the buying price must exceed the bid of the non buyer.

money to induce any farmer to sell any single plot. In the other half, they were given only one third of this amount.[33]

6.3 Data Issues

Due to the complexity of our experimental design we encountered two implementation challenges. First, our lead enumerators raised concerns that the other enumerators did not initially fully understand the rules of the experiment (we gave them three days of training including practice sessions, in retrospect we should have had more). As the enumerators were responsible for translating the instructions and teaching farmers, it is likely that farmers also did not fully understand the mechanisms in the early sessions. Sessions were block-randomized to treatment in blocks of six sessions, and in the data we observe substantially lower efficiency in the first assignment block (for instance, 63% of auctions in which efficiency was negative occurred during this first block), plus a general tendency for efficiency to be higher in later blocks. We always control for block fixed effects since treatment was stratified at this level, and in addition our preferred specification drops the first block due to these comprehension issues. Appendix Tables D2 and D3 report results for the full sample. We find qualitatively similar results, but the treatment effect on overall efficiency is somewhat weaker.

Second, we lost some data: one session due to accidental reformatting of our server computers, one session where the wrong treatment was used, and two auctions where the wrong map configuration was used. In total our main analysis dataset consists of 40 sessions, 318 auctions, 240 farmers, and 1908 farmer-auction observations.

6.4 Results

What should we expect to be the effects of our more tailored trading mechanisms? First, our package mechanisms, especially Package-4, have the potential to thicken markets. Under Package-4, a bidder can easily offer to exchange any two plots for any other pair in a different location, independent of where they started. In principle this should make it easier for potential buyers to compete for different pairs of plots. Second, they reduce exposure risk by enforcing chains: a farmer can guarantee that condition every purchase on a sale and vice-versa. Third, packages should reduce coordination frictions, because farmers only need to enter their bids (as many as they like) and let the algorithm search for the necessary chains.

[33] This should be enough to buy an unconsolidated low-quality plot from a high-ability farmer, a medium- or low-quality plot from a medium-ability farmer, or a high-, medium-, or low-quality plot from a low-ability farmer.

Thus, in principle we should expect higher efficiency in the more tailored package mechanisms. Moreover, we would predict that Package-2 reduces the complexity associated with consolidation, while Package-4, by facilitating wholesale relocations, can unlock sorting gains.

However there are good reasons to be concerned that these theoretical gains will not be realized in practice. It is very unlikely that any of our participants had ever participated in a computer-based auction before, and this unfamiliarity could lead to them making mistakes or not trading at all. The richer mechanisms may exacerbate these effects. We will particularly struggle to find chains if some participants focus on bidding on packages while others only bid on single plots. Finally, richer mechanisms may enable sophisticated participants to profit at the expense of the less-sophisticated, potentially exacerbating inequality.

Our first result paints a strongly optimistic picture for the potential of mechanisms ours to be effective in solving the land trade problem.

Result 3 *Average efficiency is 70 percent or higher in all three mechanisms. Package-4 achieves 7 percentage points higher efficiency than Package-1, and does not exacerbate inequality.*

Column (1) in Panel A of Table IX regresses overall efficiency on our package treatments. Average efficiency is high under all three mechanisms, so our concerns about the platform overwhelming our participants seems unfounded. From a base of 70% efficiency in the Package-1 treatment, Package-4 increases efficiency, by 7 percentage points, or 10%. This difference is significant. Efficiency is 3 percentage points (5%) higher in Package-2 than Package-1, but the difference is not significant.

Table X shows that the efficiency gains from more tailored designs do not come at the cost of higher inequality. We split the analysis by the amount of cash available because our inequality measure is not invariant to the initial asset level. We find no evidence that inequality was increased by our package mechanisms, and some evidence that the package mechanisms reduce inequality when cash balances are low. It could be that when buyers are credit constrained they can only offer low prices, so the distribution of surplus becomes more unequal. Packages ameliorate this effect by allowing buyers to compensate their sellers with land instead of cash.

Farmers continue to find the sorting problem harder than the consolidation problem. In the baseline Package-1 treatment they realize 86% of the potential gains from consolidation. The high consolidation rate is in line with our findings from experiment 1, especially in the centralized market. This is consistent with our theoretical observation that consolidation is more like a partnership problem and so less subject to information problems.

In contrast, participants only achieve 44% of the potential gains from sorting in Package-1. However, as predicted, our Package-4 mechanism successfully unlocks significant additional sorting gains.

Result 4 *Relative to Package-1, the Package-4 treatment unlocks an additional 12 percent of the potential gains from sorting. In contrast, Package-2 unlocks just 3 percent of the potential gains, and this increase is not statistically significant.*

Columns (2) and (3) in Panel A of Table IX report the impact of our treatments on consolidation and sorting. We see some improvements in consolidation but these are not significant. The absolute gains on the sorting dimension are larger, and substantially larger when expressed as a percentage of total potential sorting gains.

Finally, there are low levels of exposure losses in this experiment when compared to experiment 1, and no significant difference across treatments. This is similar to what we saw in the centralized market in experiment 1, where farmers effectively eliminated exposure losses.

7 Additional Results

In this section we discuss a sequence of additional results that shed further light on the frictions that constrain optimal trade, and the strategies farmers use to work around them.

Cultural constraints and nontradable plots. Experiment 1 was designed to test how nontradable plots, by complicating the problem of packing consolidated farms onto the map, would affect efficiency and the distribution of outcomes. This is important to understand because cultural considerations (Fact 4) suggest it is unlikely that all plots will be available to trade. As explained in Section 3.1, the presence of nontradable plots has the potential to exacerbate information and coordination problems and increase exposure risk.

Each village played the game twice, once with a "complex" map (as in Figure IIa and once with a "simple" map (Figure IIb), in random order. We exploit this within-village variation to estimate the effect of nontradable plots.[34]

Table VII Panel C provides suggestive evidence that simpler maps yielded higher

[34] As explained in Section 3 we had eight simple maps each constructed from one of our eight complex maps, with an identical payoff structure. Each village played one simple map and one *different* complex map.

efficiency, by 4.4 percentage points, but this is not statistically significant.[35] This is made up of a (marginally significant) 2 percentage point improvement in consolidation, a similar decline in sorting, and a 4 percentage point decrease in exposure losses. None of these components is significant after adjusting for multiple testing.

Map complexity seems to matter more for inequality. Table VIII Panel B shows that inequality of outcomes was around 15–30% lower on simple maps (depending on how we adjust for negative net asset positions). The difference is significant in three out of our four specifications. This may be driven by the reduction in exposure losses that we saw in Table VII.

We do not find that the centralization intervention helped more on the complex maps. Table VII Panel D shows that if anything it caused larger efficiency gains on simple maps (though all point estimates are small, nonsignificant, and swamped by the overall gains from centralization). For inequality, our estimates in Table VIII Panel C point to larger gains from centralization on complex maps, but again the differences are relatively small and not significant.

Overall, the impact of eliminating nontradable plots in our game was relatively modest, not contributing meaningfully to low overall efficiency, but potentially relevant for inequality.

Other measures of map complexity. In Experiment 1, our complex maps varied in how many welfare-equivalent efficient allocations existed, averaging between 1.67–5.33 ways to assign six consolidated blocks in a given quality region (one assignment is a unique combination of L- and I-shaped consolidated three-plot units). In the simple maps there are 134 such assignments. Figure B1 graphs efficiency, and its decomposition, against the number of potential solutions, within the complex treatment, based on the conjecture that maps with more potential solutions are less complex. Consistent with our finding that the main simple/complex map treatment did not have large effects, we do not see any clear relationship in these graphs.

Experiment 2 used eight hand-generated maps which we ordered according to our own judgment of complexity (see Appendix C and Figure C1). Figure D1 plots efficiency against this ordering. In general we find a decreasing but non-monotone relationship, with near-perfect efficiency on the easiest map (which could be solved by each player swapping one plot with their neighbor). The Package-4 treatment appears to have had its largest impacts on maps we judged to be more complex.

[35]We pre-specified a one-sided test; the one-sided p-value is 0.08 which we take as very weak evidence.

Credit constraints. In Experiment 2 we varied initial cash balances across auction rounds, to mimic the effects of missing credit markets, potentially worsening the exposure problem. Table IX Panel B shows that this had no discernible effect on efficiency or its subcomponents, and we see no significant interactions with the different package mechanisms. The main effects of the package mechanisms remain very similar. It seems the low cash treatment was not sufficiently severe to really constrain trade.

However, Table X shows that low cash appears to have been important for inequality, and interacts with the package mechanisms. It is hard to directly compare the *levels* of inequality across low and high cash because the Atkinson index is sensitive to total assets, but we see proportionally larger and statistically stronger reductions in inequality under Package-2 and Package-4 when cash balances are low, and significant interactions between the low cash and package treatments in the pooled specification (column (3)). This might reflect, as discussed in section 3.1, that our package mechanisms relax the effects of low cash balances on exposure risk. In the Package-2 and Package-4 mechanisms, plots can be swapped, facilitating transactions that require only small monetary transfers. By contrast, in the Package-1 mechanism, farmers are constrained to buying and selling land for cash, which could lead to sellers being unable to fully capture the value of their land from cash-constrained buyers.

Coordination frictions and verbal bargaining. Both experiments allowed participants to bargain verbally over trades, either directly or through a mediator, since this is likely to be a natural feature of any solution implementable in the field. But verbally coordinating on chains of trade may be very difficult, whereas package bids allow participants to express preferences to the system and let the algorithm search.

The transactions data in Experiment 2 provide indirect evidence for this conjecture. We see many transactions with zero cash surplus, meaning that total willingness to pay and willingness to accept were identical (e.g., farmer 1 demands 300 and farmer 2 offers exactly 300). We assume that these reflect "brokered" trades that were verbally agreed before being entered into the system. Figure D2 shows that 39% of transactions in the Package-1 treatment are brokered, but only 21% of in Package-2 and 16% in Package-4. All these differences are significant in a regression following our main analysis specification (p-value < .001 for all comparisons).

Spontaneous centralization. Besides addressing our three frictions, an additional benefit of market design may be to help *coordinate on rules.* In general, there are many potential land trading "rules" that farmers might need to agree on. How does the community feel about people who make conditional promises and then renege? Should trade be bilateral or organized in groups or by brokers? When should trade take place? Settling on such rules may add an additional level of complexity to decentralized exchange.

For example, if centralization is so effective, why didn't communities in Experiment 1 organize their own periods of centralized trade? They had good incentives to do so: the average participant's earnings from the experiment increased by 16% or 2,450 UGX during the period of centralized trade. In fact, participants did try to: 89% said that they gathered in groups during the weeks of trade. However, we see little relationship between the intensity of "endogenous centralization" and outcomes. Table B2 divides the sample into "endogeneous centralization" villages where every participant said they gathered in groups (62% of villages), versus the remainder where not everybody said so. We see no differences in efficiency between these two groups. Endogenous centralization is associated with improvements in consolidation and sorting, but more exposure losses. We conjecture that even when everybody is trying to centralize the market, they were not able to all coordinate at the same time, which led to more chains being formed that were ultimately not completed. This suggests there is value to a external market designer "exogenously" defining where and when the market will be centralized.

8 Conclusion

Fragmented, misallocated plots are a hallmark of the agricultural sector in less-developed countries, and there is evidence of high potential returns to land consolidation and reallocation. To help understand how market design might improve the land allocation, we conducted a survey and two lab-in-the-field experiments in Uganda and Kenya.

Our results suggest that the production technology and institutional environment have characteristics that theory predicts would restrict trade. Our lab-in-the-field-experiments, using a game that reflects these characteristics, confirm this conjecture. We find significant support for the hypothesis that more tailored market design can improve efficiency, without increasing inequality. We see our results as providing a first step toward better-functioning, more equitable land markets that leverage farmers' preferences and information to reshape the rural landscape.

References

Acampora, M., Casaburi, L., and Willis, J. (2022). Land rental markets: Experimental evidence from kenya. *University of Zurich mimeo.*

Adamopoulos, T., Brandt, L., Leight, J., and Restuccia, D. (2022). Misallocation, selection, and productivity: A quantitative analysis with panel data from china. *Econometrica*, 90(3):1261–1282.

Adamopoulos, T. and Restuccia, D. (2014). The size distribution of farms and international productivity differences. *The American Economic Review*, 104(6):1667–1697.

Agyei-Holmes, A., Buehren, N., Goldstein, M., Osei, R., Osei-Akoto, I., and Udry, C. (2020). The effects of land title registration on tenure security, investment and the allocation of productive resources. *World Bank Policy Research Working Paper No. 9376.*

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3):488.

Ali, D. A., Deininger, K., and Goldstein, M. (2014). Environmental and gender impacts of land tenure regularization in africa: Pilot evidence from rwanda. *Journal of Development Economics*, 110:262–275.

Ali, D. A., Deininger, K., and Ronchi, L. (2015). Costs and benefits of land fragmentation: Evidence from rwanda. *World Bank Policy Research Working Paper No. 7290.*

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495.

Aragón, F., Restuccia, D., and Rud, J. P. (2022a). Assessing misallocation in agriculture: Plots versus farms. *NBER Working Paper 29749.*

Aragón, F. M., Restuccia, D., and Rud, J. P. (2022b). Are small farms really more productive than large farms? *Food Policy*, 106:102168.

Athey, S. and Imbens, G. (2017). The econometrics of randomized experiments. In *Handbook of Field Experiments*, pages 73–140. Elsevier.

Balakrishnan, H. (2007). Techniques for reallocating airport resources during adverse weather. Technical report, in Proceedings of the 46th IEEE Conference on Decision and Control, 2949—2956.

Banerjee, A. (2003). Contracting constraints, credit markets, and economic develpment. In Dewatripont, M., Hndsen, L. P., and Turnovsky, S. J., editors, *Advances in Economics and Econometrics*, pages 1–46. Cambridge University Press.

Besley, T. and Ghatak, M. (2010). Property rights and economic development. *Handbook of Development Economics*, 5:4525–4595.

Binswanger, H. P. and Elgin, M. (1988). What are the prospects for land reform? *Proceedings Of The Twentieth International Conference of Agricultural Economists.*

Binswanger, H. P. and Rosenzweig, M. R. (1986). Behavioural and material determinants of production relations in agriculture. *Journal of Development Studies*, 22(3):503–539.

Bleakley, H. and Ferrie, J. (2014). Land openings on the georgia frontier and the coase theorem in the short- and long-run. Technical report, Working Paper.

Bolhuis, M., Rachapalli, S., and Restuccia, D. (2021). Misallocation in indian agriculture. *NBER Working Paper 29363.*

Britos, B., Hernandez, M. A., Robles, M., and Trupkin, D. R. (2022). Land market distortions and aggregate agricultural productivity: Evidence from guatemala. *Journal of Development Economics*, 155.

Bulte, E., Kontoleon, A., List, J., Turley, T., and Voors, M. (2013). When economics meets hierarchy: a field experiment on the workings of the invisible hand. Technical report, Working Paper.

Carter, M. R. and Mesbah, D. (1993). Can land market reform mitigate the exclusionary aspects of rapid agro-export growth? *World Development*, 21(7):1085–1100.

Chamberlin, E. H. (1948). An experimental imperfect market. *Journal of Political Economy*, 56(2):95—108.

Chari, A., Liu, E. M., Wang, S.-Y., and Wang, Y. (2020). Property rights, land misallocation, and agricultural efficiency in china. *The Review of Economic Studies*, 88(4):1831–1862.

Chen, C., Restuccia, D., and Santaeulalia-Llopis, R. (2022a). The effects of land markets on resource allocation and agricultural productivity. *Review of Economic Dynamics*, 45:41–54.

Chen, C., Restuccia, D., and Santaeulalia-Llopis, R. (2022b). Land misallocation and productivity. *American Economic Journal: Macroeconomics, forthcoming.*

Clarke, E. H. (1971). Multipart pricing of public goods. *Public Choice*, 11(1):17–33.

Cramton, P., Gibbons, R., and Klemperer, P. (1987). Dissolving a partnership efficiently. *Econometrica*, 55(3):615–632.

D'Arcy, M., Nistotskaya, M., and Olsson, O. (2021). Land property rights, cadasters and economic growth: A cross-country panel 1000-2015 ce. *QoG Working Paper Series 2021:3.*

de Janvry, A., Emerick, K., Gonzalez-Navarro, M., and Sadoulet, E. (2015). Delinking land rights from land use: Certification and migration in mexico. *American Economic Review*, 105(10):3125–3149.

de Soto, H. (2000). *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. Basic Books, New York.

Deininger, K. and Feder, G. (2001). Land institutions and land markets. *Handbook of Agricultural Economics*, 1:287–331.

Deininger, K. and Jin, S. (2006). Tenure security and land-related investment: Evidence from ethiopia. *European Economic Review*, 50(5):1245–1277.

Deininger, K., Monchuk, D., Nagarajan, H. K., and Singh, S. K. (2016). Does land fragmentation increase the cost of cultivation? evidence from india. *The Journal of Development Studies*, 53(1):82–98.

Delacrétaz, D., Kominers, S. D., and Teytelboym, A. (2020). Matching mechanisms for refugee resettlement. Technical report, Working Paper.

Delacrétaz, D., Loertscher, S., Marx, L. M., and Wilkening, T. (2019). Two-sided allocation problems, decomposability, and the impossibility of efficient trade. *Journal of Economic Theory*, 179(1):416–454.

Eswaran, M. and Kotwal, A. (1986). Access to capital and agrarian production organisation. *The Economic Journal*, 96(382):482.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, 133(4):1645–1692.

Fenske, J. (2011). Land tenure and investment incentives: Evidence from west africa. *Journal of Development Economics*, 95(2):137–156.

Fiala, N. (2015). Skills in the marketplace: Market efficiency, social orientation, and ability in a field-based experiment. *Journal of Economic Behavior and Organization*, 120:174–188.

Field, E. (2007). Entitled to work: Urban property rights and labor supply in peru. *The Quarterly Journal of Economics*, 122(4):1561–1602.

Fine, L., Goeree, J. K., Ishikida, T., and Ledyard, J. O. (2017). Ace: A combinatorial market mechanism. In Bichler, M. and Goeree, J. K., editors, *Spectrum Auction Design*, pages 874–902. Cambridge University press.

Finley, T., Franck, R., and Johnson, N. D. (2021). The effects of land redistribution: Evidence from the french revolution. *The Journal of Law and Economics*, 64(2):233–267.

Foster, A. D. and Rosenzweig, M. R. (2022). Are there too many farms in the world? labor market transaction costs, machine capacities, and optimal farm size. *Journal of Political Economy*, 130(3):636–680.

Galiani, S. and Schargrodsky, E. (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics*, 94(9-10):700–729.

Galiani, S. and Schargrodsky, E. (2011). Land property rights and resource allocation. *The Journal of Law and Economics*, 54(S4):S329–S345.

Goeree, J. K. and Lindsay, L. (2019). The Exposure Problem and Market Design. *The Review of Economic Studies*, 87(5):2230–2255.

Goldstein, M. and Udry, C. (2008). The profits of power: Land rights and agricultural investment in ghana. *Journal of Political Economy*, 116(6):981–1022.

Gollin, D. and Udry, C. (2021). Heterogeneity, measurement error, and misallocation: Evidence from african agriculture. *Journal of Political Economy*, 129(1):1–80.

Gresik, T. and Satterthwaite, M. (1989). The rate at which a simple market converges to efficiency as the number of traders increases: An asymptotic result for optimal trading mechanisms. *Journal of Economic Theory*, 48(1):304–332.

Grether, D. M., Isaac, R. M., and Plott, C. R. (1989). *Allocation of Scarce Resources: Experimental Economics and the Problem of Allocating Airport Slots. Underground classics in economics*. Westview Press.

Grossman, Z., Pincus, J., Shapiro, P., and Yengin, D. (2019). Second-best mecahnisms for land assembly and hold-out problems. *Journal of Public Economics*, 175:1–16.

Groves, T. (1973). Incentives in teams. *Econometrica*, 41(4):617.

Hartvigsen, M. (2014). Land consolidation and land banking in denmark-tradition, multi-purpose and perspectives. *Tidsskrift for Kortlægning og Arealforvaltning*, 122(47):51–74.

Hussam, R., Rigol, N., and Roth, B. N. (2022). Targeting high ability entrepreneurs using community information: Mechanism design in the field. *American Economic Review*, 112(3):861–98.

Kenya National Bureau of Statistics, Ministry of Health/Kenya, National AIDS Control Council/Kenya, Kenya Medical Research Institute, and National Council for Population and Development/Kenya (2015). Kenya demographic and health survey 2014. Technical report, Rockville, MD, USA.

Kominers, S. D. and Weyl, E. G. (2012). Concordance among holdouts. Technical report, Harvard Institute of Economic Research Discussion Paper.

Kremer, M. (2001a). Creating markets for new vaccines part 2: Rationale. In Jaffe, A. B., Lerner, J., and Stern, S., editors, *Innovation Policy and the Economy, Volume 1*, pages 73–118. MIT Press.

Kremer, M. (2001b). Creating markets for new vaccines part i: Rationale. In Jaffe, A. B., Lerner, J., and Stern, S., editors, *Innovation Policy and the Economy, Volume 1*, pages 35–72. MIT Press.

Kremer, M. and Wilkening, T. (2015). Protecting antiquities: A role for long-term leases. Technical report, Working Paper.

Kwasnica, A. M., Ledyard, J. O., Porter, D., and DeMartini, C. (2005). A new and improved design for multiobject iterative auctions. *Management Science*, 51(3):419–434.

Lawry, S., Samii, C., Hall, R., Leopold, A., Hornby, D., and Mtero, F. (2016). The impact of land property rights interventions on investment and agricultural productivity in developing countries: a systematic review. *Journal of Development Effectiveness*, 9(1):61–81.

Libecap, G. D. and Lueck, D. (2011). The demarcation of land and the role of coordinating property institutions. *Journal of Political Economy*, 119(3):426–467.

Loertscher, S., Marx, L. M., and Wilkening, T. (2015). A long way coming: Designing centralized markets with privately informed buyers and sellers. *Journal of Economic Literature*, 55(4):857–897.

Loertscher, S. and Waser, C. (2019). Optimal structure and dissolution of partnerships. *Theoretical Economics*, 14:1063–1114.

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review*, 111(6):1807–1844.

McAfee, R. P. (1992). A dominant strategy double auction. *Journal of Economic Theory*, 56(2):434–450.

McCloskey, D. (1975). English open fields as behavior towards risk. In *Research in Economic History: an Annual Compilation*. Jai Press.

Milgrom, P. (2007). Package auctions and exchanges. *Econometrica*, 75(4):935—965.

Milgrom, P. (2017). *Discovering Prices: Auction Design in Markets with Complex Constraints*. Columbia University Press.

Milgrom, P. R. and Segal, I. R. (2017). Designing the us incentive auction. In Bichler, M. and Goeree, J. K., editors, *Spectrum Auction Design*, pages 827–873. Cambridge University press.

Myerson, R. and Satterthwaite, M. (1983). Efficient Mechanisms for Bilateral Trading. *Journal of Economic Theory*, 29(2):265–281.

Nemes, V., Plott, C. R., and Stoneham, G. (2008). Electronic bushbroker exchange: Designing a combinatorial double auction for native vegetation offsets. Technical report, Working Paper.

Nethercote, N., Stuckey, P., Becket, R., Brand, S., Duck, G., and Tack, G. (2007). Minizinc: Towards a standard cp modelling language. In Bessiéere, C., editor, *Proceedings of the 13th International Conference on Principles and Practice of Constraint Programming, volume 4741 of LNCS*, pages 529–543. Berlin, Heidelberg: Springer.

Plassman, F. and Tideman, T. N. (2010). Providing incentives for efficient land assembly. Technical report, Available at SSRN: https://ssrn.com./abstract=1-15820.

Platteau, J.-P. (2015). *Institutions, Social Norms and Economic Development.* Routledge.

Rassenti, S. J., Smith, V. L., and Bulfin, R. L. (1982). A combinational auction mechanism for airport time slot allocation. *Bell Journal of Economics*, 13(2):402–417.

Roth, A. E. (2002). The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica*, 70(4):1341–1378.

Rustichini, A., Satterthwaite, M. A., and Williams, S. R. (1994). Convergence to efficiency in a simple market with incomplete information. *Econometrica*, 62(5):1041–1063.

Sarkar, S. (2017). Mechanism design for land acquisition. *International Journal of Game Theory*, 46(3):783–812.

Sarkar, S. (2022). Optimal mechanism for land acquisition. *Review of Economic Design*, 26:87–116.

Segal, I. and Whinston, M. (2016). Property rights and the efficiency of bargaining. *Journal of the European Economic Association*, 14(6):1287–1328.

Smith, C. (2019). Land concentration and long-run development in the frontier united states. *mimeo.*

Smith, V. L. (1962). An experimental study of competitive market behavior. *Journal of Political Economy*, 70(2):111—1375.

Smith, V. L. (1964). Effect of market organization on competitive equilibrium. *The Quarterly Journal of Economics*, 78(2):181—201.

Smith, V. L. (1990). The boundaries of competitive price theory: Convergence, expectations, and transaction costs. In Green, L. and Kagel, J., editors, *Advances in Behavioral Economics*, pages 31–53. Ablex Publishing Company, Norwood.

Smith, V. L. and Williams, A. W. (1981). An experimental study of decentralized institutions of monopoly restraint. In Horwich, G. and Quirk, J., editors, *Essays in Contemporary Fields of Economics in Honor of Emmanuel T. Zweiler*, pages 83–106. Purdue University Press, West Lafayette.

Stuckey, P. J., Feydy, T., Schutt, A., Tack, G., and Fischer, J. (2014). The MiniZinc Challenge 2008–2013. *AI Magazine*, 35(2):55–60.

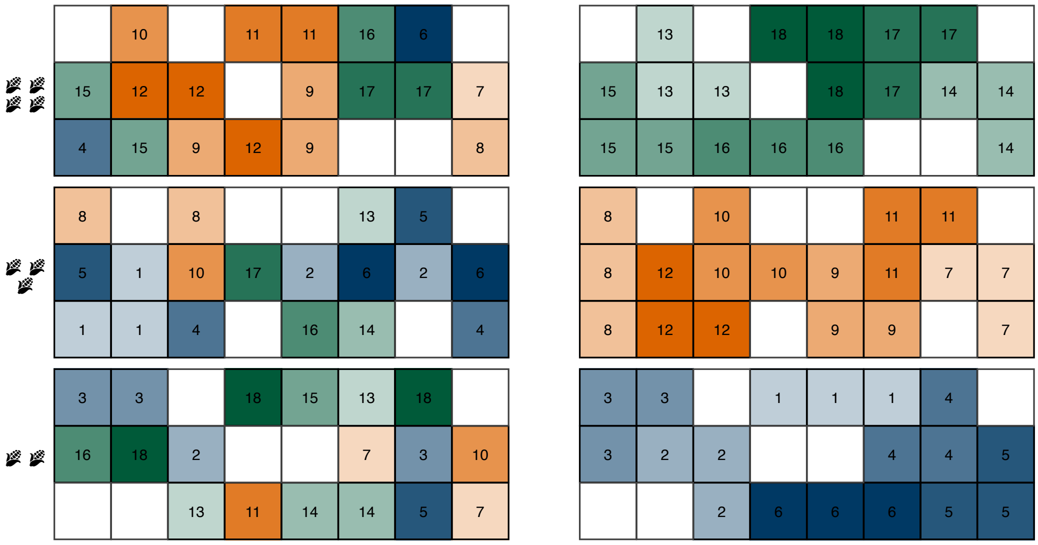
Tanaka, T. (2007). Resource allocatio with spatial externalities: Experiments on land consolidation. *The B.E. Journal of Economic Analysis & Policy*, 7(1 (Topics) Article 7):1–33.

Uganda National Bureau of Statistics (2020). Uganda National Panel Survey (UNPS) 2019-2020. Ref: UGA_2019_UNPS_v03_M. Technical report.

Vickrey, W. (1961). Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance*, 16(1):8–37.

Williams, S. R. (1999). A characterization of efficient, bayesian incentive compatible mechanisms. *Economic Theory*, 14:155–180.

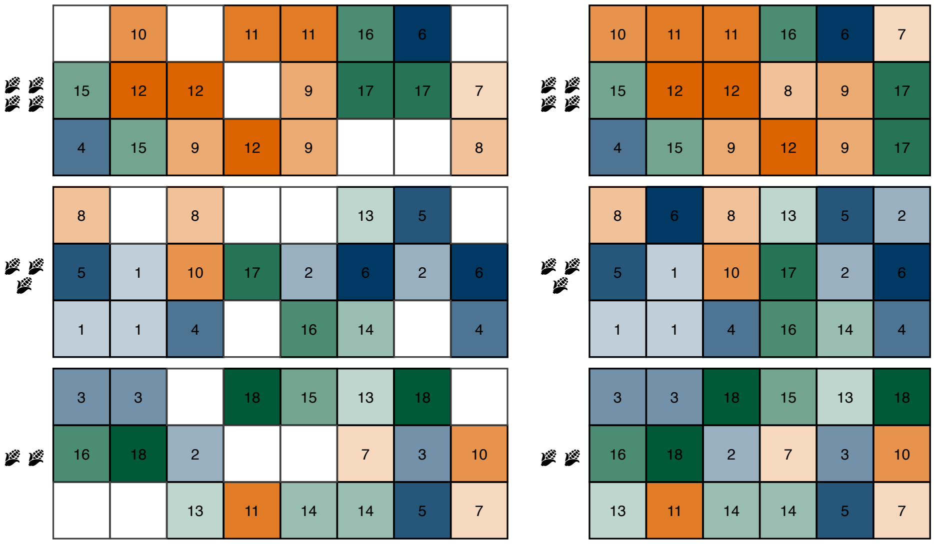
9 Figures



(a) Example initial allocation (b) Example efficient allocation

Figure I: The land trading game

Properties: (1) Players are numbered 1–18 in increasing *ability* order. Blue corresponds to the lowest ability type, Orange to the middle type, and Green to the high type. (2) There are three regions, (, ,), in increasing *land quality* order. (3) Ability and quality are complements. (4) Contiguous farms earn higher profits than fragmented farms. (5) Farmers cannot cultivate more than three plots.



(a) Complex (b) Simple

Figure II: Example map in complex and simple form

Properties: same as Figure I. Simple map is constructed from Complex by moving plot owners leftwards so as to retain all existing adjacencies between plots of the same owner, while preserving as much as possible of the relative locations of different owners.

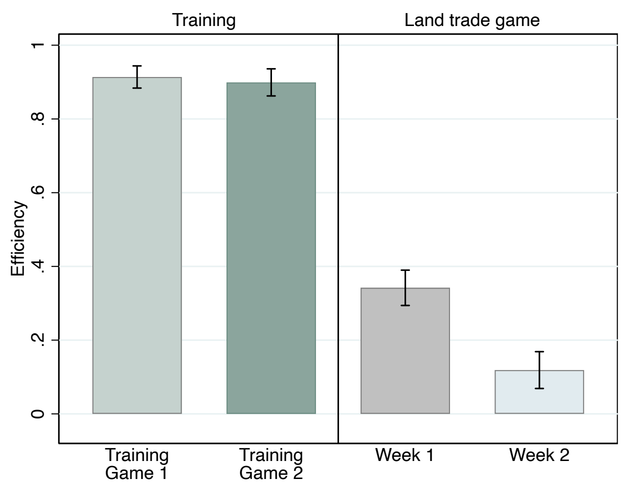


Figure III: Efficiency in Training Games and in First and Second Week of Experiment 1

The graph shows mean efficiency in each game, with 95% confidence intervals.

(a) Example initial allocation

(b) Example efficient allocation

Figure IV: Example Map from Experiment 2

Note: numbers correspond to player IDs: 1–2 are low types, 3–4 medium types, 5–6 high types. The top region is high-quality land, the middle region medium-quality, and the bottom region low-quality.

10 Tables

Table I: Fragmentation

Existence of fragmentation	mean	S.D.	obs
Owns two or more fragmented plots	0.64	0.48	1404
Distance between plots (in min)			
Lower bound	23.81	27.67	801
Upper bound	41.22	46.07	801
Distance from home to plots (in min)			
Purchased plots	19.28	26.63	1234
Inherited plots	11.85	19.22	1345
Plots received as gift	18.09	28.74	119
Costs of fragmentation (yes/no)			
Thinks having two one-acre plots better in same location	0.91	0.28	1404
Suppose that, instead of being separated, all your land was in the same place:			
Think they would earn more	0.88	0.32	896
How much more? (fraction)	0.53	0.28	792
Imagine a farmer with 2 **contiguous** acres who produces 20 bags of maize			
- He inherits one more acre **next** to the other two, so now he has three acres. Do you think he would be now be able to produce:			
Less than 30 bags	0.01	0.09	1404
30 bags	0.33	0.47	1404
More than 30 bags	0.66	0.47	1404
- He inherits one more acre **but far away** from the other two, so now he has three acres. Do you think he would be now be able to produce:			
Less than 30 bags	0.33	0.47	1404
30 bags	0.37	0.48	1404
More than 30 bags	0.30	0.46	1404

Note: Respondents are participants from Experiment 1, see section 5 for implementation details. We elicited walking distances from the home to each plot, and use these to compute a lower bound and an upper bound on the distance *between* the fragmented plots. The lower bound assumes both plots lie on the same bearing from home, so their separation is the difference between their distances from home. The upper bound assumes they are in opposite directions from home, so the distance between them is the sum of their distances from home.

Table II: Complementarity Between Farmer Ability and Land Quality

Farmer heterogeneity	mean	S.D.	obs
Are there better farmers than others in your village?	0.99	0.09	1404
Does everyone in your village agree with who they are?	0.98	0.12	1393
How do you know these are better farmers? (select all that apply)			
Produce more yield/acre	0.93	0.25	1393
Use innovative farming techniques	0.37	0.48	1393
Have better machinery	0.05	0.21	1393
Have received farm training	0.18	0.39	1393
Are older farmers	0.11	0.31	1393
Do you think your village as a whole would produce more agricultural yield if the best farmers were cultivating more of the village's land?			
Produce the same	0.00	0.06	1404
Produce somewhat more	0.12	0.32	1404
Produce much more	0.88	0.33	1404
Land-crop heterogeneity			
Is some land better quality than other land in this village?	0.99	0.12	1404
Suppose you were thinking of buying or renting in a plot in your village. To what extent do you agree with the following statement: Good land is good for all crops, bad land is bad for all crops			
Strongly disagree	0.12	0.32	1404
Disagree	0.15	0.36	1404
Neither disagree or agree	0.20	0.40	1404
Agree	0.41	0.49	1404
Strongly agree	0.13	0.34	1404
Land-farmer complementarity			
Do you think your village as a whole would produce more agricultural yield if the best farmers were cultivating the best land?			
Produce much less	0.00	0.04	1404
Produce somewhat less	0.00	0.04	1404
Produce the same	0.01	0.08	1404
Produce somewhat more	0.16	0.36	1404
Produce much more	0.83	0.37	1404

Note: Respondents are participants from Experiment 1, see section 5 for implementation details.

Table III: Decreasing Returns at the Farm Level

Could you farm more land than you have now? (yes/no)	0.60	0.49	1404
Are some people in your village better at managing large farms? (yes/no)	0.99	0.09	1404
How much land can best farmer manage (acres) w/o hired labor	4.67	2.55	1403
How much land can worst farmer manage (acres) w/o hired labor	0.73	0.40	1401
Max amount of land that respondent's hh can manage (acres) w/o hired labor	2.53	1.50	1404
Max amount of land per adult in hh (acres) w/o hired labor	1.02	0.75	1404

Note: Respondents are participants from Experiment 1, see section 5 for implementation details.

Table IV: Current Market Institutions: Land Markets

Land markets	Frequently/ very frequently	Occasionally	Rarely/ very rarely	Never
How frequently do people in your village buy/sell land?	0.23	0.33	0.43	0.00

	mean	S.D.	obs
Have you sold any land during the past 12 months?	0.02	0.14	1404
– How many acres did you sell in total?	0.72	0.55	29
Have you bought any land during the past 12 months?	0.08	0.27	1404
– How many acres did you buy in total?	0.92	1.03	110
Have you rented OUT any land during the past 12 months?	0.04	0.19	1404
– How many acres did you rent out in total?	1.44	1.32	53
Have you rented IN any land during the past 12 months?	0.17	0.38	1404
– How many acres did you rent in in total?	1.24	0.86	242
Would you know if someone is selling/buying/renting land?	0.89	0.31	1404
Attempts to consolidate			
Tried to consolidate	0.24	0.43	1404
– Successful? (Conditional on Trying)	0.50	0.50	341
Land owned by people from outside village			
How many out of 10 plots are owned by people outside village?	1.84	1.34	1404
Characteristics of household's own plots			
Do you cultivate this plot? yes	0.96	0.20	2726
If plot not cultivated, is it rented out? yes	0.13	0.33	111
Plot was purchased	0.45	0.50	2726
Plot was inherited	0.49	0.50	2726
Plot was given	0.04	0.20	2726

Note: Respondents are participants from Experiment 1, see section 5 for implementation details.

Table V: Current Market Institutions: Culture and Attitudes Toward Land Trade

	Var mean	Very unfair (1)	Unfair (2)	Neither unfair or fair (3)	Fair (4)	Very fair (5)
Do you think it would be fair if best farmers cultivate best land?	2.52	0.38	0.23	0.06	0.13	0.19
	Var mean	**Strongly disagree (1)**	**Disagree (2)**	**Neither disagree or agree (3)**	**Agree (4)**	**Strongly agree (5)**
A family should never sell their ancestral land	4.53	0.04	0.04	0.01	0.14	0.76
A family should be free to trade their land if it improves their situation	3.06	0.18	0.18	0.14	0.38	0.12
It is very important to me to own some land	4.75	0.01	0.00	0.01	0.20	0.79
If I could sell my land for a good price, I would move to the city	1.62	0.61	0.28	0.04	0.04	0.04
I would like to migrate	1.68	0.53	0.35	0.05	0.05	0.02
I would like my children to be farmers	3.74	0.04	0.16	0.11	0.41	0.28
People shouldn't sell land to people within the family	1.74	0.43	0.49	0.02	0.03	0.03
People shouldn't sell land to people in the village outside their family	2.00	0.26	0.62	0.03	0.05	0.04
People shouldn't sell land to people outside the village	2.38	0.15	0.59	0.05	0.11	0.09
People shouldn't sell land to people from outside the tribe	2.71	0.12	0.49	0.08	0.18	0.13
People shouldn't sell land to foreigners	3.53	0.09	0.22	0.08	0.32	0.30
People shouldn't swap land within the family	2.05	0.33	0.50	0.03	0.06	0.08
People shouldn't swap land with the people in the village outside their family	2.59	0.16	0.53	0.03	0.12	0.15
People shouldn't swap land with people outside the village	3.15	0.07	0.41	0.06	0.24	0.23
People shouldn't swap land with people from outside the tribe	3.39	0.05	0.35	0.07	0.25	0.29
Observations: 1404						

Note: Respondents are participants from Experiment 1, see section 5 for implementation details.

Table VI: Adverse Selection

Adverse selection	mean	S.D.
Knows how to locate best land	0.99	0.11
How would you check if a plot is of good quality? (select all that apply)		
Look at what is growing	0.82	0.38
Look at the soil	0.79	0.41
Look at the slope	0.14	0.35
Look at the water resources	0.12	0.32
Look at weather in location	0.06	0.24
Ask owner	0.06	0.24
Ask neighbours	0.04	0.19
Ask others in village	0.02	0.14
Look at irrigation	0.00	0.05

Suppose you were thinking of buying or renting in a plot in your village.	Var mean	Strongly disagree	Disagree	Neither disagree or agree	Agree	Strongly agree
To what extent do you agree with the following:		(1)	(2)	(3)	(4)	(5)
People only sell or rent out their worst plots	2.91	0.18	0.30	0.14	0.22	0.17
Thinks difficult to assess quality of plot owned by others	1.86	0.28	0.63	0.05	0.02	0.01
Knows how to assess quality of plot owned by others	4.27	0.00	0.01	0.04	0.60	0.34
Thinks difficult to assess quality of own plot	1.72	0.40	0.54	0.03	0.03	0.01
Knows how to assess quality of own plot	4.44	0.00	0.00	0.02	0.50	0.48
Observations: 1404						

Note: Respondents are participants from Experiment 1, see section 5 for implementation details.

Table VII: Efficiency in Experiment 1 (Uganda Decentralized Trade)

		Decomposition		
	(1) Efficiency	(2) Consolidation	(3) Sorting	(4) Avoided exposure loss
Panel A: Average efficiency in decentralized trade				
Mean	0.230	0.306	0.127	-0.202
as % of first best		*0.611*	*0.254*	
Observations	136	136	136	136
Panel B: Impact of centralization				
Centralization	0.348***	0.062***	0.001	0.286***
	(0.019)	(0.008)	(0.005)	(0.019)
as % of first best		*0.124*	*0.001*	
FDR q-value: centralization		[0.001]	[0.439]	[0.001]
Control mean	0.119	0.287	0.147	-0.315
Control mean: % of first best		*0.574*	*0.294*	
Observations	136	136	136	136
Panel C: Impact of eliminating nontradable plots				
Simple map	0.044	0.022*	-0.018	0.040
	(0.031)	(0.012)	(0.014)	(0.028)
as % of first best		*0.045*	*-0.036*	
FDR q-value: simple map		[0.232]	[0.232]	[0.232]
Control mean	0.208	0.294	0.136	-0.222
Control mean: % of first best		*0.589*	*0.272*	
Observations	136	136	136	136
Panel D: Impact of centralization and eliminating nontradable plots				
Centralization	0.342***	0.059***	-0.005	0.287***
	(0.027)	(0.008)	(0.008)	(0.028)
Centralization × simple map	0.013	0.006	0.011	-0.003
	(0.038)	(0.015)	(0.010)	(0.037)
FDR q-value: centralization		[0.001]	[0.211]	[0.001]
FDR q-value: centralization × simple map		[1.000]	[1.000]	[1.000]
Control mean	0.111	0.284	0.151	-0.324
Observations	136	136	136	136

Note: Panel A shows mean efficiency in the two weeks of decentralized trade, and its decomposition into Consolidation, Sorting, and Exposure losses (due to farmers holding too much land). Panel B examines the effect of the centralization intervention. Panel C examines the effect of "simple" maps that eliminate nontradable plots. Panel D examines how the effects of centralization interact with simple maps. Coefficients are measured in efficiency units, i.e. as a share of total potential gains from trade. "% of first best" instead expresses the consolidation and sorting coefficients as a share of the total potential gains from these components alone. Panels A and C include data from weeks 1 and 2, excluding the centralization treatment. Panels B and D include data from week 2, pre and post-centralization. Control mean in panel B corresponds to week 2 pre-centralization, in panel C: complex maps, and in panel D: week 2, complex maps, pre-centralization. Regressions in panels B and D control for village fixed effects. Regressions in panel C control for village and week 2 × map fixed effects. Standard errors (in parentheses) are clustered by village. q-values adjust for multiple testing across the three components of efficiency.

Table VIII: Inequality Experiment 1 (Uganda Decentralized Trade

	Atkinson Index (log utility)			
	(1) + 5-day wage	(2) + worst score	(3) + show-up fee	(4) rounded
Panel A: Impact of centralization				
Centralization	-0.004***	-0.007***	-0.122***	-0.286***
	(0.001)	(0.001)	(0.022)	(0.032)
Control mean	0.012	0.020	0.209	0.522
Observations	136	136	136	136
Panel B: Impact of eliminating nontradable plots				
Simple map	-0.003**	-0.011	-0.068*	-0.090**
	(0.001)	(0.007)	(0.036)	(0.043)
Control mean	0.014	0.030	0.237	0.551
Observations	136	136	136	136
Panel C: Impact of centralization and eliminating nontradable plots				
Centralization treatment	-0.005***	-0.008***	-0.146***	-0.304***
	(0.001)	(0.002)	(0.031)	(0.042)
Centralization × simple map	0.001	0.002	0.048	0.036
	(0.001)	(0.003)	(0.044)	(0.064)
Control mean	0.013	0.023	0.255	0.582
Observations	136	136	136	136

Note: Panel A examines the effect of the centralization intervention on inequality. Panel B examines the effect of "simple" maps that eliminate nontradable plots. Panel C examines how the effects of centralization interact with simple maps. The outcome variable is the Atkinson inequality index (equation (1)). Higher values mean greater inequality. Because many participants had negative net assets and the index is based on log assets, we explore four different corrections to ensure that the index is defined. Column (1) uses final net assets adding a five-day wage, column (2) uses final net assets adding the worst score in the sample, column (3) uses final net assets adding the show-up fee and rounds to 1 those with negative assets, and column (4) uses final net assets and rounds to 1 those with negative assets. 5.23% of the sample (player-period level) has negative net assets. Five-day wage is a total of 250,000 game shillings. Worst score is -181,000 game shillings. Show-up fee is a total of 50,000 game shillings. 1.3% of the sample has a negative final net assets after adding the show-up fee. Control mean in panel A: complex maps, and in panel B: week 2 pre-centralization. Regressions in panels A and C use data from week 2, pre and post-centralization, panel B uses data from week 1 and week 2, excluding the centralization treatment. Regressions in panel A control for village fixed effects. Regressions in panel B control for village and week 2 × map fixed effects. Regressions in panel C control for village fixed effects. Standard errors clustered by village in parentheses.

Table IX: Efficiency in Experiment 2 (Kenya Package Exchanges)

		Decomposition		
	(1) Efficiency	(2) Consolidation	(3) Sorting	(4) Avoided exposure loss
Panel A: Impact of package mechanisms				
Package-2	0.033	0.006	0.008	0.018
	(0.033)	(0.018)	(0.012)	(0.013)
as % of first best		*0.008*	*0.031*	
Package-4	0.068^{**}	0.019	0.032^{**}	0.017
	(0.031)	(0.016)	(0.013)	(0.014)
as % of first best		*0.026*	*0.119*	
FDR q-value: Package-2		[1.000]	[1.000]	[1.000]
FDR q-value: Package-4		[0.184]	[0.076]	[0.184]
Control mean	0.696	0.628	0.118	-0.050
Control mean: % of first best		*0.856*	*0.444*	
Observations	318	318	318	318
Panel B: Impact of package mechanisms and low cash treatment				
Package-2	0.023	-0.012	0.024	0.011
	(0.048)	(0.026)	(0.018)	(0.020)
Package-4	0.067	0.016	0.044^{**}	0.007
	(0.048)	(0.020)	(0.021)	(0.021)
Low cash	-0.003	-0.016	0.015	-0.002
	(0.049)	(0.028)	(0.023)	(0.021)
Package-2 × low cash	0.019	0.036	-0.032	0.015
	(0.056)	(0.032)	(0.031)	(0.024)
Package-4 × low cash	0.001	0.006	-0.025	0.020
	(0.061)	(0.032)	(0.029)	(0.028)
FDR q-value: Package-2		[1.000]	[1.000]	[1.000]
FDR q-value: Package-4		[0.768]	[0.137]	[0.999]
F-test p-value: all low cash effects = 0	0.948	0.471	0.717	0.469
Control mean	0.702	0.640	0.110	-0.048
Observations	318	318	318	318

Note: Panel A shows treatment effects of our package exchange treatments on overall efficiency, and its decomposition into Consolidation, Sorting, and (avoided) exposure losses due to participants holding too much/too little land. Panel B examines the effects of the low cash manipulation, designed to induce the effects of credit constraints, and how it interacts with the trading mechanism. Coefficients are measured in efficiency units, i.e. as a share of total potential gains from trade. "% of first best" expresses consolidation and sorting coefficients as a share of total potential gains from these components. Control mean in Panel A corresponds to the Package-1 treatment, in panel B: Package-1 with high cash. All regressions exclude randomization block 1 (see section 6.3 for discussion). All regressions control for auction round, map, and randomization block fixed effects. Panel A additionally controls for a "low cash" dummy. Standard errors clustered by session in parentheses. q-values adjust for multiple testing across the three components of efficiency. F-test p-values correspond to the null hypothesis that the effects of the low cash dummy and its interactions all equal zero.

Table X: Inequality in Experiment 2 (Kenya Package Exchanges)

	Atkinson Index (log utility)		
	(1) High cash	(2) Low cash	(3) High & Low
Package-2	0.0004	-0.0031***	0.0004
	(0.0006)	(0.0011)	(0.0006)
Package-4	-0.0002	-0.0019*	-0.0002
	(0.0006)	(0.0010)	(0.0006)
Package-2 × low cash			-0.0035***
			(0.0011)
Package-4 × low cash			-0.0017
			(0.0010)
F-test p-value: all low cash effects = 0			0.006
Control mean	0.012	0.035	0.024
Observations	159	159	318

Note: The table examines the effect of our package exchange treatments on inequality of final outcomes, and how they interact with the low cash treatment (designed to induce the effects of credit constraints). The outcome variable is the Atkinson inequality index (equation (1)). Higher values mean greater inequality. Regressions in column (1) use data from high cash group, in column (2) use data from low cash group and in column (3) use data from both high and low cash groups. Control means correspond to the Package-1 treatment, but are not directly comparable between low versus high cash auctions because the Atkinson inequality index is not invariant to additive changes in total wealth that are induced by the cash variation. All regressions exclude randomization block 1 (see section 6.3 for discussion). All regressions control for auction round, map, and randomization block fixed effects, and in column (3) adds interactions of each of the fixed effects with a low cash dummy to address the level effects of the cash manipulation. Standard errors clustered by session in parentheses. F-test p-values correspond to the null hypothesis that the interactions between the low cash dummy and package treatments both equal zero.

A Appendix A: Implementation Details for Experiment 1

This appendix provides additional implementation details for Experiment 1.

A.1 Selection of Villages and Participants

We sampled our villages in a two-step process. First, we chose a set of villages to visit for our experiment. Second, we selected the participants to recruit.

Villages: We worked in the Masaka district, Uganda. Masaka was selected because the majority of land in Masaka is owned under freehold, i.e. it is in principle tradable by the owner. Tenure form differs in other parts of Uganda, and landholders do not have the legal or traditional right to trade land everywhere. While our experiment does not involve real land trade, we wanted to work in a region where land trade is imaginable to participants.

We selected villages using an administrative unit-level GIS file, containing census data from 2002 and 2010. We first dropped villages not listed as being in Masaka county, then dropped subcounties that subsequently joined other districts, leaving an initial sample frame of 357 villages, belonging to two counties (Bukoto and Masaka Municipality), 10 subcounties, and 39 parishes.

1. Next, we dropped 11 villages with zero population. This left us with 346 villages.

2. We dropped 4 villages with duplicate names, that would be difficult for our field team to identify reliably. This left us with 342 villages.

3. We dropped villages that were densely populated and had limited farmland. While these villages may contain many farming households, we were concerned that recruitment and attrition would be more challenging in these areas. We do this in three ways. First, we dropped villages above the 90th centile for population or population density. Second, we dropped parishes with median village above those thresholds. Third, we dropped Masaka Municipality (the main urban area). The thresholds were tuned by visual inspection of satellite images, inspecting the “marginal” villages around the threshold for whether they had significant farmland. This left us with 274 villages (Masaka municipality accounts for 53 of the 68 villages dropped).

4. We also dropped 7 villages that were previously visited for piloting. This left us with 267 villages.

5. We dropped 26 coastal villages (identified by visual inspection) that were expected to be dominated by fishing and other activities rather than agriculture. This left us with our final sampling frame of 241 villages.

The 241 villages belonged to 31 “parishes” (the next highest administrative unit), which we used for stratification (see section A.3 below).

Participants: In each selected village we first met with the village chief, and asked them to give us a list of as many households as they could think of (excluding the chief's own family members) that would be expected to be interested to participate in a sequence of trading games on three days separated by one week each.[36] The chief was also asked to attend the meetings and assist with ensuring that selected participants attend, and were compensated for their time.

We selected households randomly from the list and sought the consent of the household head to participate in the experiment. Eligibility criteria were 1) cultivation of some land, 2) reporting that at least 50% of household income was derived from farming, 3) having access to a mobile money account. Criteria 1 and 2 were intended to ensure we sample a relevant population that might be interested in real land trade, 3 ensured that participants can be paid their study earnings. If the household head is interested but not available we allow them to send another household member in their place.

We proceed this way until 22 households had been recruited. The first 18 were our intended "primary" participants, and the remaining 4 acted as reserves. The reserves were asked to attend each session, and paid show-up fees for doing so. If a primary participant did not attend a session, they were replaced by a reserve.

A.2 Payoffs and Maps

Payoffs: Each of our games consisted of 18 players and a map. Each players was assigned a numeric ability type. This ability type was private information and farmers were asked not to share it with others. Consistent with the results from the survey, the payoff function of each farmer had three key properties:

1. Ability-quality complementarities: the return to a given piece of land was the product of the player ability and the land quality. The land quality types were Low, Medium, High: $\{2, 3, 4\}$. The player ability types were Low: $\{0.8, 0.9, 1, 1, 1.1, 1.2\}$, Medium: $\{1.3, 1.4, 1.5, 1.5, 1.6, 1.7\}$, and High: $\{1.8, 1.9, 2, 2, 2.1, 2.2\}$.

2. Spatial complementarities: players earned an "adjacency bonus" when two of their plots shared a border, and two bonuses when three plots shared two borders (either in a vertical or horizontal strip or an"L"-shaped unit). The adjacency bonus was fixed at the player level to 10% of the player's value of a high-quality plot (e.g. a player of ability type 1 had an adjacency bonus worth $1 \times 4 \times 0.1 = 0.4$). To limit the number of payoff parameters that participants had

[36] In piloting we experimented with fully random sampling of participants (we attempted to obtain a full list of households from the LC1 chief and selected randomly from that list). However, this led to several of the selected participants being quite uninterested in participation, so many would send another household member, or might simply not show up. We therefore settled on giving some guidance to the chief, to suggest "interested" households. The primary goal is to ensure successful completion of the experiment since significant attrition can prevent completion of the three stages. It means the study population may be less representative of the village population, but may conversely be *more* representative of those interested in land trade.

to keep track of, the adjacent bonus was independent of land quality. Adjacency bonuses could only be earned within a land quality region.

3. Span of control: each player could farm a maximum of three plots – if they end the game with more than three they earned the return to their best three-plot combination.

Land values on maps and plot titles were represented visually with two, three, or four icons representing heads of maize. Thus, each player had four key payoff parameters to keep track of: their value for each type of land, and their adjacency bonus.

Given the payoff function, the efficient allocation was simple to compute: each player should hold three adjacent plots, positively sorted by quality-ability type. The payoff parameters were calibrated such that the gains from trade were divided approximately 50–50 between sorting and consolidation.

To generate payoff numbers that were similar in magnitude to those used by farmers in day-to-day trades, we multiplied payoffs by 20,000 and expressed values in terms of "game shillings."[37] Each player began each game with 240,000 game shillings in printed paper bills that could be used for exchange.

We also assigned each player an initial "debt," to be deducted from their final payoff when computing earnings from the games. The debt levels were calculated such that each person began the game with net assets (land value, plus initial adjacency bonuses, plus cash, minus debt) equal to 70,000 game shillings.

Final earnings (in game shillings) were calculated as

$$\text{Earnings} = \text{Final land value} + \text{Final adjacency bonuses} + \text{Final cash} - \text{Initial debt}$$

and then converted to UGX at the rate 5 game shillings = 1 UGX.

The primary role of debt was to calibrate incentives in the game. We wanted the gains from trade (in relative and absolute local currency terms) to be sufficiently large that participants payed attention and participated fully. Final earnings depended on initial assets and gains from trade. For a given average payoff, subtracting debt from initial assets increased the contribution of gains and therefore sharpened incentives. We also used the debt to start all players with the same payoff so that inequality changes could be easily compared across games.

Maps: We used the following procedure to construct the complex map and assign players to it:

1. We began with a grid of three 3*8 blocks of plots,

2. We randomly group plots into 3 groups of 24, corresponding to 18 players and 6 "non-trading" dummy players, such that:

[37] This implied that the minimal land value was 32,000 ($0.8 \times 2 \times 20,000$) and the maximum land value was 176,000. The minimum adjacency bonus was 6,400 and the maximum was 17,600.

- Non-trading players own exactly six plots per quality region (otherwise the first-best allocation is not achievable).
- There were no simple blocking allocations, that is, a single player that holds three plots that isolate a corner, or a combination of non-trading players that hold between them two or three plots blocking a corner.
- The number of initial "adjacencies"[38] and "near adjacencies"[39] averaged 0.3–0.4 per player. These thresholds were set to ensure a realistic amount of clustering of initial ownership, based on visual inspection of real-world maps.
- The contribution of land consolidation and sorting to total gains from trade was balanced, with a relative contribution range between 47.5%–52.5%.

Typically maps generated in this way have no efficient packing solution. We manually identified 10 such that (i) feasibility could be achieved by moving a maximum of 1 plot and (ii) The resulting map had a single contiguous set of land. Of these, 2 were solvable with no edits, and the remaining 8 needed one swap (exchanging a single plot between one trading and one non-trading player). Swaps were implemented so as to avoid breaking or creating new adjacencies. Thus the initial payoffs were unaffected.

A.2.1 Making the simple maps

For our "simple" treatment we want to eliminate the non-trading players. We do this by manually "compressing" the complex maps, so as to preserve the adjacency structure of the map. We did this by shifting plots horizontally left, except where doing so would create or break an adjacency. Therefore, the initial payoffs are unaffected. Note that it is not possible to preserve the "near-adjacency" structure. See Figure II.

A.2.2 Pairing complex and simple maps

Following the above process, we generated 10 candidate maps, each with a complex and simple variant. From these we selected 8 and created four pairs, matched according to the number of possible efficient solutions in the complex form. According to our internal map numbering these are:

- Maps 69 and 148 which have on average 1.67 solutions per quality block, and 8 adjacencies among the trading players.
- Maps 74 and 149 which have on average 3 and 3.67 solutions per quality block, and 5 adjacencies among the trading players.

[38] A player with two horizontally or vertically adjacent plots within the same quality region counts 1, a player with three plots sharing two borders counts 2.

[39] Two plots owned by the same player that are diagonally adjacent, or separated by one plot, count as 1 near adjacency, so long as that player is not already fully consolidated. We allow near-adjacencies to span across quality types.

- Maps 93 and 130 which have on average 3.67 and 4.67 solutions per quality block, and 6 adjacencies among the trading players.
- Maps 28 and 193 which have on average 5 and 5.33 solutions per quality block, and 4 and 6 adjacencies respectively among the trading players.

A.3 Treatment assignment

Each village played the game twice, once on a simple map and once on a complex map. This section details how the ordering and the specific maps were assigned.

A.3.1 Possible assignments

As described in Appendix A.2 our map generation procedure yielded 8 maps (internal IDs 28, 69, 74, 93, 130, 148, 149, 193) each of which had a simple and a complex form. We grouped these 8 maps into 4 matched pairs according to the number of possible efficient packings available in their complex form. Accounting for possible map and complexity orderings this yielded 16 possible assignments. These are listed in Table A1.

Assignment	Assignment pair	Map ordering	Complexity ordering
1	1	(69, 148)	(simple, complex)
2	1	(69, 148)	(complex, simple)
3	2	(148, 69)	(simple, complex)
4	2	(148, 69)	(complex, simple)
5	3	(74, 149)	(simple, complex)
6	3	(74, 149)	(complex, simple)
7	4	(149, 74)	(simple, complex)
8	4	(149, 74)	(complex, simple)
9	5	(93, 130)	(simple, complex)
10	5	(93, 130)	(complex, simple)
11	6	(130, 93)	(simple, complex)
12	6	(130, 93)	(complex, simple)
13	7	(28, 193)	(simple, complex)
14	7	(28, 193)	(complex, simple)
15	8	(193, 28)	(simple, complex)
16	8	(193, 28)	(complex, simple)

Table A1: Possible treatment assignments

Our field plan involved two field teams working simultaneously five days per week, covering 10 villages per week. We intended to sample 68 villages, leaving two vacant "slots" for replacement villages in case a village decides to withdraw (see section

A.3.3). The 68 villages constituted four complete blocks of 16 assignments plus one randomly selected block of four (either assignments 1–4, 5–8, 9–12 or 13–16).[40]

A.3.2 Randomization

- Our primary regression specification exploit the within-village variation in complexity, but to increase power in between-village comparisons we stratified the assignment by parish and study date.
- Specifically, when selecting study villages we first randomly ordered parishes, then randomly selected pairs of villages from each parish. Each pair of villages was then assigned an *assignment pair* (see Table A1), so they differed only in their {simple, complex} ordering.
- We randomly ordered non-selected villages within each parish to act as backups in case a selected village opted not to participate.
- We had two experimental teams operating, such that each pair of villages participated in the study simultaneously, i.e. we conducted meetings 1, 2, and 3 on the same day for both villages.
- Since we have 31 parishes, we sampled all parishes once and three parishes twice.

We also stratified the assignment by four blocks of 16 assignments, i.e. we played every assignment pair once (in random order) before moving to the next block of 16.

A.3.3 Village attrition

Our protocol was designed to address attrition of individuals by replacement with reserves. We also faced two possible sources of village attrition:

1. The field team was unable to locate a sampled village at mobilization time, or the village chose not to participate. In this case the team moved to another randomly selected village from the same parish.
2. A village chose to withdraw from the study during the experiment. In this case we replaced the village with a randomly selected village from the same parish. To avoid disrupting the field work, these replacement villages were visited at the end of the experiment.

Overall, we had only one village that chose to withdraw. This village was replaced with another village from the same parish.

[40]Our original sampling plan was 64 villages, we later discovered we had sufficient budget to increase to 68.

A.3.4 Materials

Figure A1 gives examples of the handouts that participants received, showing their title cards, debt card, map of inial allocations, and payoff parameters including adjacency bonuses.

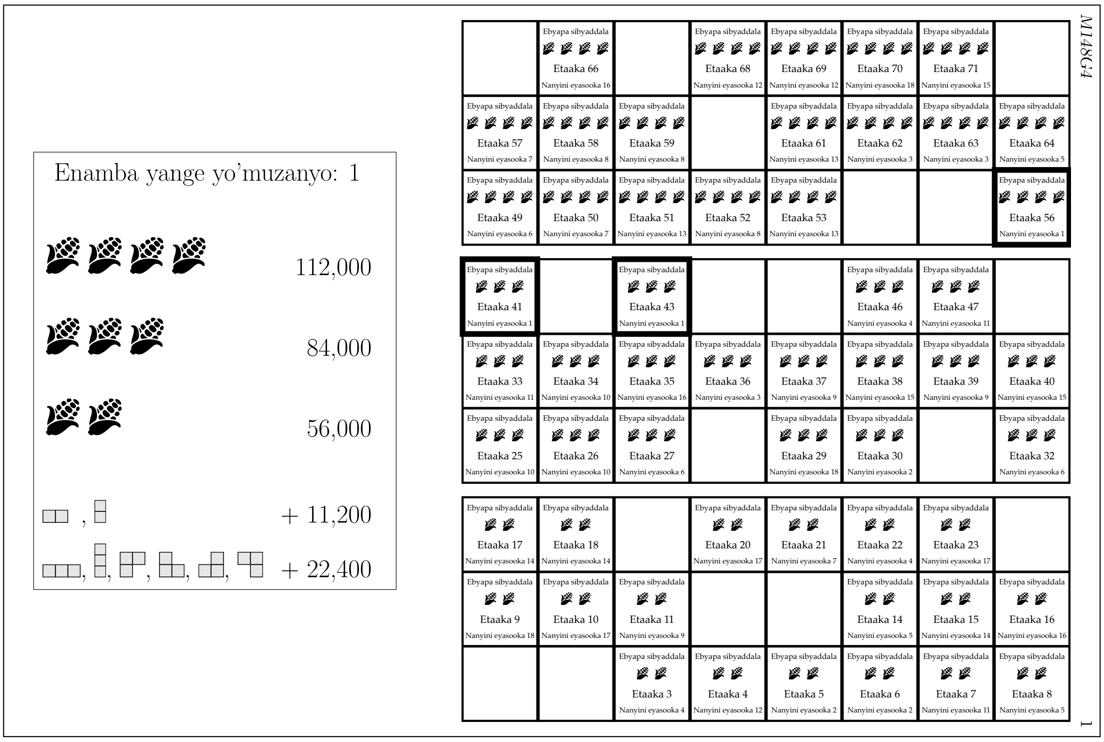


Figure A1: Participant materials in Experiment 1

B Appendix B: Additional Results from Experiment 1

This appendix contains all additional analysis that was part of the pre-analysis plan for experiment 1. The pre-analysis plan can be found at `https://doi.org/10.1257/rct.4581`.

Week-level analysis and learning Our pre-analysis plan included specifications in which we study (i) learning across the two periods and (ii) compare efficiency, consolidation, sorting, and exposure gains and losses in each week separately. Table B3 provides the details of these regressions.

As can be seen in Panel A, overall efficiency is lower in the second week relative to the first week. This reduction in efficiency is primarily due to much larger exposure losses, which suggest that individuals in the second week traded to intermediate positions that they could not trade out of. These is also weak evidence of slightly lower consolidation gains and slightly higher sorting gains. However, we note that these measures are influenced by the reassignment of unused plots to highest value users and are both lower in the second week if we do not adjust the measures for exposure losses.

As seen in Panel B, there is no strong difference between the simple and complex treatment in the first week, the second week, or after the surprise centralization treatment. As such, the map complexity appears to be second order relative to inefficiencies that exist in both the simple and complex map.

Alternative efficiency, sorting, and consolidation measures Our pre-analysis plan also specified number of alternative efficiency, sorting, and consolidation measures. These are provided in Tables B4 and B5 below.

In Table B4, we provide alternative measures of consolidation (column 2) and sorting (column 3) in which we do not reassign unused lots to their highest value use when calculating gains. There continues to be no significant difference in efficiency between the simple and complex treatments using these alternative values. In column (4), we use an alternative adjusted efficiency measure where we reassign unused plots to their highest value use. The simple treatment is again not significant in this specification.

In Table B5, we analyze efficiency, consolidation, and sorting only in the High-quality region in the first three columns. We then report on an alternative count-based consolidation and sorting measures. For the consolidation measure, we replaced "land value" with the number of plots owned by a player of the efficient type. For the sorting measure, we counted the number of adjacency bonuses rather than using the value of these bonuses.

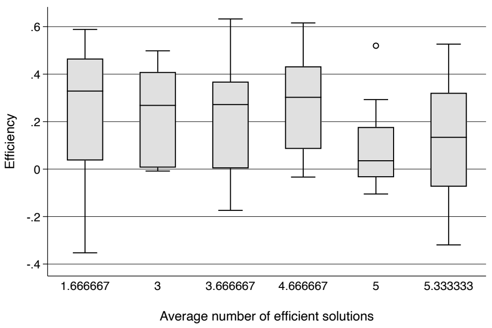
As seen in the table, the results using these alternative measures are similar to those provided in the main text. As seen in panel A, there is weak evidence that consolidation is easier in the simple treatment and no other significant differences between the simple and complex treatments. As seen in Panel B, the centralization treatment improves efficiency and consolidation, but has no impact on sorting. There is also no significant interaction between the centralization treatment and map complexity.

Alternative Inequality Measures Our pre-analysis plan discusses a potential inequality measure based on the Shapley Value. However, the Shapley analysis was sensitive to efficiency and we specified that we would only complete the analysis if efficiency was over 70%. Given the low efficiency observed in both weeks, we did not do the Shapley analysis for Experiment 1.

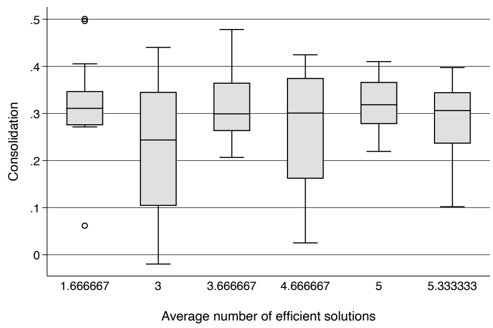
Alternative complexity measure Our pre-analysis plan proposes to descriptively analyze how efficiency and its decomposition depend on the number of (welfare-equivalent) efficient solutions in complex maps. Figure B1 plots these measures against the number of efficient solutions, averaged across quality regions. Consistent with our finding that the simple/complex map treatment – which substantially affects how many solutions exist – did not have large effects, we do not see any clear relationship in these graphs.

B.1 Non pre-specified analyses

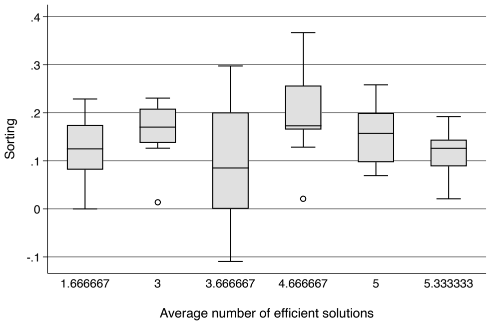
Figure B2 and Table B2 present additional analysis that we discuss in Section 7.



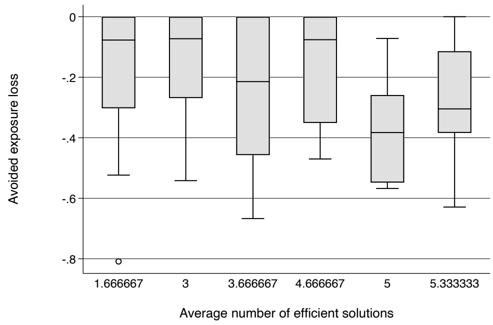
(a) Efficiency



(b) Consolidation



(c) Sorting



(d) Exposure loss

Figure B1: Analysis of alternative map complexity measure in experiment 1

Each plot graphs efficiency (or a subcomponent) against the number of ways to efficiently "pack" consolidated three-plot farms on our complex maps, averaged across the three quality regions.

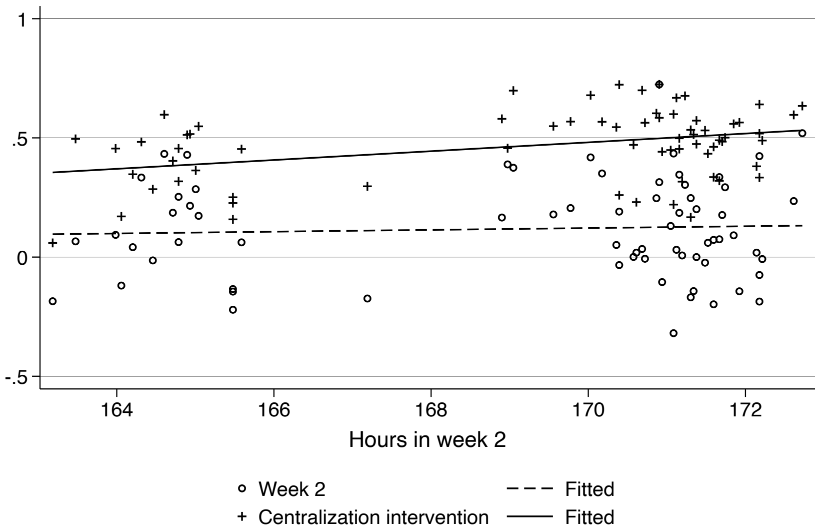


Figure B2: Relationship between time to trade and efficiency

We plot efficiency against the total time available to trade in week 2. We plot the relationship separately for efficiency measured before and after the trading day. There is no apparent improvement in week 2 outcomes for groups that had more time, and no negative relationship between the length of the week 2 window and the improvements attained during the centralization intervention.

Table B1: Summary Statistics: Experiment 1

	Our sample			Uganda		
Demographics	mean	S.D.	obs	mean	S.D.	obs
Age	43.76	13.52	1404	39.11	17.48	3338
Female	0.51	0.50	1404	0.51	0.50	3338
Head of household	0.65	0.48	1404	0.38	0.49	3338
Married: monogamous	0.63	0.48	1404	0.49	0.50	3338
Married: polygamous	0.06	0.24	1404	0.11	0.32	3338
Nr adults (inc respondent)	2.99	1.54	1404	2.60	1.27	1246
Nr children in household	3.37	2.07	1404	2.97	2.13	1246
Education						
Education (years)	7.16	3.21	1404	6.34	3.24	2551
Numeracy	0.76	0.37	1224			
Farm size and income						
How many plots do you own and cultivate?	2.10	1.15	1404	1.69	0.93	1246
Total land holdings cultivated (in acres)	2.95	3.32	1349	2.94	4.22	1244
Income from agriculture (1000 UGX/season)	1482	2174	1349	897	1995	847
Income from agriculture (USD PPP/season)	1365	2002	1349	826	1837	847
Farming ability (self-evaluated, relative to best in village)						
Farmer's total production	0.47	0.30	1403			
Max farm size (w/o hired labor)	0.59	0.35	1403			
Preferences (1-5 scale)					GPS	
Patience	4.35	0.66	1404	3.52	1.17	1000
Risk tolerance	4.09	0.90	1404	3.40	0.91	1000

Note: Comparison demographic data is from the Living Standards Measurement Study - Integrated Surveys on Agriculture 2019–2020 (Uganda National Bureau of Statistics, 2020) and the sample is restricted to respondents aged 18 and older whose main income comes from agriculture, and cultivates one or more plots. Statistics are weighed by household. Time and risk preferences are from a nationally representative sample of Uganda, and are sourced from the Global Preference Survey (GPS) (Falk et al., 2018). We thank Armin Falk and Markus Antony for sharing the GPS summary statistics needed for this comparison. Numeracy is the average between the following two numeracy questions (dummy = 1 if answered correctly) Q1) If one bottle of milk costs 2.480 and you give 2.500, how much change do you receive? and Q2) If 5 bottles cost 10.400, how much does one cost? Preference measures rescaled to 1–5 scale for comparability, with higher numbers indicating higher patience and lower risk aversion. Productivity relative to "best in village" is the farmer's total production and maximum farm size relative to what they think the best farmer in the village could produce/farm. Farmer's total production relative to best farmer is winsorized at the 99th percentile due to an extreme value. Household income from agriculture is the total production per season from all plots owned by household. USD purchase power parity (PPP) was 1085.85 at the end of 2019 (source: NASDAQ Data Link).

Table B2: Endogenous Centralization

		Decomposition		
	(1) Efficiency	(2) Consolidation	(3) Sorting	(4) Avoided exposure loss
Endogenous centralization	-0.008	0.057**	0.016	-0.081***
	(0.038)	(0.024)	(0.013)	(0.025)
Control mean	0.118	0.262	0.138	-0.282
Observations	136	136	136	136

Note: We regress efficiency and its decomposition on a dummy which equals 1 for the 62% of villages where 100% of respondents said that they got together in groups to trade during the experiment (âEndogenous Centralizationâ) relative to those where less than 100% did. Data from weeks 1 and 2, excluding the centralization treatment. All regressions control for map fixed effects. Standard errors (in parentheses) are clustered by village.

Table B3: Comparison of Efficiency Across Weeks and Week-by-week comparison of the Simple and Complex map treatments

		Decomposition		
	(1) Efficiency	(2) Consolidation	(3) Sorting	(4) Avoided exposure loss
Panel A				
Week 2	-0.223***	-0.037***	0.040**	-0.226***
	(0.032)	(0.013)	(0.016)	(0.028)
FDR q-value: Week 2		[0.005]	[0.007]	[0.001]
Control mean	0.342	0.324	0.107	-0.089
Observations	136	136	136	136
Panel B				
Simple map × Week 1	0.072*	0.039**	-0.028	0.061*
	(0.043)	(0.019)	(0.019)	(0.035)
Simple map × Week 2	0.016	0.005	-0.008	0.018
	(0.044)	(0.023)	(0.017)	(0.038)
Simple map × Centralization	0.030	0.011	0.003	0.015
	(0.032)	(0.015)	(0.017)	(0.019)
FDR q-value: simple map × Week 1		[0.148]	[0.148]	[0.148]
FDR q-value: simple map × Week 2		[1.000]	[1.000]	[1.000]
FDR q-value: simple map × Centralization		[1.000]	[1.000]	[1.000]
Control mean	0.208	0.294	0.136	-0.222
Observations	204	204	204	204

Note: Control group in Panel A: week 1. Control group in Panel B: complex maps. Column (1) shows absolute efficiency. Columns (2), (3) and (4) show the decomposition of efficiency in consolidation, sorting and exposure respectively. Panel A uses data from week 1 and week 2 (no Centralization treatment). Panel B uses data from week 1, week 2, and the Centralization treatment. Regressions in panel A control for village fixed effects. Regressions in panel B control for the centralization treatment and village pair, field team and week2 × map fixed effects. Standard errors clustered by village in parentheses.

Table B4: Alternative Consolidation, Sorting, and Efficiency Measures

	Unadjusted			Adjusted
	(1) Efficiency	(2) Consolidation	(3) Sorting	(4) Efficiency
Panel A				
Simple map	0.044	0.021*	0.023	0.004
	(0.031)	(0.012)	(0.026)	(0.021)
Control mean	0.208	0.286	-0.078	0.430
Observations	136	136	136	136
Panel B				
Centralization	0.342***	0.071***	0.270***	0.054***
	(0.027)	(0.009)	(0.024)	(0.011)
Simple map × Centralization	0.013	0.011	0.002	0.017
	(0.038)	(0.016)	(0.033)	(0.020)
Control mean	0.111	0.272	-0.161	0.435
Observations	136	136	136	136

Note: Control group in Panel A: complex maps and in Panel B: complex maps and week 2. Column (1) shows absolute efficiency. Columns (2) and (3) show the decomposition of efficiency into consolidation and sorting, unadjusted for exposure losses (this means that most avoided exposure losses are counted as sorting gains). Column (4) shows "adjusted" efficiency (efficiency after adding back all exposure losses). Panel A uses data from week 1 and week 2 (no Centralization treatment). Panel B uses data from week 2 only. Regressions in panel A control for village and week2 × map fixed effects. Regressions in panel B control for village fixed effects. Standard errors clustered by village in parentheses.

Table B5: Analysis of high-quality region and alternative count-based measures

	High quality region			Count-based	
	(1) Efficiency	(2) Consolidation	(3) Sorting	(4) Consolidation	(5) Sorting
Panel A					
Simple map	-0.010	0.007	-0.026	0.049**	-0.005
	(0.028)	(0.005)	(0.025)	(0.024)	(0.019)
Control mean	0.275	0.081	0.217	0.573	0.116
Observations	136	136	136	136	136
Panel B					
Centralization	0.050***	0.017***	0.004	0.121***	-0.006
	(0.014)	(0.003)	(0.013)	(0.017)	(0.009)
Simple map × Centralization	0.024	0.004	0.026	0.008	0.006
	(0.025)	(0.006)	(0.018)	(0.031)	(0.012)
Control mean	0.296	0.080	0.247	0.551	0.143
Observations	136	136	136	136	136

Note: Control group in Panel A: complex maps, and in Panel B: complex maps and week 2. Columns (1), (2) and (3) show measures of efficiency, consolidation and sorting for the high-quality region of the maps. Columns (4) and (5) show measures of count-based consolidation and count-based sorting for all regions of the maps. Panel A uses data from week 1 and week 2 (no Centralization treatment). Panel B uses data from week 2 only. Regressions in panel A control for village and week2 × map fixed effects. Regressions in panel B control for village fixed effects. Standard errors clustered by village in parentheses.

C Appendix C: Implementation Details for Experiment 2

This appendix contains additional implementation details for Experiment 2. Section C.1 describes the payoffs, variation in starting maps, and procedure for randomizing map orders. Section C.2 describes how we assigned treatments to sessions. Finally, C.3 describes the algorithms and provides details of the computer interfaces used in the exchange.

C.1 Payoffs and Maps

Payoffs: Each of our games consisted of 6 players and a map. Each player was assigned a numeric ability type. This ability type was private information and farmers were asked not to share it with others.

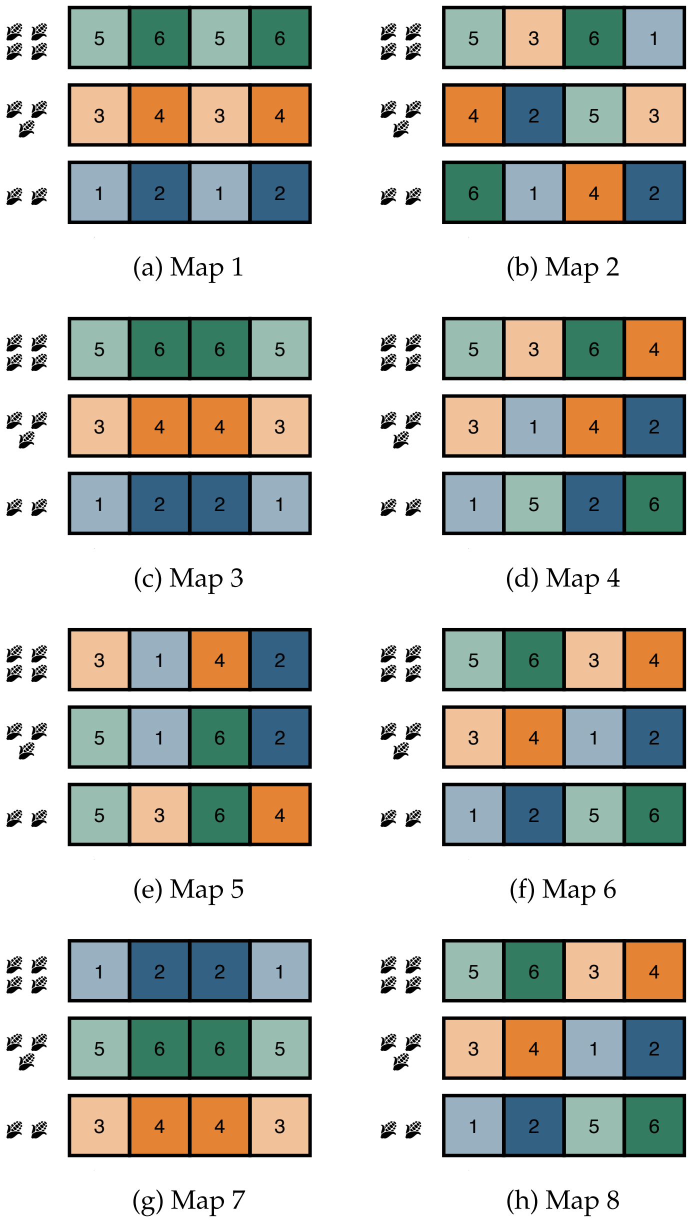
As with the first experiment, we considered an environment where land was fragmented and where additional gains could be achieved through sorting. Land was again divided into three quality regions with high-quality land being twice as valuable as low-quality land and medium quality land 1.5 times as valuable. Farmers were also divided into three farmer types: high ability, medium ability, and low quality. In all sessions there were two of each type of farmer and medium-ability and high-ability farmers were 50% and 100% more productive than low-ability farmers. Participant earnings were calculated based on their type-specific value for their two highest-quality pieces of land, plus an adjacency bonus if their two highest-value land holdings were adjacent.

As seen in Table C1, the return to a given piece of land was the product of the farmer's ability and the land type. Adjacency bonuses were set at 40% of the value of a single piece of land for the farmer and therefore scaled with both the quality of the land and the ability of the farmer.

	Panel A: Land values			Panel B: Adjacency bonuses		
	High Quality	Medium Quality	Low Quality	High Quality	Medium Quality	Low Quality
High Ability	400	300	200	160	120	80
Medium Ability	300	225	150	120	90	60
Low Ability	200	150	100	80	60	40

Table C1: Land and Farmer Types in Experiment 2

Maps: We conjectured that the the initial allocation of plots would affect the ease of achieving consolidation and efficient sorting. To study this issue, we created eight different initial land allocations, which are shown in Figure C1. In each diagram, players 1 & 2 are low, 3 & 4 are medium and 5 & 6 are high ability types.



(a) Map 1 (b) Map 2 (c) Map 3 (d) Map 4 (e) Map 5 (f) Map 6 (g) Map 7 (h) Map 8

Figure C1: Initial Land Allocations, Experiment 2

Note: numbers correspond to player IDs: 1–2 are low types, 3–4 medium types, 5–6 high types. The top region is high-quality land, the middle region medium-quality, and the bottom region low-quality.

The allocations are ordered according to our pre-experimental assessment of how difficult it would be to reach full efficiency. We considered four dimensions of difficulty. First, for each player, we determined how many *Package-1* trades were necessary to get to their efficient allocation.[41] Second, we considered how many farmers would need to be involved in any efficient *Package-2* trade. Third, we considered whether money was required to reach an efficient outcome. Finally, we considered strategic issues, for example the extent to which one farmer could holdup another farmer.

C.2 Treatment assignment

We played 48 sessions in total. Each session consisted of 8 auctions, and was assigned to one trading mechanism: *Package-1*, *Package-2*, or *Package-4*. In each session, the first four auctions had the same cash treatment and the second four auctions had the alternative cash treatment. Hence, each session could be assigned to one of six possible treatments that varied by mechanism and the ordering of the cash treatments. These treatments were block randomized. The set of 48 sessions was divided into 8 blocks, each consisting of 6 consecutive sessions. Each of the 6 treatments was then randomly assigned to one of the sessions within each block.

Each lab session required one lead enumerator to introduce the environment and implement the computer programs, 6 bidding assistants, and one broker. Two labs (labeled red and black) ran in parallel, each playing one session in the morning and one in the afternoon. Lead enumerators were assigned to a specific lab (red or black) and stayed in that lab throughout. Bidding assistants were randomly assigned to a specific farmer and lab (e.g. farmer 4 red) on a session by session basis. Brokers were also randomly assigned on a session by session basis.

Because subjects arrived slowly over time (it was hard to get farmers to all arrive at 9am), the first session of the day alternated between the red and black lab. The first 6 farmers to arrive were randomly assigned to a player number between 1 and 6, and then played in the lab that was operating the first session. The next six farmers to arrive were similarly assigned a player number, and played in the second lab. Each farmer played four auctions as their initial player number, and was then moved to a different player number. This was done such that every subject had an equal chance of being assigned to play one of the six possible sequences $\{HM; HL; MH; ML; LH; LM\}$.

Finally, the 8 maps displayed in Figure C1 were assigned to sessions. Every session played every map, and they were played in one of 8 orders. These orders were devised to minimize ordering effects: we wanted to have difficulty approximately even across the session to minimize the impact of learning. To assign orders to sessions, we first randomly permuted the 8 map orders as shown in Table C2. We then assigned map orders 1 to 6 to the sessions in block 1 (in order), orders 2 to 7 to block 2 (in order), etc.

Overall, this method gives assignment to the main auction and cash treatments

[41] Note that if an allocation required two *Package-1* trades, it required only one *Package-2* trade. If an allocation requires four *Package-1* trades, it requires two *Package-2* trades or one *Package-4* trade.

Order 1	5	1	3	7	6	2	4	8
Order 2	7	3	1	5	8	4	2	6
Order 3	6	2	4	8	5	1	3	7
Order 4	8	4	2	6	7	3	1	5
Order 5	3	7	5	1	4	8	6	2
Order 6	1	5	7	3	2	6	8	4
Order 7	4	8	6	2	3	7	5	1
Order 8	2	6	8	4	1	5	7	3

Table C2: Map Orders

that are orthogonal to the other elements of the design, as well as maps that are assigned orthogonally to the treatments and also randomly across time and session. We also have balance across all main elements of the experimental design.

C.3 Algorithms and Interfaces

C.3.1 The Winner Determination and Surplus Division Algorithms

Winner determination and surplus division are as outlined in Goeree and Lindsay (2019) with some modifications to impose XOR bidding. Let the set of farmers, $\mathbb{F}$, be indexed by $i \in \{1,...,6\}$ and the set of plots, $\mathbb{L}$, be indexed by $l \in \{1,...,12\}$. Farmers submit orders $o = (m, x)$ consisting of the minimum amount of money they must receive, m, and a vector of demanded plots, $x \in \{-1, 0, 1\}^{12}$. A negative number indicates that a farmer is offering money or trying to sell a plot, while a positive number indicates that a farmer must receive money, or wants to buy a plot. For instance, an order $(-500, \langle 1, 0, ..., 0\rangle)$ indicates that a farmer is willing to pay up to 500 points to acquire plot 1, while an order $(0, \langle 1, -1, 0, ..., 0\rangle)$ implies that the farmer is willing to buy plot 1 and sell plot 2, as long as he pays no money.

Orders placed by a farmer must be *legal*. Denote the plots owned by farmer i at time t as $\omega_i^t \in \{0, 1\}^{12}$ and denote the cash of farmer i at time t as c_i^t. A bid (m, x) is legal if at the time of placing the order, $c_i^t + m \geq 0$ and $\omega_i^t + x$ contains only zeros and ones. A bid is thus legal if the farmer has more cash than the amount of money he offers, he sells only land that he owns, and he buys only land that he does not own. Orders placed by a farmer are also restricted by the mechanism used in each treatment, as outlined above.

Legal orders are sent to the order book in the order that they arrive, and transactions occur any time there exists a set of legal orders where: (i) supply equals or exceeds demand for all plots; (ii) only a single order is used for each farmer; and (iii) the total amount of money offered is not positive. Formally, let $\mathbb{O}^t$ denote the legal orders in the order book at time t, and index its elements $o_j = (m_j, x_j)$, by

$j = \{1, \ldots, |\mathbb{O}^t|\}$. Let $d = \{0,1\}^{|O^t|}$ be a vector of orders from the order book, where $d_j = 1$ if an order j is winning and $d_j = 0$ otherwise. Let $\mathbb{O}_i^t$ be the active orders of farmer i and let $\mathbb{W}_i = \{o_j \in \mathbb{O}_i^t | d_j = 1\}$ be the orders of farmer i that are winning. At each time t we find:

$$V^* \equiv \max_d \sum_j -m_j d_j$$

subject to

$$\sum_j x_j^l d_j \leq 0 \qquad \forall l \in \mathbb{L}, \quad \text{and}$$
$$|\mathbb{W}_i| \leq 1 \qquad \forall i \in \mathbb{F}.$$

Trade is triggered if $V^* \geq 0$.[42]

When a transaction is triggered, we return plots that were not demanded to their original owners, and transfer all other plots according to the set of winning orders. If there is a positive surplus (i.e., $V^* > 0$), we divide the remaining surplus amongst the winning farmers as follows: let $\mathbb{W} = \{o_j \in \mathbb{O}^t | d_j = 1\}$ be the set of winning orders and $\widehat{\mathbb{W}} = \{o_j \in \mathbb{O}^t | o_j \in \mathbb{O}_i^t, |\mathbb{W}_i| = 1\}$ be the set of all orders made by the winning farmers. Likewise, denote the set of orders made by non-winners by $\mathbb{NW} = \mathbb{O}^t \setminus \widehat{\mathbb{W}}$. Let $p \in \{0, \ldots, 10000\}^{12}$ be a vector of (integer) prices, and denote the surplus generated by order j at prices p as $s_j(p) = -m_j - p \cdot x_j$.[43] As is standard in these problems, we find the set of prices that lexicographically maximizes the minimum surplus of winning farmers, subject to the revealed preference constraints of the losing orders.[44] The revealed preference constraints ensure that a losing farmer would not prefer to have won once the surplus is reallocated given the information that was submitted to the market. Finding these prices is equivalent to solving:

$$\min_p \sum_j d_j \left(s_j(p) - \frac{V^*}{|\mathbb{W}|} \right)^2$$

[42]Note that the restriction to legal trades ensures that there is no short selling, and that all budget constraints are met. We handle these on the client side to minimize the computation time required to solve the winner allocation problem, and to make farmers aware of attempted bids that could not be exercised. Relative to Goeree and Lindsay (2019), the additional cardinality constraint prevents more than one order from a farmer being used in each transaction. This constraint ensures that orders submitted by each farmer are considered XOR. Further, we only use the bids of non-winners to set prices, while Goeree and Lindsay (2019) use all non-winning bids. This change avoids a situation that can arise in our setting, where bidders impose revealed preference constraints on themselves, and reduce their own surplus.

[43]We use integer prices in the experiment in the range of 1 and 10000 so that trade prices are similar to ones that farmers are likely to encounter when trading in Kenya Shillings on a day-to-day basis.

[44]See Kwasnica et al. (2005) for a broader discussion of revealed preference constraints.

subject to:

$$\begin{aligned} s_j(p) &\geq 0 \quad \forall o_j \in \mathbb{W}, \\ s_j(p) &\leq 0 \quad \forall o_j \in \mathbb{NW}, \quad \text{and} \\ \sum_j d_j s_j(p) &= V^*. \end{aligned}$$

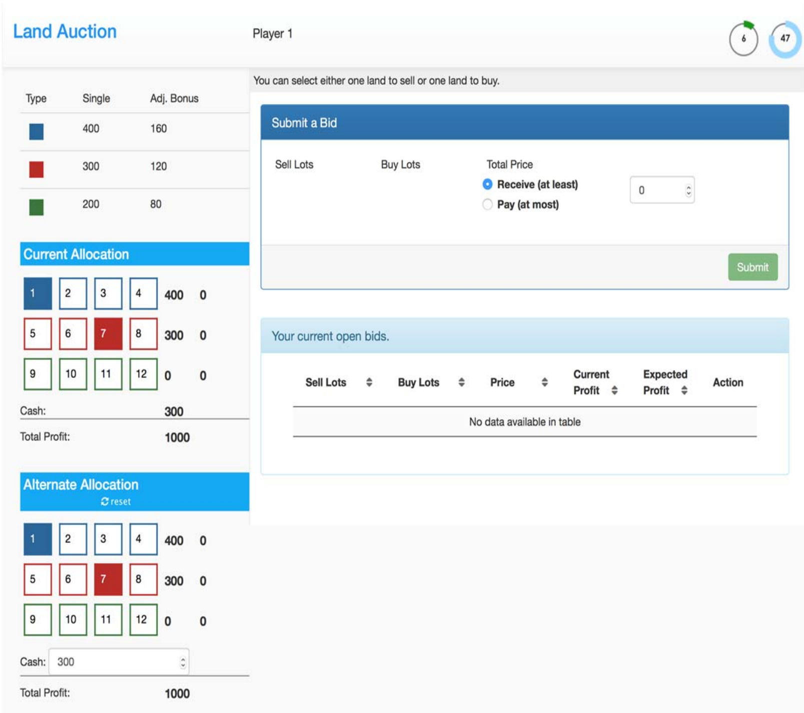
Each winner pays or receives $p \cdot x_j$ and losing farmers pay and receive nothing. In the case of ties, we use the first solution found by the solver.[45]

As can be seen in the optimization rule above, lexicographically maximizing the minimum surplus is equivalent to minimizing the squared difference between the surplus of each winner and the equal split subject to an additional constraint that all surplus is allocated. We explain our surplus division rule using this logic. Farmers are told that we try to split the surplus as evenly as possible between the farmers but that we want to make sure that farmers who do not trade are not disadvantaged. In training our enumerators we gave two main examples — one where there is a single buy order and a single sell order and where the surplus is divided equally, and one where there are two buy orders and a single sell order and where the non-winning buy order pins down prices.

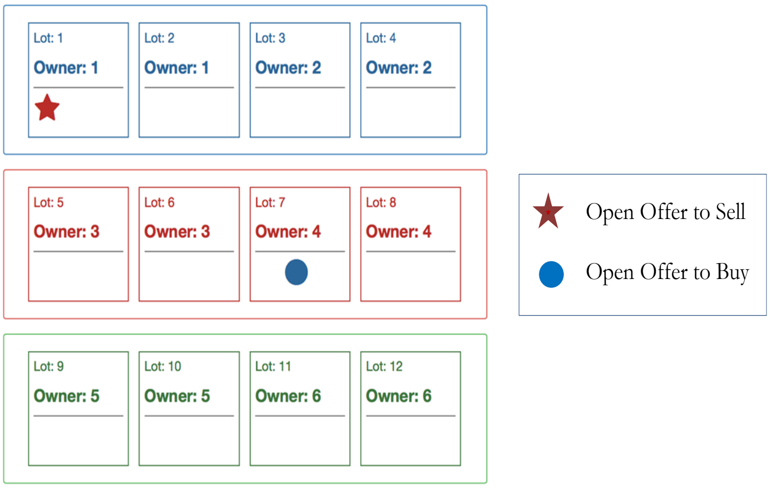
After a transaction is triggered, all non-winning orders made by farmers in the winning coalition become inactive, and we allow farmers to renew any legal orders if they wish. Orders that are made illegal (for instance, orders that contain sale offers of objects no longer owned) are hidden from a farmer's offer book, but can be renewed if later transactions make them legal. Farmers have the ability to withdraw legal orders at any time.

Interfaces All bids were entered through a computer interface. The interface displayed the farmer's valuations and current allocation on a geospatial map as in Panel (a) of Figure C2, and provided a calculator that could be used to determine the value of different allocations. Players (or their bidding assistant) could click on sets of plots on the map (depending on the treatment) and enter a willingness to pay, or willingness to accept to make the trade. Only legal bids were accepted by the computer. The interface also showed a list of all current bids placed by the farmer. In addition to the individual interfaces, a projector showed a map which indicated who owned each plot of land and when a plot of land was offered for sale, or had an offer to purchase. Combinatorial bids showed up on the projected interface as separate components. A screenshot of the individual and projected interfaces is shown in Figure C2.

[45]The underlying algorithms were written in Minizinc, a free open-source constraint modelling language, and solved using GECODE (Nethercote et al. 2007; Stuckey et al. 2014). In general, the winner determination problem could be solved in under 200 milliseconds for order books containing under 100 legal orders. The surplus division rule was slightly slower, but usually completed in under 600 milliseconds.



Panel (a): Computer Interface Used by Farmers



Panel (b): Projected Land Market Interface

Figure C2: Computer Interfaces

D Appendix D: Additional Results from Experiment 2

D.1 Full sample results

Our preferred regression specification for this experiment drops the first randomization block of six sessions, due to comprehension issues we encountered in these early sessions. Tables D2 and D3 report the full sample results. The efficiency results are qualitatively similar to our main results but the treatment effects are smaller, due to very low efficiency in block 1. Our inequality findings are qualitatively and quantitatively similar to the main results.

D.2 Efficiency and Initial Land Allocation

As discussed above, we conjectured that the ability to achieve full efficiency would depend on the initial allocation of plots, and we tentatively ranked our 8 initial allocations in order of perceived difficulty. Figure D1 shows that efficiency gains depend on the initial allocation of plots, but are not monotonically decreasing in our pre-experimental assessment of difficulty.

We ranked maps by a conjecture on whether or not *full* efficiency would be reached. As shown above, however, full efficiency was rarely reached, and so ease of reaching partial efficiency was more important. For example, on the basis of full efficiency, we believed that Map 8 was very hard, and Map 5 less difficult. Inspection of Figure D1, however, implies that this was not the case. One possible explanation is that for map 5, consolidation (and efficiency) requires a *Package-2* chain with three people involved. On the other hand, while full efficiency in Map 8 requires a *Package-2* chain with at least 4 people, consolidation requires only a *Package-2* chain with 2 players. Thus 8 is easy to consolidate and hard to improve sorting, but 5 is hard to consolidate. Because our auctions mostly reduced consoldiation, Map 8 turned out to be easier than Map 5.

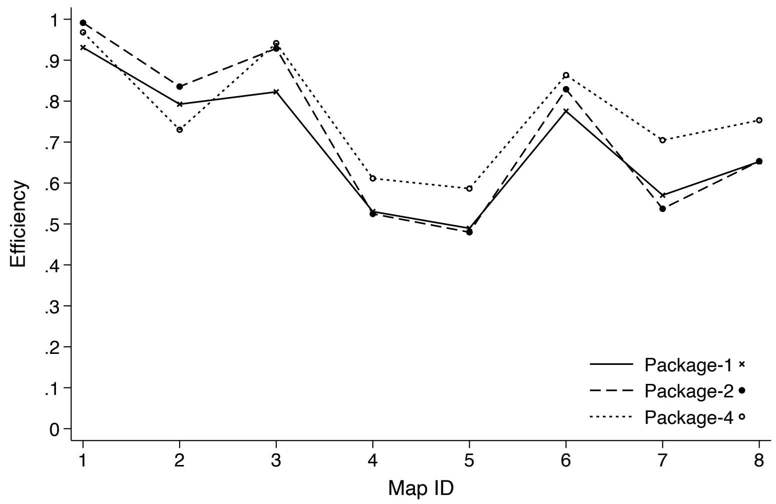


Figure D1: Mean efficiency by map and treatment

F-test for no difference by map: $F(7,39) = 23.67$, $p < 0.001$. Figure excludes data from randomization block 1 (see section 6.3 for discussion).

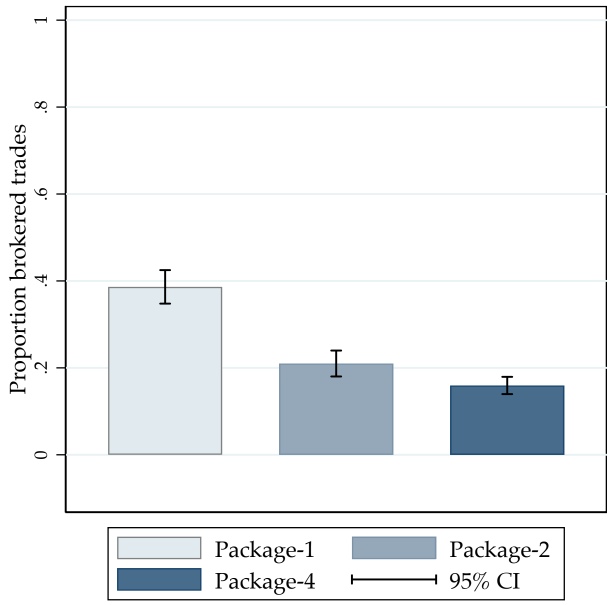


Figure D2: Proportion of Brokered Trades in Each Treatment

The figure plots the fraction of trades that had zero monetary surplus under each mechanism in experiment 2. We interpret these as "brokered" trades that were verbally agreed before being entered into the system. Figure excludes data from randomization block 1 (see section 6.3 for discussion).

Table D1: Summary Statistics: Experiment 2

	Our sample			Kenya		
Demographics	mean	S.D.	obs	mean	S.D.	obs
Age	42.65	10.45	263	38.73	16.61	51535
Female	0.58	0.50	264	0.52	0.50	51535
Married	0.77	0.42	264	0.63	0.48	51535
Nr of people in household	4.06	1.71	264	4.31	2.48	23785
Education						
Education (years)	9.75	2.94	264	8.01	4.23	51416
Land tenure						
Owns two or more plots	0.22	0.41	264			
Total land ownership in acres	1.01	1.52	237	2.56	3.79	23230
Land trade						
Fraction of plots with joint ownership	0.61	0.49	303			
Fraction of plots that are far from home	0.24	0.43	303			
Fraction of plots with a title	0.64	0.48	303			
Fraction who bought a plot (last 12 months)	0.05	0.22	264			
If has bought land: How many acres	0.83	1.42	11			
Fraction who sold a plot (last 12 months)	0.02	0.14	264			
If has sold land: How many acres	7.62	11.80	4			
Fraction of sales due to emergencies	0.40	0.55	5			
Consolidation						
How important is it to have all your plots together? (1–10, 1 is better to have spread out)						
1	0.43	0.50	264			
2 – 9	0.08	0.27	264			
10	0.47	0.50	264			
Why?						
Why fragment? Less risky	0.25	0.43	264			
Why consolidate? More productive	0.38	0.49	264			
Preferences (1–5)					GPS	
Risk tolerance	3.95	1.42	264	3.49	0.93	998

Comparison demographic data is from the Kenya Demographic and Health Survey 2014 (Kenya National Bureau of Statistics et al., 2015), for individuals aged 18 and older that own land suitable for agriculture. Time and risk preferences are from a nationally representative sample of Kenya, and are sourced from the Global Preference Survey (GPS) (Falk et al., 2018). We thank Armin Falk and Markus Antony for sharing the GPS summary statistics needed for this comparison.

Table D2: Efficiency in Experiment 2, including block 1

		Decomposition		
	(1) Efficiency	(2) Consolidation	(3) Sorting	(4) Avoided exposure loss
Panel A: Impact of package mechanisms				
Package-2	-0.002	0.002	0.004	-0.008
	(0.037)	(0.018)	(0.011)	(0.018)
as % of first best		*0.003*	*0.013*	
Package-4	0.042	0.008	0.024*	0.011
	(0.031)	(0.015)	(0.013)	(0.015)
as % of first best		*0.011*	*0.090*	
FDR q-value: Package-2		[1.000]	[1.000]	[1.000]
FDR q-value: Package-4		[0.693]	[0.255]	[0.693]
Control mean	0.677	0.614	0.114	-0.051
Control mean: % of first best		*0.838*	*0.428*	
Observations	366	366	366	366
Panel B: Impact of package mechanisms and low cash treatment				
Package-2	-0.016	-0.013	0.018	-0.021
	(0.048)	(0.024)	(0.017)	(0.026)
Package-4	0.061	0.016	0.033	0.012
	(0.044)	(0.019)	(0.020)	(0.020)
Low cash	-0.001	-0.019	0.008	0.011
	(0.043)	(0.025)	(0.020)	(0.020)
Package-2 × low cash	0.027	0.030	-0.029	0.026
	(0.050)	(0.029)	(0.028)	(0.027)
Package-4 × low cash	-0.038	-0.017	-0.019	-0.002
	(0.058)	(0.034)	(0.027)	(0.027)
FDR q-value: Package-2		[1.000]	[1.000]	[1.000]
FDR q-value: Package-4		[0.603]	[0.480]	[0.603]
F-test p-value: all low cash effects = 0	0.570	0.333	0.603	0.230
Control mean	0.678	0.624	0.110	-0.056
Observations	366	366	366	366

Note: *** denotes significance at the 1% level, ** 5% level and * 10% level. % of first best shows coefficient as share of potential gains within each category. First best: consolidation = 0.733, and sorting = 0.357. Control group: Package-1. Columns (2), (3) and (4) show the decomposition of efficiency in consolidation, sorting and avoided exposure loss respectively. Regressions in columns (1), (2) and (4) use data from all maps, sessions and randomization blocks 1-8. All columns control for low cash dummy and auction number, map and randomization block fixed effects. Standard errors clustered by session in parentheses.

Table D3: Inequality in Experiment 2, including block 1

	Atkinson Index (log utility)		
	(1) High cash	(2) Low cash	(3) High & Low
Package-2	0.0010	-0.0022**	0.0010
	(0.0006)	(0.0011)	(0.0006)
Package-4	0.0001	-0.0017*	0.0001
	(0.0006)	(0.0010)	(0.0006)
Package-2 × low cash			-0.0032***
			(0.0010)
Package-4 × low cash			-0.0018*
			(0.0010)
F-test p-value: all low cash effects = 0			0.009
Control mean	0.012	0.035	0.024
Observations	183	183	366

Note: *** denotes significance at the 1% level, ** 5% level and * 10% level. The control mean in columns (1) and (2) are not directly comparable because the Atkinson inequality index is not invariant to additive changes in total wealth. Control group: Package-1. Regressions in column (1) use data from high cash group, in column (2) use data from low cash group and in column (3) use data from both high and low cash groups. All regressions control for randomization block, auction and map fixed effects, and in column (3) adds interactions of each of the fixed effects with a low cash dummy. Standard errors clustered by session in parentheses.

Impact

This doctoral dissertation undertakes a comprehensive investigation into crucial aspects of refugee integration, labor market dynamics, and market design in Uganda and Kenya. Each chapter delves into distinct areas of inquiry, yielding valuable insights and implications for various disciplines. In Chapter 1, the focus lies on examining whether work contact between refugees and locals can foster social cohesion in refugee-receiving countries. Chapter 2 delves into the effect of matching firms with refugee workers to enhance labor market efficiency and integration. Lastly, Chapter 3 explores market design for land trade aiming to improve agricultural productivity among smallholder farmers.

The research presented in this dissertation carries substantial potential for scientific advancement and addresses pressing societal challenges. Scientifically, the findings from Chapter 1 contribute to a deeper understanding of social cohesion mechanisms, enriching research in social sciences and inter-group relations. In contrast to much of the existing experimental literature, we measure biases and behavioral change with contact both for the majority group (i.e. the local workers) and the minority group (i.e. the refugees), using the most relevant activity for adults, namely work. As far as we are concerned, there is no experimental evidence regarding the effect of workplace contact on inter-group integration. Also, by measuring both implicit bias, explicit bias, and behavior, we can use the latter to interpret the former and thus contribute to the discussion on how to measure and interpret implicit bias through implicit association tests (IATs).

Chapter 2's exploration of matching attitudes between employers and refugees advances knowledge of labor market behavior and provides key insights into promoting inclusive employment practices. Much of the existing literature focuses on improving firms' access to information about the quality of job seekers or adjusting workers' and employers' expectations. By contrast, our intervention targets firms' demand for workers from a disadvantaged group.

Chapter 3 offers innovative market design strategies, with implications for resource allocation and agricultural productivity, serving as an important reference for researchers in market design and economic studies. Our work complements prior studies that have established a set of missing markets and institutions that inhibit trade, focusing largely on property rights. Despite the success of this literature and the widespread adoption of programs such as land titling, recent work suggests that these innovations are necessary but not sufficient to achieve efficiency. We believe that addressing the frictions we identify can help unlock the benefits of land titling and that our market designs, which rely on voluntary trade and leverage farmers' preferences and information, may be preferable to governmental consolidation programs in settings with low state capacity, low trust in government, and risk of expropriation. Finally, we contribute to the thin literature on using market design for developmental challenges.

The impact of the research in this dissertation extends beyond academia and holds significant relevance for society. The first two chapters' investigation into social cohesion and refugee integration in the labor market is particularly pertinent for policymakers and organizations working with refugees, fostering inclusive communities, and promoting trust among diverse populations. It is especially relevant for low- and middle-income countries that host three-quarters of the world's refugees.

Specifically, enhancing labor market integration for disadvantaged workers, including migrants and refugees, is a crucial policy concern, especially in low-income countries where labor markets struggle and resources are limited. Despite possessing skills, refugees often encounter difficulties in securing employment due to perceived skill gaps and assessment costs for employers. Moreover, the influx of refugees might disrupt and change social dynamics, making it more difficult for refugees and natives to collaborate. Chapter 2 demonstrates that even brief exposure between refugee workers and firm owners or managers increases hiring after eight months, particularly for employers with positive experiences. Moreover, Chapter 1 shows that short work contact between local and refugee workers can also enhance their cooperation and possibly, future business collaboration.

The results presented in these two first chapters provide robust support for the adoption of policies that facilitate the issuance of labor permits to refugees, enabling them to actively engage in the workforce of their host nation. While there have been debates regarding allowing refugees to work in their host country, our research provides evidence that enabling skilled refugees to access employment opportunities can significantly aid their integration process, while at the same time boosting the economic and social prosperity of local businesses. Moreover, governments and organizations 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 preparatory training before engaging in the internship. This may assist them in adjusting their initial attitudes and improve the out-group contact experience.

Chapter 3's research holds significant societal importance by focusing on market design to effectively and equitably redistribute agricultural land, with potential implications for sustainable livelihoods, food security, and economic growth among smallholder farmers, particularly in low-income countries. A prevalent issue in many low- and middle-income countries, such as in Africa or Latin America is the fragmentation of agricultural plots into small and scattered pieces of land. This fragmentation can result in lower productivity and limited gains from trade. However, there is substantial evidence indicating that consolidating and reallocating these plots could yield substantial advantages for both individual farmers and their communities at large. Our study not only sheds light on the potential benefits but also highlights the importance of well-designed trade rules to facilitate this transformative process.

The research findings appeal to a diverse spectrum of stakeholders with varying interests. In the initial two chapters, the insights gathered from the results hold significant value for policymakers and governmental bodies responsible for refugee integration and labor market regulations. These stakeholders can leverage these findings to craft well-informed policies that foster both social cohesion and labor market efficiency. Additionally, entities such as the United Nations, non-governmental organizations (NGOs), and refugee-led organizations stand to benefit from the research findings due to their alignment with their existing programmatic endeavors.

The level of interest in the research is already evident. The findings have been presented at the United Nations High Commissioner for Refugees (UNHCR) and other high-level conferences, and these have garnered citations from influential institutions like the World Bank (Schuettler and Caron, 2020) and the Center for Global Development (CGDev) (Baseler et al., 2023). The findings were also cited in the influential World Development Report 2023: Migrants, Refugees and Societies (World Bank, 2023). Furthermore, for the first two chapters, we have collaborated with significant partners, including the Directorate of Industrial Training of the Ugandan government, as well as the International Growth Center (IGC), YARID, and BONDEKO – two refugee-led organizations operating within the country.

The strategy for disseminating the research findings covers governmental and international organizations, as well as refugee-led groups. A multifaceted approach has already been adopted, including dissemination sessions and the creation of informative blog posts. The initial two chapters have already been featured in blog posts on prominent platforms specializing in development economics research, such as J-PAL, Innovations for Poverty Action (Loiacono and Silva-Vargas, 2023a), IGC (Loiacono and Silva-Vargas, 2023c) and VoxDev (Loiacono and Silva-Vargas, 2023b). Furthermore, the refugee integration research was showcased through a photo exhibition held in Paris in June 2023. The exhibition was an integral part of the “Science and the Fight against Poverty: How Far Have We Come in 20 Years and What’s Next” conference, an event presided over by Nobel Prize laureate Prof. Esther Duflo and organized by J-PAL Europe.

Finally, this research exhibits a strong alignment with several key United Nations Sustainable Development Goals (SDGs). The content of Chapter 1, which places significant emphasis on the promotion of social cohesion and the creation of inclusive communities, directly corresponds to the principles outlined in SDG 10 (Reduced Inequalities) and SDG 16 (Peace, Justice, and Strong Institutions). Chapter 2's primary focus on enhancing the integration of refugees into the labor market is inherently tied to the objectives outlined in SDG 8 (Decent Work and Economic Growth). Lastly, the research in Chapter 3, aimed at improving agricultural productivity, resonates strongly with several SDGs. The pursuit of eradicating poverty and ensuring food security, as evident in SDG 1 (No Poverty) and SDG 2 (Zero Hunger) respectively, is closely intertwined with the chapter's goals. Furthermore, the efficient and equitable management of land resources, as explored in the chapter, aligns harmoniously with the principles of SDG 15 (Life on Land), which advocates for sustainable land-use practices and management.

In conclusion, this doctoral dissertation has far-reaching implications for both scientific understanding and societal progress. The findings from each chapter contribute significantly to various fields, including economics, development studies, political science, and psychology. Moreover, the research outcomes have practical applications for governments, policymakers, local firms, and smallholder farmers, fostering positive changes in social cohesion, labor market integration, and agricultural productivity. The active dissemination of research findings to diverse stakeholders ensures that this work will drive positive change and contribute to the achievement of sustainable development goals in the studied regions and beyond.

The exhibition can be found in this link: https://marijosilvaphotography.com/J-PAL-20-Exhibition-in-Paris

Bibliography

Baseler, T., Ginn, T., Hakiza, R., and Woldemikael, O. (2023). Can Redistribution Change Policy Views? Aid and Attitudes toward Refugees in Uganda. Working Paper, Center for Global Development (CGD).

Loiacono, F. and Silva-Vargas, M. (2023a). Improving Labor Market Opportunities for Refugees in Uganda. Policy Brief, Innovations for Poverty Action (IPA).

Loiacono, F. and Silva-Vargas, M. (2023b). Helping Refugees Integrate into the Labour Market: Evidence from Uganda. Policy Blog, VoxDev.

Loiacono, F. and Silva-Vargas, M. (2023c). Refugees and the choice between cities and settlements: Evidence from Uganda. Policy Brief, The International Growth Centre (IGC).

Schuettler, K., & Caron, L. (2020). Jobs Interventions for Refugees and Internally Displaced Persons. World Bank Jobs Working Papers. Issue N. 47.

World Bank. (2023). World Development Report 2023: Migrants, Refugees, and Societies. The World Bank.

Acknowledgments

As I embark on the culmination of a six-year journey spanning seven countries, expressing gratitude feels like an understatement. This dissertation, a tapestry woven across continents, owes its existence to a multitude of people and places.

First, my gratitude extends to my exceptional supervisors, Prof. dr. ir. Eleonora Nillesen and Prof. dr. Jonathan de Quidt, who not only provided guidance but also fostered a nurturing environment for ambitious research to flourish. Eleonora, your support and guidance allowed my ideas to take shape and become a reality. Jon, your supervision, mentorship, and friendship have been a guiding light through both the smooth and tumultuous terrains of this academic journey. Special appreciation goes to the committee members: Prof. dr. Melissa Siegel, Prof. dr. Martina Björkman Nyqvist, Prof. dr. ir. Erwin Bulte, and Dr. David McKenzie. Heartfelt thanks to David McKenzie for sharing my passion for photography and research during this journey.

I would like to extend my sincere appreciation to my dedicated paranymphs, Michelle Gonzalez Amador and Stefanie Roost, for their encouragement and friendship throughout my PhD journey.

Secondly, I would like to thank my dearest family, my unwavering support system, starting with the extraordinary force that is my mother, Machi. Her unyielding belief in the power of vision not only shaped my academic pursuits but also steered the trajectory of our lives. Migrating alone from Bolivia to Italy, she crafted a better future for her daughters. Gracias, mami; your courage is the silent force behind every line of this dissertation.

To my sister, Clow, whose intelligence, kindness, patience, and friendship provided the support and love I needed. Without your guidance, I would not have finished this dissertation. You are the pillar of my strength. And to my cuñado Roberto, thank you for being a support system to all of us.

To my dad, Jaime, who, despite the distance, consistently offered the artistic perspective that added a unique dimension to my life and career. And to Sirius White, whose silent canine companionship has seen me through the end of my bachelor's, master's, and now, a PhD. You have been a constant amidst the chaos of academia and life, and for that, I am eternally grateful.

To my co-authors, Prof. Gharad Bryan, Dr. Francesco Loiacono, and Prof. Tom Wilkening, thank you for the shared laughter, the moments of tension, and the countless lines of code that shaped our journey. Each of you has been a source of inspiration and learning.

My gratitude extends to the various research teams I had the privilege to collaborate with, especially the IGREC team, the Metajua team, and all the enumerators, field managers, and research assistants who contributed to the data collection process. Your dedication has been the backbone of this research. Special thanks to Joshua Bwiira, Lenny Gichia, Ian Kusimakwe, and Apollo Tumusiime, with whom it was a pleasure to collaborate. I am deeply grateful to all the participants for giving us their time and patience to reply to our questions, making this research possible. I am also forever thankful for the organizations that collaborated with us on such ambitious projects. Special thanks to YARID and its founder, Robert Hakiza, as well as Bondeko and its founder, Paul Kithima.

A heartfelt thank you to my "two home universities" - UNU-MERIT at Maastricht University and the IIES at Stockholm University. Professors, staff, colleagues, and friends, your support provided the ideal foundation for my PhD journey, blending academia with vibrant friendships. In Maastricht, our interactions extended beyond the realms of academia, involving discussions on econometric models combined with a lot of dancing. In Stockholm, the "Development tea" group, a professor biting a gold Nobel prize coin to test its authenticity, the tradition of "fika", the commitment to winter sports, and jumping into frozen lakes after hot saunas, even in the face of darkness, have left an indelible mark on my memory!

To the friends I made in Uganda, you transformed a foreign land into

a home, making challenging moments bearable and joyful ones unforgettable. I will never forget numerous precious moments, such as riding boda-bodas to commute to work or even to another village, dancing in the middle of the Ssese Islands, conducting a photo session inside Makerere University, engaging in insightful discussions while staying at Nakivale refugee settlement, or walking through the jungle to meet a family of gorillas. All these memories and the people who were part of them will forever be with me.

To my cherished family and friends in Bolivia and Italy, your presence, though distant, is a constant motivator. Your constant love inspires me to forge ahead. A family predominantly composed of strong, inspiring women, each of your stories propels me forward. Moreover, my family's artistic inclinations remind me never to forget to infuse art into everything I do.

To my friends, Norka, Anna V., Crespin, Cristina C., Divya, Gissel, Gege, Joakim, Sophie, and Vale, thank you for being there through thick and thin.

Finally, the move to Paris and joining J-PAL Europe was a fortuitous twist of fate, seamlessly combining all the elements I hold dear and introducing me to an extraordinary team. To my new colleagues, I extend my gratitude for transforming work into a passionate pursuit and for your unwavering friendship and support during the final stages of my PhD. To the newfound special friends, climbers, photographers, and more in Paris, though our time together has been brief, thank you for the unexpected joys that have added vibrant hues to my life. I look forward to discovering what other adventures we will have.

As I reflect on this journey, I am reminded that every word penned in this dissertation is a collective effort, a symphony of support, mentorship, and friendship, with a touch of chaos. To everyone who played a part, thank you for being integral to this academic odyssey.

Gracias desde el fondo de mi ser.

About the author

“I like to think that my work is a mix between Bolivian melancholy, Italian Renaissance and Ugandan colours.”

Mariajose is a Bolivian-born and naturalized Italian researcher and photographer. During her PhD, she worked as a principal investigator and co-investigator on various field experiments related to the socio-economic integration of refugees in Uganda and complex constraints in land markets. Moreover, her passion for photography, policy, and research led her to use photography as a means of storytelling in her research and in other development projects. She has been featured in important photo-documentary outlets and development blogs, such as Photo VOGUE, Fotografas LATAM, the World Bank's Development Impact Blog, and more.

Currently, she serves as a Research and Policy Manager at J-PAL Europe, where she supports development agencies in applying and generating evidence to inform development policy. She was a Robert S. McNamara Fellow at the World Bank from 2021 to 2022 and has also visited esteemed institutions like the IIES at Stockholm University and the Global Poverty Research Lab at Northwestern University.

Before joining J-PAL Europe, Mariajose worked in Uganda and Tanzania, managing impact evaluations on labor markets, gender, and governance for the World Bank and Innovations for Poverty Action. She also conducted research in the Ivory Coast on a UN program related to the reintegration of ex-combatants through agriculture.

Mariajose holds a Master's degree in Agricultural Economics and Rural Development from Ghent University and Humboldt University of Berlin, a Bachelor's degree in Political Science from Roma Tre University, and a Digital Photography degree from the Ettore Rolli School of Arts in Rome.